Lecture 1: Modeling Change
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Learning Goals

By the end of the lesson, readers should be able to:

1. explain why a mathematical model is a purposeful simplification of reality;
2. recognise and test a **proportional relationship**;
3. translate “future = present + change” into a **difference equation**;
4. interpret parameters, initial conditions, and equilibrium values in context;
5. distinguish unconstrained growth from constrained growth;
6. evaluate a simple model through validation, sensitivity, and limitations.

What is a mathematical model?

Core explanation, worked reasoning, and application

What is a mathematical model?

A **mathematical model** is a mathematical representation built for a particular purpose. It may help us describe a system, explain a mechanism, forecast what may happen, or compare possible decisions. A model is not a complete copy of reality: it keeps the features that matter for the question and deliberately ignores others.

Activity: Launch: What would you model?

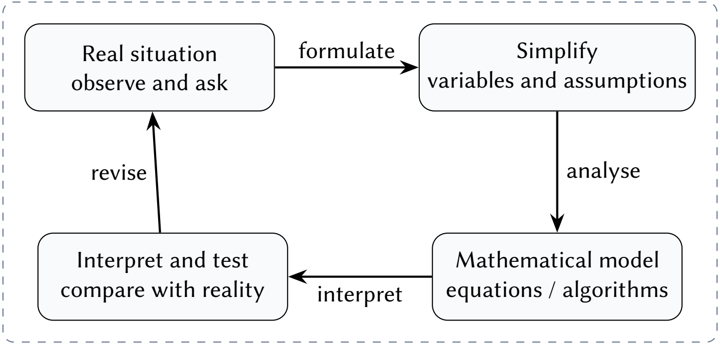
A school notices that the queue at its canteen is sometimes short and sometimes extends outside the building. In pairs, discuss:

1. What could be measured? List at least four variables.
2. What does "a bad queue" mean: long waiting time, long physical length, or too many late students?
3. Name two simplifying assumptions you might make.
4. What decision could the model support?

Modeling lesson: Before choosing mathematics, define the question and the output that matters.

Illustration: What is a mathematical model?

The modeling cycle



Modeling is iterative. A failed check is not the end; it tells us what to revise.

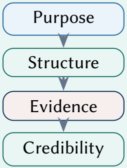
Concept: Four questions to ask throughout a modeling problem

Four guiding questions

1. **Purpose:** What decision or explanation is required?
2. **Structure:** Which variables interact, and how?
3. **Evidence:** What data or reasoning supports the relationship?
4. **Credibility:** How will we test the result and state limitations?

How to use them

- Ask them at the **start** of the task.
- Revisit them when building equations or diagrams.
- Use them again when checking results and limitations.



A first simplifying relationship: proportionality

Core explanation, worked reasoning, and application

A first simplifying relationship: proportionality

Two variables y and x are **proportional** if

$$y = kx,$$

where k is a **constant of proportionality**. The graph is a straight line through the origin. The units of k are important: they explain how much y changes for each unit of x.

Worked Example: Worked Example 1: Spring–mass data

A spring is loaded with different masses. Let m be mass in grams and e be elongation in centimetres. The measured data are shown below.

m	50	100	150	200	250	300	350	400	450	500	550
e	1.000	1.875	2.750	3.250	4.375	4.875	5.675	6.500	7.250	8.000	8.750

A least-squares line forced through the origin has slope

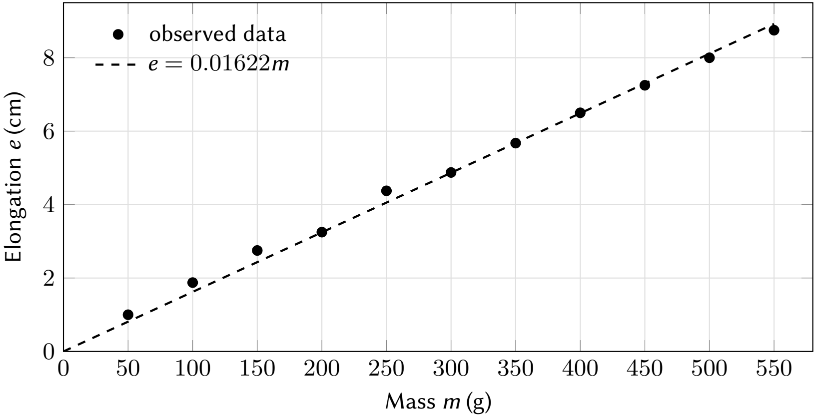
$$k = \frac{\sum m_i e_i}{\sum m_i^2} \approx 0.01622 \text{ cm/g}.$$

Thus a simple model is

$$\boxed{e = 0.01622m.}$$

For a mass of 420 g, the model predicts $e \approx 6.81$ cm.

Illustration: A first simplifying relationship: proportionality



The points lie close to a straight line through the origin, so proportionality is plausible.

Model Critique: Is the model exact?

The points do not lie exactly on the line because measurements contain error and a real spring may not obey the same relationship under every load. We should not write "the spring follows the equation exactly." A defensible statement is: *Within the tested range, elongation is approximately proportional to mass.*

Checkpoint: Proportionality

A taxi charges a fixed flag-fall fee plus a charge per kilometre. Is total fare proportional to distance? Explain in one sentence.
Hint: What should the fare be when distance is zero?

Modeling discrete change

Core explanation, worked reasoning, and application

Modeling discrete change

When a quantity is recorded at regular steps (days, months, generations), a useful starting point is

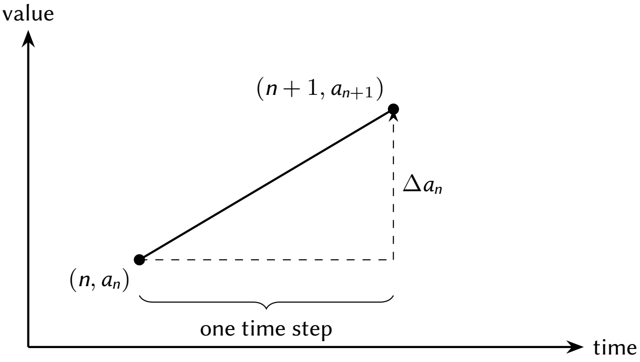
$$\boxed{\text{next value} = \text{current value} + \text{change during one step.}}$$

For a sequence $a_0, a_1, a_2, \ldots$, the **first difference** is

$$\Delta a_n = a_{n+1} - a_n.$$

A **difference equation** states how the next value depends on current values. Together with an **initial condition**, it defines a discrete **dynamical system**.

Illustration: Modeling discrete change



A first difference is the vertical change between consecutive observations.

Worked Example: Worked Example 2: Monthly compound interest

A certificate begins with $a_0 = 1000$ dollars and earns 1% per month. The monthly change equals 1% of the current balance:

$$\Delta a_n = 0.01 a_n.$$

Therefore

$$a_{n+1} = a_n + 0.01 a_n = 1.01 a_n, \qquad a_0 = 1000.$$

The first values are

$$1000,\ 1010,\ 1020.10,\ 1030.30, \ldots$$

Repeated substitution suggests the closed form

$$\boxed{a_n = 1000(1.01)^n.}$$

The recurrence explains the step-by-step mechanism; the closed form answers long-range questions quickly.

Activity: Guided practice: Drug elimination

8 min

A medicine loses 20% of the amount present each hour. Initially there are 640 mg.

1. Define a_n clearly, including its units.
2. Write the change equation $\Delta a_n =$ ____________________.
3. Write the recurrence and calculate a_1, a_2, a_3.
4. Without calculating every term, predict the long-term behavior.

Concept: Interpreting a multiplier

For $a_{n+1} = ra_n$:

Multiplier	Behavior
$r > 1$	positive growth
$0 < r < 1$	decay toward 0
$-1 < r < 0$	alternating signs with shrinking magnitude
$\lvert r \rvert > 1$	magnitude grows; signs may alternate

The meaning of negative values must be checked against the context. A recurrence can be mathematically valid but physically meaningless after it predicts a negative population or balance.

Constant input or removal: $a_{n+1} = ra_n + b$

Core explanation, worked reasoning, and application

Constant input or removal: $a_{n+1} = ra_n + b$

Many systems combine proportional change with a fixed input or output:

$$a_{n+1} = ra_n + b.$$

Examples include a regular deposit, a fixed monthly payment, medication taken at fixed intervals, and a stock replenished by a constant amount.
An **equilibrium value** a^* is a value that remains unchanged:

$$a^* = ra^* + b.$$

When $r \neq 1$,

$$\boxed{a^* = \frac{b}{1-r}.}$$

Subtracting the equilibrium from the recurrence gives a particularly useful form:

$$a_{n+1} - a^* = r(a_n - a^*).$$

Thus the distance from equilibrium is multiplied by r at every step.

Worked Example: Worked Example 3: Repeated medication

At the end of each day, 50% of a drug remains. Immediately afterwards, a patient takes another 0.1 mg. Let a_n be the amount just after the nth dose. Then

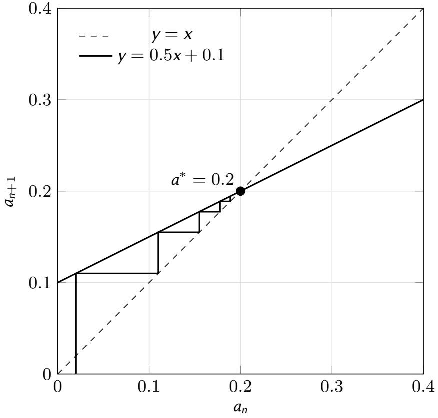
$$a_{n+1} = 0.5a_n + 0.1.$$

The equilibrium is

$$a^* = \frac{0.1}{1 - 0.5} = 0.2 \text{ mg}.$$

Because $|0.5| < 1$, the distance from 0.2 is halved each day. The equilibrium is **stable**: nearby solutions move toward it.

Illustration: Constant input or removal: $a_{n+1} = ra_n + b$



A cobweb plot: the iterations approach the intersection of $y = x$ and $y = 0.5x + 0.1$.

Worked Example: Worked Example 4: Withdrawal from an investment

An account earns 1% per month and pays out 1000 dollars at the end of each month:

$$a_{n+1} = 1.01a_n - 1000.$$

Its equilibrium is

$$a^* = \frac{-1000}{1 - 1.01} = 100000.$$

Because $1.01 > 1$, the equilibrium is **unstable**. Starting exactly at 100000 keeps the balance constant; starting above it causes growth, while starting below it causes the balance to move downward until the model eventually predicts a negative value. At that point the recurrence must stop, because an account cannot continue making withdrawals after it is empty.

Checkpoint: Equilibrium and stability

For each system, find the equilibrium and predict whether it is stable.

1. $x_{n+1} = 0.8x_n + 12$
2. $y_{n+1} = 1.1y_n - 5$

Explain the conclusion using the multiplier, not only a numerical table.

When growth slows: a logistic difference equation

Core explanation, worked reasoning, and application

When growth slows: a logistic difference equation

Exponential growth assumes that the same percentage growth can continue indefinitely. In a limited environment, resources become scarce and growth usually slows. A simple constrained model is

$$p_{n+1} = p_n + k(M - p_n)p_n,$$

where M is the **carrying capacity** and k controls the speed of growth.
The change

$$\Delta p_n = k(M - p_n)p_n$$

contains two factors:

- p_n: more organisms can produce more new organisms;
- $M - p_n$: the remaining capacity becomes small near the limit.

Worked Example: Worked Example 5: Yeast culture

Suppose

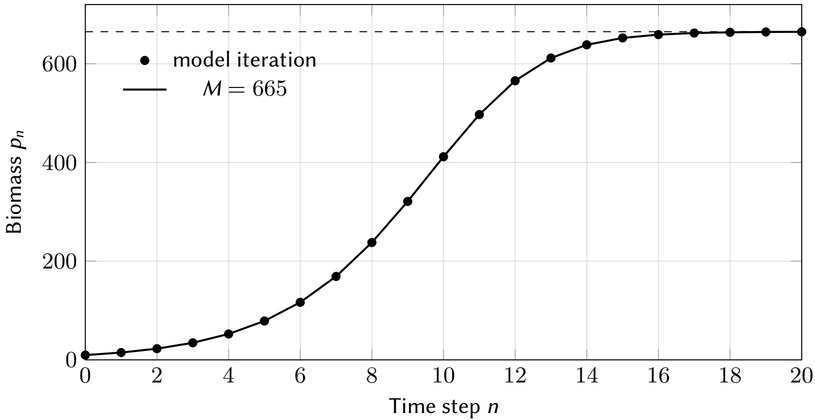
$$p_{n+1} = p_n + 0.00082(665 - p_n)p_n, \qquad p_0 = 9.6.$$

The first values are approximately

$$9.6,\ 14.8,\ 22.6,\ 34.5,\ 52.4,\ 78.7,\ 116.6, \ldots$$

Growth first accelerates, then slows as the population approaches 665. The S-shaped pattern is characteristic of **logistic growth**.

Illustration: When growth slows: a logistic difference equation



The population approaches the carrying capacity rather than growing without bound.

Model Critique: Sensitivity and interpretation

Changing k mainly changes *how quickly* the model approaches equilibrium. Changing M changes the predicted long-run level itself. Therefore, a conclusion about the final population is much more sensitive to an inaccurate estimate of M than to a moderate error in k.

A good competition report should state this in context rather than merely list percentages.

Activity: Model comparison

8 min

Two teams model the same yeast population.

$$\text{Team A: } p_{n+1} = 1.45 p_n,$$
$$\text{Team B: } p_{n+1} = p_n + 0.00082(665 - p_n)p_n.$$

Which model is more appropriate for short-term growth? Which is more credible over a long period? Give one reason for each answer.

Mini-case: a discrete SIR epidemic model

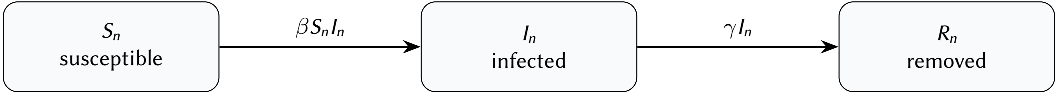
Core explanation, worked reasoning, and application

Mini-case: a discrete SIR epidemic model

This final case is a preview of a **system of difference equations**. It is included to demonstrate how the ideas above combine in a larger modeling problem; it is not intended for full derivation within this first lesson.
A dormitory contains $N = 1000$ students. At day n:

- S_n are susceptible;
- I_n are currently infected;
- R_n are removed from transmission (recovered or isolated).

Illustration: Mini-case: a discrete SIR epidemic model



Compartment flow in the SIR model.

Mini-case: a discrete SIR epidemic model

Under the simplifying assumptions of a closed population, homogeneous mixing, no reinfection, and constant rates,

$$\begin{aligned} S_{n+1} &= S_n - \beta S_n I_n, \\ I_{n+1} &= I_n + \beta S_n I_n - \gamma I_n, \\ R_{n+1} &= R_n + \gamma I_n. \end{aligned}$$

Here β is the transmission parameter and γ is the daily recovery proportion. For $\beta = 0.0005$, $\gamma = 0.1$, and $(S_0, I_0, R_0) = (999, 1, 0)$, the early reproduction indicator is

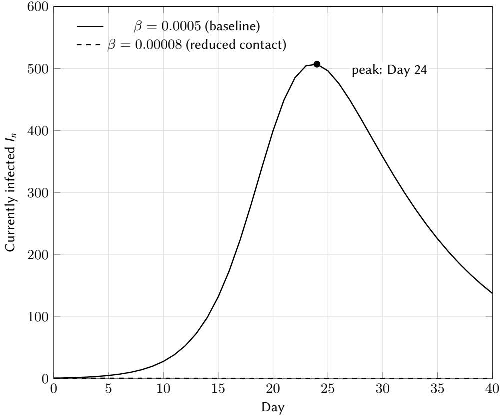
$$\mathcal{R}_0 \approx \frac{\beta S_0}{\gamma} = 4.995 > 1,$$

so infections initially increase.

Mini-case: a discrete SIR epidemic model

Day	S_n	I_n	R_n
0	999.0	1.0	0.0
5	993.6	5.4	1.1
10	965.0	28.1	6.9
15	831.8	132.1	36.1
20	448.9	399.7	151.3
24	157.7	507.0	335.2
30	33.3	357.8	608.9

Illustration: Mini-case: a discrete SIR epidemic model



Reducing the transmission parameter below the threshold prevents initial growth.

Activity: Competition-style communication

Using the table and graph, write a three-sentence conclusion:

1. one sentence reporting the baseline peak;
2. one sentence describing the intervention result;
3. one sentence stating a limitation of the model.

Avoid saying that the model “proves” what will happen.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Six ideas to retain

1. A model is built for a purpose and is always a simplification.
2. Parameter units and interpretations matter as much as algebra.
3. For discrete time, start with next $=$ current $+$ change.
4. In $a_{n+1} = ra_n + b$, the multiplier r determines stability around equilibrium.
5. Nonlinear feedback can slow growth and create a carrying capacity.
6. A credible model includes testing, sensitivity analysis, and limitations.

Checkpoint: Exit ticket

2 min

Complete the sentence:
“A mathematical model is useful when ______________________________________, but it may fail when ______________________________________.”

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A water tank contains 800 L and loses 6% of its current volume each hour.
   1.1 Write a recurrence for the volume V_n.
   1.2 Find V_5 and the long-term behavior.
   1.3 State one reason the model might eventually become unrealistic.
2. A club has 120 members. Each month, 8% of current members leave and 15 new members join.
   2.1 Formulate $M_{n+1} = rM_n + b$.
   2.2 Find the equilibrium membership.
   2.3 Decide whether the equilibrium is stable and explain why.
3. A town's recycling programme begins with 200 households. Participation grows according to

   $$H_{n+1} = H_n + 0.002(1000 - H_n)H_n.$$

   Explain the meaning of 1000 and why the model slows near that value.
4. A sequence satisfies $x_{n+1} = 1.15x_n - 30$, $x_0 = 250$.
   4.1 Find the equilibrium.
   4.2 Calculate the first four terms.

4.3 Does the context need a stopping rule if x_n represents money? Explain.

5. The following data record the number of downloads after successive advertising rounds:

$$80,\ 116,\ 169,\ 244,\ 351.$$

   Test whether a constant multiplier is a reasonable approximation. Estimate the multiplier and predict the next value. Comment on the uncertainty of your prediction.

6. In an epidemic model, β is estimated from only three days of data. Explain why sensitivity analysis with respect to β is important. Your answer should refer to decisions, not only mathematics.

Exercises: Modeling challenge

Choose one of the following and prepare a one-page modeling brief containing: purpose, variables, assumptions, recurrence(s), expected behavior, validation plan, and one limitation.

1. Number of bicycles available at two sharing stations each morning.
2. Daily amount of food waste in a school canteen after a reduction campaign.
3. Growth and saturation of users joining a new school app.

You do not need a perfect numerical answer. The quality of the formulation and justification is the main criterion.

Optional derivation: closed form for $a_{n+1} = ra_n + b$

Core explanation, worked reasoning, and application

Optional derivation: closed form for $a_{n+1} = ra_n + b$

This appendix is not part of the two-hour core lesson. It may be assigned to students who are comfortable with algebraic derivations.

Let $a^* = b/(1-r)$ for $r \neq 1$. Define $u_n = a_n - a^*$. Then

$$\begin{aligned} u_{n+1} &= a_{n+1} - a^* \\ &= ra_n + b - a^* \\ &= r(a_n - a^*) \\ &= ru_n. \end{aligned}$$

Therefore $u_n = r^n u_0$, so

$$\boxed{a_n = a^* + r^n(a_0 - a^*)} = \boxed{\frac{b}{1-r} + r^n\left(a_0 - \frac{b}{1-r}\right)}.$$

This form makes stability transparent: when $|r| < 1$, the term $r^n(a_0 - a^*)$ tends to zero.

Lecture 2: The Modeling Process and Scaling

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, readers should be able to:

1. turn a vague real-world question into a precise modeling objective;
2. identify variables, units, assumptions, and useful **submodels**;
3. distinguish a proportional relationship from a general linear relationship;
4. test relationships of the form $y \propto f(x)$ by transforming variables;
5. apply **geometric similarity** to derive area, volume, and weight scaling laws;
6. evaluate a model through interpretation, sensitivity, and limitations.

From a vague question to a modeling objective

Core explanation, worked reasoning, and application

From a vague question to a modeling objective

A real-world question rarely arrives in a form that can be solved immediately. The modeler must decide what is being predicted, for whom the answer is useful, and under what conditions the conclusion is intended to hold. This first act of narrowing the question is called **problem formulation**.

Activity: Launch: Is the scale model enough?

A school wants to build a shaded outdoor stage. Before construction, students make a 1:20 scale model and test it under a lamp.
In pairs, discuss:

1. Which measurements from the model can be scaled directly to the full stage?
2. Which effects might not scale in the same way?
3. What decision could the model support?
4. What evidence would make you trust the model?

Modeling lesson: A model can preserve some relationships while distorting others. We must know which quantities scale and why.

Concept: A practical modeling workflow

For most competition and classroom problems, the following six questions are more useful than immediately searching for a formula.

Workflow checklist

1. **Purpose:** What decision, explanation, or prediction is required?
2. **Variables and units:** What must be measured, and in what units?
3. **Assumptions:** What is held constant, neglected, or idealised?
4. **Structure:** Can the problem be separated into smaller submodels?
5. **Evidence:** How will parameters be estimated and results checked?
6. **Communication:** What conclusion is justified, and where might it fail?

Report-writing tip

A good modeling report does not begin with equations alone. It begins by clarifying **what is being asked**, **what is measurable**, and **how credibility will be checked**.

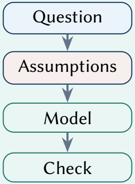
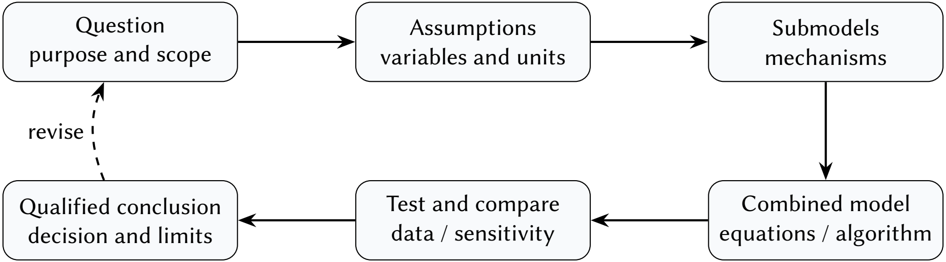


Illustration: From a vague question to a modeling objective



A model is developed through connected decisions, not by choosing a formula in isolation.

Checkpoint: A modelable question

Rewrite the vague question “How safe is this road?” as a more precise modeling question. Your version must identify an output quantity and at least one condition under which the answer applies.

Case study: vehicular stopping distance

Core explanation, worked reasoning, and application

Case study: vehicular stopping distance

Consider the question: *How much distance does a car need to stop safely?* A useful first model separates the process into two mechanisms:

$$\text{total stopping distance} = \text{reaction distance} + \text{braking distance.}$$

This is an example of **decomposition**: a complicated system is represented by simpler submodels whose outputs can be combined.

Worked Example: Worked Example 1: Building the stopping-distance model

Let v be the initial speed in metres per second, t_r the driver's reaction time in seconds, and a the magnitude of the braking deceleration in metres per second squared.

Reaction submodel. During the reaction interval, the car is assumed to travel at approximately constant speed:

$$d_r = t_r v.$$

Braking submodel. Once the brakes act, assume constant deceleration. From $0 = v^2 - 2ad_b$,

$$d_b = \frac{v^2}{2a}.$$

Combined model.

$$\boxed{d(v) = t_r v + \frac{v^2}{2a}.}$$

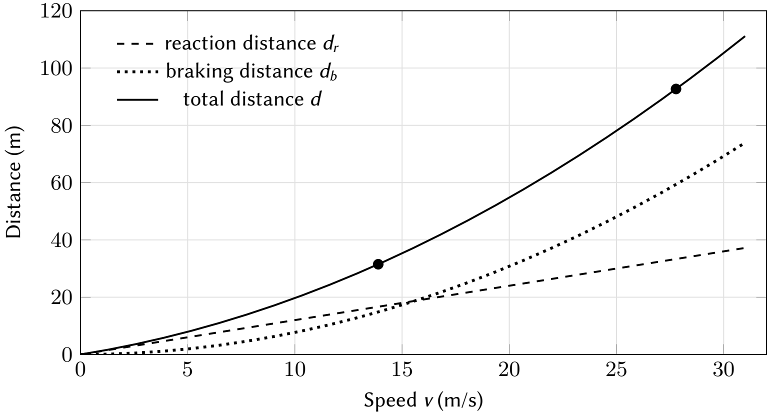
For a baseline driver with $t_r = 1.2$ s and a dry-road deceleration of $a = 6.5\ \mathrm{m/s^2}$,

$$\boxed{d(v) = 1.2v + 0.0769v^2.}$$

Worked Example: Worked Example 1: Building the stopping-distance model

At 50 km/h ($v = 13.89$ m/s), the predicted stopping distance is about 31.5 m. At 100 km/h ($v = 27.78$ m/s), it is about 92.7 m. Doubling speed therefore requires almost three times the distance, not merely twice the distance.

Illustration: Case study: vehicular stopping distance



The quadratic braking term becomes increasingly important as speed rises.

Checkpoint: When does braking dominate?

Find the positive speed at which the braking distance first equals the reaction distance. Express the answer in both m/s and km/h. What does this threshold mean for high-speed driving?

Case study: vehicular stopping distance

A model is not complete when an equation has been obtained. We must ask whether its conclusions are stable under plausible changes in the assumptions.

Worked Example: Worked Example 2: Sensitivity at 80 km/h

At 80 km/h, $v = 22.22$ m/s. Compare four scenarios:

Scenario	t_r (s)	a (m/s^2)	Predicted distance (m)
Baseline: alert driver, dry road	1.2	6.5	64.7
Slower reaction	1.8	6.5	78.0
Wet road	1.2	3.5	97.2
Slower reaction and wet road	1.8	3.5	110.5

For this speed, the conclusion is more sensitive to the available deceleration than to a moderate change in reaction time. This does not mean reaction time is unimportant; it means road and tyre conditions dominate this particular comparison.

Model Critique: What the model leaves out

The model assumes level ground, constant deceleration, immediate full braking after the reaction interval, and no change in direction. It also treats t_r and a as known constants. In reality, they vary with fatigue, distraction, weather, tyre condition, road slope, and braking technology.

A defensible conclusion is therefore conditional: *under the stated reaction time and deceleration assumptions, the required stopping distance increases nonlinearly with speed.*

Activity: Refine the model

Choose one omitted factor—road slope, rain, tyre wear, or driver fatigue.

1. Which parameter or submodel would it affect?
2. Would you represent the factor by a new variable or by changing an existing parameter?
3. What data would be needed to estimate the change?

Generalised proportionality and transformed variables

Core explanation, worked reasoning, and application

Generalised proportionality and transformed variables

A proportional model need not have the form $y = kx$. More generally,

$$y \propto f(x) \qquad \Longleftrightarrow \qquad y = kf(x).$$

If the relationship is plausible, plotting y against the transformed variable $f(x)$ should produce an approximately straight line through the origin.

Concept: Linear is not always proportional

A relationship $y = mx + c$ is **linear** in the school-algebra sense, but it is proportional only when $c = 0$. A taxi fare with a flag-fall charge is linear in distance but not proportional to distance.

Likewise, $y = kx^2$ is not a straight line when plotted against x, yet it is proportional to x^2 and becomes a straight line when plotted against the transformed variable x^2.

Worked Example: Worked Example 3: Kepler's third law

Let R be a planet's mean orbital radius measured in astronomical units (AU), and let T be its orbital period in Earth years. Kepler's third law suggests

$$T \propto R^{3/2}.$$

Using Earth, $R = 1$ and $T = 1$, so the proportionality constant is approximately 1:

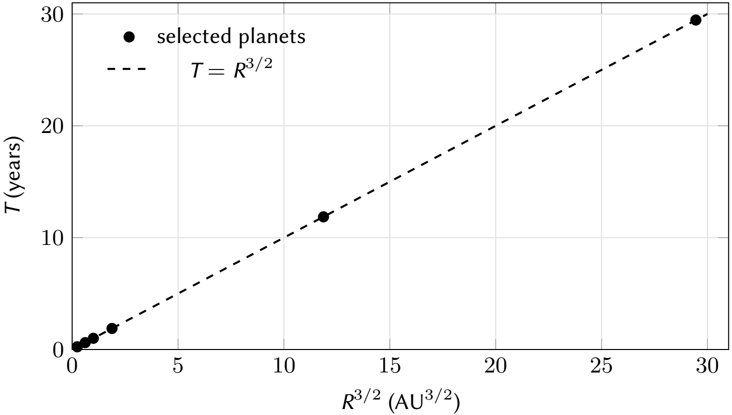
$$\boxed{T \approx R^{3/2}.}$$

For Jupiter, $R = 5.203$ AU, giving

$$T \approx 5.203^{3/2} \approx 11.87 \text{ years},$$

which is close to the observed 11.86 years.

Illustration: Generalised proportionality and transformed variables



After transforming the horizontal variable, Kepler's relationship is approximately proportional.

Checkpoint: Choose the useful transformation

For each proposed relationship, state what should be plotted on the horizontal axis to test proportionality.

1. $F \propto v^2$
2. $P \propto 1/V$
3. $N \propto e^{rt}$, where r is known

Geometric similarity and scaling laws

Core explanation, worked reasoning, and application

Geometric similarity and scaling laws

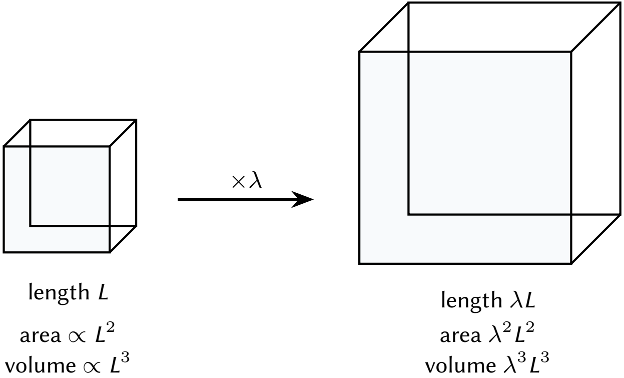
Two objects are **geometrically similar** if they have the same shape and every corresponding length is multiplied by the same **scale factor** λ.
If all lengths are multiplied by λ, then every area is multiplied by λ^2 and every volume is multiplied by λ^3:

$$\boxed{L \propto \lambda, \qquad S \propto \lambda^2, \qquad V \propto \lambda^3.}$$

For objects of the same material and density, mass and weight also scale as volume:

$$W \propto V \propto L^3.$$

Illustration: Geometric similarity and scaling laws



Length, area, and volume do not scale at the same rate.

Activity: Scaling table

Complete the table without using a formula sheet.

Scale factor λ	Length factor	Area factor	Volume factor
2			
3			
$\frac{1}{2}$			

Then explain why a twice-as-tall storage tank does not hold merely twice as much water.

Worked Example: Worked Example 4: A submarine scale model

A 7-m model and a 70-m submarine are geometrically similar, so the full-size vessel has scale factor $\lambda = 10$.

- Corresponding lengths are multiplied by 10.
- Surface areas are multiplied by $10^2 = 100$.
- Volumes and masses are multiplied by $10^3 = 1000$, assuming equal density.

If heat loss is mainly proportional to surface area, the full-size vessel loses about 100 times as much heat under comparable temperature conditions—not 1000 times as much. If storage capacity is proportional to volume, it is about 1000 times as large.

Model Critique: Similarity is an assumption, not an observation

The area and volume laws are exact for mathematically similar shapes. Their application to real objects is only as credible as the similarity assumption. A juvenile animal may not have the same proportions as an adult; a large bridge model may require thicker supports; and a raindrop changes shape as it grows.

Model refinement through a scaling model

Core explanation, worked reasoning, and application

Model refinement through a scaling model

A useful model is often developed in stages. We begin with the simplest plausible relationship, identify what it misses, and then refine it.

Worked Example: Worked Example 5: Estimating the weight of a fish

Suppose the weight W of a bass must be estimated without a scale.
Model I: length only. If all fish are geometrically similar and have approximately equal density, then

$$W \propto L^3.$$

This model is convenient, but two fish with the same length may have very different body thicknesses.
Model II: length and girth. Approximate the fish as a sequence of similar cross-sections. Cross-sectional area scales with the square of girth G, so

$$\boxed{W \propto LG^2.}$$

Suppose a fish with $L = 40$ cm, $G = 25$ cm weighs 1.25 kg. Then

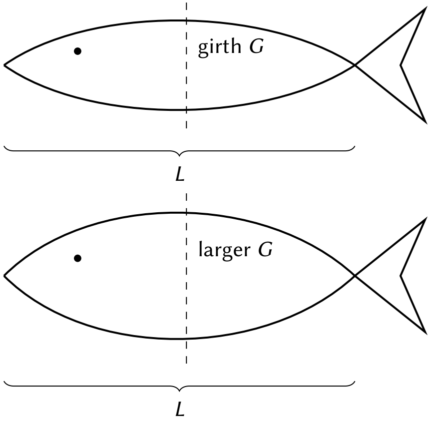
$$1.25 = k(40)(25)^2 \quad \Longrightarrow \quad k = 0.00005 \text{ kg/cm}^3.$$

A fish with $L = 50$ cm and $G = 30$ cm is predicted to weigh

$$W = 0.00005(50)(30)^2 = 2.25 \text{ kg}.$$

The second model uses one additional measurement to represent body condition more directly.

Illustration: Model refinement through a scaling model



Length alone cannot distinguish fish with different body thicknesses.

Model Critique: Does the refined model solve everything?

The model $W = kLG^2$ still assumes comparable species, density, and body shape. Girth measurements may be inconsistent, and a single calibration fish gives no estimate of prediction error. Refinement improves relevance, but it also introduces an additional measurement and another source of uncertainty.

Checkpoint: Compare two models

For a fish whose length increases by 10% while girth remains unchanged:

1. By what percentage does Model I, $W \propto L^3$, increase its prediction?
2. By what percentage does Model II, $W \propto LG^2$, increase its prediction?
3. Why do the models disagree?

A final scaling consequence: strength versus weight

Core explanation, worked reasoning, and application

A final scaling consequence: strength versus weight

Scaling laws can reveal qualitative limits even when no numerical constant is known. Suppose geometrically similar animals are made of comparable material.

- Muscle strength is approximately proportional to muscle cross-sectional area: strength $\propto L^2$.
- Body weight is proportional to volume: $W \propto L^3$.

Therefore

$$\frac{\text{strength}}{\text{weight}} \propto \frac{L^2}{L^3} = \frac{1}{L} \propto W^{-1/3}.$$

Activity: Can a giant ant keep its relative strength?

An ant is enlarged so that every length becomes 10 times as large, while shape and material remain unchanged.

1. By what factor does muscle cross-sectional area increase?
2. By what factor does body weight increase?
3. By what factor does strength per unit weight change?
4. State one reason why the geometric-similarity assumption may fail for a real giant animal.

Concept: What scaling arguments can and cannot do

A scaling argument often predicts an exponent or qualitative trend without determining the proportionality constant. It is especially useful for comparing sizes and identifying which effects grow faster. It does not replace calibration, validation, or knowledge of the underlying mechanism.

Optional Extension: Terminal velocity of similar objects

Suppose geometrically similar falling objects experience quadratic drag. At terminal velocity, weight and drag balance:

$$mg \propto A v_t^2.$$

For constant density, $m \propto L^3$ and frontal area $A \propto L^2$, so

$$L^3 \propto L^2 v_t^2 \quad \Longrightarrow \quad v_t \propto L^{1/2} \propto m^{1/6}.$$

The small exponent means terminal velocity changes relatively slowly with mass. This conclusion depends on quadratic drag, geometric similarity, and comparable density; it is not a universal law for every falling object.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Six ideas to retain

1. A vague question must be narrowed to a purpose, output, and set of conditions.
2. Complicated systems can often be built from interpretable submodels.
3. A proportional relationship $y = kf(x)$ becomes linear through the origin when y is plotted against $f(x)$.
4. In geometrically similar objects, length, area, and volume scale as λ, λ^2, and λ^3.
5. Model refinement should address a specific weakness, not merely add complexity.
6. Every conclusion should be accompanied by assumptions, sensitivity, and a statement of where the model may fail.

Checkpoint: Exit ticket

Complete both statements:

1. "The most important assumption in today's stopping-distance model is ________________________________ because ________________________________."
2. "When every length is doubled, ____________________ grows faster than ____________________ because ___________________________."

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A driver has reaction time 1.4 s and can decelerate at 5.0 m/s^2.
   1.1 Write a stopping-distance model in terms of speed v in m/s.
   1.2 Estimate the distance required at 72 km/h.
   1.3 Identify one parameter that is uncertain and explain how it could be measured.
2. A delivery fee is 25 dollars plus 4 dollars per kilometre.
   2.1 Is fee proportional to distance? Explain.
   2.2 Write the linear model and interpret both constants.
3. For each relationship, identify a transformed horizontal variable that would produce a line through the origin if the model were correct:

$$y \propto x^3, \qquad y \propto \sqrt{x}, \qquad y \propto \frac{1}{x^2}.$$

4. Planet X has a mean orbital radius of 4 AU. Use $T \approx R^{3/2}$ to estimate its orbital period. If the observed period is 8.5 years, calculate the relative error and suggest one possible reason for the difference.
5. A cube-shaped tank is redesigned so that each side is 30% longer.

5.1 By what factor does its surface area change?
5.2 By what factor does its capacity change?
5.3 Which quantity has the larger percentage increase?

6. A 1:25 model of a monument uses 0.80 kg of material. Estimate the material required for a geometrically similar full-size monument of equal density. State why the estimate may be unrealistic for an actual structure.

7. A fish model has the form $W = kLG^2$. A measured fish has $L = 45$ cm, $G = 28$ cm, and $W = 1.80$ kg.

   7.1 Estimate k.
   7.2 Predict the weight of a fish with $L = 52$ cm and $G = 31$ cm.
   7.3 State two sources of prediction uncertainty.

8. Two geometrically similar animals have weights in the ratio $8{:}1$.

   8.1 Find the ratio of their characteristic lengths.
   8.2 Find the ratio of their surface areas.
   8.3 Compare their surface-area-to-volume ratios.

Exercises: Modeling challenge

Choose one question and prepare a one-page modeling brief.

1. How should the dimensions of a school water tank change if student population increases by 50%?
2. Can a small wind-tunnel model predict the aerodynamic force on a full-size vehicle?
3. How can the weight of a fruit be estimated from measurements that do not require a scale?

The emphasis is on the logic of the model. A fully calibrated numerical answer is not required.

Concept: One-page modeling brief planner

Purpose and output	What decision or prediction will the model support? What quantity will be reported?
Variables and units	Define the measurable quantities and distinguish inputs from outputs.
Assumptions	State what is treated as constant, neglected, or geometrically similar.
Submodels / scaling	Give at least two relationships and explain why each is plausible.
Calibration	Identify the data needed to estimate any unknown constant.
Validation	Explain what observation or comparison would test the model.
Sensitivity	Name one assumption or parameter that should be varied.
Limitation	State where the conclusion should not be used.

Checkpoint: Submission check

Before submitting, underline the sentence in your brief that gives the main conclusion. Circle every numerical value that has a unit. Put a star beside the assumption that most strongly affects the result.

Lecture 3: Model Fitting and Residual Analysis
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, readers should be able to:

1. distinguish **model fitting** from **interpolation**;
2. calculate and interpret a **residual**;
3. fit a proportional model and a straight-line model by the **least-squares method**;
4. linearise simple exponential or power relationships using transformed variables;
5. compare fitted models using graphs, residual patterns, error measures, and withheld data;
6. communicate why a numerically good fit may still be an unsuitable model.

From a proposed relationship to a fitted model

Core explanation, worked reasoning, and application

From a proposed relationship to a fitted model

In Lecture 2, physical reasoning suggested relationships such as

$$d_b \propto v^2$$

for braking distance and speed. Such a statement describes the *form* of a model, but it does not yet determine the proportionality constant. Real observations are also unlikely to lie exactly on a mathematical curve. We therefore need a systematic way to estimate parameters and judge whether the proposed relationship is credible.

Activity: Launch: Three lines, three stories

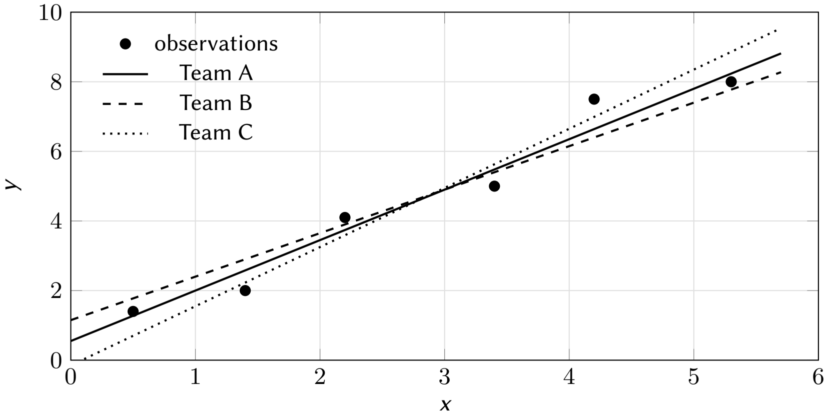
10 min

A set of six data points is shown with three possible fitted lines drawn by different teams.
In pairs, discuss:

1. What could "best fit" mean?
2. Should a positive error cancel a negative error?
3. Should one very inaccurate point matter more than several small inaccuracies?
4. What information would you need before choosing among the lines?

Modeling lesson: A fitted model is never selected by appearance or a single number alone. The criterion must match the purpose of the model.

Illustration: From a proposed relationship to a fitted model



Different fitting criteria can prefer different models. The purpose of fitting must be stated.

Concept: Fitting, interpolation, and prediction

- In **model fitting**, a model family is proposed from theory or previous reasoning, and its parameters are estimated from data. The curve need not pass through every point.
- In **interpolation**, a curve is constructed to pass through specified data points, usually to estimate values between them. This is the main subject of Lecture 4.
- A fitted model may be used for **prediction**, but prediction outside the observed range is **extrapolation** and requires additional caution.

Residuals and sources of discrepancy

Core explanation, worked reasoning, and application

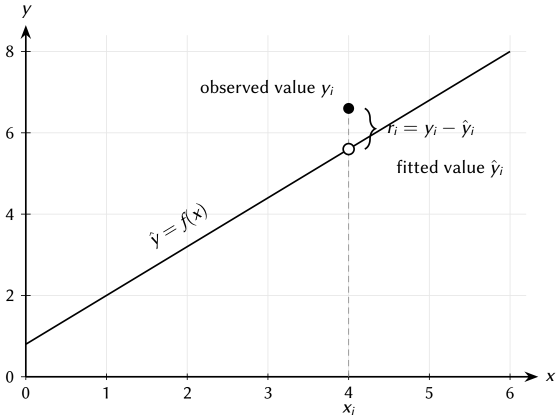
Residuals and sources of discrepancy

Suppose the observed value at x_i is y_i and the model predicts $\hat{y}_i = f(x_i)$. The **residual** is

$$\boxed{r_i = y_i - \hat{y}_i.}$$

A positive residual means the model underpredicts the observation; a negative residual means the model overpredicts it.

Illustration: Residuals and sources of discrepancy



At x_i, the residual is the signed vertical difference between the observed value y_i and the fitted value $\hat{y}_i = f(x_i)$.

Concept: A discrepancy is not always a fitting failure

Differences between data and model can arise from several sources:

1. **measurement error**: limited precision, recording error, or natural variation;
2. **formulation error**: omitted variables or an unsuitable functional form;
3. **parameter uncertainty**: the available data do not determine the parameters exactly;
4. **numerical error**: rounding or approximation during computation.

A very precise fitting algorithm cannot repair a poor model structure or unreliable data.

Checkpoint: Signs and units

A model predicts a stopping distance of 42.3 m, while the observed distance is 45.0 m.

1. Calculate the residual.
2. Does the model overpredict or underpredict?
3. What are the units of the residual?

Fitting a proportional model through the origin

Core explanation, worked reasoning, and application

Fitting a proportional model through the origin

If theory suggests

$$y = kx,$$

we should not automatically fit a general line $y = ax + b$. A nonzero intercept would contradict the proposed mechanism. For data (x_i, y_i), the least-squares value of k minimises

$$S(k) = \sum_{i=1}^{m} (y_i - kx_i)^2.$$

Differentiating and setting $S'(k) = 0$ gives

$$\boxed{k = \frac{\sum x_i y_i}{\sum x_i^2}.}$$

Worked Example: Worked Example 1: Calibrating the braking-distance model

Lecture 2 proposed $d_b = kv^2$, where v is speed in m/s and d_b is braking distance in m. Let

$$X = v^2.$$

Then the model becomes $d_b = kX$, a proportional relationship in the transformed variable X.

v (m/s)	5	10	15	20	25	30
$X = v^2$	25	100	225	400	625	900
d_b (m)	2.0	7.6	17.9	30.1	49.2	68.5

Using the least-squares formula,

$$k = \frac{\sum X_i d_i}{\sum X_i^2} = \frac{109277.5}{1421875} \approx 0.07685.$$

Therefore

$$\boxed{d_b \approx 0.07685v^2.}$$

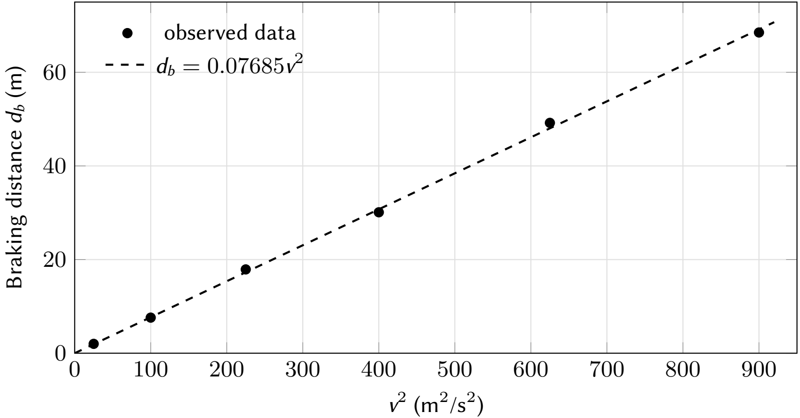
Worked Example: Worked Example 1: Calibrating the braking-distance model

Because $d_b = v^2/(2a)$ under constant deceleration, the fitted coefficient corresponds to

$$a \approx \frac{1}{2(0.07685)} \approx 6.51 \text{ m/s}^2.$$

The fitted parameter therefore has a physical interpretation rather than being merely a numerical constant.

Illustration: Fitting a proportional model through the origin



Plotting d_b against v^2 turns the proposed quadratic relationship into a line through the origin.

Model Critique: Should the intercept be forced to zero?

For $v = 0$, physical reasoning requires $d_b = 0$. Forcing the line through the origin is therefore justified. In another problem, a nonzero baseline may be real: an instrument can have an offset, or a delivery fee can include a fixed charge. The fitting constraint must come from the context, not from convenience.

Activity: Residual check

8 min

For $v = 25$ m/s, the fitted model predicts

$$\hat{d}_b = 0.07685(25)^2 \approx 48.03 \text{ m}.$$

The observation is 49.2 m.

1. Calculate the residual.
2. Explain whether the model underpredicts or overpredicts.
3. Suggest one physical reason for the discrepancy.

Fitting a general straight line

Core explanation, worked reasoning, and application

Fitting a general straight line

When a nonzero intercept is plausible, fit

$$y = ax + b.$$

The least-squares line minimises

$$S(a, b) = \sum_{i=1}^{m} (y_i - ax_i - b)^2.$$

A convenient form of the solution is

$$\boxed{a = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}, \qquad b = \bar{y} - a\bar{x}.}$$

Worked Example: Worked Example 2: Calibrating a force sensor

Known masses x in kg are placed on a sensor, producing output voltages y.

x (kg)	0	1	2	3	4	5
y (V)	0.18	1.72	3.19	4.76	6.25	7.81

Here $\bar{x} = 2.5$, $\bar{y} = 3.985$,

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = 26.655, \qquad \sum (x_i - \bar{x})^2 = 17.5.$$

Thus

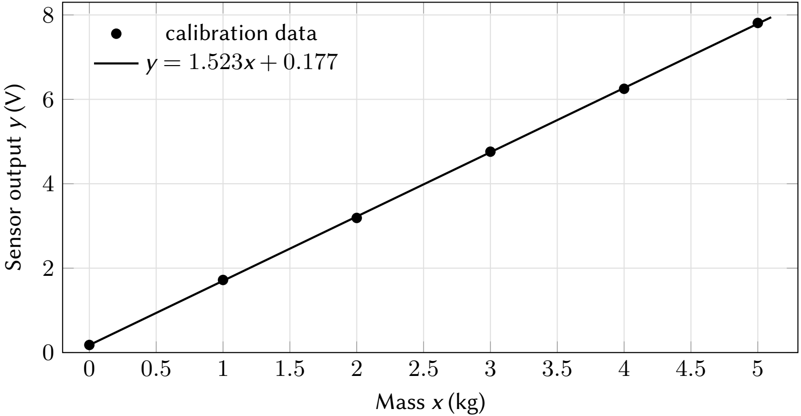
$$a \approx 1.523, \qquad b \approx 0.177,$$

and the calibration model is

$$\boxed{y \approx 1.523x + 0.177.}$$

The intercept is meaningful: the unloaded sensor reads about 0.18 V rather than zero. Forcing the model through the origin would hide this offset.

Illustration: Fitting a general straight line



The fitted intercept represents sensor offset. A parameter should be interpreted in the original context.

Concept: Why squares?

Least squares minimises

$$\mathrm{SSE} = \sum r_i^2,$$

the **sum of squared errors**. Squaring prevents positive and negative residuals from cancelling and penalises large residuals strongly. This makes the method mathematically convenient, but also sensitive to **outliers**.

Error measures and residual plots

Core explanation, worked reasoning, and application

Error measures and residual plots

No single error measure answers every question. Three common summaries are

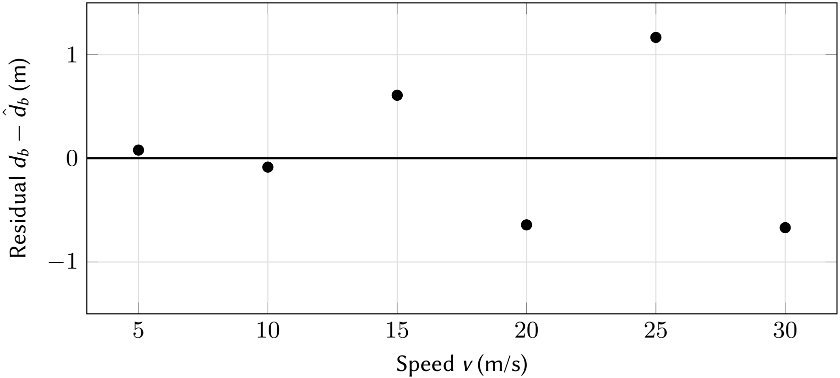
$$\mathsf{MAE} = \frac{1}{m}\sum |r_i|, \qquad \mathsf{RMSE} = \sqrt{\frac{1}{m}\sum r_i^2}, \qquad \max |r_i|.$$

The **mean absolute error** is easy to interpret; the **root mean squared error** penalises large errors more strongly; the maximum error focuses on the worst case. All three have the same units as the response variable. For the braking-distance fit,

$$\mathsf{MAE} \approx 0.54 \text{ m}, \qquad \mathsf{RMSE} \approx 0.66 \text{ m}, \qquad \max |r_i| \approx 1.17 \text{ m}.$$

These values describe the observed calibration range only. They do not prove that the model remains accurate at higher speeds or on different roads.

Illustration: Error measures and residual plots



Residuals should be examined for systematic patterns, not merely averaged into one statistic.

Checkpoint: Reading a residual plot

Suppose residuals are positive at low x, negative in the middle, and positive again at high x.

1. Is the scatter random?
2. What does the curved pattern suggest about a fitted straight line?
3. Would collecting more decimal places solve the main problem?

Linearising an exponential model

Core explanation, worked reasoning, and application

Linearising an exponential model

Some model families become linear after a transformation. For an exponential decay model

$$C = Ae^{-kt},$$

taking natural logarithms gives

$$\ln C = \ln A - kt.$$

Thus a plot of $\ln C$ against t should be approximately linear, with slope $-k$ and intercept $\ln A$.

Worked Example: Worked Example 3: Battery capacity decay

A battery is repeatedly tested under the same load. Capacity C is recorded as a percentage of its initial value.

Cycle block t	0	1	2	3	4	5
Capacity C (%)	100	82	68	55	46	37

A least-squares line fitted to $(t, \ln C)$ is approximately

$$\ln C = 4.607 - 0.19765t.$$

Exponentiating gives

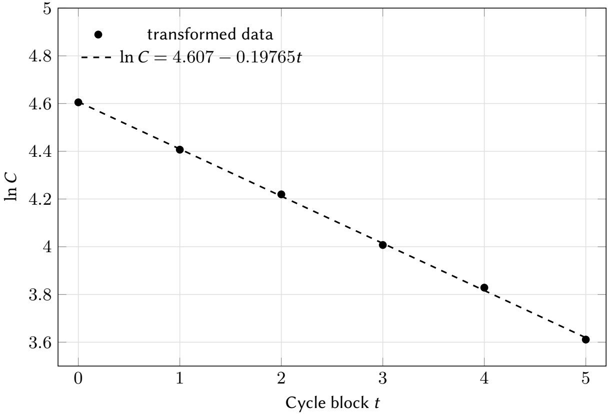
$$\boxed{C \approx 100.2e^{-0.19765t}.}$$

The fitted model predicts that capacity is multiplied by

$$e^{-0.19765} \approx 0.821$$

per cycle block, corresponding to an average decrease of about 17.9% of the remaining capacity each block.

Illustration: Linearising an exponential model



Linearisation makes parameter estimation convenient, but the resulting model must be checked in the original capacity scale.

Model Critique: Transformation changes the fitting criterion

Least squares in $\ln C$ minimises squared differences between logarithms, not squared differences between capacities. It gives greater relative importance to smaller values. Therefore:

1. transform only when the model form and data permit it;
2. convert the fitted model back to the original variables;
3. calculate residuals and error measures in the original scale;
4. compare alternative models using the same response variable.

Activity: Choose a transformation

7 min

State the variables that should be plotted to test each proposed model.

1. $y = Ae^{bx}$
2. $y = Ax^n$
3. $y = A/x$

For each case, identify the slope or intercept that has a parameter interpretation.

Choosing between models: fit, structure, and validation

Core explanation, worked reasoning, and application

Choosing between models: fit, structure, and validation

A model should not be selected only because it has the smallest calibration error. We also ask whether it respects the mechanism, whether residuals show structure, and whether it predicts data that were not used to fit the parameters.

Worked Example: Worked Example 4: Linear or exponential decay?

Using the six battery observations from $t = 0$ to $t = 5$:

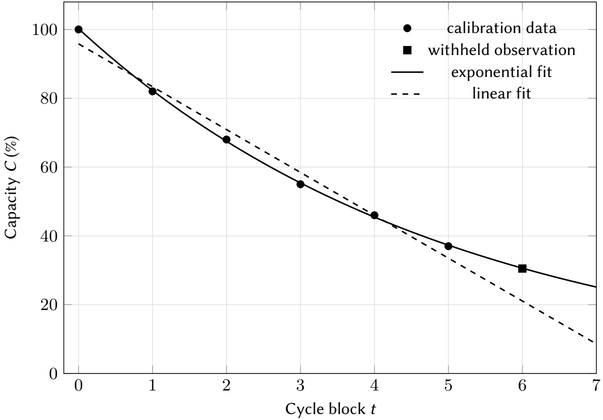
- the exponential fit is $C \approx 100.2e^{-0.19765t}$ with calibration RMSE about 0.39 percentage points;
- the least-squares straight line is $C \approx 95.81 - 12.46t$ with calibration RMSE about 2.93 percentage points.

The value at $t = 6$, $C = 30.5\%$, is withheld during fitting and used for **validation**.

$$\hat{C}_{\text{exp}}(6) \approx 30.61, \qquad \hat{C}_{\text{lin}}(6) \approx 21.07.$$

The exponential model performs much better on the withheld point and never predicts negative capacity. Both the numerical evidence and the mechanism support exponential decay over this range.

Illustration: Choosing between models: fit, structure, and validation



Calibration error, validation performance, and long-term plausibility all favour the exponential model.

Concept: A practical model-selection checklist

Before accepting a fitted model, ask:

1. **Mechanism:** Does the functional form agree with the assumptions or governing relationships?
2. **Scale:** Were errors assessed in the original response variable?
3. **Residuals:** Are they small and free of systematic pattern?
4. **Validation:** Does the model predict observations not used for fitting?
5. **Range:** Is the intended prediction inside or outside the observed range?
6. **Interpretability:** Do the parameters and conclusions make sense in context?

Activity: Competition-style communication

10 min

Write a four-sentence model-selection paragraph for the battery example:

1. identify the two models compared;
2. report one calibration measure;
3. explain the withheld-data result;
4. state one limitation or condition on the conclusion.

Avoid saying that a high-quality fit “proves” the model.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Seven ideas to retain

1. Model fitting estimates parameters within a proposed model family; interpolation is a different task.
2. A residual is observed minus predicted, with the same units as the response variable.
3. The fitting constraint—through the origin or with an intercept—must be justified by context.
4. Least squares minimises squared residuals and is sensitive to large errors and outliers.
5. Residual plots reveal systematic structure that a single error statistic can hide.
6. Transformations simplify fitting but alter the error criterion; always return to the original variables.
7. A credible model combines mechanism, calibration, validation, interpretation, and stated limitations.

Checkpoint: Exit ticket

2 min

Complete both statements:

1. "A small RMSE is not sufficient because ____."
2. "I would force a fitted line through the origin only when ____."

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A model predicts water demand of 82, 91, and 106 L, while the corresponding observations are 85, 88, and 110 L.
   1.1 Calculate the three residuals.
   1.2 Find the MAE and RMSE.
   1.3 Which observation contributes most strongly to the SSE?
2. The data below relate force F to extension x for a spring:

x (cm)	1	2	3	4	5
F (N)	1.9	4.2	6.0	8.1	10.3

   Fit $F = kx$ by least squares. Interpret the units of k and calculate the residual at $x = 4$ cm.
3. A temperature sensor gives readings

x (actual °C)	0	10	20	30	40
y (sensor °C)	1.2	10.8	20.5	30.6	40.1

   Explain why an intercept should be allowed. Fit a straight line using a spreadsheet or calculator and interpret both parameters.

Exercises: Core practice

4. State the transformation for each model and explain what a straight transformed plot would support:
   4.1 $y = Ae^{bx}$;
   4.2 $y = Ax^n$;
   4.3 $y = A/(B + x)$.
5. Two models have the following residuals, measured in minutes:

$$A : (-1.0, -0.8, 0.2, 0.4, 1.2), \qquad B : (-0.2, -0.1, 0.0, 0.1, 3.0).$$

   Compare their MAE, RMSE, and maximum error qualitatively. Which model would you prefer for average performance? Which for avoiding a severe worst case?
6. A fitted quadratic has residuals that alternate randomly around zero, while a fitted line has residuals forming a clear U-shape. Explain what this suggests. Why is "the quadratic has more parameters" not by itself a sufficient conclusion?
7. A population model is fitted using years 2018–2024 and then used to predict 2045. Explain why a small training RMSE does not establish that the 2045 forecast is trustworthy. Name two additional checks.

Exercises: Modeling challenge

Choose one dataset that your team can collect or obtain: cooling temperature, queue waiting time, plant growth, device battery decay, transport travel time, or another approved context.
Prepare a one-page fitting brief containing:

1. the question, variables, units, and proposed model family;
2. a justification for allowing or excluding an intercept;
3. fitted parameter values and one graph in the original variables;
4. a residual plot and at least two error measures;
5. one validation strategy and one limitation;
6. a short conclusion written for a non-specialist decision maker.

The goal is not to obtain the lowest possible error by adding terms. The goal is to produce a defensible model.

Optional derivation: the least-squares line

Core explanation, worked reasoning, and application

Optional derivation: the least-squares line

This appendix is not part of the two-hour core lesson. It may be assigned to students who are comfortable with partial differentiation.

For $y = ax + b$, minimise

$$S(a, b) = \sum_{i=1}^{m} (y_i - ax_i - b)^2.$$

Setting both partial derivatives equal to zero gives

$$\frac{\partial S}{\partial a} = -2 \sum x_i (y_i - ax_i - b) = 0,$$
$$\frac{\partial S}{\partial b} = -2 \sum (y_i - ax_i - b) = 0.$$

Therefore

$$a \sum x_i^2 + b \sum x_i = \sum x_i y_i,$$
$$a \sum x_i + mb = \sum y_i.$$

These are the **normal equations**. Solving them is equivalent to the mean-centred formulas used in the lesson.

Optional extension: alternatives to least squares

Core explanation, worked reasoning, and application

Optional extension: alternatives to least squares

Least squares is not the only defensible fitting criterion.

- The L^1 criterion minimises $\sum |r_i|$ and is less sensitive to a single large outlier.
- The L^∞ or **minimax criterion** minimises $\max |r_i|$ and is useful when the worst error matters most.

The appropriate criterion depends on the decision. A medical safety model, a forecasting model, and an engineering calibration problem may not value errors in the same way.

Lecture 4: Empirical Models, Interpolation, and Splines

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Learning Goals

By the end of the lesson, readers should be able to:

1. distinguish **interpolation**, **smoothing**, and **extrapolation**;
2. construct and interpret a linear interpolant between two observations;
3. build a quadratic interpolating polynomial in Lagrange form;
4. use a difference table to recognise simple polynomial structure;
5. explain why an exact high-degree interpolant may behave poorly between data points;
6. compare piecewise linear interpolation with a natural cubic spline and select an appropriate empirical method.

When the data come before the formula

Core explanation, worked reasoning, and application

When the data come before the formula

Lecture 3 began with a proposed model family and estimated its parameters from data. In many practical situations, however, no convincing mechanism is yet available. We may only have measurements at selected times or locations and need a reasonable value between them. This leads to an **empirical model**: a mathematical representation constructed mainly from observed data.

Activity: Launch: Should the curve pass through every point?

A classroom sensor records carbon-dioxide concentration once every five minutes:

time (min)	0	5	10	15	20
CO_2 (ppm)	612	648	701	694	760

In pairs, discuss:

1. How would you estimate the concentration at minute 12?
2. Should the estimated curve pass through every recorded value?
3. Which matters more here: measurement noise or exact agreement with the observations?
4. Would the same method be defensible for predicting minute 60?

Modeling lesson: The purpose of the model and the reliability of the data determine whether exact interpolation is desirable.

Concept: Interpolation, smoothing, and extrapolation

- **Interpolation** constructs a function that agrees with specified data values and estimates values *inside* the observed range.
- **Smoothing** allows small deviations from noisy data to reveal an underlying trend.
- **Extrapolation** predicts *outside* the observed range. It depends on assumptions that the data alone cannot verify.

An interpolant can be mathematically exact and still be scientifically untrustworthy if the observations contain noise or the curve behaves implausibly between them.

Model Critique: Exact agreement is not the same as credibility

If a thermometer is accurate only to the nearest degree, forcing a complicated curve through every reported value may model rounding noise rather than temperature. Conversely, if the data are exact geometric coordinates in computer graphics, passing through every point may be essential. The data-generating process must guide the method.

Linear interpolation and piecewise linear models

Core explanation, worked reasoning, and application

Linear interpolation and piecewise linear models

Suppose two trusted observations are (x_0, y_0) and (x_1, y_1), with $x_0 < x_1$. For a point x between them, define the relative position

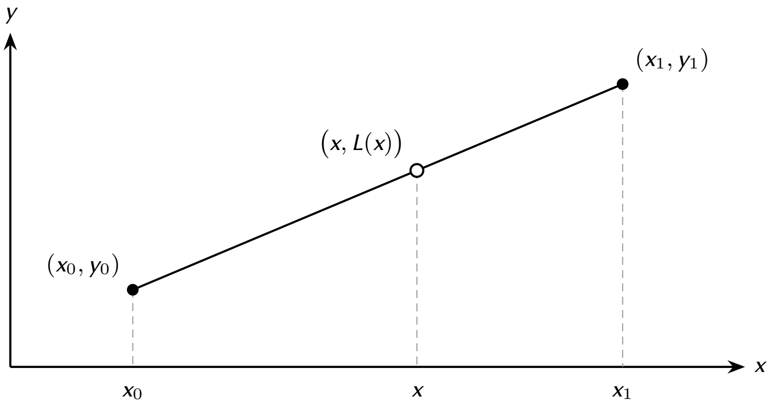
$$\lambda = \frac{x - x_0}{x_1 - x_0}, \qquad 0 \leq \lambda \leq 1.$$

The **linear interpolant** is

$$\boxed{L(x) = (1-\lambda)y_0 + \lambda y_1 = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0).}$$

The formula is a weighted average: when x is closer to x_0, more weight is placed on y_0.

Illustration: Linear interpolation and piecewise linear models



The interpolated point lies on the line segment joining the two observations; its horizontal position determines the weights assigned to y_0 and y_1.

Worked Example: Worked Example 1: Estimating temperature between readings

A heating experiment records $38.0°\mathrm{C}$ at $t = 4$ minutes and $62.0°\mathrm{C}$ at $t = 10$ minutes. Estimate the temperature at $t = 7.5$ minutes.
The relative position is

$$\lambda = \frac{7.5 - 4}{10 - 4} = \frac{3.5}{6}.$$

Hence

$$T(7.5) = 38.0 + \frac{3.5}{6}(62.0 - 38.0) = 38.0 + 14.0 = \boxed{52.0°\mathrm{C}}.$$

The estimate assumes that temperature changed at an approximately constant rate over this six-minute interval. It does not claim that the entire heating process is linear.

Linear interpolation and piecewise linear models

With several observations, we may connect each consecutive pair by a line. The resulting **piecewise linear interpolant** is continuous and simple to compute, but its slope changes abruptly at the data points.

Activity: Guided practice: Build a piecewise model

A library records the number of occupied seats:

$$(0, 20), \qquad (2, 44), \qquad (5, 68),$$

where x is hours after opening.

1. Write the linear formula on $0 \leq x \leq 2$.
2. Write the linear formula on $2 \leq x \leq 5$.
3. Estimate occupancy at $x = 1.5$ and $x = 3.5$.
4. Compare the slopes of the two pieces. What does the change of slope mean in context?

Checkpoint: Interpolation or extrapolation?

For the temperature observations at $t = 4$ and $t = 10$ minutes, classify each request:

1. estimate $T(8)$;
2. estimate $T(15)$;
3. estimate the maximum temperature after one hour.

Explain why only one request can be answered by linear interpolation alone.

Quadratic interpolation through three points

Core explanation, worked reasoning, and application

Quadratic interpolation through three points

Two points determine a line; three points with distinct x-coordinates determine a unique polynomial of degree at most two. The **Lagrange form** constructs this polynomial directly.
For three points (x_0, y_0), (x_1, y_1), and (x_2, y_2),

$$P_2(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x),$$

where

$$\begin{aligned} L_0(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}, \\ L_1(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}, \\ L_2(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}. \end{aligned}$$

Each basis polynomial satisfies $L_i(x_i) = 1$ and is zero at the other two nodes.

Worked Example: Worked Example 2: A quadratic interpolant

Construct the polynomial through

$$(0,1), \qquad (1,3), \qquad (2,2).$$

The basis functions are

$$L_0(x) = \frac{(x-1)(x-2)}{2}, \qquad L_1(x) = -x(x-2), \qquad L_2(x) = \frac{x(x-1)}{2}.$$

Therefore

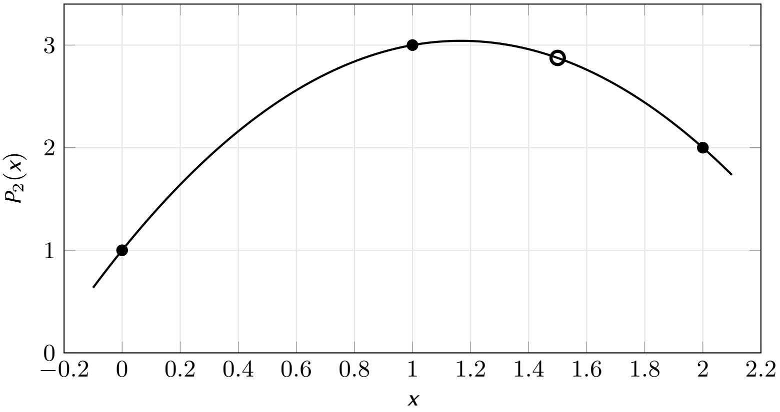
$$\begin{aligned} P_2(x) &= L_0(x) + 3L_1(x) + 2L_2(x) \\ &= -\frac{3}{2}x^2 + \frac{7}{2}x + 1. \end{aligned}$$

At $x = 1.5$,

$$P_2(1.5) = \boxed{2.875}.$$

Substitution verifies $P_2(0) = 1$, $P_2(1) = 3$, and $P_2(2) = 2$ exactly.

Illustration: Quadratic interpolation through three points



The solid curve is $P_2(x) = -1.5x^2 + 3.5x + 1$. Filled points are the interpolation nodes; the open point is the estimate at $x = 1.5$.

Model Critique: What does the curve between the points mean?

The interpolation conditions determine the quadratic uniquely, but they do not prove that the real process is quadratic. A different physical mechanism could produce the same three measurements and different values between them. Interpolation supplies a convenient curve; scientific interpretation still requires context.

Difference tables as a diagnostic

Core explanation, worked reasoning, and application

Difference tables as a diagnostic

For equally spaced values of x, successive differences can reveal low-degree polynomial structure:

- constant first differences suggest a linear model;
- constant second differences suggest a quadratic model;
- constant third differences suggest a cubic model.

The result is exact for perfect polynomial data and only approximate for noisy observations.

Worked Example: Worked Example 3: Recognising a quadratic pattern

Consider

x	0	1	2	3	4
y	2	5	10	17	26

The difference table is

y	2	5	10	17	26
Δy		3	5	7	9
$\Delta^2 y$			2	2	2

The second differences are constant, so the data lie on a quadratic. For unit spacing, the leading coefficient is half the constant second difference, giving $a = 1$. Writing

$$y = x^2 + bx + c$$

and using $y(0) = 2$ and $y(1) = 5$ gives $c = 2$ and $b = 2$. Thus

$$\boxed{y = x^2 + 2x + 2.}$$

Concept: What a difference table can and cannot do

A difference table is a diagnostic, not a certificate of truth. With measured data, higher-order differences often amplify noise. The method is most useful for identifying a simple pattern in accurate, equally spaced data. For unequally spaced nodes, use **divided differences**; an optional introduction appears in Appendix 10.

Activity: Pattern or noise?

8 min

For each sequence, calculate enough differences to suggest a simple model.

1. $4,\ 9,\ 16,\ 25,\ 36$
2. $3.0,\ 6.1,\ 11.9,\ 21.2,\ 34.0$

For the second sequence, explain why "approximately quadratic" is more defensible than "exactly quadratic."

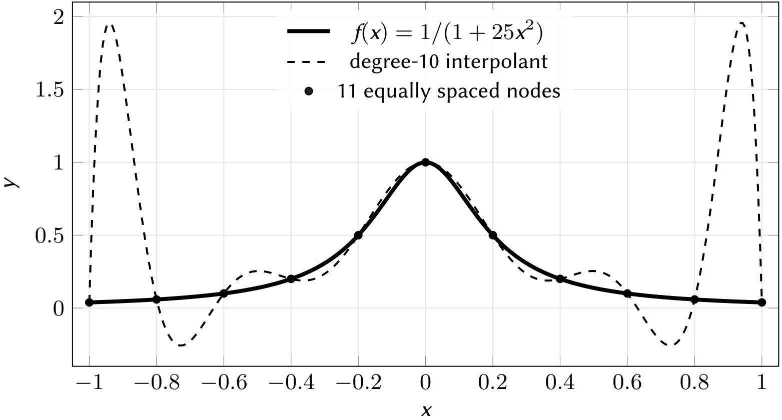
Why a high-degree polynomial can fail

Core explanation, worked reasoning, and application

Why a high-degree polynomial can fail

Given $n + 1$ distinct points, a unique polynomial of degree at most n passes through all of them. This theorem guarantees existence, not good behavior. As the degree increases, a global polynomial can oscillate strongly between nodes, especially near the endpoints.

Illustration: Why a high-degree polynomial can fail



The polynomial matches every node but oscillates badly between them. This is **Runge's phenomenon**.

Model Critique: The zero-residual trap

At the interpolation nodes, every residual is zero. Therefore SSE, MAE, and RMSE computed only at those nodes are also zero, even though the curve is poor between them. Model assessment must consider behavior between observations, sensitivity to data perturbations, and the intended use of the curve.

Why a high-degree polynomial can fail

Possible remedies include using fewer degrees of freedom, smoothing noisy data, choosing better-distributed nodes, or replacing one global polynomial with piecewise low-degree polynomials.

Piecewise interpolation and cubic splines

Core explanation, worked reasoning, and application

Piecewise interpolation and cubic splines

A **spline** uses a separate polynomial on each interval between consecutive **knots**. A linear spline joins the points by straight segments. A **cubic spline** uses cubic pieces while imposing smooth joins.
For a cubic spline S:

1. S passes through every data point;
2. adjacent pieces have the same value at each interior knot;
3. their first derivatives agree, so there is no corner;
4. their second derivatives agree, so curvature changes smoothly.

A **natural cubic spline** additionally satisfies $S'' = 0$ at the two endpoints.

Worked Example: Worked Example 4: Natural spline through three points

Construct the natural cubic spline through

$$(0,0), \qquad (1,1), \qquad (2,0).$$

Let $M_i = S''(x_i)$. Natural endpoint conditions give $M_0 = M_2 = 0$. With equal spacing $h = 1$, the single interior spline equation is

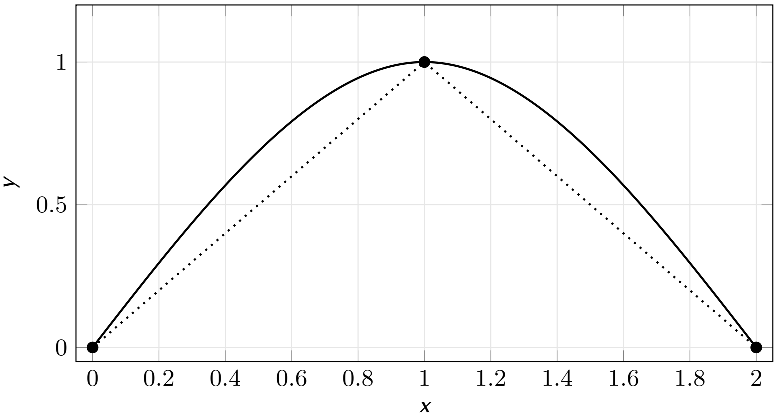
$$M_0 + 4M_1 + M_2 = 6(y_2 - 2y_1 + y_0) = -12,$$

so $M_1 = -3$. The resulting spline is

$$S(x) = \begin{cases} -\frac{1}{2}x^3 + \frac{3}{2}x, & 0 \le x \le 1, \\ 1 - \frac{3}{2}(x-1)^2 + \frac{1}{2}(x-1)^3, & 1 \le x \le 2. \end{cases}$$

At $x = 0.5$, the spline gives $S(0.5) = 0.6875$, while linear interpolation gives 0.5. Neither value is automatically "correct"; the smoother cubic estimate encodes stronger assumptions about how the slope changes.

Illustration: Piecewise interpolation and cubic splines



The dotted line is the linear spline, the solid curve is the natural cubic spline, and the filled points are the knots. Both interpolate the same data, but only the cubic spline has continuous slope and curvature at $x = 1$.

Choosing an empirical method

Core explanation, worked reasoning, and application

Choosing an empirical method

Method	Strength	Main caution
Linear interpolation	Transparent, local, easy to compute	Corners at knots; assumes constant rate within each interval
Quadratic or low-degree interpolation	Captures curvature with few points	A small dataset does not prove the polynomial mechanism
High-degree global polynomial	Passes through every point	Oscillation, sensitivity, and poor extrapolation
Low-degree smoothing model	Handles noisy data and reveals trend	Does not reproduce every observation exactly
Cubic spline	Local, smooth, accurate between many trusted points	May overshoot; boundary conditions matter

Checkpoint: Which property matters?

A road profile is measured at survey points and will be used to estimate vehicle motion. Would you prefer a linear spline or a cubic spline? State one benefit and one risk of your choice.

Model Critique: Cubic splines are not automatically shape-preserving

A cubic spline is smooth and usually stable, but it may overshoot between data points or create negative values when the modeled quantity must remain nonnegative. For monotone or bounded data, specialised shape-preserving splines may be more appropriate.

Activity: Method-selection studio

Choose an appropriate method for each situation and justify it in one or two sentences.

1. Exact coordinates define the edge of a manufactured part.
2. Daily demand data contain substantial random variation, and the goal is to reveal the weekly trend.
3. A temperature sensor failed for two minutes between two reliable readings.
4. Ten trusted measurements describe a smooth road elevation profile.
5. A team wants to forecast ten years beyond a five-year dataset.

For the final situation, explain why none of the interpolation methods is sufficient by itself.

Concept: A practical empirical-modeling workflow

1. Clarify whether the task is interpolation, smoothing, or extrapolation.
2. Plot the original data and check units, ordering, missing values, and outliers.
3. Decide whether observations should be treated as exact or noisy.
4. Start with the simplest local method that meets the purpose.
5. Inspect the curve between points, not only at the observations.
6. Test sensitivity to removing or perturbing a data point.
7. State the valid interval and any shape constraints explicitly.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Six ideas to retain

1. Interpolation estimates inside the data range; extrapolation requires additional assumptions.
2. Linear interpolation is a weighted average between two observations.
3. Three distinct nodes determine a unique quadratic, but not necessarily the true mechanism.
4. Difference tables can reveal simple polynomial patterns in equally spaced data.
5. Zero residuals at interpolation nodes do not guarantee good behavior between them.
6. Splines replace one unstable high-degree polynomial with smooth, local, low-degree pieces.

Checkpoint: Exit ticket

Complete both statements:

1. "I would interpolate rather than smooth when ____."
2. "A curve passing through every point can still be poor because ____."

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A sensor reads 18.2 units at $t = 3$ and 25.4 units at $t = 9$.
   1.1 Use linear interpolation to estimate the value at $t = 5.5$.
   1.2 State the assumption behind the calculation.
   1.3 Explain why the same formula should not automatically be used at $t = 15$.
2. Construct the piecewise linear interpolant through $(0, 2)$, $(2, 6)$, and $(5, 3)$. Evaluate it at $x = 1$, $x = 3$, and $x = 4.5$. Identify the point where the slope changes.
3. Find the quadratic interpolating polynomial through $(0, 2)$, $(1, 1)$, and $(3, 5)$ using the Lagrange form. Evaluate the polynomial at $x = 2$ and verify all three interpolation conditions.
4. Construct a difference table for
$$7,\ 12,\ 19,\ 28,\ 39,\ 52.$$
   Suggest a polynomial degree and find a formula in terms of $n = 0, 1, 2, \ldots$.
5. A degree-12 polynomial passes through 13 experimental measurements exactly. Give three reasons why a lower-degree smoother or spline may nevertheless be more credible.
6. For the points $(0, 0)$, $(1, 2)$, and $(2, 0)$:
   6.1 write the linear spline;

Exercises: Core practice

6.2 describe the continuity properties required of a cubic spline;
6.3 without calculating the full cubic spline, predict how its graph would differ near $x = 1$.

7. Decide whether interpolation, smoothing, or neither is the main task:

   7.1 estimating rainfall at an unmonitored hour between two observations;
   7.2 finding the long-run trend in noisy monthly sales;
   7.3 predicting population in 2070 using records ending in 2025;
   7.4 reconstructing a computer-drawn curve from exact control points.

Exercises: Modeling challenge

A school records pedestrian flow at a corridor entrance every five minutes, but several readings are missing. Some recorded values may also contain counting errors.
Prepare a one-page empirical-modeling brief containing:

1. the purpose of reconstructing the missing data;
2. which observations you would treat as reliable or noisy;
3. a comparison of piecewise linear interpolation, cubic splines, and smoothing;
4. one diagnostic graph or calculation;
5. the interval on which your result is valid;
6. one sensitivity test and one limitation.

The quality of the justification matters more than using the most complicated method.

Optional extension: Newton divided differences

Core explanation, worked reasoning, and application

Optional extension: Newton divided differences

For unequally spaced nodes, ordinary differences are replaced by **divided differences**:

$$f[x_i, x_{i+1}] = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i},$$

$$f[x_i, \ldots, x_{i+k}] = \frac{f[x_{i+1}, \ldots, x_{i+k}] - f[x_i, \ldots, x_{i+k-1}]}{x_{i+k} - x_i}.$$

The Newton interpolating polynomial is

$$P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + \cdots + f[x_0, \ldots, x_n] \prod_{j=0}^{n-1} (x - x_j).$$

Unlike the Lagrange form, the Newton form can be extended efficiently when a new data point is added.

Worked Example: Unequally spaced data

For $(0, 1)$, $(1, 3)$, and $(3, 7)$,

$$f[0, 1] = 2, \qquad f[1, 3] = 2,$$

so

$$f[0, 1, 3] = \frac{2 - 2}{3 - 0} = 0.$$

Hence $P_2(x) = 1 + 2x$, showing that the three points are actually collinear even though the spacing is unequal.

Optional extension: General natural-spline system

Core explanation, worked reasoning, and application

Optional extension: General natural-spline system

Let $x_0 < \cdots < x_n$, $h_i = x_{i+1} - x_i$, and $M_i = S''(x_i)$. For a natural cubic spline, $M_0 = M_n = 0$, and the interior second derivatives satisfy

$$h_{i-1}M_{i-1} + 2(h_{i-1} + h_i)M_i + h_iM_{i+1} = 6\left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right).$$

This is a tridiagonal linear system, so a spline with many knots can be computed efficiently. The formula is useful for implementation but is not part of the two-hour core lesson.

Lecture 5: Monte Carlo and Discrete-Event Simulation

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Learning Goals

By the end of the lesson, readers should be able to:

1. distinguish a **deterministic model** from a **stochastic model**;
2. map uniform random numbers to outcomes from a discrete or empirical probability distribution;
3. construct and interpret a simple **Monte Carlo estimator**;
4. explain why repeated simulation runs are needed and why simulation error decreases slowly;
5. trace a single-server queue using arrival, service-start, waiting, and departure times;
6. design a simulation experiment that includes outputs, replications, validation, sensitivity, and limitations.

Why simulate?

Core explanation, worked reasoning, and application

Why simulate?

Previous lectures represented a system using equations or curves. That approach remains valuable, but some systems contain uncertain events whose exact sequence cannot be predicted: customers arrive at irregular times, service durations vary, and equipment may fail unexpectedly. A **simulation** imitates the operation of such a system and records what happens under specified assumptions.

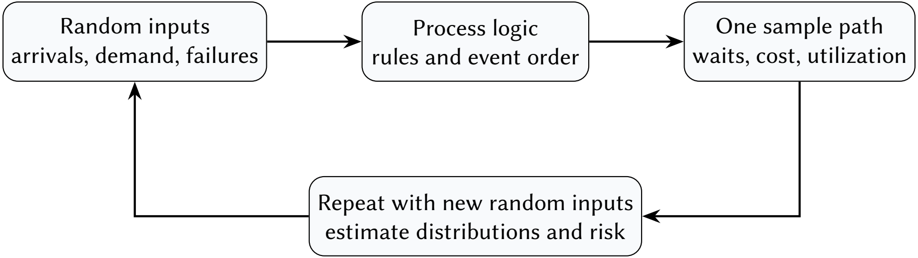
Activity: Launch: The canteen queue

At lunchtime, students arrive at one canteen counter at irregular times. Some orders take much longer than others. The school is considering opening a second counter.
In pairs, identify:

1. two quantities that vary unpredictably from student to student;
2. two quantities that describe the current *state* of the system;
3. three outputs that the school should examine before making a decision;
4. one factor that might make a simulation unrealistic.

Modeling lesson: A simulation requires more than random numbers. It needs explicit input assumptions, update rules, outputs, and a plan for checking credibility.

Illustration: Why simulate?



A simulation combines random inputs with deterministic update rules. One run gives one possible history; repeated runs reveal the range of possible outcomes.

Concept: Core simulation language

- A **state variable** describes the system at a given time, such as queue length or inventory level.
- An **event** changes the state, such as an arrival, a service completion, or a delivery.
- A **sample path** is one simulated history of the system.
- A **replication** is an independent run under the same policy and assumptions.
- A **random seed** reproduces the same pseudorandom sequence, which is useful for debugging and fair comparisons.

The computer performs deterministic calculations after the random inputs have been generated. Simulation is therefore structured experimentation, not random guessing.

From uniform random numbers to model inputs

Core explanation, worked reasoning, and application

From uniform random numbers to model inputs

A **random variable** assigns a numerical value to each possible outcome of a random experiment. To simulate a discrete random variable, partition the interval $[0, 1)$ according to cumulative probabilities and generate $U \sim \mathrm{Uniform}(0, 1)$.
Suppose a service time S has possible values $s_1, \ldots, s_k$ with probabilities $p_1, \ldots, p_k$. Let

$$F_j = p_1 + \cdots + p_j.$$

Assign $S = s_j$ when $F_{j-1} \leq U < F_j$, where $F_0 = 0$. This is the discrete **inverse transform method**.

Worked Example: Worked Example 1: Simulating service times

Historical observations at a snack counter suggest

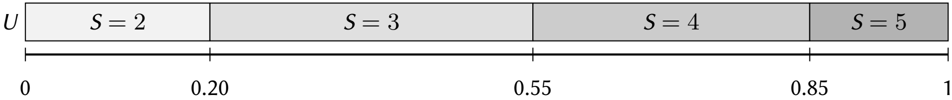
service time S (min)	2	3	4	5
$P(S = s)$	0.20	0.35	0.30	0.15

The cumulative probabilities are 0.20, 0.55, 0.85, and 1.00, so

$$S = \begin{cases} 2, & 0 \leq U < 0.20, \\ 3, & 0.20 \leq U < 0.55, \\ 4, & 0.55 \leq U < 0.85, \\ 5, & 0.85 \leq U < 1. \end{cases}$$

For $U = 0.07, 0.48, 0.61, 0.93$, the simulated service times are $2, 3, 4, 5$ minutes respectively.

Illustration: From uniform random numbers to model inputs



The length of each interval equals the probability of its outcome. A larger probability occupies more of $[0, 1)$.

Activity: Guided practice: Interarrival times

Student interarrival times have the distribution

A (min)	2	3	4	5
$P(A = a)$	0.15	0.35	0.35	0.15.

1. Construct the cumulative probability table and the interval mapping.
2. Convert $U = 0.08, 0.32, 0.71, 0.94, 0.51$ into interarrival times.
3. Explain why a fixed seed is helpful when comparing one counter with two counters.

Model Critique: The input distribution is part of the model

A simulation can produce precise-looking output from a poor input model. Historical frequencies may be unrepresentative of examination days, special menus, or changing student behavior. The probabilities, dependence between variables, and time period represented by the data must all be justified.

Monte Carlo estimation

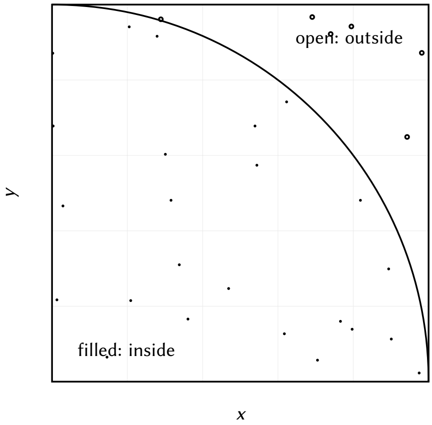
Core explanation, worked reasoning, and application

Monte Carlo estimation

A **Monte Carlo simulation** estimates a quantity by repeating a random experiment. Consider points drawn uniformly from the unit square. The probability of falling inside the quarter circle $x^2 + y^2 \leq 1$ is its area, $\pi/4$. If C of n points fall inside, then

$$\frac{C}{n} \approx \frac{\pi}{4}, \qquad \boxed{\widehat{\pi} = 4\frac{C}{n}.}$$

Illustration: Monte Carlo estimation



Thirty reproducible sample points: 24 fall inside the quarter circle and 6 outside. The picture represents one sample path, not a guarantee of accuracy.

Worked Example: Worked Example 2: One estimate of π

In Figure 3, $C = 24$ and $n = 30$, so

$$\widehat{\pi} = 4\left(\frac{24}{30}\right) = \boxed{3.20}.$$

The estimate is plausible but rough. A different seed would produce a different count. The value 3.20 should therefore be reported together with n, the random-number procedure, and preferably results from repeated runs.

Monte Carlo estimation

For independent samples, Monte Carlo uncertainty usually decreases at the rate

$$\text{typical error} \propto \frac{1}{\sqrt{n}}.$$

Thus reducing the typical error by a factor of 2 requires about 4 times as many samples; reducing it by a factor of 10 requires about 100 times as many.

Checkpoint: Convergence

A simulation uses $2{,}500$ independent samples. Approximately how many samples are required to reduce its typical Monte Carlo error to one-fifth of the current value? Explain using the square-root rule.

Concept: Replication and uncertainty

Increasing the number of observations within one run and increasing the number of independent replications answer different questions. A long run may estimate steady performance; several replications reveal run-to-run variability. Report at least a mean and a measure of spread, and retain the individual replication results rather than only the grand average.

Discrete-event simulation of a queue

Core explanation, worked reasoning, and application

Discrete-event simulation of a queue

In a **discrete-event simulation**, the state changes only when an event occurs. For a single-server queue, let

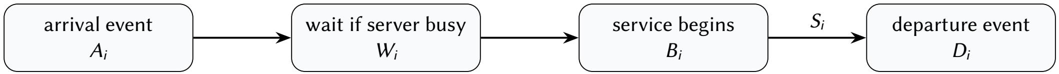
- A_i be the arrival time of customer i;
- S_i be the required service time;
- B_i be the service-start time;
- D_i be the departure time;
- $W_i = B_i - A_i$ be the waiting time.

The update equations are

$$\boxed{B_i = \max(A_i, D_{i-1}), \qquad D_i = B_i + S_i, \qquad W_i = B_i - A_i,}$$

with $D_0 = 0$. The maximum expresses the queue logic: service begins only after both the customer has arrived and the server is free.

Illustration: Discrete-event simulation of a queue



A queue is governed by event order. Randomness enters through arrival and service times; the update logic is deterministic.

Worked Example: Worked Example 3: Manual trace of one counter

Six customers have arrival times

$$A_i = (0, 3, 5, 10, 12, 16)$$

and service times

$$S_i = (4, 5, 3, 4, 2, 5)$$

in minutes. Applying the update equations gives:

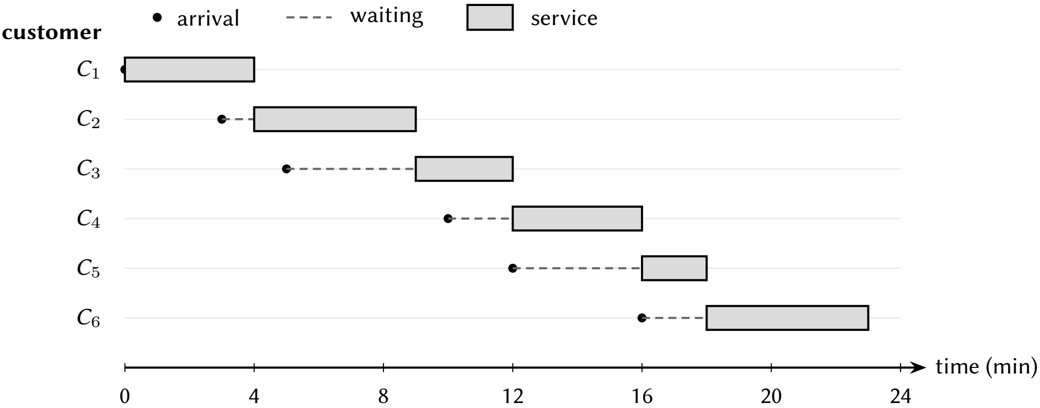
Customer i	1	2	3	4	5	6
Arrival A_i	0	3	5	10	12	16
Service S_i	4	5	3	4	2	5
Start B_i	0	4	9	12	16	18
Wait W_i	0	1	4	2	4	2
Departure D_i	4	9	12	16	18	23

The average waiting time is

$$\bar{W} = \frac{0 + 1 + 4 + 2 + 4 + 2}{6} = \boxed{2.17 \text{ min}},$$

and the maximum waiting time is 4 minutes. These values describe only this six-customer sample path.

Illustration: Discrete-event simulation of a queue



Customer-centred queue timeline. A filled point marks arrival, a dashed segment shows waiting, and a shaded block shows the service interval.

Activity: Policy comparison: Add a second counter

Use the same arrivals and service times as Worked Example 3. Assign each arriving customer to the counter that becomes available first.

1. Complete a table containing counter number, start time, waiting time, and departure time.
2. Find the average and maximum waiting times.
3. Explain why this single sample path is insufficient evidence for permanently staffing a second counter.

From one run to a decision

Core explanation, worked reasoning, and application

From one run to a decision

A simulation study is an experiment on a model. Before running it, define the policy alternatives, input distributions, performance measures, simulation length, number of replications, and comparison method. A useful preliminary check for a queue is the **traffic intensity**

$$\rho \approx \frac{\text{mean service time}}{c \times \text{mean interarrival time}},$$

where c is the number of servers. If $\rho > 1$, work arrives faster than the system can process it, so the queue tends to grow rather than settle to a stable long-run distribution.

Worked Example: Worked Example 4: Capacity before detailed simulation

Suppose the mean interarrival time is 3.5 minutes and the mean service time is 4.7 minutes.
For one counter,

$$\rho_1 = \frac{4.7}{1(3.5)} \approx 1.34 > 1.$$

The counter lacks sufficient average capacity. A long simulation will show increasing waits, but the key conclusion is structural: one counter cannot keep up under these assumptions.
For two counters,

$$\rho_2 = \frac{4.7}{2(3.5)} \approx 0.67 < 1.$$

A stable system is possible, although waiting-time variability still requires simulation. Capacity checks should accompany, not replace, detailed simulation.

Model Critique: Do not mistake a simulated average for a universal fact

A result such as "average wait $= 2.17$ minutes" is conditional on the input distributions, queue discipline, staffing policy, operating period, and random sample. A credible report should provide:

1. the assumptions and data source for each random input;
2. the number and length of replications;
3. the mean together with variability or percentiles;
4. sensitivity to uncertain assumptions;
5. verification of the algorithm and validation against real observations.

Activity: Simulation study design

The school wants to compare one counter, two counters, and one counter with a shorter menu.
Write a short experiment plan specifying:

1. the random inputs and how they would be estimated;
2. at least four outputs;
3. the simulation horizon and number of replications;
4. one fair comparison technique using common random seeds;
5. one validation check and one sensitivity test.

Concept: Verification and validation

Verification asks, “Did we implement the model correctly?” Trace small cases by hand, check conservation rules, and reproduce results with a fixed seed.

Validation asks, “Does this model represent the real system adequately for its purpose?” Compare simulated and observed queue lengths, waits, utilization, and arrival patterns. A program can be verified but still model the wrong system.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Six ideas to retain

1. Simulation combines random inputs with deterministic process rules.
2. Uniform random numbers can be mapped to empirical outcomes through cumulative probabilities.
3. One Monte Carlo run is one possible outcome, not the answer.
4. Typical Monte Carlo error decreases only as $1/\sqrt{n}$.
5. In a queue, event order is captured by $B_i = \max(A_i, D_{i-1})$.
6. Decisions require replications, uncertainty summaries, verification, validation, and sensitivity analysis.

Checkpoint: Exit ticket

Complete both statements:
“A fixed random seed is useful because ______________________________.”
“A second seed or independent replication is necessary because ______________________________.”

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A repair time T has distribution

T (h)	1	2	3	5
$P(T=t)$	0.10	0.40	0.35	0.15.

   Construct the interval mapping and convert $U = 0.04, 0.36, 0.71, 0.92$ into repair times.

2. A quarter-circle simulation produces $C = 7{,}824$ inside points out of $n = 10{,}000$.
   2.1 Estimate π.
   2.2 If the sample size is multiplied by 25, by what factor should the typical Monte Carlo error change?
   2.3 Explain why the actual error need not decrease in every individual run.

3. For a single-server queue, arrivals occur at

$$A = (0, 2, 6, 7, 11),$$

   and service times are

$$S = (3, 5, 2, 4, 3).$$

   Construct a complete table of start times, waits, and departures. Find the average wait, maximum wait, and average time in the system.

Exercises: Core practice

4. The mean interarrival time at a help desk is 6 minutes and the mean service time is 8 minutes.
   - 4.1 Calculate ρ for one server and two servers.
   - 4.2 What can be concluded before simulation?
   - 4.3 What important information is still missing?
5. Explain why each statement is inadequate:
   - 5.1 "The simulation was run once and gave an average wait of 4.2 minutes."
   - 5.2 "The program has no coding errors, so the model is valid."
   - 5.3 "The histogram fits last month's data, so it will remain correct all year."
   - 5.4 "The two-counter system has a lower mean wait, so it is automatically the best policy."
6. A team compares two staffing policies using exactly the same sequence of arrivals and service requirements. Explain why this **common random numbers** design can make the comparison clearer. Why should the team still use several seeds?

Exercises: Modeling challenge

Prepare a one-page simulation brief for one of the following systems:

1. a school canteen queue;
2. bicycle availability at a sharing station;
3. daily inventory of a popular stationery item;
4. patient arrivals at a small clinic.

Include: purpose, state variables, events, random inputs, input-data plan, process logic, outputs, policy alternatives, replication plan, validation, sensitivity, and one limitation. A complete computer program is not required.

Optional algorithm: Single-server queue

Core explanation, worked reasoning, and application

Optional algorithm: Single-server queue

This appendix is not part of the two-hour core lesson. It translates the event equations into pseudocode.

Concept: Queue simulation pseudocode

1. Set previous departure $D \leftarrow 0$ and initialize output lists.
2. For customer $i = 1, \ldots, n$:
   2.1 generate or read interarrival and service times;
   2.2 update the arrival time A_i;
   2.3 set $B_i \leftarrow \max(A_i, D)$;
   2.4 set $W_i \leftarrow B_i - A_i$;
   2.5 set $D_i \leftarrow B_i + S_i$ and update $D \leftarrow D_i$;
   2.6 record waiting time, time in system, queue length, and server state.
3. Compute summary statistics for the replication.
4. Repeat with independent seeds and summarize across replications.

Optional note: Standard error for the quarter-circle estimator

Core explanation, worked reasoning, and application

Optional note: Standard error for the quarter-circle estimator

Let I_j equal 1 when sample point j lies inside the quarter circle and 0 otherwise. Then I_j is Bernoulli with $p = \pi/4$, and

$$\widehat{\pi} = 4\bar{I}.$$

Therefore

$$\mathrm{SE}(\widehat{\pi}) = 4\sqrt{\frac{p(1-p)}{n}} \approx \frac{1.64}{\sqrt{n}}.$$

This is a standard deviation of the estimator, not a guaranteed error bound. For example, increasing n from 100 to $10{,}000$ reduces the standard error by a factor of 10.

Lecture 6: Linear Programming and Decision Optimization
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Learning Goals

By the end of the lesson, readers should be able to:

1. identify **decision variables**, an **objective function**, and **constraints** from a real situation;
2. formulate a two-variable **linear program**;
3. graph the **feasible region** and locate its **extreme points**;
4. use corner evaluation and objective level lines to solve both maximization and minimization models;
5. recognise infeasible, unbounded, and multiple-optimum cases from geometry;
6. explain why integrality, linearity, certainty, and divisibility are modeling assumptions rather than facts;
7. formulate a small **binary program** with budget and logical constraints;
8. interpret **binding constraints**, **slack**, **shadow prices**, and a **range of optimality** in context.

From a situation to a decision model

Core explanation, worked reasoning, and application

From a situation to a decision model

Optimization begins with a decision, not an algorithm. A modeler must first decide what can be controlled, what outcome matters, and what limits the possible choices. A general constrained optimization model has the form

$$\text{optimize } f(\mathbf{x}) \qquad \text{subject to} \qquad g_i(\mathbf{x}) \le b_i.$$

When the objective and every constraint are linear, the model is a **linear program**.

Activity: Launch: Planning a production week

A school workshop makes display stands and storage boxes for an exhibition. Each product earns a different contribution, but both use limited wood and labor.

In pairs, decide:

1. what the decision variables should represent;
2. what quantity should be maximized;
3. which resource limits must become constraints;
4. whether fractional answers such as 12.4 boxes are meaningful.

Modeling lesson: The mathematical structure should reflect the decision that must actually be implemented.

Concept: Four ingredients of an optimization model

1. **Decision variables:** quantities that the decision maker controls.
2. **Objective function:** a numerical measure of success to maximize or minimize.
3. **Constraints:** resource, policy, physical, or logical restrictions.
4. **Domain restrictions:** nonnegative, integer, binary, or bounded values required by context.

A model is incomplete if any of these ingredients is left implicit.

Worked Example: Worked Example 1: Formulating the carpenter's problem

A carpenter earns a contribution of \$25 per table and \$30 per bookcase. A table requires 20 board-feet of lumber and 5 labor hours; a bookcase requires 30 board-feet and 4 labor hours. Each week, 690 board-feet and 120 labor hours are available.
Let

$$x = \text{number of tables}, \qquad y = \text{number of bookcases}.$$

The objective is weekly contribution

$$\boxed{\text{maximize } P = 25x + 30y.}$$

The resource constraints are

$$\begin{aligned} 20x + 30y &\leq 690 && \text{(lumber)},\\ 5x + 4y &\leq 120 && \text{(labor)},\\ x, y &\geq 0 && \text{(nonnegativity)}. \end{aligned}$$

If only whole products can be made, the additional restrictions $x, y \in \mathbb{Z}$ should be stated. We first solve the continuous model, then examine whether the resulting decision is implementable.

Checkpoint: Formulation

A bakery makes regular cakes and premium cakes. Contributions are \$18 and \$26. Each regular cake requires 2 hours of preparation and 1 hour of decoration; each premium cake requires 3 preparation hours and 2 decoration hours. Weekly capacities are 36 preparation hours and 20 decoration hours.

Define the variables and write the complete linear program. Include units and domain restrictions.

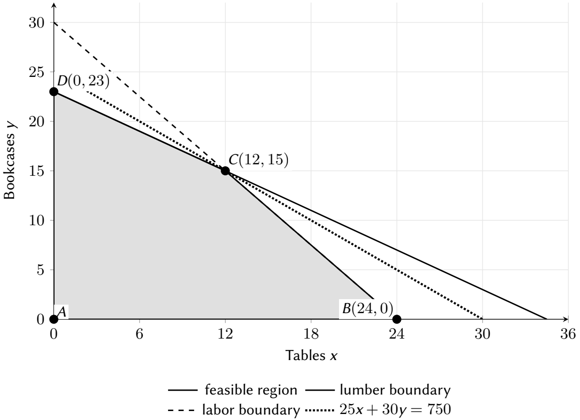
The feasible region

Core explanation, worked reasoning, and application

The feasible region

Each inequality describes a half-plane. The **feasible region** is the set of points satisfying all constraints simultaneously. A feasible point is an implementable plan under the assumptions of the model; an infeasible point violates at least one restriction.

Illustration: The feasible region



The feasible region is the overlap of all constraints. The dotted objective line is the highest parallel level line that still touches the feasible region.

The feasible region

A set is **convex** if the line segment joining any two points in the set remains inside the set. Intersections of linear half-planes are convex, so every linear-programming feasible region is convex.

Concept: Why do corners matter?

A linear objective has parallel level lines. Moving a level line in the improving direction, the last contact with a bounded convex polygon occurs at a corner, or along an entire edge when there are multiple optimal solutions. Therefore, for a two-variable linear program, it is enough to evaluate the objective at the feasible extreme points. This conclusion depends on linearity; it is not generally true for nonlinear objectives.

Worked Example: Worked Example 2: Solving by extreme points

The feasible extreme points are

$$A = (0, 0), \quad B = (24, 0), \quad C = (12, 15), \quad D = (0, 23).$$

Evaluate $P = 25x + 30y$:

point	P (dollars)
A	0
B	600
C	750
D	690

Hence the continuous optimum is

$$\boxed{x = 12, \qquad y = 15, \qquad P = 750.}$$

Both resource constraints are equalities at C, so all lumber and all labor are used.

Activity: Guided graphical solution

For the bakery model from the checkpoint:

1. draw each boundary line using intercepts;
2. identify the feasible side of each line;
3. list all feasible extreme points;
4. evaluate the objective at each extreme point;
5. state the optimal plan in a complete sentence.

Minimization and minimum requirements

Core explanation, worked reasoning, and application

Minimization and minimum requirements

Not every optimization problem seeks the largest profit. A school may minimize staffing cost, a transport company may minimize fuel use, or a diet model may minimize cost while meeting minimum nutritional requirements. Constraints of the form “at least” usually produce $\geq$ inequalities, so the feasible region may lie *above* the boundary lines and may not contain the origin.

Worked Example: Worked Example 3: Minimum-cost delivery fleet

A delivery service can hire trucks and vans for one day. A truck costs \$120 and carries 8 pallet units; a van costs \$70 and carries 4 pallet units. At least 48 pallet units must be carried, and at least 8 vehicles must be available for route coverage.

Let x be the number of trucks and y the number of vans. The continuous model is

$$\text{minimize } C = 120x + 70y \quad \text{subject to} \quad 8x + 4y \geq 48, \qquad x + y \geq 8, \qquad x, y \geq 0.$$

Dividing the first constraint by 4 gives $2x + y \geq 12$. The relevant lower-bound extreme points are

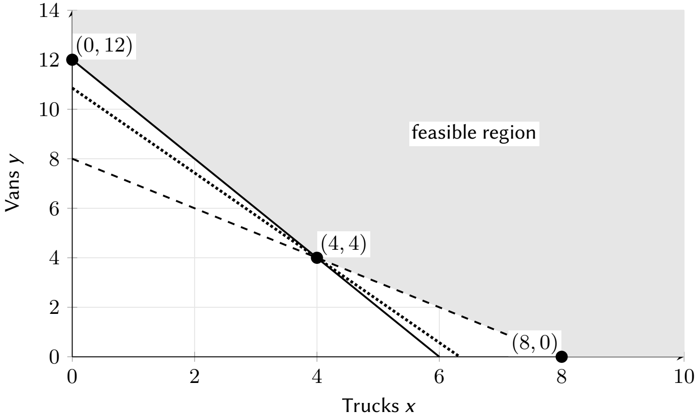
$$(0, 12), \qquad (4, 4), \qquad (8, 0).$$

Their costs are \$840, \$760, and \$960 respectively. Hence the continuous optimum is

$$\boxed{x = 4, \qquad y = 4, \qquad C = 760.}$$

Here the objective line is moved *toward smaller cost* until it first touches the feasible region.

Illustration: Minimization and minimum requirements



For minimum requirements, the feasible region lies above the constraints. The minimum-cost level line first touches the region at $(4, 4)$.

Checkpoint: Maximize or minimize?

For each situation, identify a plausible objective and state whether it should be maximized or minimized:

1. scheduling revision sessions while meeting minimum subject coverage;
2. allocating an advertising budget to increase expected registrations;
3. choosing emergency supplies while keeping total weight as small as possible.

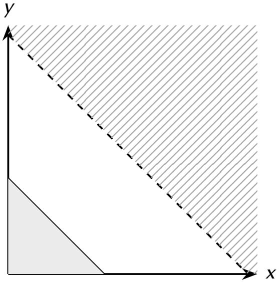
Three special outcomes

Core explanation, worked reasoning, and application

Three special outcomes

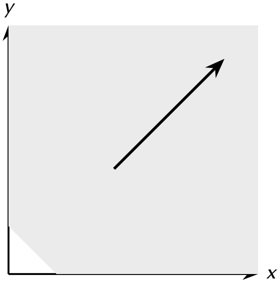
A correctly formulated linear program does not always have one finite optimal corner. The geometry should be checked before an answer is reported.

Illustration: Three special outcomes



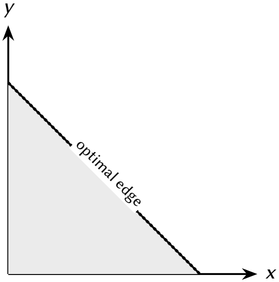
Infeasible

required regions do not overlap



Unbounded

objective can improve indefinitely



Multiple optima

objective is parallel to an edge

Before interpreting a solution, check whether the model is infeasible, unbounded, or has an entire edge of optimal solutions.

Concept: How to report special cases

- **Infeasible:** the requirements cannot all be satisfied. Recheck data or negotiate constraints.
- **Unbounded:** the model permits indefinite improvement. A missing upper limit, cost, or capacity is often the cause.
- **Multiple optima:** several decisions are mathematically equivalent, so secondary criteria such as risk, fairness, or ease of implementation can decide among them.

Model Critique: A feasible point is not automatically a sensible decision

The mathematics checks only the stated constraints. A feasible production plan may still be unacceptable because the model omitted demand limits, storage, contractual minimums, setup costs, uncertainty, or fairness. Optimization does not replace judgment; it makes the stated trade-offs explicit.

Binding constraints, slack, and what-if questions

Core explanation, worked reasoning, and application

Binding constraints, slack, and what-if questions

At a proposed solution, a constraint is **binding** if it holds with equality. Otherwise it has **slack**. For a resource constraint

$$a_1 x + a_2 y \leq b,$$

the slack is $b - (a_1 x + a_2 y)$.

Worked Example: Worked Example 4: Interpreting slack

At $B = (24, 0)$:

$$\text{lumber used} = 20(24) = 480, \quad \text{lumber slack} = 690 - 480 = 210,$$
$$\text{labor used} = 5(24) = 120, \qquad \text{labor slack} = 120 - 120 = 0.$$

Labor is binding, while 210 board-feet remain unused. Adding lumber alone cannot improve this plan locally because labor is already exhausted.

At the optimum $C = (12, 15)$, both slacks are zero. Both resources are bottlenecks under the current coefficients.

Binding constraints, slack, and what-if questions

A **shadow price** estimates the increase in the optimal objective value from one additional unit of a binding resource, provided the change is small enough that the same pair of constraints remains active.

Worked Example: Worked Example 5: Marginal value of resources

At the carpenter optimum, the two binding constraints are

$$20x + 30y = b_L, \qquad 5x + 4y = b_H.$$

The shadow prices u and v for lumber and labor satisfy

$$20u + 5v = 25, \qquad 30u + 4v = 30,$$

because the value assigned to the resources used by one product must match its contribution at the margin. Solving gives

$$\boxed{u = \frac{5}{7} \approx 0.71 \text{ dollars per board-foot}}, \qquad \boxed{v = \frac{15}{7} \approx 2.14 \text{ dollars per labor hour}}.$$

Thus one extra labor hour is locally more valuable than one extra board-foot. Paying \$3 for one extra labor hour would not be worthwhile under this local model, while paying \$1 might be.

Worked Example: Worked Example 6: Range of optimality for table profit

Suppose the table contribution changes from \$25 to c_T, while bookcase contribution remains \$30. The objective level lines have slope

$$-\frac{c_T}{30}.$$

At the current optimum $C = (12, 15)$, the slopes of the two binding edges are $-5/4$ (labor) and $-2/3$ (lumber). The same corner remains optimal while the objective slope lies between them:

$$-\frac{5}{4} \leq -\frac{c_T}{30} \leq -\frac{2}{3}.$$

Reversing signs and solving gives

$$\boxed{20 \leq c_T \leq 37.5.}$$

Within this interval, the recommended production quantities remain $(12, 15)$ even though total contribution changes. At an endpoint, the objective line is parallel to one edge and multiple optimal solutions appear. Outside the interval, a different corner becomes optimal.

Activity: Sensitivity studio

Use slopes rather than recomputing every objective value.

1. If table contribution rises to \$34, does the production mix change?
2. If it rises to \$40, which neighboring corner should become preferable?
3. Why is a **range of optimality** more useful than testing only a single alternative coefficient?

Checkpoint: Slack and marginal value

A solution leaves 40 kg of material unused but uses every available labor hour.

1. Which constraint is binding?
2. Which resource should have zero shadow price locally?
3. Why does "buy the resource with the larger shadow price" still require a validity range and a cost comparison?

Model Critique: Sensitivity is local

A shadow price is not a universal market value. If a resource changes enough, the optimal corner may change and the old shadow price no longer applies. Sensitivity analysis should report both the marginal value and the range over which the interpretation remains valid.

When whole-number decisions matter

Core explanation, worked reasoning, and application

When whole-number decisions matter

A continuous linear program allows fractional values. This is appropriate for divisible quantities such as kilograms, hours, or proportions. It may be inappropriate for people, vehicles, projects, or indivisible products. Adding $x_j \in \mathbb{Z}$ gives an **integer program**; restricting $x_j \in \{0, 1\}$ gives a **binary program**.

Worked Example: Worked Example 7: Why rounding can fail

Consider

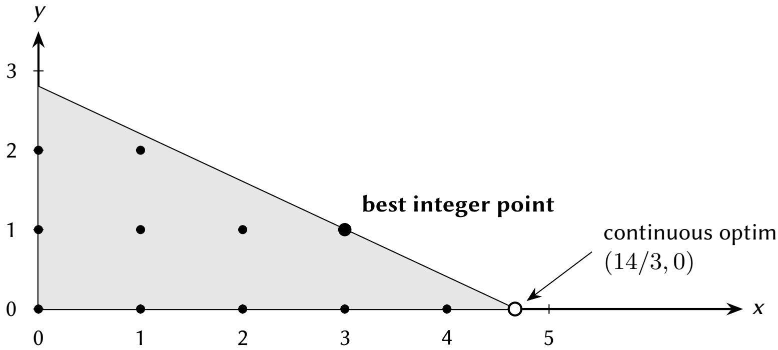
$$\text{maximize } 7x + 10y \quad \text{subject to} \quad 3x + 5y \leq 14, \qquad x, y \geq 0,$$

where x and y count indivisible packages.

The continuous model prefers $(14/3, 0)$, with objective $98/3 \approx 32.67$. Rounding x to 5 violates the capacity constraint because $3(5) = 15 > 14$. Rounding down to $(4, 0)$ gives value 28, but the integer-feasible point $(3, 1)$ uses all capacity and gives value 31.

Conclusion: Solving continuously and rounding afterward can produce either an infeasible or a suboptimal decision.

Illustration: When whole-number decisions matter



Integer feasibility replaces a continuous region by a set of lattice points. The best integer point need not be obtained by rounding the continuous optimum.

Worked Example: Worked Example 8: Binary project selection

A student council considers five projects. Costs and estimated benefit scores are:

Project	Cost	Benefit score
A: Study hub	4	8
B: Booking system	5	12
C: Peer mentoring	3	7
D: Wellbeing event	4	9
E: Equipment fund	6	11

Let $z_A, \ldots, z_E \in \{0, 1\}$ indicate selection. The council has budget 12, may choose at most three projects, Project B requires A, and C and D cannot both be chosen:

$$\text{maximize } B = 8z_A + 12z_B + 7z_C + 9z_D + 11z_E,$$

$$\begin{aligned} 4z_A + 5z_B + 3z_C + 4z_D + 6z_E &\leq 12, \\ z_A + z_B + z_C + z_D + z_E &\leq 3, \\ z_B &\leq z_A, \\ z_C + z_D &\leq 1. \end{aligned}$$

Worked Example: Worked Example 8: Binary project selection

For this small problem, feasible combinations can be enumerated. The strongest candidates are

selection	cost	benefit
$A + B + C$	12	27
$D + E$	10	20
$A + E$	10	19
$C + E$	9	18
$A + D$	8	17

Hence the optimal selection is $A + B + C$, with benefit score 27.

Activity: Differentiated optimization studio

Choose one route.
Core route: Verify every constraint for the proposed choice $A + B + C$, then explain why choosing B alone is prohibited.
Challenge route: Suppose the budget falls to 10. Find the new optimal combination by systematic enumeration and report which logical constraint becomes important.
Extension route: Add a condition "exactly one of D and E must be selected" and write the corresponding binary equation. Discuss whether the original optimum remains feasible.

Assumptions and communication

Core explanation, worked reasoning, and application

Assumptions and communication

The standard linear-programming assumptions are useful but demanding:

- **Proportionality:** doubling an activity doubles its resource use and contribution.
- **Additivity:** total effects are sums; there are no interaction or setup effects.
- **Divisibility:** continuous variables may take fractional values.
- **Certainty:** all coefficients are treated as known constants.

Model Critique: When linearity becomes questionable

Bulk discounts, overtime premiums, fixed setup costs, uncertain demand, product interactions, congestion, and economies of scale all violate simple linearity. A useful first model may still be linear, but its assumptions and intended range must be stated explicitly.

Activity: Competition-style conclusion

Write a four-sentence recommendation for the carpenter:

1. report the optimal production plan and contribution;
2. identify the binding resources;
3. state one useful sensitivity result;
4. acknowledge one assumption that could change the decision.

Avoid writing that the model “proves” the decision is best in reality.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Nine ideas to retain

1. Optimization translates a decision into variables, an objective, constraints, and domain restrictions.
2. A feasible solution satisfies every stated constraint; feasibility does not guarantee practical suitability.
3. A two-variable linear program can be solved by evaluating feasible extreme points, whether the objective is maximized or minimized.
4. A model may be infeasible, unbounded, or have multiple optimal solutions; each outcome has a different interpretation.
5. Binding constraints identify current bottlenecks; slack measures unused capacity.
6. Shadow prices provide local marginal values, not permanent prices.
7. A range of optimality describes how far an objective coefficient can change before the recommended corner changes.
8. Whole-number or yes/no decisions require integer or binary variables; naive rounding is unsafe.
9. Logical requirements such as dependence and incompatibility can be written as linear constraints on binary variables.
10. The optimal answer is conditional on assumptions about linearity, divisibility, certainty, and omitted factors.

Checkpoint: Exit ticket

Complete both statements:
"The mathematical optimum is ______________________________."
"Before implementing it, I would check ____________________________________."

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A farmer plants wheat and corn. Wheat earns \$200 per acre and uses 3 labor-days and 2 fertilizer units. Corn earns \$300 per acre and uses 2 labor-days and 4 fertilizer units. At most 45 acres, 100 labor-days, and 120 fertilizer units are available.
   - 1.1 Define the variables and formulate the linear program.
   - 1.2 Graph the feasible region and find all extreme points.
   - 1.3 Determine the optimal planting plan and identify binding constraints.
2. For the carpenter model, calculate the slack in each resource at A, B, C, and D. Explain why unused lumber at B does not imply that B can be improved by adding more tables.
3. A delivery plan uses trucks x and vans y:

   $$\text{minimize } 120x + 70y \quad \text{subject to} \quad 8x + 4y \geq 48, \quad x + y \geq 8, \quad x, y \geq 0.$$

   Interpret every coefficient, graph the feasible region, and find the continuous optimum. Then discuss whether integrality matters.
4. A student claims: “The point with the greatest x-coordinate must maximize every profit function.” Explain the error using level lines.

Exercises: Core practice

5. For each of the following, sketch a feasible region and explain the outcome: (a) inconsistent constraints, (b) an unbounded maximization problem, and (c) multiple optimal solutions.
6. For the carpenter model, derive the range of bookcase contribution c_B for which $(12, 15)$ remains optimal. Interpret what happens at the endpoints.
7. Re-solve the binary project model when the budget is 10. Show a compact enumeration table rather than listing all 2^5 possibilities.
8. Construct an example in which a linear program has multiple optimal solutions. State the geometric condition that causes this.
9. A report states: “One extra labor hour is worth \$2.14, so 1000 additional hours are worth \$2140.” Critique the statement carefully.

Exercises: Modeling challenge

A school must choose how many standard and intensive revision workshops to offer. Each type uses teacher hours, rooms, and a limited printing budget. Student benefit is estimated differently for the two formats. Prepare a one-page optimization brief containing:

1. decision purpose and stakeholders;
2. variables with units and domain restrictions;
3. an objective and at least three constraints;
4. assumptions behind linearity and benefit scores;
5. a graphical or computational solution plan;
6. one sensitivity question and one implementation limitation.

The quality of the formulation and interpretation is more important than obtaining a particular numerical optimum.

Optional extension: What the Simplex Method does

Core explanation, worked reasoning, and application

Optional extension: What the Simplex Method does

The **Simplex Method** is an algebraic method for larger linear programs. It starts at one feasible extreme point and repeatedly moves to an adjacent extreme point with a better objective value. **Slack variables** convert inequalities into equations, and a **pivot** changes the current basis.
For the carpenter model, Simplex follows the path

$$(0,0) \longrightarrow (0,23) \longrightarrow (12,15),$$

then stops because no adjacent move improves the objective. A full tableau treatment is best taught as a separate lesson after students understand formulation, geometry, special cases, and interpretation.

Optional extension: Derivative-free search

Core explanation, worked reasoning, and application

Optional extension: Derivative-free search

A one-dimensional nonlinear objective may be optimized by **golden-section search** when it is unimodal on a known interval. This is a different topic from linear programming: it narrows an interval containing one optimum rather than moving among corners of a feasible polytope. It is reserved for a later numerical-optimization lesson.

Lecture 7: Decision Theory and Risk
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Learning Goals

By the end of the lesson, readers should be able to:

1. distinguish a **decision alternative** from a **state of nature**;
2. compute and interpret **expected value** for repeated risky decisions;
3. build and solve a simple **decision tree** using the fold-back method;
4. use **Bayes' theorem** to update probabilities after new information;
5. compute the **expected value of perfect information**;
6. apply Laplace, Maximin, Maximax, Hurwicz, and Minimax Regret criteria when probabilities are unknown;
7. explain why the mathematically optimal decision may depend on risk tolerance, repetition, and consequences.

Why decision theory?

Core explanation, worked reasoning, and application

Why decision theory?

Previous lectures often assumed that the model inputs were known. In real decisions, some inputs are uncertain: the weather may change, market demand may be high or low, a bid may be accepted or rejected, and a medical test may be correct or misleading. **Decision theory** provides a disciplined way to choose actions when outcomes depend on uncertain future states.
A decision model separates three ingredients:

1. **Alternatives**: actions the decision maker can choose.
2. **States of nature**: uncertain conditions not controlled by the decision maker.
3. **Payoffs**: consequences produced by each alternative-state combination.

Activity: Launch: What should the stall sell?

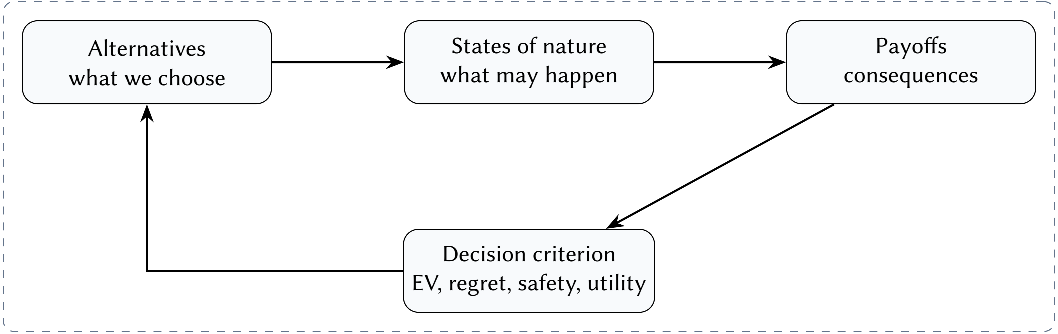
A school fundraising stall can sell either hot coffee or iced tea. Coffee earns more on a cold day; iced tea earns more on a warm day. Tomorrow's weather is uncertain.
In pairs, discuss:

1. What are the alternatives? What are the states of nature?
2. What information would change the decision?
3. Is this a repeated decision or a one-time decision?
4. Would you rather maximize average profit or avoid a very bad outcome?

Modeling lesson: Before computing anything, identify who chooses, what is uncertain, and what consequence matters.

Illustration: Why decision theory?

Decision modeling structure



A decision model combines choices, uncertain states, consequences, and a criterion for judging them.

Concept: Repeated decisions versus one-time decisions

Expected value is a long-run average. It is most convincing when the same kind of decision is repeated many times, such as an insurance company selling many policies. For a one-time decision with large downside risk, the highest expected value may not be the most acceptable choice.

This distinction is not a minor philosophical issue. It determines whether we should use expected value, a risk-averse criterion, or a regret-based criterion.

Expected value for repeated risky decisions

Core explanation, worked reasoning, and application

Expected value for repeated risky decisions

If a random outcome has payoffs $w_1, w_2, \ldots, w_n$ with probabilities $p_1, p_2, \ldots, p_n$, where $\sum p_i = 1$, its **expected value** is

$$E = \sum_{i=1}^{n} p_i w_i.$$

The value E is not necessarily a possible outcome. It is the average payoff expected over many repetitions under the same probability model.

Worked Example: Worked Example 1: A dice game

A game costs \$1 to play. If two fair dice sum to 7, the player receives \$6 total; otherwise the player receives nothing. The net payoff is therefore \$5 for a sum of 7 and $-\$1$ otherwise.
There are 36 equally likely dice outcomes and 6 of them sum to 7, so

$$P(\text{sum} = 7) = \frac{6}{36} = \frac{1}{6}.$$

The expected net payoff is

$$E = 5\left(\frac{1}{6}\right) + (-1)\left(\frac{5}{6}\right) = 0.$$

The game is **fair** in the expected-value sense. This does not mean every player breaks even; it means that the average net gain approaches zero over many plays.

Model Critique: Expected value is not a promise

A student might say, "The expected value is zero, so I will neither win nor lose." This is wrong. A single play gives either \$5 or $-\$1$. Expected value describes the centre of the distribution over repeated trials, not the outcome of one trial.

Worked Example: Worked Example 2: Concession decision with known weather probability

A concession firm must choose one product for tomorrow's school sports day.

	Cold day	**Warm day**
Coffee	\$4000	\$1000
Iced tea	\$1500	\$5000

The forecast gives $P(\text{cold}) = 0.30$, so $P(\text{warm}) = 0.70$.

$$E(\text{coffee}) = 0.30(4000) + 0.70(1000) = 1900.$$

$$E(\text{iced tea}) = 0.30(1500) + 0.70(5000) = 3950.$$

The expected-value decision is to sell iced tea.
The **break-even probability** for cold weather satisfies

$$4000p + 1000(1 - p) = 1500p + 5000(1 - p).$$

Solving gives $p = \frac{4000}{6500} \approx 0.615$. Coffee becomes preferable only if the probability of a cold day exceeds about 61.5%.

Checkpoint: Expected value

A bid costs \$2000 to prepare. If accepted, it earns a net profit of \$18,000 after subtracting preparation cost. If rejected, the company loses the \$2000 preparation cost.

1. Write the expected value in terms of the acceptance probability p.
2. Find the break-even value of p.
3. Explain whether the company should bid if $p = 0.15$.

Decision trees and the fold-back method

Core explanation, worked reasoning, and application

Decision trees and the fold-back method

A **decision tree** is useful when a decision is followed by one or more uncertain events. We use square nodes for decisions, circular nodes for uncertainty, and terminal values for payoffs.

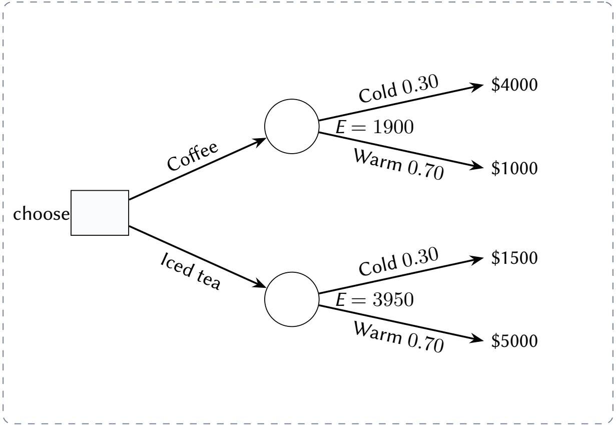
Concept: Fold-back method

Solve a decision tree from right to left:

1. At each chance node, compute the expected value of its branches.
2. At each decision node, keep the branch with the best expected value.
3. Replace the solved node by its value and continue moving left.

This method is called **fold-back**, because we simplify the future part of the tree before choosing the earlier action.

Illustration: Decision trees and the fold-back method



Decision tree for the concession example. Fold-back chooses the branch with the larger expected value.

Worked Example: Worked Example 3: Plant-size decision

A company may build a large plant, build a small plant, or build nothing. Demand may be high, moderate, or low.

Alternative	**High** (0.25)	**Moderate** (0.40)	**Low** (0.35)
Large plant	200000	100000	-120000
Small plant	90000	50000	-20000
No plant	0	0	0

$$E(\text{large}) = 0.25(200000) + 0.40(100000) + 0.35(-120000) = 48000.$$

$$E(\text{small}) = 0.25(90000) + 0.40(50000) + 0.35(-20000) = 35500.$$

$$E(\text{none}) = 0.$$

By expected value, the large plant is best. However, the low-demand loss is \$120,000, so the recommendation should mention risk, not only the average.

Model Critique: Risk profile

Two alternatives can have similar expected values but very different downside risk. A complete decision analysis should report:

1. expected value;
2. best-case and worst-case payoff;
3. probability of loss;
4. whether the decision is repeated or one-time;
5. the decision maker's risk tolerance.

Information, Bayes' theorem, and EVPI

Core explanation, worked reasoning, and application

Information, Bayes' theorem, and EVPI

Information is valuable only if it can change the decision. Decision theory allows us to place an upper bound on the value of information and, when the information is imperfect, to compute whether it is worth buying.

Concept: Perfect information

The **expected value of perfect information** is

$$\text{EVPI} = E(\text{best action after knowing the state}) - E(\text{best action now}).$$

It is the most we should ever pay for flawless information. Any imperfect forecast or survey must be worth no more than EVPI.

Worked Example: Worked Example 4: EVPI for the plant-size decision

Without information, the best expected value is \$48,000 from the large plant.
With perfect information, the company would choose the best action after seeing the true demand state:

State	Best action	Payoff
High	Large plant	\$200000
Moderate	Large plant	\$100000
Low	No plant	\$0

Therefore,

$$E(\text{perfect information}) = 0.25(200000) + 0.40(100000) + 0.35(0) = 90000.$$

$$\boxed{\text{EVPI} = 90000 - 48000 = 42000.}$$

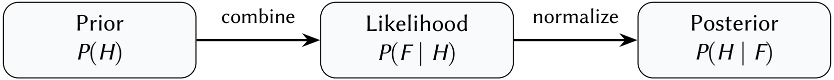
No market research report, even a perfect one, is worth more than \$42,000 for this decision.

Information, Bayes' theorem, and EVPI

When information is imperfect, we update probabilities using **Bayes' theorem**. If $H_1, \ldots, H_n$ are possible states and F is evidence, then

$$P(H_i \mid F) = \frac{P(F \mid H_i)P(H_i)}{\sum_j P(F \mid H_j)P(H_j)}.$$

Illustration: Information, Bayes' theorem, and EVPI



Bayes update: new evidence changes the probability model before the decision is evaluated.

Bayes' theorem updates beliefs from prior probabilities to posterior probabilities using evidence.

Worked Example: Worked Example 5: Market research with Bayes update

For the plant-size problem, suppose a market report is either favorable F or unfavorable U. Its historical accuracy is:

State	$P(F \mid \text{state})$	$P(U \mid \text{state})$
High	0.70	0.30
Moderate	0.50	0.50
Low	0.20	0.80

First compute

$$P(F) = 0.70(0.25) + 0.50(0.40) + 0.20(0.35) = 0.445.$$

The posterior probabilities after a favorable report are

$$P(H \mid F) = \frac{0.70(0.25)}{0.445} \approx 0.393, \quad P(M \mid F) = \frac{0.50(0.40)}{0.445} \approx 0.449, \quad P(L \mid F) = \frac{0.20(0.35)}{0.445} \approx 0.157.$$

Re-evaluating the large plant after a favorable report gives

$$E(\text{large} \mid F) \approx 0.393(200000) + 0.449(100000) + 0.157(-120000) = 104640.$$

Worked Example: Worked Example 5: Market research with Bayes update

Small plant gives about \$54,690, so a favorable report supports the large plant.
For an unfavorable report, $P(U) = 0.555$ and

$$P(H \mid U) \approx 0.135, \quad P(M \mid U) \approx 0.360, \quad P(L \mid U) \approx 0.505.$$

Then

$$E(\text{large} \mid U) \approx 2400, \qquad E(\text{small} \mid U) \approx 20020,$$

so an unfavorable report supports the small plant. The report is useful because it changes the decision in some cases.

Checkpoint: Value of information

In the market-research example, after a favorable report the decision is large plant, and after an unfavorable report the decision is small plant.

1. Compute the expected value if the company uses the report: $0.445E_F + 0.555E_U$.
2. Subtract the best no-report expected value \$48,000 to obtain EVSI.
3. Should the company buy the report if it costs \$10,000? What if it costs \$25,000?

When probabilities are unknown

Core explanation, worked reasoning, and application

When probabilities are unknown

Sometimes we cannot assign credible probabilities. In that case, expected value may give a false impression of precision. We use alternative criteria that represent different attitudes toward uncertainty.

Worked Example: Running Example: Investment payoffs

A student club has funds to invest for one year. Payoffs are in thousand dollars.

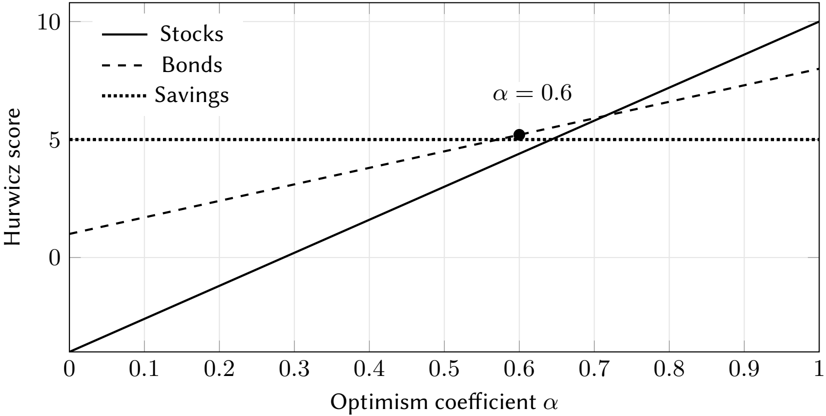
Alternative	**Fast growth**	**Normal growth**	**Slow growth**
Stocks	10	6.5	-4
Bonds	8	6	1
Savings	5	5	5

No reliable probabilities are available for the three economic states.

When probabilities are unknown

Criterion	Rule	Decision in example
Laplace	Treat all states as equally likely and average each row.	Bonds/Savings tie at 5.0.
Maximin	For each row, find the worst payoff; choose the best worst payoff.	Savings.
Maximax	For each row, find the best payoff; choose the best best payoff.	Stocks.
Hurwicz	Use $\alpha(\text{best}) + (1-\alpha)(\text{worst})$.	Depends on α; at $\alpha = 0.6$, Bonds.
Minimax regret	Build a regret matrix and minimize the largest regret.	Bonds.

Illustration: When probabilities are unknown



Hurwicz scores change with the optimism coefficient. The chosen alternative depends on the decision maker's attitude toward risk.

Worked Example: Worked Example 6: Minimax regret

For each state, subtract each payoff from the best payoff in that column.

	Fast	**Normal**	**Slow**	**Max regret**
Stocks	$10 - 10 = 0$	$6.5 - 6.5 = 0$	$5 - (-4) = 9$	9
Bonds	$10 - 8 = 2$	$6.5 - 6 = 0.5$	$5 - 1 = 4$	4
Savings	$10 - 5 = 5$	$6.5 - 5 = 1.5$	$5 - 5 = 0$	5

The smallest maximum regret is 4, so the minimax-regret decision is Bonds. This criterion asks: "If the true state is later revealed, which decision limits how much we will regret not having chosen the best action for that state?"

Activity: Decision studio

15 min

Use the investment example.

1. **Core route:** Apply Laplace, Maximin, Maximax, and Minimax Regret. Write one sentence explaining each attitude.
2. **Challenge route:** Find the α values where Hurwicz changes from Savings to Bonds and from Bonds to Stocks.
3. **Extension route:** Suppose the club cannot tolerate any possibility of loss. Which criteria respect this constraint? Which criterion ignores it?

Integrated mini-case: stocking under uncertain demand

Core explanation, worked reasoning, and application

Integrated mini-case: stocking under uncertain demand

Decision theory becomes most useful when the modeler must connect a business rule, uncertain demand, and a recommendation. The following example is intentionally small enough to compute by hand, but it contains the same structure as larger inventory decisions.

Worked Example: Worked Example 7: Seasonal T-shirt stocking

A school club can order 10, 20, or 30 festival T-shirts. Each shirt costs $50 and sells for $80. Unsold shirts can be cleared for $20 each. If demand exceeds stock, each unmet demand unit causes a goodwill loss of $10. Demand is estimated as:

$$P(D = 10) = 0.30, \qquad P(D = 20) = 0.40, \qquad P(D = 30) = 0.30.$$

For an order quantity Q and demand D, the payoff is

$$80 \min(Q, D) + 20 \max(Q - D, 0) - 50Q - 10 \max(D - Q, 0).$$

This gives the payoff matrix:

Order quantity	$D = 10$	$D = 20$	$D = 30$	**Expected value**
$Q = 10$	300	200	100	$0.30(300) + 0.40(200) + 0.30(100) = 200$
$Q = 20$	0	600	500	$0.30(0) + 0.40(600) + 0.30(500) = 390$
$Q = 30$	-300	300	900	$0.30(-300) + 0.40(300) + 0.30(900) = 300$

By expected value, the best order is $Q = 20$.
Different risk criteria tell a more nuanced story:

Worked Example: Worked Example 7: Seasonal T-shirt stocking

- Maximin chooses $Q = 10$, because its worst payoff is still \$100.
- Maximax chooses $Q = 30$, because it can earn \$900 if demand is high.
- Minimax regret chooses $Q = 20$: its maximum regret is \$400, smaller than the maximum regret for $Q = 10$ or $Q = 30$.

Thus the same payoff matrix can support different recommendations depending on whether the club prioritizes average profit, downside protection, or regret control.

Activity: Sensitivity check

8 min

In the T-shirt example, suppose the probability of high demand $D = 30$ increases while the probability of low demand $D = 10$ decreases. Without recalculating every criterion, predict which order quantity becomes more attractive. Then compute the expected values when

$$P(D = 10) = 0.20, \qquad P(D = 20) = 0.40, \qquad P(D = 30) = 0.40.$$

Explain whether the decision changes.

Model Critique: Hidden assumptions in payoff matrices

A payoff matrix looks precise, but each number depends on assumptions: selling price, clearance price, lost goodwill cost, and the demand probabilities. If these inputs are estimated weakly, the report should include sensitivity checks rather than present the final order quantity as certain.

Writing a defensible decision recommendation

Core explanation, worked reasoning, and application

Writing a defensible decision recommendation

Decision theory does not remove judgment. It makes the judgment explicit. In a modeling report, the recommendation should not simply state the branch with the largest expected value; it should state the assumptions under which the recommendation holds.

Model Critique: Common mistakes in decision models

1. Treating guessed probabilities as if they were measured facts.
2. Using expected value for a one-time catastrophic decision without discussing risk.
3. Forgetting to subtract the cost of information from EVSI.
4. Confusing a decision node with a chance node: the modeler chooses at one; nature decides at the other.
5. Reporting a criterion result without explaining the risk attitude represented by the criterion.

Worked Example: Competition-style conclusion

For the plant-size case, a concise conclusion might be:
Using the prior demand probabilities, the large plant has the highest expected value, \$48,000, and is therefore the recommended decision if management is willing to accept a possible \$120,000 loss in the low-demand state. Perfect information would be worth at most \$42,000, so any market research costing more than this should be rejected immediately. Because the large-plant recommendation is sensitive to the low-demand loss and to the estimated probability of high demand, we also recommend reporting a risk profile and considering the small plant if management is strongly risk-averse.

Summary: Lesson summary

1. A decision model separates alternatives, states of nature, probabilities, and payoffs.
2. Expected value is a long-run average and is most appropriate for repeated decisions.
3. Decision trees are solved by fold-back: expected values at chance nodes, best choices at decision nodes.
4. Bayes' theorem updates probabilities after evidence.
5. EVPI is an upper bound on the value of information; EVSI must be compared with the cost of the information.
6. When probabilities are unknown, different criteria represent different attitudes toward risk and uncertainty.

Checkpoint: Exit ticket

2 min

Complete the sentence:
"Expected value is a useful decision criterion when ______________________________, but it may be misleading when ______________________________."

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. A game costs \$2 to play. A player wins \$12 with probability 0.20 and wins nothing otherwise. Compute the expected net payoff. Is the game fair?
2. A food stall must choose noodles or sandwiches. Profits are:

	Rainy ($p = 0.40$)	Sunny ($p = 0.60$)
Noodles	3200	1600
Sandwiches	1800	3600

   Compute both expected values and find the break-even probability of rain.
3. For the plant-size problem in Worked Example 3, compute the probability of losing money under each alternative. Does this change how you would present the recommendation?
4. A medical screening test has $P(+ \mid D) = 0.95$, $P(+ \mid \text{no } D) = 0.08$, and disease prevalence $P(D) = 0.02$. Compute $P(D \mid +)$. Explain why a highly sensitive test can still have a moderate positive predictive value.
5. For the investment payoff matrix, apply Hurwicz with $\alpha = 0.3$ and $\alpha = 0.8$. Explain why the decision changes.
6. For the following payoff matrix, construct the regret matrix and apply minimax regret:

Exercises: Core practice

	S_1	S_2	S_3
A	20	5	8
B	12	12	12
C	4	18	10

Exercises: Modeling challenge

Choose one of the following and prepare a one-page decision analysis brief containing: alternatives, states of nature, payoff table, probability assumptions if any, chosen criterion, recommendation, sensitivity comment, and one limitation.

1. A school club deciding whether to hold an outdoor fundraising event, indoor event, or cancel.
2. A student team deciding whether to submit a safe project, a risky innovative project, or a hybrid project to a modeling competition.
3. A small shop deciding whether to stock 10, 20, or 30 units of a seasonal item.

Your answer should explain why the chosen criterion is appropriate for the situation.

Optional extension: expected utility

Core explanation, worked reasoning, and application

Optional extension: expected utility

Expected value treats each dollar equally. In reality, the pain of losing \$100,000 may be more than twice the pain of losing \$50,000. **Expected utility** replaces payoff w by a utility function $u(w)$ and chooses the alternative with largest

$$E[u(W)] = \sum_i p_i u(w_i).$$

A concave utility function represents risk aversion: gaining an extra dollar matters less when the decision maker is already wealthy, while losses near a critical threshold may be unacceptable. This is why a risk-averse manager may choose the small plant even when the large plant has a larger expected monetary value.

Optional extension: expected value of sample information

Core explanation, worked reasoning, and application

Optional extension: expected value of sample information

If a report can be favorable or unfavorable, the expected value with sample information is

$$E(\text{with report}) = P(F)\max_a E(a \mid F) + P(U)\max_a E(a \mid U).$$

Then

$$\text{EVSI} = E(\text{with report}) - E(\text{best decision without report}).$$

The report should be purchased only when EVSI exceeds its cost. If the cost is included in terminal payoffs, do not subtract it a second time.

Lecture 8: Game Theory and Strategic Modeling

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, readers should be able to:

1. explain why strategic models differ from ordinary decision models;
2. identify players, strategies, payoffs, and assumptions in a game;
3. find a pure-strategy solution using **maximin**, **minimax**, and movement diagrams;
4. solve a 2×2 zero-sum game without a saddle point using mixed strategies;
5. interpret a game value as a guaranteed long-run payoff, not as a promise in one play;
6. analyse a partial-conflict game using **Nash equilibrium** and explain the Prisoner's Dilemma;
7. decide when game theory is an appropriate modeling tool and state its limitations.

Why game theory?

Core explanation, worked reasoning, and application

Why game theory?

In Lecture 7, a decision maker faced uncertain states of nature: weather, market demand, or a research report. Nature did not try to outsmart the decision maker. In **game theory**, at least one other participant is also making a decision, and that participant may change strategy after thinking about what we will do.

A **game** is a mathematical model of strategic interaction. It must specify:

- the **players**;
- each player's possible **strategies**;
- the **payoff** produced by every combination of strategies;
- what each player knows when making a decision.

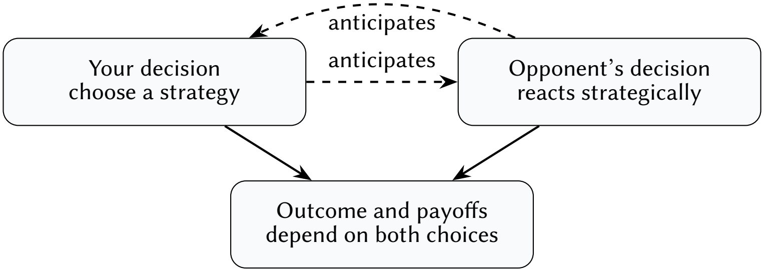
Activity: Launch: ordinary decision or game?

For each situation, decide whether a decision-theory model or a game-theory model is more appropriate.

1. A shop chooses how many umbrellas to order before the rainy season.
2. Two restaurants decide whether to offer lunch discounts.
3. A goalkeeper chooses which side to dive during a penalty kick.
4. A school decides whether to hire one or two extra canteen staff.

For the game situations, identify the players and their possible strategies.

Illustration: Why game theory?



A game is different from an ordinary decision because the other side is also reasoning.

Concept: Total conflict and partial conflict

A **total-conflict game**, often called a **zero-sum game**, is one in which one player's gain is exactly the other player's loss. A penalty kick, rock-paper-scissors, or a search-versus-hide problem is often modeled this way.

A **partial-conflict game** allows both players to gain or lose together. Advertising wars, arms races, pricing decisions, and environmental agreements often have this structure.

Pure strategies in total-conflict games

Core explanation, worked reasoning, and application

Pure strategies in total-conflict games

In a two-player zero-sum game, we write a payoff matrix from the row player's point of view. The **row player** wants larger numbers; the **column player** wants smaller numbers.

Concept: Pure strategies, maximin, minimax

A **pure strategy** is a fixed choice, such as always choosing Row 1. For a payoff matrix $A = (a_{ij})$:

$$\text{row player's maximin} = \max_i \min_j a_{ij}, \qquad \text{column player's minimax} = \min_j \max_i a_{ij}.$$

If these two numbers are equal, the game has a **saddle point**. The corresponding cell is a stable pure-strategy solution: neither player can improve by switching alone.

Worked Example: Worked Example 1: Location competition with a saddle point

Two competing stores choose either a large district L or a small district S. The entry gives Store A's market share percentage.

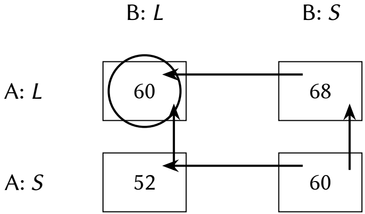
Store A \ Store B	**B:** L	**B:** S	**Row minimum**
A: L	60	68	60
A: S	52	60	52
Column maximum	60	68	

The row player guarantees at least 60 by choosing L. The column player can hold A to at most 60 by choosing L. Hence

$$\max\{60, 52\} = 60 = \min\{60, 68\}.$$

There is a saddle point at (L, L) with game value $V = 60$. The conclusion is not merely "both choose L"; it is: A can guarantee 60% market share, and B can prevent A from exceeding 60%.

Illustration: Pure strategies in total-conflict games



arrows point into
the saddle point

Movement diagram for a pure-strategy saddle point. Vertical arrows show A's preference; horizontal arrows show B's preference.

Checkpoint: Pure strategy check

For the matrix below, find the row minima, column maxima, and decide whether a saddle point exists.

$$\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

Explain the result in words, not only by calculation.

Worked Example: Worked Example 2: Dominated strategies

Before solving a game, look for strategies that a rational player should never use. A row is **dominated** if another row gives the row player at least as much payoff in every column and strictly more in at least one column. A column is dominated if another column gives the row player at most as much payoff in every row. Consider the row player's payoff matrix:

	C_1	C_2	C_3
R_1	4	2	5
R_2	3	1	4
R_3	1	6	2

Row R_1 dominates R_2 because $4 > 3$, $2 > 1$, and $5 > 4$. Thus the row player should never choose R_2. After removing it, compare columns C_1, C_2, C_3 from the column player's viewpoint. Since the column player wants smaller payoffs, a column with larger entries is unattractive.

Dominance does not always solve the game, but it simplifies the model and often reveals strategic structure before calculation.

Model Critique: Do not remove strategies too quickly

A dominated strategy can be removed only when dominance is clear across every possible response of the opponent. Removing a strategy because it is "usually bad" is not valid. In modeling reports, state the dominance comparison explicitly.

When no pure strategy is stable: mixed strategies

Core explanation, worked reasoning, and application

When no pure strategy is stable: mixed strategies

If there is no saddle point, any fixed strategy can be exploited. The solution is to randomize. A **mixed strategy** assigns probabilities to pure strategies.

The central idea is not "being unpredictable" in a vague sense. The mathematical purpose is to make the opponent indifferent between their pure strategies, so that they cannot exploit a predictable bias.

Worked Example: Worked Example 3: Batter versus pitcher

A batter guesses Fastball F or Curve C. The pitcher throws Fastball or Curve. The entries are batting averages.

Batter \ Pitcher	**Pitcher: F**	**Pitcher: C**	**Row min**
Guess F	.400	.200	.200
Guess C	.100	.300	.100
Column max	.400	.300	

The maximin is .200, while the minimax is .300. Since they are unequal, no pure-strategy saddle point exists. Let x be the probability that the batter guesses Fastball. Then the batter's expected average if the pitcher throws Fastball is

$$E_F(x) = .400x + .100(1-x) = .100 + .300x.$$

If the pitcher throws Curve,

$$E_C(x) = .200x + .300(1-x) = .300 - .100x.$$

The batter wants to maximize the lower of these two lines. The best choice occurs where the two lines meet:

$$.100 + .300x = .300 - .100x \quad \Rightarrow \quad x = 0.5.$$

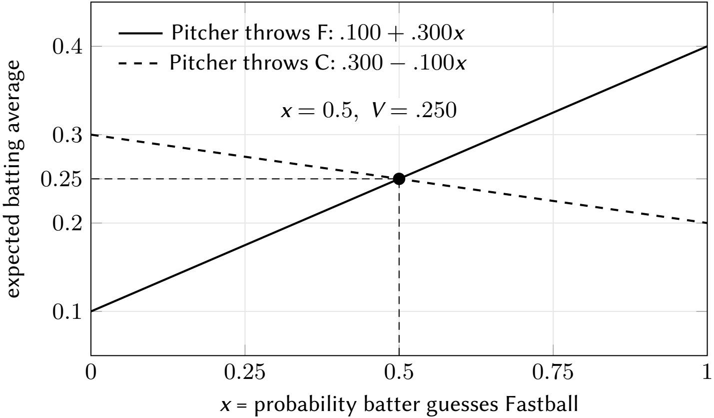
Worked Example: Worked Example 3: Batter versus pitcher

The game value is

$$V = .100 + .300(0.5) = .250.$$

So the batter should guess Fastball with probability 0.5 and Curve with probability 0.5. This guarantees a batting average of $.250$ in the long run.

Illustration: When no pure strategy is stable: mixed strategies



The batter chooses the mixture that maximizes the lower envelope of the two opponent-response lines.

Concept: Solving a 2×2 zero-sum game by indifference

For a payoff matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

let x be the probability that the row player chooses Row 1. To find the row player's optimal mix, equate the column player's two pure responses:

$$ax + c(1-x) = bx + d(1-x).$$

To find the column player's optimal mix, let y be the probability that the column player chooses Column 1, and equate the row player's two pure responses:

$$ay + b(1-y) = cy + d(1-y).$$

This method works after checking that no saddle point exists.

Worked Example: Worked Example 4: Predator and prey

A prey animal chooses Hide or Run. A predator chooses Ambush or Pursue. Entries are the prey's survival probabilities.

Prey \ Predator	**Ambush**	**Pursue**
Hide	0.30	0.60
Run	0.50	0.40

There is no saddle point. Let x be the probability the prey Hides. If the predator Ambushes,

$$E_A(x) = 0.30x + 0.50(1 - x) = 0.50 - 0.20x.$$

If the predator Pursues,

$$E_P(x) = 0.60x + 0.40(1 - x) = 0.40 + 0.20x.$$

Equating them gives

$$0.50 - 0.20x = 0.40 + 0.20x \quad \Rightarrow \quad x = 0.25.$$

The prey should Hide 25% and Run 75%. The guaranteed survival probability is

$$V = 0.50 - 0.20(0.25) = 0.45.$$

Worked Example: Worked Example 4: Predator and prey

For the predator, let y be the probability of Ambush. Equating the prey's two rows:

$$0.30y + 0.60(1 - y) = 0.50y + 0.40(1 - y) \quad \Rightarrow \quad y = 0.5.$$

The predator should Ambush 50% and Pursue 50%.

Model Critique: What does a mixed strategy mean?

A mixed strategy does not mean that the player is confused. It means that the player deliberately randomizes to prevent exploitation. The result is meaningful when the interaction is repeated many times, or when unpredictability itself has strategic value. In a single high-stakes play, a player may still choose a pure action based on information not represented in the matrix.

Games against nature: conservative guarantees

Core explanation, worked reasoning, and application

Games against nature: conservative guarantees

Sometimes the “opponent” is not intelligent: market demand, economy, weather, or machine failure. If probabilities are known, Lecture 7 expected value is appropriate. If probabilities are unknown and the decision maker wants a guaranteed floor, we may treat nature as adversarial.

Worked Example: Worked Example 5: Small or large production

A firm chooses Small or Large production. The economy may be Poor or Good. Payoffs are in thousands of dollars.

Firm \ Economy	**Poor**	**Good**	**Worst case**
Small	500	300	300
Large	100	900	100

A pure maximin decision chooses Small and guarantees 300. But a mixed strategy may guarantee more. Let x be the probability of Small. The guaranteed payoff against Poor is

$$E_P(x) = 500x + 100(1 - x) = 100 + 400x.$$

Against Good,

$$E_G(x) = 300x + 900(1 - x) = 900 - 600x.$$

Equate them:

$$100 + 400x = 900 - 600x \quad \Rightarrow \quad x = 0.8.$$

Worked Example: Worked Example 5: Small or large production

The firm should use Small 80% and Large 20%, guaranteeing

$$V = 100 + 400(0.8) = 420.$$

This is better than the pure maximin guarantee of 300.

Checkpoint: Decision theory or game against nature?

In the production example, suppose reliable probabilities are available: $P(\text{Poor}) = 0.30$ and $P(\text{Good}) = 0.70$. Should the firm still use the conservative mixed strategy? Compute the expected value of Small and Large and compare with the game-theoretic guarantee.

Partial conflict and Nash equilibrium

Core explanation, worked reasoning, and application

Partial conflict and Nash equilibrium

Zero-sum games are only one type of strategic model. In many real situations, both players may benefit from cooperation, but each also has an incentive to defect.

For partial-conflict games, we write payoffs as ordered pairs (row payoff, column payoff).

Concept: Nash equilibrium

A strategy pair is a **Nash equilibrium** if no player can improve their own payoff by switching strategy unilaterally. In a movement diagram, a Nash equilibrium is a cell into which all relevant preference arrows point.

Worked Example: Worked Example 6: Prisoner's Dilemma as an arms race

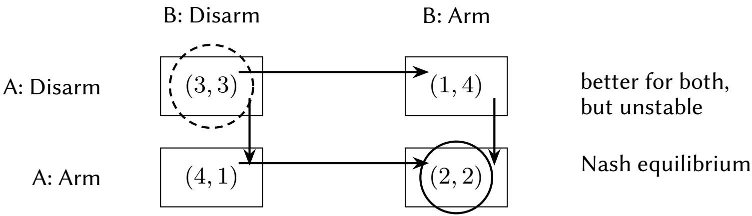
Two countries choose Disarm or Arm. Payoffs are preference scores, where 4 is best and 1 is worst.

Country A $\setminus$ Country B	**B Disarms**	**B Arms**
A Disarms	$(3,3)$	$(1,4)$
A Arms	$(4,1)$	$(2,2)$

If B disarms, A prefers to arm because $4 > 3$. If B arms, A still prefers to arm because $2 > 1$. Thus Arm is A's dominant strategy. By symmetry, Arm is also B's dominant strategy. The Nash equilibrium is therefore (Arm, Arm) with payoff $(2,2)$.

However, (Disarm, Disarm) gives $(3,3)$, which is better for both. This is the dilemma: individually rational choices lead to a collectively worse outcome.

Illustration: Partial conflict and Nash equilibrium



In the Prisoner's Dilemma, the Nash equilibrium is not the socially best outcome.

Model Critique: A common modeling mistake

Do not call every conflict a Prisoner's Dilemma. A true Prisoner's Dilemma requires: each player has a dominant strategy to defect; mutual defection is the Nash equilibrium; mutual cooperation is better for both than mutual defection. Many conflicts are coordination games, chicken games, or bargaining games instead.

Differentiated strategic modeling studio

Core explanation, worked reasoning, and application

Activity: Strategic modeling studio

Choose one route according to readiness. All groups should finish with a short written recommendation.

Core route: Pricing competition. Two drink stalls choose Low or High prices. The payoffs are weekly profits in hundreds of dollars.

	Stall B Low	Stall B High
Stall A Low	$(6, 6)$	$(10, 3)$
Stall A High	$(3, 10)$	$(8, 8)$

Draw a movement diagram and identify any Nash equilibrium. Is the result socially efficient?

Challenge route: Inspection game. A student may Study or Not Study. A teacher may Check or Not Check. From the student's point of view, payoffs are:

	Check	Not Check
Study	4	3
Not Study	-2	5

Treat this as a zero-sum simplification. Find the student's optimal mixed strategy and the value of the game.

Extension route: General formulation. For the matrix

$$\begin{pmatrix} 5 & 2 & 4 \\ 1 & 6 & 3 \end{pmatrix},$$

Activity: Strategic modeling studio

write the row player's linear program to maximize the guaranteed payoff V. You do not need to solve it, but each constraint must have a clear meaning.

Checkpoint: Competition-style communication

Write a four-sentence conclusion for one game in this lecture:

1. identify the recommended strategy;
2. state the game value or equilibrium outcome;
3. explain the strategic reason behind the recommendation;
4. state one limitation of the model.

Avoid writing only “we choose Row 1” without explaining what the opponent is expected to do.

Lesson summary

Core explanation, worked reasoning, and application

Summary: Eight ideas to retain

1. Game theory models decisions in which other rational players react.
2. A payoff matrix must state whose payoff is being recorded.
3. In a zero-sum game, the row player maximizes and the column player minimizes.
4. A pure-strategy saddle point exists when maximin equals minimax.
5. Without a saddle point, mixed strategies prevent exploitation.
6. In a 2×2 zero-sum game, optimal mixing is found by making the opponent indifferent.
7. In partial-conflict games, Nash equilibria can be stable but inefficient.
8. Game theory gives strategic insight only if the payoff matrix and rationality assumptions are credible.

Checkpoint: Exit ticket

Complete the sentence:
“A game-theory model is useful when ______________________________, but it may fail when ______________________________.”

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. For each zero-sum matrix, find row minima, column maxima, and decide whether a saddle point exists.

$$A = \begin{pmatrix} 5 & 2 \\ 3 & 6 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 5 \\ 3 & 2 \end{pmatrix}.$$

2. A goalkeeper dives Left or Right. A kicker shoots Left or Right. Entries are scoring probabilities:

	Goalkeeper L	Goalkeeper R
Kick L	0.55	0.90
Kick R	0.85	0.60

   Check for a saddle point. If none exists, find the kicker's optimal mixed strategy and the game value.

3. A firm faces uncertain demand and chooses Conservative or Aggressive production. Payoffs are in thousands:

	Low demand	High demand
Conservative	40	60
Aggressive	10	100

   Treat demand as nature. Find the best pure maximin decision. Then find whether a mixed strategy can guarantee more.

Exercises: Core practice

4. Draw the movement diagram for the pricing game in the studio. Is there a dominant strategy for either stall? Is the Nash equilibrium Pareto efficient?
5. In the Prisoner's Dilemma matrix from this lecture, explain why (Disarm, Disarm) is not a Nash equilibrium even though it is better for both countries.
6. Give one real-world situation that should *not* be modeled as zero-sum. Explain what would be lost if it were forced into a zero-sum payoff matrix.

Exercises: Modeling challenge

Choose one scenario and prepare a one-page strategic modeling brief containing players, strategies, payoff table, assumptions, equilibrium or recommended strategy, and limitations.

1. Two schools decide whether to host their open day on the same weekend or different weekends.
2. Two delivery companies choose whether to enter a new district.
3. A student team chooses a competition topic while anticipating what other teams may choose.

Your brief should justify the payoff values qualitatively even if exact numerical data are unavailable.

Exercises: Extension practice

1. **Battle of the search routes.** A rescue team searches North or South. A lost hiker chooses, unknowingly, one of two paths. The payoff to the rescue team is probability of finding the hiker within one hour:

	Hiker North	Hiker South
Search North	0.80	0.30
Search South	0.40	0.70

   Treat this as a game against nature. Find the pure maximin strategy and the mixed strategy that maximizes the guaranteed probability of success.

2. **Penalty kick interpretation.** A team claims, "Our kicker should always kick right because the average scoring rate is higher." Explain why this argument may fail in a strategic game. Use the goalkeeper's possible response in your explanation.
3. **Dominance with context.** In a pricing game, one row is dominated mathematically. Give a possible real-world reason why a firm might still consider it, and explain whether that means the payoff matrix is incomplete.
4. **Repeated interaction.** Revisit the drink-stall pricing game. Explain how repeated play, reputation, or a school regulation might change the equilibrium outcome. State which assumption of the one-shot game has changed.

Exercises: Strategic modeling brief template

For a competition-style response, students may use the following structure:

1. **Players:** Who makes strategic decisions?
2. **Strategies:** What actions are realistically available to each player?
3. **Payoffs:** What does each player value, and how are numerical payoffs justified?
4. **Solution concept:** Is the game zero-sum, a game against nature, or partial conflict? Which method is appropriate?
5. **Result:** State the saddle point, mixed strategy, game value, or Nash equilibrium.
6. **Interpretation:** What does the result recommend in the real situation?
7. **Limitations:** Which assumptions are most fragile?

Optional extension: Linear programming formulation for larger zero-sum games

Core explanation, worked reasoning, and application

Optional extension: Linear programming formulation for larger zero-sum games

For an $m \times n$ zero-sum matrix $A = (a_{ij})$, let x_i be the probability that the row player chooses row i. The row player maximizes the guaranteed payoff V:

$$
\begin{aligned}
\text{maximize} \quad & V \\
\text{subject to} \quad & V \le \sum_{i=1}^{m} a_{ij} x_i, \qquad j = 1, \ldots, n, \\
& \sum_{i=1}^{m} x_i = 1, \\
& x_i \ge 0.
\end{aligned}
$$

Each inequality says: even if the column player chooses column j, the expected payoff must be at least V. The column player's dual problem minimizes V subject to analogous constraints. The equality of the two optimal values is closely related to the minimax theorem.

Optional extension: Repeated Prisoner’s Dilemma

Core explanation, worked reasoning, and application

Optional extension: Repeated Prisoner's Dilemma

If the Prisoner's Dilemma is played once, defection is individually rational. If it is repeated many times, cooperation can become stable. A simple strategy is **Tit-for-Tat**: cooperate first, then copy the opponent's previous move. It rewards cooperation and punishes defection, but only when players expect future interaction and can observe one another's actions reliably.

Lecture 9: Differential Equations and Dynamic Models

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, readers should be able to:

1. explain when a **differential equation** is more appropriate than a difference equation;
2. build first-order models from the idea “rate in minus rate out” or “rate proportional to current amount”;
3. solve and interpret exponential, logistic, cooling, and simple pollution models;
4. use a **phase line** to classify **equilibria** as stable or unstable;
5. apply Euler, improved Euler/Heun, and RK4 ideas to approximate dynamic models;
6. read a basic two-population **phase plane** using nullclines;
7. communicate assumptions, parameters, validation evidence, and limitations of a dynamic model.

Summary: Two-hour lesson route

0–12 min	Launch: from discrete change to continuous rate.
12–32 min	Exponential growth and decay: build, solve, and criticize.
32–52 min	Logistic refinement and phase-line reasoning.
52–70 min	Cooling and pollution as rate-balance models.
70–98 min	Numerical solution: Euler as a baseline, improved Euler, RK4, step-size checks, and stability.
98–110 min	Differentiated studio: dose interval, lake cleanup, or predator-prey phase plane.
110–120 min	Competition-style conclusion, exit ticket, and homework briefing.

Why differential equations?

Core explanation, worked reasoning, and application

Why differential equations?

A **difference equation** updates a system at fixed steps:

$$P_{n+1} = P_n + kP_n.$$

A **differential equation** describes the instantaneous rate of change:

$$\frac{dP}{dt} = kP.$$

The second form is appropriate when the change is naturally continuous, or when the time step is so small that a continuous approximation is easier to interpret.

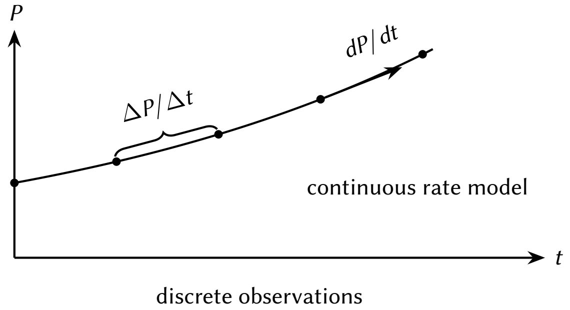
Activity: Launch: choosing the time language

For each situation, decide whether a difference equation or a differential equation is more natural. Give one reason.

1. A bank account receives interest once per month.
2. A hot drink cools continuously after being placed on a table.
3. A school records the number of absent students each morning.
4. A pollutant is continuously flushed out of a lake by a river.

Modeling lesson: The mathematics should match the timing of the mechanism, not merely the data table we happen to have.

Illustration: Why differential equations?



A differential equation idealizes many small updates as an instantaneous rate law.

Concept: Core vocabulary

A **first-order ordinary differential equation** has the form

$$\frac{dy}{dt} = f(t, y).$$

An **initial value problem** adds a starting value $y(t_0) = y_0$. A value y^* with $f(y^*) = 0$ is an **equilibrium**. If nearby solutions move toward it, it is **stable**; if they move away, it is **unstable**.

Exponential growth and decay

Core explanation, worked reasoning, and application

Exponential growth and decay

The simplest rate law says that the rate of change is proportional to the present amount:

$$\frac{dP}{dt} = kP.$$

If $k > 0$, the quantity grows exponentially; if $k < 0$, it decays exponentially.

Worked Example: Worked Example 1: Population growth and the danger of extrapolation

Suppose a population is 203.2 million in 1970 and 248.7 million in 1990. Let t be years after 1970 and assume

$$\frac{dP}{dt} = kP, \qquad P(0) = 203.2.$$

Separation gives

$$\frac{dP}{P} = k\,dt \quad \Rightarrow \quad \ln P = kt + C \quad \Rightarrow \quad P(t) = 203.2e^{kt}.$$

Use the 1990 value to estimate k:

$$248.7 = 203.2e^{20k} \quad \Rightarrow \quad k = \frac{1}{20}\ln\left(\frac{248.7}{203.2}\right) \approx 0.0101.$$

The 2000 prediction is

$$P(30) = 203.2e^{0.0101(30)} \approx 274.9 \text{ million}.$$

This is a reasonable short-term estimate, but the same model predicts unlimited growth. The equation is useful only while resources, policy, and demographics remain close to the calibration period.

Model Critique: Model criticism: what does k hide?

The parameter k compresses birth rates, death rates, immigration, age structure, economic conditions, policy, and culture into one number. A good report should say not only that $k \approx 0.0101$, but also why a constant k is a simplification and how far into the future the assumption remains credible.

Logistic growth and phase-line reasoning

Core explanation, worked reasoning, and application

Logistic growth and phase-line reasoning

Exponential growth ignores limited resources. A common refinement is the **logistic model**:

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{M}\right),$$

where M is the **carrying capacity**. The factor $1 - P/M$ reduces growth as P approaches M.

Worked Example: Worked Example 2: Choosing between exponential and logistic models

A fish population in a closed pond is initially $P(0) = 80$. Ecologists believe the pond can support at most $M = 500$ fish. They estimate $r = 0.40$ per month.

$$\frac{dP}{dt} = 0.40P\left(1 - \frac{P}{500}\right).$$

The logistic solution is

$$P(t) = \frac{M}{1 + Ae^{-rt}}, \qquad A = \frac{M - P_0}{P_0}.$$

Here $A = (500 - 80)/80 = 5.25$, so

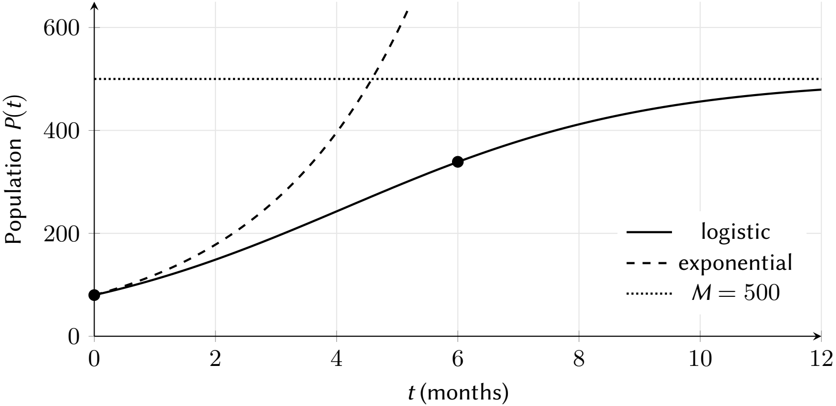
$$\boxed{P(t) = \frac{500}{1 + 5.25e^{-0.40t}}.}$$

At $t = 6$ months,

$$P(6) = \frac{500}{1 + 5.25e^{-2.4}} \approx 339.$$

The exponential model with the same initial growth rate gives $80e^{0.40(6)} \approx 882$, which already exceeds the pond's capacity. The logistic model is more credible because it encodes the physical constraint.

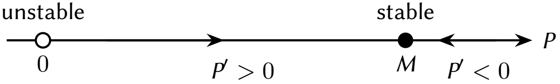
Illustration: Logistic growth and phase-line reasoning



The logistic model behaves like exponential growth at first, but bends toward the carrying capacity.

Concept: Phase line for the logistic model

For the autonomous equation $P' = rP(1 - P/M)$, equilibria occur at $P = 0$ and $P = M$. If $0 < P < M$, then $P' > 0$; if $P > M$, then $P' < 0$. Therefore 0 is unstable and M is stable.



Checkpoint: Phase-line check

For $y' = y(3 - y)$:

1. find all equilibria;
2. determine the sign of y' on each interval;
3. classify each equilibrium as stable or unstable;
4. explain the long-run behavior if $y(0) = 1$, $y(0) = 5$, and $y(0) = -1$.

Rate-balance models: cooling and pollution

Core explanation, worked reasoning, and application

Rate-balance models: cooling and pollution

Many differential equation models are built from the principle

$$\text{rate of change} = \text{rate in} - \text{rate out}.$$

This principle is easier for students to remember than a list of formula types.

Worked Example: Worked Example 3: Newton's law of cooling

A cup of tea is 90°C when placed in a room at 25°C. Newton's law of cooling assumes that the cooling rate is proportional to the temperature difference:

$$\frac{dT}{dt} = -k(T-25), \qquad T(0) = 90.$$

The solution is

$$T(t) = 25 + 65e^{-kt}.$$

If after 10 minutes the temperature is 55°C, then

$$55 = 25 + 65e^{-10k} \quad \Rightarrow \quad e^{-10k} = \frac{30}{65} \quad \Rightarrow \quad k \approx 0.0773.$$

To find when the tea reaches 40°C:

$$40 = 25 + 65e^{-0.0773t} \quad \Rightarrow \quad t \approx 19.0 \text{ minutes}.$$

The stable equilibrium is the ambient temperature 25°C.

Worked Example: Worked Example 4: Lake cleanup as rate in minus rate out

A lake has volume $V = 10^9$ L. Clean water flows in and out at $r = 5 \times 10^7$ L/year. Let $p(t)$ be the amount of pollutant in grams. If no new pollutant enters, the pollutant leaves at concentration p/V:

$$\frac{dp}{dt} = 0 - r\frac{p}{V} = -\frac{r}{V}p = -0.05p.$$

So

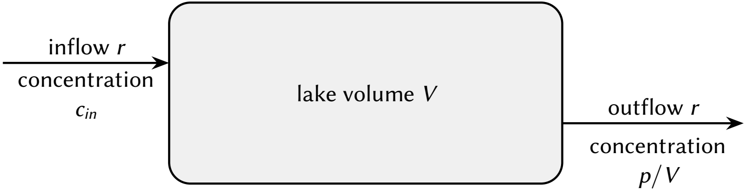
$$p(t) = p_0 e^{-0.05t}.$$

The time to reach 10% of the initial pollutant is

$$0.10p_0 = p_0 e^{-0.05t} \quad \Rightarrow \quad t = \frac{\ln 10}{0.05} \approx 46.1 \text{ years.}$$

The model explains why environmental cleanup can be slow even after pollution input stops: the relevant time scale is the **flushing time** $V/r = 20$ years.

Illustration: Rate-balance models: cooling and pollution



$$p'(t) = rc_{in} - r\,p(t)/V$$

A lake pollution model follows directly from rate in minus rate out.

Model Critique: Validation and limitations

A rate-balance model assumes perfect mixing: the pollutant concentration is the same everywhere in the lake. In a large lake, this may be false. A report should discuss whether the lake is well mixed, whether the inflow and outflow rates vary seasonally, and whether pollutant can settle into sediment and later re-enter the water.

Worked Example: Worked Example 5: Repeated drug dosage

A medicine is effective above $L = 2$ mg/L and unsafe above $H = 8$ mg/L. After a dose, the blood concentration decays according to

$$C' = -kC, \qquad k = 0.18\ \mathrm{hr}^{-1}.$$

If each dose raises concentration by $H - L = 6$ mg/L, the residual concentration just before the next dose should approach L. For repeated dosing at interval T, the limiting residual is

$$R = \frac{H - L}{e^{kT} - 1}.$$

Setting $R = L$ gives

$$L = \frac{H - L}{e^{kT} - 1} \quad \Rightarrow \quad e^{kT} = \frac{H}{L} \quad \Rightarrow \quad \boxed{T = \frac{1}{k}\ln\left(\frac{H}{L}\right)}.$$

Here

$$T = \frac{1}{0.18}\ln 4 \approx 7.70 \text{ hours}.$$

Interpretation: an 8-hour schedule is close but slightly slower than the ideal schedule; it may be acceptable only if clinical safety margins and patient variation are considered.

Checkpoint: Dosage interpretation

If k increases because the body eliminates the drug faster, should the recommended interval T increase or decrease? Explain using the formula and in ordinary language.

Numerical solution: from Euler to reliable computation

Core explanation, worked reasoning, and application

Numerical solution: from Euler to reliable computation

Analytic formulas are useful, but many modeling ODEs cannot be solved neatly. In competitions and applied projects, it is often more realistic to state the differential equation clearly and then compute the solution numerically. The issue is not merely “press a button”. Students must understand what the numerical method is doing, how accurate it is likely to be, and how to check whether the output is trustworthy.

Concept: Numerical solution as a modeling tool

A **numerical solution** approximates the values of an unknown solution at a grid of times

$$t_0, t_1 = t_0 + h, t_2 = t_0 + 2h, \ldots$$

where h is the **step size**. A numerical method turns the rate law

$$y' = f(t, y)$$

into an update rule for y_{n+1} from information near (t_n, y_n). The smaller the step size, the more often the model is updated, but the more computation is required.

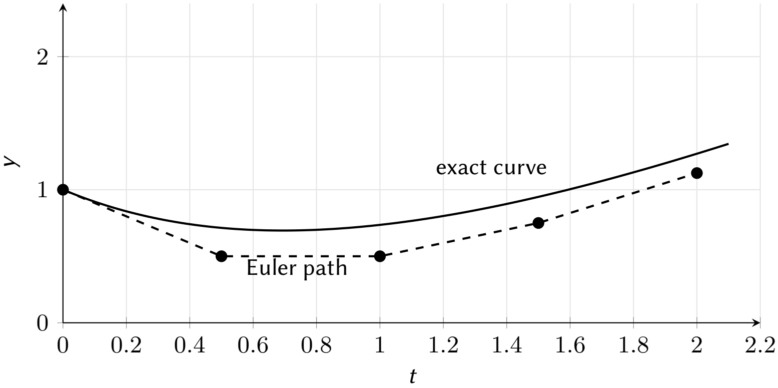
Euler's method as the baseline

Euler's method follows the tangent line at the current point:

$$\boxed{t_{n+1} = t_n + h, \qquad y_{n+1} = y_n + hf(t_n, y_n).}$$

It is valuable because it reveals the mechanism of numerical integration. However, it is only first-order accurate: the **global error** is usually $O(h)$, so halving h often roughly halves the error.

Illustration: Euler's method as the baseline



Euler's method replaces a curved solution by short tangent-line steps.

Worked Example: Worked Example 6: Euler approximation and step-size check

Approximate the solution of

$$y' = t - y, \qquad y(0) = 1,$$

on $0 \le t \le 2$ using $h = 0.5$.

$$y_{n+1} = y_n + 0.5(t_n - y_n).$$

n	t_n	y_n	slope $t_n - y_n$	y_{n+1}
0	0.0	1.000	-1.000	0.500
1	0.5	0.500	0.000	0.500
2	1.0	0.500	0.500	0.750
3	1.5	0.750	0.750	1.125

The exact solution is $y = t - 1 + 2e^{-t}$. At $t = 2$, exact $y(2) \approx 1.271$, so Euler gives error about 0.146.
Step-size check. If we repeat Euler with $h = 0.25$, the estimate at $t = 2$ is approximately 1.200, with error about 0.070. The error is roughly halved, which is consistent with first-order convergence.

Checkpoint: Numerical method check

Why is a single Euler computation not enough evidence? State two checks a modeler should perform before trusting a numerical ODE result.

Improved Euler / Heun's method

Euler uses only the slope at the beginning of the step. If the slope changes across the interval, this can be crude. The **improved Euler method**, also called **Heun's method**, first predicts a trial endpoint, then averages the starting and ending slopes:

$$\begin{aligned} k_1 &= f(t_n, y_n), \\ y_{\text{pred}} &= y_n + hk_1, \\ k_2 &= f(t_n + h, y_{\text{pred}}), \\ y_{n+1} &= y_n + \frac{h}{2}(k_1 + k_2). \end{aligned}$$

This is a second-order method: its global error is typically $O(h^2)$.

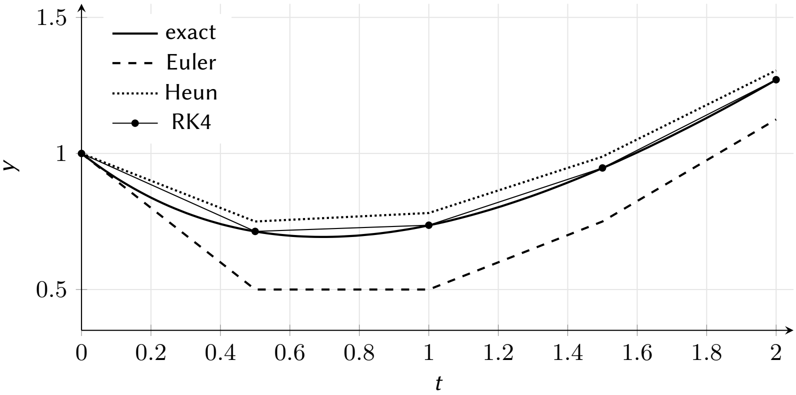
Worked Example: Worked Example 7: Euler versus Heun versus RK4

Use the same IVP $y' = t - y$, $y(0) = 1$, with $h = 0.5$ up to $t = 2$. The exact value is $y(2) = 1.27067$.

Method	Update idea	Estimate at $t = 2$	Error
Euler	starting slope only	1.12500	0.14567
Improved Euler / Heun	average of start and predicted-end slopes	1.30518	0.03451
RK4	weighted four-slope average	1.27110	0.00043

The main lesson is not that one method is always best, but that numerical method choice matters. Euler is excellent for explaining the idea; Heun and RK4 are better default tools for producing values in a report.

Illustration: Improved Euler / Heun's method



Different numerical methods may produce visibly different paths even with the same step size.

Fourth-order Runge-Kutta: the practical default

The classical **fourth-order Runge-Kutta method** balances accuracy and simplicity. It samples four slopes inside one step:

$$\begin{aligned}
k_1 &= f(t_n, y_n), \\
k_2 &= f\left(t_n + \frac{h}{2},\ y_n + \frac{h}{2}k_1\right), \\
k_3 &= f\left(t_n + \frac{h}{2},\ y_n + \frac{h}{2}k_2\right), \\
k_4 &= f(t_n + h,\ y_n + hk_3), \\
y_{n+1} &= y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4).
\end{aligned}$$

RK4 has global error $O(h^4)$ for sufficiently smooth problems. In student modeling reports, it is often enough to state that the ODE was solved by RK4 or by a standard adaptive solver, and then report a step-size or tolerance check.

Concept: A numerical ODE workflow for modeling reports

1. **Write the rate law clearly.** Define variables, units, parameters, and initial conditions.
2. **Choose a method.** Use Euler for classroom illustration; use Heun, RK4, or an adaptive solver for reported numerical results.
3. **Check convergence.** Repeat with h and $h/2$, or with a tighter solver tolerance. The conclusion should not change materially.
4. **Check against known behavior.** The numerical path should respect equilibria, signs, conservation, non-negativity, or carrying capacities.
5. **Separate numerical error from modeling error.** A better solver cannot fix a wrong mechanism.

Stability of a numerical method

Accuracy is not the only issue. A method can be unstable even for a simple stable differential equation. Consider

$$y' = -10y, \qquad y(0) = 1.$$

The true solution $y = e^{-10t}$ decays smoothly to 0. Euler's method gives

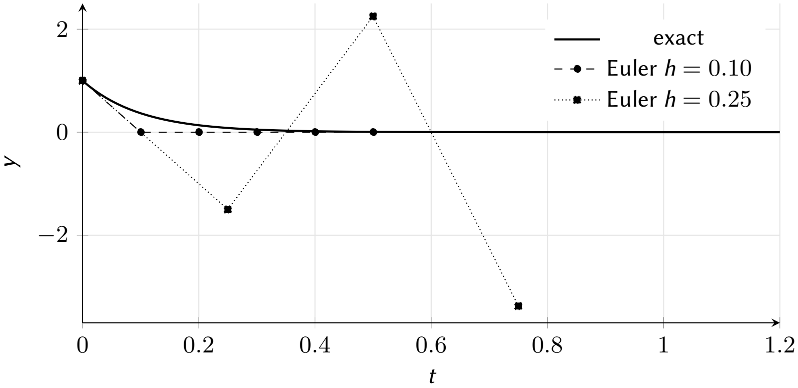
$$y_{n+1} = y_n + h(-10y_n) = (1 - 10h)y_n.$$

For the numerical solution to decay rather than explode, we need

$$|1 - 10h| < 1 \quad \Rightarrow \quad 0 < h < 0.2.$$

Thus a large step size may create artificial oscillation or blow-up even though the true model is stable.

Illustration: Stability of a numerical method



For a stable decay equation, too large a step size can create artificial numerical instability.

Model Critique: What to write in a competition report

Avoid writing only "we used Euler's method". A stronger statement is:

> *We solved the ODE numerically using RK4 with step size* $h = 0.05$*. Repeating the computation with* $h = 0.025$ *changed the predicted peak by less than 1%, so numerical error is small relative to parameter uncertainty. The model preserves non-negative populations and approaches the expected stable equilibrium.*

This sentence shows method, convergence, qualitative validation, and limitation.

Systems and phase planes: predator and prey

Core explanation, worked reasoning, and application

Systems and phase planes: predator and prey

For two interacting quantities, a single phase line is not enough. We use a **phase plane**. The key visual tools are **nullclines**: curves where one derivative is zero.

Worked Example: Worked Example 8: Predator-prey nullclines

Let $x(t)$ be prey and $y(t)$ be predator. A simple Lotka-Volterra model is

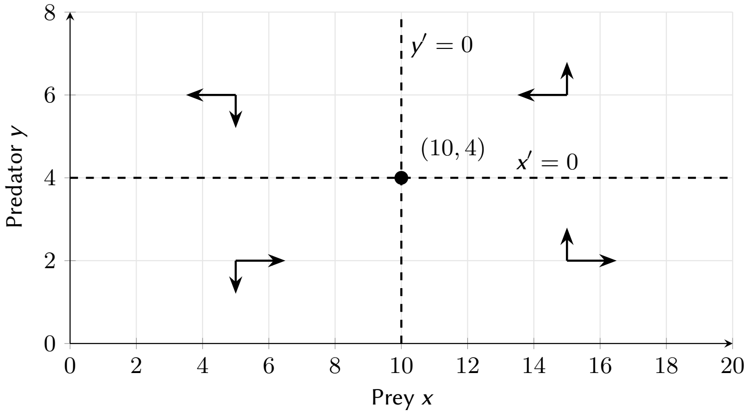
$$x' = x(2 - 0.5y), \qquad y' = y(-1 + 0.1x).$$

The x-nullclines satisfy $x' = 0$, so $x = 0$ or $y = 4$. The y-nullclines satisfy $y' = 0$, so $y = 0$ or $x = 10$. The positive equilibrium is

$$(x^*, y^*) = (10, 4).$$

If $y < 4$, prey increase; if $y > 4$, prey decrease. If $x > 10$, predators increase; if $x < 10$, predators decrease.

Illustration: Systems and phase planes: predator and prey



Nullclines divide the predator-prey phase plane into regions where each population increases or decreases.

Activity: Differentiated modeling studio

Choose one route.

1. **Core route: dosing interval.** A medicine has $L = 2$, $H = 8$, and $k = 0.18$ hr^{-1}. Use $T = (1/k)\ln(H/L)$ and explain the result in plain language.
2. **Challenge route: lake cleanup.** A lake has $V = 3 \times 10^9$ L and $r = 10^8$ L/year. Find the time to reduce pollutant to 25% of its initial amount, then discuss one assumption that may fail.
3. **Extension route: predator-prey.** For $x' = x(1 - y)$ and $y' = y(-2 + x)$, find the nullclines, the positive equilibrium, and the signs of x' and y' in the four regions.

Summary: Choosing a differential-equation method

Situation	Recommended method	Modeling question answered
$y' = ky$ or $y' = -ky$	separation / known formula	How fast does the amount grow or decay?
$y' = rP(1 - P/M)$	logistic formula + phase line	What long-run capacity is approached?
$y' = f(y)$	phase line	Which equilibria are stable?
$y' + P(t)y = Q(t)$	integrating factor	What is the balance between input and removal?
No clean formula, classroom calculation	Euler	What does one tangent-line approximation do?
No clean formula, report-quality computation	Heun / RK4 / adaptive solver	Is the numerical prediction stable under step-size refinement?
Two interacting variables	nullclines / phase plane	Which population increases or decreases in each region?

Model Critique: Numerical error versus modeling error

If Euler's method gives a poor estimate, there are two possible causes. **Numerical error** means the step size is too large for the differential equation. **Modeling error** means the differential equation itself is not a good representation of the real system. Reducing h helps only with numerical error; it cannot fix a bad assumption such as constant temperature, perfect mixing, or unlimited growth.

Writing dynamic-model conclusions

Core explanation, worked reasoning, and application

Model Critique: Competition-style communication

A dynamic-model paragraph should contain four parts:

1. **Mechanism:** what rate law or balance law was assumed?
2. **Calibration:** how were parameters such as k, r, M, or V/r obtained?
3. **Prediction and interpretation:** what does the model imply in context?
4. **Credibility check:** what evidence supports the model, and where does it likely fail?

Weak conclusion: "The logistic model predicts 339 fish."
Stronger conclusion: "Using a carrying capacity of 500 fish and an estimated monthly growth rate of 0.40, the logistic model predicts about 339 fish after six months. The exponential model exceeds the pond capacity over the same period, so it is unsuitable for long-range prediction. The main uncertainty is the carrying capacity estimate, which should be checked against pond area, food supply, and historical population counts."

Summary: Lesson summary

Concept	Modeling meaning
Differential equation	Relates a quantity to its instantaneous rate of change.
Initial condition	Selects one solution curve from a family.
Exponential model	Rate proportional to amount; useful locally, dangerous for long extrapolation.
Logistic model	Growth slows near carrying capacity; stable equilibrium at M.
Phase line	Visual tool for one-variable autonomous DEs; shows stability.
Rate-balance model	rate in − rate out; used for pollution, mixing, and flow problems.
Euler's method	First-order tangent-line approximation; useful as a baseline but often too crude for final results.
Heun / RK4	Higher-order numerical solvers; compare step sizes before reporting quantitative predictions.
Phase plane	Visual tool for systems; nullclines locate equilibria and direction regions.
Model criticism	State assumptions, parameter sources, validation evidence, and limits.

Checkpoint: Exit ticket

In three sentences, explain the difference between an exponential model and a logistic model. Then give one real-world situation where the logistic model is more credible.

Exercises: Core practice

1. Solve $y' = 0.3y$, $y(0) = 12$. Find $y(5)$ and interpret the parameter 0.3.
2. A chemical decays according to $C' = -0.12C$, $C(0) = 40$. When does it fall below 5?
3. For $P' = 0.2P(1 - P/1000)$ and $P(0) = 100$, write the logistic solution and estimate $P(10)$.
4. Draw the phase line for $y' = y(y - 4)$ and classify the equilibria.
5. A room is kept at 22°C. An object cools from 80°C according to $T' = -0.05(T - 22)$. Find $T(t)$ and $T(20)$.
6. Apply Euler's method with $h = 1$ to $y' = 2 - y$, $y(0) = 0$ for three steps. Then repeat with Heun's method and compare.

Exercises: Modeling challenge

A small reservoir initially contains 12 million litres of water with pollutant concentration 0.08 g/L. Clean water enters and exits at 0.6 million litres per day. A nearby factory may continue to add pollutant at an unknown constant rate s grams/day.

1. Define $p(t)$ and write a differential equation for pollutant amount.
2. Solve the model when $s = 0$.
3. Find the long-run pollutant level when $s > 0$.
4. Suppose measurements after 10 days show concentration 0.052 g/L. Estimate whether s is closer to 0 or significantly positive.
5. Write a short recommendation for environmental officers. Include limitations of the model.

Exercises: Extension practice

1. Use separation to solve $y' = ty^2$, $y(0) = 1$. Identify whether the solution exists for all $t > 0$.
2. For $y' = y(1-y)(y-3)$, draw the phase line and classify all equilibria.
3. For $x' = x(3-y)$, $y' = y(-2+x)$, find the positive equilibrium and the nullclines. Give the direction of motion in each region.
4. Explain why Euler's method may become inaccurate or unstable if h is too large, even when the differential equation is correct. Use $y' = -10y$ as an example.

Summary: Optional appendix: integrating factor method

A first-order linear equation has standard form

$$y' + P(t)y = Q(t).$$

The integrating factor is

$$\mu(t) = e^{\int P(t)\,dt}.$$

Multiplying the equation by $\mu(t)$ turns the left side into a product derivative:

$$(\mu y)' = \mu Q.$$

Thus

$$\mu(t)y(t) = \int \mu(t)Q(t)\,dt + C.$$

This method is useful for pollution, mixing, heating/cooling with input, and many first-order engineering models.

Exercises: Additional problem bank

1. **Coffee cooling.** Coffee cools from $85°$C to $60°$C in 12 minutes in a $24°$C room. Estimate k and predict when it reaches $45°$C.
2. **Choosing a model.** A population grows from 100 to 180 to 250 to 310 over four equal time intervals, and experts believe the maximum is about 400. Explain why an exponential model may overpredict and propose a logistic model structure.
3. **Parameter sensitivity.** For the lake cleanup model $p(t) = p_0 e^{-rt/V}$, explain how the cleanup time changes if the inflow-outflow rate r doubles.
4. **Euler design.** A student uses $h = 5$ days to simulate a fast-changing epidemic model. What checks should be performed before trusting the output?
5. **Model comparison.** Write two sentences comparing a phase-line conclusion and a numerical simulation conclusion. What can each reveal that the other may hide?

Lecture 10: Continuous Optimization and Model Design

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, readers should be able to:

1. distinguish **continuous optimization** from the discrete and linear-programming models in Lecture 6;
2. construct a single-variable objective function from submodels, especially in the **economic order quantity** model;
3. use first- and second-derivative tests to find and interpret an optimum;
4. solve a two-variable unconstrained optimization problem using **partial derivatives** and the **Hessian**;
5. explain the geometric idea behind **Lagrange multipliers**;
6. interpret a multiplier as a **shadow price**;
7. describe when numerical optimization is needed and how to report solver checks in a modeling paper.

Summary: Two-hour lesson route

0–12 min	Launch: what is a "best" continuous decision? Identify variable, objective, constraints.
12–37 min	Single-variable optimization through the EOQ inventory model.
37–57 min	Marginal reasoning, convexity, second derivative, and sensitivity.
57–78 min	Two-variable optimization: profit surface, gradient, Hessian, and model interpretation.
78–98 min	Constrained optimization: Lagrange multipliers and shadow-price meaning.
98–112 min	Numerical optimization: gradient search, line search, local optima, and solver reporting.
112–120 min	Differentiated studio, model criticism, and exit ticket.

From choosing among options to choosing a best value

Core explanation, worked reasoning, and application

From choosing among options to choosing a best value

Lecture 6 optimized decisions over a polygon or a list of integer choices. Here the decision variable can vary continuously. A delivery interval might be 2.8 days, a price might be \$37.40, and a mixing proportion might be 0.62. The modeling task is not simply to differentiate a function: it is to build the function that deserves to be optimized.

Activity: Launch: a decision is not yet a model

A school shop wants to decide how many bottles of water to order at a time. Ordering too often creates administrative cost; ordering too many creates storage cost and risk of waste.

1. What is the decision variable?
2. What should the objective be: total cost per order, cost per day, profit, or service reliability?
3. Which assumptions would make the problem continuous rather than integer?
4. What data would you need before trusting the recommendation?

Modeling lesson: continuous optimization starts before calculus. The most important step is deciding what quantity is being optimized and what simplifications make the objective computable.

Concept: Core vocabulary

A **decision variable** is a controllable quantity. An **objective function** measures what we want to maximize or minimize. A **constraint** restricts feasible decisions. A **critical point** is a point where the derivative or gradient is zero. A **global optimum** is best among all feasible choices; a **local optimum** is only best nearby.

Single-variable optimization: the EOQ model

Core explanation, worked reasoning, and application

Single-variable optimization: the EOQ model

The **economic order quantity** model is useful because it is simple, teachable, and rich in modeling interpretation. It balances two costs that move in opposite directions: frequent ordering and excessive storage.

Worked Example: Worked Example 1: Deriving the EOQ formula

A station sells fuel at a constant rate r gallons per day. Every delivery has fixed cost d, and holding one gallon for one day costs s. Let T be the delivery cycle length in days.
Step 1: connect the decision to the order size. If the station orders exactly enough fuel for T days, then

$$Q = rT.$$

Step 2: build cost per cycle. Delivery cost per cycle is d. The inventory decreases approximately linearly from Q to 0, so the average inventory is $Q/2 = rT/2$. The storage cost per cycle is

$$s \cdot \frac{rT}{2} \cdot T = \frac{srT^2}{2}.$$

Step 3: optimize average daily cost.

$$c(T) = \frac{d + srT^2/2}{T} = \frac{d}{T} + \frac{srT}{2}, \qquad T > 0.$$

Then

$$c'(T) = -\frac{d}{T^2} + \frac{sr}{2}.$$

Worked Example: Worked Example 1: Deriving the EOQ formula

Setting $c'(T) = 0$ gives

$$T^2 = \frac{2d}{sr}, \qquad \boxed{T^* = \sqrt{\frac{2d}{sr}}}.$$

Since

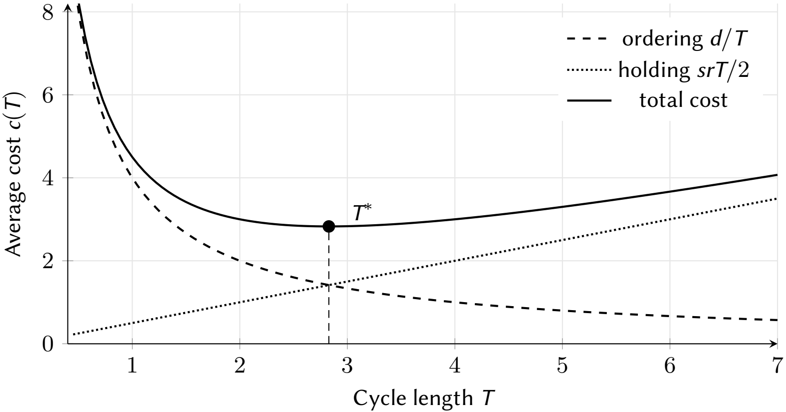
$$c''(T) = \frac{2d}{T^3} > 0,$$

this critical point is a minimum. The optimal order quantity is

$$\boxed{Q^* = rT^* = \sqrt{\frac{2dr}{s}}}.$$

Interpretation: high delivery cost pushes the model toward larger, less frequent orders; high storage cost pushes it toward smaller, more frequent orders.

Illustration: Single-variable optimization: the EOQ model



The EOQ optimum occurs where the decreasing order-cost curve and increasing holding-cost curve balance.

Checkpoint: Equal-cost principle

Show that at T^*,

$$\frac{d}{T^*} = \frac{srT^*}{2}.$$

Then explain why this is a useful way to check whether a numerical answer is reasonable.

Worked Example: Worked Example 2: Applying EOQ and checking sensitivity

A bookstore sells $r = 200$ textbooks per month. Each order costs $d = \$120$. Holding cost is $s = \$0.50$ per book per month.

$$T^* = \sqrt{\frac{2(120)}{0.50(200)}} = \sqrt{2.4} \approx 1.55 \text{ months},$$

so

$$Q^* = rT^* = 200(1.55) \approx 310 \text{ books}.$$

The minimum monthly cost is

$$c^* = \frac{120}{1.55} + \frac{0.50(200)(1.55)}{2} \approx \$154.9.$$

Sensitivity: if storage cost doubles to $s = 1.00$, then

$$T^*_{new} = \sqrt{\frac{2(120)}{1.00(200)}} = \sqrt{1.2} \approx 1.10 \text{ months}.$$

The cycle length is multiplied by $1/\sqrt{2} \approx 0.707$, not cut in half. Square-root laws are less sensitive than linear laws.

Model Critique: Model criticism: what EOQ assumes

The EOQ model assumes constant demand, no stockouts, instant delivery, no quantity discount, and divisible order sizes. In a modeling report, the EOQ result should be treated as a baseline. If demand is seasonal, delivery has delay, or order quantity must be an integer carton size, the final recommendation may require simulation or integer optimization.

Two-variable optimization: gradient and Hessian

Core explanation, worked reasoning, and application

Two-variable optimization: gradient and Hessian

For a function $f(x, y)$, the **gradient**

$$\nabla f = (f_x, f_y)$$

points in the direction of fastest increase. A critical point satisfies

$$f_x = 0, \qquad f_y = 0.$$

The **Hessian**

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

helps classify the point. For two variables, define

$$D = f_{xx} f_{yy} - f_{xy}^2.$$

If $D > 0$ and $f_{xx} < 0$, the point is a local maximum. If $D > 0$ and $f_{xx} > 0$, it is a local minimum. If $D < 0$, it is a saddle point.

Worked Example: Worked Example 3: Two-product profit optimization

A company produces two products. Let x and y be daily production quantities. Suppose the estimated profit is

$$P(x, y) = 100x + 120y - x^2 - 2y^2 - xy - 500.$$

The negative quadratic terms represent price pressure, congestion, or diminishing returns.

Find the critical point.

$$P_x = 100 - 2x - y, \qquad P_y = 120 - 4y - x.$$

Set both to zero:

$$2x + y = 100, \qquad x + 4y = 120.$$

Solving gives

$$x^* = 40, \qquad y^* = 20.$$

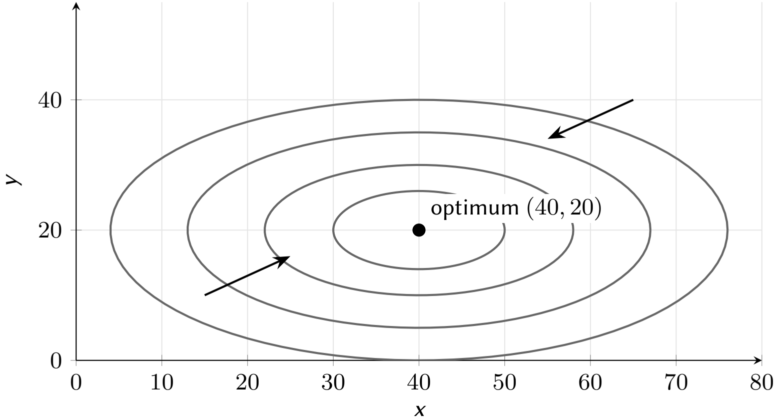
Classify.

$$P_{xx} = -2, \quad P_{yy} = -4, \quad P_{xy} = -1, \quad D = (-2)(-4) - (-1)^2 = 7 > 0.$$

Since $D > 0$ and $P_{xx} < 0$, the point is a local maximum. Because the quadratic form is concave, it is the global maximum.

Interpretation. At $(40, 20)$, marginal profit from one more unit of either product is zero. Producing more of one product reduces the marginal profitability of the other through the cross-term $-xy$.

Illustration: Two-variable optimization: gradient and Hessian



Contour lines of the profit function. The arrows indicate uphill directions toward the maximum.

Checkpoint: Hessian check

For

$$f(x, y) = 48xy - 32x^3 - 24y^2,$$

find the interior critical point and use the Hessian test to classify it. Then explain why a boundary check would still be needed if the feasible domain were a closed rectangle.

Constraints and Lagrange multipliers

Core explanation, worked reasoning, and application

Constraints and Lagrange multipliers

Many continuous optimization problems have a constraint. The geometry is simple: at a constrained optimum, the best achievable level curve of the objective is tangent to the constraint curve. This means the gradients are parallel.

Concept: Lagrange multiplier method

To optimize $f(x, y)$ subject to $g(x, y) = c$, form

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda \left[g(x, y) - c\right].$$

Solve

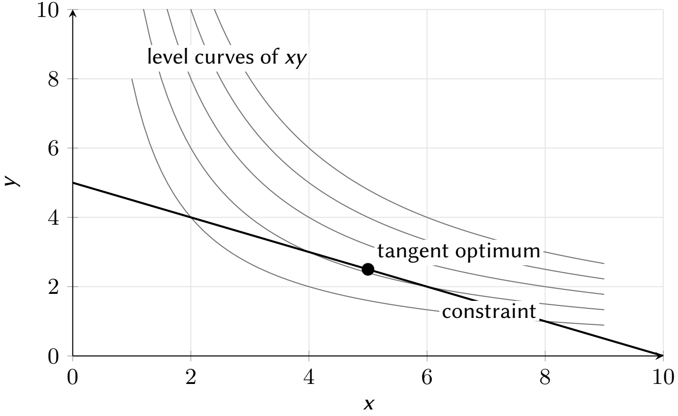
$$\mathcal{L}_x = 0, \qquad \mathcal{L}_y = 0, \qquad \mathcal{L}_\lambda = 0.$$

Equivalently,

$$\nabla f = -\lambda \nabla g.$$

The multiplier λ measures marginal value: how much the optimal objective changes when the constraint is relaxed by one unit, with sign depending on convention and whether we maximize or minimize.

Illustration: Constraints and Lagrange multipliers



At a constrained optimum, the objective level curve just touches the feasible constraint curve.

Worked Example: Worked Example 4: Maximum area with fixed fencing

A farmer has 100 meters of fencing for three sides of a rectangular garden against a wall. Let x be the side perpendicular to the wall and y the side parallel to the wall. The constraint is

$$2x + y = 100.$$

The area is

$$A = xy.$$

Use Lagrange multipliers:

$$\mathcal{L} = xy + \lambda(2x + y - 100).$$

Then

$$\mathcal{L}_x = y + 2\lambda = 0, \qquad \mathcal{L}_y = x + \lambda = 0, \qquad 2x + y = 100.$$

From $\lambda = -x$ and $y + 2\lambda = 0$, we get

$$y = 2x.$$

Substitute into the constraint:

$$2x + 2x = 100 \Rightarrow x = 25, \qquad y = 50.$$

The maximum area is

$$A^* = 25 \cdot 50 = 1250 \text{ m}^2.$$

Worked Example: Worked Example 4: Maximum area with fixed fencing

Shadow-price interpretation: if the total fencing increases slightly, the maximum area increases at approximately the marginal value of the constraint. From the single-variable form $A(x) = x(100 - 2x)$, the optimal value is $A^*(F) = F^2/8$, so $dA^*/dF = F/4 = 25$ at $F = 100$. One extra meter of fencing adds about $25\ \mathrm{m}^2$ of area near the optimum.

Worked Example: Worked Example 5: Budget-constrained production

A small workshop has a production function

$$Q(L, K) = L^{0.6}K^{0.4},$$

where L is labor units and K is capital units. Labor costs \$50 per unit and capital costs \$80 per unit. The budget is \$4000:

$$50L + 80K = 4000.$$

Maximize Q subject to this budget.
Use

$$\mathcal{L} = L^{0.6}K^{0.4} + \lambda(50L + 80K - 4000).$$

The first-order conditions imply

$$\frac{Q_L}{Q_K} = \frac{50}{80}.$$

Since

$$Q_L = 0.6L^{-0.4}K^{0.4}, \qquad Q_K = 0.4L^{0.6}K^{-0.6},$$

we get

$$\frac{Q_L}{Q_K} = \frac{0.6}{0.4}\frac{K}{L} = 1.5\frac{K}{L} = \frac{50}{80} = 0.625.$$

Worked Example: Worked Example 5: Budget-constrained production

Thus

$$K = \frac{0.625}{1.5} L \approx 0.4167L.$$

Use the budget:

$$50L + 80(0.4167L) = 4000 \Rightarrow 83.33L = 4000 \Rightarrow L^* = 48,$$

so

$$K^* = 20.$$

Interpretation: the optimal mix does not spend equally on labor and capital. It balances marginal product per dollar.

Model Critique: Important limitation: equality constraints are not the whole story

Lagrange multipliers handle smooth equality constraints. Many real models also have inequalities, such as $x \geq 0$, budget caps, capacity limits, and minimum service requirements. If an inequality is binding, it behaves like an equality; if it is slack, its multiplier is zero. The full theory is called the **Karush–Kuhn–Tucker conditions**, which is useful but too technical for the core lesson.

Numerical optimization: when equations do not solve neatly

Core explanation, worked reasoning, and application

Numerical optimization: when equations do not solve neatly

Calculus gives equations, but not always a closed-form solution. In competition work, students often minimize a nonlinear error function or maximize a model score using software. The key is not only to run a solver, but to know how to check and report it.

Concept: Gradient search as a modeling tool

For maximizing $f(\mathbf{x})$, a basic gradient ascent step is

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \nabla f(\mathbf{x}_k),$$

where α_k is a step size. For minimizing an error function, use gradient descent:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k).$$

In practice, α_k may be chosen by a **line search**, such as golden-section search, or by a solver routine.

Worked Example: Worked Example 6: Gradient ascent for a profit model

Use the profit model from Worked Example 3:

$$P(x, y) = 100x + 120y - x^2 - 2y^2 - xy - 500.$$

Starting from $(0, 0)$,

$$\nabla P = (100 - 2x - y,\ 120 - 4y - x).$$

With step size $\alpha = 0.10$, the first iteration is

$$(x_1, y_1) = (0, 0) + 0.10(100, 120) = (10, 12).$$

At $(10, 12)$,

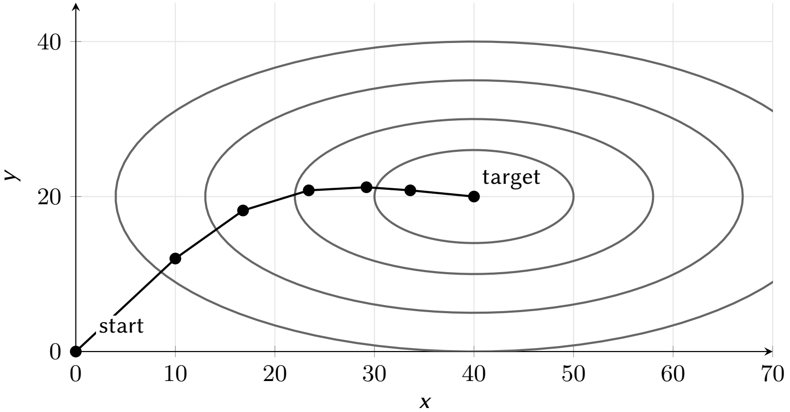
$$\nabla P = (100 - 20 - 12,\ 120 - 48 - 10) = (68, 62),$$

so

$$(x_2, y_2) = (10, 12) + 0.10(68, 62) = (16.8, 18.2).$$

The path moves toward $(40, 20)$, but a fixed step size may overshoot or converge slowly. A good numerical study should compare step sizes or use a reliable line search.

Illustration: Numerical optimization: when equations do not solve neatly



A numerical optimizer follows improvement directions, but convergence must be checked.

Model Critique: How to report numerical optimization

A competition report should not simply say "we used a solver." A stronger description is:

We optimized the nonlinear objective using gradient-based search. To reduce dependence on initial values, we used five starting points and obtained the same optimum to three decimal places. We also checked the result by plotting objective contours and by verifying that the gradient norm at the solution was below 10^{-5}*. A step-size refinement check changed the objective by less than* 0.2%*.*

This tells the reader that the numerical result is not a black box.

Applied extension: renewable-resource optimization

Core explanation, worked reasoning, and application

Applied extension: renewable-resource optimization

This section is suitable for the final part of the lesson or for homework. It connects optimization to sustainability.

Worked Example: Worked Example 7: Biological versus economic fishery optimum

Let N be fish population and let annual natural growth be

$$g(N) = aN(N_u - N), \qquad 0 < N < N_u.$$

The **biological optimum** maximizes sustainable yield $g(N)$. Since

$$g'(N) = a(N_u - 2N),$$

we get

$$N_b = \frac{N_u}{2}.$$

Suppose price per fish is p and harvest cost per fish $c(N)$ decreases as fish become easier to catch in a larger population. Profit from sustainable harvest is

$$\mathrm{TP}(N) = g(N)[p - c(N)].$$

At N_b,

$$\mathrm{TP}'(N_b) = g'(N_b)[p - c(N_b)] - g(N_b)c'(N_b).$$

Worked Example: Worked Example 7: Biological versus economic fishery optimum

Because $g'(N_b) = 0$, $g(N_b) > 0$, and $c'(N_b) < 0$,

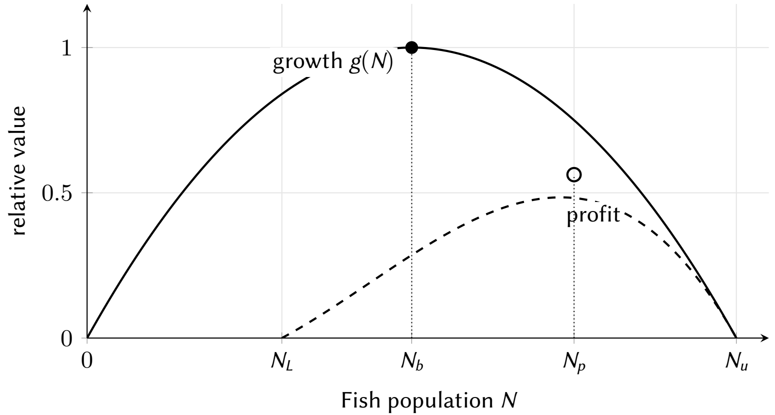
$$\mathrm{TP}'(N_b) = -g(N_b)c'(N_b) > 0.$$

Therefore the profit function is still increasing at N_b, so the **economic optimum** N_p satisfies

$$N_p > N_b.$$

Interpretation: maximizing profit may mean keeping a larger stock and harvesting less than the maximum sustainable yield, because catching each fish is cheaper when the stock is abundant.

Illustration: Applied extension: renewable-resource optimization



The biological optimum maximizes yield; the economic optimum may occur at a larger population.

Model Critique: Policy interpretation

A mathematical optimum is not automatically a policy recommendation. A fishery model should consider uncertainty in the stock estimate, illegal catch, ecosystem effects, enforcement cost, and risk of collapse. A cautious policy may prefer a population above the economic optimum when measurement error is large.

Classroom studio and communication

Core explanation, worked reasoning, and application

Activity: Differentiated optimization studio

Choose one route.

1. **Core route: EOQ.** For $d = 80$, $r = 120$, $s = 0.40$, compute T^* and Q^*. Check the equal-cost principle.
2. **Challenge route: constrained area.** Maximize xy subject to $x + 2y = 12$. Solve by Lagrange multipliers and interpret the geometry.
3. **Extension route: numerical optimization.** Starting from $(0, 0)$, perform three gradient-ascent steps with $\alpha = 0.05$ for $P(x, y) = 100x + 120y - x^2 - 2y^2 - xy - 500$. Compare with the exact optimum $(40, 20)$.

Model Critique: Competition-style conclusion

A continuous-optimization conclusion should include:

1. **Decision:** state the recommended value of the decision variable(s).
2. **Reason:** identify the objective and why the optimum occurs there.
3. **Sensitivity:** describe how the recommendation changes when key parameters change.
4. **Feasibility:** check integer, capacity, safety, or practical constraints.
5. **Limitations:** state assumptions that may break in real use.

Weak conclusion: "The optimal order quantity is 310 books."
Stronger conclusion: "Under constant monthly demand of 200 books, an ordering cost of \$120 and holding cost of \$0.50 per book-month, the EOQ model recommends ordering about 310 books every 1.55 months. This balances ordering and holding costs at about \$77.5 each per month. Because Q^* depends on a square root, moderate cost estimation errors have limited effect, but the recommendation should be rounded to carton size and reconsidered if demand is seasonal."

Summary: Lesson summary

Concept	Modeling meaning
Continuous optimization	Choose the best value from a continuum, not only among listed alternatives.
EOQ	Balances fixed order cost and holding cost; optimum follows a square-root law.
Second derivative	Confirms whether a critical point is a minimum or maximum.
Gradient	Points in the direction of fastest increase.
Hessian	Uses curvature to classify multivariable critical points.
Lagrange multiplier	Finds constrained optima where level curves are tangent to constraints.
Shadow price	Marginal value of relaxing a constraint.
Numerical optimization	Needed when equations cannot be solved neatly; requires convergence and robustness checks.
Model criticism	The mathematical optimum must be checked against assumptions, data quality, and practical constraints.

Checkpoint: Exit ticket

In three sentences, explain the difference between an unconstrained optimum and a constrained optimum. Then give one example where the calculus optimum is not directly usable because of a practical constraint.

Exercises: Core practice

1. For $c(T) = 90/T + 30T$, find T^* and verify that $c''(T^*) > 0$.
2. A shop has $d = \$60$, $r = 150$ units/week, and $s = \$0.20$ per unit-week. Find the EOQ cycle length and order quantity.
3. Find and classify the critical point of $f(x, y) = 60x + 80y - x^2 - y^2 - xy$.
4. Maximize $A = xy$ subject to $2x + y = 40$ using Lagrange multipliers.
5. Explain why the point where $f_x = f_y = 0$ might fail to be a global maximum.

Exercises: Modeling challenge

A school canteen sells a snack. If it orders too little, it loses potential sales; if it orders too much, leftovers have no resale value. A simplified continuous model estimates daily profit as

$$\Pi(q) = 8q - 0.03q^2 - 120 - 0.5\max(q - 180, 0),$$

where q is the number of snacks prepared. The term $8q - 0.03q^2$ models demand saturation; the fixed cost is 120; the final term penalizes leftovers beyond 180.

1. Explain why this is a continuous optimization approximation, even though snacks are discrete.
2. Optimize $\Pi(q)$ separately on $q \leq 180$ and $q \geq 180$.
3. Determine the best feasible q and interpret.
4. State two practical checks before giving the recommendation to the canteen manager.

Exercises: Extension practice

1. For $f(x, y) = x^2 - y^2$, show that $(0, 0)$ is a saddle point using both the Hessian test and a contour argument.
2. A container must have volume $V = xyz = 1000$ and rectangular dimensions $x, y, z > 0$. Set up the Lagrange multiplier problem to minimize surface area $2xy + 2xz + 2yz$. What shape do you expect?
3. For the EOQ model, show that a 10% error in d changes T^* by only about 5%. Use either calculus or the square-root relationship.
4. Describe a numerical optimization validation plan for a nonlinear model with many local optima.

Summary: Optional appendix: KKT idea for inequalities

For a maximization problem with inequality constraint $g(x) \leq c$, the multiplier is active only when the constraint binds. Informally:

$$\lambda \geq 0, \qquad g(x) \leq c, \qquad \lambda[g(x) - c] = 0.$$

The last condition says: either the constraint is binding and may have positive value, or it is slack and its shadow price is zero. This is the bridge from Lagrange multipliers to modern nonlinear programming.

Lecture 11: Dimensional Analysis and Similitude
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, students should be able to:

1. explain why **dimensional analysis** reduces experimental complexity;
2. check whether an equation is **dimensionally compatible**;
3. represent physical quantities using the MLT system: **mass**, **length**, and **time**;
4. construct a **dimensionless product** by solving exponent equations;
5. apply **Buckingham's Pi Theorem** to reduce a variable list;
6. use dimensional analysis to derive scaling laws such as wind force, pendulum period, and roasting time;
7. explain **similitude** and why model experiments must match key dimensionless numbers;
8. write a modeling paragraph that states variables, assumptions, Π-groups, validation, and limitations.

Why dimensional analysis?

Core explanation, worked reasoning, and application

Why dimensional analysis?

A physical model often begins with a long list of possible variables. If an experimenter studies n variables and tests k values of each, a full grid requires k^n trials. This is the **curse of dimensionality** in experimental design. Dimensional analysis does not find the whole law by magic. Instead, it says: physical laws cannot depend on whether we measure in metres or feet, kilograms or pounds. Therefore the important relationships can be expressed using *dimensionless* combinations of variables. This can reduce n original variables to far fewer dimensionless groups.

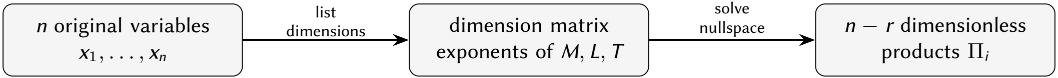
Activity: Launch: the experimental burden

12 min

A team wants to study the drag force on a small underwater robot. They suspect that drag depends on velocity v, length L, water density ρ, viscosity μ, and frontal area A.

1. If each variable is tested at 5 values, how many experiments are required in a full grid?
2. Why might this be unrealistic for a school or competition project?
3. What kind of mathematical trick would help reduce the number of experiments?

Illustration: Why dimensional analysis?



Experiments vary Π-groups, not every original variable independently.

Dimensional analysis converts a long variable list into a smaller set of dimensionless groups.

Dimensions and compatibility

Core explanation, worked reasoning, and application

Dimensions and compatibility

In mechanics, most quantities can be expressed using three primary dimensions:

$$M = \text{mass}, \qquad L = \text{length}, \qquad T = \text{time}.$$

A velocity has dimension LT^{-1}; a force has dimension MLT^{-2}; an energy has dimension ML^2T^{-2}.

Concept: Dimensional compatibility

An equation is **dimensionally compatible** if every term being added or subtracted has the same dimension. Products and powers may have different dimensions, but sums must match.
For example, in

$$s = s_0 + v_0 t - \frac{1}{2} g t^2,$$

each term has dimension L. Therefore the equation passes the compatibility check.

Dimensions and compatibility

Quantity	Dimension	Typical units
Velocity v	LT^{-1}	m/s
Acceleration a	LT^{-2}	m/s^2
Force F	MLT^{-2}	N
Density ρ	ML^{-3}	kg/m^3
Pressure p	$ML^{-1}T^{-2}$	Pa
Dynamic viscosity μ	$ML^{-1}T^{-1}$	Pa s
Energy E	ML^2T^{-2}	J
Power P	ML^2T^{-3}	W

Worked Example: Worked Example 1: catching a wrong model before collecting data

A student proposes the model

$$F = mv + v^2$$

for the force needed to move an object.

$$[F] = MLT^{-2}, \qquad [mv] = M \cdot LT^{-1} = MLT^{-1},$$

while

$$\left[v^2\right] = L^2 T^{-2}.$$

The two terms on the right cannot be added, and neither has the dimension of force. The model is structurally wrong before any numerical fitting begins.

Modeling lesson: dimensional checks are a cheap error detector. They do not prove a model is correct, but they quickly reject many impossible models.

Checkpoint: Quick compatibility check

For each equation, decide whether it is dimensionally compatible.

1. $E = \frac{1}{2}mv^2$.
2. $p = \rho g h$.
3. $F = \rho A + mv$.
4. $t = 2\pi\sqrt{L/g}$.

Constructing dimensionless products

Core explanation, worked reasoning, and application

Constructing dimensionless products

A **dimensionless product** is a product of powers of variables whose total dimension is $M^0 L^0 T^0$.
Suppose variables $x_1, \ldots, x_n$ have dimensions. We search for exponents $a_1, \ldots, a_n$ such that

$$x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$$

is dimensionless. Equating powers of M, L, T gives a homogeneous linear system.

Worked Example: Worked Example 2: the simple pendulum

Let the period t of a pendulum depend on length ℓ, mass m, gravity g, and release angle θ. The variables are:

$$t(T), \qquad \ell(L), \qquad m(M), \qquad g(LT^{-2}), \qquad \theta(1).$$

Consider

$$t^a \ell^b m^c g^d \theta^e.$$

Its dimension is

$$T^a L^b M^c (LT^{-2})^d = M^c L^{b+d} T^{a-2d}.$$

For this to be dimensionless,

$$c = 0, \qquad b + d = 0, \qquad a - 2d = 0.$$

There are two free choices. Taking $d = 1$ gives $a = 2$, $b = -1$, $c = 0$, so

$$\Pi_1 = \frac{gt^2}{\ell}.$$

Taking $e = 1$ gives

$$\Pi_2 = \theta.$$

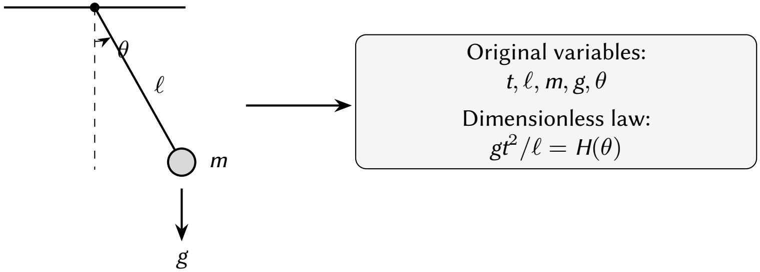
Worked Example: Worked Example 2: the simple pendulum

Therefore the period law must have the form

$$\frac{gt^2}{\ell} = H(\theta), \qquad \text{or} \qquad \boxed{t = \sqrt{\frac{\ell}{g}}\, h(\theta).}$$

Interpretation: mass disappears. Dimensional analysis predicts that pendulum period does not depend on the bob's mass.

Illustration: Constructing dimensionless products



For the pendulum, dimensional analysis removes mass and identifies the correct scale $\sqrt{\ell/g}$.

Worked Example: Worked Example 3: wind force and a missing variable

A first attempt might say wind force on a sign depends only on wind speed v and area A:

$$F = kv^a A^b.$$

But

$$\left[v^a A^b\right] = (LT^{-1})^a (L^2)^b = L^{a+2b} T^{-a},$$

which has no mass dimension. Since $[F] = MLT^{-2}$, a mass-bearing variable is missing.
Include air density ρ with dimension ML^{-3}:

$$F = k\rho^c v^a A^b.$$

Equating dimensions gives

$$M: 1 = c, \qquad T: -2 = -a \Rightarrow a = 2,$$

$$L: 1 = -3c + a + 2b = -3 + 2 + 2b \Rightarrow b = 1.$$

Thus

$$\boxed{F = k\rho v^2 A.}$$

Interpretation: doubling wind speed quadruples the force; doubling exposed area doubles the force. The dimensionless constant k must be found experimentally.

Buckingham’s Pi theorem

Core explanation, worked reasoning, and application

Concept: Buckingham's Pi theorem

If a model uses n dimensional variables and the dimension matrix has rank r, then the physical relationship can be expressed using

$$n - r$$

independent dimensionless products:

$$F(\Pi_1, \Pi_2, \ldots, \Pi_{n-r}) = 0.$$

For most mechanics problems using M, L, T, the rank is often $r = 3$, so the number of groups is usually $n - 3$.

Summary: A teachable seven-step workflow

1. State the dependent variable and plausible independent variables.
2. Write the MLT dimension of every variable.
3. Build products of powers and solve for dimensionless combinations.
4. Check that every Π-group is truly dimensionless.
5. Use Buckingham's theorem to write the reduced relationship.
6. Interpret the scaling law and identify what still needs data.
7. Validate using plots of dimensionless variables, not only original variables.

Worked Example: Worked Example 4: projectile range

The horizontal range R of a projectile depends on initial speed v, gravity g, and launch angle θ.

$$R(L), \qquad v(LT^{-1}), \qquad g(LT^{-2}), \qquad \theta(1).$$

There are four variables and rank two here because only L and T appear. Therefore there are $4-2=2$ dimensionless products.
Let

$$R^a v^b g^c \theta^d.$$

Its dimension is

$$L^a(LT^{-1})^b(LT^{-2})^c = L^{a+b+c}T^{-b-2c}.$$

For dimensionless products:

$$a+b+c=0, \qquad -b-2c=0.$$

Choose $a=1$. Then $b=-2$, $c=1$, giving

$$\Pi_1 = \frac{Rg}{v^2}.$$

Worked Example: Worked Example 4: projectile range

The angle is already dimensionless, so $\Pi_2 = \theta$. Hence

$$\frac{Rg}{v^2} = H(\theta), \qquad \text{or} \qquad \boxed{R = \frac{v^2}{g} H(\theta).}$$

The exact mechanics result $R = v^2 \sin(2\theta)/g$ is consistent with the dimensional result. Dimensional analysis finds the scale v^2/g but not the particular function $\sin(2\theta)$.

Model Critique: What dimensional analysis can and cannot do

Dimensional analysis can reveal missing variables, impossible equations, scaling exponents, and useful dimensionless combinations. It cannot determine dimensionless constants or arbitrary functions such as $H(\theta)$. Those require theory, data, or simulation.

Scaling laws and model refinement

Core explanation, worked reasoning, and application

Worked Example: Worked Example 5: turkey roasting and the $w^{2/3}$ law

Suppose roasting time t depends on a characteristic length ℓ, thermal diffusivity κ with dimension L^2T^{-1}, and a dimensionless temperature ratio. The only way to combine t, ℓ, and κ dimensionlessly is

$$\Pi = \frac{\kappa t}{\ell^2}.$$

Therefore, for fixed oven conditions,

$$t \propto \frac{\ell^2}{\kappa}.$$

If turkeys are geometrically similar, weight w is proportional to volume, so

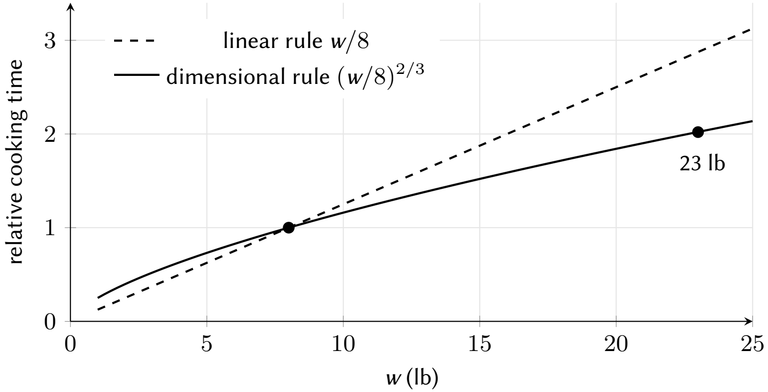
$$w \propto \ell^3, \qquad \ell \propto w^{1/3}.$$

Thus

$$\boxed{t \propto w^{2/3}.}$$

This differs from the folk rule "minutes per pound", which implies $t \propto w$. The folk rule tends to overcook large turkeys.

Illustration: Scaling laws and model refinement



The $w^{2/3}$ law grows more slowly than a linear minutes-per-pound rule.

Worked Example: Worked Example 6: raindrop terminal velocity

Let terminal velocity v depend on radius r, gravity g, air density ρ, and dynamic viscosity μ:

$$v(LT^{-1}), \quad r(L), \quad g(LT^{-2}), \quad \rho(ML^{-3}), \quad \mu(ML^{-1}T^{-1}).$$

There are 5 variables and rank 3, so we expect 2 dimensionless products. One convenient set is

$$\Pi_1 = \frac{v}{\sqrt{rg}}, \qquad \Pi_2 = \frac{r^{3/2}\sqrt{g}\,\rho}{\mu}.$$

Therefore

$$\boxed{v = \sqrt{rg}\,H\left(\frac{r^{3/2}\sqrt{g}\,\rho}{\mu}\right).}$$

Interpretation: terminal velocity is not just proportional to $\sqrt{rg}$; viscosity and density determine the drag regime through the second dimensionless group. This is why small mist droplets, raindrops, and hailstones do not follow exactly the same simple law.

Checkpoint: Scaling interpretation

In the wind-force model $F = k\rho v^2 A$, what happens to force if:

1. wind speed triples while ρ and A are unchanged?
2. a sign is replaced by a geometrically similar sign twice as tall and twice as wide?
3. the same car drives at the same speed at high altitude where air density is lower?

Similitude: from model to prototype

Core explanation, worked reasoning, and application

Similitude: from model to prototype

Similitude means that a laboratory model and the real prototype have the same relevant dimensionless products. Geometric similarity alone is not enough. A small model submarine is not dynamically similar to a real submarine unless the key force ratios also match.

Concept: Classical dimensionless numbers

Name	Formula	Physical meaning
Reynolds number Re	$\frac{vL\rho}{\mu}$	inertia / viscosity
Froude number Fr	$\frac{v^2}{gL}$	inertia / gravity
Mach number Ma	$\frac{v}{c}$	speed / sound speed
Weber number We	$\frac{\rho v^2 L}{\sigma}$	inertia / surface tension
Drag coefficient C_D	$\frac{F_D}{\frac{1}{2}\rho v^2 A}$	dimensionless drag

Worked Example: Worked Example 7: Froude similarity for a spillway model

A 1:20 scale model of a dam spillway is tested in a lab. The prototype flow speed is $v_p = 8$ m/s. Gravity effects dominate, so match the Froude number

$$Fr = \frac{v^2}{gL}.$$

Let $L_m = L_p/20$. Froude similarity requires

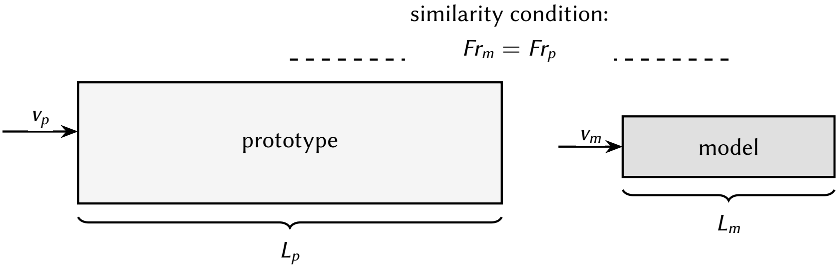
$$\frac{v_m^2}{gL_m} = \frac{v_p^2}{gL_p}.$$

Thus

$$v_m = v_p\sqrt{\frac{L_m}{L_p}} = 8\sqrt{\frac{1}{20}} \approx 1.79 \text{ m/s}.$$

Interpretation: model velocity is not $8/20 = 0.4$ m/s. For gravity-controlled free-surface flow, velocity scales with the square root of length.

Illustration: Similitude: from model to prototype



A scaled model predicts a prototype only when the relevant dimensionless numbers match.

Model Critique: Partial similitude

In real experiments, it may be impossible to match all dimensionless numbers at once. A ship model may need Froude similarity for waves and Reynolds similarity for viscosity, but the required model speeds may conflict. A strong report should state which Π-groups were matched, which were not, and how this affects confidence in the prediction.

Differentiated modeling studio

Core explanation, worked reasoning, and application

Activity: Dimensional-analysis studio

18 min

Choose one route.

1. **Core route: projectile range.** Derive $R = (v^2/g)H(\theta)$ using dimensions, then explain what dimensional analysis leaves unknown.
2. **Challenge route: Stokes drag.** Assume drag force F on a slow sphere depends on radius r, speed v, and viscosity μ. Show that $F = k\mu rv$.
3. **Extension route: blast-wave scaling.** Suppose blast radius R depends on energy E, air density ρ, and time t. Derive the scaling law $R \propto (Et^2/\rho)^{1/5}$.

Model Critique: Competition-style communication

A good dimensional-analysis paragraph should include:

1. the dependent variable and the assumed independent variables;
2. the dimension table and the number of resulting Π-groups;
3. the reduced dimensionless relationship;
4. the physical interpretation of the scaling;
5. a validation plan and a limitation.

Weak: "The answer is $t \propto w^{2/3}$."
Stronger: "Assuming geometrically similar turkeys and fixed oven conditions, the cooking time must scale as $t \propto \ell^2/\kappa$. Since weight is proportional to volume, $\ell \propto w^{1/3}$ and hence $t \propto w^{2/3}$. This predicts that large turkeys need less time per pound than small turkeys. The main limitation is that real turkeys are not perfectly similar and ovens do not heat uniformly."

Summary: Lesson summary

Concept	Modeling meaning
MLT system	Expresses physical quantities using mass, length, and time.
Compatibility	Additive terms must have the same dimension; a fast model-error check.
Dimensionless product	A product of powers whose total dimension is $M^0 L^0 T^0$.
Buckingham Π theorem	Reduces n dimensional variables to $n - r$ dimensionless groups.
Scaling law	Reveals how changing size, speed, density, or time changes outcomes.
Similitude	Model and prototype must match relevant Π-groups.
Partial similitude	Match the dominant physics; disclose unmatched effects.
Validation	Fit or compare dimensionless variables, not only raw variables.

Checkpoint: Exit ticket

In three sentences, explain why dimensional analysis is useful before collecting data. Include one example of a dimensionless product.

Practice problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. Find the dimension of pressure, power, density, viscosity, and surface tension in the MLT system.
2. Check whether $p = \rho g h$, $F = ma$, and $E = mv + gh$ are dimensionally compatible.
3. For a pendulum, verify that gt^2/ℓ is dimensionless.
4. Use dimensional analysis to show that the period of a mass-spring oscillator must have the form $t = C\sqrt{m/k}$, where k is spring stiffness with dimension MT^{-2}.
5. If $F = k\rho v^2 A$, by what factor does F change when v is doubled and A is tripled?
6. For Froude similarity, a 1:25 scale model represents a river flow with prototype speed 5 m/s. Find the model speed.

Exercises: Modeling challenge

A team wants to design a small-scale experiment for the drag force F_D on a toy car moving through air. They believe F_D depends on speed v, frontal area A, air density ρ, viscosity μ, and characteristic length L.

1. List the dimensions of all variables.
2. Predict how many dimensionless groups should appear.
3. One likely group is the Reynolds number $Re = vL\rho/\mu$. Verify it is dimensionless.
4. Propose a second dimensionless group involving F_D.
5. Explain how the team should use these groups to compare a small model car and a real car.
6. State one limitation of the experiment.

Exercises: Extension practice

1. **Projectile range.** Derive $R = (v^2/g)H(\theta)$ and verify the exact formula $R = v^2 \sin(2\theta)/g$ is consistent.
2. **Terminal velocity.** Starting with v, r, g, ρ, μ, derive two independent dimensionless products.
3. **Blast wave.** Assume blast radius R depends on energy E, air density ρ, and time t. Show that $R = C(Et^2/\rho)^{1/5}$.
4. **Choosing variables.** A student models cooking time using only turkey weight w. Explain why this variable list is incomplete, then propose a better list.
5. **Partial similitude.** Explain why matching both Reynolds and Froude numbers may be impossible in a small water-channel experiment.

Summary: Optional appendix: choosing Π-groups

Different complete sets of Π-groups are possible. For example, $Re = vL\rho/\mu$ and $1/Re$ contain the same information. A good set should be:

- dimensionless;
- independent;
- physically interpretable;
- convenient for plotting and experimentation;
- chosen so the dependent variable appears in only one group when possible.

Lecture 12: Graphical and Qualitative Modeling
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of the lesson, readers should be able to:

1. use the shape of a graph as a mathematical model when exact equations are unknown;
2. identify and interpret **equilibria** as intersections of response curves;
3. reason from **slope**, **concavity**, and curve shifts to predict policy effects;
4. build a simple **discrete dynamical system** and test stability;
5. explain the economic condition $MR = MC$ and connect it to supply curves;
6. analyze a per-unit tax using supply and demand graphs;
7. write a qualitative modeling conclusion with assumptions, diagrams, and limitations.

Summary: Two-hour lesson route

0–12 min	Launch: when a graph is a model.
12–28 min	Graphical vocabulary: increasing, decreasing, concavity, intersection and shift.
28–52 min	Case study I: arms-race response curves and equilibrium.
52–68 min	Discrete arms-race update and stability check.
68–88 min	Case study II: firm behavior and $MR = MC$.
88–108 min	Gasoline tax, supply-demand shifts, and tax incidence.
108–120 min	Differentiated studio, model criticism, exit ticket, and homework briefing.

Graphs as models

Core explanation, worked reasoning, and application

Graphs as models

A model does not always have to begin with a fully specified formula. In many policy and competition problems, we know qualitative features before we know numbers. A demand curve is decreasing. A cost curve may first flatten and then steepen. A deterrence requirement may increase but with diminishing additional effect. These features are mathematical information.

A **qualitative graphical model** uses the shape of one or more functions to infer behavior. The central questions are:

- Where do curves intersect?
- Is an intersection stable or unstable?
- How does a parameter shift a curve?
- Which conclusion is robust even if the exact formula changes?

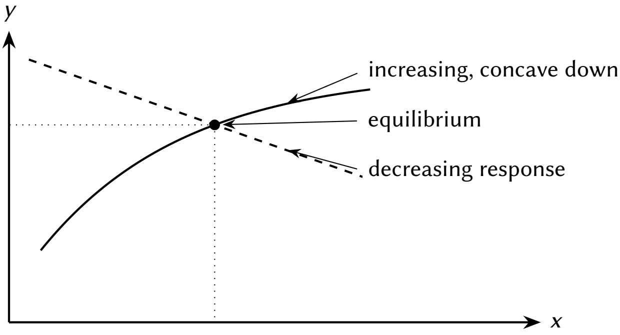
Activity: Launch: what can a graph tell us?

Consider a school canteen. Let q be the number of lunch sets prepared and let $C(q)$ be total cost.

1. Sketch a plausible cost curve. Should it be increasing? Should it be linear?
2. Sketch a plausible demand curve between price and quantity demanded.
3. State one conclusion you can make from the shapes, even without exact equations.

Modeling lesson: a graph records assumptions. A line says constant marginal effect; a curved graph says the marginal effect changes.

Illustration: Graphs as models



A graphical model turns qualitative assumptions about direction and curvature into predictions about intersections and comparative statics.

Concept: Core vocabulary

A **response curve** shows the value one part of a system requires in response to another part. An **equilibrium** is an intersection where all response requirements are simultaneously satisfied. A **comparative-static analysis** asks how the equilibrium changes after a parameter shifts a curve.

Case study I: an arms-race graphical model

Core explanation, worked reasoning, and application

Case study I: an arms-race graphical model

The arms-race model is not included because of its political content, but because it is an excellent example of qualitative modeling. The modeler does not need the exact military formula. The shape of the response curves already gives useful conclusions.

Let x be the number of weapons held by Country X and y be the number held by Country Y. Suppose

$$y = f(x)$$

means: the minimum number of Y weapons needed to maintain deterrence when X has x weapons. Similarly,

$$x = g(y)$$

means: the minimum number of X weapons needed when Y has y weapons.

Concept: Deterrence curve assumptions

For the core lesson we use three qualitative assumptions.

1. f and g are increasing: more weapons by one side require a stronger response by the other.
2. They are concave down: the first enemy weapons matter a lot; later additions may have smaller marginal impact.
3. Their intercepts represent minimum retaliation requirements even if the opponent has very few weapons.

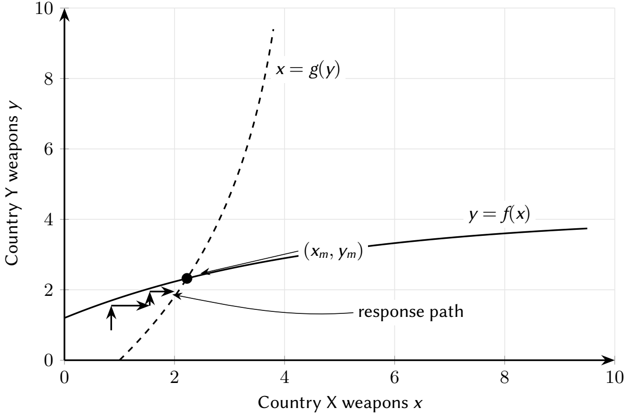
These assumptions are weaker than a full formula, but strong enough to support policy reasoning.

Worked Example: Worked Example 1: Reading the deterrence equilibrium

Suppose the two response curves intersect at (x_m, y_m). At this point, Country X has enough weapons relative to Y, and Country Y has enough weapons relative to X. The point is therefore a mutual deterrence equilibrium.

If the current state lies below $y = f(x)$, then Y judges its arsenal insufficient and builds more weapons. If the current state lies left of $x = g(y)$, then X judges its arsenal insufficient and builds more weapons. Repeated responses tend to move the system toward the intersection if each reaction is moderate.

Illustration: Case study I: an arms-race graphical model



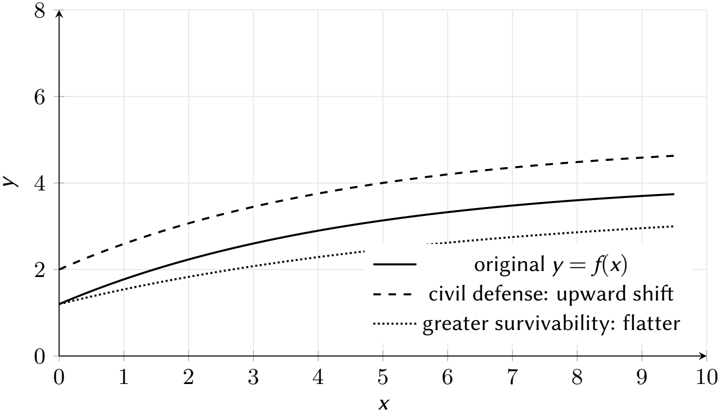
Two response curves determine a mutual deterrence equilibrium. The diagram is qualitative: the conclusions depend on shape and shifts, not exact numerical coordinates.

Worked Example: Worked Example 2: Policy shifts

The graphical model becomes useful when we ask how a policy changes a curve.

1. **Civil defense by X.** If X protects its population better, Y needs more surviving weapons to create the same deterrent effect. Y's response curve $y = f(x)$ shifts upward. The new equilibrium has a larger y, and usually a larger x as X responds to Y's buildup.
2. **Mobile missiles by X.** If X's weapons are harder to destroy, X needs fewer weapons for the same survivability. X's response curve $x = g(y)$ shifts left or becomes flatter. The equilibrium can move downward for both sides.
3. **Multiple warheads.** If one weapon can attack many targets, the exchange ratio changes. The intercept effect may reduce required launchers, but the offensive effect may steepen the response curves. A purely qualitative diagram may become ambiguous; the modeler must state that more data or a more explicit formula is needed.

Illustration: Case study I: an arms-race graphical model



Policy analysis often reduces to curve shifting. The critical question is which parameter changes and in which direction.

Checkpoint: Graphical policy check

A country builds better missile shelters, so a larger fraction of its weapons survive a first strike. Which response curve shifts: its own curve or the opponent's curve? Does the curve move upward, downward, leftward, or become flatter? Explain in two sentences.

A discrete arms-race update

Core explanation, worked reasoning, and application

A discrete arms-race update

A graph gives qualitative direction. A **discrete dynamical system** gives a sequence of numerical updates.

Worked Example: Worked Example 3: A staged arms-race model

Suppose Country Y believes it needs 120 weapons even when X has none, and then adds one additional weapon for every two weapons X has. Country X believes it needs 60 weapons even when Y has none, and then adds one additional weapon for every three weapons Y has:

$$y_{n+1} = 120 + \frac{1}{2}x_n, \qquad x_{n+1} = 60 + \frac{1}{3}y_n.$$

At equilibrium,

$$y^* = 120 + \frac{1}{2}x^*, \qquad x^* = 60 + \frac{1}{3}y^*.$$

Substitute the second equation into the first:

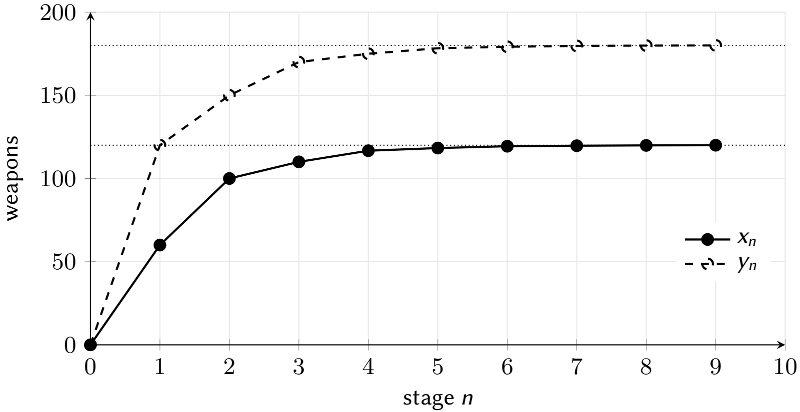
$$y^* = 120 + \frac{1}{2}\left(60 + \frac{1}{3}y^*\right) = 150 + \frac{1}{6}y^*.$$

Thus $\frac{5}{6}y^* = 150$, so

$$\boxed{y^* = 180}, \qquad \boxed{x^* = 120}.$$

The product of reaction coefficients is $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} < 1$, so repeated updates converge to the equilibrium.

Illustration: A discrete arms-race update



When the product of reaction coefficients is below one, the staged arms-race model converges geometrically to equilibrium.

Model Critique: Model criticism: what does the update hide?

The update equations assume that each side reacts only to the opponent's previous stockpile, with fixed coefficients and no negotiation, budget limit, domestic politics, or uncertainty. A good modeling report should present the recurrence as a simplified warning model, not as a complete theory of political behavior.

Case study II: firm behavior and $MR = MC$

Core explanation, worked reasoning, and application

Case study II: firm behavior and $MR = MC$

Graphical modeling also explains economic decisions. Let q be output quantity. A competitive firm accepts the market price p. Its **marginal revenue** is therefore $MR = p$. Its **marginal cost** is $MC(q) = TC'(q)$.

Concept: Profit-maximizing rule

Profit is

$$\Pi(q) = TR(q) - TC(q).$$

The first-order condition is

$$\Pi'(q) = MR(q) - MC(q) = 0.$$

For a competitive firm, $MR = p$, so the firm produces where

$$\boxed{MR = MC.}$$

The firm should produce more while $MR > MC$, and stop increasing output once the next unit costs more than it earns.

Worked Example: Worked Example 4: A firm's optimal output

A small producer sells at price $p = 10$ dollars per unit and has total cost

$$TC(q) = 100 + 2q + 0.1q^2.$$

Then

$$TR(q) = 10q, \qquad \Pi(q) = 10q - (100 + 2q + 0.1q^2) = 8q - 0.1q^2 - 100.$$

The marginal cost is

$$MC(q) = TC'(q) = 2 + 0.2q.$$

Set $MR = MC$:

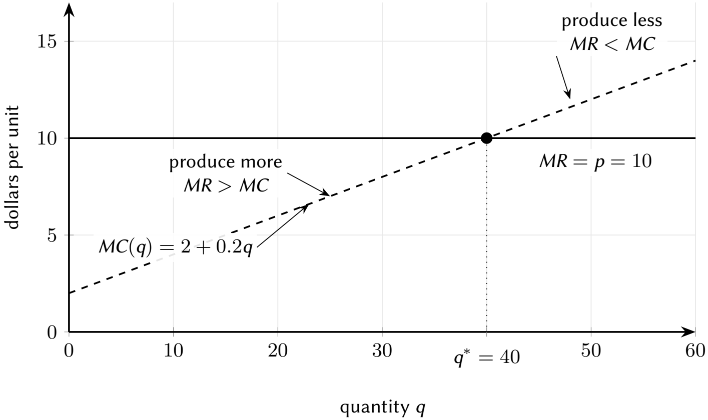
$$10 = 2 + 0.2q \quad \Rightarrow \quad q^* = 40.$$

The maximum profit is

$$\Pi(40) = 8(40) - 0.1(40)^2 - 100 = 60.$$

The graph shows the same conclusion: produce until the horizontal price line meets the marginal-cost curve.

Illustration: Case study II: firm behavior and $MR = MC$



For a competitive firm, the profit-maximizing output occurs where price equals marginal cost.

Checkpoint: Economic interpretation

If a per-unit tax of 2 dollars is imposed on the producer, the marginal cost becomes $MC(q) = 4 + 0.2q$. What is the new optimal output? Explain why the output falls.

Gasoline tax: supply, demand, and tax incidence

Core explanation, worked reasoning, and application

Gasoline tax: supply, demand, and tax incidence

Now move from one firm to a market. The **supply curve** gives the price at which producers are willing to supply quantity q. The **demand curve** gives the price at which consumers are willing to buy quantity q. The intersection is the **market equilibrium**.

A per-unit tax shifts the supply curve upward by the tax amount. The price paid by consumers rises, the quantity falls, and the tax burden is shared between consumers and producers.

Worked Example: Worked Example 5: Linear tax incidence

Suppose

$$S(q) = 1 + 0.5q, \qquad D(q) = 5 - 0.5q.$$

The original equilibrium solves

$$1 + 0.5q = 5 - 0.5q \quad \Rightarrow \quad q^* = 4, \qquad p^* = 3.$$

A per-unit tax $\tau = 1$ shifts supply to

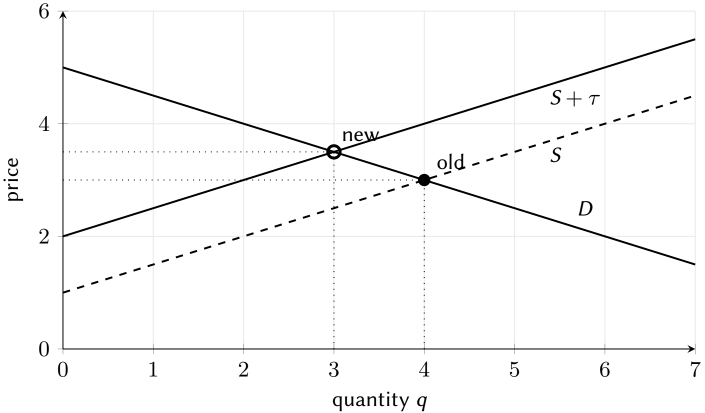
$$S_1(q) = S(q) + 1 = 2 + 0.5q.$$

The new equilibrium solves

$$2 + 0.5q = 5 - 0.5q \quad \Rightarrow \quad q_1 = 3, \qquad p_1 = 3.5.$$

Consumers pay $p_1 - p^* = 0.5$ more per unit. Producers receive $p_1 - \tau = 2.5$, which is 0.5 less than before. The one-dollar tax is split equally in this symmetric example.

Illustration: Gasoline tax: supply, demand, and tax incidence

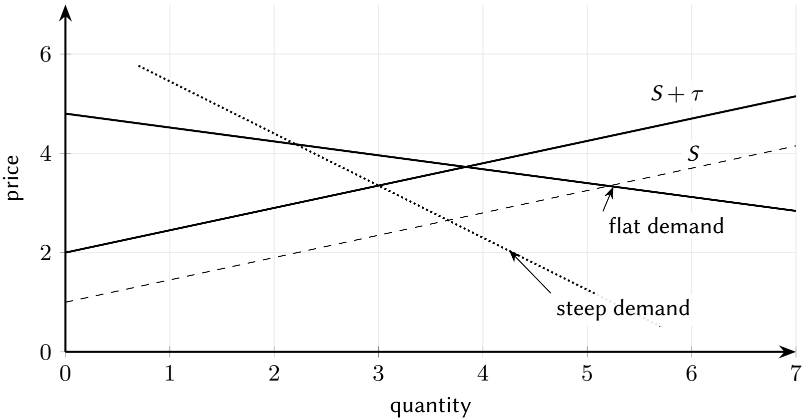


A per-unit tax shifts the supply curve upward. Price rises by less than the tax; quantity falls.

Concept: Tax incidence and elasticity

The side of the market with less flexibility bears more of the tax. If demand is steep, consumers cannot reduce quantity easily, so consumers pay most of the tax. If demand is flat, consumers can switch away easily, so producers absorb more of the tax.

Illustration: Gasoline tax: supply, demand, and tax incidence



The same tax has different effects depending on the slope of demand. This is a qualitative conclusion that does not require exact functional forms.

Model Critique: Short-run versus long-run gasoline demand

Immediately after an oil shock, demand is often inelastic: people still need to commute, cars cannot be replaced instantly, and public transport may be unavailable. A gasoline tax then raises prices but may reduce quantity only slightly. Over several years, demand becomes more elastic as people buy efficient cars, change routes, carpool, or use public transport. The same tax can then reduce consumption more effectively. The policy conclusion depends on the time horizon.

Differentiated studio: choosing a graph-based model

Core explanation, worked reasoning, and application

Activity: Differentiated modeling studio

Choose one route.

1. **Core route: tax graph.** Draw supply and demand. Shift supply upward by a tax. Label old and new equilibria. State who pays more when demand is steep.
2. **Challenge route: staged update.** For

$$y_{n+1} = 200 + \frac{1}{3}x_n, \qquad x_{n+1} = 80 + \frac{1}{4}y_n,$$

   find the equilibrium and test stability.
3. **Extension route: curve-shift argument.** Explain, using deterrence curves, why increasing survivability may reduce equilibrium weapons, while increasing offensive exchange ratio may increase them.

Summary: Choosing a graph-based modeling tool

Situation	Graphical tool	Question answered
Two sides respond to each other	Response curves	Where is mutual satisfaction?
Repeated strategic reaction	Difference equations	Does the process converge?
Firm chooses output	MR and MC curves	What quantity maximizes profit?
Market clears	Supply and demand	What price and quantity occur?
Policy intervention	Curve shift	How does equilibrium change?
Unknown exact formula	Shape and slope reasoning	Which conclusion is robust?

Writing qualitative-model conclusions

Core explanation, worked reasoning, and application

Model Critique: Competition-style communication

A qualitative graphical model should not sound vague. A strong paragraph should include:

1. **Variables:** what the axes represent.
2. **Shape assumptions:** increasing/decreasing, concavity, intercepts, or stability.
3. **Mechanism:** why those shapes are reasonable.
4. **Conclusion:** what happens to the equilibrium after a policy change.
5. **Limitation:** what cannot be decided without data.

Weak conclusion: "The tax makes price go up."
Stronger conclusion: "A per-unit gasoline tax shifts the supply curve upward. The new equilibrium has higher consumer price and lower quantity. Since short-run gasoline demand is steep, most of the tax burden is borne by consumers and quantity falls only modestly; therefore the tax is poor as an immediate conservation policy but may be more effective in the long run when demand becomes more elastic."

Summary: Lesson summary

Concept	Modeling meaning
Qualitative graphical model	Uses direction, curvature, and intersections when exact equations are unknown.
Response curve	Shows what one part of a system requires in response to another.
Equilibrium	Intersection where all requirements are satisfied simultaneously.
Curve shift	Represents a change in policy, technology, cost, or preference.
Staged arms race	Difference-equation version of response curves; stable if reaction product is below one.
$MR = MC$	Competitive firm produces until the next unit costs what it earns.
Supply-demand equilibrium	Intersection of willingness to supply and willingness to buy.
Tax incidence	Tax burden falls more on the less flexible side of the market.
Model criticism	State shape assumptions and identify when a conclusion is only qualitative.

Checkpoint: Exit ticket

In three sentences, explain why graphical modeling can be useful even when exact formulas are unknown. Use either the arms-race model or the gasoline-tax model as your example.

Exercises: Core practice

1. A response curve $y = f(x)$ shifts upward. Explain what happens to the intersection with an unchanged increasing curve $x = g(y)$, using a diagram.
2. For $y_{n+1} = 50 + 0.4x_n$, $x_{n+1} = 30 + 0.2y_n$, find the equilibrium and test stability.
3. A firm has $TC(q) = 200 + 4q + 0.05q^2$ and sells at price $p = 12$. Find q^* and maximum profit.
4. Supply and demand are $S(q) = 2 + 0.3q$, $D(q) = 8 - 0.7q$. Find equilibrium. Then impose a tax $\tau = 1$ and find the new equilibrium.
5. Draw two demand curves, one steep and one flat. For the same supply shift, explain which case gives a larger quantity reduction.

Exercises: Modeling challenge

A city wants to reduce private-car use near a school. It can either impose a parking fee, improve bus service, or redesign the school gate area. Build a qualitative graphical model.

1. Define the axes for your graph. Possible choices: cost of driving vs number of cars, waiting time vs number of students, or supply-demand for parking spaces.
2. State the shape assumptions for each curve.
3. Show how one policy shifts one curve.
4. Predict the direction of change in equilibrium.
5. Write a short recommendation and state what data would be needed to make the prediction quantitative.

Exercises: Extension practice

1. For the staged arms-race system

$$y_{n+1} = a + bx_n, \qquad x_{n+1} = c + dy_n,$$

   derive the equilibrium and show that convergence requires $|bd| < 1$.

2. For linear supply and demand $S(q) = \alpha + \beta q$, $D(q) = \gamma - \delta q$, derive the consumer share of a unit tax τ.
3. Construct a graphical argument for why a subsidy to public transport may reduce gasoline use more in the long run than in the short run.
4. Explain one limitation of the arms-race response-curve model that cannot be fixed by drawing a more detailed graph.

Summary: Optional appendix: smooth deterrence formula

One possible analytic form for a deterrence curve is

$$y = \frac{y_0}{s^{x/y}}, \qquad 0 < s < 1.$$

It says that if X can target Y's missiles more times, Y needs more weapons to retain enough survivors. The formula is less important than its qualitative properties: it is increasing and concave down. In a student report, this formula should be introduced only after the assumptions have been explained in words.

Supplementary 1: Sets, Functions, and Mathematical Language

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this supplementary lesson, readers should be able to:

1. use the language of **sets**, **elements**, **subsets**, **union**, **intersection**, and **complement**;
2. decide whether a relation is a **function** by checking whether each input has exactly one output;
3. determine the **domain** and **range** of common real functions;
4. interpret a graph using the **vertical line test**;
5. compute sums, products, quotients, and **compositions** of functions with attention to domain restrictions;
6. explain why precise mathematical language matters in mathematical modeling.

Why this supplement matters for modeling

Core explanation, worked reasoning, and application

Why this supplement matters for modeling

A mathematical model begins before any equation is written. We must decide what objects are under discussion, which values are allowed, what counts as an input, and what output is produced. These decisions are expressed using the language of sets and functions.

Activity: Launch: define the modeling universe

A school wants to model student lunch waiting time. Discuss:

1. What is the **universal set** of people under discussion: all students, all diners, or everyone on campus?
2. What input could define a function: arrival time, queue length, meal type, or payment method?
3. What output should the function return: waiting time, service time, total time, or probability of waiting more than 10 minutes?

Modeling lesson: Vague sets create vague models. A good model states the population, variables, and valid inputs explicitly.

Sets as modeling categories

Core explanation, worked reasoning, and application

Concept: Core definitions

A **set** is a collection of objects. The objects are called **elements**. If x is an element of A, write $x \in A$; if not, write $x \notin A$.

If every element of A is also in B, then A is a **subset** of B, written $A \subseteq B$. If $A \subseteq B$ but $A \neq B$, then A is a **proper subset**, written $A \subsetneq B$.

The **empty set** is written $\varnothing$. It contains no elements.

Worked Example: Worked Example 1: Classifying students using sets

Let U be the set of students in a modeling club. Define

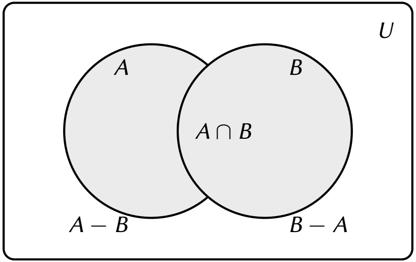
$$A = \{\text{students who know Python}\}, \qquad B = \{\text{students who have joined a modeling contest}\}.$$

Then:

- $A \cap B$ is the set of students who both know Python and have contest experience.
- $A \cup B$ is the set of students who know Python or have contest experience, or both.
- A^c is the set of students in U who do not know Python.
- $B - A$ is the set of students with contest experience but without Python experience.

These are not abstract symbols only; they are categories a teacher may use to form balanced teams.

Illustration: Sets as modeling categories



A Venn diagram is a compact way to reason about categories, overlap, and exclusions.

Concept: Set operations

Given sets A and B inside a universal set U,

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}, \qquad A \cap B = \{x \mid x \in A \text{ and } x \in B\},$$

$$B - A = \{x \mid x \in B \text{ and } x \notin A\}, \qquad A^c = U - A.$$

The word “or” in mathematics is usually inclusive: $x \in A \cup B$ allows x to be in both sets.

Checkpoint: Set-language check

Let $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $A = \{2, 4, 6, 8, 10\}$, and $B = \{1, 2, 3, 5, 8\}$. Find $A \cap B$, $A \cup B$, $B - A$, and A^c. Then write one sentence interpreting each set if A means "even-numbered students" and B means "students selected for a pilot test."

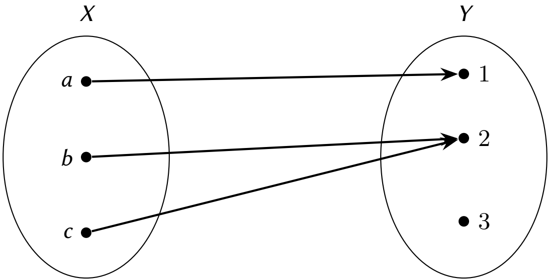
Relations and functions

Core explanation, worked reasoning, and application

Relations and functions

A **relation** from X to Y is any set of ordered pairs (x, y) with $x \in X$ and $y \in Y$. A **function** is a special relation: each input is paired with at most one output. In most calculus contexts, we require every input in the stated domain to have exactly one output.

Illustration: Relations and functions



Function: each input has exactly one output.
Different inputs may share the same output.

A many-to-one relation can still be a function. The forbidden pattern is one input producing two different outputs.

Worked Example: Worked Example 2: Is it a function?

Let $X = \{1, 2, 3\}$ and $Y = \{a, b, c\}$.

1. $R_1 = \{(1, a), (2, a), (3, b)\}$ is a function from X to Y because each input has exactly one output.
2. $R_2 = \{(1, a), (1, b), (2, c)\}$ is not a function because input 1 has two different outputs.
3. $R_3 = \{(1, a), (2, b)\}$ is a function on domain $\{1, 2\}$, but not a function with domain X if every element of X must be used.

Modeling interpretation: If the input is "student ID", then a model for "class level" should not output two different class levels for the same student.

Concept: Domain, codomain, and range

For a function $f : X \to Y$,

- the **domain** is the input set where the function is defined;
- the **codomain** is the declared output universe Y;
- the **range** is the set of outputs that actually occur.

Symbolically,

$$\mathrm{Dom}(f) = \{x \mid f(x) \text{ is defined}\}, \qquad \mathrm{Range}(f) = \{f(x) \mid x \in \mathrm{Dom}(f)\}.$$

Real functions, domain, and graph

Core explanation, worked reasoning, and application

Real functions, domain, and graph

In calculus and modeling, a common case is a **real function**, where both inputs and outputs are real numbers. If a formula is given without a stated domain, the implied domain is the largest set of real numbers for which the formula makes sense.

Worked Example: Worked Example 3: Domain and range from a formula

Consider

$$f(x) = \sqrt{x-2}.$$

For a real-valued output, the expression inside the square root must be non-negative:

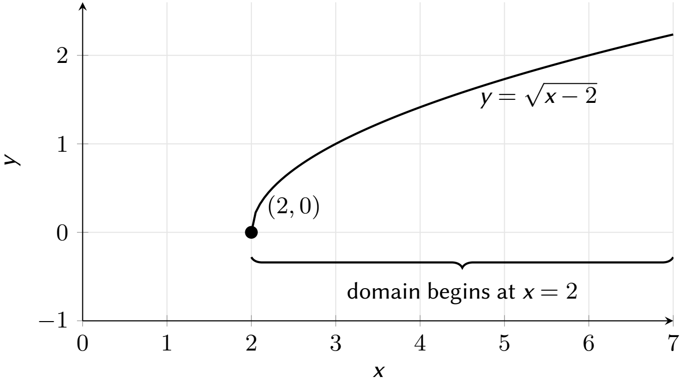
$$x - 2 \geq 0 \quad \Rightarrow \quad x \geq 2.$$

Thus

$$\mathrm{Dom}(f) = [2, \infty), \qquad \mathrm{Range}(f) = [0, \infty).$$

The starting point of the graph is $(2, 0)$.

Illustration: Real functions, domain, and graph

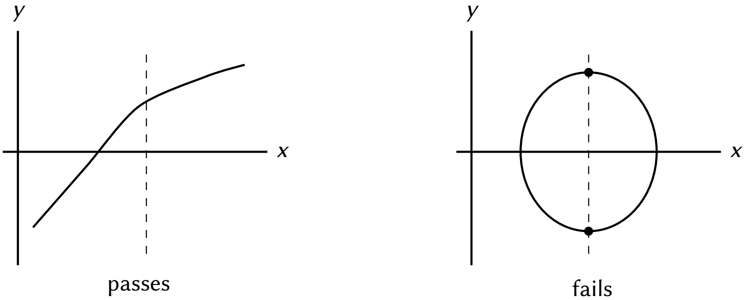


The graph of $f(x) = \sqrt{x-2}$ begins only where the formula produces real values.

Concept: Vertical line test

A curve is the graph of a function of x if every vertical line intersects it at most once. This is the geometric version of the rule: one input, one output.

Illustration: Real functions, domain, and graph



The curve on the left represents a function; the circle on the right does not, because one x gives two y values.

Checkpoint: Domain check

Find the domain of each real function:

1. $f(x) = \dfrac{1}{\sqrt{x-2}}$;
2. $g(x) = \sqrt{9 - x^2}$;
3. $h(x) = \dfrac{x+1}{x^2 - 4}$.

For each, state the restriction in words before writing the interval or set notation.

Operations and composition

Core explanation, worked reasoning, and application

Operations and composition

Functions can be added, subtracted, multiplied, divided, and composed. These operations are useful because many models are built from submodels.

Concept: Algebra of functions

If f has domain A and g has domain B, then the sum, difference, and product are defined on $A \cap B$:

$$(f+g)(x) = f(x) + g(x), \quad (f-g)(x) = f(x) - g(x), \quad (fg)(x) = f(x)g(x).$$

The quotient is defined where both functions are defined and $g(x) \neq 0$:

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}.$$

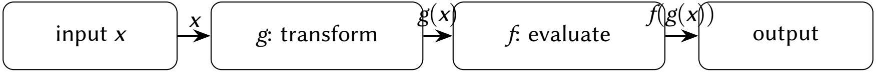
Concept: Composition

The **composite function** $f \circ g$ is

$$(f \circ g)(x) = f(g(x)).$$

The input x must first be valid for g, and then the output $g(x)$ must be valid as an input for f.

Illustration: Operations and composition



A composition is a modeling pipeline: output from one step becomes input to the next.

Composition represents a sequence of transformations.

Worked Example: Worked Example 4: Composition and domain

Let

$$f(u) = \sqrt{u}, \qquad g(x) = 4 - x^2.$$

Then

$$(f \circ g)(x) = f(4 - x^2) = \sqrt{4 - x^2}.$$

The domain is not all real numbers. Since the square root requires $4 - x^2 \geq 0$,

$$-2 \leq x \leq 2.$$

Therefore

$$\mathrm{Dom}(f \circ g) = [-2, 2].$$

Modeling interpretation: If $g(x)$ computes a predicted quantity that must be non-negative before another formula can use it, domain checking prevents meaningless outputs.

Worked Example: Worked Example 5: Iterating a function

Let

$$f(x) = \frac{1}{1+x}.$$

Compute $f(f(f(2)))$:

$$f(2) = \frac{1}{3}, \qquad f(f(2)) = f\left(\frac{1}{3}\right) = \frac{1}{1+1/3} = \frac{3}{4},$$

$$f(f(f(2))) = f\left(\frac{3}{4}\right) = \frac{1}{1+3/4} = \frac{4}{7}.$$

Iteration is important in modeling because a difference equation is just repeated function application:

$$x_{n+1} = f(x_n).$$

Model Critique: Common mistake: ignoring the domain

A formula may look algebraically correct but fail as a real function. For example,

$$\frac{1}{\sqrt{x - |x|}}$$

is never defined for any real x: if $x \geq 0$, then $x - |x| = 0$ and the denominator is zero; if $x < 0$, then $x - |x| = 2x < 0$, so the square root is not real.

Short extension: induction and binomial coefficients

Core explanation, worked reasoning, and application

Short extension: induction and binomial coefficients

This supplementary note originally contains induction and binomial coefficients. They are valuable, but in a two-hour class they should be treated as a short extension or homework pathway. They support discrete models, counting, probability, and binomial distributions later in modeling.

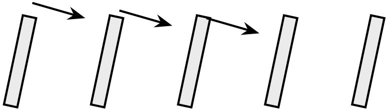
Concept: Mathematical induction

To prove a statement $P(n)$ for all positive integers n:

1. Prove the **base case**: $P(1)$ is true.
2. Prove the **inductive step**: if $P(k)$ is true, then $P(k+1)$ is true.

Then $P(n)$ is true for every $n \in \mathbb{Z}^+$.

Illustration: Short extension: induction and binomial coefficients



Base case starts the chain; inductive step passes truth from one case to the next.

Induction is a logical chain reaction, not a numerical pattern check.

Worked Example: Worked Example 6: Sum of the first n integers

Prove

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Base case: for $n = 1$, left side $= 1$ and right side $= 1(2)/2 = 1$.
Inductive step: assume

$$1 + 2 + \cdots + k = \frac{k(k+1)}{2}.$$

Then

$$1 + 2 + \cdots + k + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{(k+1)(k+2)}{2}.$$

This is exactly the formula for $n = k + 1$. Hence the formula holds for all $n \in \mathbb{Z}^+$.

Concept: Binomial coefficient

The **factorial** is

$$n! = n(n-1)\cdots 2\cdot 1, \qquad 0! = 1.$$

The **binomial coefficient**

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

counts the number of ways to choose k objects from n objects. The **binomial theorem** is

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Checkpoint: Extension check

1. Compute $\binom{8}{3}$ and explain what it counts.
2. Expand $(1+x)^4$ using the binomial theorem.
3. Explain why $\sum_{k=0}^{n}\binom{n}{k} = 2^n$ by interpreting both sides as counting subsets of an n-element set.

Using these ideas in modeling reports

Core explanation, worked reasoning, and application

Model Critique: Mathematical-language checklist

When writing a model, check the following:

1. What is the universal set or population under study?
2. What is the domain of each function or variable?
3. Are there values that must be excluded, such as negative quantities or zero denominators?
4. Does each input determine exactly one output?
5. If two formulas are composed, is the output of the first formula valid as the input of the second?
6. Does the graph match the claimed function and domain?

Summary: Lesson summary

Concept	Modeling meaning
Set	Defines the population, category, or collection being modeled.
Subset	Describes a narrower category within a larger modeling universe.
Union/intersection	Combines or overlaps categories, useful for grouping and filtering data.
Complement	Represents everything in the universe not satisfying a condition.
Function	A deterministic input-output rule; each input has one output.
Domain	States which input values are valid and meaningful.
Range	States which output values the model can actually produce.
Graph	Visual representation of ordered pairs $(x, f(x))$.
Composition	Represents a sequence of model transformations.
Induction/binomial	Supports discrete reasoning, counting, and probability models.

Checkpoint: Exit ticket

A school models total cafeteria time by

$$T(n) = \sqrt{25 - n} + \frac{60}{n},$$

where n is the number of open service counters.

1. What real-number restrictions does the formula impose on n?
2. What additional real-world restrictions should be imposed?
3. Why is domain checking important before optimizing this model?

Exercises: Core practice

1. Let $U = \{1, 2, \ldots, 12\}$, $A = \{2, 4, 6, 8, 10, 12\}$, and $B = \{3, 6, 9, 12\}$. Find $A \cup B$, $A \cap B$, $A - B$, and $(A \cup B)^c$.
2. Decide whether each relation is a function: (a) $\{(1, 2), (2, 2), (3, 4)\}$; (b) $\{(1, 2), (1, 3), (2, 4)\}$; (c) $\{(x, y) \in \mathbb{R}^2 : y = x^2\}$; (d) $\{(x, y) \in \mathbb{R}^2 : x = y^2\}$.
3. Find the domain of each real function: $f(x) = \sqrt{x + 5}$, $g(x) = 1/(x^2 - 9)$, $h(x) = \sqrt{4 - x}/(x + 1)$.
4. Let $f(x) = x^2 + 1$ and $g(x) = 2x - 3$. Find $f \circ g$, $g \circ f$, $(f + g)(x)$, and $(fg)(x)$.
5. Use induction to prove $1 + 3 + 5 + \cdots + (2n - 1) = n^2$.

Exercises: Modeling challenge

A school wants to form modeling teams. Let U be all students in the training program. Define sets for students who know Python, students who are strong in writing, students with previous contest experience, and students available on Saturday.

1. Define four sets using clear notation.
2. Write an expression for the set of students who know Python and are available on Saturday.
3. Write an expression for students who have writing strength but no previous contest experience.
4. Suppose each student is assigned exactly one role: coder, writer, analyst, or presenter. Is "student $\mapsto$ role" a function? Explain.
5. Suppose each student may have several strengths. Is "student $\mapsto$ strength" a function? Explain.
6. Write a short paragraph explaining how this set-and-function language can help a teacher form balanced teams.

Exercises: Extension practice

1. Prove by induction that
$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$
2. Prove that $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$ using the binomial theorem and then explain it using subsets.
3. Let $f(x) = 1/(1+x)$. Find a formula for the first five iterates $f(1), f^{(2)}(1), \ldots, f^{(5)}(1)$. What pattern do you notice?
4. Find the domain of $F(x) = \sqrt{\dfrac{x-1}{x+2}}$. Give your answer in interval notation and justify it with a sign chart.

Supplementary 2: Numbers, Coordinates, and Geometric Models

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this supplementary lesson, readers should be able to:

1. distinguish **rational numbers**, **irrational numbers**, and **real numbers**;
2. interpret **absolute value** as distance and use it to describe tolerances or errors;
3. represent intervals on the **real number line**;
4. use the coordinate plane to model location, distance, midpoint, and coverage;
5. derive and interpret equations of lines and circles;
6. connect slope, distance, and geometry to data fitting, routing, optimization, and modeling constraints.

Why numbers and coordinates matter in modeling

Core explanation, worked reasoning, and application

Why numbers and coordinates matter in modeling

A mathematical model turns a real situation into objects that can be computed, compared, and communicated. Before calculus, optimization, or simulation can be used, students need a precise language for three simple ideas:

$$\text{quantity}, \qquad \text{location}, \qquad \text{distance}.$$

Numbers describe quantities, coordinates describe location, and distance formulas describe closeness, error, travel, or coverage. This supplementary note is therefore not a separate algebra review; it is the coordinate language behind many modeling tasks.

Activity: Launch: choosing the right number system

For each modeling quantity, decide whether integers, rational numbers, or real numbers are the most natural language. Give one sentence of explanation.

1. Number of students assigned to a team.
2. Proportion of students who chose bus travel.
3. Distance between two locations on a map.
4. Error between an observed value and a predicted value.
5. Number of delivery robots required for a school route.

Modeling lesson: Choosing the wrong number system can create wrong decisions. For example, a linear program may output 2.7 robots, but the actual decision must be an integer.

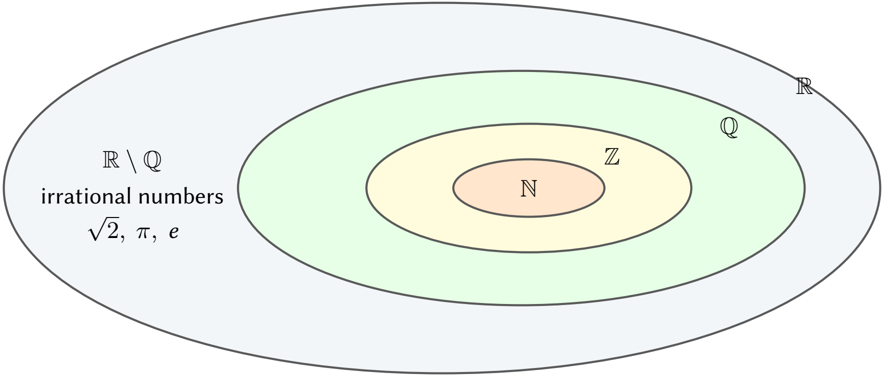
Number systems: from counting to the real line

Core explanation, worked reasoning, and application

Concept: Core vocabulary

Symbol	Name	Modeling meaning
$\mathbb{N}$	natural numbers	counting objects: $1, 2, 3, \ldots$
$\mathbb{Z}$	integers	signed counts or net changes: $\ldots, -2, -1, 0, 1, 2, \ldots$
$\mathbb{Q}$	rational numbers	ratios, fractions, exact proportions: m/n with $n \neq 0$
$\mathbb{R}$	real numbers	continuous quantities: length, time, mass, temperature, error

Illustration: Number systems: from counting to the real line



The hierarchy of number systems. The natural numbers are contained in the integers, the integers are contained in the rational numbers, and the rational numbers are contained in the real numbers. The irrational numbers are the real numbers that are not rational.

Worked Example: Worked Example 1: Classifying modeling quantities

Classify each quantity.

1. A school needs 7 buses. This is in $\mathbb{N}$ because it counts objects.
2. A team completed $\frac{3}{4}$ of its survey. This is in $\mathbb{Q}$ because it is a ratio of integers.
3. The diagonal of a 1×1 square is $\sqrt{2}$. This is irrational, hence real but not rational.
4. A fitted parameter $k = 0.03716\ldots$ from regression is usually treated as real.

Modeling interpretation: Some models are continuous approximations of discrete decisions. That is allowed, but the report must explain how the continuous result is converted into a feasible decision.

Concept: Rational numbers and the gap problem

A **rational number** is a number of the form m/n, where $m, n \in \mathbb{Z}$ and $n \neq 0$. Rational numbers are useful for exact proportions, but they do not fill the number line completely. For example, there is no rational number whose square is 2.

$$\sqrt{2} \notin \mathbb{Q}, \qquad \sqrt{2} \in \mathbb{R}.$$

The real number system is needed because geometry, distance, limits, and calculus naturally produce irrational values.

Worked Example: Worked Example 2: Why $\sqrt{2}$ is not rational

Suppose $\sqrt{2} = m/n$ in lowest terms. Then

$$m^2 = 2n^2.$$

Thus m^2 is even, so m is even. Write $m = 2k$. Then

$$4k^2 = 2n^2 \quad \Rightarrow \quad n^2 = 2k^2,$$

so n is also even. This contradicts the assumption that m/n was in lowest terms. Hence $\sqrt{2}$ is irrational.
Modeling meaning: Even the distance between $(0, 0)$ and $(1, 1)$ is irrational. A coordinate model therefore requires real numbers, not only fractions.

Checkpoint: Number-system check

A route-planning model returns a shortest distance of $5\sqrt{5}$ km, a travel time of 11.2 minutes, and a need for 3.4 vans. Which quantities are continuous model outputs? Which must be converted into integer decisions?

Absolute value, distance, and intervals

Core explanation, worked reasoning, and application

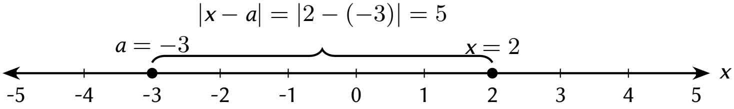
Concept: Absolute value as distance

For a real number a,

$$|a| = \begin{cases} a, & a \geq 0, \\ -a, & a < 0. \end{cases}$$

Geometrically, $|a|$ is the distance from a to 0 on the number line. More generally, $|x - a|$ is the distance between x and a.

Illustration: Absolute value, distance, and intervals



Absolute value measures distance on the real line.

Worked Example: Worked Example 3: Measurement tolerance

A temperature sensor is acceptable if its reading x is within 0.4°C of the true value 25°C. Write this condition using absolute value and interval notation.

$$|x - 25| < 0.4.$$

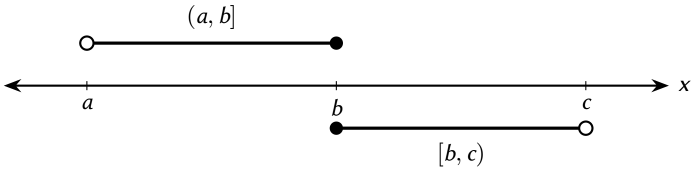
This means

$$24.6 < x < 25.4,$$

so the acceptable interval is $(24.6, 25.4)$.

Modeling interpretation: Absolute value inequalities are a compact way to state tolerance, error bounds, residual thresholds, and safety margins.

Illustration: Absolute value, distance, and intervals



Open endpoints are drawn as hollow circles; closed endpoints are filled.

Checkpoint: Interval check

Rewrite each condition in interval notation.

1. $|x - 6| \leq 2$.
2. A residual r satisfies $|r| < 0.1$.
3. A model parameter k is positive but less than 1.

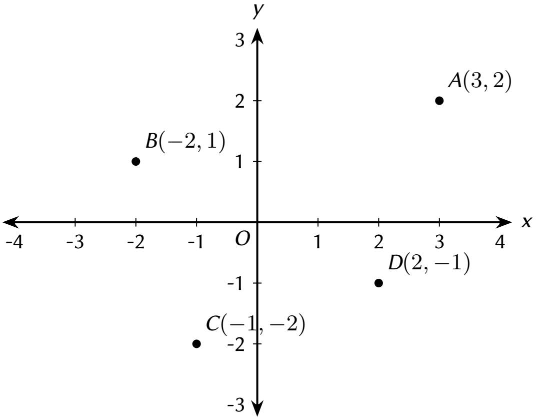
Coordinates, distance, and midpoint

Core explanation, worked reasoning, and application

Coordinates, distance, and midpoint

A point in the plane is represented by an ordered pair (x, y). In modeling, coordinates may represent a physical map, a feature space, or two variables measured from data.

Illustration: Coordinates, distance, and midpoint



Coordinates locate points relative to the *x*- and *y*-axes.

Concept: Distance and midpoint formulas

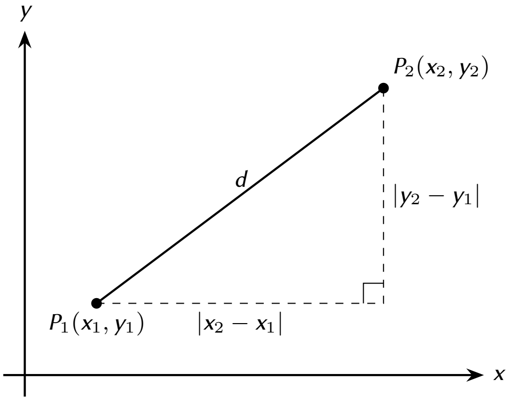
For $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$,

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

The midpoint is

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right).$$

Illustration: Coordinates, distance, and midpoint



The distance formula is the Pythagorean theorem applied to coordinate differences.

Worked Example: Worked Example 4: Choosing a meeting point

Two students live at $A = (1, 2)$ and $B = (7, 6)$ on a simplified map measured in km. A fair meeting point halfway between them is

$$M = \left(\frac{1+7}{2}, \frac{2+6}{2} \right) = (4, 4).$$

The direct distance between them is

$$d(A, B) = \sqrt{(7-1)^2 + (6-2)^2} = \sqrt{36+16} = \sqrt{52} \approx 7.21 \text{ km}.$$

Modeling interpretation: The midpoint is fair under straight-line distance. If actual walking routes follow roads, the Euclidean midpoint may no longer be fair.

Model Critique: Model criticism: Euclidean distance versus real travel distance

The distance formula assumes a flat coordinate plane and straight-line movement. In route planning, road networks, stairs, slopes, traffic lights, rivers, gates, and restricted areas may make Euclidean distance a poor proxy. A modeler should say whether Euclidean distance is a reasonable approximation or only a first estimate.

Lines as simple models

Core explanation, worked reasoning, and application

Lines as simple models

A line is the simplest model of constant rate of change.

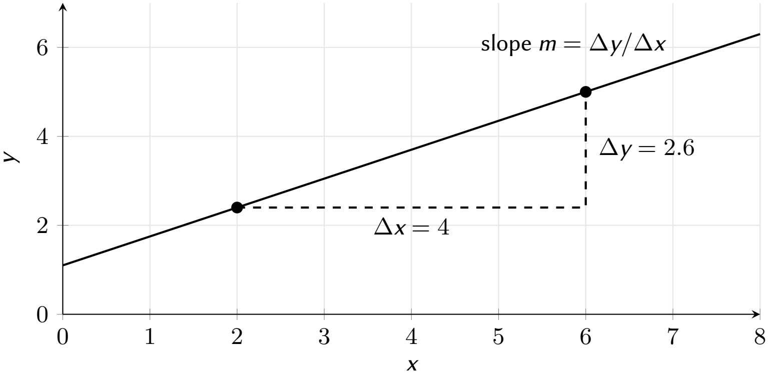
$$\text{slope} = \frac{\text{change in output}}{\text{change in input}} = \frac{\Delta y}{\Delta x}.$$

In modeling, slope carries units. If x is time in minutes and y is distance in metres, then slope is metres per minute.

Concept: Forms of a line

Form	Equation	Use
Slope-intercept	$y = mx + b$	slope and starting value are known
Point-slope	$y - y_1 = m(x - x_1)$	one point and slope are known
Two-point	$m = \dfrac{y_2 - y_1}{x_2 - x_1}$	two observations are known
General form	$Ax + By + C = 0$	useful for point-to-line distance

Illustration: Lines as simple models



The slope of a line measures constant change per unit input.

Worked Example: Worked Example 5: A linear travel-time model

A delivery robot takes 4 minutes to travel 120 m and 10 minutes to travel 300 m along a corridor. Let x be distance in metres and y be travel time in minutes. Assuming a linear model,

$$m = \frac{10-4}{300-120} = \frac{6}{180} = \frac{1}{30}\ \text{min/m}.$$

Using point-slope form through $(120, 4)$,

$$y - 4 = \frac{1}{30}(x - 120),$$

so

$$y = \frac{x}{30}.$$

The intercept is 0, suggesting that the robot has negligible fixed start-up delay in this simplified case.
Interpretation: The speed is the reciprocal of the slope: 30 m/min. If later data show a nonzero intercept, that may represent loading time or waiting time.

Checkpoint: Line check

Find the line through $(2, 5)$ and $(8, 17)$. Interpret the slope if *x* is the number of hours studied and *y* is a test score.

Concept: Parallel, perpendicular, and point-to-line distance

For non-vertical lines, parallel lines have the same slope. Perpendicular lines have slopes satisfying

$$m_1 m_2 = -1.$$

The distance from (x_0, y_0) to the line $Ax + By + C = 0$ is

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}.$$

This formula measures the shortest perpendicular distance from a point to a linear boundary.

Worked Example: Worked Example 6: Distance to a safety boundary

A restricted area is approximated by the line

$$3x - 4y + 12 = 0.$$

A drone is at $P = (2, 5)$. Its shortest distance to the boundary is

$$d = \frac{|3(2) - 4(5) + 12|}{\sqrt{3^2 + (-4)^2}} = \frac{|6 - 20 + 12|}{5} = \frac{2}{5} = 0.4.$$

If coordinates are measured in kilometres, the drone is 0.4 km from the boundary.

Circles as distance constraints

Core explanation, worked reasoning, and application

Circles as distance constraints

A circle is a set of points at a fixed distance from a centre. This makes circles useful for coverage, range, safety, and service-radius models.

Concept: Equation of a circle

A circle with centre (h, k) and radius r has equation

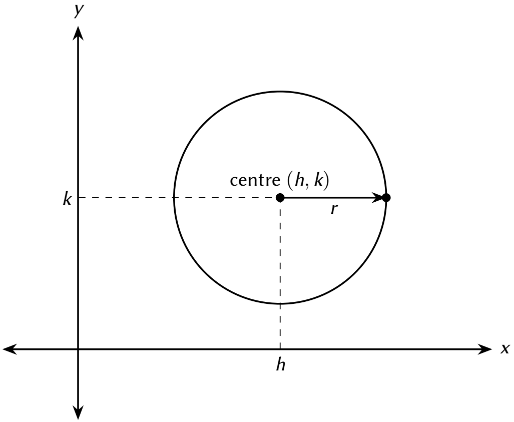
$$(x - h)^2 + (y - k)^2 = r^2.$$

The general form

$$x^2 + y^2 + Dx + Ey + F = 0$$

can be converted to standard form by completing the square.

Illustration: Circles as distance constraints



A circle models all locations within a fixed distance from a centre.

Worked Example: Worked Example 7: Coverage radius

A Wi-Fi router is located at $(3, -2)$ and has effective range 5 m. The coverage boundary is

$$(x - 3)^2 + (y + 2)^2 = 25.$$

A device at $(7, 1)$ has squared distance

$$(7 - 3)^2 + (1 + 2)^2 = 16 + 9 = 25,$$

so it lies exactly on the boundary.

Modeling interpretation: In real life, walls and interference make the boundary non-circular. The circle is a first approximation based on distance alone.

Worked Example: Worked Example 8: Completing the square

Find the centre and radius of

$$x^2 + y^2 - 6x + 4y - 12 = 0.$$

Group terms and complete the square:

$$(x^2 - 6x) + (y^2 + 4y) = 12,$$

$$(x - 3)^2 + (y + 2)^2 = 12 + 9 + 4 = 25.$$

Hence the centre is $(3, -2)$ and the radius is 5.

Activity: Differentiated geometric-modeling studio

Choose one route.

1. **Core route: facility distance.** A help desk is at $(2, 1)$. Find which of $A = (5, 5)$, $B = (-1, 3)$, and $C = (4, -2)$ is closest.
2. **Challenge route: coverage model.** A sensor at $(4, 3)$ covers radius 3. Write the coverage inequality and test whether $(6, 5)$ and $(8, 3)$ are covered.
3. **Extension route: boundary distance.** A road is modeled by $2x - y + 1 = 0$. A school at $(3, 4)$ must be at least 1 km from the road. Check whether the location is acceptable and suggest a direction to move.

Optional appendix: completeness and why calculus works

Core explanation, worked reasoning, and application

Summary: Completeness in one page

A set $S \subset \mathbb{R}$ is **bounded above** if there is a number M with $x \leq M$ for all $x \in S$. The **supremum** is the least upper bound.
The real numbers satisfy the **completeness axiom**: every nonempty subset of $\mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$.
This is deeper than what is needed for ordinary coordinate geometry, but it explains why limits and calculus are possible. Rational numbers have gaps; real numbers fill those gaps.

Model Critique: Competition-style communication

When using coordinates in a report, do not merely write formulas. State:

1. what the axes represent and what units are used;
2. why Euclidean distance is acceptable, or what it ignores;
3. whether variables are continuous or must be rounded to integers;
4. how distance, slope, or circles support the final recommendation.

Weak statement: "The distance is 5."
Stronger statement: "Using a coordinate map measured in kilometres, the Euclidean distance between the station and Building A is 5 km. This is a lower-bound estimate because actual walking routes must follow campus roads."

Summary: Lesson summary

Concept	Modeling meaning
$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$	Different number systems support counting, ratios, and continuous measurement.
Irrational number	Natural geometric quantities such as $\sqrt{2}$ cannot be expressed as fractions.
Absolute value	Measures distance from zero or error from a target.
Interval	Describes acceptable ranges, constraints, and uncertainty bands.
Coordinate plane	Converts location into computable ordered pairs.
Distance formula	Measures straight-line closeness, travel lower bounds, and error.
Slope	Constant rate of change; units matter.
Line	Simplest model of linear relationship or boundary.
Circle	Set of points at fixed distance; useful for coverage and range models.
Completeness	The hidden property of $\mathbb{R}$ that supports limits and calculus.

Checkpoint: Exit ticket

A school map uses coordinates in metres. In three sentences, explain how you would model whether a new water station covers all classrooms within 80 m. Mention the centre, the radius, and one limitation of the model.

Exercises: Core practice

1. Classify each number as natural, integer, rational, irrational, or real: $-4, 0, 7/3, \sqrt{9}, \sqrt{7}, \pi$.
2. Solve and write in interval notation: $|x - 4| < 1.5$ and $|2x + 1| \leq 5$.
3. Find the distance and midpoint between $A = (-2, 3)$ and $B = (4, -1)$.
4. Find the equation of the line through $(1, 5)$ and $(4, 11)$.
5. Find the line through $(2, -1)$ perpendicular to $y = 3x + 4$.
6. Find the centre and radius of $x^2 + y^2 + 2x - 8y + 8 = 0$.

Exercises: Modeling challenge

A campus planner wants to place a temporary information booth. Buildings are represented by points

$$A = (0,0), \quad B = (6,2), \quad C = (3,7), \quad D = (-2,4).$$

1. Compute the midpoint of A and B. Explain why this may or may not be a good booth location.
2. Find the distance from the proposed booth $P = (2,3)$ to each building.
3. If the booth can serve students within radius 5, write the coverage circle and determine which buildings are covered.
4. A main walkway is approximated by the line $x + y = 5$. Find the distance from P to the walkway.
5. Write a short recommendation. Include one limitation of using coordinates and Euclidean distance.

Exercises: Extension practice

1. Prove that the product of a nonzero rational number and an irrational number is irrational.
2. Prove the triangle inequality $|a + b| \le |a| + |b|$ using the fact that $-|a| \le a \le |a|$.
3. A set $S = (0, 1)$ has no maximum but has supremum 1. Explain the difference between maximum and supremum.
4. Find all points on the x-axis that are equidistant from $(2, 3)$ and $(-1, 4)$.
5. A circle passes through $(0, 0)$, $(4, 0)$, and $(0, 3)$. Find its equation.

Supplementary 3: Graphs, Transformations, and Periodic Models

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this supplementary lesson, readers should be able to:

1. read a **graph** as a model of a relationship between two quantities;
2. distinguish a **relation** from a **function** using the vertical-line test;
3. determine **domain**, **range**, intercepts, symmetry, and key features from formulas and graphs;
4. use graph transformations to sketch related functions without re-plotting from scratch;
5. interpret elementary graph families as modeling shapes: linear, quadratic, reciprocal, absolute value, square-root, and sinusoidal;
6. build a simple **periodic model** using amplitude, period, phase shift, and vertical shift;
7. explain what a graph reveals, what it hides, and what assumptions are needed before using it for prediction.

Graphs as modeling language

Core explanation, worked reasoning, and application

Graphs as modeling language

A graph is not merely a picture of a formula. In mathematical modeling, a graph is a compact way to communicate a relationship: increase, decrease, saturation, threshold, periodicity, outliers, and breakdown of assumptions can often be seen before they are calculated.

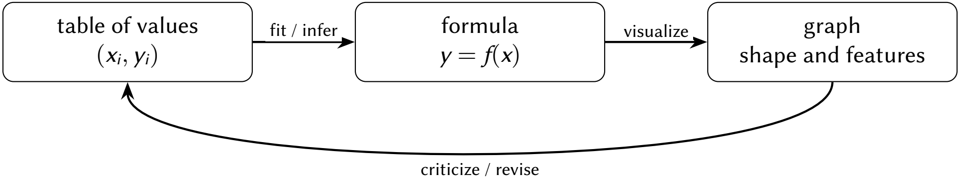
Activity: Launch: what can a graph say?

A school records the waiting time at a canteen from 12:00 to 12:40. The curve rises quickly, reaches a peak, then slowly returns to nearly zero.

1. What does the peak mean in context?
2. What would a flatter curve mean?
3. What information is missing if we only know the curve shape but not the units?
4. What intervention could move the peak downward or earlier?

Modeling lesson: a graph should be read as a statement about a system, not only as a drawing.

Illustration: Graphs as modeling language

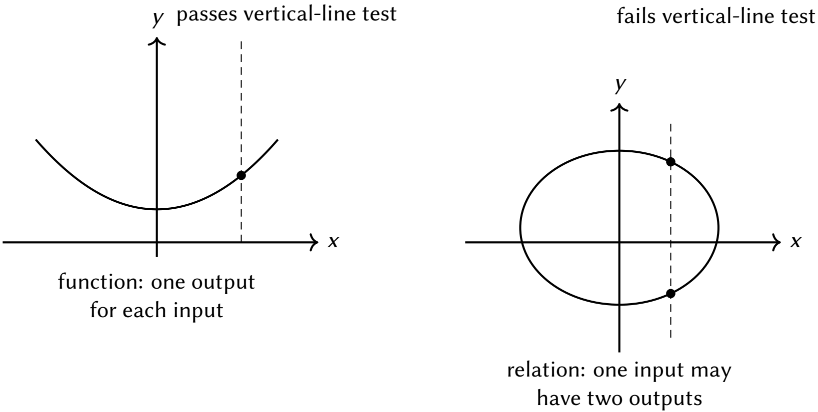


In modeling, data, formula, and graph should support one another.

Concept: Core vocabulary

A **relation** is a set of ordered pairs (x, y). A **function** is a special relation in which each input x corresponds to exactly one output y. The **domain** is the set of allowed inputs; the **range** is the set of outputs produced. The graph of $y = f(x)$ is the set of all points $(x, f(x))$.

Illustration: Graphs as modeling language



The vertical-line test checks whether a relation can be written as $y = f(x)$.

Domain, range, intercepts, and symmetry

Core explanation, worked reasoning, and application

Domain, range, intercepts, and symmetry

Before fitting or interpreting a graph, a modeler must decide which values are meaningful. In school mathematics, a formula may have a natural algebraic domain. In modeling, the realistic domain may be smaller.

Worked Example: Worked Example 1: Natural domain versus modeling domain

A simplified stopping-distance model is

$$D(v) = 0.75v + 0.006v^2,$$

where v is speed in km/h and D is stopping distance in metres.
Algebraic domain: the formula accepts all real v. **Modeling domain:** speed cannot be negative, and the model was calibrated only for ordinary road speeds, say $0 \leq v \leq 120$.
The y-intercept is $D(0) = 0$, meaning no stopping distance when speed is zero. There is no meaningful negative-speed branch, even though the algebraic formula could be evaluated there. The graph should therefore be drawn only on the modeling domain.

Worked Example: Worked Example 2: Domain and range from a formula

Find the natural domain and range of

$$f(x) = \sqrt{4 - x^2}.$$

The square root requires $4 - x^2 \geq 0$, so $-2 \leq x \leq 2$. Hence

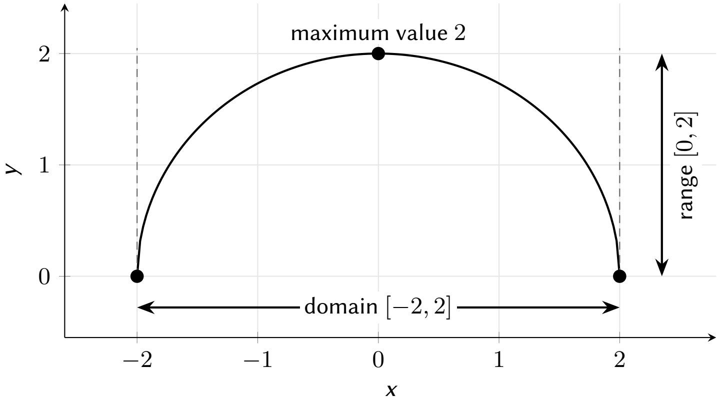
$$\mathrm{Dom}(f) = [-2, 2].$$

Since $f(x)$ is the upper half of the circle $x^2 + y^2 = 4$, the output satisfies $0 \leq y \leq 2$:

$$\mathrm{Range}(f) = [0, 2].$$

This is a function because each x in $[-2, 2]$ gives exactly one nonnegative y.

Illustration: Domain, range, intercepts, and symmetry



The graph of $y = \sqrt{4 - x^2}$ is the upper semicircle of radius 2, so its domain and range are visible.

Checkpoint: Feature check

For each function, state the natural domain, one intercept, and whether the graph is even, odd, or neither.

1. $f(x) = x^4 - 3x^2$
2. $g(x) = \dfrac{1}{x-2}$
3. $h(x) = \sqrt{x+1}$

Graph transformations and standard shapes

Core explanation, worked reasoning, and application

Graph transformations and standard shapes

Graph transformations allow us to sketch related functions quickly. This is important in modeling because fitted or theoretical curves often differ from a standard graph only by scaling or shifting.

Concept: Transformations of $y = f(x)$

New function	Effect on the graph
$y = f(x) + c$	shift upward by c
$y = f(x - c)$	shift right by c
$y = af(x)$	vertical stretch by factor $\lvert a \rvert$; reflect in x-axis if $a < 0$
$y = f(bx)$	horizontal compression by factor b if $b > 1$
$y = -f(x)$	reflect in the x-axis
$y = f(-x)$	reflect in the y-axis

Worked Example: Worked Example 3: Transforming a graph

Describe the graph of

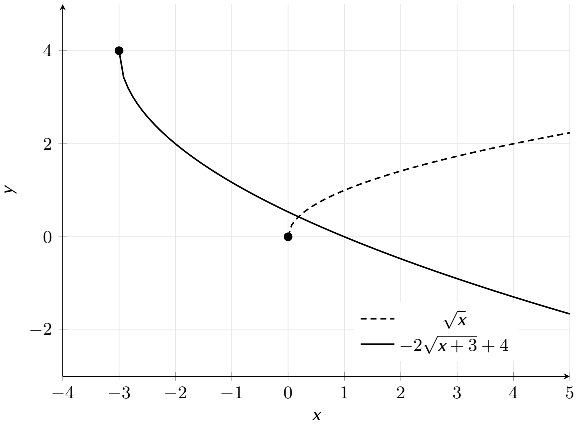
$$y = -2\sqrt{x+3} + 4.$$

Start with $y = \sqrt{x}$.

1. $\sqrt{x+3}$ shifts the graph left by 3.
2. $-2\sqrt{x+3}$ reflects it in the x-axis and stretches vertically by factor 2.
3. $-2\sqrt{x+3} + 4$ shifts the graph upward by 4.

The starting point moves from $(0,0)$ to $(-3,4)$. The domain is $[-3,\infty)$ and the range is $(-\infty,4]$.

Illustration: Graph transformations and standard shapes



Transformations let us sketch a new model from a familiar parent graph.

Summary: Standard graph families for modeling

Family	Typical formula	Modeling meaning
Linear	$mx + b$	constant rate of change; calibration line
Quadratic	$ax^2 + bx + c$	accelerating change, projectile-like shapes, cost curves
Absolute value	$a\|x - h\| + k$	distance from a target, penalty around a desired value
Square root	$a\sqrt{x - h} + k$	fast initial growth then slowing; diminishing gains
Reciprocal	$a/(x - h) + k$	inverse relationship, asymptote, capacity or congestion effects
Sinusoidal	$A \sin(B(t - C)) + D$	periodic or seasonal behaviour

Model Critique: Choosing a graph family is an assumption

If a student chooses a quadratic graph only because it fits six points well, the model may extrapolate badly. A graph family should be justified by a mechanism: constant rate, saturation, inverse dependence, periodic forcing, or geometric constraint. A competition report should state why the chosen shape is meaningful, not only that it fits the data.

Trigonometric graphs as periodic models

Core explanation, worked reasoning, and application

Trigonometric graphs as periodic models

The most useful part of trigonometry for modeling is not memorizing identities; it is recognizing repeated behaviour. Many systems have a repeating pattern: daylight, tides, traffic cycles, temperature, electricity demand, and school attendance.

Concept: Sinusoidal model

A basic periodic model has the form

$$y = A\sin(B(t - C)) + D.$$

Here $|A|$ is the **amplitude**, $2\pi/|B|$ is the **period**, C is the **phase shift**, and D is the **midline**. The same interpretation applies to cosine models.

Worked Example: Worked Example 4: A seasonal attendance model

Suppose the average number of students absent from a school follows a yearly cycle. The minimum is about 8 students in September, the maximum is about 28 students in February, and the period is one year. Let t be months after September.

The midline is

$$D = \frac{28+8}{2} = 18,$$

and the amplitude is

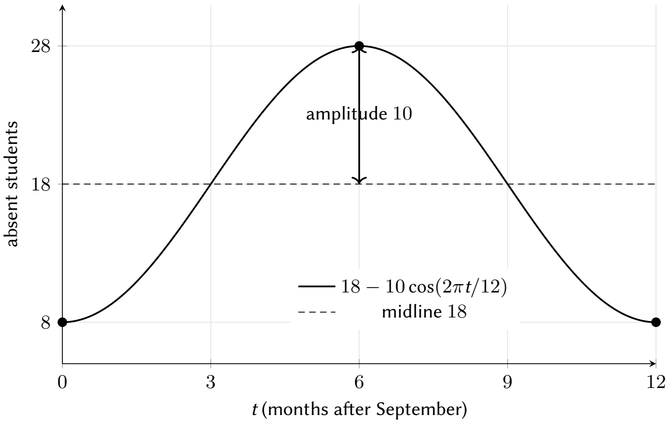
$$A = \frac{28-8}{2} = 10.$$

Because the minimum occurs at $t = 0$, a convenient cosine model is

$$\boxed{A(t) = 18 - 10\cos\left(\frac{2\pi}{12}t\right).}$$

At $t = 6$, the model gives $A(6) = 28$, representing the winter peak. This model should be validated against several years of records before being used for staffing decisions.

Illustration: Trigonometric graphs as periodic models



A sinusoidal graph turns a repeating pattern into interpretable parameters.

Checkpoint: Periodic model check

A tide height ranges from 1.2 m to 3.8 m and repeats every 12 hours. At $t = 0$, the tide is at its maximum.

1. Find the amplitude and midline.
2. Write a cosine model.
3. Explain one limitation of this model.

Asymptotes and conics as reserve modeling shapes

Core explanation, worked reasoning, and application

Asymptotes and conics as reserve modeling shapes

Some graphs are useful because they show boundaries rather than ordinary turning points. An **asymptote** describes a value the curve approaches but does not cross in a simple way. A **conic section** describes geometric constraints such as fixed distance, fixed sum of distances, or fixed difference of distances.

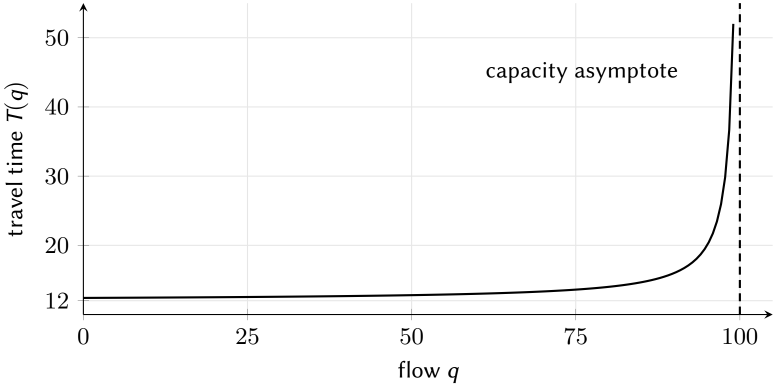
Worked Example: Worked Example 5: Reading asymptotes in a congestion model

Suppose a simple travel-time model is

$$T(q) = 12 + \frac{40}{100 - q}, \qquad 0 \leq q < 100,$$

where q is the number of vehicles entering a short road segment per minute. The graph has a vertical asymptote at $q = 100$. As q approaches capacity, the predicted travel time increases rapidly. The asymptote is a warning: operating close to capacity is unstable and small demand changes can cause large delays.

Illustration: Asymptotes and conics as reserve modeling shapes



An asymptote can represent a practical capacity boundary.

Summary: Conics in one paragraph

Conic	Defining idea	Modeling use
Circle	fixed distance from a centre	coverage, safety radius, sensor range
Parabola	equal distance from focus and directrix	projectile shape, reflector geometry
Ellipse	fixed sum of distances to two foci	orbit-like shapes, service regions
Hyperbola	fixed difference of distances to two foci	localization from time-difference signals

Writing graph-based modeling conclusions

Core explanation, worked reasoning, and application

Model Critique: Competition-style communication

A graph-based conclusion should include four parts:

1. **Feature:** What graph property matters? For example: peak, intercept, slope, asymptote, period.
2. **Meaning:** What does that feature represent in the real system?
3. **Decision:** What action follows from this interpretation?
4. **Limitation:** What would make the graph unreliable?

Weak conclusion: “The graph goes up and then down.”
Stronger conclusion: “The waiting-time curve peaks at about 12:18, so the canteen bottleneck is concentrated in the first half of lunch. A second service counter should be opened before the peak rather than after it. This conclusion assumes that the recorded day is typical and that waiting time was measured consistently.”

Activity: Differentiated graph-modeling studio

Choose one route.

1. **Core route:** Given the graph of a quadratic cost model, identify intercepts, vertex, and the meaningful domain.
2. **Challenge route:** Fit a sinusoidal model from maximum, minimum, and period; then explain the phase shift.
3. **Extension route:** Choose between a reciprocal model and a square-root model for a congestion or learning-curve dataset. Justify by mechanism and residual pattern.

Summary: Lesson summary

Concept	Modeling meaning
Graph	Visual representation of a relation or function; reveals qualitative structure.
Vertical-line test	Checks whether each input has only one output.
Domain and range	Specify realistic inputs and possible outputs.
Intercepts	Describe starting values, thresholds, or break-even points.
Symmetry	Reduces work and reveals structural assumptions.
Transformation	Reuses known graph shapes through shifts, stretches, and reflections.
Sinusoidal model	Represents periodic behaviour through amplitude, period, phase, and midline.
Asymptote	Indicates a limiting value or practical boundary.
Conic	Encodes geometric constraints such as distance, sum of distances, or coverage.

Checkpoint: Exit ticket

Pick one graph feature from today's lesson and explain how it could affect a real decision in a modeling problem. Your answer must include the feature, its context meaning, and one limitation.

Exercises: Core practice

1. Find the natural domain of $f(x) = \dfrac{\sqrt{9 - x^2}}{x - 1}$.
2. Determine whether each function is even, odd, or neither: $x^5 - x$, $x^4 + 2x^2$, $x^2 + x$.
3. Describe the transformations from $y = x^2$ to $y = -3(x + 2)^2 + 5$.
4. State the amplitude, period, phase shift, and midline of $y = 4\sin(2t - \pi) + 3$.
5. Find the vertical and horizontal asymptotes of $f(x) = \dfrac{2x + 1}{x - 3}$.

Exercises: Modeling challenge

A small theme park records the number of visitors entering per hour from 9:00 to 18:00. The data rise slowly in the morning, peak around 14:00, then decline. A student proposes a quadratic model.

1. Define a time variable and a visitor-count function.
2. Explain why a quadratic graph may be reasonable over one day.
3. State the modeling domain and explain why extrapolating beyond it may be misleading.
4. Suggest one graph feature that would help staffing decisions.
5. Write a short graph-based conclusion with one limitation.

Exercises: Extension practice

1. Prove that the graph of $x^2 + y^2 = 4$ is not the graph of a function of x, but can be split into two functions.
2. A signal is modeled by $S(t) = 2.5\cos(\pi t/6) + 7$. Find its maximum, minimum, period, and value at $t = 3$.
3. Explain how the asymptote in $T(q) = 12 + 40/(100 - q)$ changes if the road capacity increases from 100 to 130 vehicles per minute.
4. Choose a conic to model the set of points within 50 m of a Wi-Fi router. Write the equation if the router is at $(20, 15)$.

Summary: Optional appendix: trigonometric identities

The identities below are useful for algebra and calculus, but they need not all be taught in the two-hour core.

$$
\begin{aligned}
\sin^2\theta + \cos^2\theta &= 1,\\
\sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta,\\
\cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta,\\
\sin(2\theta) &= 2\sin\theta\cos\theta,\\
\cos(2\theta) &= 2\cos^2\theta - 1 = 1 - 2\sin^2\theta.
\end{aligned}
$$

Supplementary 4: Limits, Continuity, and Asymptotic Thinking

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this supplementary lesson, readers should be able to:

1. explain a **limit** as the value approached by a function near a point, not necessarily at the point;
2. evaluate common limits by substitution, factorisation, rationalisation, and standard trigonometric limits;
3. use **one-sided limits** to decide whether a two-sided limit exists;
4. interpret **infinite limits**, **limits at infinity**, and **asymptotes**;
5. classify removable, jump, and infinite discontinuities;
6. apply **continuity**, the **Intermediate Value Theorem**, and the **Extreme Value Theorem** in modeling contexts;
7. explain numerical and graphical evidence for a limit without confusing it with proof.

Limits as local behaviour

Core explanation, worked reasoning, and application

Limits as local behaviour

A **limit** describes what a function is approaching. It is local: it looks near a point, not necessarily at the point. This idea is central to calculus, but it is also a modeling idea. A sensor may fail at one time, a formula may be undefined at one value, or a process may approach a capacity without ever reaching it exactly.

Activity: Launch: the missing sensor reading

A temperature sensor records readings every minute. At $t = 5$, the sensor fails, but nearby readings suggest the curve should pass near 23.8°C.

1. Should we say the temperature at $t = 5$ is known?
2. Should we say the *trend* near $t = 5$ is known?
3. What evidence would make the estimated limiting value more credible?

Modeling lesson: a limit is about approaching behaviour. It can be meaningful even when the value at the exact input is missing or unreliable.

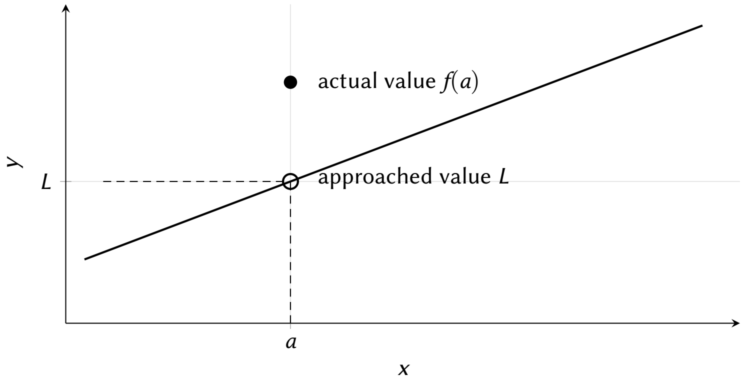
Concept: Core definition, stated informally

We write

$$\lim_{x \to a} f(x) = L$$

when $f(x)$ can be made as close as we like to L by choosing x sufficiently close to a, with $x \neq a$. The value $f(a)$ may equal L, differ from L, or fail to exist.

Illustration: Limits as local behaviour



A limit describes the value approached by the graph near $x = a$. The actual value at a can be different.

Worked Example: Worked Example 1: A removable hole

Evaluate

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1}.$$

At $x = 1$, the expression is undefined. But for $x \neq 1$,

$$\frac{x^2 - 1}{x - 1} = \frac{(x - 1)(x + 1)}{x - 1} = x + 1.$$

Therefore

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = \lim_{x \to 1}(x + 1) = 2.$$

Interpretation: the original formula has a hole at $x = 1$, but the nearby trend approaches 2.

Worked Example: Worked Example 2: Rationalisation

Evaluate

$$\lim_{x\to 0} \frac{\sqrt{x+4}-2}{x}.$$

Direct substitution gives $0/0$, so we rationalise:

$$\frac{\sqrt{x+4}-2}{x}\cdot\frac{\sqrt{x+4}+2}{\sqrt{x+4}+2} = \frac{x}{x(\sqrt{x+4}+2)} = \frac{1}{\sqrt{x+4}+2}, \qquad x \neq 0.$$

Hence

$$\lim_{x\to 0} \frac{\sqrt{x+4}-2}{x} = \frac{1}{2+2} = \frac{1}{4}.$$

Checkpoint: Algebraic limit check

Evaluate:

1. $\displaystyle\lim_{x\to 2} \frac{x^2 - 4}{x - 2}$;
2. $\displaystyle\lim_{x\to 9} \frac{\sqrt{x} - 3}{x - 9}$;
3. $\displaystyle\lim_{x\to 1} \frac{x^3 - 1}{x^2 - 1}$.

For each one, state the obstacle first: direct substitution, $0/0$, square root, or factorisation.

From informal limits to epsilon-delta language

Core explanation, worked reasoning, and application

From informal limits to epsilon-delta language

In a first modeling course, students do not need to master every ε-δ **proof**. They should, however, understand the meaning of the definition: output tolerance controls input tolerance.

Concept: The formal definition

We say

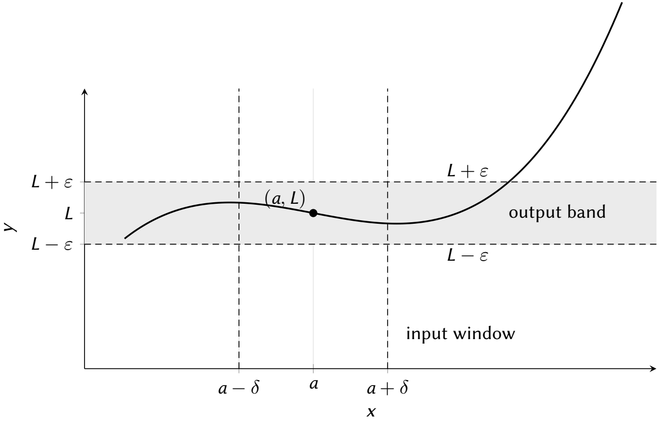
$$\lim_{x \to a} f(x) = L$$

if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

Here ε is the allowed error in the output, while δ is the allowed distance from the input a.

Illustration: From informal limits to epsilon-delta language



The epsilon-delta definition of a limit: if x is sufficiently close to a, then $f(x)$ is within ε of L.

Worked Example: Worked Example 3: A simple epsilon-delta proof

Prove that $\lim_{x \to 3}(2x - 1) = 5$.
We want

$$|(2x - 1) - 5| = |2x - 6| = 2\,|x - 3| < \varepsilon.$$

So it is enough to require $|x - 3| < \varepsilon/2$. Let $\delta = \varepsilon/2$. Then whenever $0 < |x - 3| < \delta$,

$$|(2x - 1) - 5| = 2\,|x - 3| < 2\delta = \varepsilon.$$

Therefore $\lim_{x \to 3}(2x - 1) = 5$.

Model Critique: Proof versus evidence

A numerical table can suggest a limit; a graph can illustrate a limit; an epsilon-delta proof establishes the limit.
In mathematical modeling reports, students usually need to explain the trend and justify the computation.
Formal proof is useful when the result will be used as a theorem or when the graph is misleading.

One-sided and infinite limits

Core explanation, worked reasoning, and application

One-sided and infinite limits

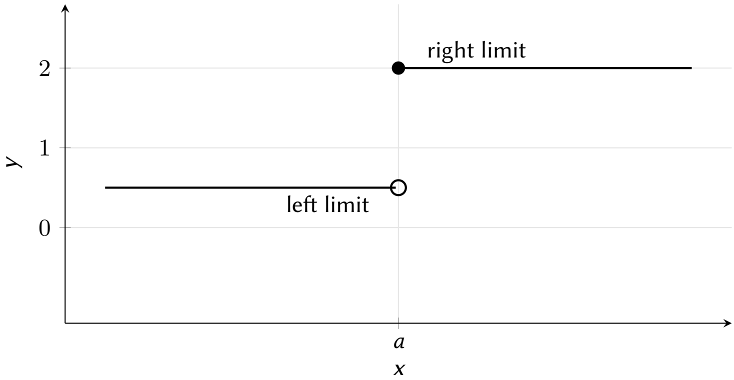
Some functions approach different values from the left and from the right. A two-sided limit exists only when the two one-sided limits agree.

Concept: One-sided limits

The **right-hand limit** is $\lim_{x \to a^+} f(x)$; it approaches a using $x > a$. The **left-hand limit** is $\lim_{x \to a^-} f(x)$; it approaches a using $x < a$.

$$\lim_{x \to a} f(x) = L \quad \iff \quad \lim_{x \to a^-} f(x) = L \text{ and } \lim_{x \to a^+} f(x) = L.$$

Illustration: One-sided and infinite limits



A jump occurs when the left-hand and right-hand limits are different.

Worked Example: Worked Example 4: Absolute value quotient

Evaluate

$$\lim_{x \to 0} \frac{|x|}{x}.$$

If $x > 0$, then $|x| = x$, so $|x|/x = 1$. If $x < 0$, then $|x| = -x$, so $|x|/x = -1$. Therefore

$$\lim_{x \to 0^+} \frac{|x|}{x} = 1, \qquad \lim_{x \to 0^-} \frac{|x|}{x} = -1.$$

Since the two one-sided limits are unequal, the two-sided limit does not exist.

Concept: Infinite limits and vertical asymptotes

Writing $\lim_{x \to a} f(x) = +\infty$ means that $f(x)$ grows without bound as x approaches a. It does not mean that $+\infty$ is an ordinary real-number limit. If at least one one-sided limit is infinite, then the line $x = a$ is a **vertical asymptote**.

Worked Example: Worked Example 5: Capacity blow-up

A simple congestion model for waiting time is

$$W(q) = \frac{1}{1-q}, \qquad 0 \le q < 1,$$

where q is demand as a fraction of capacity. Then

$$\lim_{q \to 1^-} W(q) = +\infty.$$

Interpretation: as demand approaches full capacity, the model predicts waiting time grows extremely fast. The vertical asymptote $q = 1$ is a warning that operating close to capacity is unstable.

Limits at infinity and asymptotic thinking

Core explanation, worked reasoning, and application

Limits at infinity and asymptotic thinking

A **limit at infinity** describes long-run behaviour. In modeling, it answers questions such as: Does the prediction stabilise? Does a cost approach a baseline? Does a probability approach one?

Worked Example: Worked Example 6: Long-run behaviour of a rational model

Evaluate

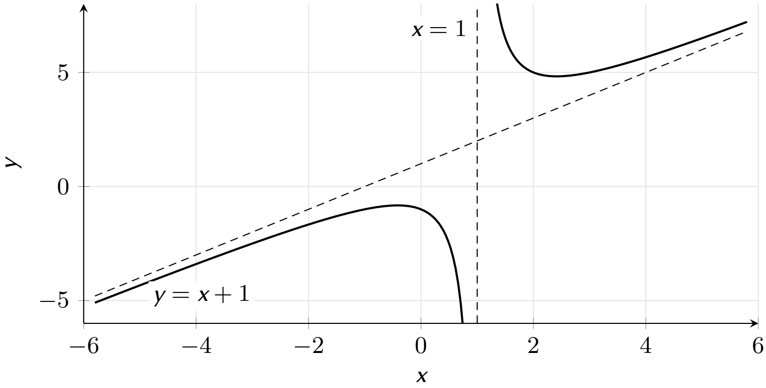
$$\lim_{x \to \infty} \frac{3x^2 - 5x + 2}{2x^2 + x - 1}.$$

Divide numerator and denominator by x^2:

$$\frac{3 - 5/x + 2/x^2}{2 + 1/x - 1/x^2} \to \frac{3}{2}.$$

Therefore the model has horizontal asymptote $y = 3/2$.

Illustration: Limits at infinity and asymptotic thinking



Asymptotes summarise behaviour near forbidden inputs and far from the origin.

Summary: Limit techniques for class use

Situation	Technique	Example
Continuous expression	Substitute directly	polynomial, square root on domain
$0/0$ with polynomial	Factor and cancel	$(x^2-4)/(x-2)$
$0/0$ with radicals	Rationalise	$(\sqrt{x+4}-2)/x$
Oscillating bounded factor	Squeeze theorem	$x^2\sin(1/x)$
Piecewise or absolute value	Check one-sided limits	$\lvert x\rvert/x$
Rational function at infinity	Divide by highest power	leading coefficient ratio
Vertical blow-up	One-sided infinite limits	$1/(x-a)$ or $1/(x-a)^2$

Continuity: no jump, no hole, no blow-up

Core explanation, worked reasoning, and application

Continuity: no jump, no hole, no blow-up

Continuity means that the function value matches the limiting trend.

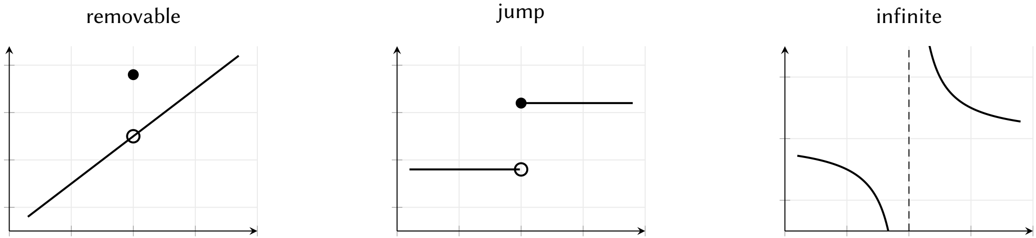
Concept: Continuity at a point

A function f is **continuous** at $x = a$ if all three conditions hold:

1. $f(a)$ is defined;
2. $\lim_{x\to a} f(x)$ exists;
3. $\lim_{x\to a} f(x) = f(a)$.

If any condition fails, the function is discontinuous at a.

Illustration: Continuity: no jump, no hole, no blow-up



Three common discontinuities: a hole, a jump, and a vertical blow-up.

Worked Example: Worked Example 7: Make a piecewise model continuous

Find k so that

$$f(x) = \begin{cases} kx^2 - 1, & x < 2, \\ 3x + k, & x \geq 2 \end{cases}$$

is continuous at $x = 2$.
Left-hand limit:

$$\lim_{x \to 2^-} f(x) = 4k - 1.$$

Right-hand value and limit:

$$\lim_{x \to 2^+} f(x) = f(2) = 6 + k.$$

Set them equal:

$$4k - 1 = 6 + k \quad \Rightarrow \quad 3k = 7 \quad \Rightarrow \quad k = \frac{7}{3}.$$

Model Critique: Modeling interpretation of continuity

Continuity is an assumption. A physical temperature curve is often continuous; a fare system, school grade boundary, or traffic-light rule may have jumps. Before applying a continuous model, ask whether the real system changes smoothly or by thresholds.

Existence theorems: IVT and EVT

Core explanation, worked reasoning, and application

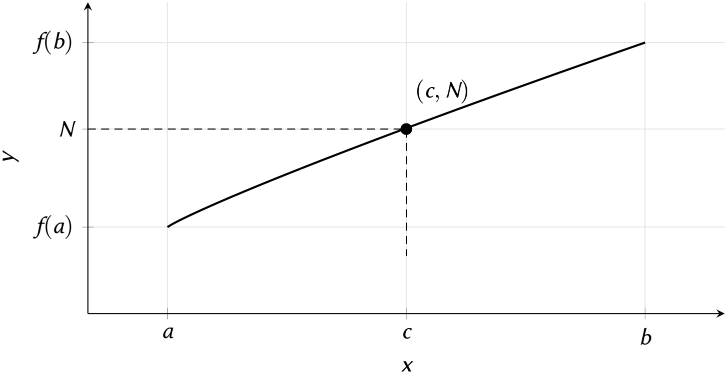
Existence theorems: IVT and EVT

The **Intermediate Value Theorem** and **Extreme Value Theorem** are important because they guarantee existence even when we cannot solve exactly.

Concept: Intermediate Value Theorem

If f is continuous on $[a, b]$ and N lies between $f(a)$ and $f(b)$, then there exists $c \in (a, b)$ such that $f(c) = N$.
In particular, if $f(a)$ and $f(b)$ have opposite signs, then f has a root in (a, b).

Illustration: Existence theorems: IVT and EVT



A continuous graph cannot move from $f(a)$ to $f(b)$ without taking every intermediate value.

Worked Example: Worked Example 8: Existence of a solution

Show that

$$x^3 - x - 1 = 0$$

has a solution between 1 and 2.
Let $f(x) = x^3 - x - 1$. Since f is a polynomial, it is continuous. Compute

$$f(1) = 1 - 1 - 1 = -1 < 0, \qquad f(2) = 8 - 2 - 1 = 5 > 0.$$

By the IVT, there exists $c \in (1, 2)$ such that $f(c) = 0$.

Concept: Extreme Value Theorem

If f is continuous on a closed bounded interval $[a, b]$, then f attains an absolute maximum and an absolute minimum on $[a, b]$.
For modeling, this theorem justifies the phrase "there is an optimal value" before we try to find it. The hypotheses matter: continuous function, closed interval, bounded interval.

Squeeze theorem and trigonometric limits

Core explanation, worked reasoning, and application

Squeeze theorem and trigonometric limits

The **Squeeze Theorem** handles functions trapped between two simpler functions.

Concept: Squeeze theorem

If

$$g(x) \leq f(x) \leq h(x)$$

near a, and

$$\lim_{x \to a} g(x) = \lim_{x \to a} h(x) = L,$$

then $\lim_{x \to a} f(x) = L$.

Worked Example: Worked Example 9: Oscillation controlled by a vanishing factor

Evaluate

$$\lim_{x \to 0} x^2 \sin\left(\frac{1}{x}\right).$$

Since $-1 \leq \sin(1/x) \leq 1$ for $x \neq 0$,

$$-x^2 \leq x^2 \sin\left(\frac{1}{x}\right) \leq x^2.$$

Both outer functions approach 0, so by the Squeeze Theorem,

$$\lim_{x \to 0} x^2 \sin\left(\frac{1}{x}\right) = 0.$$

Summary: Two standard trigonometric limits

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x\to 0} \frac{1-\cos x}{x} = 0.$$

These are foundational for derivatives of trigonometric functions. They also show why radians, not degrees, are the natural angle unit in calculus.

Worked Example: Worked Example 10: Using the standard trig limit

Evaluate

$$\lim_{x\to 0} \frac{\sin(3x)}{x}.$$

Rewrite:

$$\frac{\sin(3x)}{x} = 3 \cdot \frac{\sin(3x)}{3x}.$$

As $x \to 0$, also $3x \to 0$, so

$$\lim_{x\to 0} \frac{\sin(3x)}{x} = 3 \cdot 1 = 3.$$

Differentiated limit studio

Core explanation, worked reasoning, and application

Activity: Differentiated studio

Choose one route.

1. **Core route: algebraic limits.** Evaluate

$$\lim_{x\to -2}\frac{x^3+8}{x+2},\qquad \lim_{x\to 0}\frac{\sqrt{x+9}-3}{x}.$$

2. **Challenge route: one-sided and continuity.** For

$$f(x)=\begin{cases} ax+1, & x<1,\\ x^2+a, & x\geq 1,\end{cases}$$

   find a so that f is continuous at $x=1$.

3. **Extension route: asymptotic model.** A cost model is

$$C(n)=\frac{5n^2+30n}{2n^2+1}.$$

   Find the long-run cost and explain what the horizontal asymptote means.

Model Critique: Competition-style communication

When using limits in a modeling report, write in context.

Weak statement: "The limit is infinity."

Stronger statement: "As demand approaches full capacity from below, the waiting-time model $W(q) = 1/(1-q)$ grows without bound. This suggests that operating near $q = 1$ is unsafe: small increases in demand may produce very large waiting times. The conclusion depends on the model's assumption that service capacity remains fixed."

Summary: Lesson summary

Concept	Modeling meaning
Limit	Describes approaching behaviour near a point.
Hole	Formula may be undefined at one input while the trend remains clear.
One-sided limit	Detects thresholds, jumps, and directional behaviour.
Infinite limit	Signals blow-up or capacity failure.
Limit at infinity	Describes long-run behaviour and horizontal asymptotes.
Continuity	Assumes smooth change; may fail in threshold systems.
IVT	Guarantees existence of a solution between two signs or values.
EVT	Guarantees a maximum/minimum when continuity and closed interval hold.
Squeeze theorem	Controls a complicated function between simpler ones.

Checkpoint: Exit ticket

In three sentences, explain the difference between $f(a)$ and $\lim_{x\to a} f(x)$. Then give one modeling situation where the limit is useful even if $f(a)$ is missing.

Exercises: Core practice

1. Evaluate $\lim_{x \to 4} (5x - 3)$ using direct substitution.
2. Evaluate $\lim_{x \to 2} \frac{x^2 - x - 2}{x - 2}$.
3. Evaluate $\lim_{x \to 0} \frac{\sqrt{x+1} - 1}{x}$.
4. Evaluate $\lim_{x \to 0} \frac{\tan x}{x}$.
5. Find the one-sided limits of $f(x) = \lfloor x \rfloor$ at $x = 3$.
6. Classify the discontinuity of $f(x) = (x^2 - 9)/(x - 3)$ at $x = 3$.

Exercises: Modeling challenge

A school cafeteria uses

$$W(q) = \frac{4q}{1-q}, \qquad 0 \le q < 1,$$

to estimate average waiting time in minutes, where q is demand divided by service capacity.

1. Find $\lim_{q\to 0^+} W(q)$ and interpret.
2. Find $\lim_{q\to 1^-} W(q)$ and interpret.
3. Find the value of q at which the model predicts $W = 12$ minutes.
4. State one limitation of this model.
5. Write a short recommendation for cafeteria managers based on the limiting behaviour.

Exercises: Extension practice

1. Prove using the epsilon-delta definition that $\lim_{x \to 1}(3x + 2) = 5$.
2. Use the sequential criterion to show that $\lim_{x \to 0} \sin(1/x)$ does not exist.
3. Evaluate $\lim\limits_{x \to \infty} \left(\sqrt{x^2 + 2x + 3} - x \right)$.
4. Find all vertical and horizontal asymptotes of $f(x) = \dfrac{3x^2 - 2}{x^2 - 4}$.
5. Use the IVT to prove that $e^x = 3 - 2x$ has at least one real solution.

Summary: Optional appendix: L'Hopital's Rule preview

Some limits produce indeterminate forms such as $0/0$ or ∞/∞. After derivatives are introduced, one powerful tool is L'Hopital's Rule: under suitable conditions,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

This rule should not replace algebraic thinking. Factorisation, rationalisation, and standard trigonometric limits remain important because they reveal the structure of the expression.

Supplementary 5: Derivatives, Rates, and Local Linear Models

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this supplementary lesson, readers should be able to:

1. interpret the **derivative** as a local slope, instantaneous rate, and sensitivity coefficient;
2. compute simple derivatives from the **difference quotient**;
3. use the power, product, quotient, trigonometric, exponential, logarithmic, and **chain rule** correctly;
4. identify common reasons why a function is not differentiable;
5. use implicit, parametric, and related-rate differentiation in geometric or physical models;
6. use the **Mean Value Theorem** to connect average and instantaneous rates;
7. explain derivative calculations in a modeling report instead of presenting them as isolated algebra.

Derivative as slope, rate, and sensitivity

Core explanation, worked reasoning, and application

Derivative as slope, rate, and sensitivity

In modeling, a derivative is not just a calculus operation. It answers a practical question: *how fast does the output change when the input changes slightly?* If $y = f(x)$, then $f'(a)$ measures the local change in y per unit change in x near $x = a$.

Activity: Launch: average speed versus speed at one instant

A robot travels along a corridor. Its distance from the start after t seconds is modeled by

$$s(t) = 0.4t^2 + 1.2t \qquad (0 \leq t \leq 10),$$

where s is in metres.

1. Compute the average speed from $t = 3$ to $t = 5$.
2. Compute the average speed from $t = 4$ to $t = 4.1$.
3. What value should the speedometer display at exactly $t = 4$?

Discussion point: smaller time windows give more local information. The derivative is the limiting value of these average rates.

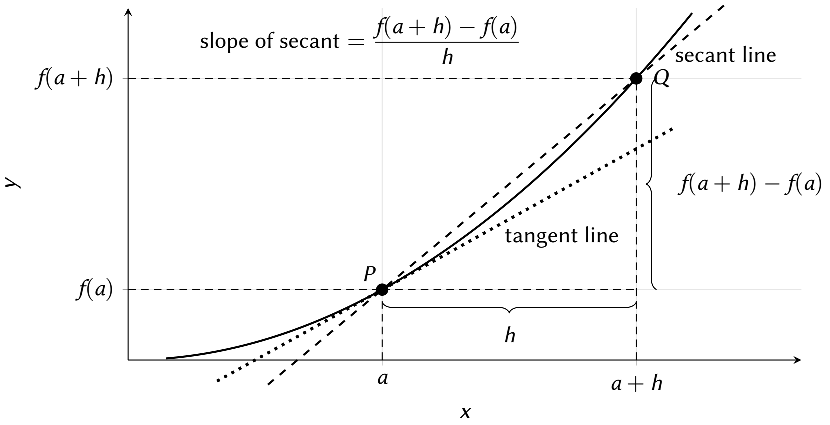
Concept: Core definition

For a function f, the **derivative at a** is

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h},$$

provided the limit exists. The fraction is the **difference quotient**. It is the slope of the secant line through $(a, f(a))$ and $(a+h, f(a+h))$.

Illustration: Derivative as slope, rate, and sensitivity



The derivative at a is the limit of the slopes of secant lines: $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$. As h becomes smaller, the secant line through P and Q approaches the tangent line at P.

Worked Example: Worked example 1: speed from a position model

For $s(t) = 0.4t^2 + 1.2t$, the instantaneous speed at $t = 4$ is

$$s'(4) = \lim_{h \to 0} \frac{s(4+h) - s(4)}{h}.$$

Compute

$$\begin{aligned} s(4+h) &= 0.4(4+h)^2 + 1.2(4+h) \\ &= 0.4(16 + 8h + h^2) + 4.8 + 1.2h \\ &= 11.2 + 4.4h + 0.4h^2, \end{aligned}$$

while $s(4) = 11.2$. Hence

$$\frac{s(4+h) - s(4)}{h} = 4.4 + 0.4h \to 4.4.$$

So the robot's speed at $t = 4$ is 4.4 m/s.

Checkpoint: Checkpoint 1

Explain the units of $s'(4)$. Why is it not measured in metres?

First-principles derivatives

Core explanation, worked reasoning, and application

First-principles derivatives

Before using rules, students should experience why the rules are true. This prevents derivatives from becoming a memorised table detached from modeling.

Worked Example: Worked example 2: derivative of x^2 from the definition

Let $f(x) = x^2$. Then

$$f'(x) = \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h\to 0} \frac{2xh + h^2}{h} = \lim_{h\to 0}(2x + h) = 2x.$$

At $x = 3$, the local slope is 6. This means that near $x = 3$, increasing x by 0.01 increases x^2 by about 0.06.

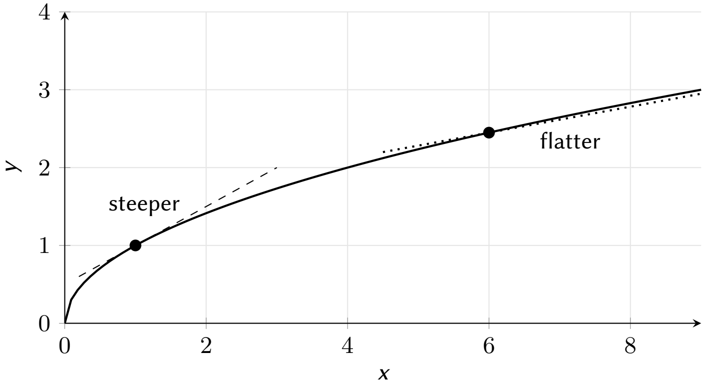
Worked Example: Worked example 3: derivative of $\sqrt{x}$

For $x > 0$,

$$\begin{aligned}\frac{d}{dx}\sqrt{x} &= \lim_{h\to 0}\frac{\sqrt{x+h}-\sqrt{x}}{h}\\ &= \lim_{h\to 0}\frac{(x+h)-x}{h(\sqrt{x+h}+\sqrt{x})}\\ &= \lim_{h\to 0}\frac{1}{\sqrt{x+h}+\sqrt{x}} = \frac{1}{2\sqrt{x}}.\end{aligned}$$

The derivative is larger near $x = 0$, so the square-root graph is steep near the origin and gradually flattens.

Illustration: First-principles derivatives



The derivative of $\sqrt{x}$ decreases as x increases. This is a useful visual interpretation of $1/(2\sqrt{x})$.

Model Critique: Model critique: derivative as local linear model

A derivative gives a local approximation:

$$f(a+\Delta x) \approx f(a) + f'(a)\Delta x.$$

This is powerful, but it is local. It may be poor when Δx is large or when the graph bends sharply. In a modeling report, always state the interval over which the local approximation is being used.

Differentiation rules for efficient calculation

Core explanation, worked reasoning, and application

Concept: Core derivative rules

If f and g are differentiable, then

$$
\begin{aligned}
&\frac{d}{dx}(c) = 0, && \frac{d}{dx}(x^n) = nx^{n-1}, && \frac{d}{dx}[cf] = cf', \\
&\frac{d}{dx}(f+g) = f' + g', && \frac{d}{dx}(fg) = f'g + fg', && \frac{d}{dx}\left(\frac{f}{g}\right) = \frac{f'g - fg'}{g^2}.
\end{aligned}
$$

The product rule is not $(fg)' = f'g'$. The quotient rule is not the derivative of numerator divided by derivative of denominator.

Worked Example: Worked example 4: product and quotient rules

Differentiate

$$F(x) = x^2 \sin x + \frac{x^2+1}{x-1}.$$

For the product term,

$$\frac{d}{dx}(x^2 \sin x) = 2x \sin x + x^2 \cos x.$$

For the quotient term,

$$\frac{d}{dx}\left(\frac{x^2+1}{x-1}\right) = \frac{2x(x-1)-(x^2+1)}{(x-1)^2} = \frac{x^2-2x-1}{(x-1)^2}.$$

Therefore

$$F'(x) = 2x \sin x + x^2 \cos x + \frac{x^2-2x-1}{(x-1)^2}.$$

Worked Example: Worked example 5: trigonometric and exponential derivatives

For a seasonal demand model

$$D(t) = 120 + 30\sin\left(\frac{2\pi t}{12}\right) + 5e^{0.03t},$$

where t is measured in months, compute the instantaneous rate of change:

$$D'(t) = 30\cos\left(\frac{2\pi t}{12}\right)\cdot\frac{2\pi}{12} + 5e^{0.03t}\cdot 0.03.$$

Thus

$$D'(t) = 5\pi\cos\left(\frac{\pi t}{6}\right) + 0.15e^{0.03t}.$$

The first term describes seasonal increase/decrease; the second term describes long-term growth.

Checkpoint: Checkpoint 2

Why is the derivative of $\sin(\pi t/6)$ not simply $\cos(\pi t/6)$? Name the rule being used.

The chain rule as a modeling pipeline

Core explanation, worked reasoning, and application

The chain rule as a modeling pipeline

The **chain rule** is essential for modeling because many models are pipelines: an input affects an intermediate quantity, which affects the output.

Concept: Chain rule

If $y = f(u)$ and $u = g(x)$, then

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

In function notation,

$$\frac{d}{dx} f(g(x)) = f'(g(x))g'(x).$$

Illustration: The chain rule as a modeling pipeline

$$
\begin{array}{c} \text{input} \\ x \end{array}
\xrightarrow{\frac{du}{dx} = g'(x)}
\begin{array}{c} \text{inner function} \\ u = g(x) \end{array}
\xrightarrow{\frac{dy}{du} = f'(u)}
\begin{array}{c} \text{outer function} \\ y = f(u) = f(g(x)) \end{array}
$$

Overall rate from x to y

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = f'(g(x))\,g'(x)$$

The chain rule for a composite function $y = f(g(x))$. The input x first changes the intermediate variable $u = g(x)$, and then u changes the final output $y = f(u)$. Thus the overall rate is the product of the two rates.

Worked Example: Worked example 6: pollution concentration through a transformation

A concentration index is modeled by

$$I(t) = \sqrt{1 + 0.4C(t)}, \qquad C(t) = 20e^{-0.08t}.$$

Find $I'(t)$ and interpret the sign.

$$I'(t) = \frac{1}{2\sqrt{1 + 0.4C(t)}} \cdot 0.4C'(t).$$

Since $C'(t) = -1.6e^{-0.08t}$,

$$I'(t) = \frac{0.2(-1.6e^{-0.08t})}{\sqrt{1 + 8e^{-0.08t}}} < 0.$$

The index is decreasing. The derivative becomes closer to zero over time because the exponential decay slows in absolute size.

Worked Example: Worked example 7: logarithmic differentiation

Differentiate $y = x^x$ for $x > 0$.

$$\ln y = x \ln x.$$

Differentiate both sides:

$$\frac{y'}{y} = \ln x + 1.$$

Therefore

$$y' = x^x(\ln x + 1).$$

This method is useful when both the base and exponent vary.

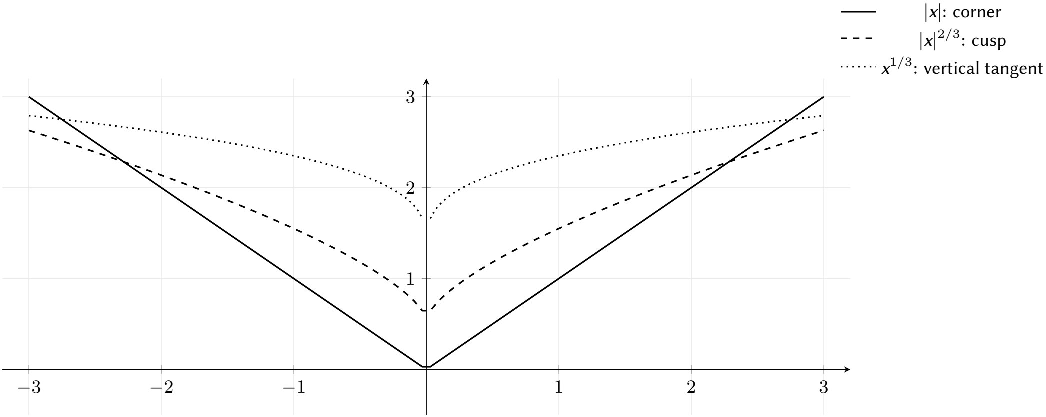
When derivatives do not exist

Core explanation, worked reasoning, and application

When derivatives do not exist

A derivative may fail to exist even when a graph looks simple. This matters in modeling: a non-differentiable point may represent a switch, threshold, collision, capacity boundary, or change of policy.

Illustration: When derivatives do not exist



Three common non-differentiable behaviours. Each can have a modeling interpretation.

Checkpoint: Checkpoint 3

For $f(x) = |x|$, compute the right-hand derivative and left-hand derivative at 0. Why does the derivative not exist there?

Model Critique: Model critique: smoothing versus preserving a threshold

In applications, replacing a kink by a smooth curve may make calculus easier, but it may erase a real threshold. For example, a penalty fee may start exactly when demand exceeds capacity. A good modeler should ask: is the corner a mathematical inconvenience, or is it the phenomenon we need to preserve?

Implicit, parametric, and related-rate differentiation

Core explanation, worked reasoning, and application

Implicit, parametric, and related-rate differentiation

Not all models are naturally written as $y = f(x)$. Circles, ellipses, supply-demand equilibria, and geometric constraints often appear as equations involving several variables.

Worked Example: Worked example 8: implicit differentiation on a circle

For a circle

$$x^2 + y^2 = 25,$$

we differentiate both sides with respect to x:

$$2x + 2y\frac{dy}{dx} = 0,$$

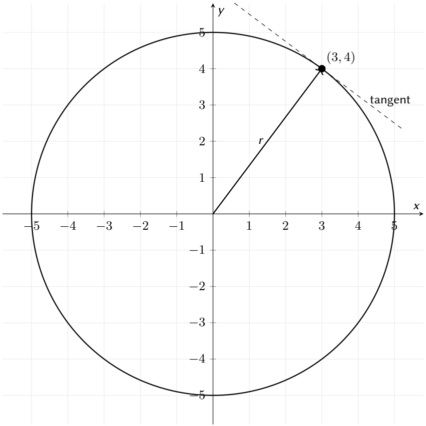
so

$$\frac{dy}{dx} = -\frac{x}{y}.$$

At $(3, 4)$, the tangent slope is $-3/4$, so the tangent line is

$$y - 4 = -\frac{3}{4}(x - 3).$$

Illustration: Implicit, parametric, and related-rate differentiation



Implicit differentiation finds tangent slopes even when the curve is not a function of x globally.

Concept: Parametric derivative

If a curve is given by $x = g(t)$ and $y = h(t)$, then

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt}, \qquad dx/dt \neq 0.$$

This is useful for motion models because t is the natural independent variable.

Worked Example: Worked example 9: related rates for a sliding ladder

A 10 m ladder leans against a wall. Its base moves away at 0.5 m/s. How fast is the top moving down when the base is 6 m from the wall?
Let x be the distance of the base from the wall and y the height of the top. Then

$$x^2 + y^2 = 100.$$

Differentiate with respect to time:

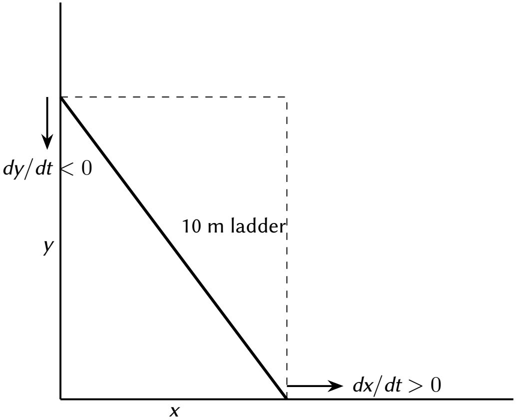
$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0.$$

When $x = 6$, $y = 8$. Hence

$$\frac{dy}{dt} = -\frac{x}{y}\frac{dx}{dt} = -\frac{6}{8}(0.5) = -0.375 \text{ m/s}.$$

The negative sign means the height is decreasing.

Illustration: Implicit, parametric, and related-rate differentiation



Related rates translate a geometric constraint into a relationship between rates.

Mean Value Theorem and local linear thinking

Core explanation, worked reasoning, and application

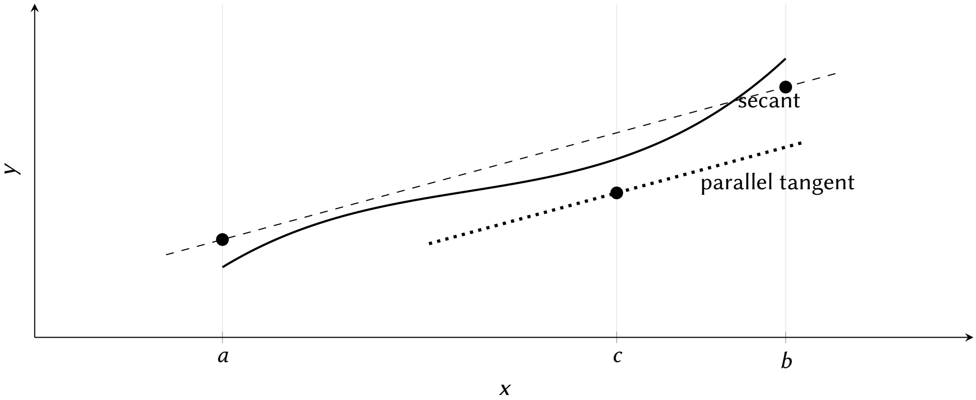
Concept: Mean Value Theorem

If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

The instantaneous rate somewhere equals the average rate over the interval.

Illustration: Mean Value Theorem and local linear thinking



The Mean Value Theorem links an average rate to at least one instantaneous rate.

Worked Example: Worked example 10: interpreting an average speed claim

A car travels 120 km in 1.5 hours. Assuming the position function is continuous and differentiable, the Mean Value Theorem says that at some instant the car's speed was exactly

$$\frac{120}{1.5} = 80 \text{ km/h}.$$

This does not say the speed was always 80 km/h. It says an instantaneous value equal to the average must have occurred at least once.

Summary: What students should take away

- A derivative is a local rate, slope, and sensitivity.
- The difference quotient is the bridge from average rate to instantaneous rate.
- Rules speed up calculation, but the chain rule is the main rule for composite models.
- Non-differentiability is often meaningful, not merely a technical failure.
- Implicit and related-rate methods are essential when the model is constrained geometrically.
- The MVT connects local and average behaviour, which is why derivatives support real-world interpretation.

Optional extension: L’Hôpital’s Rule

Core explanation, worked reasoning, and application

Concept: Reserve material

L'Hôpital's Rule belongs after students understand derivatives. If a limit has indeterminate form $0/0$ or ∞/∞, and the relevant derivative quotient limit exists, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

It should not be used before checking the indeterminate form.

Worked Example: Optional example: exponential dominates polynomial

Evaluate

$$\lim_{x \to \infty} \frac{x^2}{e^x}.$$

The form is ∞/∞. Applying L'Hôpital twice gives

$$\lim_{x \to \infty} \frac{2x}{e^x} o \lim_{x \to \infty} \frac{2}{e^x} = 0.$$

Thus exponential growth eventually dominates quadratic growth.

Exercises

Core explanation, worked reasoning, and application

Exercises: Core practice

1. Use the definition of derivative to find the derivative of $f(x) = 3x^2 - 2x$.
2. Differentiate:
$$f(x) = 5x^4 - 3x^{-2} + 7\sqrt{x} - 9.$$
3. Differentiate $g(x) = (x^2 + 1)(\sin x + x)$.
4. Differentiate $h(x) = \dfrac{x^2 + 3x}{x - 2}$.
5. Differentiate $p(x) = \sin(3x^2 + 1)$ and explain which part is the inner function.
6. Find the tangent line to $x^2 + y^2 = 13$ at $(2, 3)$.

Exercises: Modeling practice

1. A queue waiting-time model is $W(\rho) = \dfrac{4\rho}{1-\rho}$ for $0 < \rho < 1$. Compute $W'(\rho)$ and explain what happens as $\rho \to 1^-$. Why is this important for capacity planning?
2. A pollutant concentration follows $C(t) = 60e^{-0.12t} + 8$. Compute $C'(t)$ and determine when the magnitude of the rate first falls below 1 unit/day.
3. A circular infection front has radius $r(t)$ and area $A(t) = \pi r(t)^2$. If $dr/dt = 0.3$ km/day, find dA/dt when $r = 5$ km.
4. A seasonal sales model is $S(t) = 200 + 40\cos(\pi t/6)$. Find when sales are increasing fastest and interpret your answer.

Exercises: Challenge and extension

1. Show that if $f'(x) = 0$ for all x in an interval, then f is constant on that interval. State clearly where the Mean Value Theorem is used.
2. Differentiate $y = x^x \sin x$ for $x > 0$.
3. For the parametric curve $x = t^2 + 1$, $y = t^3 - t$, find dy/dx and the slope at $t = 2$.
4. Use L'Hôpital's Rule to evaluate
$$\lim_{x \to 0} \frac{e^x - 1 - x}{x^2}.$$
5. Write a short modeling paragraph explaining why a derivative-based approximation may fail near a kink or threshold.

Supplementary 6: Applications of Derivatives and Optimization

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this two-hour supplementary lesson, students should be able to:

- use the sign of f' to determine **increasing/decreasing behaviour**;
- locate and classify **critical points**;
- use f'' to discuss **concavity** and **inflection points**;
- sketch a curve from domain, asymptotes, derivative sign, and concavity;
- formulate and solve a one-variable **optimization** model;
- communicate derivative-based conclusions as modeling evidence, not just computation.

Derivative evidence for optimization and model design

Core explanation, worked reasoning, and application

Derivative evidence for optimization and model design

In a modeling report, a derivative is rarely an end in itself. It is evidence for a statement such as:

> *"The waiting time increases slowly at first, then rises sharply near capacity."*

The first derivative tells us the *direction and rate of change*; the second derivative tells us whether the change is accelerating or slowing down. These ideas turn calculus into a practical language for decisions: when to stop production, where a maximum occurs, whether a policy has diminishing returns, and whether a model becomes unstable near a boundary.

Activity: Launch activity: reading a derivative without calculating

A robot delivery system has travel-time function $T(q)$, where q is the number of orders per hour. Suppose that at $q = 40$, $T'(40) > 0$ and $T''(40) > 0$.

What can you say in words? A good answer is not "the derivative is positive." It should say:

> *At 40 orders per hour, increasing demand will increase travel time, and the increase is itself accelerating. The system is moving toward congestion.*

Extreme Values and Critical Points

Core explanation, worked reasoning, and application

Concept: Extreme values

Let f be defined on a domain D.

- $f(c)$ is an **absolute maximum** if $f(c) \geq f(x)$ for all $x \in D$.
- $f(c)$ is an **absolute minimum** if $f(c) \leq f(x)$ for all $x \in D$.
- $f(c)$ is a **local maximum** or **local minimum** if the comparison only needs to hold near c.

A point c is a **critical point** if $f'(c) = 0$ or $f'(c)$ does not exist.

Concept: Closed-interval method

If f is continuous on $[a, b]$, then the **Extreme Value Theorem** guarantees an absolute maximum and an absolute minimum. To find them:

1. find all critical points in (a, b);
2. evaluate f at the critical points and endpoints;
3. compare the values.

Worked Example: Worked example 1: absolute extrema on a closed interval

Find the absolute maximum and minimum of

$$f(x) = 2x^3 - 3x^2 - 12x + 5 \qquad \text{on } [-2, 3].$$

First,

$$f'(x) = 6x^2 - 6x - 12 = 6(x-2)(x+1).$$

Critical points in $(-2, 3)$ are $x = -1$ and $x = 2$. Now compare the candidate values:

x	-2	-1	2	3
$f(x)$	1	12	-15	-4

Therefore the absolute maximum is 12 at $x = -1$, and the absolute minimum is -15 at $x = 2$.

Checkpoint: Checkpoint

Why do endpoints matter? Because an absolute maximum or minimum on a closed interval may occur at the boundary even when $f'(x) \neq 0$ there.

First Derivative Test: Direction of Change

Core explanation, worked reasoning, and application

Concept: Monotonicity from the sign of f'

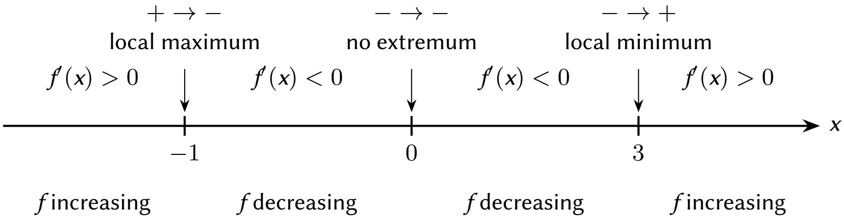
If $f'(x) > 0$ on an interval, then f is **increasing** there. If $f'(x) < 0$, then f is **decreasing**. This follows from the **Mean Value Theorem**.

Concept: First derivative test

Let c be a critical point where f is continuous.

- If f' changes from $+$ to $-$ at c, then $f(c)$ is a local maximum.
- If f' changes from $-$ to $+$ at c, then $f(c)$ is a local minimum.
- If f' does not change sign, then c is not a local extremum.

Illustration: First Derivative Test: Direction of Change



A first-derivative sign chart. Since $f'(x) > 0$ means f is increasing and $f'(x) < 0$ means f is decreasing, a change from $+$ to $-$ gives a local maximum, while a change from $-$ to $+$ gives a local minimum. At $x = 0$, the sign of f' does not change, so there is no extremum.

Worked Example: Worked example 2: first derivative test

Analyse local extrema of

$$f(x) = x^4 - 4x^3.$$

We compute

$$f'(x) = 4x^3 - 12x^2 = 4x^2(x-3).$$

Critical points are $x = 0$ and $x = 3$. Since $4x^2 \geq 0$, the sign of $f'(x)$ away from 0 is determined by $x - 3$:

x	$(-\infty, 0)$	0	$(0, 3)$	3	$(3, \infty)$
$f'(x)$	$-$	0	$-$	0	$+$
f	$\searrow$		$\searrow$		$\nearrow$

At $x = 0$, f' does not change sign, so there is no extremum. At $x = 3$, f' changes from negative to positive, so f has a local minimum, with $f(3) = -27$.

Concavity and the Second Derivative

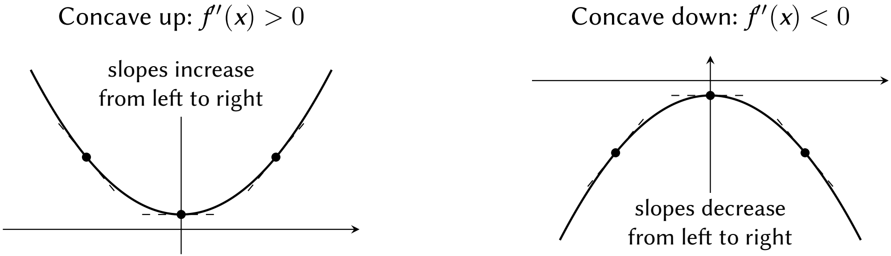
Core explanation, worked reasoning, and application

Concept: Concavity and inflection

The second derivative measures how the slope changes.

- If $f''(x) > 0$, then f' is increasing and the graph is **concave up**.
- If $f''(x) < 0$, then f' is decreasing and the graph is **concave down**.
- An **inflection point** occurs where concavity changes.

Illustration: Concavity and the Second Derivative



Concavity describes how tangent slopes change. For a concave-up graph, $f''(x) > 0$, the tangent slopes increase from left to right. For a concave-down graph, $f''(x) < 0$, the tangent slopes decrease from left to right.

Concept: Second derivative test

Suppose $f'(c) = 0$ and f'' is continuous near c.

- If $f''(c) > 0$, then f has a local minimum at c.
- If $f''(c) < 0$, then f has a local maximum at c.
- If $f''(c) = 0$, the test is inconclusive. Use the first derivative test.

Worked Example: Worked example 3: concavity and inflection

Find the inflection points of

$$f(x) = x^4 - 6x^2 + 1.$$

We have

$$f'(x) = 4x^3 - 12x, \qquad f''(x) = 12x^2 - 12 = 12(x-1)(x+1).$$

The possible inflection points are $x = -1$ and $x = 1$. Check the sign of f'':

x	$(-\infty, -1)$	$(-1, 1)$	$(1, \infty)$
$f''(x)$	$+$	$-$	$+$
concavity	up	down	up

Concavity changes at both $x = -1$ and $x = 1$. Since $f(-1) = f(1) = -4$, the inflection points are $(-1, -4)$ and $(1, -4)$.

Curve Sketching as Model Diagnosis

Core explanation, worked reasoning, and application

Curve Sketching as Model Diagnosis

A curve sketch is not only a drawing exercise. In modeling, it diagnoses where a formula is valid, where it has unrealistic blow-up, where the response saturates, and where the model predicts a turning point.

Concept: Curve sketching checklist

For $y = f(x)$, analyse:

1. domain and excluded values;
2. intercepts and symmetry;
3. vertical, horizontal, or oblique asymptotes;
4. sign chart for f' and local extrema;
5. sign chart for f'' and inflection points;
6. interpretation in the original context.

Worked Example: Worked example 4: rational curve sketch

Sketch and interpret

$$f(x) = \frac{x^2}{x^2 - 1}.$$

The domain is $x \neq \pm 1$. The function is even. There are vertical asymptotes at $x = -1$ and $x = 1$, and

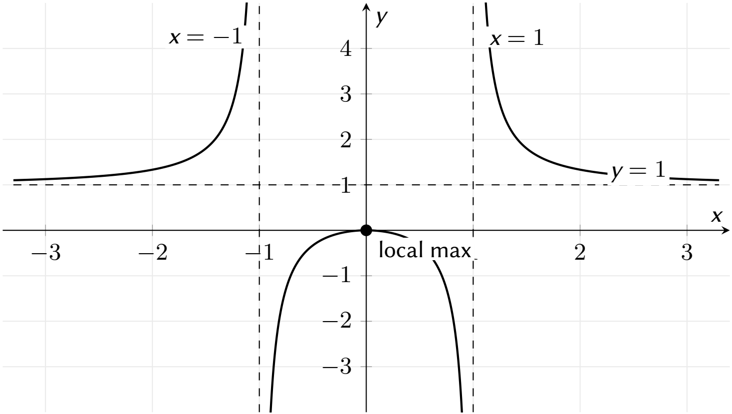
$$\lim_{x \to \pm\infty} \frac{x^2}{x^2 - 1} = 1,$$

so $y = 1$ is a horizontal asymptote. Also,

$$f'(x) = \frac{-2x}{(x^2 - 1)^2},$$

so f increases for $x < 0$ and decreases for $x > 0$ on each domain interval. The point $(0, 0)$ is a local maximum on the middle branch.

Worked Example: Worked example 4: rational curve sketch



A clean curve sketch for $f(x) = x^2/(x^2 - 1)$. Labels are placed outside dense regions and boxed with white background to avoid overlap with the branches and asymptotes.

Worked Example: Worked example 4: rational curve sketch

Model interpretation. If this formula represented a response ratio, then $x = \pm 1$ would be dangerous thresholds: the model predicts unbounded behaviour there. The horizontal asymptote $y = 1$ suggests the response stabilises far from the thresholds.

Model Critique: Model critique: derivative evidence

A derivative-based sketch is only as meaningful as the formula and domain. When using calculus in a modeling report, state:

- the domain used in the real problem;
- whether excluded points are physical boundaries or mathematical artefacts;
- whether maxima/minima occur inside the feasible range or at a boundary;
- whether a vertical asymptote indicates real danger, capacity failure, or a model breakdown.

Applied Optimization

Core explanation, worked reasoning, and application

Concept: Optimization workflow

A good optimization solution has four layers:

1. define variables with units;
2. write the objective and constraints;
3. reduce to a one-variable function on a realistic domain;
4. verify the optimum and interpret it in context.

Skipping the domain is one of the most common modeling mistakes.

Worked Example: Worked example 5: fencing beside a river

A farmer has 200 m of fencing and wants to enclose a rectangular field beside a straight river. No fence is needed along the river. Find the dimensions that maximize the area.
Let x be the side parallel to the river and y the width. Then

$$x + 2y = 200, \qquad A = xy.$$

Using $x = 200 - 2y$,

$$A(y) = (200 - 2y)y = 200y - 2y^2, \qquad 0 < y < 100.$$

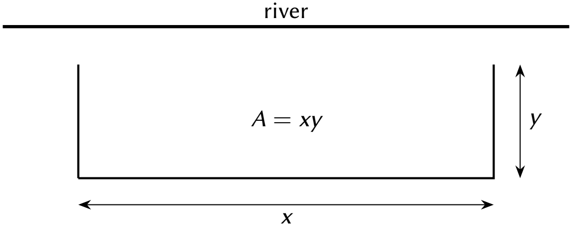
Now

$$A'(y) = 200 - 4y,$$

so $A'(y) = 0$ gives $y = 50$. Since $A''(y) = -4 < 0$, this gives a maximum. Therefore

$$y = 50 \text{ m}, \qquad x = 100 \text{ m}, \qquad A_{\max} = 5000 \text{ m}^2.$$

Worked Example: Worked example 5: fencing beside a river



$$\text{fencing used} = x + 2y$$

A rectangular field is built against a river, so only three sides need fencing. Thus $A = xy$ and the fencing constraint is $x + 2y = 200$.

Worked Example: Worked example 6: profit-maximizing production

A manufacturer sells q units per week. Cost is

$$C(q) = 0.01q^2 + 10q + 400,$$

and the price-demand equation is

$$p(q) = 50 - 0.02q.$$

Find the weekly production level maximizing profit.
Revenue is

$$R(q) = qp(q) = 50q - 0.02q^2.$$

Profit is

$$P(q) = R(q) - C(q) = (50q - 0.02q^2) - (0.01q^2 + 10q + 400) = 40q - 0.03q^2 - 400.$$

Then

$$P'(q) = 40 - 0.06q.$$

Set $P'(q) = 0$:

$$q = \frac{40}{0.06} \approx 666.7.$$

Since $P''(q) = -0.06 < 0$, this is a maximum for the continuous model. In practice, choose either $q = 666$ or $q = 667$ and compare integer profit.

Checkpoint: Checkpoint

Why might a continuous optimum be inappropriate for production quantity? Because real production may require an integer number of units, batch sizes, machine capacity, or demand uncertainty. The calculus result is a guide, not automatically the final decision.

Local Approximation and Newton's Method

Core explanation, worked reasoning, and application

Local Approximation and Newton's Method

For a differentiable function, the tangent line gives a local linear approximation:

$$f(x) \approx f(a) + f'(a)(x - a).$$

This is useful when exact calculation is difficult, or when a model needs a quick sensitivity estimate.

Worked Example: Worked example 7: local linear approximation

Approximate $\sqrt{4.1}$ using $f(x) = \sqrt{x}$ near $a = 4$.

$$f(4) = 2, \qquad f'(x) = \frac{1}{2\sqrt{x}}, \qquad f'(4) = \frac{1}{4}.$$

Therefore

$$\sqrt{x} \approx 2 + \frac{1}{4}(x - 4).$$

At $x = 4.1$,

$$\sqrt{4.1} \approx 2 + \frac{1}{4}(0.1) = 2.025.$$

This is close to the actual value $2.0248\ldots$.

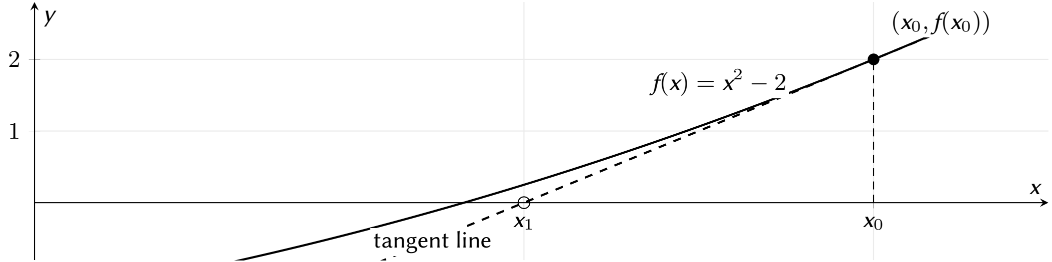
Concept: Newton's method

To solve $f(x) = 0$, start with a guess x_0 and iterate

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Geometrically, x_{n+1} is the x-intercept of the tangent line at x_n.

Illustration: Local Approximation and Newton's Method



One step of Newton's method for $f(x) = x^2 - 2$. Starting from x_0, the tangent line meets the x-axis at the next approximation x_1.

Worked Example: Worked example 8: approximating $\sqrt{2}$

Solve $f(x) = x^2 - 2 = 0$ with $x_0 = 1$.

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

n	x_n
0	1.000000
1	1.500000
2	1.416667
3	1.414216
4	1.414214

Newton's method converges very quickly here because the initial guess is reasonable and the root is simple.

Differentiated Optimization Studio

Core explanation, worked reasoning, and application

Activity: 15-minute studio

Choose one route.
Core route. A box with square base and open top must have volume $32\ \mathrm{m}^3$. Find the dimensions minimizing surface area.
Challenge route. Find the largest rectangle that can be inscribed in the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Extension route. Use Newton's method to approximate the positive solution of $\cos x = x$ starting from $x_0 = 0.7$. Explain why your iteration is credible.

Model Critique: How to write the modeling conclusion

A strong derivative-based conclusion should contain:

1. the objective being optimized;
2. the feasible domain;
3. the critical point or boundary comparison;
4. the verification method;
5. the practical interpretation and limitation.

For example:

> *Within the continuous production model, the profit function is concave because $P''(q) < 0$, so the critical point gives the global maximum. The optimal continuous production level is about 667 units per week. In an operational plan, this should be rounded only after checking integer profit and capacity constraints.*

Summary

Core explanation, worked reasoning, and application

Summary: Key takeaways

Derivative information	Modeling interpretation
$f'(x) > 0$	output increases as input increases.
$f'(x) < 0$	output decreases as input increases.
f' changes $+ \to -$	local maximum.
f' changes $- \to +$	local minimum.
$f''(x) > 0$	slope is increasing; response accelerates; concave up.
$f''(x) < 0$	slope is decreasing; diminishing returns or concave down.
Vertical asymptote	possible capacity limit, singularity, or model breakdown.
Closed interval comparison	absolute optimum may occur at a boundary.
Newton iteration	tangent-line method for solving equations numerically.

Exercises

Core explanation, worked reasoning, and application

Exercises: Core practice

1. Find the absolute maximum and minimum of $f(x) = x^3 - 3x^2 + 1$ on $[-1/2, 4]$.
2. For $f(x) = 2x^3 + 3x^2 - 36x + 5$, find critical points, intervals of increase/decrease, and local extrema.
3. Find all inflection points of $f(x) = x^5 - 5x^4 + 5x^3$.
4. Sketch $f(x) = \dfrac{x^2 - 4x + 3}{x - 2}$ using the curve-sketching checklist.
5. A manufacturer has cost $C(q) = 0.01q^2 + 10q + 400$ and price $p = 50 - 0.02q$. Rework the profit example and compare the integer choices $q = 666$ and $q = 667$.

Exercises: Modeling challenge

A cafeteria counter can serve at most 60 students per minute. A simple congestion model for average waiting time is

$$W(q) = \frac{q}{60 - q}, \qquad 0 < q < 60,$$

where q is arrival rate.

1. Find $W'(q)$ and $W''(q)$.
2. Interpret the signs of W' and W''.
3. What happens as $q \to 60^-$?
4. Write a short recommendation explaining why operating near capacity is risky.

Exercises: Extension practice

1. Use Newton's method to find the positive root of $x^3 + 2x - 5 = 0$ starting from $x_0 = 1$.
2. Find the Maclaurin polynomial $P_4(x)$ for $\cos x$ and use it to approximate $\cos(0.2)$.
3. Find the dimensions of the largest rectangle inscribed under the semicircle $x^2 + y^2 = R^2$, $y \geq 0$.

Supplementary 7: Definite Integrals, Accumulation, and Numerical Integration

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this two-hour supplementary lesson, students should be able to:

- interpret a **definite integral** as signed area, accumulated change, and average value;
- construct and read a **Riemann sum** using left, right, and midpoint rectangles;
- use the **Fundamental Theorem of Calculus** to evaluate basic definite integrals;
- apply **substitution** and **integration by parts** in simple modeling cases;
- recognise when an integral is **improper** and whether it converges;
- compare rectangle, trapezoidal, Simpson, and Monte Carlo thinking as numerical integration tools.

From area to accumulated change

Core explanation, worked reasoning, and application

From area to accumulated change

A definite integral is not just a way to compute geometric area. In modeling it often represents an accumulated quantity: total water entering a tank, total cost over time, average inventory, probability mass, or total pollutant load. The same notation

$$\int_a^b f(x)\,dx$$

may mean area, total change, or average effect depending on what f represents.

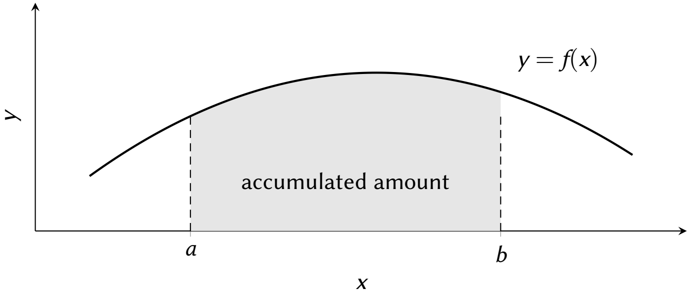
Activity: Launch activity: what does the integral measure?

Suppose $r(t)$ is the arrival rate of students at a cafeteria, measured in students per minute. What does

$$\int_0^{20} r(t)\, dt$$

represent? It is not a rate. It is the total number of students who arrive during the first 20 minutes.

Illustration: From area to accumulated change



The definite integral accumulates the vertical quantity $f(x)$ across an interval. If f is a rate, the integral is total change. If f is a density, the integral is total mass.

Riemann sums: building the integral

Core explanation, worked reasoning, and application

Concept: Partition and Riemann sum

A **partition** of $[a, b]$ is a list

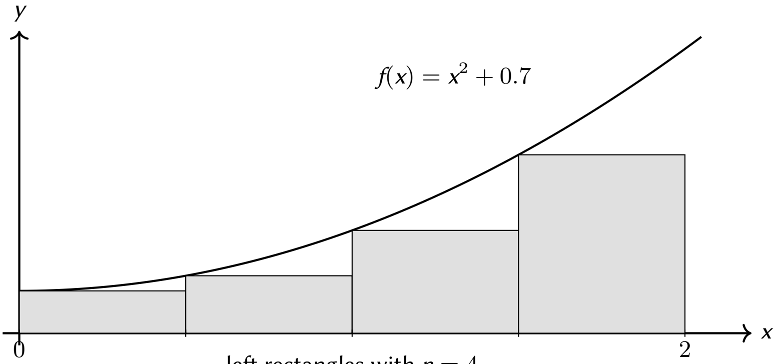
$$a = x_0 < x_1 < \cdots < x_n = b.$$

The width of the k-th subinterval is $\Delta x_k = x_k - x_{k-1}$. If x_k^* is a sample point in the subinterval, the corresponding **Riemann sum** is

$$\sum_{k=1}^{n} f(x_k^*)\Delta x_k.$$

For a regular partition, $\Delta x = (b - a)/n$. Common choices are left endpoint, right endpoint, and midpoint.

Illustration: Riemann sums: building the integral



left rectangles with $n = 4$

A left Riemann sum uses the height at the left endpoint of each subinterval. For an increasing function, it underestimates the area.

Worked Example: Worked example 1: a left Riemann sum

Approximate

$$\int_0^2 (x^2 + 1)\, dx$$

using a left Riemann sum with $n = 4$ equal subintervals.
Here $\Delta x = 0.5$, and the left sample points are $0, 0.5, 1, 1.5$. Thus

$$S_L = 0.5\,[f(0) + f(0.5) + f(1) + f(1.5)] = 0.5(1 + 1.25 + 2 + 3.25) = 3.75.$$

The exact value will later be $14/3 \approx 4.667$, so the left sum underestimates.

Concept: The definite integral

If the Riemann sums approach a unique number as the mesh of the partition tends to zero, we define

$$\int_a^b f(x)\,dx = \lim_{\|P\|\to 0} \sum_{k=1}^{n} f(x_k^*)\Delta x_k.$$

When $f \geq 0$, this is area. When f changes sign, it is **signed area**: area above the x-axis contributes positively, and area below contributes negatively.

Checkpoint: Checkpoint

For an increasing positive function on $[a, b]$, which is usually larger: the left sum or the right sum? Explain using rectangles, not formulas.

The Fundamental Theorem of Calculus

Core explanation, worked reasoning, and application

The Fundamental Theorem of Calculus

The Fundamental Theorem of Calculus connects the two ideas from Supplementary 5 and this lesson: derivative and integral.

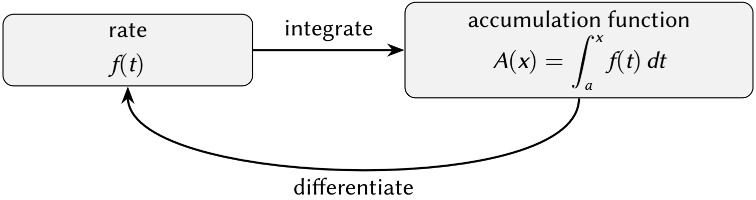
Concept: Fundamental Theorem of Calculus

Let f be continuous on $[a, b]$.

- **Part I:** If $A(x) = \int_a^x f(t)\, dt$, then $A'(x) = f(x)$.
- **Part II:** If $F'(x) = f(x)$, then

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

The first part says accumulation has rate $f(x)$; the second part says definite integrals can be evaluated by antiderivatives.



$$\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x)$$

FTC Part I says that integration and differentiation undo each other in the accumulation-function setting.

Worked Example: Worked example 2: evaluating by the FTC

Evaluate

$$\int_0^2 (x^2 + 1)\, dx.$$

An antiderivative is $F(x) = x^3/3 + x$. Hence

$$\int_0^2 (x^2 + 1)\, dx = \left[\frac{x^3}{3} + x\right]_0^2 = \frac{8}{3} + 2 = \frac{14}{3}.$$

This explains why the $n = 4$ left sum 3.75 was only an approximation.

Worked Example: Worked example 3: differentiating an integral

Compute

$$\frac{d}{dx}\left(\int_{2}^{x^2} \sin(t^2)\,dt\right).$$

Let $u = x^2$. By FTC Part I and the chain rule,

$$\frac{d}{dx}\int_{2}^{u} \sin(t^2)\,dt = \sin(u^2)\frac{du}{dx} = 2x\sin(x^4).$$

The integrand need not have an elementary antiderivative; the derivative is still easy.

Substitution and integration by parts

Core explanation, worked reasoning, and application

Concept: Substitution as reverse chain rule

If $u = g(x)$ and $du = g'(x)\,dx$, then

$$\int f(g(x))g'(x)\,dx = \int f(u)\,du.$$

For a definite integral, change the limits: $x = a \mapsto u = g(a)$ and $x = b \mapsto u = g(b)$.

Worked Example: Worked example 4: substitution with definite limits

Evaluate

$$\int_0^2 xe^{x^2}\,dx.$$

Let $u = x^2$, so $du = 2x\,dx$ and $x\,dx = \frac{1}{2}\,du$. The limits become $0 \mapsto 0$ and $2 \mapsto 4$. Therefore

$$\int_0^2 xe^{x^2}\,dx = \frac{1}{2}\int_0^4 e^u\,du = \frac{1}{2}(e^4 - 1).$$

Concept: Integration by parts as reverse product rule

From $(uv)' = u'v + uv'$, we obtain

$$\int u\,dv = uv - \int v\,du.$$

A useful choice rule is LIATE: logarithmic, inverse-trigonometric, algebraic, trigonometric, exponential. The function earlier in the list is often a good choice for u.

Worked Example: Worked example 5: integration by parts

Evaluate

$$\int_0^1 xe^x \, dx.$$

Choose $u = x$ and $dv = e^x \, dx$. Then $du = \, dx$ and $v = e^x$. Thus

$$\int_0^1 xe^x \, dx = [xe^x]_0^1 - \int_0^1 e^x \, dx = e - (e-1) = 1.$$

Checkpoint: Checkpoint

In substitution, the main question is "what inner function has its derivative nearby?" In integration by parts, the main question is "which part becomes simpler after differentiating?"

Modeling applications of definite integrals

Core explanation, worked reasoning, and application

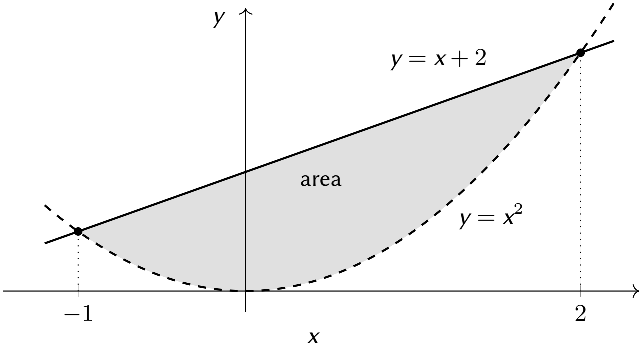
Concept: Area between two curves

If $f(x) \geq g(x)$ on $[a, b]$, the area between the curves is

$$A = \int_a^b [f(x) - g(x)]\, dx.$$

The order matters: top curve minus bottom curve.

Illustration: Modeling applications of definite integrals



$$\text{Area} = \int_{-1}^{2} \left[(x+2) - x^2\right] dx$$

For area between curves, integrate upper minus lower. The shaded region is bounded by $y = x + 2$ and $y = x^2$ from $x = -1$ to $x = 2$.

Worked Example: Worked example 6: area between curves

Find the area enclosed by $y = x^2$ and $y = x + 2$.
Intersections satisfy $x^2 = x + 2$, so $x = -1$ or $x = 2$. On this interval, $x + 2 \geq x^2$. Hence

$$A = \int_{-1}^{2} [(x+2) - x^2]\, dx = \left[\frac{x^2}{2} + 2x - \frac{x^3}{3}\right]_{-1}^{2} = \frac{9}{2}.$$

Concept: Average value and accumulated change

For a continuous function f on $[a, b]$, its **average value** is

$$\bar{f} = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

If $Q'(t) = r(t)$, then

$$Q(b) - Q(a) = \int_a^b r(t)\,dt.$$

This is the form used in flow, pollution, population, inventory, and energy models.

Worked Example: Worked example 7: average inventory

In the EOQ model, suppose inventory decreases linearly from Q to 0 over a cycle of length T:

$$I(t) = Q\left(1 - \frac{t}{T}\right), \qquad 0 \le t \le T.$$

The average inventory is

$$\bar{I} = \frac{1}{T}\int_0^T Q\left(1 - \frac{t}{T}\right)\,dt = \frac{Q}{T}\left[t - \frac{t^2}{2T}\right]_0^T = \frac{Q}{2}.$$

This familiar result is an integral statement, not only a geometric triangle argument.

Model Critique: Model criticism: units and sign

Before reporting an integral, check its units. If $r(t)$ is litres/hour and t is measured in hours, then $\int r(t)\,dt$ is litres. Also check whether signed area makes sense: negative values may represent outflow, loss, or correction rather than literal negative area.

Improper integrals and convergence

Core explanation, worked reasoning, and application

Improper integrals and convergence

Some modeling integrals extend over an infinite interval or include an unbounded integrand. These are **improper integrals**. They must be defined using limits.

Concept: Two common improper integrals

For an infinite interval,

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

For a vertical blow-up at a,

$$\int_a^b f(x)\,dx = \lim_{c\to a^+} \int_c^b f(x)\,dx.$$

If the limit exists, the integral **converges**. Otherwise it **diverges**.

Worked Example: Worked example 8: convergence versus divergence

Compare

$$\int_1^\infty \frac{1}{x^2}\,dx \qquad \text{and} \qquad \int_1^\infty \frac{1}{x}\,dx.$$

For the first,

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{b\to\infty}\left[-\frac{1}{x}\right]_1^b = 1,$$

so it converges. For the second,

$$\int_1^\infty \frac{1}{x}\,dx = \lim_{b\to\infty}\ln b = \infty,$$

so it diverges. Infinite interval does not automatically mean infinite area.

Checkpoint: Checkpoint

Why is the computation $\int_{-1}^{1} 1/x^2\, dx = [-1/x]_{-1}^{1}$ invalid? Identify the hidden discontinuity.

Numerical integration

Core explanation, worked reasoning, and application

Numerical integration

Many integrals in modeling do not have elementary antiderivatives. For example, $\int e^{-x^2}\,dx$ appears in statistics, but cannot be expressed using elementary functions. We then approximate.

Concept: Trapezoidal and Simpson rules

For a regular partition of $[a, b]$ with n subintervals and $h = (b - a)/n$:

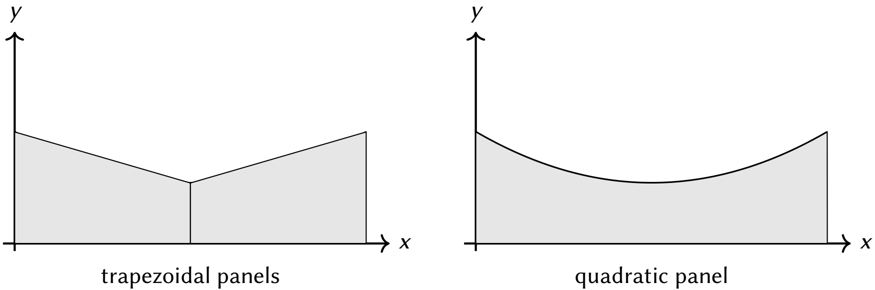
$$T_n = \frac{h}{2}\left[f(x_0) + 2f(x_1) + \cdots + 2f(x_{n-1}) + f(x_n)\right].$$

If n is even, Simpson's rule is

$$S_n = \frac{h}{3}\left[f(x_0) + 4f(x_1) + 2f(x_2) + \cdots + 4f(x_{n-1}) + f(x_n)\right].$$

Trapezoidal uses straight-line panels; Simpson uses quadratic panels.

Illustration: Numerical integration



Trapezoidal panels approximate the curve by straight segments. Simpson's rule approximates over pairs of subintervals by a quadratic curve.

Worked Example: Worked example 9: comparing numerical rules

Estimate

$$\int_0^1 e^{-x^2}\,dx$$

using $n = 4$ subintervals. With $h = 0.25$, the function values are approximately

x_k	0	0.25	0.50	0.75	1.00
$e^{-x_k^2}$	1.00000	0.93941	0.77880	0.56978	0.36788

Thus

$$T_4 = 0.74298, \qquad S_4 = 0.74685.$$

The true value is approximately 0.746824, so Simpson's rule is much more accurate here.

Model Critique: Numerical integration in a modeling report

A good report should not only state "we used Simpson's rule." It should specify the interval, step size, stopping criterion, and a validation check. For example:

> *We estimated the integral using Simpson's rule with step size $h = 0.01$. Doubling the number of subintervals changed the value by less than 10^{-4}, so numerical error is small relative to data uncertainty.*

Summary and practice

Core explanation, worked reasoning, and application

Summary: Key ideas

- Riemann sums approximate integrals by finite pieces.
- The definite integral is an accumulation operator: area, total change, average value, probability, or mass.
- FTC connects accumulated change and instantaneous rate.
- Substitution is reverse chain rule; integration by parts is reverse product rule.
- Improper integrals require limits and may converge or diverge.
- Numerical integration is a modeling choice; accuracy and cost should be justified.

Exercises: Core practice

1. Compute left, right, and midpoint sums for $\int_0^2 x^2\,dx$ with $n = 4$. Compare with the exact value.
2. Evaluate $\int_1^4 (3x^2 - 2x + 5)\,dx$ using the FTC.
3. Evaluate $\int_0^1 2x(1 + x^2)^5\,dx$ using substitution.
4. Evaluate $\int_1^e \ln x\,dx$ using integration by parts.
5. Determine whether $\int_1^\infty x^{-3/2}\,dx$ converges.
6. Find the average value of $f(x) = x^2$ on $[0, 3]$.

Exercises: Modeling challenge: cafeteria inflow

During the first 30 minutes of lunch, the arrival rate of students is modeled by

$$r(t) = 18 + 0.9t - 0.03t^2, \qquad 0 \leq t \leq 30,$$

where t is measured in minutes and $r(t)$ in students per minute.

1. Compute the total number of students arriving in the first 30 minutes.
2. Find the average arrival rate.
3. Explain one assumption behind using a smooth rate function for student arrivals.
4. Suggest one way to validate the model using observed queue data.

Exercises: Extension practice

1. Show that if f is odd and integrable on $[-a, a]$, then $\int_{-a}^{a} f(x)\,dx = 0$.
2. Use Simpson's rule with $n = 4$ to approximate $\int_0^1 \frac{1}{1+x^2}\,dx$ and compare with $\pi/4$.
3. Verify that $\int_0^\infty \lambda e^{-\lambda x}\,dx = 1$ for $\lambda > 0$, and explain why this matters for probability density models.

Supplementary 8: Probability Essentials for Modeling and Simulation

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this two-hour supplementary lesson, students should be able to:

- describe a random experiment using a **sample space**, **event**, and probability model;
- compute simple probabilities using counting, complements, and inclusion–exclusion;
- distinguish **conditional probability**, **independence**, and disjointness;
- use **Bayes' theorem** to update probabilities from evidence;
- model uncertainty using **random variables**, PMF/PDF, CDF, expectation, and variance;
- recognise when binomial, Poisson, exponential, normal, or Markov-chain models are appropriate;
- explain how the Law of Large Numbers and Central Limit Theorem justify Monte Carlo simulation.

Why probability belongs in mathematical modeling

Core explanation, worked reasoning, and application

Why probability belongs in mathematical modeling

A deterministic model gives one output when the input is fixed. Real systems often require a different language. Customers do not arrive exactly every five minutes; a sensor reading contains noise; a production line occasionally creates defective products; weather changes from day to day. A modeling report should therefore be able to say not only "what will happen", but also "how likely", "with what variability", and "how confident are we?"

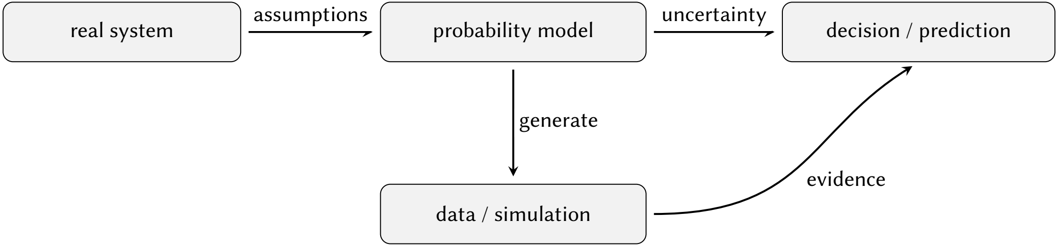
Activity: Launch activity: deterministic or stochastic?

Classify each situation as mainly deterministic, stochastic, or mixed.

1. A robot travels 120 m at a constant speed of 1.5 m/s.
2. Customers enter a canteen during lunch time.
3. A component has a 2% chance of failing during one month.
4. A delivery route is planned, but travel time depends on weather and crowding.

A strong modeling answer usually separates the predictable structure from the random variation.

Illustration: Why probability belongs in mathematical modeling



Probability enters modeling when assumptions must describe uncertainty, variability, and evidence rather than a single fixed outcome.

Sample spaces, events, and counting

Core explanation, worked reasoning, and application

Concept: Sample space and event

A **random experiment** is a procedure whose outcome is uncertain but whose possible outcomes can be described. The **sample space** Ω is the set of all possible outcomes. An **event** is a subset $A \subseteq \Omega$.

Worked Example: Worked example 1: two dice

For rolling two fair dice,

$$\Omega = \{(i,j) : 1 \leq i,j \leq 6\}, \qquad |\Omega| = 36.$$

The event “sum is 7” is

$$\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\},$$

so $\mathbb{P}(\text{sum} = 7) = 6/36 = 1/6$.

Concept: Probability rules

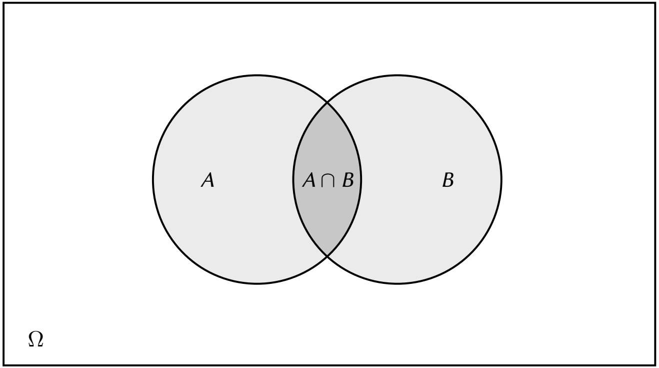
For events A and B,

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A), \qquad \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

If all outcomes are equally likely and Ω is finite, then

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}.$$

Illustration: Sample spaces, events, and counting



$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$$

Inclusion–exclusion subtracts the overlap once because it is counted in both A and B.

Worked Example: Worked example 2: birthday problem

Assume birthdays are uniformly distributed over 365 days. For $n = 23$ people, compute the probability that at least two people share a birthday.
It is easier to use the complement. The probability all birthdays are distinct is

$$\frac{365 \cdot 364 \cdot 363 \cdots (365 - 22)}{365^{23}} \approx 0.4927.$$

Therefore

$$\mathbb{P}(\text{at least one shared birthday}) \approx 1 - 0.4927 = 0.5073.$$

Only 23 people are enough for a probability slightly above one half.

Checkpoint: Checkpoint

A bag has 5 red balls and 3 blue balls. Two balls are drawn without replacement. Is "first ball red" independent of "second ball red"? Explain using conditional probability.

Conditional probability, independence, and Bayes

Core explanation, worked reasoning, and application

Concept: Conditional probability

For events A and B with $\mathbb{P}(B) > 0$,

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

This means: restrict the sample space to B, then renormalise.

Concept: Independence

Events A and B are **independent** if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B),$$

or equivalently $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ when $\mathbb{P}(B) > 0$. Independence is not the same as disjointness. If two positive-probability events are disjoint, knowing one happened makes the other impossible, so they are dependent.

Concept: Total probability and Bayes' theorem

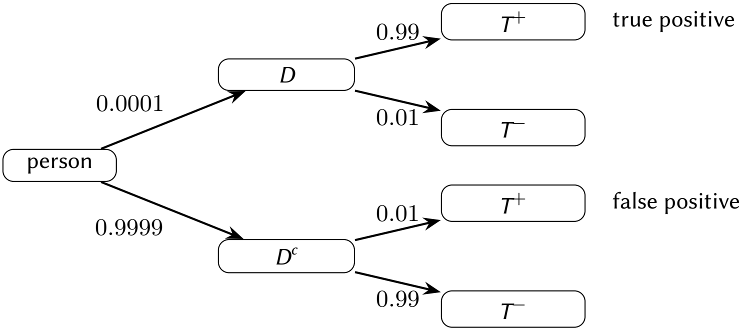
If $B_1, \ldots, B_n$ partition Ω, then

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i).$$

Bayes' theorem updates the probability of a cause after observing evidence:

$$\mathbb{P}(B_k \mid A) = \frac{\mathbb{P}(A \mid B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}.$$

Illustration: Conditional probability, independence, and Bayes



A probability tree separates prior probability from test reliability. The false-positive branch can dominate when the base rate is very small.

Worked Example: Worked example 3: medical testing and base rate

A disease affects 1 in 10,000 people. A test has 99% sensitivity and 99% specificity. A randomly selected person tests positive. What is the probability they have the disease?
Let D be "has disease" and T^+ be "positive test". Then

$$\mathbb{P}(D) = 10^{-4}, \quad \mathbb{P}(T^+ \mid D) = 0.99, \quad \mathbb{P}(T^+ \mid D^c) = 0.01.$$

Therefore

$$\mathbb{P}(D \mid T^+) = \frac{0.99(10^{-4})}{0.99(10^{-4}) + 0.01(0.9999)} \approx 0.0098.$$

Even after a positive test, the probability is under 1%. This is not because the test is useless; it is because the disease is rare.

Model Critique: Model criticism: Bayes in reports

When using Bayes' theorem, a report should state the prior probabilities, evidence reliability, and what changes after observing evidence. Many wrong conclusions come from ignoring the base rate.

Random variables, distributions, expectation, and variance

Core explanation, worked reasoning, and application

Concept: Random variable

A **random variable** is a function $X : \Omega \to \mathbb{R}$ that assigns a numerical value to each outcome. It is **discrete** if its possible values are finite or countable, and **continuous** if probabilities are described by intervals and integrals.

Concept: PMF, PDF, and CDF

For a discrete random variable, the **probability mass function** is

$$p_X(x) = \mathbb{P}(X = x), \qquad \sum_x p_X(x) = 1.$$

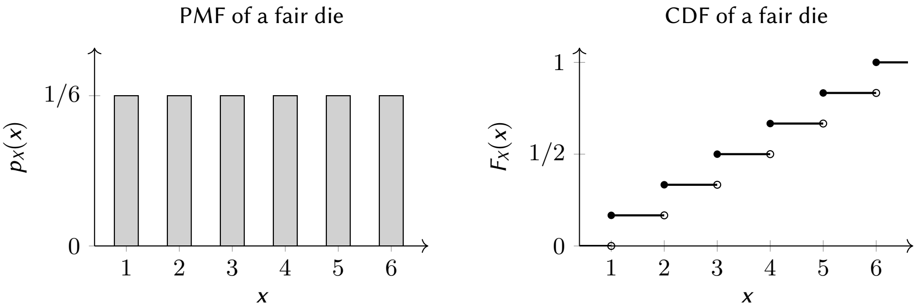
For a continuous random variable, the **probability density function** satisfies

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\, dx.$$

For any random variable, the **cumulative distribution function** is

$$F_X(x) = \mathbb{P}(X \leq x).$$

Illustration: Random variables, distributions, expectation, and variance



A discrete distribution has mass at points. Its CDF increases in jumps.

Concept: Expectation and variance

The **expected value** is the long-run average:

$$\mathbb{E}[X] = \sum_x x p_X(x) \quad \text{(discrete)}, \qquad \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx \quad \text{(continuous)}.$$

The **variance** measures spread:

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

The standard deviation is $\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$.

Worked Example: Worked example 4: die roll mean and variance

Let X be the result of a fair six-sided die. Then

$$\mathbb{E}[X] = \frac{1+2+3+4+5+6}{6} = 3.5,$$

$$\mathbb{E}[X^2] = \frac{1+4+9+16+25+36}{6} = \frac{91}{6},$$

so

$$\mathrm{Var}(X) = \frac{91}{6} - (3.5)^2 = \frac{35}{12} \approx 2.917.$$

The standard deviation is about 1.708.

Worked Example: Worked example 5: expected profit

An insurer sells a one-year policy for a premium of \$1,500. The claim is \$200,000 if death occurs, and the probability of death in the year is $p = 0.01$. The insurer's expected profit is

$$0.99(1500) + 0.01(1500 - 200000) = 1485 - 1985 = -500.$$

At this premium the policy loses \$500 in expectation. This connects directly to decision theory and risk modeling.

Named distributions for modeling

Core explanation, worked reasoning, and application

Summary: When to use which distribution

Distribution	Modeling question	Key quantities
Bernoulli(p)	one success/failure trial	mean p, variance $p(1-p)$
Binomial(n, p)	number of successes in n independent trials	mean np, variance $np(1-p)$
Poisson(λ)	number of random arrivals in a fixed interval	mean λ, variance λ
Exponential(λ)	waiting time between Poisson arrivals	mean $1/\lambda$, memoryless
Normal(μ, σ^2)	measurement error, averages, natural variability	mean μ, variance σ^2

Worked Example: Worked example 6: binomial quality control

A factory has defect rate $p = 0.02$. In a batch of $n = 100$ items, let X be the number of defective items. Then $X \sim \mathrm{Bin}(100, 0.02)$ and

$$\mathbb{E}[X] = 100(0.02) = 2.$$

The probability of at most 2 defects is

$$\mathbb{P}(X \leq 2) = \sum_{k=0}^{2} \binom{100}{k} (0.02)^k (0.98)^{100-k} \approx 0.677.$$

Worked Example: Worked example 7: Poisson arrivals and exponential waiting

Customers arrive at a service window at average rate 12 per hour. If arrivals are modeled as Poisson, then the number of arrivals in one hour is $X \sim \mathrm{Pois}(12)$:

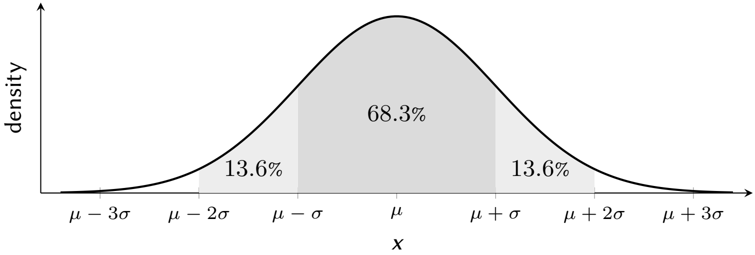
$$\mathbb{P}(X = 10) = e^{-12}\frac{12^{10}}{10!} \approx 0.105.$$

The waiting time until the next customer is $T \sim \mathrm{Exp}(12)$ hours, so

$$\mathbb{E}[T] = \frac{1}{12}\text{ hour} = 5\text{ minutes}.$$

This is the probability foundation of queue simulation.

Illustration: Named distributions for modeling



The normal distribution is used for measurement error, natural variation, and sampling distributions. About 68.3% of values lie within one standard deviation of the mean.

Checkpoint: Checkpoint

Which distribution would you use for each case?

1. Number of students arriving in the next 10 minutes.
2. Whether a randomly selected component is defective.
3. Waiting time until the next customer.
4. Measurement error of repeated sensor readings.

Monte Carlo, LLN, and CLT

Core explanation, worked reasoning, and application

Monte Carlo, LLN, and CLT

The link between probability and simulation is simple: if we cannot compute a quantity exactly, we can sometimes estimate it by repeated random trials. The theory tells us why the estimate improves.

Concept: Law of Large Numbers

Let $X_1, X_2, \ldots$ be independent and identically distributed with mean μ. The sample mean

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

converges to μ in probability. In modeling language: repeated simulation averages approach the expected value.

Concept: Central Limit Theorem

If $X_1, \ldots, X_n$ are i.i.d. with mean μ and variance σ^2, then for large n,

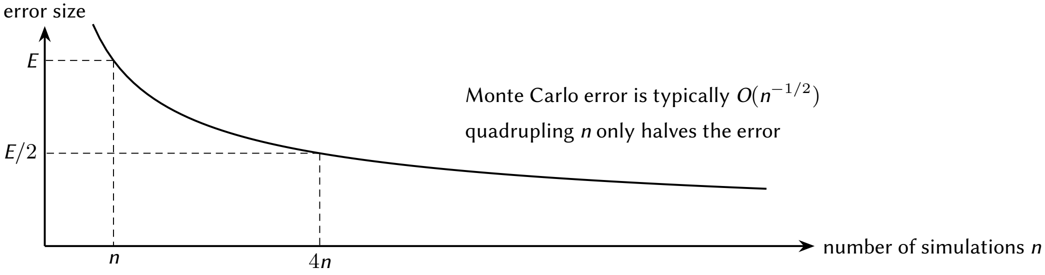
$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \approx Z \sim \mathcal{N}(0,1).$$

Therefore an approximate 95% confidence interval for μ is

$$\bar{X}_n \pm 1.96\frac{s}{\sqrt{n}},$$

where s is the sample standard deviation.

Illustration: Monte Carlo, LLN, and CLT



The $1/\sqrt{n}$ convergence rate explains why Monte Carlo is flexible but often slow.

Worked Example: Worked example 8: sample size for Monte Carlo

Suppose a Monte Carlo estimator is an average of Bernoulli indicators, so its standard deviation is at most $1/2$. How large should n be for a 95% confidence interval half-width at most 0.005?

$$1.96\frac{1/2}{\sqrt{n}} \leq 0.005 \quad \Longrightarrow \quad \sqrt{n} \geq 196 \quad \Longrightarrow \quad n \geq 38416.$$

This explains why Monte Carlo may need many samples in one dimension, even though it becomes valuable in high dimensions.

Model Critique: Reporting simulation uncertainty

A competition report should not only give a simulated estimate. It should also report the number of replications, the standard error, a confidence interval, and whether the conclusion changes under reasonable parameter perturbations.

A first Markov-chain model

Core explanation, worked reasoning, and application

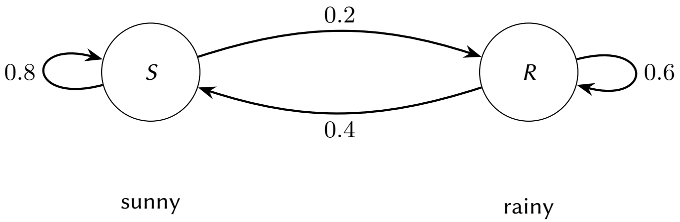
Concept: Discrete-time Markov chain

A sequence $X_0, X_1, X_2, \ldots$ on a finite state space is a **Markov chain** if the next state depends on the present state but not on the full past:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1}, \ldots, X_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i) = P_{ij}.$$

The matrix $P = (P_{ij})$ is the **transition matrix**; each row sums to 1.

Illustration: A first Markov-chain model



A two-state Markov chain for weather. The probabilities leaving each state add to 1.

Worked Example: Worked example 9: two-day weather prediction

States are S = sunny and R = rainy. Suppose

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix},$$

where rows are today's weather and columns are tomorrow's weather. If today is sunny, the distribution after two days is

$$(1,0)P^2 = (1,0)\begin{pmatrix} 0.72 & 0.28 \\ 0.56 & 0.44 \end{pmatrix} = (0.72, 0.28).$$

Thus the probability of rain two days from now is 0.28.

Worked Example: Worked example 10: stationary distribution

A stationary distribution $\pi = (\pi_S, \pi_R)$ satisfies $\pi P = \pi$ and $\pi_S + \pi_R = 1$. From the weather matrix,

$$\pi_S = 0.8\pi_S + 0.4\pi_R \quad \Longrightarrow \quad 0.2\pi_S = 0.4\pi_R \quad \Longrightarrow \quad \pi_S = 2\pi_R.$$

Together with $\pi_S + \pi_R = 1$, we get

$$\pi = (2/3, 1/3).$$

In the long run, about two-thirds of days are sunny.

Differentiated modeling studio

Core explanation, worked reasoning, and application

Activity: Studio task: canteen queue uncertainty

A canteen receives students at an average rate of 18 per 10 minutes. During peak time, arrivals are modeled as Poisson.

1. Compute the probability of exactly 20 arrivals in the next 10 minutes.
2. Compute or approximate the probability of more than 24 arrivals.
3. If one server can handle 20 students per 10 minutes, what risk does this create?
4. Write one paragraph explaining whether the Poisson assumption is reasonable.

Summary: Competition-style reporting paragraph

We model arrivals in each 10-minute period as $X \sim \mathrm{Pois}(18)$, assuming independent arrivals at an approximately constant rate during the short peak interval. This gives $\mathbb{E}[X] = 18$ and $\mathrm{Var}(X) = 18$, so the typical fluctuation is about $\sqrt{18} \approx 4.24$ students. Since the server capacity is 20 students per 10 minutes, the system may become overloaded in a non-negligible fraction of intervals. We therefore recommend evaluating service capacity using not only the mean arrival rate, but also the upper tail probability $\mathbb{P}(X > 20)$ and sensitivity to the estimated rate.

Model Critique: Model criticism

Probability models are assumptions, not facts. A Poisson arrival model assumes independent arrivals and a roughly constant rate. If students arrive in groups after a bell rings, the independence assumption is weak. If the rate changes sharply during lunch, the constant-rate assumption is weak. Validation should compare both the mean and the variability of observed arrivals against the model.

Core practice and extension problems

Core explanation, worked reasoning, and application

Exercises: Core practice

1. Two fair dice are rolled. Find $\mathbb{P}(\text{sum} = 8)$ and $\mathbb{P}(\text{sum} \geq 10)$.
2. A bag contains 5 red and 3 blue balls. Two are drawn without replacement. Find the probability both are red.
3. A factory has two machines. Machine A produces 70% of items with defect rate 1%; machine B produces 30% with defect rate 4%. A randomly chosen item is defective. Find the probability it came from B.
4. Let X be the number of heads in 4 tosses of a fair coin. Write its distribution and compute $\mathbb{E}[X]$.
5. If $X \sim \text{Pois}(5)$, compute $\mathbb{P}(X = 3)$ and $\mathbb{P}(X = 0)$.
6. If $T \sim \text{Exp}(6)$ hours, find $\mathbb{E}[T]$ and $\mathbb{P}(T > 0.5)$.

Exercises: Modeling challenge

A small help desk receives requests according to a Poisson process with rate $\lambda = 10$ per hour. Service time is approximately exponential with mean 5 minutes.

1. Convert the mean service time into a service rate μ per hour.
2. Compute the utilisation ratio $\rho = \lambda/\mu$.
3. Explain what it means if ρ is close to 1.
4. Suggest one data source that could validate the arrival-rate assumption and one source that could validate the service-time assumption.

Exercises: Extension practice

1. Prove that if A and B are independent, then A and B^c are independent.
2. Let $X \sim U(0,1)$ and $Y = -\ln X$. Derive the CDF of Y and identify its distribution.
3. Let $X_1, \ldots, X_n$ be independent Bernoulli(p). Use linearity of expectation to show $\mathbb{E}[\sum X_i] = np$.
4. In the weather Markov chain, if today is rainy, compute the probability of a sunny day three days from now.
5. Give an example where $\mathrm{Cov}(X, Y) = 0$ but X and Y are not independent.

Supplementary 9: Linear Algebra Essentials for Modeling
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this two-hour supplementary lesson, students should be able to:

- represent data, decisions, and states using **vectors**;
- compute and interpret **dot products**, norms, linear combinations, and span;
- read a **matrix** as a table, a linear transformation, or a system of equations;
- solve small systems by **Gaussian elimination**;
- interpret **rank**, free variables, and the solution structure of homogeneous systems;
- connect linear algebra to least squares, Markov chains, game theory, and dimensional analysis;
- explain why eigenvectors and stationary distributions matter in dynamic modeling.

Why linear algebra belongs in modeling

Core explanation, worked reasoning, and application

Why linear algebra belongs in modeling

Calculus describes how one quantity changes. Linear algebra describes how many quantities combine. In mathematical modeling, this is unavoidable: a delivery route has many travel times, a regression model has many residuals, a Markov chain has many transition probabilities, and a dimensional-analysis problem has many physical exponents.

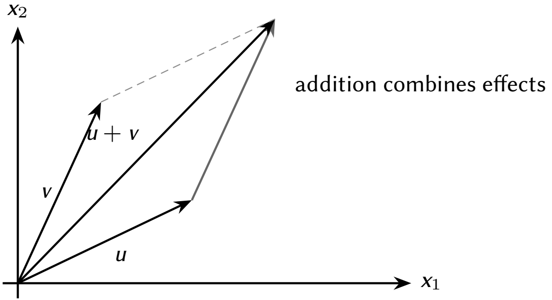
Activity: Launch activity: where are the vectors?

For each modeling situation, identify a natural vector.

1. A robot route visits five buildings and records five travel times.
2. A school canteen records arrivals in six ten-minute intervals.
3. A company tracks market share among three brands over many weeks.
4. A regression model compares predicted and observed values for ten data points.

The goal is not to make notation heavier. It is to make multi-variable information readable and computable.

Illustration: Why linear algebra belongs in modeling



Vector addition models combined effects, such as displacement, resource use, or accumulated changes across components.

Vectors as data packages

Core explanation, worked reasoning, and application

Concept: Vectors, dot products, and norms

A vector in $\mathbb{R}^n$ is an ordered list

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

For $u, v \in \mathbb{R}^n$,

$$u \cdot v = u_1 v_1 + \cdots + u_n v_n, \qquad \|v\| = \sqrt{v \cdot v}.$$

The dot product measures alignment. If $u \cdot v = 0$, the vectors are **orthogonal**.

Worked Example: Worked example 1: resource-use vector

A delivery robot uses battery, time, and staff attention. For one route, suppose

$$x = \begin{pmatrix} 18 \\ 42 \\ 3 \end{pmatrix}$$

means 18 battery units, 42 minutes, and 3 staff interventions. A cost vector

$$c = \begin{pmatrix} 0.4 \\ 1.2 \\ 8 \end{pmatrix}$$

assigns weights to each component. The total weighted cost is

$$c \cdot x = 0.4(18) + 1.2(42) + 8(3) = 81.6.$$

This is the same idea as an LP objective $c \cdot x$ in Lecture 6.

Concept: Linear combination and span

A **linear combination** of $v_1, \ldots, v_k$ is

$$c_1 v_1 + \cdots + c_k v_k.$$

The set of all such combinations is the **span**, written

$$\mathrm{span}(v_1, \ldots, v_k).$$

Vectors are **linearly independent** if the only way to make $\mathbf{0}$ is to use all zero coefficients.

Worked Example: Worked example 2: detecting dependence

Let

$$v_1 = \begin{pmatrix}1\\2\\0\end{pmatrix}, \quad v_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}, \quad v_3 = \begin{pmatrix}1\\4\\2\end{pmatrix}.$$

Since

$$v_1 + 2v_2 = \begin{pmatrix}1\\2\\0\end{pmatrix} + 2\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}1\\4\\2\end{pmatrix} = v_3,$$

the three vectors are linearly dependent. In modeling terms, the third feature adds no new direction; it is already explained by the first two.

Checkpoint: Checkpoint

In $\mathbb{R}^3$, can four vectors be linearly independent? Explain without computation.

Matrices as structured models

Core explanation, worked reasoning, and application

Concept: Matrix multiplication

An $m \times n$ matrix $\mathbf{A}$ maps vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$ by

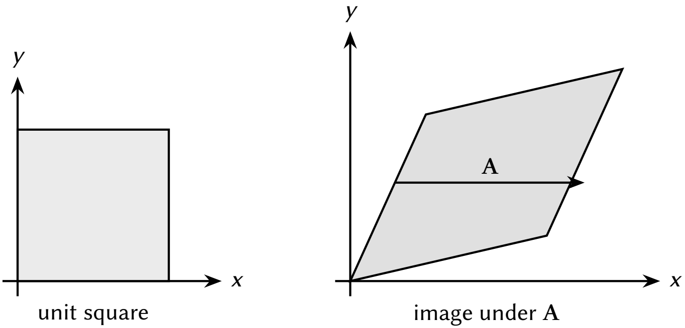
$$y = \mathbf{A}x.$$

If $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times p}$, then

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

The product $\mathbf{AB}$ represents composition: do $\mathbf{B}$ first, then $\mathbf{A}$.

Illustration: Matrices as structured models



A matrix can be viewed as a transformation. The columns of $\mathbf{A}$ show where the unit basis vectors move.

Worked Example: Worked example 3: matrix product as a transition

Suppose a two-state weather model uses the transition matrix

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix},$$

where rows represent today's state and columns represent tomorrow's state: sunny S or rainy R. If today is certainly sunny, the row vector is

$$x_0^\top = (1, 0).$$

After one day,

$$x_1^\top = x_0^\top \mathbf{P} = (0.8, 0.2).$$

After two days,

$$x_2^\top = x_0^\top \mathbf{P}^2 = (0.72, 0.28).$$

Matrix powers describe repeated transitions, which is the core of Markov-chain modeling.

Model Critique: Model critique: what does a matrix assume?

A transition matrix assumes that the next state depends on the current state in a fixed way. For weather, customer movement, or disease progression, ask:

- Are transition probabilities stable over time?
- Does tomorrow depend only on today, or also on the previous week?
- Are all relevant states included?

A matrix is compact, but compact notation can hide strong assumptions.

Solving systems by Gaussian elimination

Core explanation, worked reasoning, and application

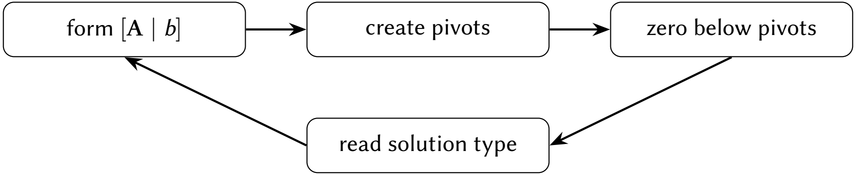
Concept: Linear system

A system of m linear equations in n unknowns can be written as

$$\mathbf{A}x = b.$$

Gaussian elimination uses row operations on the augmented matrix $[\mathbf{A} \mid b]$ to reveal pivots, free variables, and consistency.

Summary: Gaussian elimination route



Worked Example: Worked example 4: a 3×3 system

Solve

$$\begin{cases} x + 2y + z = 6, \\ 2x + 3y - z = 4, \\ -x + y + 4z = 9. \end{cases}$$

The augmented matrix reduces as follows:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 6 \\ 2 & 3 & -1 & 4 \\ -1 & 1 & 4 & 9 \end{array}\right] \xrightarrow{R_2 - 2R_1,\ R_3 + R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 6 \\ 0 & -1 & -3 & -8 \\ 0 & 3 & 5 & 15 \end{array}\right]$$

$$\xrightarrow{R_3 + 3R_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 6 \\ 0 & -1 & -3 & -8 \\ 0 & 0 & -4 & -9 \end{array}\right].$$

Back substitution gives $z = 9/4$, $y = 5/4$, and $x = 5/4$. Therefore

$$(x, y, z) = \left(\frac{5}{4}, \frac{5}{4}, \frac{9}{4}\right).$$

Concept: Rank and solution type

The **rank** of a matrix is the number of pivots in row echelon form. A system has:

- no solution if row reduction gives $[0\ 0\ \cdots\ 0 \mid c]$ with $c \neq 0$;
- exactly one solution if every variable column has a pivot;
- infinitely many solutions if the system is consistent but has free variables.

For $\mathbf{A} \in \mathbb{R}^{m \times n}$,

$$\dim(\text{null space}) = n - \operatorname{rank}(\mathbf{A}).$$

Worked Example: Worked example 5: null space and dimensionless products

For a pendulum model, suppose the period T may depend on length L, mass m, and gravitational acceleration g. A dimensionless product has the form

$$T^a L^b m^c g^d.$$

Using dimensions $[T] = (0, 0, 1)$, $[L] = (0, 1, 0)$, $[m] = (1, 0, 0)$, and $[g] = (0, 1, -2)$ in rows M, L, T, the dimension equations are

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \mathbf{0}.$$

The equations are $c = 0$, $b + d = 0$, and $a - 2d = 0$. Taking $d = 1$ gives

$$(a, b, c, d) = (2, -1, 0, 1),$$

so one dimensionless product is

$$\Pi = \frac{T^2 g}{L}.$$

This is exactly the null-space calculation behind Buckingham's Π theorem.

Least squares as linear algebra

Core explanation, worked reasoning, and application

Concept: Residual vector and normal equations

For data (x_i, y_i) and a line $y = a + bx$, define

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

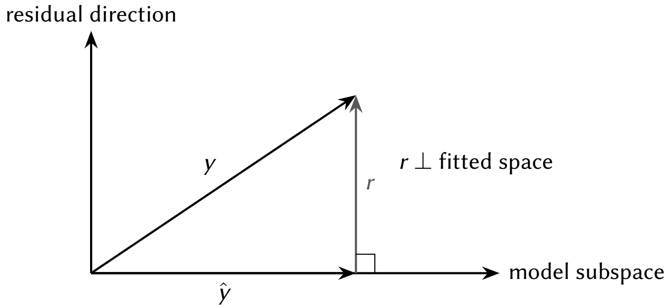
Predictions are $\hat{y} = \mathbf{X}\boldsymbol{\beta}$ and residuals are

$$r = y - \mathbf{X}\boldsymbol{\beta}.$$

Least squares chooses $\boldsymbol{\beta}$ to minimise $\|r\|^2$. The normal equations are

$$\mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} = \mathbf{X}^\top y.$$

Illustration: Least squares as linear algebra



Least squares chooses the fitted vector $\hat{y}$ so that the residual vector is orthogonal to the model subspace. This gives $\mathbf{X}^\top r = \mathbf{0}$.

Worked Example: Worked example 6: fitting a line

Fit $y = a + bx$ to the data

$$(0,1),\ (1,2),\ (2,2),\ (3,4).$$

Then

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \qquad y = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 4 \end{pmatrix}.$$

Compute

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}, \qquad \mathbf{X}^\top y = \begin{pmatrix} 9 \\ 18 \end{pmatrix}.$$

Solve

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 9 \\ 18 \end{pmatrix}.$$

The solution is $a = 0.9$, $b = 0.9$. The fitted model is

$$\hat{y} = 0.9 + 0.9x.$$

Model Critique: Model critique: matrix calculation is not the whole model

The normal equations compute the best line under a squared-error criterion. A modeling report still needs to discuss residual patterns, units, outliers, whether the relationship should be linear, and whether the prediction range is appropriate.

Eigenvalues, stationary states, and long-run behavior

Core explanation, worked reasoning, and application

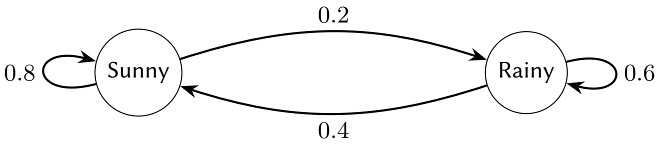
Concept: Eigenvalue and eigenvector

For a square matrix $\mathbf{A}$, a non-zero vector v is an **eigenvector** with **eigenvalue** λ if

$$\mathbf{A}v = \lambda v.$$

The direction of v is preserved by the transformation; only its scale changes.

Illustration: Eigenvalues, stationary states, and long-run behavior



A two-state Markov chain. The stationary distribution is a left eigenvector of the transition matrix with eigenvalue 1.

Worked Example: Worked example 7: stationary distribution

For

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix},$$

a stationary distribution $\boldsymbol{\pi} = (\pi_1, \pi_2)$ satisfies

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}, \qquad \pi_1 + \pi_2 = 1.$$

The first equation gives

$$0.8\pi_1 + 0.4\pi_2 = \pi_1 \quad \Rightarrow \quad 0.4\pi_2 = 0.2\pi_1 \quad \Rightarrow \quad \pi_1 = 2\pi_2.$$

Together with $\pi_1 + \pi_2 = 1$, we get

$$\boldsymbol{\pi} = \left(\frac{2}{3}, \frac{1}{3}\right).$$

In the long run, about two thirds of days are sunny in this model.

Worked Example: Worked example 8: population matrix

Let J_n and A_n be juvenile and adult populations in year n. Suppose each adult produces 0.6 juveniles, 30% of juveniles mature, and 80% of adults survive. Then

$$\begin{pmatrix} J_{n+1} \\ A_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0.6 \\ 0.3 & 0.8 \end{pmatrix}}_{\mathbf{M}} \begin{pmatrix} J_n \\ A_n \end{pmatrix}.$$

The dominant eigenvalue of $\mathbf{M}$ describes the long-run growth factor. Solving

$$\det(\mathbf{M} - \lambda \mathbf{I}) = \det \begin{pmatrix} -\lambda & 0.6 \\ 0.3 & 0.8 - \lambda \end{pmatrix} = \lambda^2 - 0.8\lambda - 0.18 = 0$$

gives $\lambda \approx 0.983$. The model predicts a slow long-run decline, because the dominant growth factor is slightly below 1.

Differentiated modeling studio

Core explanation, worked reasoning, and application

Activity: Modeling studio

Choose one route according to readiness.

Core route. Solve a 3×3 linear system by elimination and explain the solution type.

Challenge route. Fit a line to five data points using the normal equations and compute the residual vector.

Extension route. Find the stationary distribution of a three-state transition matrix and interpret the long-run state proportions.

Each group should write one short paragraph explaining what the matrix represents before doing the computation.

Summary: Competition-style communication

A concise modeling report paragraph might read:

We represented the system state as a vector and encoded transitions by a matrix. Repeated behavior was computed using matrix multiplication, while the long-run distribution was obtained by solving the stationary equation $\boldsymbol{\pi} P = \boldsymbol{\pi}$ *with* $\sum_i \pi_i = 1$*. This interpretation assumes transition probabilities are stable and that the current state contains enough information to predict the next step.*

Checkpoint: Exit ticket

Answer in three sentences:

1. What does a pivot tell you in Gaussian elimination?
2. Why is a residual vector useful in least squares?
3. What does a stationary distribution mean in a Markov-chain model?

Exercises

Core explanation, worked reasoning, and application

1. Find a unit vector in the direction of $(3, 4, 0)$.
2. Compute the dot product of $(1, 2, 3)$ and $(4, -1, 2)$, then find the angle between them.
3. Determine whether $(1, 0, 2)$, $(3, 1, 4)$, and $(2, 1, 0)$ are linearly independent.
4. In a delivery model, $x = (8, 20, 2)$ records distance, time, and number of delays. If $c = (1.5, 0.8, 10)$, interpret $c \cdot x$.

Exercises: Exercise 2: Matrices and systems

1. Compute **AB** and **BA** for
$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 5 & 0 \\ 1 & 2 \end{pmatrix}.$$
2. Solve by Gaussian elimination:
$$\begin{cases} 2x + y - z = 3, \\ x - y + 2z = 1, \\ 3x + 2y + z = 10. \end{cases}$$
3. Explain why the system
$$x + 2y + 3z = 1, \qquad 2x + 4y + 6z = 3$$
has no solution.
4. Find all solutions of
$$\begin{cases} x_1 + x_2 - x_3 + 2x_4 = 0, \\ 2x_1 - x_2 + x_3 + x_4 = 0. \end{cases}$$

1. Fit $y = a + bx$ to $(1,2), (2,3), (3,5), (4,4), (5,6)$ using $\mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} = \mathbf{X}^\top y$.
2. For

$$P = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{pmatrix},$$

   find the stationary distribution.
3. A zero-sum game has payoff matrix

$$\mathbf{A} = \begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix}.$$

   Find Row's mixed strategy by making Column indifferent.
4. In dimensional analysis, explain why the number of independent dimensionless groups equals "number of variables minus rank of the dimension matrix."

Exercises: Exercise 4: Extension

1. Find eigenvalues and eigenvectors of $\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.
2. Let $\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Explain geometrically why $\mathbf{R}(\alpha)\mathbf{R}(\beta) = \mathbf{R}(\alpha+\beta)$.
3. Prove that if v is an eigenvector of $\mathbf{A}$ with eigenvalue λ, then v is an eigenvector of $\mathbf{A}^k$ with eigenvalue λ^k.
4. For a square matrix, list three equivalent ways to say that it is invertible.

Exercises

Appendix A: Algorithm summary

Core explanation, worked reasoning, and application

Appendix A: Algorithm summary

Task	**Practical method**
Solve $\mathbf{A}x = b$	Row-reduce $[\mathbf{A} \mid b]$, then back-substitute.
Find a null space	Row-reduce $\mathbf{A}$, assign parameters to free variables, write basis vectors.
Test independence	Put vectors as columns; check whether every column has a pivot.
Find least-squares line	Form $\mathbf{X}$, compute $\mathbf{X}^\top\mathbf{X}$ and $\mathbf{X}^\top y$, solve normal equations.
Find stationary distribution	Solve $\boldsymbol{\pi} P = \boldsymbol{\pi}$ plus $\sum_i \pi_i = 1$.
Find eigenvalues	Solve $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.
Find eigenvectors	For each eigenvalue, solve $(\mathbf{A} - \lambda\mathbf{I})v = \mathbf{0}$.

Appendix B: Connection map to the main lectures

Core explanation, worked reasoning, and application

Appendix B: Connection map to the main lectures

Tool	Where it appears
Dot product and norm	Lecture 3 residual vector; Lecture 6 linear objective $c \cdot x$.
Matrix multiplication	Lecture 5 transition matrices; Lecture 8 mixed strategies.
Gaussian elimination	Lecture 6 simplex thinking; Lecture 11 dimensionless products.
Rank and null space	Lecture 11 Buckingham Π theorem; Lecture 3 design matrix.
Least squares	Lecture 3 model fitting and residual analysis.
Eigenvalues	Lecture 5 Markov chains; Lecture 9 linear dynamical systems; Lecture 12 discrete systems.

Supplementary 10: Multivariable Calculus and Lagrange Multipliers

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Learning Goals

By the end of this two-hour supplementary lesson, students should be able to:

- interpret a function of several variables through formulas, level curves, and modeling context;
- compute and interpret **partial derivatives** as one-variable-at-a-time sensitivities;
- use the **gradient** to identify steepest change and directional rates;
- form a local linear approximation and state the units of each sensitivity coefficient;
- locate critical points and classify them using the **Hessian matrix**;
- solve one-constraint optimization problems using **Lagrange multipliers**;
- interpret the multiplier as a local **shadow price**;
- communicate an optimization result with assumptions, verification, and limitations.

Why several variables change the modeling problem

Core explanation, worked reasoning, and application

Why several variables change the modeling problem

A one-variable model asks how an outcome changes when one input moves. A multivariable model asks which input moved, in which direction, and under which constraints. For example, a canteen manager may choose staffing x, number of service counters y, and preparation level z. Waiting time, cost, and service quality all depend on the vector (x, y, z) rather than on one number.

Activity: Launch activity: what is being controlled?

A school event has expected net benefit

$$B(x, y) = 80x + 60y - 2x^2 - y^2,$$

where x and y are activity levels. The total staff-time constraint is $x + y = 20$.

1. Which expression is the objective? Which expression is the constraint?
2. Why can we not maximize B by examining x alone?
3. What does it mean to increase x while "holding y fixed"?

This lesson builds the tools needed to answer these questions systematically.

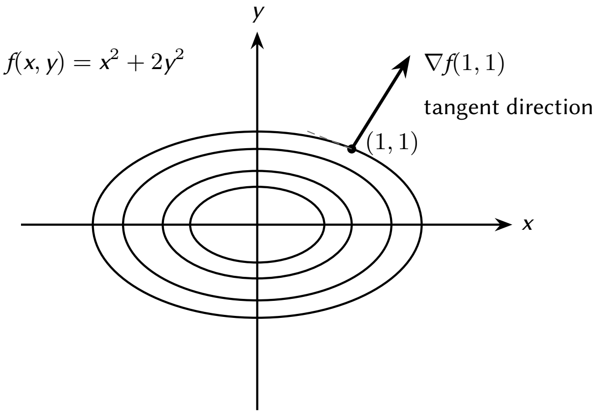
Concept: Functions and level curves

A function of two variables is a rule

$$f : D \subseteq \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto f(x, y).$$

A **level curve** is the set $f(x, y) = c$. Points on the same level curve give the same output. Contour maps use many level curves to describe a surface without drawing a fragile three-dimensional picture.

Illustration: Why several variables change the modeling problem



The gradient crosses a level curve at right angles. Moving along the curve changes direction but not the function value.

Partial derivatives and local sensitivity

Core explanation, worked reasoning, and application

Concept: Partial derivatives

For a function $f(x, y)$,

$$f_x(x, y) = \frac{\partial f}{\partial x}, \qquad f_y(x, y) = \frac{\partial f}{\partial y}.$$

To compute f_x, hold y fixed and differentiate with respect to x. To compute f_y, hold x fixed. In modeling language, these are **marginal effects**.

Worked Example: Worked example 1: interpreting units

Suppose the daily operating cost of a canteen is

$$C(x, y) = 400 + 30x + 45y + 0.8x^2 + 0.5xy,$$

where x is staff-hours and y is the number of prepared meal batches. Then

$$C_x = 30 + 1.6x + 0.5y, \qquad C_y = 45 + 0.5x.$$

At $(x, y) = (10, 8)$,

$$C_x(10, 8) = 50, \qquad C_y(10, 8) = 50.$$

Interpretation: near this operating point, one extra staff-hour adds about 50 currency units to daily cost, while one extra batch also adds about 50. The two numbers are equal here, but their units are different: cost per staff-hour versus cost per batch.

Concept: Local linear approximation

For small changes Δx and Δy,

$$f(x+\Delta x, y+\Delta y) \approx f(x,y) + f_x(x,y)\Delta x + f_y(x,y)\Delta y.$$

The change is approximated by

$$\Delta f \approx f_x \Delta x + f_y \Delta y.$$

This is the multivariable version of a tangent-line approximation and is often the first sensitivity calculation in a modeling report.

Worked Example: Worked example 2: a sensitivity estimate

Using the canteen cost model at $(10, 8)$, estimate the cost change if staffing rises by 0.5 hour while meal batches fall by 1:

$$\Delta C \approx C_x(0.5) + C_y(-1) = 50(0.5) - 50 = -25.$$

The model predicts a decrease of about 25 currency units. This is a local estimate; it should not be extrapolated to large changes without checking curvature.

Checkpoint: Checkpoint

If $f_x(a, b) = 0$ but $f_y(a, b) > 0$, is (a, b) necessarily a local minimum? Explain geometrically.

Gradient and directional change

Core explanation, worked reasoning, and application

Concept: Gradient and directional derivative

The gradient collects all first partial derivatives:

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix}.$$

For a unit vector u, the **directional derivative** is

$$D_u f = \nabla f \cdot u.$$

The gradient points in the direction of steepest increase, $-\nabla f$ in the direction of steepest decrease, and $\|\nabla f\|$ is the maximum instantaneous rate of increase.

Worked Example: Worked example 3: moving through a temperature field

Let

$$T(x, y) = 30 - 0.5x^2 - y^2 + 2x$$

be temperature in $^\circ$C, with x, y measured in kilometres. At $(1, 1)$,

$$\nabla T = (-x + 2, -2y) = (1, -2).$$

The steepest warming direction is

$$\frac{(1, -2)}{\sqrt{5}},$$

and the maximum warming rate is $\sqrt{5}$ $^\circ$C per kilometre. If a student walks east, $u = (1, 0)$, then $D_u T = 1$ $^\circ$C/km.

Model Critique: Model criticism: the gradient is local

A gradient is computed from the model at one point. It does not promise that the same direction remains best far away. For a long route, update the gradient along the path or solve a path-planning problem rather than following one fixed arrow.

Concept: Chain rule along a path

If a moving point follows $r(t) = (x(t), y(t))$, then

$$\frac{d}{dt}f(x(t), y(t)) = f_x\frac{dx}{dt} + f_y\frac{dy}{dt} = \nabla f \cdot r'(t).$$

This separates the field sensitivity ∇f from the movement $r'(t)$.

Unconstrained optimization and the Hessian

Core explanation, worked reasoning, and application

Concept: Critical points

At an interior local maximum or minimum of a differentiable function,

$$\nabla f = \mathbf{0}.$$

Such points are **critical points**. As in one variable, the equation is necessary but not sufficient: a critical point may be a minimum, maximum, or saddle.

Concept: Hessian test in two variables

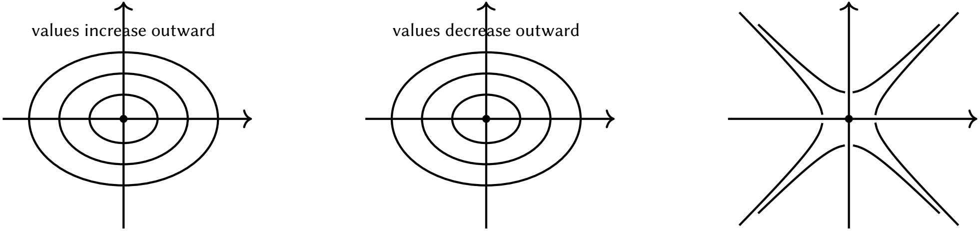
The Hessian is

$$\mathrm{H}f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}.$$

At a critical point, let $D = f_{xx}f_{yy} - f_{xy}^2$.

Condition	**Classification**
$D > 0$ and $f_{xx} > 0$	local minimum
$D > 0$ and $f_{xx} < 0$	local maximum
$D < 0$	saddle point
$D = 0$	inconclusive

Illustration: Unconstrained optimization and the Hessian



minimum: bowl   maximum: hill   saddle: mixed curvature

Contour patterns distinguish a local minimum, local maximum, and saddle. The Hessian records this local curvature algebraically.

Worked Example: Worked example 4: classify a critical point

For

$$f(x,y) = x^2 + 2y^2 - 2x - 8y + 5,$$

$$\nabla f = (2x-2, 4y-8) = \mathbf{0} \quad \Rightarrow \quad (x,y) = (1,2).$$

The Hessian is

$$\mathbf{H}f = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix},$$

so $D = 8 > 0$ and $f_{xx} = 2 > 0$. Therefore $(1,2)$ is a local minimum. Completing squares shows it is also global:

$$f = (x-1)^2 + 2(y-2)^2 - 4.$$

Worked Example: Worked example 5: a saddle point

For $g(x, y) = x^2 - y^2$, $\nabla g = (2x, -2y)$, so the only critical point is $(0, 0)$. The Hessian is $\mathrm{diag}(2, -2)$ and $D = -4 < 0$, hence the origin is a saddle: moving along the x-axis increases g, while moving along the y-axis decreases it.

Lagrange multipliers: optimizing on a constraint

Core explanation, worked reasoning, and application

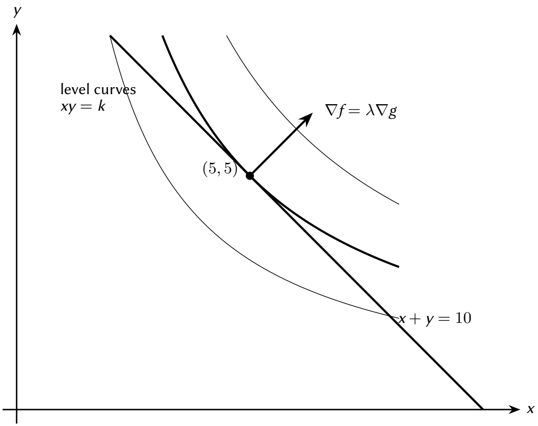
Concept: Geometric principle

Suppose we optimize $f(x, y)$ subject to $g(x, y) = c$. At a regular constrained extremum, a level curve of f is tangent to the constraint. Their normal vectors are parallel:

$$\nabla f = \lambda \nabla g, \qquad g(x, y) = c.$$

The scalar λ is the Lagrange multiplier.

Illustration: Lagrange multipliers: optimizing on a constraint



The maximum of $f(x, y) = xy$ subject to $x + y = 10$ occurs where the highest attainable level curve is tangent to the constraint, namely at $(5, 5)$.

Summary: Lagrange method

To optimize $f(x, y)$ subject to $g(x, y) = c$:

1. compute ∇f and ∇g;
2. solve $\nabla f = \lambda \nabla g$ together with $g = c$;
3. evaluate f at all candidates;
4. check domain boundaries, positivity restrictions, and whether the constraint is regular;
5. interpret the result in context, including units and sensitivity.

Worked Example: Worked example 6: maximum area with fixed perimeter

Let a rectangle have side lengths $x, y > 0$ and perimeter 20. Maximize

$$f(x,y) = xy \quad \text{subject to} \quad g(x,y) = x + y = 10.$$

Since

$$\nabla f = (y, x), \qquad \nabla g = (1, 1),$$

the Lagrange equations are

$$y = \lambda, \qquad x = \lambda, \qquad x + y = 10.$$

Hence $x = y = 5$ and the maximum area is 25. The positivity boundary gives area tending to 0, confirming the interior candidate is the constrained maximum.

Checkpoint: Checkpoint

Why is solving $\nabla f = \mathbf{0}$ inappropriate for the rectangle problem? What would it miss about the feasible set?

Shadow prices and a production-allocation model

Core explanation, worked reasoning, and application

Worked Example: Worked example 7: allocating a limited resource

A workshop chooses production levels x and y. Weekly benefit is

$$P(x, y) = 60x + 40y - 2x^2 - y^2,$$

subject to a resource limit

$$x + y = 20, \qquad x, y \geq 0.$$

Set $g(x, y) = x + y$. Then

$$\nabla P = (60 - 4x, 40 - 2y), \qquad \nabla g = (1, 1).$$

The equations

$$60 - 4x = \lambda, \qquad 40 - 2y = \lambda, \qquad x + y = 20$$

give $x = 10$, $y = 10$, and $\lambda = 20$. The optimal benefit is

$$P(10, 10) = 700.$$

Since the objective is concave (Hessian $\mathrm{diag}(-4, -2)$), this is the global constrained maximum.

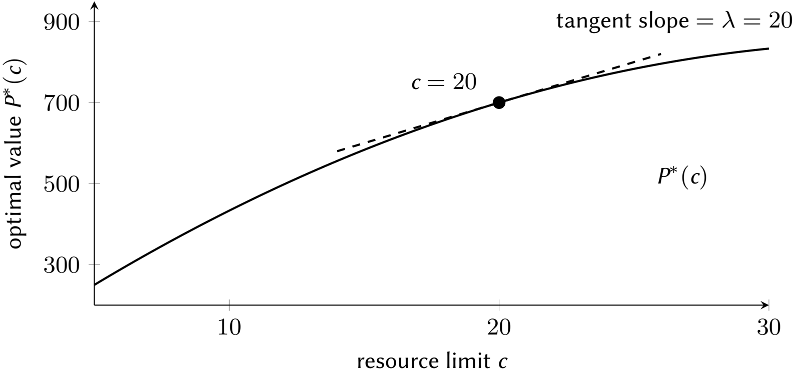
Concept: Multiplier as a shadow price

Write the resource constraint as $x+y=c$ and let $P^*(c)$ be the optimal value. Near $c=20$,

$$\frac{dP^*}{dc}=\lambda\approx 20.$$

Therefore one additional resource unit is worth about 20 benefit units locally. This does not mean every extra unit is always worth 20: the multiplier changes as the operating point changes.

Illustration: Shadow prices and a production-allocation model



The shadow price is the slope of the optimal-value curve with respect to the resource limit. It is a local sensitivity, not a permanent price.

Model Critique: What the optimization has not proved

Before recommending $(x, y) = (10, 10)$, check:

- Are x and y divisible, or must they be integers?
- Is the resource relation exactly $x + y = 20$, or only $x + y \leq 20$?
- Were benefit coefficients estimated from data, and how uncertain are they?
- Does the quadratic model remain realistic near the proposed solution?
- Are there omitted constraints such as minimum service levels or staff capacity?

A mathematically optimal point can still be a poor real-world recommendation if the model is incomplete.

Differentiated optimization studio

Core explanation, worked reasoning, and application

Activity: Core route: gradient and Hessian

For

$$f(x, y) = 2x^2 + y^2 - 8x + 4y + 10,$$

find the critical point, classify it using the Hessian, and write f in completed-square form to verify the classification.

Activity: Challenge route: constrained geometry

Find the maximum and minimum values of

$$f(x, y) = 3x + 4y$$

on the circle $x^2 + y^2 = 25$. Explain the answer using both Lagrange multipliers and the geometry of the dot product.

Activity: Extension route: modeling and sensitivity

A farm allocates land x to crop A and y to crop B. Profit is

$$\Pi(x, y) = 100x + 80y - x^2 - 2y^2,$$

with $x + y = 50$ and $x, y \geq 0$.

1. Find the continuous optimum and multiplier.
2. Interpret the multiplier with units.
3. Test what happens when the land limit changes from 50 to 51.
4. Name one assumption that could make the mathematical optimum impractical.

Communicating the result

Core explanation, worked reasoning, and application

Summary: Competition-style conclusion template

A strong conclusion should state:

1. **Decision:** the recommended variable values and objective value;
2. **Feasibility:** the constraints satisfied by the recommendation;
3. **Verification:** Hessian, concavity, boundary comparison, or numerical confirmation;
4. **Sensitivity:** the meaning of the multiplier or a small perturbation test;
5. **Limitation:** at least one assumption that may affect implementation.

Example: "Under the continuous quadratic model and the binding 20-unit resource constraint, the recommended allocation is $(10, 10)$, giving benefit 700. The negative-definite Hessian confirms global concavity. The multiplier 20 indicates that one additional resource unit is worth about 20 benefit units locally. Integer restrictions and parameter uncertainty should be checked before implementation."

Checkpoint: Exit ticket

In two sentences, explain the difference between $\nabla f = \mathbf{0}$ and $\nabla f = \lambda \nabla g$. Then state one reason why a Lagrange candidate may fail to be a valid final recommendation.

Practice problems

Core explanation, worked reasoning, and application

Exercises: Exercise 1: partial derivatives and local change

1. For $f(x, y) = x^2 y + e^{xy}$, compute f_x and f_y.
2. For $T(x, y) = 25 - 0.2x^2 - 0.5y^2$, find the gradient at $(2, 1)$ and the directional derivative toward $(5, 5)$.
3. Use local linearization to estimate $\sqrt{x^2 + y^2}$ near $(3, 4)$ when (x, y) changes to $(3.03, 3.98)$.
4. Explain the units of f_x and f_y for a cost function whose inputs are labour-hours and kilograms of material.

Exercises: Exercise 2: unconstrained optimization

Find and classify all critical points:

1. $f(x, y) = x^2 + 3y^2 - 4x + 6y$;
2. $g(x, y) = x^2 - y^2 + 2x$;
3. $h(x, y) = x^3 + y^3 - 3xy$;
4. $q(x, y) = x^4 + y^4 - 4xy$.

For each, explain whether the Hessian test is conclusive.

Exercises: Exercise 3: Lagrange multipliers

Use Lagrange multipliers and verify the result:

1. maximize xy subject to $x^2 + 4y^2 = 8$;
2. find the point on $x + 2y + 3z = 6$ closest to the origin;
3. maximize xyz subject to $x + y + z = 12$, with $x, y, z > 0$;
4. maximize $5x + 8y - x^2 - y^2$ subject to $2x + y = 12$ and $x, y \geq 0$.

Exercises: Exercise 4: modeling challenge

A manufacturer chooses two continuous production levels. Profit is

$$\Pi(x, y) = 90x + 70y - 1.5x^2 - y^2 - 0.5xy,$$

subject to $2x + y \leq 60$ and $x, y \geq 0$.

1. Find the unconstrained optimum and test whether it is feasible.
2. If the resource constraint binds, solve the equality-constrained problem.
3. Compare the objective values at all relevant candidates and boundaries.
4. Interpret the multiplier and discuss whether integer production changes the recommendation.
5. Write a five-sentence modeling conclusion using the template above.

Practice problems

Optional extension A: several constraints

Core explanation, worked reasoning, and application

Optional extension A: several constraints

For constraints $g_1(x) = c_1, \ldots, g_k(x) = c_k$, the necessary condition becomes

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_k \nabla g_k.$$

The constraint gradients must be linearly independent at the candidate. Each multiplier measures local sensitivity to its corresponding constraint level, subject to regularity conditions.

Optional extension B: inequality constraints

Core explanation, worked reasoning, and application

Optional extension B: inequality constraints

If the feasible set contains inequalities such as $g(x) \leq c$, the constraint may be active or inactive. When active, the solution can often be studied by setting $g = c$ and using Lagrange multipliers. A systematic treatment uses the Karush–Kuhn–Tucker conditions, which are beyond this supplementary. In a report, explicitly check whether the inequality binds rather than assuming equality from the start.

Writing Class 1: Paper Structure and Rubric Decoded

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Why This Class Exists

Core explanation, worked reasoning, and application

Why This Class Exists

By now you have spent months learning the mathematics of modeling: difference equations, regression, simulation, optimisation, dynamical systems. You are technically prepared. And yet, when judges sit down to evaluate your IMMC or HiMCM submission, they will not see your difference equation. They will see *your paper*.

A judge will spend approximately 5–10 minutes on the first read of your paper. In those minutes they decide whether to keep reading carefully, skim the rest, or move on. Within a 36-page submission, the first page (your **Summary**) will receive disproportionate attention. The mathematics behind it might be brilliant, but if the paper does not communicate that brilliance — clearly, in the structure judges expect — it will not be recognised.

This first class establishes:

- What the judges are scoring (their published rubric).
- The standard paper structure they expect.
- How to read a winning paper analytically, so you can imitate the moves that work.

The next four classes (W2–W5) drill into specific sections, language, and tools. Everything starts here.

Note: The Competition Reality

- **IMMC** (International Mathematical Modeling Challenge): teams of up to 4 students; 96 hours (Continental round) or 120 hours (International); papers up to ~ 25–30 pages; English required for International.
- **HiMCM** (High School Mathematical Contest in Modeling): teams of up to 4; 36 hours within a 14-day window; English; one of two problems (A or B).
- In both, the *paper itself* is what is judged — not your code, your simulations, or your conversations. Everything must reach the judge through the document.

The Standard Paper Structure

Core explanation, worked reasoning, and application

The Standard Paper Structure

There is a near-universal structure for modeling competition papers. It is essentially the structure of a scientific research paper, adapted for the time and scope of a competition.

Method: Method: Standard IMMC/HiMCM Paper Skeleton

1. **Title page** — problem identifier, team number, date.
2. **Summary** (1 page; sometimes called Executive Summary or Abstract). *The single most important page.*
3. **Table of Contents** (optional but common).
4. **Introduction / Restatement of the Problem**
   - Background / motivation
   - Problem clarification
   - Our approach (high-level)
5. **Assumptions and Justifications**
6. **Variables, Parameters, Notation** (table form)
7. **Model Development**
   - Model 1 (often the simplest baseline)
   - Model 2 (refinement / extension)
   - Each with derivation, equations, intuition
8. **Solution / Results**

Method: Method: Standard IMMC/HiMCM Paper Skeleton

- Analytical results (if any)
- Numerical results (with figures and tables)

9. **Model Validation**
   - Consistency checks, limiting-case checks
   - Comparison with data or with literature
10. **Sensitivity Analysis**
    - How does the answer change when key parameters or assumptions change?
11. **Strengths and Weaknesses**
12. **Conclusion and Future Work**
13. **References** (formatted consistently)
14. **Appendices** (code, derivations, supplementary figures)

Warning: Warning: The Order is Not Optional

Judges expect the sections in approximately this order. They use it to navigate quickly: when they want to check ”did you sanity-check the model?”, they jump to *Validation*. If you put validation inside Model Development, or split it across three places, they will conclude you did not understand the conventions — which is a worse signal than not validating at all.

You may merge adjacent sections (e.g., *Strengths and Weaknesses* inside *Conclusion*) or rename them slightly, but **do not invent your own structure**.

Method: Method: Section-by-Section Purpose Statement

1. **Summary:** convince the judge to read further. State problem, approach, results, in one page.
2. **Introduction:** prove you understood the problem. Restate the problem in your own words; do not copy.
3. **Assumptions:** prove you thought about the real world. List what you assumed and why it is reasonable.
4. **Variables:** let the judge read equations without flipping back. Tabulate every symbol.
5. **Model Development:** prove you built something non-trivial. Show derivation, not just final equations.
6. **Solution:** answer the actual question that was asked. Numbers, with units, in context.
7. **Validation:** prove the model is not nonsense. Limiting cases, data comparison.
8. **Sensitivity Analysis:** prove the answer is robust. Vary parameters, report effect on conclusions.
9. **Strengths and Weaknesses:** prove self-awareness. Critique your own work honestly.
10. **Conclusion:** answer the question one more time, in plain language.

The Rubric: How Judges Actually Score You

Core explanation, worked reasoning, and application

The Rubric: How Judges Actually Score You

Both IMMC and HiMCM publish (or have judges informally communicate) the criteria. The exact wording differs by year and contest, but the four pillars are stable.

Method: Method: The Four Pillars of Modeling Paper Evaluation

1. **Modeling Approach** — Is the modeling thought process clear? Are assumptions reasonable? Are the chosen methods appropriate? Is there originality or insight?
2. **Mathematical Sophistication** — Is the mathematics correct and at an appropriate level? Are the techniques used well? Does the team demonstrate command of the methods?
3. **Validation and Sensitivity** — Does the team check whether their model makes sense? Do they report how sensitive results are to inputs?
4. **Communication** — Is the paper clearly written, well organised, properly figured and cited? Is the Summary informative? Do the figures clarify?

Note: A Realistic Weighting (Approximate)

Although exact weights vary, a useful mental model:

Pillar	Approx. weight
Modeling Approach	25–30%
Mathematical Sophistication	25–30%
Validation & Sensitivity	15–20%
Communication (Writing)	25–30%

Communication is roughly equal to all the mathematics. This is the central insight that surprises most newcomers, and the reason for this entire Writing Class.

Method: Method: Things That Lose Points (and Why)

1. **Modeling Approach failures**
    - Skipping the simplest model and jumping straight to a fancy one.
    - Treating assumptions as filler ("we assume the world is real") rather than as genuine modeling choices.
    - No clear connection between assumptions, model, and the question being answered.
2. **Mathematical Sophistication failures**
    - Equations appear without derivation or motivation.
    - Using a "big" method (neural network, fuzzy logic, AHP) without justifying why a simpler method would not work.
    - Numerical errors in worked examples that propagate through the paper.
3. **Validation & Sensitivity failures**
    - Validation section omitted entirely (alarmingly common).
    - "Sensitivity analysis" that is one paragraph of words with no numbers.
    - Reporting only that the result is "robust" without saying *to what* and *by how much*.
4. **Communication failures**

Method: Method: Things That Lose Points (and Why)

- Summary that says what you *did* but not what you *found*.
- Equations without numbered references; figures without captions; tables without units.
- Walls of text with no headings; pages of code in the body of the paper.
- Inconsistent notation (using r for radius in one section and growth rate in the next).

Anatomy of a Winning Paper

Core explanation, worked reasoning, and application

Anatomy of a Winning Paper

Let us walk through how the four pillars apply, using a concrete (illustrative) example. Imagine the problem: *model the throughput of a single-server queue at a coffee shop and recommend a staffing strategy.*

Worked Example: Example 1: A Strong Modeling Approach Section (Sketch)

We begin with the simplest defensible model and progressively refine.

Model 1 (baseline): Assume customers arrive deterministically every 5 minutes and service takes exactly 4 minutes. Then no queue ever forms, and the throughput is 12 customers/hour. This baseline gives a sanity-check ceiling.
Model 2 (stochastic arrivals): Customers arrive as a Poisson process with rate λ. Service times remain deterministic at 4 minutes. Under this assumption, queues form occasionally; we use simulation to estimate average wait.
Model 3 (full M/M/1): Service times exponential with rate $\mu = 1/4$ per minute. Now both standard queuing theory results and simulation apply, and the two can validate each other.

Note: Why this Modeling Approach is strong

- Starts simple; each model is motivated by what the previous missed.
- Each model is a *stepping stone*, not a side trip.
- Validation is baked in: Model 1 gives a sanity ceiling; Models 2 and 3 cross-check.

Worked Example: Example 2: A Strong Assumptions Section (Sketch)

Assumption 1. Customer arrivals follow a Poisson process with rate λ.
Justification: Coffee shops in similar settings have been modeled this way (Larson & Odoni, 1981), and the assumption of independent arrivals is reasonable for unaffiliated customers at random times.
Effect on model: Allows use of standard queuing results and inter-arrival time generation by inverse-transform sampling.
Assumption 2. Service times are exponentially distributed with rate μ.
Justification: Choice driven by the memorylessness property; arguably less realistic than a deterministic time but vastly easier to analyse.
Effect: The system becomes an $M/M/1$ queue. We will check the result against simulation with deterministic service in §6.
Assumption 3. The shop operates from 7:00 to 10:00 with constant λ.
Justification: Simplifies analysis at the cost of accuracy; in reality, λ peaks at 8:00.
Effect: Our results give an average estimate; peak conditions may need separate treatment.

Note: Why this Assumptions section is strong

Each assumption has three parts:

1. **What** is assumed.
2. **Why** it is reasonable (citation, simplification, or domain knowledge).
3. **What effect** it has on the model.

Many teams write only (1). A judge can tell instantly.

Worked Example: Example 3: A Strong Sensitivity Analysis (Sketch)

We test sensitivity of the predicted average wait time to three parameters:

- Arrival rate λ: vary from $\pm 20\%$ of nominal.
- Service rate μ: vary $\pm 20\%$.
- Assumption of exponential service: replace with deterministic service.

Findings.

Scenario	λ	μ	Avg wait (min)	Change
Baseline	12/hr	15/hr	4.00	—
Higher λ	14.4/hr	15/hr	9.60	+140%
Lower μ	12/hr	12/hr	∞	system unstable
Higher μ	12/hr	18/hr	1.33	-67%
Deterministic service	12/hr	15/hr	2.00	-50%

Worked Example: Example 3: A Strong Sensitivity Analysis (Sketch)

Interpretation. The model is highly sensitive to λ and μ near the stability boundary $\lambda \approx \mu$. The deterministic-service comparison shows our exponential-service assumption *overstates* the queue by roughly 50%, so our recommendation will be conservative.

Note: Why this Sensitivity Analysis is strong

- Identifies *which* parameters or assumptions matter.
- Reports actual numbers, not just words.
- Connects findings back to the conclusion (”our recommendation is conservative”).
- Includes a structural sensitivity (the exponential vs deterministic comparison) — not just numerical perturbation.

Reading a Real Outstanding Paper

Core explanation, worked reasoning, and application

Reading a Real Outstanding Paper

The single best preparation for writing well is to read good papers analytically. COMAP publishes **Outstanding Papers** from HiMCM and MCM annually; IMMC publishes its top-scoring papers as well. Reading three or four of these — not for content but for structure — is more valuable than reading any guide, including this one.

Method: Method: How to Read an Outstanding Paper

Allocate 90 minutes per paper. Read with a notebook and answer these questions:

1. **Summary** (5 min): Does it tell you the problem, the approach, the result, and the key insight, in roughly that order? How long is each part?
2. **Section Headers** (5 min): What is their section structure? How does it differ from the skeleton in §2 of this class?
3. **Assumptions** (10 min): How many do they list? How are they justified? Do they explicitly call out the effect of each assumption?
4. **Model Development** (20 min): How do they introduce equations? Always with words first? With derivations or just statements? How do they use figures?
5. **Results** (15 min): Where do the numerical answers appear? Are they highlighted (boxed, bolded, or in a result table)?
6. **Validation & Sensitivity** (15 min): How much space is allocated? What types of checks did they perform?
7. **Writing Style** (10 min): Average paragraph length? Active or passive voice? Use of ”we”? Tone?

Method: Method: How to Read an Outstanding Paper

8. **Figures & Tables** (10 min): How many? Captioned? Referenced from the text? Use of colour, fonts, captions?

Build a one-page **template extract** based on your findings, then compare across multiple papers to find the patterns.

Note: Where to Find Outstanding Papers

- **COMAP.** The COMAP website publishes Outstanding papers for HiMCM and MCM each year. They are freely accessible.
- **UMAP Journal.** The *UMAP Journal* prints selected Outstanding papers with judges' commentaries — the commentaries are particularly valuable.
- **IMMC Chinese region.** The IMMC China Hong Kong page publishes lists and sometimes selected papers from the Continental round.
- **Caution.** Older papers (pre-2010) have different writing conventions; focus on the last 5 years.

Common Mistakes Made by First-Time Teams

Core explanation, worked reasoning, and application

Warning: Warning: The Mistakes That Most Hurt First-Time Teams

Read these now. Re-read them the day before submission.

1. **Writing the Summary last (and rushing it).** The Summary determines the first impression; treat it like the second-most-important section, not an afterthought.
2. **Copying the problem statement instead of restating it.** Judges read this immediately and conclude you did not engage. Rewrite the problem in your own words.
3. **Listing assumptions without effects.** "We assume the Earth is flat" is not an assumption; an assumption needs a reason and a consequence.
4. **Hiding the answer.** The numerical recommendation or final result must be **boxed, bolded, or otherwise impossible to miss**. Many strong papers fail because the judge cannot find the answer.
5. **No connection between models.** If you have three models, explain how each relates to the others. Are they alternatives? Refinements? Are results consistent?
6. **Equations without numbers.** Number every important equation. Reference them by number from the text.
7. **Figures without captions.** Every figure needs a caption that stands on its own; the reader should not need the body text to understand the figure.

Warning: Warning: The Mistakes That Most Hurt First-Time Teams

8. **Code in the body.** Code goes in the Appendix. The body text describes *what* the code does and *why*; the code itself is supporting material.
9. **Skipping Sensitivity Analysis ”because we ran out of time”.** Better to truncate Model Development than to skip Sensitivity. Judges flag the omission immediately.
10. **No references.** A paper with zero citations signals isolation from the literature. Even 3–4 citations to standard textbooks (Giordano; Hillier & Lieberman; Bertsekas) is dramatically better than none.

Workshop Exercises

Core explanation, worked reasoning, and application

Exercises: Exercise W1.1: Diagnose a Section

Below is a one-paragraph Assumptions section. Identify all the problems and rewrite it in the three-part format from Example 2.

> *We assume that the population is constant and there is no migration. We also assume the disease is contagious and people get sick from each other. We assume the recovery rate is constant and people can recover. There is no death due to other causes.*

Exercises: Exercise W1.2: Identify the Pillar Failures

For each of the following short excerpts from imagined papers, identify which of the four pillars (Modeling, Math, Validation, Communication) is being undermined and explain why.

1. "We applied a fuzzy-logic neural network with 24 hidden layers to determine the optimal car park entry fee."
2. "Our model predicts the answer is $\$3.47$ per hour. This concludes our analysis."
3. "Eq. 12 gives the average waiting time, where everything is as defined above."
4. "After running our simulation, we conclude that the queue length is robust to all parameter changes."
5. "We assume the price elasticity is -0.5. We assume the population grows. We assume the year has 365 days."

Exercises: Exercise W1.3: Skeleton Practice

Take any of the following short modeling questions. For each, sketch a section-by-section outline (1–2 sentences per section) following the standard skeleton from §2. *Do not solve the problem.* Just structure how you would present a solution.

1. How should a school schedule its bus routes to minimise total travel time?
2. How long should a stoplight stay green at a single-intersection city centre?
3. What is the carrying capacity of a small island for a deer population?

Exercises: Exercise W1.4: Outstanding Paper Reading

Choose one Outstanding paper from HiMCM (any year 2020 or later) from the COMAP website. Spend 90 minutes reading it analytically per Method 5.1. Produce:

- A bullet-list of the section headers (with subsections).
- Counts: number of equations, number of figures, number of tables, number of references.
- A one-paragraph summary of how the team handled Sensitivity Analysis.
- A one-paragraph judgment: what one move from this paper will you copy in your own?

What to Carry Into W2

Core explanation, worked reasoning, and application

Note: Class W1 Take-aways

1. Papers are scored on four pillars; **Communication is roughly** $25–30\%$ — equal to all the mathematics.
2. The standard structure is rigid; do not innovate at the section-organisation level.
3. Every section has a **specific purpose**; if you cannot state in one sentence what a paragraph contributes, cut it.
4. Assumptions need three parts: *what*, *why reasonable*, *effect on model*.
5. Sensitivity must include numbers, not just words.
6. Reading 3–4 Outstanding papers is the highest-leverage preparation you can do.

What to Carry Into W2

In **W2: The Summary** we attack the single most important page of your paper: the Summary. We will look at what a great Summary looks like, decompose its standard four-block structure, and write three versions of a Summary against the same problem.

Appendix A: One-Page Cheat Sheet

Core explanation, worked reasoning, and application

Appendix A: One-Page Cheat Sheet

Section	One-sentence purpose
Summary	Convince the judge your paper is worth reading.
Introduction	Show you understood the question.
Assumptions	Show you thought about the real world.
Variables	Let the judge read equations smoothly.
Model	Show you built something non-trivial.
Solution	Answer the question with numbers, units, context.
Validation	Show the model is not nonsense.
Sensitivity	Show the answer is robust.
Strengths/Weaknesses	Show self-awareness.
Conclusion	Answer one more time, in plain language.
References	Show grounding in the literature.
Appendix	Code, derivations, supplementary figures.

Appendix B: Pre-Submission Checklist

Core explanation, worked reasoning, and application

Appendix B: Pre-Submission Checklist

- Summary fits on **one page** and follows a clear four-block structure (Problem → Approach → Results → Insight).
- Every assumption has *what*, *why*, *effect*.
- Every variable used in an equation is in the Variables table.
- Every equation is numbered.
- Every figure has a caption that stands alone.
- Every table has a title and units in column headers.
- Validation section is present and includes at least *one* concrete check (limiting case, data comparison, or alternative-model comparison).
- Sensitivity Analysis includes a *table* with parameter changes and result changes in numbers.
- Strengths and Weaknesses are both present; at least two of each.
- Conclusion answers the actual question asked.
- References use a consistent format (APA, IEEE, or similar). At least 4 entries.
- Code is in the Appendix, not in the body.
- Page numbers visible, team number on every page (if required by contest).

Writing Class 2: The Summary: The Most Important Page

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Why the Summary Decides Everything

Core explanation, worked reasoning, and application

Why the Summary Decides Everything

In W1 we noted that judges spend disproportionate time on the first page. Now we say it more directly: the Summary often determines whether your paper is read carefully at all.
Consider the judging process for a major competition. A single judge may evaluate 30–50 papers in a few days. They cannot read every page of every paper with equal attention. What they do is the following:

- Read the Summary carefully (5–10 minutes).
- Based on the Summary, form a tentative ranking.
- Skim the rest of the paper to confirm or revise.
- Final ranking is heavily anchored by Summary impressions.

This is not laziness. It is the only feasible procedure given time constraints. The Summary serves as both an **advertisement** and a **filter**: it advertises what your paper accomplishes and filters whether the rest deserves close attention.

Note: The Disproportionate-Effort Rule

In a 36-page paper, the Summary is approximately 2.7% of the length but plausibly determines $40–60\%$ of the final outcome. Spend a corresponding share of your writing time on it. Many teams spend 30 minutes on the Summary at the end of the 96-hour competition. **This is the single most leverage-rich mistake to fix.**

The Four-Block Structure

Core explanation, worked reasoning, and application

The Four-Block Structure

Although every Summary is different, almost every *good* Summary follows a four-block structure. Internalising this structure is the single most useful Summary-writing skill.

Method: Method: The Four-Block Summary Structure

1. **Problem block** (2–4 sentences): What is the problem? Why does it matter? What is asked of you specifically?
2. **Approach block** (3–5 sentences): What is your approach? Which models did you build, in what order, why? Mention 2–3 specific techniques.
3. **Results block** (3–5 sentences): What did you find? Give *numbers*. Make the headline answer impossible to miss.
4. **Insight / Recommendation block** (2–3 sentences): What does this mean? What is your recommendation? Why should the reader believe you?

Total: 10–17 sentences, fitting on one page with single spacing or comfortable double spacing.

Note: A Useful Diagnostic

After drafting your Summary, label every sentence with one of P (Problem), A (Approach), R (Result), I (Insight). The labels should appear roughly in this order. If your Summary is mostly A's with no R's, you have described *what you did* but not *what you found* — a fatal flaw.

Block 1: Problem

This block proves you understood the question. Three moves to make:

- Briefly establish **context** (one sentence).
- State the **specific question** asked (one sentence; in your words, not theirs).
- Mention what makes the problem **interesting or difficult** (one sentence).

Worked Example: Example 1: Problem Block — HiMCM-Style Problem on Drone Light Shows

Weak:

We were asked to figure out how to make a drone light show. The problem is from HiMCM 2024.

Better:

Drone light shows are increasingly replacing fireworks for large public celebrations because they produce no air or sound pollution and allow programmable imagery. The central modeling challenge is to determine how many drones are needed to render a target image at a given altitude and viewing angle, and how to assign each drone an efficient flight path that avoids collisions. This combination of geometric resolution and combinatorial routing is what makes the problem non-trivial.

Note: What Made the Second Version Stronger

- Opens with context (drones replace fireworks) before diving in.
- Restates the question in modeling language (‘geometric resolution and combinatorial routing’).
- Names what makes it hard, signalling that the team thought about the problem rather than just receiving it.
- Does not mention the contest year — judges know what year it is.

Block 2: Approach

This block proves you built something. Three moves to make:

- State your **strategy** at one level of abstraction higher than the equations (one sentence).
- Briefly describe **each model** you built and how they relate (two or three sentences).
- Indicate **which techniques** were used (one or two sentences naming concrete methods).

Worked Example: Example 2: Approach Block

Weak:

> *We used MATLAB and Python to make a model. We tried different things and picked the best one. Our model uses lots of equations and computer simulations to find the answer.*

Better:

> *We approach the problem in three stages of increasing realism. Model 1 computes the minimum number of drones needed to render a target image at a fixed viewing distance, using a discrete image-sampling argument. Model 2 extends this to allow drones to move during the display, transforming the question into a min-cost assignment problem solvable by the Hungarian algorithm. Model 3 adds collision-avoidance via a velocity-obstacle method validated by Monte Carlo simulation of* 10^4 *random initial conditions.*

Note: Why the Second Version Works

- Numbered models give a clear progression and let the reader hold them in memory.
- Specific algorithms (Hungarian, velocity-obstacle, Monte Carlo) signal mathematical depth without losing the reader.
- The number 10^4 commits to a concrete simulation effort.
- Every claim is something the body of the paper must defend; do not promise more than you delivered.

Block 3: Results

This block delivers the **headline answer**. The single most important sentence of your entire paper lives here. It must:

- Be a **quantitative claim with units**.
- Be the **first thing** stated in the block, not buried after methodology recap.
- Be **visually emphasised** where appropriate (bold, italics, or a result table).

Worked Example: Example 3: Results Block — the Headline Sentence

Weak (headline buried):

> *After running our simulations and analysing the results, we found values for various parameters and arrived at a recommendation regarding the optimal configuration.*

Better (headline first):

> *We recommend deploying 312 drones, with assigned paths totalling 2,840 metres and a maximum collision rate of* 0.02% *across* 10^4 *simulated runs. Our two upper bounds (a geometric ceiling of 350 drones and an information-theoretic ceiling of 295) bracket this recommendation and indicate it is within* 6% *of the theoretical optimum.*

Warning: Warning: Hide Nothing in the Results

A judge should be able to find your answer in fewer than 10 seconds by glancing at the Summary. If they have to read three paragraphs and still are not sure what the answer is, the Summary has failed — regardless of how clever the model is.

Block 4: Insight / Recommendation

This block elevates the work from a numerical exercise to a defensible recommendation. Two moves:

- State **why** the recommendation is trustworthy (cite validation or sensitivity briefly).
- State the **broader implication** or recommendation in plain language.

Worked Example: Example 4: Insight Block

Weak:

> *In conclusion, we believe our model gives a good answer to the problem. Future work could improve the model.*

Better:

> *Sensitivity analysis shows the recommended fleet size is robust to* $\pm 15\%$ *variation in audience-distance assumptions, and that the dominant constraint at scale is communication-latency, not propulsion. Event organisers planning shows of similar scope should therefore allocate budget toward redundant communication channels rather than additional drones.*

A Complete Worked Summary

Core explanation, worked reasoning, and application

A Complete Worked Summary

Below is a full Summary against the (imagined) problem: *Model the effect of installing a network of bicycle highways on commute times in a mid-sized city.*

Worked Example: Example 5: Complete Summary — Bicycle Highway Network

PROBLEM. As mid-sized cities seek to reduce traffic congestion and emissions, several have begun investing in dedicated bicycle highways — continuous, signal-free, motor-traffic-free bike paths between major centres. We are asked to determine how such a network should be designed for a city of $200{,}000$ residents, and to estimate the resulting reduction in average commute time. The problem combines network design, mode-choice modeling, and uncertainty in commuter behaviour.

APPROACH. We attack the problem in three stages. **Model 1** treats the city as a graph and solves a minimum-spanning-network problem to identify the most cost-effective set of highway routes connecting the eight largest residential clusters to the central business district. **Model 2** introduces a **discrete-choice mode model** (logit) calibrated against transport-survey data, predicting what fraction of commuters will switch from cars to bicycles given a new highway. **Model 3** feeds the predicted switch rates back into a **traffic flow model**, allowing us to estimate residual car congestion and overall commute time savings. We validate against published European deployments (Copenhagen, Utrecht).

RESULTS. Our recommended network consists of **four bicycle highways totalling 47 km**, built at an estimated cost of **EUR 28 million**. This network is predicted to shift **18%** of car-based commuters to bicycles, reducing peak-hour car traffic by **14%** and the average commute time by **6.2 minutes per traveller**. Sensitivity

Worked Example: Example 5: Complete Summary — Bicycle Highway Network

analysis indicates the modal shift estimate is most sensitive to weather assumptions (range $14–22\%$) and least sensitive to construction cost; the savings recommendation is therefore robust.

INSIGHT. The break-even period of the investment is approximately **4.3 years** on commute-time savings alone, before considering health and emissions benefits. City planners should prioritise highways connecting the two highest-density residential clusters to the city centre first, as these account for 62% of total time savings in our model. Our results are consistent with the European data and suggest that bicycle highways are an unusually high-return urban transportation investment for mid-sized cities.

Note: How to Read This Summary

- **Problem block:** 3 sentences. Context (mid-sized cities reducing emissions), question (network design + commute reduction), difficulty (network + mode choice + behaviour).
- **Approach block:** 5 sentences. Three numbered models, two specific techniques named (logit, traffic flow), validation source mentioned.
- **Results block:** 4 sentences. *Five* bolded quantitative claims. Sensitivity touched.
- **Insight block:** 3 sentences. Break-even period, prioritisation, broader implication.

Total: 15 sentences. Fits on one page.

Diagnosing Common Summary Failures

Core explanation, worked reasoning, and application

Diagnosing Common Summary Failures

We now look at the four most common failure patterns, each with a real-style example and a fix.

Worked Example: Failure 1: The Methods-Heavy Summary (most common)

Failure:

> *We used a system of differential equations together with the Runge-Kutta method to solve them. We also performed a Monte Carlo simulation with 10,000 trials. The results were then analysed using regression and the model was validated.*

Diagnosis: The reader learns *nothing* about the problem or the answer. This is an Approach block with no Problem, Results, or Insight.

Fix: Move this content into Block 2 (Approach), shrink it to one or two sentences, and add Blocks 1, 3, and 4.

Worked Example: Failure 2: The Process Narrative

Failure:

> *Our team first read the problem and discussed possible approaches. We then divided the work and each member tackled a different aspect. After several iterations, we converged on a final model, which we then refined to produce our recommendations.*

Diagnosis: The reader learns about the team's *workflow*, not about the modeling. This is almost always the result of writing the Summary last and confusing it with a reflection or acknowledgement.

Fix: Delete the entire paragraph. Replace with content about the problem, model, and findings.

Worked Example: Failure 3: The Buried Answer

Failure:

> *This paper investigates several models for the optimal coffee-shop staffing problem. After developing three models, performing sensitivity analysis on key parameters, validating our results against the literature, and considering implementation constraints, we have produced a final recommendation that aligns with the expectations of the operational research literature.*

Diagnosis: 50 words, zero numbers. No answer.

Fix: The very next sentence must be: *The optimal staffing level is 3 baristas during peak hours and 1 during off-peak.*

Worked Example: Failure 4: The Hyped Summary

Failure:

> *Our novel and innovative model uses state-of-the-art techniques including deep learning, fuzzy systems, and genetic algorithms to perfectly solve the problem. Our results are highly accurate and our recommendations are guaranteed to optimise the situation.*

Diagnosis: Adjectives doing the work that numbers should do. Triggers immediate skepticism in judges, who have seen exactly this kind of language fail many times.

Fix: Replace every adjective with a number. ”novel” $\rightarrow$ describe what makes the approach different in one sentence. ”state-of-the-art” $\rightarrow$ drop. ”perfectly solve” $\rightarrow$ ”within 5% of the theoretical optimum”. ”highly accurate” $\rightarrow$ ”validated against 12 data points with mean error 4%”.

Writing Process: Drafting the Summary

Core explanation, worked reasoning, and application

Method: Method: When and How to Write the Summary

1. **First draft: end of Day 1** (or as soon as you have a working baseline model). This forces you to articulate the problem and approach before the details overwhelm.
2. **Second draft: end of Day 2/3** after results are available. Replace placeholder numbers with actual numbers.
3. **Third draft: in the last 4–6 hours.** Polish for language. Read aloud. Ask: would a stranger find the answer in 10 seconds?
4. **Final test:** have a teammate who did not write the Summary read it. Ask them: (a) what is the question? (b) what is the answer? (c) why should we trust it? If they cannot answer all three, revise.

Warning: Warning: Common Process Mistakes

- **Writing the Summary only at the very end.** The Summary written at hour 94 of 96 is almost always weak.
- **Letting one person write the whole Summary alone.** The Summary is too important; at minimum, two teammates should review.
- **Copying sentences from the body.** The Summary is its own document. Its sentences are tighter, more declarative, more headline-like than body text.

Method: Method: Write Three Versions, Pick the Best Pieces

For a high-stakes Summary, write three full drafts:

1. **Version A: methods-first.** Lead with the Approach. *Strength:* shows mathematical depth early.
2. **Version B: results-first.** Lead with the Result. *Strength:* maximum impact; judges immediately know your answer.
3. **Version C: problem-first.** Lead with the Problem (classic structure). *Strength:* demonstrates understanding; smooth read.

Read all three. Pick the best Problem block, the best Approach block, the best Results block, and the best Insight block, then assemble. The final draft is usually closest to C, with the Results block from B.

Workshop Exercises

Core explanation, worked reasoning, and application

Exercises: Exercise W2.1: Label the Sentences

Take any of the example Summaries in this class. For every sentence, label P (Problem), A (Approach), R (Result), or I (Insight). Then count the sentences in each block. Compare against the recommended 2–4, 3–5, 3–5, 2–3.

Exercises: Exercise W2.2: Diagnose and Rewrite

The following is a (constructed) weak Summary. Diagnose every problem and rewrite it as a strong four-block Summary.

> *Our team studied the problem of optimal bus routing for our school district. We developed a model using mathematical techniques. After many iterations, we determined that our model works well and produces results that are useful for the school district. We hope our recommendations will be considered. Future work could extend our approach.*

Assume (for the rewrite) that the actual findings were: 4 buses needed, total daily route length 87 km, average student commute reduced from 43 to 28 minutes, sensitivity test shows the recommendation is stable to ± 10 students per route.

Choose any modeling problem from your past coursework (or a HiMCM problem you have read). Write three Summary drafts of approximately 15 sentences each, following Method 5.1 (methods-first, results-first, problem-first). Have a teammate read all three blindly and rank them.

Exercises: Exercise W2.4: The 10-Second Test

Find one Outstanding paper online. Open it to the Summary page. Time yourself: how long does it take you to identify the team's headline numerical answer? Now do the same with your most recent Summary. If yours is slower, redesign.

Exercises: Exercise W2.5: The Length-Cut Drill

Take your most recent Summary draft. Without losing any content, cut it by 30%. Methods to try:

- Replace verbose phrases ("In order to do this, we then performed an analysis of...") with terse equivalents ("We analysed...").
- Remove every adjective that does not earn its place ("complex", "novel", "innovative", "comprehensive").
- Combine adjacent sentences with semicolons or commas where appropriate.
- Move detail to the body. The Summary is not the place for derivations.

A good Summary at 400 words is usually better than the same content at 600.

What to Carry Into W3

Core explanation, worked reasoning, and application

Note: Class W2 Take-aways

1. The Summary disproportionately determines your score.
2. Use the four-block structure: Problem $\rightarrow$ Approach $\rightarrow$ Results $\rightarrow$ Insight.
3. The single most important sentence is the headline Results sentence. Make it impossible to miss.
4. Adjectives do not score points. *Numbers* score points.
5. Write the first Summary draft early; iterate at least three times.
6. Two writers always beat one writer on a high-stakes Summary.

What to Carry Into W3

In **W3: Assumptions, Validation, and Sensitivity Analysis** we attack the three sections that distinguish a thoughtful paper from a mathematical exercise. These are sections that average teams skim and great teams use to separate themselves.

Appendix A: Summary One-Page Template

Core explanation, worked reasoning, and application

Appendix A: Summary template (1/2)

Summary

[Problem block: 2–4 sentences]

Context.

Specific question asked.

Why the problem is interesting/hard.

[Approach block: 3–5 sentences]

One-sentence strategy at a high level.

Model 1 (one sentence, named technique).

Model 2 (one sentence, refinement of Model 1).

Model 3 if any.

Validation source briefly mentioned.

Appendix A: Summary template (2/2)

Summary (continued)

[Results block: 3–5 sentences]

Headline numerical answer.

Supporting numerical findings.

One-line sensitivity note.

[Insight block: 2–3 sentences]

Why the result is trustworthy (validation/sensitivity).

Broader recommendation.

Appendix B: Headline-Sentence Templates

Core explanation, worked reasoning, and application

Appendix B: Headline-Sentence Templates

- **Recommendation form:** *We recommend [doing X], using [N units of resource Y], at an estimated [cost/benefit Z].*
- **Estimate form:** *Our model predicts [quantity] = [value with units], with [error margin] across [scenario set].*
- **Comparison form:** *[Option A] outperforms [Option B] by [X%] on [metric], assuming [conditions].*
- **Conditional form:** *If [parameter] exceeds [threshold], we recommend [action 1]; otherwise [action 2].*

Use a template; modify; do not invent from scratch under competition pressure.

Appendix C: 25 Adjectives to Avoid in Summaries

Core explanation, worked reasoning, and application

Appendix C: 25 Adjectives to Avoid in Summaries

- novel
- innovative
- state-of-the-art
- cutting-edge
- revolutionary
- powerful
- robust (without qualifier)
- accurate (without number)
- efficient (without number)
- complex
- sophisticated
- comprehensive
- rigorous (without proof)
- thorough
- extensive
- insightful
- meaningful
- significant (in non-statistical sense)
- optimal (without proof)
- perfect
- ideal
- best-in-class
- world-class
- ground-breaking
- seamless

Appendix C: 25 Adjectives to Avoid in Summaries

Replace each with either (a) a number or (b) a verb. "Our novel approach..." $\rightarrow$ "We combine X and Y in a way not previously applied to Z."

Writing Class 3: Assumptions, Validation, and Sensitivity Analysis
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

Why These Three Sections Win Papers

Core explanation, worked reasoning, and application

Why These Three Sections Win Papers

Assumptions, Validation, and Sensitivity Analysis are the three sections that:

- Together count for 15–20% of the score (the *Validation & Sensitivity* pillar).
- Are the most commonly weak in first-time-team papers.
- Are the cheapest to fix — they require almost no additional mathematics, only structured thinking and careful writing.

The combined leverage is large. A team that can write these three sections at **Outstanding-paper quality** can move from Honourable Mention to Meritorious, or from Meritorious to Outstanding, on these alone.

Note: A useful frame: the three trust questions

Judges read these sections to answer three trust questions about your work:

1. **Assumptions:** *Did you think about the real world before you wrote equations?*
2. **Validation:** *Are your results plausible at all, or are they nonsense in disguise?*
3. **Sensitivity:** *Are your results stable enough to act on, or are they an artifact of your parameter choices?*

A confident ”yes” on all three is the signal that distinguishes a publication-quality paper from a homework exercise.

Assumptions: The Three-Part Structure

Core explanation, worked reasoning, and application

Assumptions: The Three-Part Structure

We met this structure briefly in W1. Here we develop it fully.

Method: Method: The Three-Part Assumption

Every assumption in your paper should have three parts:

1. **Statement (What).** What exactly are you assuming? Be specific; an assumption is a claim that can be true or false.
2. **Justification (Why reasonable).** Why is this assumption appropriate? Cite literature, domain knowledge, simplification rationale, or empirical observation.
3. **Implication (What it does for the model).** What does this assumption let you do mathematically? What does it cost you in realism?

Worked Example: Example 1: Three-Part Assumptions for an SIR Epidemic Model

Assumption 1. The total population $N = S + I + R$ remains constant during the modeled period.
Justification: The disease in question (a seasonal influenza) is rarely fatal, and the modeled period is two months, far shorter than demographic timescales. Births, deaths from other causes, and migration thus produce $< 0.5\%$ change in N and can be neglected.
Implication: Reduces the system from three coupled differential equations to two ($S' = -\beta SI/N$, $I' = \beta SI/N - \gamma I$, with $R = N - S - I$). This is the standard SIR reduction.
Assumption 2. The population mixes homogeneously: every infected person is equally likely to contact every susceptible person.
Justification: For a single school community of $\sim 1{,}000$ students in continuous daily contact, this is approximately reasonable, though weaker than for general populations. Network-based models would refine this.
Implication: Allows the bilinear infection term $\beta SI/N$. Without it, we would need a contact matrix or network model, increasing model complexity by an order of magnitude.
Assumption 3. The recovery rate γ is constant (i.e., individuals stay infectious for an exponentially distributed duration with mean $1/\gamma$).

Worked Example: Example 1: Three-Part Assumptions for an SIR Epidemic Model

Justification: Empirical data on influenza recovery times approximate a gamma distribution; the exponential is a one-parameter simplification. The mean is well-measured at approximately 5 days, so $\gamma \approx 0.2$/day.
Implication: Makes I's dynamics first-order, supporting closed-form expressions for the basic reproduction number $R_0 = \beta/\gamma$. A more realistic but more complex alternative would be a gamma-distributed infectious period.

Note: Format conventions for the Assumptions section

- **Number** the assumptions: "Assumption 1.", "Assumption 2.", etc.
- **Bold or italicise** the three parts (Statement, Justification, Implication) for visual scanning.
- Aim for **4–7 assumptions** for a typical paper. Fewer than 4 suggests you have not thought about the modeling; more than 8 suggests you are padding.
- Order assumptions from **strongest to weakest** (most certain first). Judges read down the list with diminishing patience; do not bury your best on page 4.

Worked Example: Example 2: Strong vs Weak Assumptions

Weak: *We assume our model is reasonable.*
Diagnosis: Not a claim. Says nothing.
Weak: *We assume there are 7 days in a week.*
Diagnosis: Trivially true. Wastes space and signals padding.
Weak: *We assume the function is differentiable, integrable, and continuous.*
Diagnosis: Mathematical convenience assumptions stated together. Not a modeling assumption.
Strong: *We assume the daily car arrivals at the toll booth follow a Poisson process. This is supported by the chi-squared goodness-of-fit test ($p = 0.34$) on the 30-day data from the highway authority, and it lets us use exponential inter-arrival times and standard $M/M/1$ queue results.*
Diagnosis: Specific, justified, and connected to model consequence. ✓

Warning: Warning: Distinguish Assumptions from Definitions

- "We define the utilisation as $\rho = \lambda/\mu$" is a *definition*, not an assumption.
- "We assume utilisation is below 1" is a *model constraint*, not really an assumption either (it is needed for stability).
- "We assume that customers' arrival pattern is independent of the service rate" *is* an assumption, because it could be empirically false (longer waits might deter arrivals).

Definitions go in the Variables/Notation section. Assumptions are claims about the world that you could in principle test.

Method: Method: A Checklist of Assumption Categories

Use this list to make sure you have not omitted a relevant assumption:

1. **Boundary / scope:** What system, time period, geography are we considering?
2. **Conservation:** What stays constant? (population, energy, money, mass)
3. **Homogeneity:** Are individuals/units treated as interchangeable?
4. **Independence:** Are events / individuals statistically independent?
5. **Distributional:** What probability distributions are assumed?
6. **Functional form:** Linear, logistic, exponential growth?
7. **Stationarity:** Do parameters change over time?
8. **Equilibrium:** Are we at steady state, or in transient?
9. **Behavioural:** How do humans/agents respond? (Rational? Myopic?)
10. **Data quality:** Is the input data accurate, complete, representative?

Validation: Proving the Model is Not Nonsense

Core explanation, worked reasoning, and application

Validation: Proving the Model is Not Nonsense

Validation answers: *Is your model output plausible?* It is distinct from sensitivity analysis (which asks how output changes with input). Validation comes *before* sensitivity in a well-written paper.

Method: Method: The Five Standard Validation Checks

For each model in your paper, perform as many of the following as feasible:

1. **Limiting-case check.** What does the model predict in extreme cases that you can solve in your head? (Zero customers, zero growth, no friction, etc.)
2. **Conservation check.** If the model should conserve something (population, money), does it?
3. **Dimensional check.** Are the units of every equation consistent?
4. **Literature check.** Does the result agree with published values for similar systems?
5. **Cross-method check.** Does an alternative method (e.g., simulation vs. analytical solution) give the same answer?

Worked Example: Example 3: Limiting-Case Validation for an EOQ Model

The Economic Order Quantity formula $Q^* = \sqrt{2Dk/s}$ predicts the optimal order quantity given annual demand D, ordering cost k, and holding cost s.

Limiting case 1: $k \to 0$ **(no ordering cost).** Then $Q^* \to 0$: order infinitely often in infinitesimally small batches. ✓(Matches intuition: if ordering is free, never hold inventory.)

Limiting case 2: $s \to 0$ **(no holding cost).** Then $Q^* \to \infty$: order all annual demand at once. ✓(Matches intuition: if storage is free, minimise the number of orders.)

Limiting case 3: $D \to 0$ **(no demand).** Then $Q^* \to 0$: do not order. ✓

All three limiting cases check out, increasing confidence in the formula.

Worked Example: Example 4: Cross-Method Validation for a Queuing Model

Our queueing analysis gives the average wait time in an $M/M/1$ queue with $\lambda = 12$/hr and $\mu = 15$/hr as:

$$W_q = \frac{\rho}{\mu(1-\rho)} \text{ where } \rho = \lambda/\mu = 0.8.$$

$$W_q = \frac{0.8}{15 \cdot 0.2} = 0.2667 \text{ hr} = 16 \text{ min}.$$

We cross-check this against a discrete-event simulation of 10^5 customer arrivals, implemented in Python (`simpy`). The simulation gives average wait $W_q^{\text{sim}} = 15.94 \pm 0.21$ min (95% CI).
The analytical formula and simulation agree to within 0.4%, well inside the simulation's confidence interval. ✓

Worked Example: Example 5: Literature Comparison

Our model predicts that a single bicycle highway between the city centre and the Eastern residential district will shift 18% of commuters from cars to bicycles. We compare against three published European deployments:

Source	City	Mode shift	Notes
Pucher & Buehler (2008)	Copenhagen	22%	After 10 years of network expansion
Hellinga (2017)	Utrecht	16%	After single highway opening
Aldred et al. (2019)	London CS3	14%	After single highway
Our model	–	**18%**	Single highway, mid-sized city

Our prediction falls within the range of comparable single-corridor deployments (14–22%) and lies between the values for single-corridor cases (Utrecht, London) and the long-term, network-scale Copenhagen result, as we would expect.

Note: Format conventions for Validation

- Validation appears as a numbered section (often §6 or §7), *not* inside Model Development.
- Each check is a sub-subsection or a clearly labelled paragraph.
- If a check fails, **do not hide it** — discuss what failed and what it implies. Honest discussion of a failed check is far stronger than silent omission.

Sensitivity Analysis: How Robust is Your Answer?

Core explanation, worked reasoning, and application

Sensitivity Analysis: How Robust is Your Answer?

Sensitivity analysis answers: *How does the output change as inputs change?*
This is the single most commonly weak section in first-time-team papers. Many teams write a few sentences claiming their results are "robust". Strong teams build sensitivity tables, tornado diagrams, or sensitivity heatmaps that constitute a visual narrative on their own.

Method: Method: Three Levels of Sensitivity Analysis

From simplest to most thorough:

1. **One-at-a-time (OAT).** Hold all parameters at their nominal values; vary one parameter at a time ($\pm 10\%$ or $\pm 20\%$); record effect on output. Simple, fast, but misses interactions.
2. **Tornado diagram.** A horizontal bar chart showing the effect of each parameter when varied within its plausible range. Bars sorted by magnitude. Immediately identifies the most influential parameters.
3. **Monte Carlo / global sensitivity.** Sample all parameters jointly from plausible distributions; compute the distribution of outputs. Captures interactions. Best for high-credibility papers.

For competition papers, OAT plus a tornado diagram is usually sufficient. Reserve Monte Carlo for problems where parameter interactions matter (e.g., queuing under load).

Worked Example: Example 6: OAT Sensitivity Table

Our model predicts a baseline of 312 drones needed. We vary each input parameter $\pm 20\%$ from its nominal value and report the effect on the recommendation.

Parameter	Nominal	-20%	$+20\%$	Range
Viewing distance d (m)	200	287 drones	339 drones	$\pm 9\%$
Image height h (m)	50	256 drones	369 drones	$\pm 18\%$
Drone spacing s (m)	1.0	488 drones	217 drones	$\pm 56\%$
Refresh rate f (Hz)	30	312 drones	312 drones	0%

Interpretation. The recommendation is most sensitive to drone spacing s, which dominates the geometric resolution constraint. Image height has secondary effect. Viewing distance has modest effect. Refresh rate has no effect (the model is quasi-static).

Implication for recommendation. Drone spacing should be measured/specified precisely. Refresh rate can be left to operational judgement.

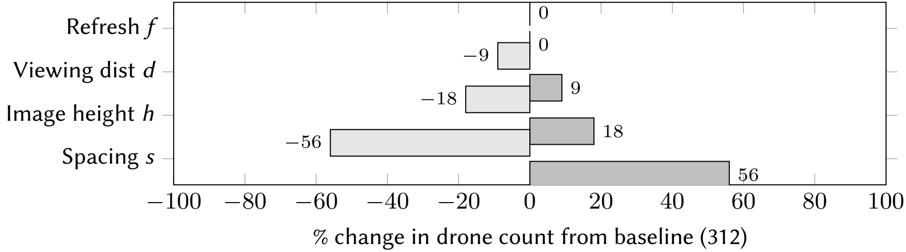
Note: Why tables work better than prose for sensitivity

Compare the table above with: "Our model is relatively sensitive to drone spacing, which causes large changes, and to a lesser extent image height, while viewing distance has only modest effect and refresh rate is essentially neutral."

The table delivers the same information at a glance *and* with numbers *and* with a clearer ranking. Tables almost always beat prose for sensitivity. Use prose to interpret the table, not to replace it.

Worked Example: Example 7: Tornado Diagram

The tornado diagram below visualises Example 6's OAT data:



Sorted by absolute effect; the dominance of s is immediately visible.

Structural Sensitivity: Beyond Numerical Perturbation

Numerical sensitivity tests parameter changes within an assumed model. **Structural sensitivity** tests whether the conclusion survives when the *model itself* changes.

Method: Method: Structural Sensitivity Checks

Ask the following:

1. If we replace the exponential service distribution with deterministic, what happens?
2. If we replace logistic growth with exponential growth, what happens?
3. If we relax linearity in the dose-response curve, what happens?
4. If we add second-order effects to a first-order model, what happens?

Even one structural check, well-executed, signals model maturity.

Worked Example: Example 8: Structural Sensitivity Check

Our queuing model assumes exponential service times (the $M/M/1$ assumption). To test the structural sensitivity of our recommendation, we re-run the analysis with deterministic service times (an $M/D/1$ queue).
Under $M/M/1$: average wait $W_q = 16$ min, recommended staffing $= 3$ baristas.
Under $M/D/1$: average wait $W_q = 8$ min, recommended staffing $= 2$ baristas.
The headline recommendation changes by one barista. We thus report a recommendation **range** of 2–3 baristas with 3 as the upper bound, and we note that more accurate service-time data would allow tighter recommendation.

Putting Them Together: The Trust Chain

Core explanation, worked reasoning, and application

Putting Them Together: The Trust Chain

A great paper builds an **evidence chain**: assumptions $\rightarrow$ model $\rightarrow$ validation $\rightarrow$ sensitivity $\rightarrow$ recommendation. Each link supports the next.

Method: Method: The Evidence Chain Self-Check

Before submitting your paper, work through this chain backward from your recommendation:

1. **Recommendation:** What am I recommending?
2. **Sensitivity:** Is this recommendation stable to plausible changes in inputs?
3. **Validation:** Has the model that produced this recommendation passed at least one independent check?
4. **Assumptions:** Are the assumptions behind the model explicit and defensible?
5. **Coverage:** Does the recommendation actually answer the question that was asked?

Break in any link of the chain weakens the entire paper. Strong papers have all five visible to the judge.

Worked Example: Example 9: A Complete Evidence Chain

Recommendation. The bicycle highway network should be built (47 km, EUR 28M, predicted 18% modal shift).
Sensitivity. Modal shift varies from 14–22% across plausible weather and demographic assumptions; the recommendation is robust within this range. The most influential parameter is winter rainfall frequency.
Validation. The 18% prediction falls within the range of comparable European deployments (14–22%) per three published studies; our value lies sensibly between single-corridor and network-scale data points.
Assumptions. Key modelling assumptions: logit mode-choice form with parameters from European studies; constant winter weather; no major changes to public transit during the implementation period.
Coverage. The original problem asked for network design and commute time reduction; both are answered.

Every link in the chain is short, specific, numbered, and visible. This is what makes the paper trustworthy.

Workshop Exercises

Core explanation, worked reasoning, and application

Exercises: Exercise W3.1: Rewriting Assumptions

The following assumptions are taken from a (constructed) weak paper. Rewrite each using the three-part structure (Statement / Justification / Implication).

1. *We assume the road is straight.*
2. *We assume the population grows logistically.*
3. *We assume customers are rational.*
4. *We assume there is no friction.*
5. *We assume the data is accurate.*

Exercises: Exercise W3.2: Choose the Validation Method

For each of the following models, suggest two appropriate validation checks (chosen from the five in Method 4.1). Justify your choice.

1. A discrete-event simulation of a hospital emergency department.
2. An optimisation model for warehouse layout.
3. A regression model for sales prediction.
4. A network model for transport corridor routing.
5. A stochastic model for inventory under uncertain demand.

Exercises: Exercise W3.3: Build a Sensitivity Table

Consider the following hypothetical setting. A model for daily revenue at an ice-cream shop is

$$R = (a - bT) \cdot p \cdot N$$

where T is the day's temperature (°C), p is the price per cone (\$), N is the number of potential customers, and a, b are calibrated constants. Nominal values: $a = 5$, $b = 0.05$, $T = 25$°C, $p = 4$, $N = 200$, giving $R = (5 - 0.05 \cdot 25) \cdot 4 \cdot 200 = \$3000$.

Build a one-at-a-time sensitivity table varying each of T, p, N, a, b by $\pm 20\%$, and identify the parameter the recommendation is most sensitive to.

Exercises: Exercise W3.4: Diagnose the Evidence Chain

Read any one Outstanding paper. Identify the team's:

- Top-line recommendation.
- Validation checks performed.
- Sensitivity analysis methodology and main findings.
- Key assumptions.

Now trace the chain backward from recommendation to assumptions. Where are the strongest and weakest links?

Exercises: Exercise W3.5: Write a Full Trio

For your most recent modeling project (course or competition), draft:

1. An Assumptions section with 5–6 three-part assumptions.
2. A Validation section using at least two of the five standard checks.
3. A Sensitivity Analysis section with one OAT table and one structural check.

Have a teammate read all three and ask: *do you trust the recommendation more after reading these than you did before?* If the answer is ”yes, much more”, you have succeeded.

What to Carry Into W4

Core explanation, worked reasoning, and application

Note: Class W3 Take-aways

1. These three sections sit at the heart of the *Validation & Sensitivity* pillar (15–20% of score), but their leverage is greater because they distinguish thoughtful from formulaic work.
2. Every assumption needs Statement, Justification, Implication.
3. Validation has five standard methods; do at least two; include cross-method when possible.
4. Sensitivity analysis must include numbers and a table or figure — not just prose.
5. Structural sensitivity (changing the model) is rare and high-value; do at least one when feasible.
6. The evidence chain from recommendation back to assumptions must hold link-by-link.

What to Carry Into W4

In **W4: English Academic Writing Style** we shift from *structure* to *language*: how to write sentences that sound like scientific papers rather than like translations from another language. We address voice, hedging, connectives, and common Chinglish patterns.

Appendix A: Sensitivity Analysis Quick-Start

Core explanation, worked reasoning, and application

Appendix A: Sensitivity Analysis Quick-Start

Step	Action
1. Identify parameters	List every numerical input to your model.
2. Pick a range	For each, choose a plausible $\pm$ range (often $\pm 10\%$ or $\pm 20\%$).
3. Vary one at a time	Re-run the model with each parameter perturbed.
4. Tabulate effects	Record output change for each parameter.
5. Rank by influence	Sort the table by absolute effect.
6. Plot tornado diagram	Visualise the ranking.
7. Discuss	In one paragraph, identify the most/least influential parameters and what this implies for the recommendation.
8. (Optional) Structural	Replace one model component with a different functional form; re-run.

Appendix B: Validation Checklist

Core explanation, worked reasoning, and application

Appendix B: Validation Checklist

For each model in your paper, check:

- ☐ Does the model behave correctly in at least one limiting case?
- ☐ Are quantities that should be conserved actually conserved?
- ☐ Are the units consistent throughout?
- ☐ Does the result match (or fall in the range of) published values for similar systems?
- ☐ Does an alternative method give a comparable result?
- ☐ If a check failed, is that discussed honestly in the paper?

Appendix C: Strong Assumption Templates

Core explanation, worked reasoning, and application

Appendix C: Strong Assumption Templates

Use these as starting points; modify and extend.

- *We assume X is approximately constant during the modeled period [Y], because [reason]. This allows us to [model consequence], at the cost of [realism trade-off].*
- *We assume X follows a [distribution] with parameters estimated from [data source]. We confirm this fit via [test/visual], obtaining [p-value or descriptive metric]. This supports the use of [standard model relying on this distribution].*
- *We assume X and Y are independent, which is justified by [data, theory, or convention in the field]. This allows the joint probability factorisation* $P(X, Y) = P(X)P(Y)$*, simplifying the calculation of [downstream quantity].*
- *We restrict our analysis to the regime [condition]. Outside this regime, model assumption [A] would fail, requiring a [different approach]. We address that regime briefly in Section [Z].*

Writing Class 4: English Academic Writing Style
Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch

Question

Motivating context, key vocabulary, and the first modelling decision.

Build

Method

Worked example, diagram or table, calculation, and parameter meaning.

Test

Critique

Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

From Structure to Language

Core explanation, worked reasoning, and application

In W1–W3 we focused on *structure*: what sections appear, in what order, with what content. In this class we focus on *language*: how a sentence in a competition paper should look and sound.
Two facts motivate this class:

- For IMMC International and HiMCM, English is required. Many teams write English that is grammatically correct but does not sound like a scientific paper. Judges notice immediately.
- Even a paper with strong content can be marked down for awkward prose. Conversely, prose that sounds confident and conventional flatters the content.

This class is *not* a remedial English class. It will not teach articles, tenses, or grammar; we assume that level.
The aim is to bridge the gap from "correct English" to "the English of a scientific paper".

Note: An honest framing

Most members of a Macau/Hong Kong team are not native English writers. That is fine. **Judges do not expect native-level fluency; they expect register-appropriate prose.** A consistent, plain, scientific register beats a flowery, ambitious one every time.

Plain is good. Plain is professional. Plain is what scientific writing actually looks like.

The Two-Voice Rule

Core explanation, worked reasoning, and application

The Two-Voice Rule

Scientific English makes heavy use of *both* active and passive voice. Choosing well is half the battle.

Method: Method: When to Use Each Voice

- **Active voice** when *the team* is the agent and the team's choice matters: "We assume...", "We define...", "We recommend...", "We model the population as...". Use whenever the team is acting.
- **Passive voice** when the *process or result* matters more than who did it: "The system is described by...", "Data was collected from...", "Parameters are estimated using...". Use for procedures, definitions, and standard operations.

Bad writing comes from overusing either: all-passive sounds stiff and bureaucratic; all-active reads like a personal diary.

Worked Example: Example 1: Mixing Voices Well

We model the population as a logistic growth process (active). The carrying capacity K is estimated from satellite data over the period 2010–2020 (passive). We choose this estimation window because earlier data are less reliable (active). Standard least-squares regression is applied, yielding K $\approx 1{,}240$ *animals (passive). We then compare this estimate against a published 2018 census of 1,195 animals (active).*

Note: Why this paragraph works

Notice the rhythm: *active, passive, active, passive, active.* The team's choices and judgments are in active voice; the data-processing and standard operations are in passive. This is the natural rhythm of a scientific paragraph.

Method: Method: How to Refer to Yourself

- **Use "We"**: the convention in 99% of modeling papers. Comfortable, unpretentious, accurate for a team paper.
- **Do not use "I"** unless the paper is single-authored.
- **Do not use "the authors"** in self-reference. It reads as evasive and dated. Reserve it for citing other papers ("the authors of ...").
- **Avoid the impersonal** "It is decided that..." or "It was felt...". These hide the agent and sound bureaucratic.

Hedging: How to Sound Honest Without Sounding Weak

Core explanation, worked reasoning, and application

Hedging: How to Sound Honest Without Sounding Weak

Hedging is the practice of qualifying claims. In scientific writing, hedging signals integrity (you know the limits of your work). Done badly, it sounds evasive or insecure. Done well, it sounds rigorous.

Method: Method: The Hedging Vocabulary

Common scientific-paper hedges, from weak to strong:

Modal verbs	may, might, could, can
Adverbs	approximately, roughly, broadly, plausibly, possibly, likely, presumably
Verbs of estimation	we estimate, we suggest, we infer, we conjecture, we expect, we find that
Conditionals	under the assumption that, in the regime where, provided that
Quantitative hedges	within a 5% margin, to first order, in most scenarios tested

Worked Example: Example 2: Three Levels of Hedging

Too strong (overclaims):

> *Our model proves that the optimal staffing level is 3 baristas.*

Problem: "Proves" is the language of mathematics. A model does not *prove* anything; it *suggests* or *indicates*. Judges will note the overclaim.

Too weak (overcautious):

> *Our model might possibly suggest that maybe the staffing level could perhaps be approximately around 3 baristas in some cases.*

Problem: Five hedges. The reader cannot tell what is being claimed.

Right (one calibrated hedge):

> *Under our assumptions, the model recommends 3 baristas, with sensitivity analysis confirming this recommendation across a* $\pm 20\%$ *range in arrival rate.*

Why it works: "Under our assumptions" is one explicit hedge, immediately backed by a quantitative robustness claim. The recommendation reads as confident yet honest.

Warning: Warning: When NOT to Hedge

Do not hedge in:

- Definitions: "*The utilisation is defined as...*" (not "*may be defined as...*").
- Mathematical statements: "*By Eq. (3), $W_q = 16$ min*" (not "*may approximately equal 16 min*").
- Recommendations after sensitivity: be honest about uncertainty in the analysis section, then make a clean recommendation. Hedging the recommendation itself reads as a failure to commit.

Connecting Sentences: The Logic of Flow

Core explanation, worked reasoning, and application

Connecting Sentences: The Logic of Flow

Scientific writing is heavy with *connective words*: words that show how one idea relates to the next. Good connections make the paper feel inevitable; bad connections (or none) make the reader feel lost.

Method: Method: Connective Vocabulary by Purpose

Purpose	Connectives
Addition	moreover, in addition, further, furthermore
Contrast	however, in contrast, conversely, on the other hand, yet, but
Cause / Consequence	therefore, hence, thus, consequently, as a result, accordingly, it follows that
Specification	specifically, in particular, for instance, namely
Generalisation	in general, broadly, overall, on balance
Sequence	first, next, finally, subsequently, ultimately
Concession	although, while, despite, even though, nevertheless, nonetheless
Restatement	in other words, that is, equivalently, put differently

Worked Example: Example 3: With and Without Connectives

Without:

Our baseline model predicts an average wait time of 4 minutes. We tested the sensitivity to arrival rate. The wait time changed substantially. We added a second server. The wait time dropped to 1.5 minutes. We recommend two servers during peak hours.

With:

Our baseline model predicts an average wait time of 4 minutes. However, sensitivity analysis to arrival rate revealed substantial fluctuations. We therefore explored a two-server configuration, which reduced average wait time to 1.5 minutes. Accordingly, we recommend deploying two servers during peak hours.

Warning: Warning: Overusing Connectives

Every sentence does not need a connective. Connectives are like joints: they go where direction *changes*. If consecutive sentences continue the same thread, no connective is needed (and one would feel intrusive). Use connectives where the logic turns or where the reader might otherwise get lost.

Common Chinglish Patterns and Their Fixes

Core explanation, worked reasoning, and application

Common Chinglish Patterns and Their Fixes

This section is the one where we are most direct. We list common patterns produced by Chinese-/Cantonese-speaking writers in English scientific contexts, and we give specific fixes.

Warning: Warning: Why Chinglish Hurts You Specifically in Competitions

Judges of international competitions read hundreds of papers per year, many by non-native writers. They recognise Chinglish patterns instantly. The problem is not that the prose is technically incorrect — it is that it signals the team did not have a native (or near-native) reader review the paper. Teams that fix the patterns below are immediately read more carefully.

Pattern 1: Long, Multi-Clause Sentences

Chinese tolerates very long sentences (commas as clause separators); English does not.

Worked Example: Example 4: Long Sentence → Three Sentences

Chinglish:

> *In our model, we use the SIR framework to describe the spread of the disease, and we assume that the infection rate is constant, which is a simplification, and we then solve the system using the Runge-Kutta method, obtaining the time evolution of S, I, and R.*

English:

> *We model the disease using the SIR framework with a constant infection rate. While this is a simplification, it permits closed-form analysis. We solve the system using the Runge-Kutta method and obtain the time evolution of S, I, and R.*

Pattern 2: Overuse of ”Which”

Chinese has no relative pronouns; English writers from Chinese backgrounds compensate by overusing “which”.

Worked Example: Example 5: ”Which” Overuse

Chinglish: *We used the optimal value, which is calculated from the equation, which gives the result, which we present in Table 3.*
English: *We computed the optimal value from Eq. (3) and report the result in Table 3.*

Method: Method: Wordy Phrases → Concise Replacements

Wordy (avoid)	Concise (prefer)
in order to	to
due to the fact that	because
at the present time	now
in the event that	if
in spite of the fact that	although
a large number of	many
the majority of	most
on a daily basis	daily
with regard to / with respect to	regarding / about
it should be noted that	note that
as a matter of fact	in fact
it is important to mention that	importantly,
for the purpose of	to / for
in close proximity to	near
prior to	before
subsequent to	after

Worked Example: Example 6: Wordiness → Concision

Wordy:

In order to test the validity of our model, we performed a large number of simulations. Due to the fact that the variance was small, we may say that the results are robust. It should be noted that prior to the analysis, we cleaned the data with regard to outliers.

Concise:

To validate our model, we performed many simulations. Because the variance was small, the results are robust. Note that before the analysis we removed outliers from the data.

Pattern 4: *Missing or Misused Articles*

Chinese has no articles (“a”, “an”, “the”). This is the single most common Chinglish flag.

Method: Method: A Three-Question Article Test

For every singular countable noun, ask:

1. Is the noun *specific* (the reader knows which one)? Use **the**.
2. Is it *any one of a class*, mentioned for the first time? Use **a** or **an**.
3. Is the noun in its general sense, abstract, or uncountable? Use **no article**.
4. Plural countables in general sense: **no article**.

Examples:

- *A model is built. The model has three parameters. We modified the model.* (First instance: any one. Subsequent: that specific one.)
- *Simulation is a powerful tool.* (General sense: no article.)
- *Data was collected.* (“Data” is treated as uncountable: no article.)
- *The data we collected showed...* (specific data: “the”.)

Method: Method: Common Connective Confusions

- **“Besides”** in academic English means “in addition to (a previously mentioned thing)”. Avoid using it as a generic “moreover” because it sounds informal.
  - ✗ *Besides, we ran more simulations.*
  - ✓ *In addition, we ran more simulations.* or simply *We also ran more simulations.*
- **“On the other hand”** should be paired with “on the one hand” (or implied). Otherwise, use “in contrast” or “however”.
- **“At the same time”** usually means temporal simultaneity, not logical concurrence. Use “meanwhile” or “also”.
- **“As we all know”** is rarely appropriate in a paper. Replace with a citation or simply state the fact.

Worked Example: Example 7: Idiomatic Verb-Noun Pairs

Some verb-noun pairs translate idiomatically in English in ways that differ from their Chinese counterparts. Some traps:

- ✗ "*do a research*" ✓ "conduct research" or "perform research"
- ✗ "*open a meeting*" ✓ "hold a meeting"
- ✗ "*learn a knowledge*" ✓ "acquire knowledge" (uncountable)
- ✗ "*study at home*" (when meaning university) ✓ "study domestically" or "study at home university"
- ✗ "*problem solving*" as a verb ✓ "solve the problem"
- ✗ "*make a decision to do*" ✓ "decide to do" (verb is shorter than noun)

Pattern 7: Topic-Comment Sentences

Chinese sentences often start with a topic, even where the topic is not the grammatical subject. English needs a clear subject-verb-object spine.

Worked Example: Example 8: Topic-First → Subject-First

Chinglish: *For our model, the assumption of independence is important.*
English: *Our model requires the assumption of independence.* or *Independence is a crucial assumption of our model.*
Chinglish: *In the simulation, the result was robust.*
English: *The simulation result was robust.* or *Our simulation produced robust results.*

Diagnosing and Polishing a Paragraph

Core explanation, worked reasoning, and application

Diagnosing and Polishing a Paragraph

This section illustrates the kind of careful revision a strong team applies to every paragraph in the final 12 hours of a competition.

Worked Example: Example 9: A Full Paragraph Polish

First-draft Chinglish:

> *In order to do the validation of our model, we have done a lot of work. First of all, we collected the data from the Hong Kong government website, which we then cleaned, which took us a long time. After that, we run our model with the data, which produced the output, which we compared with what the literature says. The result is, our model is generally consistent with the literature, which we think is a good thing. Besides, we also did sensitivity analysis, which is in the next section.*

Polished version:

> *To validate our model, we collected daily reports from the Hong Kong Department of Health for the period 2019–2021. After removing duplicate entries, we ran our model on the cleaned dataset and compared the predicted values against published estimates. The model agreed with the literature to within 7% across the period, supporting the validity of our approach. Sensitivity analysis is treated separately in §7.*

Word count: 86 words $\rightarrow$ 59 words. The polished version is shorter, clearer, and more confident.

Method: Method: The ”Polish Pass” Procedure

In the last 8–12 hours, do a polish pass on every paragraph:

1. Read each sentence aloud. If it sounds awkward, rewrite.
2. Cut every wordy phrase (use the table in §5.3).
3. Replace “which” chains with two simple sentences.
4. Ensure each paragraph has *one* main idea.
5. Check articles in every singular countable noun.
6. Verify the connective at the start of each paragraph still makes sense in context.

Allocate one teammate as the dedicated polisher for the last 6 hours — ideally someone with the strongest English. This role is as important as the modeler or coder during this final phase.

Workshop Exercises

Core explanation, worked reasoning, and application

Exercises: Exercise W4.1: Active/Passive Identification

Take any of the example paragraphs in W1–W3. Label each verb as active or passive. Count the ratio. Is it natural? Does it follow the rhythm described in §2?

Exercises: Exercise W4.2: Polish a Paragraph

The following paragraph is from a constructed weak paper. Polish it using the techniques from this class. Aim to cut 30% of words without losing any content.

> *In order to do the analysis of our problem in a thorough way, we have decided to make use of the method which is known as Monte Carlo simulation, which is widely used in the modeling field, which has many advantages such as its flexibility and which can be applied to a large number of different situations. Besides, we also chose to use Python as the programming language, because of the fact that Python has many libraries, which are useful, and which we are familiar with. The result of our simulation is shown in Figure 3, which can be seen above.*

Exercises: Exercise W4.3: Hedging Calibration

For each statement, rewrite with appropriate hedging (neither overclaiming nor over-hedging).

1. Our model proves the optimal staffing is 3 servers.
2. This result is definitely correct under all conditions.
3. Maybe our analysis might possibly suggest some kind of optimum exists.
4. Our recommendation will definitely solve the problem completely.
5. This conclusion might possibly perhaps be approximately right.

Exercises: Exercise W4.4: Detect the Chinglish

Read each sentence and identify which Chinglish pattern (from §6) it exhibits. Then rewrite.

1. *For our research, the most important thing is the data.*
2. *We did a research on the bus routing problem, which is interesting, which has many applications.*
3. *In order to in the future make our analysis more comprehensive, we plan to do additional studies.*
4. *There are a large number of citizens who are using the bus on a daily basis.*
5. *Besides, our model also can be used to predict the future trend.*

Exercises: Exercise W4.5: Find a Native Reader

During practice runs, ask a teacher (ideally a native English speaker, or someone who reads English papers regularly) to read your Summary and identify three sentences they would rewrite. Compare with your own rewrite of those sentences. Where do you and the reviewer disagree, and why?

What to Carry Into W5

Core explanation, worked reasoning, and application

Note: Class W4 Take-aways

1. Mix active and passive voice; use active for team choices, passive for procedures.
2. Hedge enough to be honest, not so much that meaning vanishes.
3. Use connectives where logic changes; do not use them in every sentence.
4. English does not tolerate the long, comma-spliced sentences that Chinese tolerates. Break them.
5. Articles, "which"-chains, and topic-first sentences are the three most common Chinglish patterns; fix them on every pass.
6. Allocate one teammate as the dedicated polisher for the final hours.

What to Carry Into W5

In **W5: LaTeX, Figures and Visual Communication** we move beyond words to the visual presentation of the paper: typesetting, figures, tables, and the workflow that makes everything come together at submission.

Appendix A: Quick-Reference Style Sheet

Core explanation, worked reasoning, and application

Appendix A: Quick-Reference Style Sheet

- First person plural "we"; never "I", "the authors", or "one".
- Mix active and passive voice; alternate by paragraph rhythm.
- Hedge with quantitative qualifiers ("within 5%") rather than vague adverbs ("somewhat").
- Cite numerically: "In our simulation we observe X (see Fig. 3)".
- Equations: punctuate as parts of sentences, with commas and periods after them as needed.
- Numbers: spell out below ten in narrative prose ("three models"); use digits in technical contexts ("$N = 3$").
- Units always with a non-breaking space: `200~m` renders as "200 m".
- Avoid contractions: "do not", not "don't".

Appendix B: Top 30 Phrases to Replace

Core explanation, worked reasoning, and application

Appendix B: Top 30 Phrases to Replace

Appendix B: Top 30 Phrases to Replace

- “in order to” → “to”
- “due to the fact that” → “because”
- “in spite of the fact that” → “although”
- “it should be noted that” → “note that”
- “a large number of” → “many”
- “the majority of” → “most”
- “a number of” → “several” or a count
- “the fact that” → usually deletable
- “regarding” / “with respect to” → “about” or “on”
- “in close proximity to” → “near”
- “prior to” → “before”
- “subsequent to” → “after”
- “utilise” → “use”
- “commence” → “start”
- “terminate” → “end”
- “demonstrate” (often) → “show”
- “investigate” → “study”
- “in the event that” → “if”
- “at the present time” → “now”
- “at some future date” → “later”
- “in light of the fact that” → “because”
- “it is necessary that” → “must”
- “in many cases” → “often”
- “in most cases” → “usually”
- “serves to” → usually deletable
- “has the ability to” → “can”
- “is in a position to” → “can”
- “does not have the ability to” → “cannot”
- “in the same way” → “similarly”
- “with the result that” → “so”

Appendix C: A Recommended Final-Hours Editing Procedure

Core explanation, worked reasoning, and application

Appendix C: A Recommended Final-Hours Editing Procedure

1. **Hours** $\mathrm{T}-12$ **to** $\mathrm{T}-8$**:** Read the entire paper top to bottom. List sections needing improvement.
2. **Hours** $\mathrm{T}-8$ **to** $\mathrm{T}-4$**:** Address the list. Polish prose. Re-check all numbers, units, figure captions.
3. **Hour** $\mathrm{T}-4$ **to** $\mathrm{T}-2$**:** Polish the Summary specifically. Read it aloud. Confirm a stranger could find the headline answer in 10 seconds.
4. **Hour** $\mathrm{T}-2$ **to** $\mathrm{T}-1$**:** Spell check; reference check; verify cross-references ("see Fig. 3", "see Eq. (12)") resolve.
5. **Final hour:** Compile to PDF, verify rendering, check page numbers, name file according to contest convention.
6. **Submit with one hour to spare.** Never aim for the deadline. Internet failures during the final hour are routine.

Writing Class 5: LaTeX, Figures, Tables, and Visual Communication

Mathematical Modeling Lecture

Mathematical Modeling

Kenneth Sok Kin Cheng

Session Route

Launch	Build	Test
Question Motivating context, key vocabulary, and the first modelling decision.	**Method** Worked example, diagram or table, calculation, and parameter meaning.	**Critique** Checkpoint practice, limitation check, and communication of results.

Columns are used where comparison helps; explanation slides stay clean and spacious.

From Words to Pages

Core explanation, worked reasoning, and application

From Words to Pages

In W1–W4 we developed structure, content, and language. This final class is about the *presentation* of all of that on the page: how a paper is typeset, how figures and tables are made, and how a team's workflow makes everything come together at the submission deadline.

There are two reasons to take this seriously:

- Visual quality is part of the *Communication* pillar of the rubric. A well-typeset paper with clear figures *looks* like an Outstanding paper before the judge has read a single sentence.
- In the final 24 hours of a competition, integration issues (figures not exporting, references broken, page count overflow) consume disproportionate time. A team that has practised the tools in advance can use those hours for content; a team that has not will spend them debugging the document.

This class will not turn anyone into a typography expert. Its purpose is to establish a competent baseline, identify the highest-leverage visual choices, and lock in a workflow that survives competition stress.

LaTeX: The Competition Standard

Core explanation, worked reasoning, and application

LaTeX: The Competition Standard

LaTeX is the de facto standard for mathematical-modeling competition papers. We recommend it for three reasons:

- **Mathematical typesetting.** Equations look professional with no extra effort. Word-processor equations are flagged immediately by judges.
- **Cross-references.** Equations, figures, tables, and citations stay linked correctly through rewrites and reorderings.
- **Templates.** Mature templates exist for IMMC, HiMCM, MCM that handle layout decisions for you.

Method: Method: Setting Up Overleaf for a Competition

1. Create a single shared Overleaf project before the competition starts. All teammates should be added as collaborators.
2. Start from a known-good template (IMMC, HiMCM, or MCM-style; many examples exist on Overleaf's template gallery).
3. In the first hour of the competition, finalise:
   - Document class, font size, margin settings.
   - Section structure (insert empty sections with placeholders).
   - Citation style (use BibTeX, not manual references).
4. Set up auto-compile on save, and commit to a single "main" file structure.

Warning: Warning: Common Overleaf Pitfalls

- **Compilation timeout.** Overleaf free tier limits compile time. Heavy TikZ figures may time out. Compile externally (TeX Live) if you anticipate problems.
- **Concurrent edits on the same line.** Coordinate via teammates rather than via Overleaf's auto-merge — it sometimes loses content silently.
- **Version history.** Keep an external backup (a daily download of the source as a zip) in case Overleaf has an outage at submission time.

Code or LaTeX example: A Minimal Preamble for a Modeling Paper

Listing 1: A clean preamble adequate for any IMMC/HiMCM paper.

```
\documentclass[12pt,a4paper]{article}
\usepackage[margin=2.5cm]{geometry}
\usepackage{amsmath, amssymb, amsthm}
\usepackage{graphicx}       % include figures
\usepackage{booktabs}       % nicer tables (\toprule, \midrule, \bottomrule)
\usepackage{caption}        % better caption formatting
\usepackage{subcaption}     % side-by-side subfigures
\usepackage{float}          % stricter figure placement (H option)
\usepackage{hyperref}       % clickable references
\usepackage{cleveref}       % "Fig. 1" instead of just "1" in cross-refs
\usepackage[numbers]{natbib} % numerical citation style

\title{Team N12345: Modeling the Throughput of\\
       a Coffee Shop Queue}
\author{}                   % no individual names, per contest rules
\date{November 2026}
```

Note: On font size and margins

Contest rules specify font size (usually 12pt) and margins. Verify before the competition. The most common defaults are 12pt Times New Roman or Computer Modern, 1-inch margins. Do *not* shrink either to gain space; judges check.

Method: Method: Mathematics That Reads Well

1. **Display long equations; inline short ones.** An equation that names quantities and will be referenced later (`\label`, `\ref`) goes on its own line in an `equation` environment. A short relation like $x \geq 0$ goes inline.
2. **Number every equation you reference.** Use `equation` with `\label`; reference with `\eqref` or `Cref`. Equations cited zero times can be `equation*` (un-numbered).
3. **Punctuate displayed equations as parts of sentences.** An equation that ends a sentence ends with a period. An equation followed by “where” or “and” ends with a comma.

Worked Example: Example 1: Punctuated Displayed Equations

The dynamics are governed by

$$\frac{dS}{dt} = -\beta SI, \qquad \frac{dI}{dt} = \beta SI - \gamma I, \tag{1}$$

where β is the infection rate and γ the recovery rate. Eq. (1) is the classical SIR system, named in Section 4.

Note: the comma after the second equation, the "where" clause that explains symbols, and the cross-reference `Eq.\ \eqref{eq:sir}`. This is the conventional style.

Figures: The Single Most Underrated Investment

Core explanation, worked reasoning, and application

Figures: The Single Most Underrated Investment

A single well-made figure can communicate a key insight more effectively than three pages of text. Conversely, a sloppy figure tells the judge you ran out of time — or worse, that you don't understand your own model.

Method: Method: The Five Criteria for a Publication-Quality Figure

Every figure in your paper should satisfy:

1. **Self-contained.** A reader who skips the body and looks only at the figure plus caption should understand what the figure shows.
2. **Labelled axes with units.** "Time" is not a label; "Time (s)" or "Time (days)" is.
3. **Legend present.** Multiple curves on one plot need a legend distinguishing them.
4. **Readable text.** Tick labels and legends should be readable at the printed size (typically 10pt minimum).
5. **Honest scales.** Linear scale by default; logarithmic only when there is a reason (range spans orders of magnitude, or theory predicts a power law). Never truncate axes to exaggerate differences.

Note: On colour

Use colour sparingly and meaningfully. Greyscale plus one accent colour (often a dark blue or burgundy) is sufficient for most figures. Avoid rainbow colour schemes; they imply ordering even where none exists, and they fail under greyscale printing.

For technical reasons many judges print papers in greyscale. *Your figures must remain interpretable in greyscale.* Distinguish curves by line style (solid, dashed, dotted) as well as by colour.

Method: Method: Writing a Strong Caption

A figure caption should:

1. Start with a **short label phrase** that names the figure (often a sentence fragment, period-terminated).
2. Continue with one or two sentences of explanation.
3. Identify the **key takeaway** the reader should see.
4. Always end with a period.

Worked Example: Example 2: Strong vs Weak Captions

Weak: "Figure 3: Simulation results."
Problem: Says nothing. Does not stand alone.
Weak: "Figure 3: Number of customers vs. time."
Problem: Better, but still does not say what the reader should conclude.
Strong: "Figure 3: Customers in the queue over a simulated 4-hour service window. Peak queue length of 17 occurs around minute 90 and remains above 10 from minute 60 to minute 120, supporting the recommendation to deploy a second server during this period."
Why it works: Names what the figure is, what the reader should observe (peak around 90, sustained above 10), and connects to the recommendation. A judge skimming captions still gets the headline result.

Method: Method: Recommended Tools by Figure Type

Figure type	Recommended tool
Time series / scatter / regression	Python `matplotlib` or Excel; export PDF or PNG at 300+ dpi.
Bar / pie / tornado	Python `matplotlib`; or Excel with care taken to clean the default style.
Network diagrams	TikZ (for full integration), or Graphviz / NetworkX.
Schematics / flow charts	TikZ; or Inkscape; or `draw.io` (export to SVG/PDF).
Mathematical diagrams (with formulas)	TikZ exclusively.
3D surfaces	`matplotlib`'s `Axes3D`, or `plotly`; consider showing a 2D contour plot instead.
Maps	QGIS or Python `geopandas`; for simple maps, manual TikZ.

Code or LaTeX example: Tools for Generating Figures

Listing 2: A minimal matplotlib snippet that produces a paper-quality figure.

```
import matplotlib.pyplot as plt
import numpy as np

t = np.linspace(0, 50, 500)
beta, gamma = 0.4, 0.1
# (SIR solution omitted; use scipy.integrate.solve_ivp)

fig, ax = plt.subplots(figsize=(6, 4))
ax.plot(t, S, label='Susceptible',  color='black', linestyle='-')
ax.plot(t, I, label='Infected',     color='black', linestyle='--')
ax.plot(t, R, label='Recovered',    color='black', linestyle=':')
ax.set_xlabel('Time (days)')
ax.set_ylabel('Number of individuals')
ax.set_title('Time evolution of an SIR epidemic ($\\beta=0.4$, $\\gamma=0.1$)')
ax.legend
ax.grid(True, alpha=0.3)
plt.tight_layout
plt.savefig('sir_dynamics.pdf', dpi=300, bbox_inches='tight')
```

Note: Why `savefig` settings matter

`dpi=300` ensures sharpness on print. `bbox_inches='tight'` prevents whitespace around the figure. Saving as PDF (vector) is preferred over PNG (raster) for line plots; PNG is appropriate only for images that are inherently rasterised (photographs, satellite images).

Tables: Often Better Than Figures

Core explanation, worked reasoning, and application

Tables: Often Better Than Figures

If you have exactly four data points to compare across three categories, a table is almost always clearer than a figure. Tables are underused in competition papers because students learn early that “figures are more impressive”. Judges, in fact, value tables: they let the reader see the actual numbers.

Method: Method: Five Rules for Strong Tables

1. **Use `booktabs`.** The `\toprule`, `\midrule`, `\bottomrule` commands produce typographically correct horizontal rules. Vertical rules look amateur; avoid them.
2. **Units in column headers, not in every cell.** “Time (s)” as a header is better than “5.2 s” in every row.
3. **Consistent decimal places** in numerical columns. “5.20” looks better than “5.2” in a column where most entries have two decimals.
4. **Right-align numbers; left-align text.** Aligned decimal points make comparison visual.
5. **Caption above the table.** Unlike figures (caption below), tables conventionally carry their caption above.

Worked Example: Example 3: A Strong Table

Validation of the queuing model: predicted vs simulated average wait times across four arrival-rate scenarios.

Arrival rate λ (/hr)	Predicted W_q (min)	Simulated W_q (min)	Relative error
8	0.67	0.65	-3.0%
10	1.33	1.31	-1.5%
12	4.00	3.94	-1.5%
14	18.67	18.21	-2.5%

Caption above the table; `booktabs` *rules; aligned decimals; units in headers.*

Multi-Panel Layouts

For figures with two or three related panels (e.g., baseline vs intervention), use `subcaption` for proper labelling (a), (b), (c).

Code or LaTeX example: Multi-Panel Layouts

Listing 3: A two-panel figure with sub-captions.

```
\begin{figure}[H]
    \centering
    \begin{subfigure}[t]{0.45\textwidth}
        \includegraphics[width=\textwidth]{baseline.pdf}
        \caption{Baseline configuration ($N=1$ server).}
        \label{fig:baseline}
    \end{subfigure}
    \hfill
    \begin{subfigure}[t]{0.45\textwidth}
        \includegraphics[width=\textwidth]{intervention.pdf}
        \caption{Two-server configuration.}
        \label{fig:intervention}
    \end{subfigure}
    \caption{Queue dynamics under (a) baseline staffing and (b) the recommended two-server
     intervention. The peak queue length drops by 76\% under intervention.}
    \label{fig:queue-comparison}
\end{figure}
```

Cross-References, Citations, and the Bibliography

Core explanation, worked reasoning, and application

Method: Method: Use `cleveref` for Cross-References

- Label every figure, table, equation, and section that you intend to reference.
- Use `\cref` from `cleveref`; it automatically picks the right word: "Fig. 3", "Table 2", "Eq. (5)", "Section 4".
- Never write "see the figure above" — after a rewrite, "above" may no longer be true. Always reference by number.

Method: Method: Citation Practice

1. Use BibTeX. Maintain a single `references.bib` file.
2. For each citation, capture: author, title, journal/conference/publisher, year, volume/pages, DOI/URL.
3. Cite at the point of claim: "...logistic growth (Murray, 2002)" or "...Murray [3] develops this further".
4. Include 4–10 references for an IMMC paper. Fewer than 4 reads as isolation; more than 15 (without good reason) reads as padding.

Note: On citation style

Use IEEE numerical or APA author-date; pick one and be consistent. The choice matters less than consistency. Common acceptable sources:

- **Textbooks.** Giordano *A First Course in Mathematical Modeling*; Hillier & Lieberman *Introduction to Operations Research*; Bertsekas *Nonlinear Programming*.
- **Peer-reviewed papers.** Search via Google Scholar; verify the journal exists.
- **Government / authoritative sources.** Census data, WHO, transport authorities — cite the URL and access date.
- **Wikipedia.** Acceptable for definitions only; for any substantive claim, find the underlying source Wikipedia cites and cite that instead.

The Submission-Day Workflow

Core explanation, worked reasoning, and application

The Submission-Day Workflow

The final 6–12 hours of a competition are not for new analysis; they are for integration, polishing, and submission. A team that has rehearsed the workflow finishes calmly; one that has not, finishes in panic.

Method: Method: A Rehearsable Final-Day Workflow

1. **T**-12 **hours:** Freeze the modeling. No new models from this point. All effort goes to writing and presentation.
2. **T**-12 **to T**-8**:** Polish prose pass. Apply W4 techniques. One teammate is the dedicated polisher.
3. **T**-10 **to T**-6**:** Finalise figures and tables. Regenerate any low-resolution images; verify captions; cross-check numbers between figures and text.
4. **T**-8 **to T**-4**:** Verify cross-references. Use `\cref`; ensure every Fig./Tab./Eq./Sec. number renders.
5. **T**-4 **to T**-2**:** Rewrite the Summary one last time (see W2). This deserves its own dedicated session.
6. **T**-2 **to T**-1**:** Bibliography check, spell check, page-count check (HiMCM has page limits; IMMC has soft conventions).
7. **T**-1 **to T**-0.5**:** Compile final PDF. Verify it opens correctly on a different machine.
8. **T**-0.5**:** Submit. *Never* aim for the deadline; aim to submit half an hour early.

Warning: Warning: The Common Final-Hour Failures

Each of these has cost real teams their submission:

- **LaTeX compilation timeout** from a heavy TikZ figure introduced late. *Solution:* compile after every major addition; never add a 500-line TikZ figure in the last hour.
- **Bibliography format errors** that look fine in the editor but break the compile. *Solution:* compile the bibliography on Day 2, not Day 4.
- **Page-count overflow** discovered too late. *Solution:* monitor page count daily; tighten language preemptively.
- **File-name violations** (contest requires Team_ID_Problem_X.pdf; team submits paper.pdf). *Solution:* read the submission rules at hour 1 and again at hour T-2.
- **Internet outage during submission.** *Solution:* submit early. Always have a backup of the PDF on a personal device.

Method: Method: Role Allocation for the Final Phase

A 4-person team should explicitly assign final-phase roles:

- **Polisher:** reads every section line-by-line for prose quality. The strongest writer.
- **Validator:** re-checks every number, equation, and figure caption against the source.
- **Submitter:** handles compilation, file naming, and the submission portal. Owns the deadline.
- **Floater:** handles whatever comes up; final read of the Summary; sanity checks.

A 2–3 person team merges these. The important point: the roles are explicit, not implicit. Otherwise multiple people end up doing the same thing while no-one fixes the bibliography.

Workshop Exercises

Core explanation, worked reasoning, and application

Exercises: Exercise W5.1: Build a Template

Before any practice runs, build a complete LaTeX template containing all section headings, dummy variables/captions, and your team's intended formatting. The template should compile cleanly out of the box with no warnings. Save it; reuse it for every practice.

Exercises: Exercise W5.2: Reproduce a Strong Figure

Find one figure from an Outstanding paper that you think is exceptional. Recreate it in Python `matplotlib` or TikZ from scratch. Match every element: axes, units, font sizes, legend placement, colour. This is the fastest way to learn what publication-quality looks like.

Exercises: Exercise W5.3: Write Five Captions

For each of the following hypothetical figures, write a paper-quality caption (1–3 sentences):

1. A scatter plot of car arrivals over a day, fitted by a Poisson model.
2. A tornado diagram of sensitivity to five parameters.
3. A network graph showing the proposed bicycle highway routes.
4. A heatmap of predicted commute times across the city.
5. Side-by-side simulated outcomes (baseline / intervention).

Exercises: Exercise W5.4: Mock Submission

Run a one-hour mock submission drill. Starting from a known LaTeX project:

- Add one new figure, one new equation, one new section.
- Re-verify all cross-references.
- Recompile cleanly.
- Rename the file according to a (made-up) contest naming convention.
- Upload to a shared drive folder named “submission”.

Time each step. The full drill should take under one hour; any step that takes longer is a process problem to fix before the real competition.

Exercises: Exercise W5.5: Find Three Visual Improvements in Your Paper

Take your most recent practice paper. Identify three visual choices that could be improved: a figure that needs a better caption, a table that should use `booktabs`, an equation that should be displayed instead of inline, etc. Implement the fixes and check that the paper still compiles.

Closing the Writing Class

Core explanation, worked reasoning, and application

Note: The Five-Class Arc

1. **W1:** The four pillars of the rubric, the standard paper structure, the way Outstanding papers read.
2. **W2:** The Summary: the single most important page.
3. **W3:** Assumptions, Validation, Sensitivity — the trust chain.
4. **W4:** English academic writing style, voice, hedging, connectives, Chinglish patterns.
5. **W5:** LaTeX, figures, tables, the submission workflow.

Closing the Writing Class

The five classes amount to roughly thirty hours of work to internalise. By comparison, learning a new modeling method (regression, simulation, optimisation) takes longer and contributes a smaller share of the score. Writing is the most leverage-rich investment the team can make.

Note: Three Pieces of Final Advice

1. **Read at least three Outstanding papers before your first competition.** Nothing else compresses learning so quickly.
2. **Practise the workflow under time pressure.** A two-day mock submission a month before the real competition reveals every weak link.
3. **Designate the polisher early.** The teammate who will be responsible for prose quality should know that role from Day 1. They should keep the W4 reference sheet open.

Closing the Writing Class

The Writing Class ends here. The next step — the only step that matters now — is to write papers. Many papers. Imperfect papers. Improved papers. Polished papers. Submitted papers.

Appendix A: One-Page Compilation Checklist

Core explanation, worked reasoning, and application

Appendix A: One-Page Compilation Checklist

- ☐ Document compiles with no errors and no critical warnings.
- ☐ All cross-references resolve (no "Fig. ??").
- ☐ Page count within contest limit (HiMCM: 25 pages; IMMC: check yearly).
- ☐ Font and margin per contest rules.
- ☐ Team ID on every page (if required).
- ☐ Page numbers visible.
- ☐ Title page formatted per contest.
- ☐ Summary on its own page (one page).
- ☐ All figures captioned (caption below figure).
- ☐ All tables captioned (caption above table).
- ☐ All equations referenced from text are numbered.
- ☐ Bibliography in consistent style (APA, IEEE, etc.).
- ☐ File named per contest convention.
- ☐ PDF opens correctly on a non-team machine.
- ☐ Upload completed at least 30 minutes before deadline.

Appendix B: A Minimal Template Snippet

Core explanation, worked reasoning, and application

Code or LaTeX example: Appendix B: A Minimal Template Snippet

```latex
\documentclass[12pt,a4paper]{article}
\usepackage[margin=2.5cm]{geometry}
\usepackage{amsmath, amssymb, amsthm, graphicx, booktabs,
            caption, subcaption, float, hyperref, cleveref}
\usepackage[numbers]{natbib}

\title{Team N12345: [Your Paper Title]}
\author{}
\date{}

\begin{document}


\begin{abstract}
\end{abstract}

\section{Introduction}
\section{Assumptions and Notation}
\section{Model Development}
\section{Results}
\section{Validation}
\section{Sensitivity Analysis}
\section{Strengths and Weaknesses}
\section{Conclusion}
```

Code or LaTeX example: Appendix B: A Minimal Template Snippet

```latex
\bibliographystyle{IEEEtran}
\bibliography{references}

\appendix
\section{Code Listing}

\end{document}
```

Appendix C: Tools Quick-Reference

Core explanation, worked reasoning, and application

Appendix C: Tools Quick-Reference

Task	Tool
LaTeX editor	Overleaf (online, free tier sufficient); TeXstudio (local).
Version control	Overleaf's Git integration; or Git + GitHub if working locally.
Figure generation	Python `matplotlib`; TikZ for math diagrams.
Data analysis	Python `pandas`, `numpy`, `scipy`.
Reference management	BibTeX + `references.bib`.
Spell-check	Overleaf's built-in; or pass through Grammarly.
PDF preview	Built-in Overleaf preview; SumatraPDF for desktop.
Backup	Daily zip download of the Overleaf project.