

Lectures in Mathematical Modeling

Lecture 1: Modeling Change

Kenneth Sok Kin Cheng

Revised 20 August 2026

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.

1 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.

Launch: What would you model?

[10 min]

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.

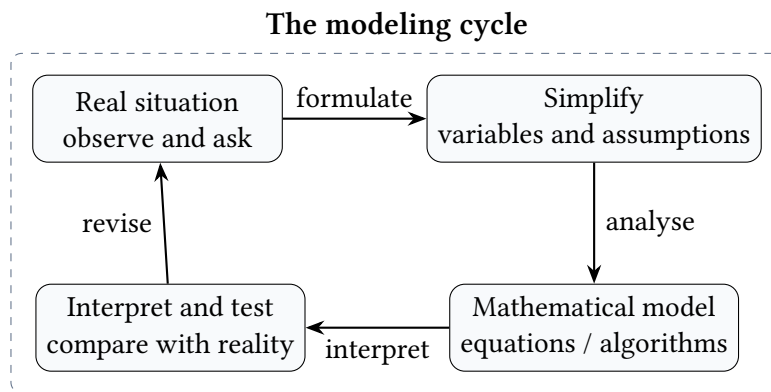


Figure 1: Modeling is iterative. A failed check is not the end; it tells us what to revise.

Four questions to ask throughout a modeling problem

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?

2 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 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

$$e = 0.01622m.$$

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

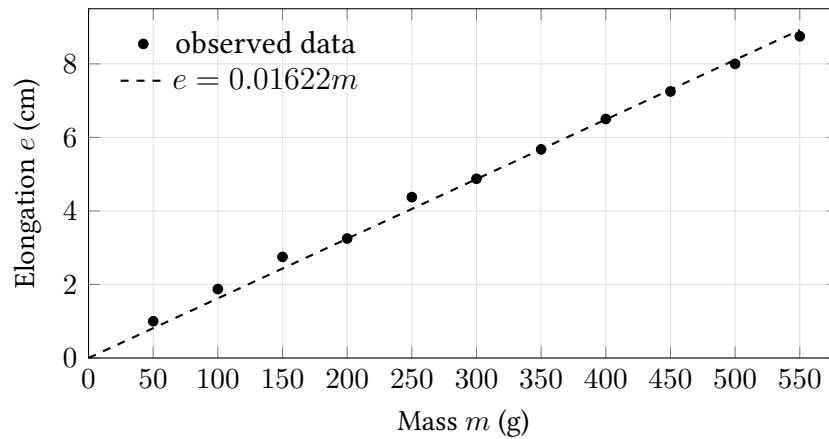


Figure 2: The points lie close to a straight line through the origin, so proportionality is plausible.

Model criticism: 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?

3 Modeling discrete change

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

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

For a sequence $a_0, a_1, a_2, \dots$, 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.

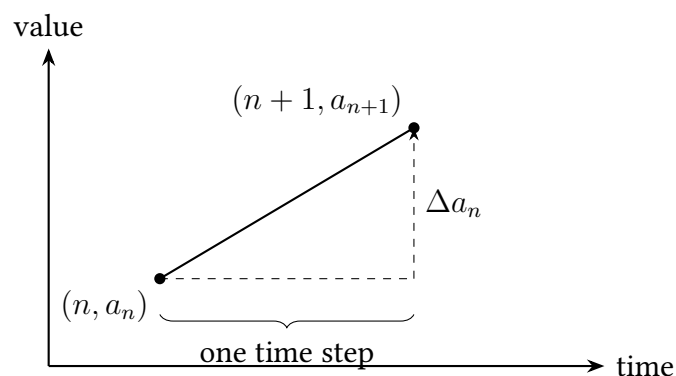


Figure 3: A first difference is the vertical change between consecutive observations.

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:

$$\Delta a_n = 0.01a_n.$$

Therefore

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

The first values are

$$1000, 1010, 1020.10, 1030.30, \dots$$

Repeated substitution suggests the closed form

$$a_n = 1000(1.01)^n.$$

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

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 $\Delta a_n = \underline{\hspace{2cm}}$.
3. Write the recurrence and calculate a_1, a_2, a_3 .
4. Without calculating every term, predict the long-term behavior.

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
$ r > 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.

4 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 $r \neq 1$,

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

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 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 n th 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.

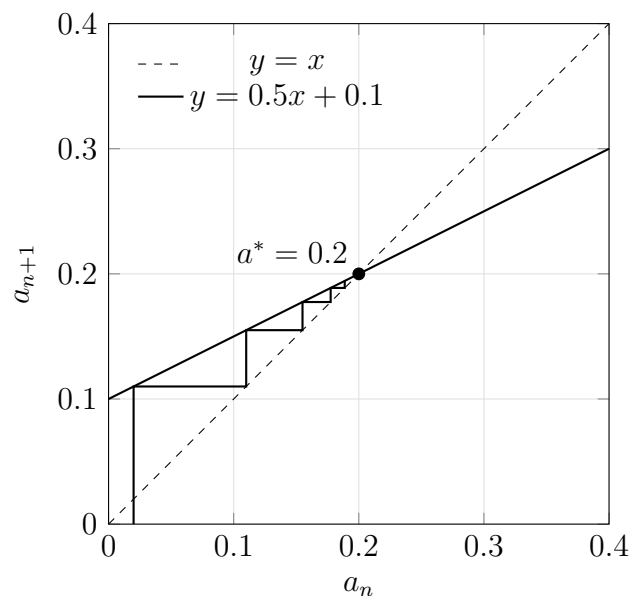


Figure 4: A cobweb plot: the iterations approach the intersection of $y = x$ and $y = 0.5x + 0.1$.

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.

5 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 5: Yeast culture

Suppose

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

The first values are approximately

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

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

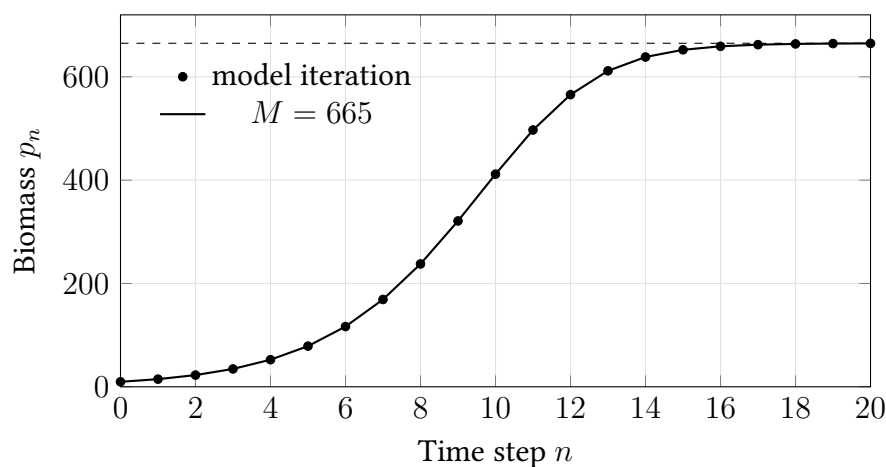


Figure 5: The population approaches the carrying capacity rather than growing without bound.

Model criticism: 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.

Model comparison**[8 min]**

Two teams model the same yeast population.

$$\text{Team A: } p_{n+1} = 1.45p_n,$$

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

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

6 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).

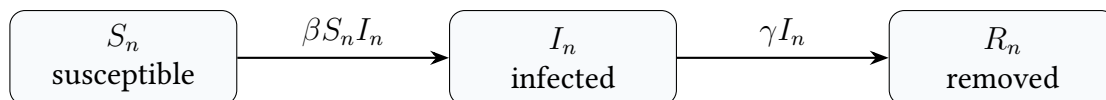


Figure 6: Compartment flow in the SIR model.

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

$$S_{n+1} = S_n - \beta S_n I_n,$$

$$I_{n+1} = I_n + \beta S_n I_n - \gamma I_n,$$

$$R_{n+1} = R_n + \gamma I_n.$$

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.

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

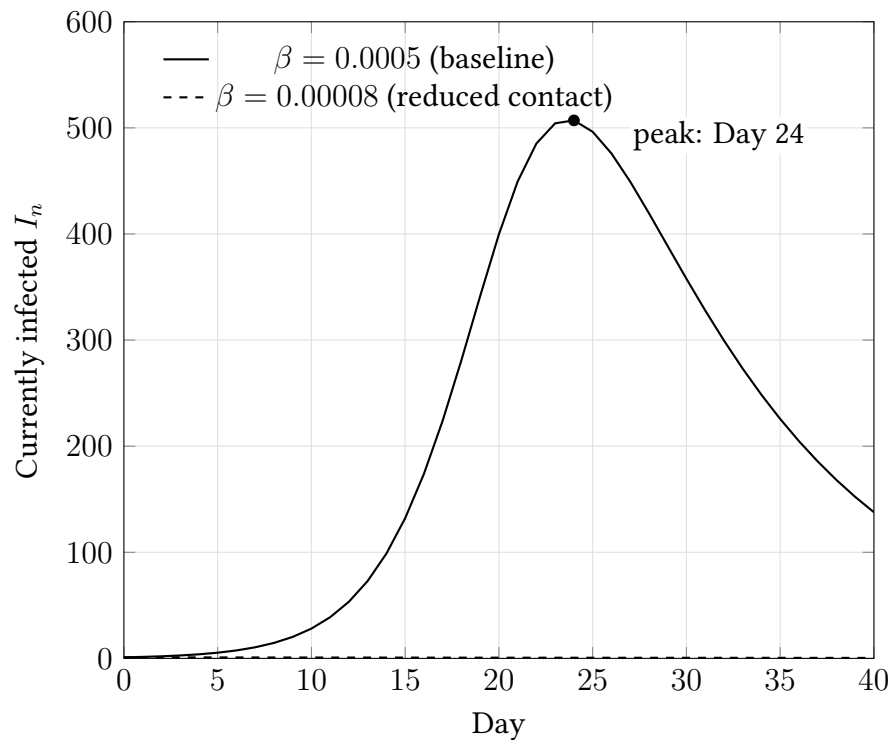


Figure 7: Reducing the transmission parameter below the threshold prevents initial growth.

Competition-style communication

[10 min]

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.

7 Lesson 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 $\text{next} = \text{current} + \text{change}$.
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 _____.”

8 Practice problems

Core practice

1. A water tank contains 800 L and loses 6% of its current volume each hour.
 - (a) Write a recurrence for the volume V_n .
 - (b) Find V_5 and the long-term behavior.
 - (c) 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.
 - (a) Formulate $M_{n+1} = rM_n + b$.
 - (b) Find the equilibrium membership.
 - (c) 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 $x_{n+1} = 1.15x_n - 30$, $x_0 = 250$.
 - (a) Find the equilibrium.
 - (b) Calculate the first four terms.
 - (c) 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.

Practice check and hints

For Questions 1–5, useful checks are:

1. $V_{n+1} = 0.94V_n$, $V_5 \approx 587.12$ L, and $V_n \rightarrow 0$ unless a physical floor or different loss mechanism is imposed.
2. $M_{n+1} = 0.92M_n + 15$ has stable equilibrium $M^* = 187.5$.
3. The value 1000 is the saturation level represented by the model.
4. $x^* = 200$ and $(x_1, x_2, x_3, x_4) = (257.5, 266.125, 276.044, 287.450)$. A money model still needs domain and stopping rules, even though this initial value moves upward.
5. Successive download ratios are close to 1.45; a defensible next estimate is about 508, conditional on that pattern continuing.

For Question 6, a complete response links plausible changes in β to forecast trajectories, intervention timing or allocation, and a stated sensitivity range.

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.

A 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

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

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

Model audit card

Audit before trusting or communicating a model			
Purpose	What question or decision does the model support?	Data	Which inputs are observed, derived, or simulated?
Assumptions	Which simplifications may change the answer?	Units	Are dimensions and scales consistent?
Verification	Were equations, code, constraints, and limiting cases checked?	Validation	Is performance tested against independent evidence?
Uncertainty	How do measurement, parameter, solver, and structural errors affect conclusions?	Ethics	Who benefits, who bears risk, and what fairness criterion matters?
Reproducibility	Are data provenance, versions, seeds, and transformations recorded?	Communication	Are claims conditional, traceable, and decision-relevant?

Further reading

- GAIMME: Guidelines for Assessment and Instruction in Mathematical Modeling Education (SIAM / COMAP)
- 18.03SC Differential Equations (MIT OpenCourseWare)

Sources, provenance, and licence

These notes follow the iterative modeling and assessment perspective in the SIAM/COMAP GAIMME report. Competition-specific remarks are a snapshot checked on 20 August 2026 against the official COMAP HiMCM instructions; readers should verify current rules before submission. Unless a source is explicitly identified, classroom datasets and numerical scenarios are synthetic or simulated teaching examples and are not empirical claims about a real institution. Third-party references retain their own terms. Original teaching material by Kenneth Sok Kin Cheng is released under CC BY-NC-SA 4.0.

Student response guide and checking criteria

For closed calculations, show the definition, substitution, units, and at least one independent check; the worked examples in this unit are method hints, but you must complete the arithmetic yourself. Open modeling responses are checked on five dimensions: a clear purpose and output; stated data and provenance; defensible assumptions and scope; traceable computational verification and model validation; and alignment among uncertainty, limitations, and the recommendation. Support each dimension with concrete evidence rather than writing only “completed.”