Lectures in Mathematical Modeling

Supplementary 9: Linear Algebra Essentials for Modeling

Kenneth Sok Kin Cheng

Revised 20 August 2026

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.

1 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.

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.

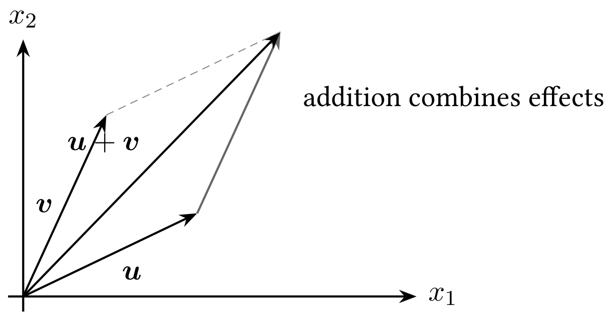


Figure 1: Vector addition models combined effects, such as displacement, resource use, or accumulated changes across components.

2 Vectors as data packages

Vectors, dot products, and norms

A vector in $\mathbb{R}^n$ is an ordered list

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$,

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + \cdots + u_n v_n, \qquad \|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}.$$

The dot product measures alignment. If $\boldsymbol{u} \cdot \boldsymbol{v} = 0$, the vectors are **orthogonal**.

Worked example 1: resource-use vector

A delivery robot uses battery, time, and staff attention. For one route, suppose

$$\boldsymbol{x} = \begin{pmatrix} 18 \\ 42 \\ 3 \end{pmatrix}$$

means 18 battery units, 42 minutes, and 3 staff interventions. A cost vector

$$\boldsymbol{c} = \begin{pmatrix} 0.4 \\ 1.2 \\ 8 \end{pmatrix}$$

assigns weights to each component. The total weighted cost is

$$\boldsymbol{c} \cdot \boldsymbol{x} = 0.4(18) + 1.2(42) + 8(3) = 81.6.$$

This is the same idea as an LP objective $\boldsymbol{c} \cdot \boldsymbol{x}$ in Lecture 6.

Linear combination and span

A **linear combination** of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ is

$$c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k.$$

The set of all such combinations is the **span**, written

$$\mathrm{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k).$$

Vectors are **linearly independent** if the only way to make $\mathbf{0}$ is to use all zero coefficients.

Worked example 2: detecting dependence

Let

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}.$$

Since

$$\boldsymbol{v}_1 + 2\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} = \boldsymbol{v}_3,$$

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.

3 Matrices as structured models

Matrix multiplication

An $m \times n$ matrix $\mathbf{A}$ maps vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$ by

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x}.$$

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 $\mathbf{A}$.

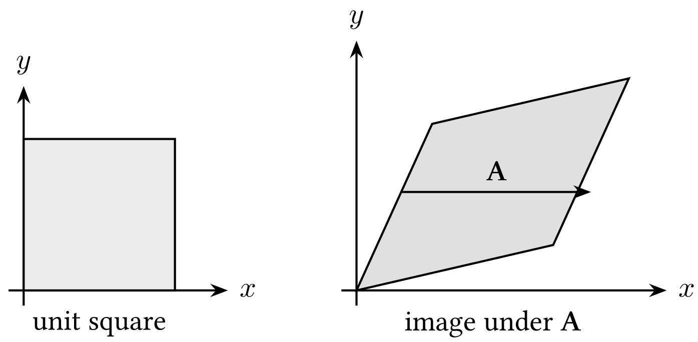


Figure 2: A matrix can be viewed as a transformation. The columns of $\mathbf{A}$ show where the unit basis vectors move.

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

$$\boldsymbol{x}_0^\top = (1, 0).$$

After one day,

$$\boldsymbol{x}_1^\top = \boldsymbol{x}_0^\top \mathbf{P} = (0.8, 0.2).$$

After two days,

$$\boldsymbol{x}_2^\top = \boldsymbol{x}_0^\top \mathbf{P}^2 = (0.72, 0.28).$$

Matrix powers describe repeated transitions, which is the core of Markov-chain modeling.

Model criticism: 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.

4 Solving systems by Gaussian elimination

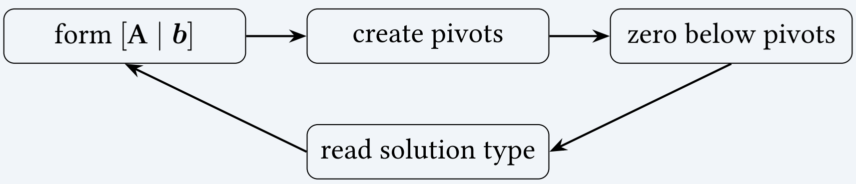
Linear system

A system of m linear equations in n unknowns can be written as

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}.$$

Gaussian elimination uses row operations on the augmented matrix $[\mathbf{A} \mid \boldsymbol{b}]$ to reveal pivots, free variables, and consistency.

Gaussian elimination route



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

$$(x, y, z) = \left(\frac{5}{4}, \frac{5}{4}, \frac{9}{4}\right).$$

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 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.

5 Least squares as linear algebra

Residual vector and normal equations

For data (x_i, y_i) and a line $y = a + bx$, define

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Predictions are $\hat{\boldsymbol{y}} = \mathbf{X}\boldsymbol{\beta}$ and residuals are

$$\boldsymbol{r} = \boldsymbol{y} - \mathbf{X}\boldsymbol{\beta}.$$

Least squares chooses $\boldsymbol{\beta}$ to minimise $\|\boldsymbol{r}\|^2$. The normal equations are

$$\mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} = \mathbf{X}^\top \boldsymbol{y}.$$

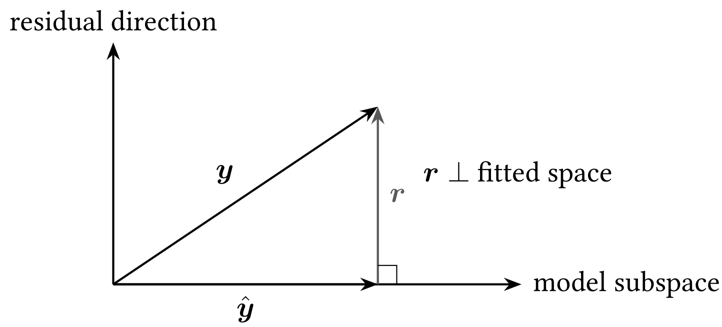


Figure 3: Least squares chooses the fitted vector $\hat{\boldsymbol{y}}$ so that the residual vector is orthogonal to the model subspace. This gives $\mathbf{X}^\top \boldsymbol{r} = \mathbf{0}$.

Worked example 6: fitting a line

Fit $y = a + bx$ to the data

$$(0, 1),\ (1, 2),\ (2, 2),\ (3, 4).$$

Then

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 4 \end{pmatrix}.$$

Compute

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}, \qquad \mathbf{X}^\top \boldsymbol{y} = \begin{pmatrix} 9 \\ 18 \end{pmatrix}.$$

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 criticism: 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.

6 Eigenvalues, stationary states, and long-run behavior

Eigenvalue and eigenvector

For a square matrix $\mathbf{A}$, a non-zero vector $\boldsymbol{v}$ is an **eigenvector** with **eigenvalue** λ if

$$\mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

The direction of $\boldsymbol{v}$ is preserved by the transformation; only its scale changes.

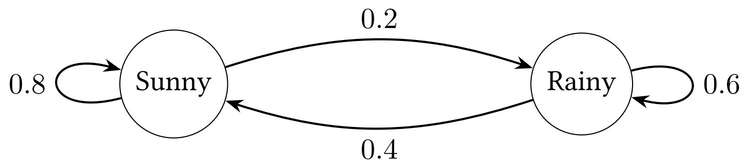


Figure 4: A two-state Markov chain. The stationary distribution is a left eigenvector of the transition matrix with eigenvalue 1.

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 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.

7 Differentiated modeling studio

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.

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?

8 Exercises

Exercise 1: Vectors

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, $\boldsymbol{x} = (8, 20, 2)$ records distance, time, and number of delays. If $\boldsymbol{c} = (1.5, 0.8, 10)$, interpret $\boldsymbol{c} \cdot \boldsymbol{x}$.

Vector checks

The unit vector is $(3/5, 4/5, 0)$. The dot product in Question 2 is 8, so $\cos\theta = 8/\sqrt{294}$ and $\theta \approx 62.19°$. The determinant formed from the three vectors in Question 3 is -2, so they are independent. The weighted delivery cost is 48 in the cost scale defined by $\boldsymbol{c}$.

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

$$\begin{cases} x_1 + x_2 - x_3 + 2x_4 = 0, \\ 2x_1 - x_2 + x_3 + x_4 = 0. \end{cases}$$

Matrix and system checks

$$\mathbf{AB} = \begin{pmatrix} 7 & 4 \\ 19 & 8 \end{pmatrix}, \qquad \mathbf{BA} = \begin{pmatrix} 5 & 10 \\ 7 & 10 \end{pmatrix}.$$

Question 2 has $(x, y, z) = (4/5, 3, 8/5)$. In Question 3 the left side of the second equation is twice the first, but $3 \neq 2(1)$, so the system is inconsistent. For Question 4, with free parameters $s, t \in \mathbb{R}$,

$$(x_1, x_2, x_3, x_4) = s(0, 1, 1, 0) + t(-1, -1, 0, 1).$$

Exercise 3: Modeling connections

1. Fit $y = a + bx$ to $(1, 2), (2, 3), (3, 5), (4, 4), (5, 6)$ using $\mathbf{X}^\top \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^\top \boldsymbol{y}$.

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."

Modeling checks

The fitted line is $\hat{y} = 1.3 + 0.9x$. The stationary distribution is $(19/41, 13/41, 9/41)$. Row's mixed strategy is $(1/3, 2/3)$, obtained by equating the two column payoffs. For dimensional analysis, connect the number of free exponent directions to the nullity of the dimension matrix via rank–nullity.

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 $\boldsymbol{v}$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ, then $\boldsymbol{v}$ 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.

Extension hints

The eigenpairs are $\lambda = 5$ with direction $(1, 1)$ and $\lambda = 2$ with direction $(1, -2)$. For rotations, compose the geometric actions or use angle-addition identities. Question 3 follows by induction on k. Three valid invertibility equivalents are nonzero determinant, full rank, and trivial null space.

Appendix A: Algorithm summary

Task	Practical method
Solve $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$	Row-reduce $[\mathbf{A} \mid \boldsymbol{b}]$, 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 $\mathbf{X}^\top\boldsymbol{y}$, 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 $(\mathbf{A} - \lambda\mathbf{I})\boldsymbol{v} = \mathbf{0}$.

Appendix B: Connection map to the main lectures

Tool	**Where it appears**
Dot product and norm	Lecture 3 residual vector; Lecture 6 linear objective $\boldsymbol{c} \cdot \boldsymbol{x}$.
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.

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

- 18.06SC Linear Algebra (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.

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.”