Lectures in Mathematical Modeling

Supplementary 10: Multivariable Calculus and Lagrange Multipliers

Kenneth Sok Kin Cheng

Revised 20 August 2026

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.

1 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.

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.

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.

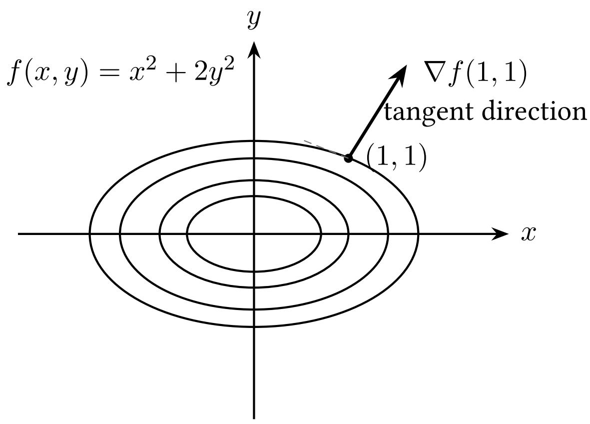


Figure 1: The gradient crosses a level curve at right angles. Moving along the curve changes direction but not the function value.

2 Partial derivatives and local sensitivity

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 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.

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 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.

3 Gradient and directional change

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 $\boldsymbol{u}$, the **directional derivative** is

$$D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u}.$$

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 3: moving through a temperature field

Let

$$T(x, y) = 30 - 0.5x^2 - y^2 + 2x$$

be temperature in °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 $\sqrt{5}$ °C per kilometre. If a student walks east, $\boldsymbol{u} = (1, 0)$, then $D_{\boldsymbol{u}} T = 1$ °C/km.

Model criticism: 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.

Chain rule along a path

If a moving point follows $\boldsymbol{r}(t) = (x(t), y(t))$, then

$$\frac{d}{dt} f(x(t), y(t)) = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} = \nabla f \cdot \boldsymbol{r}'(t).$$

This separates the field sensitivity ∇f from the movement $\boldsymbol{r}'(t)$.

4 Unconstrained optimization and the Hessian

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.

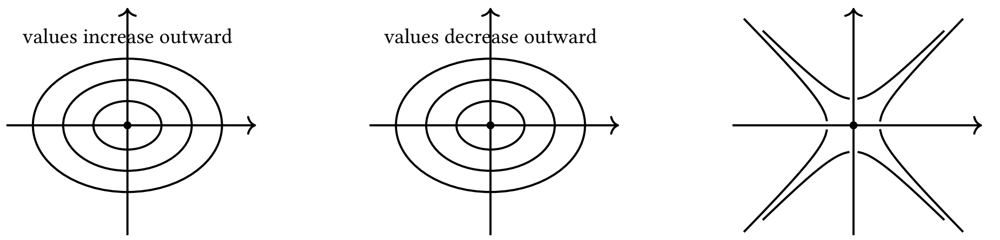
Hessian test in two variables

The Hessian is

$$\mathbf{H}f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}.$$

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



minimum: bowl   maximum: hill   saddle: mixed curvature

Figure 2: Contour patterns distinguish a local minimum, local maximum, and saddle. The Hessian records this local curvature algebraically.

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 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 $\operatorname{diag}(2,-2)$ 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.

5 Lagrange multipliers: optimizing on a constraint

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.

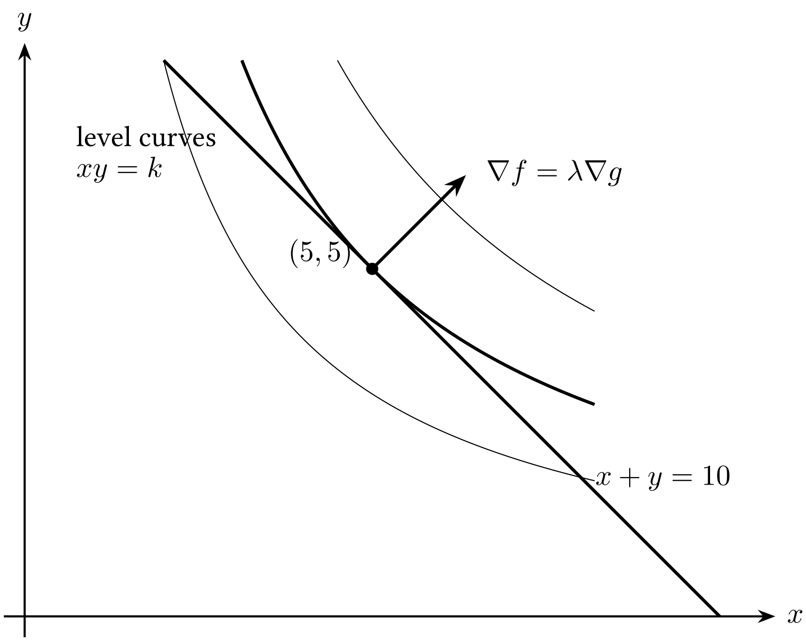


Figure 3: The maximum of $f(x,y) = xy$ subject to $x + y = 10$ occurs where the highest attainable level curve is tangent to the constraint, namely at $(5,5)$.

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 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?

6 Shadow prices and a production-allocation model

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.

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.

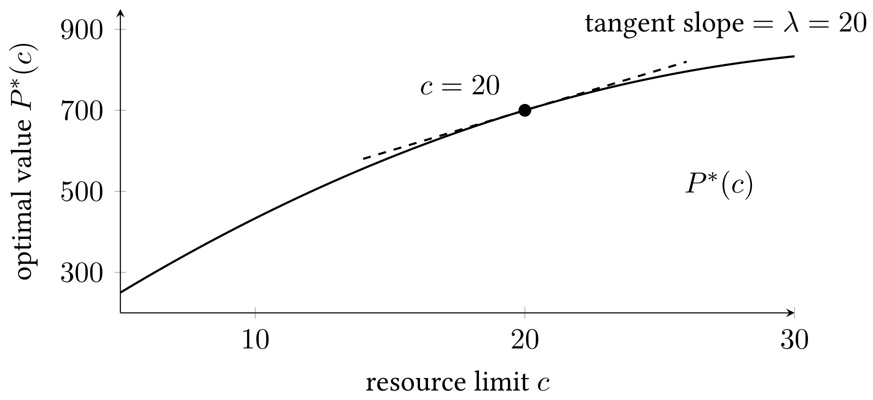


Figure 4: The shadow price is the slope of the optimal-value curve with respect to the resource limit. It is a local sensitivity, not a permanent price.

Model criticism: 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.

7 Differentiated optimization studio

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.

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.

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.

8 Communicating the result

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.

9 Practice problems

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.

Partial-derivative checks

For Question 1,

$$f_x = 2xy + ye^{xy}, \qquad f_y = x^2 + xe^{xy}.$$

At $(2, 1)$, $\nabla T = (-0.8, -1)$; the unit direction toward $(5, 5)$ is $(3/5, 4/5)$, giving directional derivative -1.28. Linearization of $\sqrt{x^2 + y^2}$ about $(3, 4)$ gives 5.002. The final answer should attach cost per labour-hour and cost per kilogram to the two partial derivatives.

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.

Critical-point checks

The critical points and classifications are: $(2, -1)$, a minimum; $(-1, 0)$, a saddle; for h, $(0, 0)$ is a saddle and $(1, 1)$ is a local minimum; for q, $(0, 0)$ is a saddle and $(1, 1)$ and $(-1, -1)$ are local minima. In every listed case the Hessian determinant is nonzero at the critical point, so the test is conclusive.

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 $x, y \geq 0$.

Lagrange-multiplier checks

The maxima in Question 1 are $(2, 1)$ and $(-2, -1)$, with value 2. The closest point in Question 2 is $(3/7, 6/7, 9/7)$. Question 3 has $(x, y, z) = (4, 4, 4)$ and maximum 64. Question 4 has $(x, y) = (37/10, 23/5)$ and maximum $409/20$; compare with the feasible endpoints when verifying it.

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.

Modeling checks

The unconstrained point $(580/23, 660/23)$ violates $2x + y \leq 60$. On the binding line, the continuous optimum is

$$(x, y) = \left(\frac{160}{9}, \frac{220}{9}\right), \qquad \Pi = \frac{18200}{9} \approx 2022.22.$$

It exceeds the best axis candidates, $(0, 35)$ with value 1225 and $(30, 0)$ with value 1350. With the convention $\nabla\Pi = \lambda\nabla(2x + y)$, $\lambda = 110/9$ is the local value of one extra resource unit. For integer production, checking nearby feasible points selects $(18, 24)$ with profit 2022. A complete conclusion states the continuous assumption, boundary comparison, multiplier locality, integer check, and one validation or uncertainty limitation.

Optional extension A: several constraints

For constraints $g_1(\boldsymbol{x}) = c_1, \ldots, g_k(\boldsymbol{x}) = c_k$, 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

If the feasible set contains inequalities such as $g(\boldsymbol{x}) \le c$, 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.

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

- Calculus Volume 3 (OpenStax)
- Nonlinear Constrained Optimization (NEOS Guide)

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.”