Lecture 10: Continuous Optimization and Model Design

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. 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

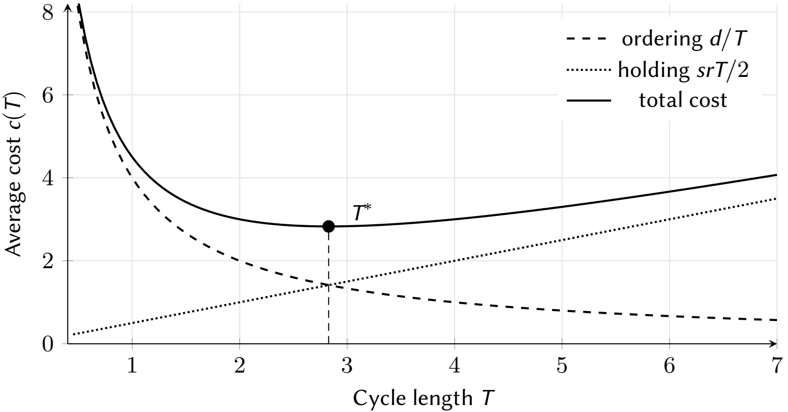
$$c''(T) = \frac{2d}{T^3} > 0,$$

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

$$x^* = 40, \qquad y^* = 20.$$

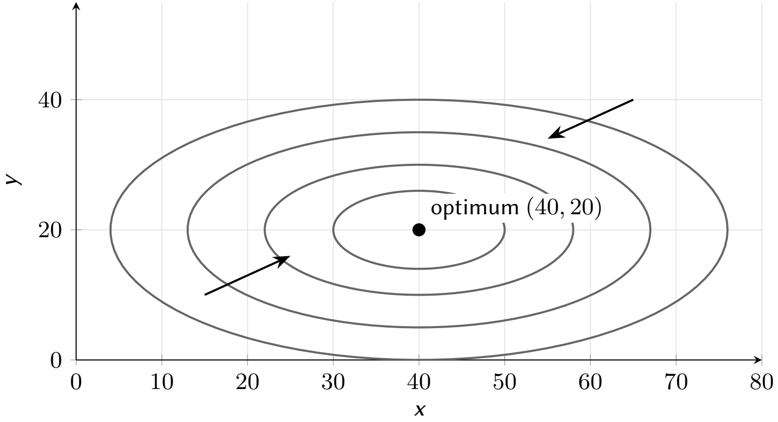
Classify.

$$P_{xx} = -2, \quad P_{yy} = -4, \quad P_{xy} = -1, \quad D = (-2)(-4) - (-1)^2 = 7 > 0.$$

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

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda\,[g(x, y) - c].$$

Solve

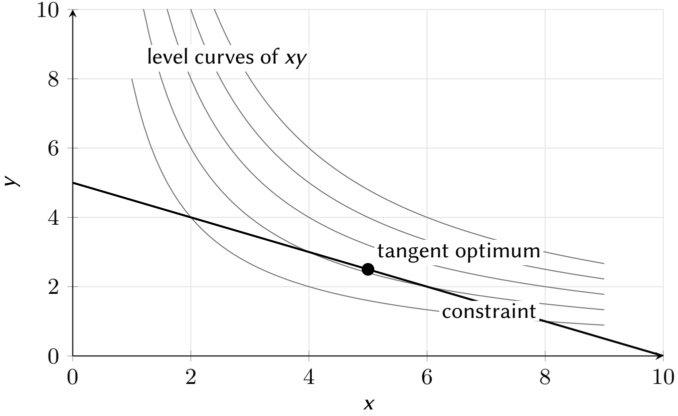
$$\mathcal{L}_x = 0, \qquad \mathcal{L}_y = 0, \qquad \mathcal{L}_\lambda = 0.$$

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

$$A^* = 25 \cdot 50 = 1250 \text{ m}^2.$$

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 $(10, 12)$,

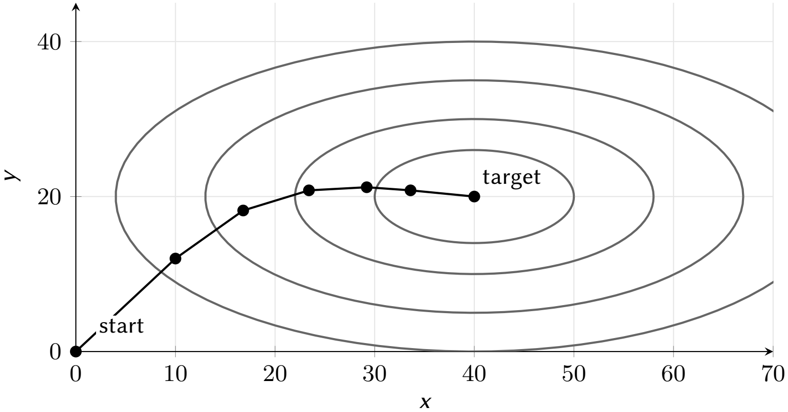
$$\nabla P = (100 - 20 - 12,\ 120 - 48 - 10) = (68, 62),$$

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 $c'(N_b) < 0$,

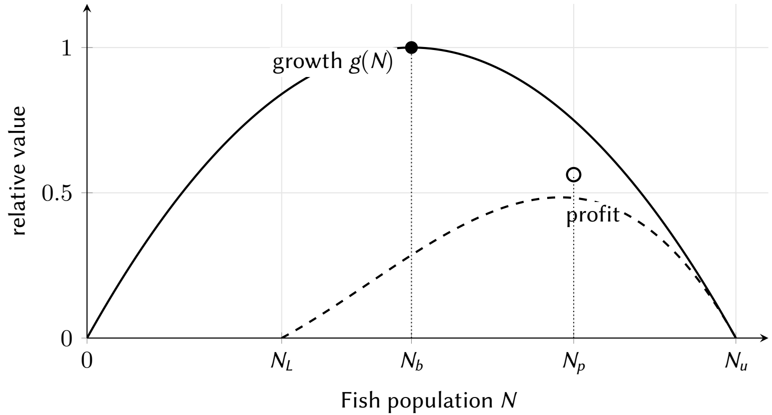
$$\mathrm{TP}'(N_b) = -g(N_b)c'(N_b) > 0.$$

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.