General Optimization

Lecture 6

Maria Titova

August 11, 2026

Econ 8000 — Math Camp

Motivation

Solving an optimization problem requires three separate steps.

- First, we show that a solution exists.
- Second, we find the points that satisfy the necessary conditions for a solution.
- Third, we determine whether any of those points is a global solution.

The problem and existence

Optimization problem

Definition 1.1 (Optimization problem)

Let Θ be a parameter set, let $U \subseteq \mathbb{R}^n$, and let $f : U \times \Theta \to \mathbb{R}$.

For a fixed $\theta \in \Theta$, consider the problem

$$\max_{x \in D(\theta)} f(x;\theta), \qquad D(\theta) \subseteq U.$$

- x is the *choice variable*, θ is the *parameter*, $D(\theta)$ is the *feasible set*, and $f(\,\cdot\,;\theta)$ is the *objective*.
- The *solution set* is $X^*(\theta) = \arg\max_{x \in D(\theta)} f(x;\theta)$; when it is nonempty, its elements share one objective value, the *value* $V(\theta) = f(x^*;\theta)$ for any $x^* \in X^*(\theta)$.

An optimization problem from economics

Example 1.2 (Utility maximization)

A consumer with utility $u : \mathbb{R}^n_+ \to \mathbb{R}$ faces prices $p \in \mathbb{R}^n_{++}$ and wealth $w > 0$:

$$\max_{x \in B(p,w)} u(x), \qquad B(p, w) = \{x \in \mathbb{R}^n_+ : p \cdot x \leq w\}.$$

In the notation of the definition:

- the bundle x is the choice variable and $\theta = (p, w)$ is the parameter;
- the budget set $B(p, w)$ is the feasible set and u is the objective;
- the solution set $X^*(p, w)$ is the demand and the value $V(p, w)$ is the indirect utility.

max **presumes attainment**

Writing max instead of sup presumes that the supremum is attained.

A supremum can be finite without being attained.

- Then $X^*(\theta)$ is empty and $V(\theta)$ is undefined.

The first goal is thus is establishing that a solution exists.

- Otherwise using the notation X^* and V does not make sense.

Local and global maximizers

Definition 1.3 (Local and global maximizers)

Let $D \subseteq \mathbb{R}^n$ and $f : D \to \mathbb{R}$, and let $x^* \in D$.

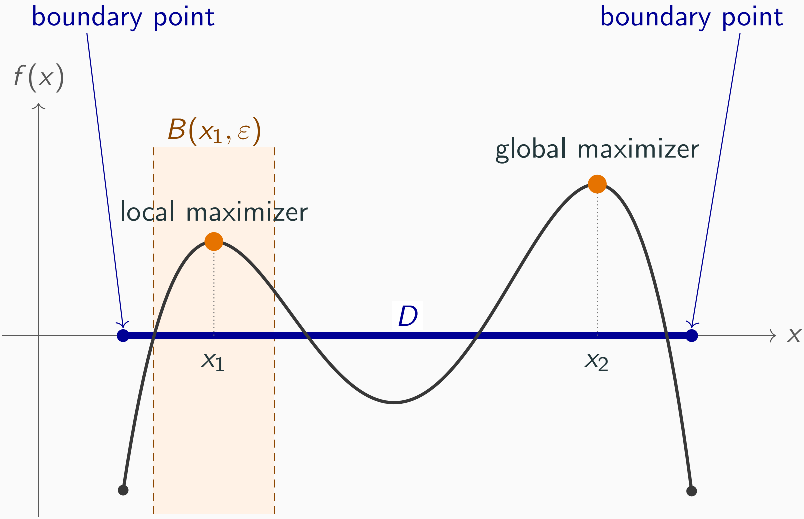
- x^* is a *global maximizer* if $f(x^*) \geq f(x)$ for every $x \in D$.
- x^* is a *local maximizer* if there exists $\varepsilon > 0$ with

  $$f(x^*) \geq f(x) \qquad \text{for every } x \in D \cap B(x^*, \varepsilon),$$

  and a *strict* local maximizer if that inequality is strict for every such $x \neq x^*$.
- x^* is *interior* if $x^* \in \operatorname{int} D$, meaning that $B(x^*, \varepsilon) \subseteq D$ for some $\varepsilon > 0$, and a *boundary point* if every ball $B(x^*, \varepsilon)$ meets both D and its complement.

Minimizers are defined by reversing every inequality.

The definitions, drawn



Weierstrass theorem

Our first existence result comes from of Lecture 2.

Theorem 1.4 (Weierstrass)

Let $D \subseteq \mathbb{R}^n$ be nonempty and compact and let $f : D \to \mathbb{R}$ be continuous.

Then f attains both a maximum and a minimum on D.

Existence on an unbounded set

Compactness is sufficient for existence but not necessary.

A nonnegative orthant and all of $\mathbb{R}^n$ are closed but unbounded.

- A maximum is still attained if the objective becomes sufficiently small far away.

Coercivity for maximization

Definition 1.5 (Coercivity for maximization)

Let $D \subseteq \mathbb{R}^n$ be unbounded.

A function $f : D \to \mathbb{R}$ is *coercive for maximization* on D if

$$x_k \in D, \quad \|x_k\| \to \infty \quad \implies \quad f(x_k) \to -\infty$$

for every sequence (x_k).

Equivalently, for every $M \in \mathbb{R}$ there exists $R > 0$ with

$$f(x) < M \qquad \text{whenever } x \in D \text{ and } \|x\| > R.$$

Coercive existence

Theorem 1.6 (Coercive existence)

Let $D \subseteq \mathbb{R}^n$ be nonempty, closed, and unbounded.

Let $f : D \to \mathbb{R}$ be continuous and coercive for maximization.

Then f attains a maximum on D.

Coercivity controls the tails

$$\forall M \in \mathbb{R}\ \exists R > 0 : |x| > R \implies f(x) < M$$

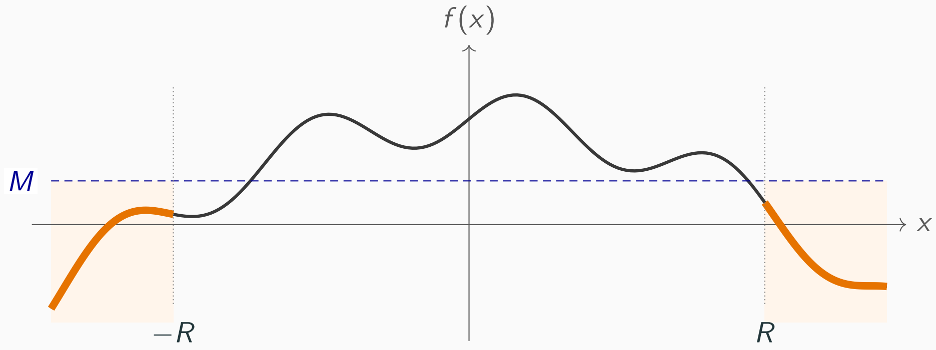
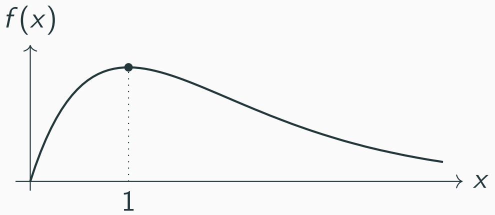


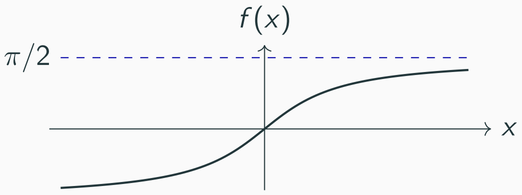
Figure 1.1

- No matter how low the level M is, both tails eventually lie below it.

Weierstrass and coercivity are not necessary



$f(x) = xe^{-x}$ on $[0, \infty)$ attains a maximum at $x^* = 1$, but neither existence result applies: $[0, \infty)$ is not compact, and $f(x) \to 0$, not $-\infty$, so f is not coercive.



$f(x) = \arctan x$ on $\mathbb{R}$ does not attain a maximum (it increases toward the supremum $\pi/2$ but never reaches it). Neither result applies: $\mathbb{R}$ is not compact, and f is not coercive because $f(x) \to \pi/2$ as $x \to \infty$.

Necessary local conditions

Restricting to a line

First-order conditions for functions in $\mathbb{R}^n$ are proved via one-variable calculus.

Lemma 2.1 (One-variable optimality)

Let $\delta > 0$ and let $g : (-\delta, \delta) \to \mathbb{R}$ have a local maximum at 0.

1. *If g is differentiable at 0, then $g'(0) = 0$.*
2. *If g is twice continuously differentiable, then $g''(0) \leq 0$.*

- Part (1) is the interior extremum theorem, proved in Lecture 4.
- For (2), write the Taylor expansion $g(t) - g(0) = \frac{1}{2}g''(0)t^2 + r(t)$ with $|r(t)|/t^2 \longrightarrow 0$; if $g''(0) > 0$, the quadratic term dominates $r(t)$ for small $t \neq 0$, so $g(t) > g(0)$, contradicting the local maximum at 0.

Interior first-order condition

Theorem 2.2 (Interior first-order condition)

Let $U \subseteq \mathbb{R}^n$ be open, let $D \subseteq U$, and let $f : U \to \mathbb{R}$ be differentiable at $x^* \in \operatorname{int} D$.

If x^* is a local maximizer or a local minimizer of f on D, then $\nabla f(x^*) = 0$.

Interior first-order condition

Theorem 2.2 (Interior first-order condition)

Let $U \subseteq \mathbb{R}^n$ be open, let $D \subseteq U$, and let $f : U \to \mathbb{R}$ be differentiable at $x^* \in \operatorname{int} D$.

If x^* is a local maximizer or a local minimizer of f on D, then $\nabla f(x^*) = 0$.

Proof.

Fix $v \in \mathbb{R}^n$ and put $g(t) = f(x^* + tv)$, defined for small $|t|$ since $x^* \in \operatorname{int} D$.

At a local maximizer x^*, $g(t) \leq g(0)$ for small $|t|$, so 0 is a local maximum of g.

By part (1) of the one-variable lemma and the chain rule, $0 = g'(0) = \nabla f(x^*) \cdot v$.

This holds for every $v \in \mathbb{R}^n$; in particular it holds for $v = \nabla f(x^*)$.

For $v = \nabla f(x^*)$, we have $\|\nabla f(x^*)\|^2 = 0 \iff \nabla f(x^*) = 0$. $\square$

Critical points are candidates, not solutions

Definition 2.3 (Critical point)

A point x at which f is differentiable and $\nabla f(x) = 0$ is a *critical point* of f.

It is also called a *stationary point*.

The interior first-order condition is a necessary condition, not a sufficient one:

$$\text{interior local maximizer or interior local minimizer} \implies \text{critical point.}$$

- Its converse fails even at a differentiable interior point.
- The condition itself fails if we drop interiority.
- It does not apply without differentiability.

Three counterexamples

1. **Critical does not imply optimal.** The origin is critical for $f(x) = x^3$ on $\mathbb{R}$, yet $f(t) > 0 > f(-t)$ for every $t > 0$.

Three counterexamples

1. **Critical does not imply optimal.** The origin is critical for $f(x) = x^3$ on $\mathbb{R}$, yet $f(t) > 0 > f(-t)$ for every $t > 0$.
2. **Optimal does not imply critical at the boundary.** For $f(x) = x$ on $D = [0, 1]$, the maximizer is the boundary point $x^* = 1$, yet $f'(1) = 1 \neq 0$.

Three counterexamples

1. **Critical does not imply optimal.** The origin is critical for $f(x) = x^3$ on $\mathbb{R}$, yet $f(t) > 0 > f(-t)$ for every $t > 0$.
2. **Optimal does not imply critical at the boundary.** For $f(x) = x$ on $D = [0, 1]$, the maximizer is the boundary point $x^* = 1$, yet $f'(1) = 1 \neq 0$.
3. **Without differentiability the condition does not apply.** On $\mathbb{R}$, $f(x) = -|x|$ has a strict global maximizer at the interior point 0, where no derivative exists.

The Hessian classifies a critical point

A critical point can be a local maximizer, a local minimizer, or neither.

The Hessian classifies it whenever the associated quadratic form is definite.

The tool is the second-order Taylor expansion of Lecture 5: if f is C^2 near x^*, then

$$f(x^* + h) - f(x^*) = \nabla f(x^*) \cdot h + \frac{1}{2} h^\top H_f(x^*) h + r(h), \qquad \frac{|r(h)|}{\|h\|^2} \to 0.$$

Second-order conditions

Theorem 2.5 (Second-order conditions)

Let f be C^2 on an open set containing x^*.

1. *If x^* is a local maximizer, then $H_f(x^*)$ is negative semidefinite.*

Suppose in addition that $\nabla f(x^*) = 0$.

2. *If $H_f(x^*)$ is negative definite, then x^* is a strict local maximizer.*
3. *If $H_f(x^*)$ is indefinite, then x^* is a saddle point (neither a local max nor min).*

For local minimizers, reverse the signs in the first two statements.

The second derivative of a line restriction $g(t) = f(x^* + tv)$ is the quadratic form

$$g''(t) = v^\top H_f(x^* + tv)v,$$

and $g''(0) \leq 0$ at a local max my the one-variable lemma part (2).

A saddle point classified

Example 2.7 (A point classified)

Let $f(x_1, x_2) = x_1^2 - x_2^2$. Classify $x^* = (0, 0)$.

A saddle point classified

Example 2.7 (A point classified)

Let $f(x_1, x_2) = x_1^2 - x_2^2$. Classify $x^* = (0, 0)$.

Solution.

1. Check the first-order condition: $\nabla f(x_1, x_2) = (2x_1, -2x_2)^\top$; the point $(0, 0)$ is critical because $\nabla f(0, 0) = 0$.
2. Compute curvature: $H_f = \left(\begin{smallmatrix} 2 & 0 \\ 0 & -2 \end{smallmatrix}\right)$ has eigenvalues 2 and -2. By Lecture 5's eigenvalue test, it is indefinite.
3. Apply the second-order condition: the origin is a saddle point.

$\square$

The saddle drawn

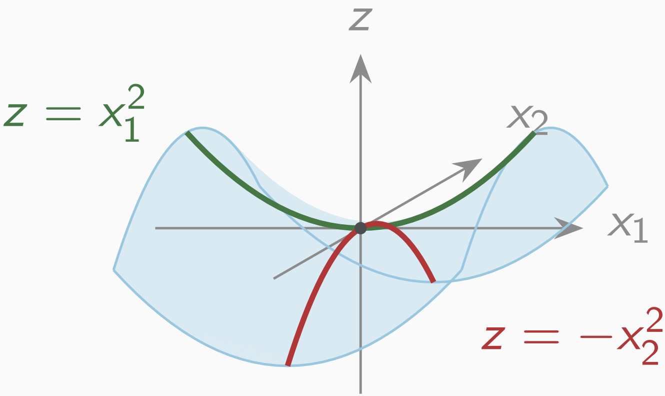


Figure 2.1

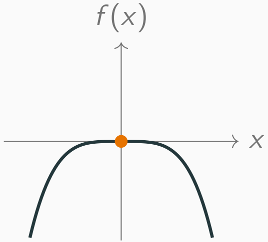
- For $f(x_1, x_2) = x_1^2 - x_2^2$ the point $(0, 0)$ is a saddle point.

A semidefinite Hessian is inconclusive

A semidefinite Hessian does not classify a critical point.

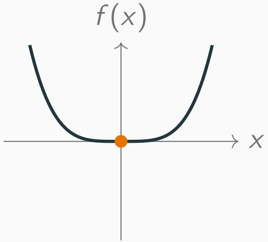
Each function below has $f'(0) = f''(0) = 0$, but the origin has a different classification.

$f(x) = -x^4$



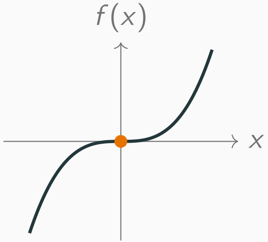
0: strict local maximizer

$f(x) = x^4$



0: strict local minimizer

$f(x) = x^3$



0: neither

Feasible direction

It is a good idea to treat boundary points separately from interior points.

- A boundary point does not need to be critical to be a local maximizer.

Definition 2.8 (Feasible direction)

Let $x \in D \subseteq \mathbb{R}^n$.

A vector $v \in \mathbb{R}^n$ is a *feasible direction* from x if there exists $\delta > 0$ with

$$x + tv \in D \qquad \text{for every } t \in (0, \delta).$$

If D is convex, then $y - x$ is a feasible direction from x for every $y \in D$.

Boundary first-order condition

Theorem 2.9 (Boundary first-order condition)

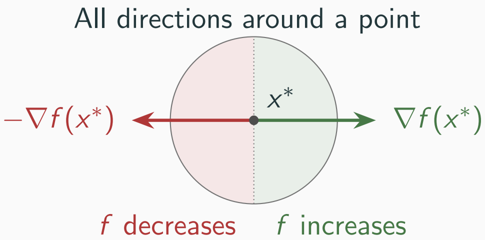
Let $U \subseteq \mathbb{R}^n$ be open, let $D \subseteq U$, and let $f : U \to \mathbb{R}$ be differentiable.

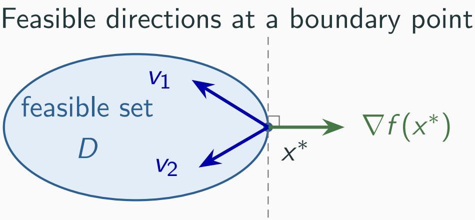
If $x^* \in D$ is a local maximizer of f on D, then

$$\nabla f(x^*) \cdot v \leq 0 \qquad \text{for every feasible direction } v \text{ from } x^*.$$

- $\nabla f(x^*) \cdot v$ is the initial slope along the feasible move $x^* + tv$; at a local maximum, that slope cannot be positive.
- Interior point: v and $-v$ are feasible, so $\nabla f(x^*) = 0$.

Boundary FOC: feasible moves cannot point uphill





$$\nabla f(x^*) \cdot v \le 0 \qquad \text{for every feasible direction } v.$$

- $\nabla f(x^*) \cdot v \le 0$ means the angle between $\nabla f(x^*)$ and v is right or obtuse: every feasible direction lies in the red half-disc or on the dotted line; a green direction would raise f, which is not allowed at a local max.

Componentwise form on the orthant

On $\mathbb{R}^n_+$, the boundary condition splits coordinate by coordinate.

Theorem 2.10 (Componentwise form on the orthant)

Let $U \subseteq \mathbb{R}^n$ be open with $\mathbb{R}^n_+ \subseteq U$, let $f : U \to \mathbb{R}$ be differentiable, and let $x^* \in \mathbb{R}^n_+$.

Then the boundary condition with $D = \mathbb{R}^n_+$ holds if and only if

$$\frac{\partial f}{\partial x_i}(x^*) \le 0, \qquad x_i^* \ge 0, \qquad x_i^* \frac{\partial f}{\partial x_i}(x^*) = 0 \qquad (i = 1, \dots, n).$$

- In $\mathbb{R}^n_+$, the direction e_i is always feasible, so $\partial f / \partial x_i(x^*) \le 0$; if $x_i^* > 0$, then $-e_i$ is feasible too, so $\partial f / \partial x_i(x^*) = 0$.
- Hence the product $x_i^* \, \partial f / \partial x_i(x^*)$ is always zero: a negative partial derivative can occur only at a corner, where $x_i^* = 0$.

Global certification

The tangent inequality

Recall from Lecture 5 that a differentiable **concave** f on a convex set D satisfies

$$f(y) \leq f(x) + \nabla f(x) \cdot (y - x) \qquad \text{for every } x, y \in D.$$

- The tangent plane at x lies above the graph everywhere, not just near x.
- So the gradient at one point bounds f at every other point.
- At a point x^* with $\nabla f(x^*) \cdot (y - x^*) \leq 0$ for every $y \in D$, we get $f(y) \leq f(x^*)$.

Thus, if f is concave and differentiable, a local max is automatically a global one.

Concave maximization

Theorem 3.1 (Concave maximization)

Let $D \subseteq \mathbb{R}^n$ be convex and let $f : D \to \mathbb{R}$ be concave.

1. *Every local maximizer of f on D is a global maximizer.*
2. *The set $X^* = \arg\max_{x \in D} f(x)$ is either empty or convex.*
3. *If f is strictly concave, then X^* contains at most one point.*

- For (1), if x^* is a local maximizer and y beats it ($f(y) > f(x^*)$), then points on the segment near x^* would beat x^* too.
- For (3), suppose $x \neq y$ both attain the maximum M. Their midpoint is feasible, and strict concavity puts it above $\frac{1}{2}f(x) + \frac{1}{2}f(y) = M$, which is impossible.

First-order characterization

Theorem 3.2 (First-order characterization)

Let $D \subseteq \mathbb{R}^n$ be convex.

Let f be differentiable and concave on an open convex set containing D.

A point $x^* \in D$ is a global maximizer of f on D if and only if

$$\nabla f(x^*) \cdot (y - x^*) \leq 0 \qquad \text{for every } y \in D.$$

Theorem 3.2 (First-order characterization)

Let $D \subseteq \mathbb{R}^n$ be convex.

Let f be differentiable and concave on an open convex set containing D.

A point $x^* \in D$ is a global maximizer of f on D if and only if

$$\nabla f(x^*) \cdot (y - x^*) \leq 0 \qquad \text{for every } y \in D.$$

Proof.

Necessity. A global max is a local one, so this is the boundary first-order condition.

Sufficiency. For $y \in D$, the tangent inequality at x^* and then the hypothesis give

$$f(y) \leq f(x^*) + \nabla f(x^*) \cdot (y - x^*) \leq f(x^*).$$

□

Example 3.3 (A quadratic on a half-line)

For $\theta \in \mathbb{R}$, find the solution set $X^*(\theta)$ and value $V(\theta)$ of

$$\max_{x \in \mathbb{R}_+} f(x;\theta), \qquad f(x;\theta) = \theta x - \frac{1}{2}x^2.$$

A quadratic on a half-line

Example 3.3 (A quadratic on a half-line)

For $\theta \in \mathbb{R}$, find the solution set $X^*(\theta)$ and value $V(\theta)$ of

$$\max_{x \in \mathbb{R}_+} f(x;\theta), \qquad f(x;\theta) = \theta x - \frac{1}{2}x^2.$$

Answer.

$$X^*(\theta) = \big\{\max\{\theta, 0\}\big\}, \qquad V(\theta) = \begin{cases} 0, & \theta \leq 0, \\ \frac{1}{2}\theta^2, & \theta > 0. \end{cases}$$

$\square$

A workflow

1. Specify the objective, its domain, the feasible set, and the parameters.
2. Establish existence, by compactness, by coercivity, or by a direct argument.
3. Generate every candidate: interior critical points, boundary points meeting the feasible-direction condition, and points of nondifferentiability.
4. Classify locally with the second-order conditions when the Hessian is definite.
5. Certify global maximality, by an exhaustive value comparison or by concavity (only if the objective is concave on a convex feasible set).
6. Report X^* and V.

Example 3.5 (A non-concave problem in $\mathbb{R}^2$)

Find the solution set X^* and the value V of

$$\max_{x \in \mathbb{R}^2} f(x_1, x_2), \qquad f(x_1, x_2) = -x_1^4 - x_2^4 + 4x_1^2 + 2x_2^2.$$

Step 1: the objective is f, the domain and the feasible set are both $\mathbb{R}^2$, and there is no parameter.

Non-concave, step 2: existence

The feasible set $\mathbb{R}^2$ is nonempty, closed, and unbounded, and f is continuous, so we check coercivity.

Split the objective as $f(x) = g(x_1) + h(x_2)$ with

$$g(t) = -t^4 + 4t^2 \leq 4, \qquad h(t) = -t^4 + 2t^2 \leq 1, \qquad g(t), h(t) \to -\infty \ \text{ as } |t| \to \infty.$$

If $\|x\| \to \infty$, then $|x_1| \to \infty$ or $|x_2| \to \infty$, so one term falls to $-\infty$ while the other stays below its bound. Hence f is coercive for maximization, and coercive existence gives a maximizer.

Non-concave, step 3: candidates

Every point of $\mathbb{R}^2$ is interior and f is differentiable everywhere, so critical points are the only candidates:

$$\nabla f(x) = \begin{pmatrix} -4x_1^3 + 8x_1 \\ -4x_2^3 + 4x_2 \end{pmatrix} = \begin{pmatrix} -4x_1(x_1^2 - 2) \\ -4x_2(x_2^2 - 1) \end{pmatrix} = 0.$$

We thus have $x_1 \in \{0, \pm\sqrt{2}\}$ and $x_2 \in \{0, \pm 1\}$, nine critical points total.

Non-concave, step 4: local classification

The Hessian is diagonal, so its eigenvalues are the diagonal entries:

$$H_f(x) = \begin{pmatrix} 8 - 12x_1^2 & 0 \\ 0 & 4 - 12x_2^2 \end{pmatrix}.$$

point	eigenvalues	definiteness	classification
$(0,0)$	$8,\ 4$	positive definite	strict local min
$(0,\pm 1)$	$8,\ -8$	indefinite	neither
$(\pm\sqrt{2},0)$	$-16,\ 4$	indefinite	neither
$(\pm\sqrt{2},\pm 1)$	$-16,\ -8$	negative definite	strict local max

Non-concave, step 5: globality

f is not concave so we compare values at the nine candidate points to find the global maximizers:

$$f(0,0) = 0, \qquad f(0,\pm 1) = 1, \qquad f(\pm\sqrt{2},0) = 4, \qquad f(\pm\sqrt{2},\pm 1) = 5.$$

Step 2 guarantees that a maximizer exists, and it must be one of these nine points, so the four attaining 5 are global maximizers.

Non-concave, step 6: report

$$X^* = \{(\sqrt{2}, 1),\ (\sqrt{2}, -1),\ (-\sqrt{2}, 1),\ (-\sqrt{2}, -1)\}, \qquad V = 5.$$

Example 3.6 (A concave problem in $\mathbb{R}^2$)

Find the solution set X^* and the value V of

$$\max_{x \in \mathbb{R}^2} f(x_1, x_2), \qquad f(x_1, x_2) = 3x_1 + 3x_2 - x_1^2 - x_1x_2 - x_2^2.$$

Step 1: the objective is f, the domain and the feasible set are both $\mathbb{R}^2$, and there is no parameter.

Concave, step 2: existence

The feasible set $\mathbb{R}^2$ is nonempty, closed, and unbounded, and f is continuous, so we check coercivity.

Completing the square in the quadratic part,

$$x_1^2 + x_1x_2 + x_2^2 = \tfrac{1}{2}\|x\|^2 + \tfrac{1}{2}(x_1 + x_2)^2 \geq \tfrac{1}{2}\|x\|^2,$$

while $3(x_1 + x_2) \leq 3\sqrt{2}\,\|x\|$ by Cauchy–Schwarz. Together,

$$f(x) \leq 3\sqrt{2}\,\|x\| - \tfrac{1}{2}\|x\|^2 \longrightarrow -\infty \qquad \text{as } \|x\| \to \infty.$$

So f is coercive for maximization, and coercive existence gives a maximizer.

Concave, step 3: candidates

Every point of $\mathbb{R}^2$ is interior and f is differentiable everywhere, so there are no boundary candidates and no points of nondifferentiability.

The interior first-order condition is:

$$\nabla f(x) = \begin{pmatrix} 3 - 2x_1 - x_2 \\ 3 - x_1 - 2x_2 \end{pmatrix} = 0 \qquad \iff \qquad \begin{cases} 2x_1 + x_2 = 3, \\ x_1 + 2x_2 = 3. \end{cases}$$

Subtracting the two equations gives $x_1 = x_2$, so the only critical point is $x^* = (1, 1)$.

Concave, step 4: local classification

The Hessian is constant:

$$H_f(x) = \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}.$$

- Its leading principal minors are $-2 < 0$ and $\det H_f = 4 - 1 = 3 > 0$, so H_f is negative definite by Lecture 5's leading-minor test.
- By the second-order condition, $x^* = (1, 1)$ is a strict local maximizer.

Concave, step 5: globality and uniqueness

The Hessian is negative definite at *every* x, not just at x^*, so f is strictly concave on the convex set $\mathbb{R}^2$ by Lecture 5's curvature test.

- The first-order characterization then makes the critical point $x^* = (1, 1)$ a global maximizer.
- Part (3) of concave maximization makes it the only one.

Concave, step 6: report

$$X^* = \{(1,1)\}, \qquad V = f(1,1) = 3 + 3 - 1 - 1 - 1 = 3.$$