Mathematics for Economists

Mitri Kitti

Aalto University

Unconstrained Optimization

Unconstrained optimization

- Let $f : U \to \mathbb{R}$ be a function, with $U \subseteq \mathbb{R}^n$
- The problem of maximizing f is written as

$$\max_{\mathbf{x} \in U} f(\mathbf{x})$$

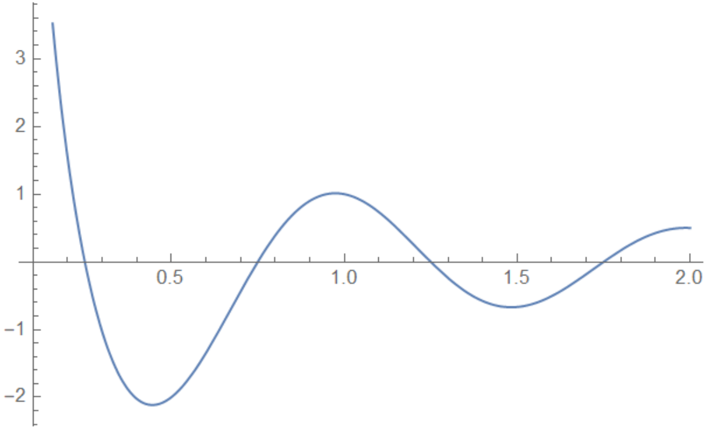
- In words, we want to find a point $\mathbf{x}^*$ in the domain of f such that $f(\mathbf{x}^*)$ is the largest value that f can attain in its domain
- f is called the objective function, $\mathbf{x}$ is the choice variable
- The maximization problem is unconstrained because the choice variable does not have to satisfy any constraint other than being in the domain U of f

Local and Global Maxima

- $\mathbf{x}^* \in U$ is a **global maximizer** of f if $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x} \in U$
- $\mathbf{x}^* \in U$ is a **strict global maximizer** of f if $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $x \in U$ such that $\mathbf{x} \neq \mathbf{x}^*$
- $\mathbf{x}^* \in U$ is a **local maximizer** of f if there is a ball $B_r(\mathbf{x}^*)$ around $\mathbf{x}^*$ such that $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^*) \cap U$
- $\mathbf{x}^* \in U$ is a **strict local maximizer** of f if there is a ball $B_r(\mathbf{x}^*)$ around $\mathbf{x}^*$ such that $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^*) \cap U$ such that $\mathbf{x} \neq \mathbf{x}^*$
- Terminology: $\mathbf{x}^*$ is the *maximizer*, $f(\mathbf{x}^*)$ is the *maximum value* of f

Local and Global Maxima

- A global maximizer is also a local maximizer, but the converse is not necessarily true
- For example, consider the function $f(x) = \frac{\cos(2\pi x)}{x}$ defined over the interval $[0.1, 2]$



Local and Global Minima

- $\mathbf{x}^* \in U$ is a **global minimizer** of f if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in U$
- $\mathbf{x}^* \in U$ is a **strict global minimizer** of f if $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $x \in U$ such that $\mathbf{x} \neq \mathbf{x}^*$
- $\mathbf{x}^* \in U$ is a **local minimizer** of f if there is a ball $B_r(\mathbf{x}^*)$ around $\mathbf{x}^*$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^*) \cap U$
- $\mathbf{x}^* \in U$ is a **strict local minimizer** of f if there is a ball $B_r(\mathbf{x}^*)$ around $\mathbf{x}^*$ such that $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^*) \cap U$ such that $\mathbf{x} \neq \mathbf{x}^*$
- Terminology: $\mathbf{x}^*$ is the *minimizer*, $f(\mathbf{x}^*)$ is the *minimum value* of f
- Note: $\mathbf{x}^*$ is a minimizer of f if and only if $\mathbf{x}^*$ is a maximizer of $-f$

Unconstrained optimization

- When looking for a (local or global) maximizer $\mathbf{x}^*$, three cases are possible:

  1. a maximizer does not exist. E.g., $\max_{(x,y)\in\mathbb{R}^2} \; x+y$
  2. a unique maximizer exists. E.g., $\max_{x\in\mathbb{R}} \; -x^2$
  3. more than one maximizer exist. E.g., $\max_{x\in\mathbb{R}} \; \sin(x)$

Weierstrass Theorem

- The fundamental result about the **existence** of maximizers and minimizers is Weierstrass's Theorem

Theorem (Weierstrass)

Let $f : C \to \mathbb{R}$ *be a* **continuous** *function whose domain is a* **compact** *set* $C \subset \mathbb{R}^n$*. Then there exists a global maximizer* $\mathbf{x}_M \in C$ *of* f*, and there exists also a global minimizer* $\mathbf{x}_m \in C$ *of* f*.*

- A set $C \subseteq \mathbb{R}^n$ is compact if and only if it is **closed** and **bounded** (Heine-Borel Theorem)

- A set $S \subseteq \mathbb{R}^n$ is bounded if there exists a number B such that $||\mathbf{x}|| \leq B$ for all $x \in S$. In other words, S is contained in some ball in $\mathbb{R}^n$

Weierstrass Theorem

- Weierstrass's Theorem:
  - gives *sufficient* conditions for the existence of global extrema (i.e. maximizers or minimizers)
  - does not say that extrema are unique
  - does not tell us how to find extrema
- We will use *first and second order conditions* to find extrema

First Order Necessary Optimality Condition

- The following proposition provides a *necessary* condition for extrema that are both *local* and *interior*

Proposition (First order necessary condition for interior extrema)

Let $f : U \to \mathbb{R}^n$ be a C^1 function, with $U \subseteq \mathbb{R}^n$. If $\mathbf{x}^*$ is a **local** *maximizer or minimizer of f and if $\mathbf{x}^*$ is an* **interior** *point of U, then*

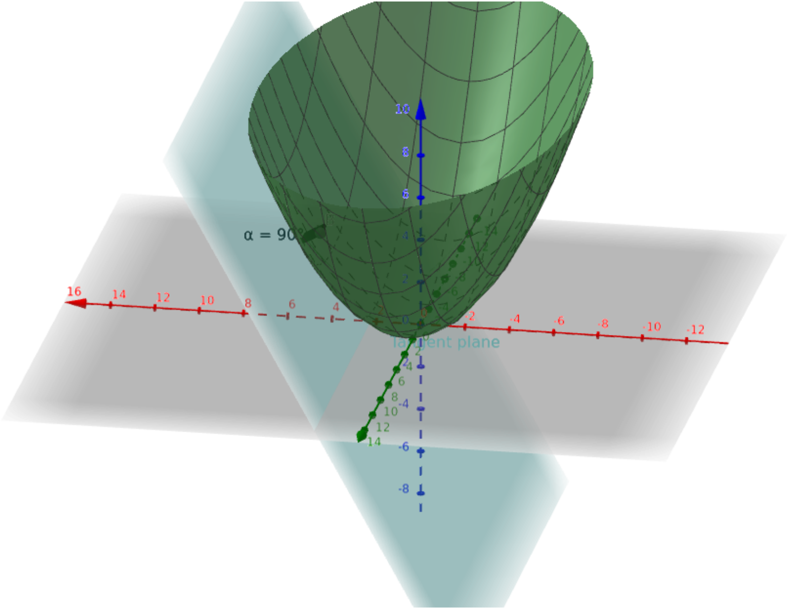
$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0, \qquad \textit{for } i = 1, \ldots, n$$

- A point $\mathbf{x} \in U \subseteq \mathbb{R}^n$ is an interior point of U if there exists an open ball $B_r(\mathbf{x})$ around $\mathbf{x}$ such that $B_r(\mathbf{x}) \subseteq U$

First Order Necessary Optimality Condition

- A point $\mathbf{x} \in U \subseteq \mathbb{R}^n$ is an interior point of U if there exists an open ball $B_r(\mathbf{x})$ around $\mathbf{x}$ such that $B_r(\mathbf{x}) \subseteq U$
- A point at which all the partial derivatives of f are equal to zero is called a *critical point* of f
- If a local maximizer is not interior, then the above proposition does not hold. E.g., $\max_{x \in [0,1]} \; x^2$

Optimality Geometrically



Second Order Sufficient Optimality Condition

- The first order condition says that a local extremum must be a critical point. To understand whether a critical point is a maximizer or a minimizer or neither, we rely on the following second order sufficient condition

Proposition (Second order sufficient condition for interior extrema)

Let $f : U \to \mathbb{R}^n$ be a C^2 function, where $U \subseteq \mathbb{R}^n$ is an **open** *set. Suppose that $\mathbf{x}^*$ is a critical point of f.*

1. *If the Hessian $D^2 f(\mathbf{x}^*)$ is a* **negative definite** *symmetric matrix, then $\mathbf{x}^*$ is a strict local maximizer of f;*
2. *If the Hessian $D^2 f(\mathbf{x}^*)$ is a* **positive definite** *symmetric matrix, then $\mathbf{x}^*$ is a strict local minimizer of f;*
3. *If the Hessian $D^2 f(\mathbf{x}^*)$ is* **indefinite**, *then $\mathbf{x}^*$ is neither a local maximizer nor a local minimizer of f. In this case, we say that $\mathbf{x}^*$ is a* **saddle** *point of f.*

Definite matrices

- Let A be an $n \times n$ symmetric matrix. Then A is:
  - **positive definite** if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$ in $\mathbb{R}^n$
  - **positive semidefinite** if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$
  - **negative definite** if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$
  - **negative semidefinite** if $\mathbf{x}^T A\mathbf{x} \leq 0$ for all $x \neq 0$ in $\mathbb{R}^n$
  - **indefinite** if $\mathbf{x}^T A\mathbf{x} > 0$ for some $x \in \mathbb{R}^n$ and $\mathbf{y}^T A\mathbf{y} < 0$ for some $\mathbf{y} \neq \mathbf{x}$ in $\mathbb{R}^n$.
- Note: In the definitions above, $\mathbf{x}$ is interpreted as a *column* vector in $\mathbb{R}^n$

Definite matrices

- **Example.** Consider the matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$$

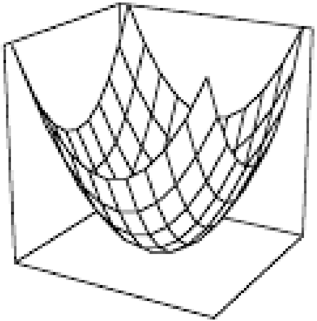
- For any $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, we have

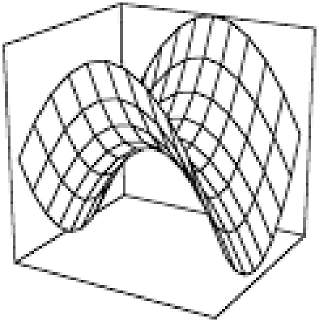
$$\begin{aligned} \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= x^2 + 4y^2 - 2xy \\ &= (x-y)^2 + 3y^2, \end{aligned}$$

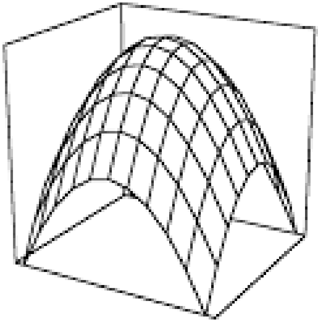
which is equal to zero if and only if $x = y = 0$, and strictly positive otherwise

- Hence A is positive definite

Graphs of quadratic functions







Definite matrices

- It is often convenient to study the definiteness of a matrix by using the *leading principal minors* of the matrix itself
- Let A be an $n \times n$ matrix:
  - The kth order **leading principal submatrix** of A is the submatrix obtained by deleting the *last* $n - k$ rows and the last $n - k$ columns from A;
  - The kth order **leading principal minor** of A is the determinant of the kth order leading principal submatrix of A.

Definite matrices

- **Example.** Let A be the 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

- The three leading principal submatrices of A are

$$A_1 = \begin{pmatrix} a_{11} \end{pmatrix}$$

$$A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Definite matrices

- Let A be an $n \times n$ symmetric matrix. Then:
  - A is **positive definite** if and only if all its n leading principal minors are strictly positive;
  - A is **negative definite** if and only if all its n leading principal minors alternate in sign as follows:

    $$\det(A_1) < 0, \ \det(A_2) > 0, \ \det(A_3) < 0, \ \det(A_4) > 0, \ldots$$

    In other words, the kth order leading principal minor should have the same sign as $(-1)^k$
  - If some kth order leading principal minor is *nonzero* but does not fit either of the above two cases, then A is **indefinite**

Definite matrices

- **Example.** Consider again the matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$$

- The two leading principal minors are

$$\det(A_1) = 1 > 0$$
$$\det(A_2) = 5 > 0$$

- Hence A is positive definite

Example

- **Example.** Let's find all the local maximizers and minimizers of the function $f(x, y) = x^3 - y^3 + 9xy$.
- First of all, we need to find the critical points of f, which turn out to be $(0, 0)$ and $(3, -3)$
- Then we use second order conditions to determine whether each of those two critical points is a maximizer or a minimizer or a saddle point

Example

- **Example (cont'd).** The Hessian of f is

$$D^2 f = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

- At the point $(0,0)$, the Hessian is

$$D^2 f(0,0) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix}$$

- The two leading principal minors are 0 and -81. This implies that $D^2 f(0,0)$ is indeterminate

- Thus we can conclude that $(0,0)$ is a saddle point of f

Example

- **Example (cont'd).** At the point $(3, -3)$, the Hessian is

$$D^2 f(3, -3) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}$$

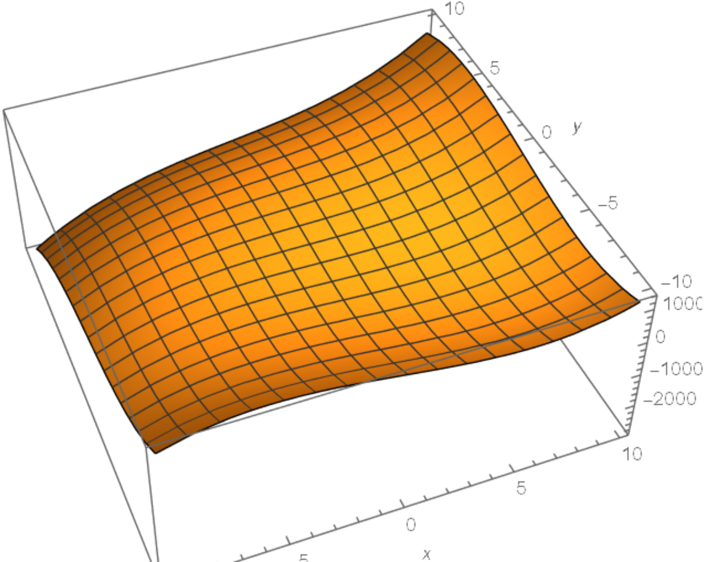
- Now the two leading principal minors are 18 and 243. This implies that $D^2 f(0, 0)$ is positive definite

- Thus we can conclude that $(3, -3)$ is a strict local minimizer of f

- Notice that $(3, -3)$ is not a global minimizer of f. Indeed, $f(3, -3) = -27$, but $\lim_{n \to \infty} f(0, n) = -\infty$

Example

- **Example (cont'd).** The graph of the function $f(x, y) = x^3 - y^3 + 9xy$



Unconstrained optimization

- In addition to the second order sufficient conditions that we already introduced, we can also give second order *necessary* conditions for local extrema

Proposition (Second order necessary condition for interior extrema)

Let $f : U \to \mathbb{R}^n$ be a C^2 function, with $U \subseteq \mathbb{R}^n$.

1. *If x^* is both an interior point of U and a local maximizer of f, then x^* is a critical point of f and the Hessian $D^2 f(x^*)$ is* **negative semidefinite***;*
2. *If x^* is both an interior point of U and a local minimizer of f, then x^* is a critical point of f and the Hessian $D^2 f(x^*)$ is* **positive semidefinite***.*

Semi-definite matrices

- We can study the semi-definiteness of a matrix by using the *principal minors* of the matrix itself (and not just its *leading* principal minors)
- Let A be an $n \times n$ matrix:
  - A kth order **principal submatrix** of A is a $k \times k$ submatrix obtained by deleting $n - k$ rows and the same $n - k$ columns from A;
  - A kth order **principal minor** of A is the determinant of a kth order principal submatrix of A.

Semi-definite matrices

- **Example.** Let A be the 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

- The matrix A itself is the only third order principal submatrix
- The second order principal submatrices of A are

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} \qquad \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

- The first order principal submatrices of A are

$$\begin{pmatrix} a_{11} \end{pmatrix} \qquad \begin{pmatrix} a_{22} \end{pmatrix} \qquad \begin{pmatrix} a_{33} \end{pmatrix}$$

Semi-definite matrices

- Let A be an $n \times n$ symmetric matrix. Then:
  - A is **positive semidefinite** if and only if *every* principal minor of A is non-negative;
  - A is **negative semidefinite** if and only if *every* principal minor of odd order is non-positive and *every* principal minor of even order is non-negative.

Semi-definite matrices

- How to check the definiteness of a matrix?

  1. Find all the leading principal minors. If the conditions for positive (respectively, negative) definiteness are satisfied, you can conclude that the matrix is positive (respectively, negative) definite and positive (respectively, negative) semidefinite too
  2. If the conditions for positive or negative definiteness are not satisfied, check if they are strictly violated. If they are, then the matrix is indefinite
  3. If the conditions for positive or negative definiteness are not strictly violated, find all the principal minors and check if the conditions for positive or negative semidefiniteness are satisfied

Unconstrained optimization

- **Exercise.** Find all the local maximizers and minimizers of the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by
$$f(x, y) = xy^2 + x^3y - xy.$$