# Ch03. Convex Sets and Concave Functions

Ping Yu

Faculty of Business and Economics
The University of Hong Kong

# Overview of This Chapter

- We will show uniqueness of the optimizer and sufficient conditions for optimization through convexity.
- To study convex functions, we need to first define convex sets.

# Convex Sets

# Convex Combination, Interval and Convex Set

- Given two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, a point $\mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y}$, where $0 \leq t \leq 1$, is called a convex combination of $\mathbf{x}$ and $\mathbf{y}$.
- The set of all possible convex combinations of $\mathbf{x}$ and $\mathbf{y}$, denoted by $[\mathbf{x}, \mathbf{y}]$, is called the interval with endpoints $\mathbf{x}$ and $\mathbf{y}$ (or, the line segment connecting $\mathbf{x}$ and $\mathbf{y}$), i.e.,

$$[\mathbf{x}, \mathbf{y}] = \{t\mathbf{x} + (1-t)\mathbf{y} \mid 0 \leq t \leq 1\}.$$

- This definition is an extension of the interval in $\mathbb{R}^1$.

### Definition

A set $S \subseteq \mathbb{R}^n$ is convex iff for any points $\mathbf{x}$ and $\mathbf{y}$ in $S$ the interval $[\mathbf{x}, \mathbf{y}] \subseteq S$. [Figure here]

- A set is convex if it contains the line segment connecting any two of its points; or
- A set is convex if for any two points in the set it also contains all points between them.

# Examples of Convex and Non-Convex Sets

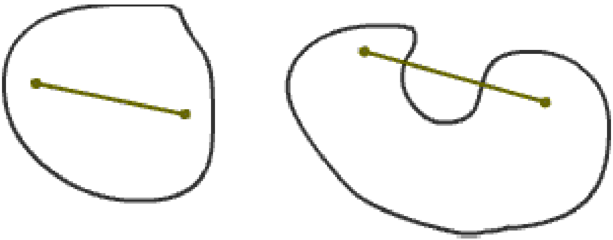

Figure: Convex and Non-Convex Set

- Convex sets in $\mathbb{R}^2$ include triangles, squares, circles, ellipses, and hosts of other sets.
- The quintessential convex set in Euclidean space $\mathbb{R}^n$ for any $n > 1$ is the $n$-dimensional open ball $B_r(\mathbf{a})$ of radius $r > 0$ about point $\mathbf{a} \in \mathbb{R}^n$, where recall from Chapter 1 that

$$B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\| < r\}.$$

- In $\mathbb{R}^3$, while a cube is a convex set, its boundary is not. (Of course, the same is true of the square in $\mathbb{R}^2$.)

## Example

Prove that the budget constraint $B = \{\mathbf{x} \in X : \mathbf{p}'\mathbf{x} \leq y\}$ is convex.

## Proof.

For any two points $\mathbf{x}_1, \mathbf{x}_2 \in B$, we have

$$\mathbf{p}'\mathbf{x}_1 \leq y \text{ and } \mathbf{p}'\mathbf{x}_2 \leq y.$$

Then for any $t \in [0,1]$, we must have

$$\mathbf{p}'\left[t\mathbf{x}_1 + (1-t)\mathbf{x}_2\right] = t\left(\mathbf{p}'\mathbf{x}_1\right) + (1-t)\left(\mathbf{p}'\mathbf{x}_2\right) \leq y.$$

This is equivalent to say that $t\mathbf{x}_1 + (1-t)\mathbf{x}_2 \in B$. So the budget constraint $B$ is convex. □

# Concave Functions

# Concave and Convex Functions

- For uniqueness, we need to know something about the shape or **curvature** of the functions $f$ and $(\mathbf{g},\mathbf{h})$.
- A function $f: S \to \mathbb{R}$ defined on a convex set $S$ is concave if for any $\mathbf{x},\mathbf{x}' \in S$ with $\mathbf{x} \neq \mathbf{x}'$ and for any $t$ such that $0 < t < 1$ we have $f(t\mathbf{x} + (1-t)\mathbf{x}') \geq tf(\mathbf{x}) + (1-t)f(\mathbf{x}')$. The function is strictly concave if $f(t\mathbf{x} + (1-t)\mathbf{x}') > tf(\mathbf{x}) + (1-t)f(\mathbf{x}')$. [Figure here]
- A function $f: S \to \mathbb{R}$ defined on a convex set $S$ is convex if for any $\mathbf{x},\mathbf{x}' \in S$ with $\mathbf{x} \neq \mathbf{x}'$ and for any $t$ such that $0 < t < 1$ we have $f(t\mathbf{x} + (1-t)\mathbf{x}') \leq tf(\mathbf{x}) + (1-t)f(\mathbf{x}')$. The function is strictly convex if $f(t\mathbf{x} + (1-t)\mathbf{x}') < tf(\mathbf{x}) + (1-t)f(\mathbf{x}')$. [Figure here]
- Why don't we check $t = 0$ and 1 in the definition? Why the domain of $f$ must be a convex set? (Exercise)
- The negative of a (strictly) convex function is (strictly) concave. (why?)
- There are both concave and convex functions, but only convex sets, no concave sets!

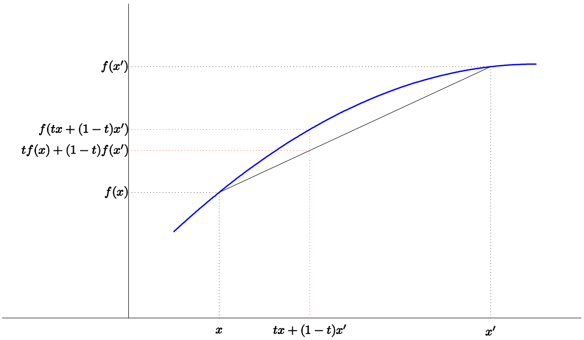


Figure: Concave Function

- A function is concave if the value of the function at the average of two points is greater than the average of the values of the function at the two points.

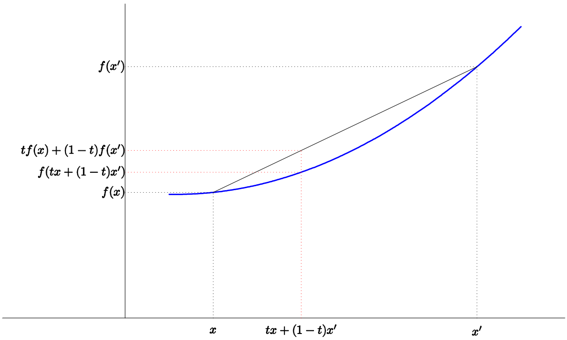


Figure: Convex Function

- A function is convex if the value of the function at the average is less than the average of the values.

# Calculus Criteria for Concavity and Convexity

### Theorem

*Let $f \in C^2(U)$, where $U \subset \mathbb{R}^n$ is open and convex. Then $f$ is concave iff the* ***Hessian***

$$D^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}$$

*is negative semidefinite for all $\mathbf{x} \in U$. If $D^2 f(\mathbf{x})$ is negative definite for all $\mathbf{x} \in U$, then $f$ is* ***strictly*** *concave on $U$. Conditions for convexity are obtained by replacing "negative" by "positive".*

- The conditions for strict concavity in the theorem are only sufficient, not necessary.
  - if $D^2 f(\mathbf{x})$ is not negative semidefinite for all $\mathbf{x} \in U$, then $f$ is not concave;
  - if $D^2 f(\mathbf{x})$ is not negative definite for all $\mathbf{x} \in U$, then $f$ may or may not be strictly concave (see the example below).
- Notations: For a matrix $\mathbf{A}$, $\mathbf{A} > 0$ means it is positive definite, $\mathbf{A} \geq 0$ means it is positive semidefinite. Similarly for $\mathbf{A} < 0$ and $\mathbf{A} \leq 0$.

# Positive (Negative) Definiteness of A Matrix

- An $n \times n$ matrix $\mathbf{H}$ is positive definite iff $\mathbf{v}'\mathbf{H}\mathbf{v} > 0$ for all $\mathbf{v} \neq \mathbf{0}$ in $\mathbb{R}^n$; $\mathbf{H}$ is negative definite iff $\mathbf{v}'\mathbf{H}\mathbf{v} < 0$ for all $\mathbf{v} \neq \mathbf{0}$ in $\mathbb{R}^n$.
- Replacing the strict inequalities above by weak ones yields the definitions of positive semidefinite and negative semidefinite.
  - Usually, positive (negative) definiteness is only defined for a symmetric matrix, so we restrict our discussions on **symmetric** matrices below. Fortunately, the Hessian is symmetric by Young's theorem.
- The positive definite matrix is an extension of the positive number. To see why, note that for any positive number $H$, and any real number $v \neq 0$, $v'Hv = v^2H > 0$. Similarly, the positive semidefinite matrix, negative definite matrix, negative semidefinite matrix are extensions of the nonnegative number, negative number and nonpositive number, respectively.

# Identifying Definiteness and Semidefiniteness

- For an $n \times n$ matrix **H**, a $k \times k$ submatrix formed by picking out $k$ columns and the **same** $k$ rows is called a $k$th order principal submatrix of **H**; the determinant of a $k$th order principal submatrix is called a $k$th order principal minor.
- The $k \times k$ submatrix formed by picking out the **first** $k$ columns and the **first** $k$ rows is called a $k$th order leading principal submatrix of **H**; its determinant is called the $k$th order leading principal minor.

- A matrix is positive definite iff its $n$ **leading principal minors** are all $> 0$.
- A matrix is negative definite iff its $n$ leading principal minors alternate in sign with the odd order ones being $< 0$ and the even order ones being $> 0$.
- A matrix is positive semidefinite iff its $2^n - 1$ **principal minors** are all $\geq 0$.
- A matrix is negative semidefinite iff its $2^n - 1$ principal minors alternate in sign so that the odd order ones are $\leq 0$ and the even order ones are $\geq 0$.

# Examples

- $f(x) = -x^4$ is strictly concave, but its Hessian is not negative definite for all $x \in \mathbb{R}$ since $D^2 f(0) = 0$.
- The particular Cobb-Douglas utility function $u(x_1, x_2) = \sqrt{x_1}\sqrt{x_2}$, $(x_1, x_2) \in \mathbb{R}^2_+$, is concave but not strictly concave. First check that it is concave.

$$D^2 f(\mathbf{x}) = \begin{pmatrix} \frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{x_2}}{\sqrt{x_1^3}} & \frac{1}{2}\frac{1}{2}\frac{1}{\sqrt{x_1}\sqrt{x_2}} \\ \frac{1}{2}\frac{1}{2}\frac{1}{\sqrt{x_1}\sqrt{x_2}} & \frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{x_1}}{\sqrt{x_2^3}} \end{pmatrix}.$$

Since

$$\frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{x_2}}{\sqrt{x_1^3}} \le 0, \frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{x_1}}{\sqrt{x_2^3}} \le 0$$

and

$$\left(\frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{x_2}}{\sqrt{x_1^3}}\right)\left(\frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{x_1}}{\sqrt{x_2^3}}\right) - \left(\frac{1}{2}\frac{1}{2}\frac{1}{\sqrt{x_1}\sqrt{x_2}}\right)^2 = 0$$

for $(x_1, x_2) \in \mathbb{R}^2_+$, $u(x_1, x_2)$ is concave.

- Let $x_2 = x_2' = 0$, $x_1 \neq x_1'$; then $u(tx_1 + (1-t)x_1', 0) = 0 = tu(x_1, 0) + (1-t)u(x_1', 0)$, so $u(x_1, x_2)$ is not strictly concave.

# Local Maximum is Global Maximum

- Consider the mixed constrained maximization problem, i.e.,

$$\max_{\mathbf{x}} f(\mathbf{x}) \text{ s.t. } \mathbf{x} \in G \equiv \left\{ \mathbf{x} \in \mathbb{R}^n | \mathbf{g}(\mathbf{x}) \geq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \right\}.$$

**Theorem**

*If $f$ is **concave**, and the feasible set $G$ is convex, then*

(i) *Any local maximum of $f$ is a global maximum of $f$.*

(ii) *The set $\arg\max \{f(\mathbf{x}) | \mathbf{x} \in G\}$ is convex.*

- In concave optimization problems, all local optima must also be global optima; therefore, to find a global optimum, it always suffices to locate a local optimum.

# The Uniqueness Theorem

**Theorem**

*If $f$ is **strictly concave**, and the feasible set $G$ is convex, then the maximizer $\mathbf{x}^*$ is unique.*

**Proof.**

Suppose $f$ has two maximizers, say, $\mathbf{x}$ and $\mathbf{x}'$; then $t\mathbf{x} + (1-t)\mathbf{x}' \in G$, and by the definition of strict concavity, for $0 < t < 1$,

$$f(t\mathbf{x} + (1-t)\mathbf{x}') > tf(\mathbf{x}) + (1-t)f(\mathbf{x}') = f(\mathbf{x}) = f(\mathbf{x}').$$

A contradiction. □

- If a strictly concave optimization problem admits a solution, the solution must be unique. So finding one solution is enough.

# Example: Consumer's Problem - Revisited

- Does the consumer's problem

$$\max_{x_1,x_2} \sqrt{x_1}\sqrt{x_2} \text{ s.t. } x_1+x_2 \leq 1, x_1 \geq 0, x_2 \geq 0$$

have a solution? Is the solution unique?

- The feasible set $G = \{x_1+x_2 \leq 1, x_1 \geq 0, x_2 \geq 0\}$ is compact (why?) and $\sqrt{x_1}\sqrt{x_2}$ is continuous, so by the Weierstrass Theorem, there exists a solution.
- The solution is unique, $(x_1^*, x_2^*) = \left(\frac{1}{2}, \frac{1}{2}\right)$. But from the discussion above, $\sqrt{x_1}\sqrt{x_2}$ is not strictly concave for $(x_1, x_2) \in \mathbb{R}_+^2$. Actually, even if we restrict $(x_1, x_2) \in \mathbb{R}_{++}^2$, where $\mathbb{R}_{++} \equiv \{x | x > 0\}$, $\sqrt{x_1}\sqrt{x_2}$ is NOT strictly concave. Check for $t \in (0,1), x_1 \neq x_1'$ and/or $x_2 \neq x_2'$,

$$\begin{aligned} & \sqrt{tx_1+(1-t)x_1'}\sqrt{tx_2+(1-t)x_2'} \geq t\sqrt{x_1x_2}+(1-t)\sqrt{x_1'x_2'} \\ \Longleftrightarrow\ & \left(tx_1+(1-t)x_1'\right)\left(tx_2+(1-t)x_2'\right) \geq \left(t\sqrt{x_1x_2}+(1-t)\sqrt{x_1'x_2'}\right)^2 \\ \Longleftrightarrow\ & x_1x_2'+x_1'x_2 \geq 2\sqrt{x_1x_2x_1'x_2'} \Longleftrightarrow \left(\sqrt{x_1x_2'}-\sqrt{x_1'x_2}\right)^2 \geq 0 \end{aligned}$$

with equality holding when $x_2/x_1 = x_2'/x_1'$ (what does this mean?).

- In summary, the theorem provides only sufficient (but not necessary) conditions.

# Sufficient Conditions for Convexity of $G$

- Problem: how to guarantee that $G$ is convex?
- Given a concave function $g$, for any $a \in \mathbb{R}$, its upper contour set $\{\mathbf{x}|g(\mathbf{x}) \geq a\}$ is convex.
- Why? Given two poitns $\mathbf{x}$ and $\mathbf{x}'$ such that $g(\mathbf{x}) \geq a$ and $g(\mathbf{x}') \geq a$, we want to show that for any $t \in [0,1]$, $g(t\mathbf{x} + (1-t)\mathbf{x}') \geq a$. Since $g$ is concave, $g(t\mathbf{x} + (1-t)\mathbf{x}') \geq tg(\mathbf{x}) + (1-t)g(\mathbf{x}') \geq ta + (1-t)a = a$.
- Given a function $h$, to guarantee that $\{\mathbf{x}|h(\mathbf{x}) = a\}$ is convex, we require $h$ to be both concave and convex.
  - A function $h$ is both concave and convex iff it is linear (or, more properly, affine), taking the form $h(\mathbf{x}) = a + \mathbf{b}'\mathbf{x}$ for some constants $a$ and $\mathbf{b}$.
- In summary, since

$$G = \left(\bigcap_{j=1}^{J} \{\mathbf{x}|g_j(\mathbf{x}) \geq 0\}\right) \bigcap \left(\bigcap_{k=1}^{K} \{\mathbf{x}|h_k(\mathbf{x}) = 0\}\right),$$

if $g_j$, $j = 1, \ldots, J$, is concave, and $h_k$, $k = 1, \cdots, K$, is affine, then $G$ is convex.[1]

[1] It is not hard to show that intersection of arbitrarily many convex sets is convex.

### Theorem (Theorem of Kuhn-Tucker under Concavity)

*Suppose $f$, $g_j$ and $h_k$, $j=1,\cdots,J$, $k=1,\cdots,K$, are all $C^1$ function, $f$ is concave, $g_j$ is concave, and $h_k$ is affine. If there exists $(\lambda^*,\mu^*)$ such that $(\mathbf{x}^*,\lambda^*,\mu^*)$ satisfies the Kuhn-Tucker conditions, then $\mathbf{x}^*$ solves the mixed constrained maximization problem.*

- We do not need the NDCQ for this sufficient condition of optimization; the NDCQ is only required for necessary conditions.

### Example

In the consumer's problem above, $g_1(\mathbf{x})=x_1, g_2(\mathbf{x})=x_2$ and $g_3(\mathbf{x})=1-x_1-x_2$ are all affine, so $G$ is convex. Since $u(x_1,x_2)=\sqrt{x_1}\sqrt{x_2}$ is concave, the solution to the Kuhn-Tucker conditions is the global maximizer.

# Second Order Conditions for Optimization

# Second Order Conditions for Optimization

- In the LN, we use the "bordered Hessians" to check a solution to the FOCs is a local maximizer or a local minimizer.
- In practice, this may be quite burdensome.
- As an easy (although less general) alternative, we can employ the concavity of the objective function $f$ to draw the conclusion.
  - if $f$ is strictly concave at $\mathbf{x}^*$ (or more restrictively, if $D^2 f(\mathbf{x}^*) < 0$), then $\mathbf{x}^*$ is a strict local maximizer.
  - if $f$ is strictly convex at $\mathbf{x}^*$ (or more restrictively, if $D^2 f(\mathbf{x}^*) > 0$), then $\mathbf{x}^*$ is a strict local minimizer.