Convex Sets and Concave Functions

Ping Yu

Department of Economics
University of Hong Kong
Convex Sets

Concave Functions
- Basics
- The Uniqueness Theorem
- Quasiconvex Functions
- Sufficient Conditions for Optimization

Second Order Conditions for Optimization
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 \( \mathbf{z} = a\mathbf{x} + (1 - a)\mathbf{y} \), where \( 0 \leq a \leq 1 \), 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.,

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

- 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

- Convex sets in $\mathbb{R}^2$ include interiors of 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 $B_r(a)$ of radius $r > 0$ about point $a \in \mathbb{R}^n$, where recall from Chapter 1 that
  \[ B_r(a) = \{ x \in \mathbb{R}^n \mid \| x - a \| < r \}. \]
- In $\mathbb{R}^3$, while the interior of 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 \( B = \{ x \in X : p'x \leq y \} \) is convex.

Proof.

For any two points \( x_1, x_2 \in B \), we have

\[
p'x_1 \leq y \quad \text{and} \quad p'x_2 \leq y.
\]

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

\[
p'[tx_1 + (1 - t)x_2] = t(p'x_1) + (1 - t)(p'x_2) \leq y.
\]

This is equivalent to say that \( tx_1 + (1 - t)x_2 \in B \). 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 $(g, h)$.

- A function $f : S \rightarrow \mathbb{R}$ defined on a convex set $S$ is **concave** if for any $x, x' \in S$ with $x \neq x'$ and for any $t$ such that $0 < t < 1$ we have
  
  \[ f(tx + (1-t)x') \geq tf(x) + (1-t)f(x') \]

  The function is **strictly concave** if
  
  \[ f(tx + (1-t)x') > tf(x) + (1-t)f(x') \]

- A function $f : S \rightarrow \mathbb{R}$ defined on a convex set $S$ is **convex** if for any $x, x' \in S$ with $x \neq x'$ and for any $t$ such that $0 < t < 1$ we have
  
  \[ f(tx + (1-t)x') \leq tf(x) + (1-t)f(x') \]

  The function is **strictly convex** if
  
  \[ f(tx + (1-t)x') < tf(x) + (1-t)f(x') \]

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

**Figure: Concave Function**
A function is convex if the value of the function at the average is less than the average of the values.
Alternative Definition of Concave and Convex Functions

- Define the subgraph and epigraph of $f$, denoted as $\text{sub}(f)$ and $\text{epi}(f)$:
  
  $$\text{sub}(f) = \{(x, y) \in S \times \mathbb{R} | f(x) \geq y\} \subset \mathbb{R}^{n+1},$$
  
  $$\text{epi}(f) = \{(x, y) \in S \times \mathbb{R} | f(x) \leq y\} \subset \mathbb{R}^{n+1}.$$

- A function $f$ is concave iff $\text{sub}(f)$ is a convex set, and is convex iff $\text{epi}(f)$ is convex.

Figure: Epigraph and Subgraph
Concave Functions

Basics

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

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

is negative semidefinite for all $x \in U$. If $D^2f(x)$ is negative definite for all $x \in U$, 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 $D^2f(x)$ is not negative semidefinite for all $x \in U$, then $f$ is not concave;
  - if $D^2f(x)$ is not negative definite for all $x \in U$, then $f$ may or may not be strictly concave (see the example below).

- Notations: For a matrix $A$, $A > 0$ means it is positive definite, $A \geq 0$ means it is positive semidefinite. Similarly for $A < 0$ and $A \leq 0$. 

Ping Yu (HKU)
Positive (Negative) Definiteness of A Matrix

- An \( n \times n \) matrix \( H \) is **positive definite** iff \( v'Hv > 0 \) for all \( v \neq 0 \) in \( \mathbb{R}^n \); \( H \) is **negative definite** iff \( v'Hv < 0 \) for all \( v \neq 0 \) 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 \( v \neq 0 \), \( v'Hv = v^2 H > 0 \). Similarly, the positive semidefinite matrix, negative definite matrix, negative semidefinite matrix are extensions of the nonnegative number, negative number and nonpositive number, respectively.

- The diagonal elements of a positive definite matrix must be positive, while the off-diagonal elements need not be. [Exercise]
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

- A linear function $f(x) = a_1x_1 + \cdots + a_nx_n$ is both concave and convex.
- $f(x) = -x^4$ is strictly concave, but its Hessian is not negative definite for all $x \in \mathbb{R}$ since $D^2f(0) = 0$.
- 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.

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

Since
$$\frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_2}}{\sqrt{x_1}} \leq 0, \quad \frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_1}}{\sqrt{x_2}} \leq 0$$

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

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.,

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

**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 $\arg \max \{ f(x) | x \in G \}$ 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 $x^*$ is unique.

**Proof.**

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

$$f(tx + (1 - t)x') > tf(x) + (1 - t)f(x') = f(x) = f(x').$$

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 \( \mathbb{R}_+^2 \equiv \{x|x > 0\}, \sqrt{x_1} \sqrt{x_2} \) is NOT strictly concave. Check for \( t \in (0, 1), x_1 \neq x'_1 \) and/or \( x_2 \neq x'_2 \),
  \[
  \sqrt{tx_1 + (1 - t)x'_1} \sqrt{tx_2 + (1 - t)x'_2} \geq t \sqrt{x_1 x_2} + (1 - t) \sqrt{x'_1 x'_2}
  \]
  \(\iff\) \( (tx_1 + (1 - t)x'_1) (tx_2 + (1 - t)x'_2) \geq \left(t \sqrt{x_1 x_2} + (1 - t) \sqrt{x'_1 x'_2}\right)^2 \)
  \(\iff\) \( x_1 x'_2 + x'_1 x_2 \geq 2 \sqrt{x_1 x_2 x'_1 x'_2} \iff \left(\sqrt{x_1 x'_2} - \sqrt{x'_1 x_2}\right)^2 \geq 0 \)

  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.
Quasiconvex and Quasiconcave Functions

- **Problem**: how to guarantee that $G$ is convex?

- A function $f : S \rightarrow \mathbb{R}$ defined on a convex set $S$ is **quasiconvex** if for any $a \in \mathbb{R}$, the lower contour set $\{x|f(x) \leq a\}$ is convex. The negative of a quasiconvex function is said to be **quasiconcave** or its upper contour sets are convex.

- An alternative definition of quasiconvexity/quasiconcavity: A function $f : S \rightarrow \mathbb{R}$ defined on a convex set $S$ is **quasiconvex** if for all $x, y \in S$ and $t \in (0, 1)$ we have

  $$f(tx + (1-t)y) \leq \max \{f(x), f(y)\};$$

  a function $f$ is **quasiconcave** if

  $$f(tx + (1-t)y) \geq \min \{f(x), f(y)\}.$$

- Replacing the equality by strict equality, we get the strictly quasiconvex/quasiconcave function.
A convex function must be a quasiconvex function (why?), but the inverse is not correct.
A function is both concave and convex iff it is linear (or, more properly, affine), taking the form \( f(x) = a + b'x \) for some constants \( a \) and \( b \).

A function that is both quasiconvex and quasiconcave is called quasilinear.
- Any monotone function on \( \mathbb{R} \) is quasilinear, and any affine function is quasilinear. For a quasilinear function \( f \), \( \{x | f(x) = c\} \) for any constant \( c \) is convex.

There are many alternative conditions to ensure \( G \) to be convex. One popular set of conditions is that \( g_j, j = 1, \ldots, J \), is quasiconcave, and \( h_k, k = 1, \ldots, K \), is quasilinear.

For example, in the consumer’s problem above, \( g_1(x) = x_1 \), \( g_2(x) = x_2 \) and \( g_3(x) = 1 - x_1 - x_2 \) are all affine, so \( G \) is convex.
Theorem (Theorem of Kuhn-Tucker under Convexity)

Suppose \( f, g_j \) and \( h_k, j = 1, \ldots, J, k = 1, \ldots, K \), are all \( C^1 \) functions, \( f \) is concave, \( g_j \) is quasiconcave, and \( h_k \) is quasilinear. If there exists \((\lambda^*, \mu^*)\) such that \((x^*, \lambda^*, \mu^*)\) satisfies the Kuhn-Tucker conditions, then \( x^* \) solves the mixed constrained maximization problem.

- In practice, check whether \( g_j \) is concave, and \( h_k \) is affine.
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 $x^*$, i.e., if $D^2 f(x^*) < 0$, then $x^*$ is a strict local maximizer.
- if $f$ is strictly convex at $x^*$, i.e., if $D^2 f(x^*) > 0$, then $x^*$ is a strict local minimizer.