Economics 7250
Advanced Mathematical Techniques for Economics
Second Semester 2014

Lecture 10
Convexity and Concavity

I. Introduction

A. We now consider convexity, concavity, and the general consequences of these concepts in economic modelling.

B. We then describe some specific economic applications.

II. Convex Sets

A. A set $S \subset \mathbb{R}^n$ is convex if it contains the line segment between any pair of points in $S$, which means that for any $x_0, x_1 \in S$,

$$\{ (1 - \alpha)x_0 + \alpha x_1 : 0 \leq \alpha \leq 1 \} \subset S.$$

B. Examples of Convex Sets.

1. Both open and closed half spaces are convex. Consider the halfspace $H = \{ x \in \mathbb{R}^n : p \cdot x \geq 0 \}$ defined by some nonzero $p \in \mathbb{R}^n$. If $x_0, x_1 \in H$ and $0 \leq \alpha \leq 1$, then

$$p \cdot ((1 - \alpha)x_0 + \alpha x_1) = (1 - \alpha)p \cdot x_0 + \alpha p \cdot x_1 \geq (1 - \alpha)0 + \alpha 0 = 0,$$

so that $(1 - \alpha)x_0 + \alpha x_1 \in H$.

2. Both the open and the closed unit balls are convex. If $\|x_0\|, \|x_1\| < 1$
and $0 \leq \alpha \leq 1$, then (applying the Cauchy-Schwartz inequality!)

$$\|(1 - \alpha)x_0 + \alpha x_1\|^2 = ((1 - \alpha)x_0 + \alpha x_1) \cdot ((1 - \alpha)x_0 + \alpha x_1)$$

$$= (1 - \alpha)^2\|x_0\|^2 + 2\alpha(1 - \alpha)x_0 \cdot x_1 + \alpha^2\|x_1\|^2$$

$$\leq (1 - \alpha)^2\|x_0\|^2 + 2\alpha(1 - \alpha)\|x_0\| \cdot \|x_1\| + \alpha^2\|x_1\|^2$$

$$= \left((1 - \alpha)\|x_0\| + \alpha\|x_1\|\right)^2$$

$$< \left((1 - \alpha) + \alpha\right)^2 = 1,$$

so that $\|(1 - \alpha)x_0 + \alpha x_1\| < 1.$

B. Properties of Convex Sets.

1. If $S$ and $T$ are convex subsets of $\mathbb{R}^n$, then $S + T = \{ x + y : x \in S, y \in T \}$ is convex. To show this we observe that for $x_0 + y_0, x_1 + y_1 \in S + T$ and $0 \leq \alpha \leq 1$,

$$(1 - \alpha)(x_0 + y_0) + \alpha(x_1 + y_1) = ((1 - \alpha)x_0 + \alpha x_1) + ((1 - \alpha)y_0 + \alpha y_1).$$

2. If $\{S_\beta\}_{\beta \in B}$ is a collection of convex sets, then the intersection $\bigcap_{\beta \in B} S_\beta$ is convex. (It is easy to draw a picture how the union of two convex sets need not be convex.)

   a. For any set $S \subset \mathbb{R}^n$ the convex hull of $S$ is the intersection of all convex sets that contain $S$.

3. We now see that any intersection of halfspaces is convex, so we are able to construct a lot of convex sets. Visually, it seems reasonable to guess that all convex sets are intersections of half spaces, and a little thought shows that this is the same as the assertion of the following, which is a twentieth century result, due to Minkowski. The generalization to infinite dimensional spaces, which is known as the Hahn-Banach theorem, is one of the most important theorems of functional analysis.

**Theorem:** If $A \subset \mathbb{R}^n$ is closed and convex, and $x \notin A$, there is $p \in \mathbb{R}^n, p \neq 0$, such that $p \cdot x < p \cdot y$ for all $y \in A$. 

2
Proof: If $A$ is empty, any nonzero $p$ will do, so we may suppose that there is a point $y_0 \in A$. Then the set $A_0 := \{ y \in A : \| y - x \| \leq \| y_0 - x \| \}$ is closed and bounded, hence compact, and since distance $d(x, y) = \| y - x \|$ is a continuous function, there must be a point $y^* \in A_0$ with $\| y^* - x \| \leq \| y - x \|$ for all $y \in A_0$ and thus for all $y \in A$. Let $p = y^* - x$. Picking $y \in A$ arbitrarily, it suffices to show that $p \cdot y \geq p \cdot y^*$, since then

$$ p \cdot y \geq p \cdot x + p \cdot (y^* - x) = p \cdot x + \| y^* - x \|^2 > p \cdot x. $$

We this goal in mind, we observe that $(1 - \lambda)y + \lambda y^* \in A$ for any $0 \leq \lambda \leq 1$, so that

$$ 0 < \| (1 - \lambda)y + \lambda y^* - x \|^2 - \| y^* - x \|^2 $$

$$ = \| (1 - \lambda)(y - y^*) + (y^* - x) \|^2 - \| y^* - x \|^2 $$

$$ = \| (1 - \lambda)(y - y^*) \|^2 + 2(1 - \lambda)(y - y^*)(y^* - x) $$

$$ = (1 - \lambda)^2 \| y - y^* \|^2 + 2(1 - \lambda)p \cdot (y - y^*). $$

Since this must hold even when $1 - \lambda$ is very small, it must be the case that $p \cdot (y - y^*) \geq 0$, as desired. 

a. In economic applications, the normal $p$ of the separating hyperplane can often be interpreted as a vector of prices.

II. Convex and Concave Functions

A. Fix a function $f : U \to \mathbb{R}$ where $U \subset \mathbb{R}^n$ is both open and convex.

1. The \textit{subgraph} of $f$ and the \textit{epigraph} of $f$ are, respectively

$$ \text{sub} f = \{ (x, y) : x \in U, y \leq f(x) \} $$

and

$$ \text{epi} f = \{ (x, y) : x \in U, y \geq f(x) \}. $$

2. We say that $f$ is \textit{convex} if the epigraph of $f$ is convex.
a. Equivalently, for any $x_0, x_1$ and any $\alpha$ between 0 and 1,

$$f((1 - \alpha)x_0 + \alpha x_1) \leq (1 - \alpha)f(x_0) + \alpha f(x_1).$$

b. We say that $f$ is strictly convex if, for any distinct $x_0, x_1 \in U$ and any $\alpha$ with $0 < \alpha < 1$,

$$f((1 - \alpha)x_0 + \alpha x_1) < (1 - \alpha)f(x_0) + \alpha f(x_1).$$

3. We say that $f$ is concave if the subgraph of $f$ is convex.

   a. Equivalently, for any $x_0, x_1 \in U$ and any $\alpha$ between 0 and 1,

$$f((1 - \alpha)x_0 + \alpha x_1) \geq (1 - \alpha)f(x_0) + \alpha f(x_1).$$

   b. We say that $f$ is strictly concave if, for any distinct $x_0, x_1 \in U$ and any $\alpha$ with $0 < \alpha < 1$,

$$f((1 - \alpha)x_0 + \alpha x_1) < (1 - \alpha)f(x_0) + \alpha f(x_1).$$

4. Example: A linear function $f(x) = mx + b$ is both concave and convex, but neither strictly concave nor strictly convex.

B. Characterizations of Concavity

1. In the following discussion we only consider concavity, since the corresponding analysis for convexity is exactly the same, except with the obvious signs and inequalities reversed.

**Proposition 1:** If $f$ is concave and differentiable at $x_0$, then, for all $x \in U$,

$$f(x_0) + Df(x_0)(x - x_0) \geq f(x).$$

**Proof:** Suppose, for some $x \in U$, that $f(x) > f(x_0) + Df(x_0)(x - x_0)$. For any $0 < \alpha < 1$
we have
\[
\frac{f(x_0 + \alpha(x - x_0)) - f(x_0) - Df(x_0)(\alpha(x - x_0))}{\|\alpha(x - x_0)\|} = \frac{f((1 - \alpha)x_0 + \alpha x) - f(x_0) - \alpha Df(x_0)(x - x_0)}{\alpha\|x - x_0\|} \geq \frac{(1 - \alpha)f(x_0) + \alpha f(x) - f(x_0) - \alpha Df(x_0)(x - x_0)}{\alpha\|x - x_0\|} = \frac{f(x) - f(x_0) - Df(x_0)(x - x_0)}{\|x - x_0\|}.
\]
Since the last expression is positive and independent of \(\alpha\), this contradicts the definition of \(Df(x_0)\). \(\blacksquare\)

c. When \(f\) is strictly concave, the inequality above is strict.

**Proposition 2:** If \(f\) is strictly concave and differentiable at \(x_0\), then
\[
f(x_0) + Df(x_0)(x - x_0) > f(x)
\]
for all \(x \in U, x \neq x_0\).

**Proof:** Aiming at a contradiction, suppose there is some \(x \in U\) such that \(f(x) - f(x_0) \geq Df(x_0)(x - x_0)\). Then the definition of strict concavity gives
\[
f\left(\frac{1}{2}x + \frac{1}{2}x_0\right) > \frac{1}{2}f(x) + \frac{1}{2}f(x_0) = f(x_0) + \frac{1}{2}(f(x) - f(x_0)) \geq f(x_0) + \frac{1}{2}Df(x_0)(x - x_0).
\]
Now Proposition 1 gives
\[
f(x_0) + \frac{1}{2}Df(x_0)(x - x_0) = f(x_0) + Df(x_0)\left(\frac{1}{2}(x - x_0)\right) \geq f(x_0) + \frac{1}{2}(x - x_0)) = f\left(\frac{1}{2}x + \frac{1}{2}x_0\right).
\]
Combining these two inequalities gives the desired contradiction. \(\blacksquare\)

D. The consequences of these results for optimization are extremely important.

**Theorem:** Suppose that \(f\) is concave and differentiable at \(x^*\). Then \(x^*\) is a global maximizer for \(f\) if and only if \(Df(x^*) = 0\). If \(f\) is strictly concave and \(Df(x^*) = 0\), then \(x^*\) is the unique global maximizer of \(f\).
Proof: We know that $Df(x^*) = 0$ if $x^*$ is a global maximizer for $f$. If $Df(x^*) = 0$, then Proposition 1 implies that $x^*$ is a global maximizer for $f$, and if $f$ is strictly concave, then Proposition 2 implies that it is the unique maximizer.

III. Concavity and the Second Derivative

A. For $x \in U$ let

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

B. For $C^2$ functions, concavity is equivalent to negative semidefiniteness of the second derivative:

Theorem: If $f$ is $C^2$ and concave, then, for every $x \in U$, $D^2 f(x)$ is negative semidefinite.

Proof: If, for some $x \in U$, $D^2 f(x)$ is not negative semidefinite, there is some $v \in \mathbb{R}^n$ such that $v'D^2 f(x)v > 0$. Let $\epsilon = v'D^2 f(x)v/2$. For sufficiently small $\alpha > 0$ we have

$$f(x + \alpha v) \geq f(x) + Df(x)(\alpha v) + (\alpha v)'D^2 f(x)(\alpha v) - \epsilon \cdot \|\alpha v\|^2$$

$$= f(x) + Df(x)(\alpha v) + \alpha^2(v'D^2 f(x)v - \epsilon \cdot \|v\|^2)$$

$$> f(x) + Df(x)(\alpha v),$$

which contradicts Proposition 1.

Theorem: If $f$ is $C^2$, then $f$ is concave if and only if, for every $x \in U$, $D^2 f(x)$ is negative semidefinite. If, for every $x \in U$, $D^2 f(x)$ is negative definite, then $f$ is strictly concave.

Proof: It is easy to see that the definition of (strict) concavity is equivalent to the following condition: for all $x_0 \in U$, $v \in \mathbb{R}^n \setminus \{0\}$, and $\epsilon > 0$ such that $x_0 + v \in U$ for all $t \in (-\epsilon, \epsilon)$, if $\gamma : (-\epsilon, \epsilon) \to U$ is the function $\gamma(t) = x_0 + tv$ and $g = f \circ \gamma$, then $g$ is (strictly) concave.

From elementary calculus we know that $g$ is concave if and only if $g''(t) \leq 0$ for all $t$, and if $g''(t) < 0$ for all $t$, then $g$ is strictly concave.
Applying the chain rule, we compute that $g''(t) = v'D^2f(x_0 + tv)v$. It follows that $g''(t) \leq (\leq) 0$ for all $x_0, v, \varepsilon, t$ as above, if and only if $D^2f(x)$ is negative semidefinite (definite) for all $x \in U$.

IV. Economic Applications

A. For a production function $y = f(L)$, concavity of $f$ is expressed economically by saying that $f$ exhibits diminishing returns. When $f$ is convex we say that it exhibits increasing returns.

1. Exercise: Prove that the profit function $\pi(L) = p \cdot f(L) - wL$ is concave when $f$ is. (Implicitly we are assuming that the price $p$ of output and the wage $w$ of labor are beyond the firm’s control.)

B. Consider a consumer whose preferences are characterized by a continuous function $u : \mathbb{R}^n_+ \to \mathbb{R}$. The associated expenditure function describes the minimum amount of money required to purchase a bundle yielding some utility level. Specifically, for a price vector $p = (p_1, \ldots, p_n) \in \mathbb{R}^n_+$ and a utility level $u \in \mathbb{R}$ let

$$e(p, u) = \min \{ p \cdot x = p_1 \cdot x_1 + \ldots + p_n \cdot x_n : u(x) \geq u \}.$$

**Theorem:** For given $u$ the expenditure function is concave as a function of prices.

**Proof:** Consider price vectors $p^0, p^1 \in \mathbb{R}^n_+$, and let $x^0$ and $x^1$ be consumption bundles that minimize expenditure at these price vectors, respectively, while attaining utility level $u$. For $0 < \alpha < 1$ let $x^\alpha$ be a consumption bundle that minimizes expenditure at the price vector $p^\alpha = (1-\alpha)p^0 + \alpha p^1 = ((1-\alpha)p^0_1 + \alpha p^1_1, \ldots, (1-\alpha)p^0_n + \alpha p^1_n)$ while attaining utility
level \( u \). Then
\[
e(p^\alpha, u) = p^\alpha \cdot x^\alpha
\]
\[
= ((1 - \alpha)p^0 + \alpha p^1) \cdot x^\alpha
\]
\[
= (1 - \alpha)p^0 \cdot x^\alpha + \alpha p^1 \cdot x^\alpha
\]
\[
\geq (1 - \alpha)p^0 \cdot x^0 + \alpha p^1 \cdot x^1
\]
\[
= (1 - \alpha)e(p^0, u) + \alpha e(p^0, u).
\]

The inequality here is derived from the fact that, at price vector \( p^0 \), \( x^\alpha \) cannot be less expensive than \( x^0 \), while \( x^\alpha \) cannot be less expensive than \( x^1 \) at price vector \( p^1 \).

IV. Valuation of Risky Prospects

A. For each of the following two problems, decide which alternative you would choose:

1. **Problem 1:** Choose between the following two probability distributions over monetary prizes:
   a. **Lottery 1A:** $2,000,000 with certainty.
   b. **Lottery 1B:** $2,000,000 with probability 0.89, $10,000,000 with probability 0.10, and nothing with probability 0.01.

2. **Problem 2:** Choose between the following two probability distributions over monetary prizes:
   a. **Lottery 2A:** $2,000,000 with probability 0.11 and nothing with probability 0.89.
   b. **Lottery 2B:** $10,000,000 with probability 0.10 and nothing with probability 0.90.

B. The expected utility theory of decision making under uncertainty asserts that, in a choice between “lotteries” over certain outcomes, lotteries are valued by taking the average (according to the probabilities) of numerical valuations of the certain outcomes.
1. The most important applications have agents valuing different levels of wealth. Let \( U: \mathbb{R}_+ \rightarrow \mathbb{R} \) be a continuous function, where \( U(w) \) might be thought of as the “lifetime psychic pleasure” of having wealth level \( w \).

2. Thus the expected utility of a lottery that pays \( w_1 \) with probability \( p_1 \), \( w_2 \) with probability \( p_2 \), \ldots, \( w_K \) with probability \( p_K \), is

\[
\sum_{k=1}^{K} p_k \cdot U(w_k).
\]

(Here we require that \( p_1 + \cdots + p_K = 1 \).)

3. The theory has the property that your choice between two lotteries does not depend on whether they are choices in and of themselves, or are presented as possible continuations in a larger “tree” of possible events.

4. The statistician Leonard “Jimmy” Savage is, perhaps, the person most responsible for showing that the expected utility theory is essentially the unique theory with this property.

5. If you choose 1A and 2B in the choice problems above, you violated this principle.

   a. Don’t feel bad: L. J. Savage made the same choices.

C. As above, consider an agent whose preferences over lotteries are governed by expected utility, with underlying utility function \( U: \mathbb{R}_+ \rightarrow \mathbb{R} \).

1. We will always assume that \( U \) is strictly increasing.

2. The agent is said to be risk averse if \( U \) is concave.

3. For any lottery ‘\( w_1 \) with probability \( p_1 \), \ldots, \( w_K \) with probability \( p_K \),’ there is a expectation: \( p_1 w_1 + \cdots + p_K w_K \).

**Theorem:** If \( U \) is concave, then any lottery is (weakly) less desirable than receiving the lottery’s expectation with certainty:

\[
p_1 U(w_1) + \cdots + p_K U(w_K) \leq U(p_1 w_1 + \cdots + p_K w_K).
\]
Proof: In the case $K = 2$ the claim amounts to the definition of concavity. For larger $K$ we argue by induction.

\[ p_1 U(w_1) + \cdots + p_K U(w_K) \]
\[ = p_1 U(w_1) + (1 - p_1) \cdot \left( \frac{p_2}{1 - p_1} U(w_2) + \cdots + \frac{p_K}{1 - p_1} U(w_K) \right) \]
\[ \leq p_1 U(w_1) + (1 - p_1) U \left( \frac{p_2}{1 - p_1} w_2 + \cdots + \frac{p_K}{1 - p_1} w_K \right) \]
\[ \leq U \left( p_1 w_1 + (1 - p_1) \left( \frac{p_2}{1 - p_1} w_2 + \cdots + \frac{p_K}{1 - p_K} w_K \right) \right) \]
\[ = U(p_1 w_1 + p_2 w_2 + \cdots + p_K w_K). \]