Mathematics for Economists
Aalto BIZ
Spring 2021
Juuso Välimäki

Convex and concave functions

In this last section of Part I of the course, we take a first look at the extremely important question of convexity and concavity of functions. These notions revolve around quite general geometric notions in $\mathbb{R}^n$ and you will see applications in many different areas of economic theory (in particular under the title of 'duality theory'. For us now, the most immediate questions relate to the curvature of non-linear functions and their extrema. The really useful observation for optimization is that for concave functions, the first-order necessary conditions for minima and maxima are also sufficient. In particular, if f is concave and $Df(\hat{x}) = 0$, then f has a global maximum at$\hat{x}$.

Basic definitions

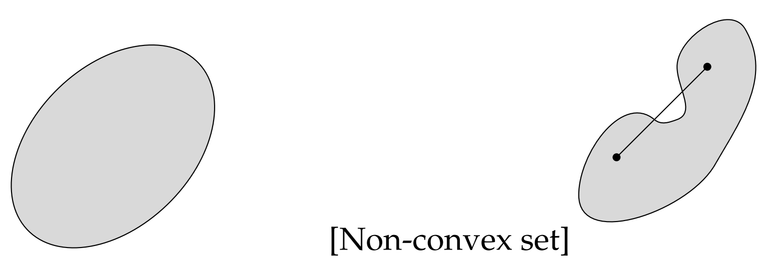
We start with a definition of convex sets.

Definition 1. A set X *is convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in X$ and for all $\lambda \in [0, 1]$, we have:

$$\lambda \boldsymbol{x} + (1 - \lambda)\, \boldsymbol{y} \in X.$$

We call $\lambda \boldsymbol{x} + (1 - \lambda)\, \boldsymbol{y}$ a *convex combination* of $\boldsymbol{x}$ and $\boldsymbol{y}$.

[Convex set]



[Non-convex set]

Figure 1: Graphical interpretation of convex sets.

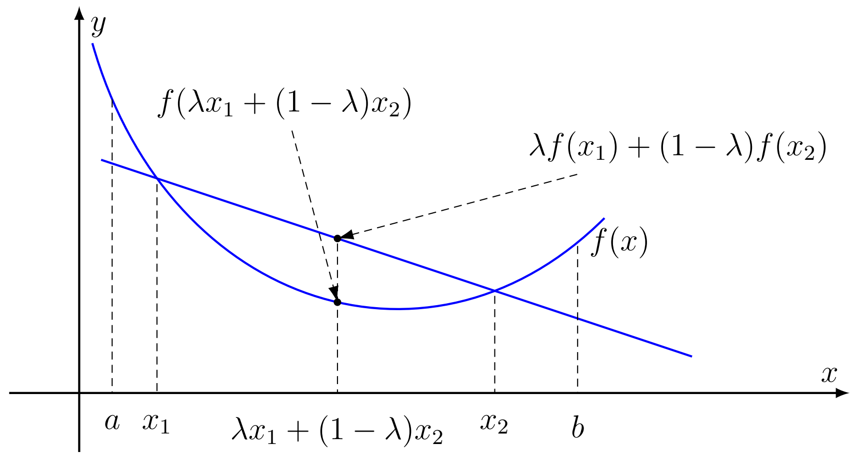


Figure 2: A convex function $f(x)$

On the real line, convex sets are intervals $a \leq x \leq b$ for some $-\infty \leq a \leq b \leq \infty$. In $\mathbb{R}^n$, convex sets are sets X with the property that when you connect linearly two points in X, the entire connecting line is also in X. Hence a disk in the plane is convex and a cube in the three dimensional space are convex, but the circle in the plane is not, a disk with the center removed is not, a doughnut in three dimensions is not etc.

Consider a real-valued function $f : X \to \mathbb{R}$, where X is a convex set.

Definition 2. The function f *is convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in X$ and for all $\lambda \in [0, 1]$, we have:

$$f\left(\lambda \boldsymbol{x} + (1 - \lambda) \boldsymbol{y}\right) \leq \lambda f\left(\boldsymbol{x}\right) + (1 - \lambda) f\left(\boldsymbol{y}\right).$$

f *is concave* if

$$f\left(\lambda \boldsymbol{x} + (1 - \lambda) \boldsymbol{y}\right) \geq \lambda f\left(\boldsymbol{x}\right) + (1 - \lambda) f\left(\boldsymbol{y}\right).$$

Observations:

- If $f(\boldsymbol{x})$ is convex, then $-f(\boldsymbol{x})$ is concave.
- If $f(\boldsymbol{x})$ is convex, then $af(\boldsymbol{x})$ is convex if $a > 0$.
- If $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are convex, then $h(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$ is convex.

- If $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are convex, then $h(\boldsymbol{x}) = f(\boldsymbol{x})\, g(\boldsymbol{x})$ is not necessarily convex. (Give an example for both cases, i.e. where the product of convex functions is convex and where it is not).

- Exercise: Assume that $f : X \to \mathbb{R}$ is convex and $g : \mathbb{R} \to \mathbb{R}$ is also convex. Is $g(f(\boldsymbol{x}))$ convex? What if g is increasing and convex?

- (Optional Exercise): Assume that $f : X \to \mathbb{R}$ is a convex function. Show that the set

  $$\{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} \quad | \quad \boldsymbol{x} \in X, y \geq f(\boldsymbol{x})\}$$

  is a convex set.

- If $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are convex, then $h(\boldsymbol{x}) = \max\{f(\boldsymbol{x}), g(\boldsymbol{x})\}$ is convex.

  Proof: Since by assumption, f and g are convex, we have:

  $$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

  and

  $$g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y}).$$

  By definition,

  $$\begin{aligned} h(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max\{f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}), g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y})\} \\ &\leq \max\{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}), \lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y})\} \\ &\leq \lambda \max\{f(\boldsymbol{x}), g(\boldsymbol{x})\} + (1-\lambda)\max\{f(\boldsymbol{y}), g(\boldsymbol{y})\} \\ &= \lambda h(\boldsymbol{x}) + (1-\lambda) h(\boldsymbol{y}). \end{aligned}$$

  The first inequality follows from the convexity of f and g.The second follows by choosing the larger of $f(\cdot), g(\cdot)$ for $\boldsymbol{x}, \boldsymbol{y}$. The last equality is just the definition of h.

- The same result is true for an arbitrary set of convex functions. Let $f(\boldsymbol{x}; \alpha)$ be convex in $\boldsymbol{x}$for all α. Then

  $$g(\boldsymbol{x}) = \max_{\alpha} f(\boldsymbol{x}; \alpha)$$

  is convex. The proof is identical to the one above.

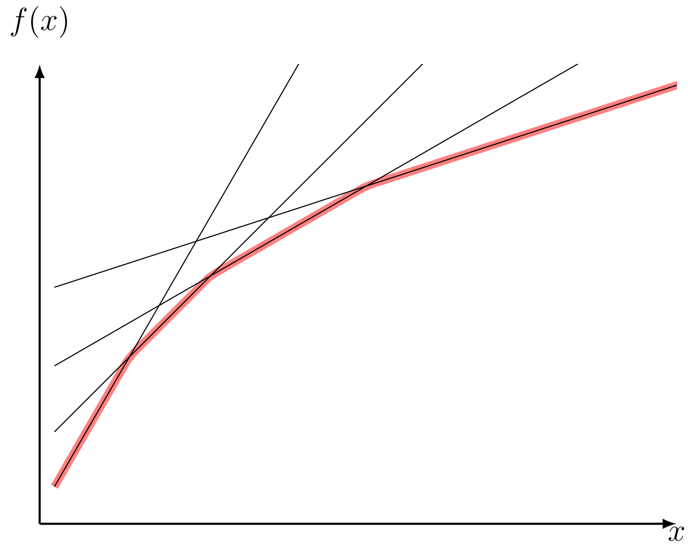


Figure 3: The minimum of linear functions (in red) is concave

- Since linear functions are convex, this result holds for any set of linear functions.
- Since
$$\max\{f(\boldsymbol{x}), g(\boldsymbol{x})\} = -\min\{-f(\boldsymbol{x}), -g(\boldsymbol{x})\},$$
and since $-f$ is concave when f is convex, we get:
$$g(\boldsymbol{x}) = \min_{\alpha} f(\boldsymbol{x}; \alpha)$$
is concave if $f(\boldsymbol{x}; \alpha)$ is concave in $\boldsymbol{x}$ for all α.

Example 1 (Profit function of a firm)**.** A competitive firm sells output y at price p_0 and buys inputs $\boldsymbol{x} = (x_1, ..., x_n)$ at input prices $(p_1, ..., p_n)$.

Its profit is
$$p_0 y - \Sigma_{i=1}^{n} p_i x_i.$$
The maximization problem is then
$$\max_{y, \boldsymbol{x} \in F} p_0 y - \Sigma_{i=1}^{n} p_i x_i,$$
where F is the feasible set determined by technological possibilities.

The profit function of the firm gives the maximum achievable profit amongst the feasible set at input and output prices p.

$$\pi(\boldsymbol{p}) = \pi(p_0, p_1, ..., p_n) = \max_{y,\boldsymbol{x}\in F} p_0 y - \Sigma_{i=1}^n p_i x_i$$

Since the profit from a fixed feasible production is a linear function of the prices p, the profit function is the maximum over linear functions and therefore convex in p.

Example 2 (Expenditure minimization)**.** Let X be the feasible set for inputs $\boldsymbol{x} = (x_1, ..., x_n)$ and $\boldsymbol{p} = (p_1, ..., p_n)$ be the input prices. The expenditure function

$$e(\boldsymbol{p}; X) = \min_{\boldsymbol{x}\in X} \boldsymbol{p}\cdot\boldsymbol{x} = \min_{\boldsymbol{x}\in X} \Sigma_{i=1}^n p_i x_i$$

is a concave function by the same argument as above.

These two examples show that convexity and concavity play a key role in economic applications. We shall see more applications when we discuss constrained optimization and value functions of optimization problems. Is there an economic intuition for the maximum of linear functions being convex? We'll return to this after some further characterizations of convex functions.

Convexity and concavity of differentiable functions

When $f : \mathbb{R} \to \mathbb{R}$, and f is convex and differentiable, it it is easy to see by drawing a picture that for all x, y we have:

$$f(y) - f(x) \geq f'(x)(y - x).$$

This just says that the graph $(x, f/x))$ of a convex function f is above all of its tangent lines.

Proposition 1. A differentiable function $f : \mathbb{R} \to \mathbb{R}$ is convex if and only if

$$f(y) - f(x) \geq f'(x)(y - x) \text{ for all } x, y.$$

Proof. i) Let f be convex. Then for all x, y:

$$\begin{gathered} f(\lambda x + (1-\lambda) y) \\ = f(x + (1-\lambda)(y - x)) \\ \leq \lambda f(x) + (1-\lambda) f(y) \\ = f(x) + (1-\lambda)(f(y) - f(x)) \end{gathered}$$

or

$$\frac{f\left(x+(1-\lambda)\left(y-x\right)\right)-f\left(x\right)}{\left(1-\lambda\right)}\leq f\left(y\right)-f\left(x\right),$$

or

$$\left(y-x\right)\frac{f\left(x+(1-\lambda)\left(y-x\right)\right)-f\left(x\right)}{\left(1-\lambda\right)\left(y-x\right)}\leq f\left(y\right)-f\left(x\right).$$

Letting $\lambda\to 1$, we get:

$$\left(y-x\right)f'\left(x\right)\leq f\left(y\right)-f\left(x\right).$$

ii) Assume that

$$f\left(y\right)-f\left(x\right)\geq f'\left(x\right)\left(y-x\right)\text{ for all }x,y.$$

Then

$$f\left(x\right)-f\left(\lambda x+\left(1-\lambda\right)y\right)\geq\left(1-\lambda\right)f'\left(\lambda x+\left(1-\lambda\right)y\right)\left(x-y\right)$$

and

$$f\left(y\right)-f\left(\lambda x+\left(1-\lambda\right)y\right)\geq-\lambda f'\left(\lambda x+\left(1-\lambda\right)y\right)\left(x-y\right).$$

Multiply the first inequality by λ and the second by $(1-\lambda)$and sum together to get the definition of convex functions. □

Consider next $f:X\to\mathbb{R}$, where X is a convex subset of $\mathbb{R}^n$.

Proposition 2. A differentiable function $f:X\to\mathbb{R}$ is convex if and only if

$$f\left(\boldsymbol{y}\right)-f\left(\boldsymbol{x}\right)\geq Df\left(\boldsymbol{x}\right)\left(\boldsymbol{y}-\boldsymbol{x}\right).$$

Proof. We start with a preliminary result: f is convex if and only if $g_{\boldsymbol{x},\boldsymbol{y}}\left(\lambda\right):=f\left(\left(1-\lambda\right)\boldsymbol{x}+\lambda y\right)$ is convex for all $\boldsymbol{x},\boldsymbol{y}$. In other words, convexity is equivalent to convexity along convex combinations. The proof of this is left as a relatively easy exercise.

By the chain rule,

$$\begin{aligned}g'_{\boldsymbol{x},\boldsymbol{y}}\left(\lambda\right) &= \Sigma_{i=1}^{n}\frac{\partial f\left(\boldsymbol{x}+\lambda\left(\boldsymbol{y}-\boldsymbol{x}\right)\right)}{\partial x_i}\left(y_i-x_i\right)\\ &= Df\left(\boldsymbol{x}+\lambda\left(\boldsymbol{y}-\boldsymbol{x}\right)\right)\left(\boldsymbol{y}-\boldsymbol{x}\right).\end{aligned}$$

By the previous theorem, $g_{\boldsymbol{x},\boldsymbol{y}}\left(\lambda\right)$ is convex if and only if

$$g_{\boldsymbol{x},\boldsymbol{y}}\left(1\right)-g_{\boldsymbol{x},\boldsymbol{y}}\left(0\right)\geq g'_{\boldsymbol{x},\boldsymbol{y}}\left(0\right).$$

In other words if and only if

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \geq Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}).$$

□

This is the multidimensional generalization to the geometric notion that the graphs of convex functions lie above their tangent lines. Can you formulate this condition in terms of level curves and gradients? What is the corresponding result to concave functions?

Exercise: Using this condition, show that if f is convex (concave) on the convex set X and $Df(\hat{\boldsymbol{x}}) = 0$, then $\hat{\boldsymbol{x}}$ is a global minumum (maximum) of f on X

Second derivatives and convexity

Start again with functions of a single variable. By Taylor's theorem,

$$f(y) = f(x) + f'(x)(y - x) + \frac{1}{2} f''(x)(y - x)^2 + \frac{1}{6} f'''(x)(y - x)^3 + \text{ h.o.t.}$$

In order to have

$$f(y) - f(x) \geq f'(x)(y - x)$$

for $|y - x|$ small, we must have

$$f''(x) \geq 0.$$

In other words, convex functions have a positive second derivative.

Taylor's theorem with a remainder term of second degree:

$$f(y) = f(x) + f'(x)(y - x) + \frac{1}{2} f''(z)(y - x)^2$$

for some $z \in [x, y]$. If f'' is everywhere non-negative, we get:

$$f(y) - f(x) \geq f'(x)(y - x)$$

for all y, x and f is therefore convex.

Let's generalize now to $f : X \to \mathbb{R}$, where X is a convex subset of $\mathbb{R}^n$ n.

We use again the function

$$g_{\boldsymbol{x},\boldsymbol{y}}(\lambda) = f((1 - \lambda)\boldsymbol{x} + \lambda \boldsymbol{y}) = f(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x}))$$

and consider the second derivatives of g.

Convexity corresponds to positive semidefiniteness of the Hessian matrix. Concavity corresponds to negative semidefiniteness of the Hessian matrix. Hence we see an immediate connection between convexity and the second order conditions for optimality.

Quasiconvex and quasiconcave functions

Even though the name suggests something extremely technical and tedious, quasiconcavity is actually one of the most important notions for functions in economic theory. We begin with the definitions and properties of quasiconcave functions, but at the end of this section, I will discuss why this is such a useful definitions for economic modeling.

Definition 3.

A function f on a convex set X is *quasiconcave* if for all $\boldsymbol{x}, \boldsymbol{y} \in X$ and for all $\lambda \in [0,1]$

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

f is *quasiconvex* is for all $\boldsymbol{x}, \boldsymbol{y} \in X$ and for all $\lambda \in [0,1]$

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}.$$

Exercise: f is quasiconcave, then $-f$ is quasiconvex.
We can make some observations:

- If f is quasiconcave, then af is quasiconcave if $a > 0$.
- If f and g are quasiconcave $f + g$ is not necessarily quasiconcave.
- All monotone (i.e. all increasing and all decreasing) functions of a single variable are both quasiconcave and quasiconvex.
- All concave functions are quasiconcave. Show this as an exercise.
- Not all quasiconcave functions are concave.
- If f is a quasiconcave function and g is a strictly increasing function, then $h(\boldsymbol{x}) = g(f(\boldsymbol{x}))$ is a quasiconcave function.

Proof:

$$\begin{aligned} h(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= g\left(f\left(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\right)\right) \\ &\geq g\left(\min\{f(\boldsymbol{x}), f(\boldsymbol{y})\}\right) \\ &= \min\{g(f(\boldsymbol{x})), g(f(\boldsymbol{y}))\} \\ &= \min\{h(\boldsymbol{x}), h(\boldsymbol{y})\}. \end{aligned}$$

Exercise: where was the increasing property of g used in the proof?

An upper contour set of function f for value α is denoted by $U(f;\alpha)$ and defined as:

$$U(f\,;\alpha) := \{\boldsymbol{x} \in X | f(\boldsymbol{x}) \geq \alpha\}.$$

Interpretation: if f is a utility function, $U(f;\alpha)$ is the better side of the indifference curve giving utility level α.

Proposition 3. A function f is quasiconcave if and only if $U(f;\alpha)$ is a convex set for all α.

Proof. i) Assume that f is quasiconcave and $\boldsymbol{x}, \boldsymbol{y} \in U(f;\alpha)$. Then $f(\boldsymbol{x}) \geq \alpha$, $f(\boldsymbol{y}) \geq \alpha$ and by quasiconcavity of f,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq \min\{f(\boldsymbol{x}), f(\boldsymbol{y})\} \geq \alpha.$$

In other words

$$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in U(f;\alpha),$$

and therefore $U(f;\alpha)$ is convex.

ii) Assume that $U(f;\alpha)$ is a convex set for all α. Then

$$\boldsymbol{x}, \boldsymbol{y} \in U(f, \min\{f(\boldsymbol{x}), f(\boldsymbol{y})\}),$$

and

$$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in U(f, \min\{f(\boldsymbol{x}), f(\boldsymbol{y})\}).$$

But then by the definition of $U(f;\alpha)$:

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

□

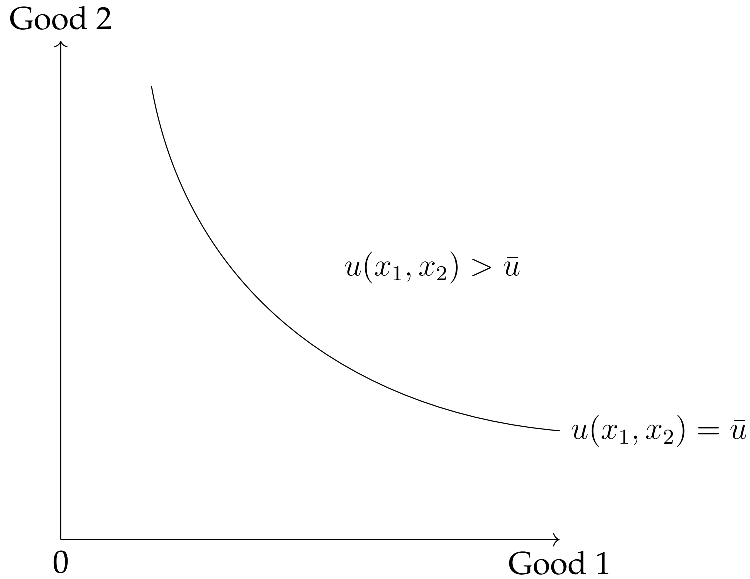


Figure 4: The better than set for a quasiconcave. utility function is convex

Let me make a methodological point here. For economic modeling, the exact mathematical form of the utility function is unimportant since many different functions represent the same preferences as discussed before. The meaningful properties for an economic model relate to the preferences, i.e. to the indifference curves. As a result, an assumption on the shape of these curves or their upper contour sets are meaningful. The convexity of upper contour sets of utility functions is a meaningful property and it is often assumed in models of consumer choice. Notice that the shapes of the upper contour sets remain unchanged when going from $u(x)$ to $v(u(x))$ for a strictly increasing v. This follows from the observation that

$$U(v(u); v(\alpha)) = U(u; \alpha) \text{ for all } \alpha.$$

We end this subsection with a useful special case of quasiconcave functions:

Definition 4. A function f on a convex set X is *strictly quasiconcave* if for all $\boldsymbol{x}, \boldsymbol{y} \in X$ and for all $\lambda \in (0, 1)$

$$f(\lambda \boldsymbol{x} + (1 - \lambda) \boldsymbol{y}) > \min\{f(\boldsymbol{x}), f(\boldsymbol{y})\}.$$

The following exercise shows why strict quasiconcavity is very useful for optimization problems.

Exercise: show that if a strictly quasiconcave function has a maximum, then the maximum is unique.

Quasiconcavity and differentiability

A differentiable function f on a convex set X is quasiconcave if and only if for all $\boldsymbol{x}, \boldsymbol{y} \in X$:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) \Rightarrow Df(\boldsymbol{x})(\boldsymbol{y}-\boldsymbol{x}) \geq 0.$$

Exercise: Compare this to the definition of concavity for differentiable functions and relate this condition to the geometry of upper contour sets and tangent planes to the upper contour sets.

The second order conditions for quasiconcavity based on bordered Hessian matrices are extremely complex and contain little economic intuition. The textbook on pages 527-531 gives an introduction to this.

With these preliminaries, we are ready for constrained optimization in Part II of this course.