Convexity (and Homogeneity)

Stéphane Dupraz

This chapter deals with both convex sets and convex functions. Convexity and concavity (and quasi-convexity and quasi-concavity) of functions play an essential role in economics first because they play an essential role in optimization: assuming the objective of an optimization problem is convex/concave guarantee that a local minimum/maximum is a global minimum/maximum. Besides, convexity assumptions on sets are central to the second welfare theorem and the Arrow-Debreu result on the existence of a walrasian equilibrium. We end the chapter with a distinct topic: homogeneous functions.

1 Convex Sets

1.1 Definition

The notion of a convex set is easy to grasp in the plane $\mathbb{R}^2$. There, a subset S is said to be convex if when taking any two points x and y in S, the whole straight line segment $[x, y]$ joining x and y—all the points "in-between" x and y—is also included in the set S. More generally, the only thing we need to define a segment—hence a convex subset—is for X to be a vector space. In a vector space, the **line segment between** x **and** y, noted $[x, y]$ is defined as the set of points $\{\lambda x + (1-\lambda)y, 0 \leq \lambda \leq 1\}$. (We note (x, y) the open line segment $\{\lambda x + (1-\lambda)y, 0 < \lambda < 1\}$). The definition of convexity follows:

Definition 1.1. *Let V be a vector space. A subset S of V is* ***convex*** *iff for any $x, y \in S$, the segment $[x, y] \subseteq S$, i.e. iff*

$$\forall x, y \in S, \forall \lambda \in [0, 1], \lambda x + (1-\lambda)y \in S$$

Note the parallel with a vector subspace, which require more strongly that for all λ and μ, $\lambda x + \mu y \in V$: the only difference is that for a convex set, we require only that linear combinations such that $\lambda, \mu \geq 0$ and $\lambda + \mu = 1$ be in S. In words, we require that only the segment $[x, y]$, and not the whole plane going through x, y and 0 be in S. So any vector subspace of V (in particular V) is convex. Pursuing the analogy, we can define convex combinations, in parallel to linear combinations:

Definition 1.2. *Let* $x_1, ..., x_n$ *be* n *vectors of* V*. A* ***convex combination*** *of* $x_1, ..., x_n$ *is a vector* $\lambda_1 x_1 + ... + \lambda_n x_n$ *for* n *scalars* $\lambda_1, ..., \lambda_n \in \mathbb{R}$ *such that* $\lambda_i \geq 0$ *for all* i *and* $\sum \lambda_i = 1$.

You can think of the λ_i's as weights and of a convex combination as a weighted average of the points. Still on our analogy with vector subspaces, we can check that any (possibly infinite) intersection of convex sets is convex, so that we can define the smallest convex set that contains a subset $S \subseteq \mathbb{R}^n$; we call it the convex hull of S.

Definition 1.3. *Let* V *be a vector space and* S *a subset of* V.

The ***convex hull of*** S*, noted* $Co(S)$*, is the smallest convex set that contains* S*, that is the intersection of all the convex sets that contain* S*:*

$$Co(S) = \bigcap \{C, C \text{ a convex set, and } S \subseteq C\}$$

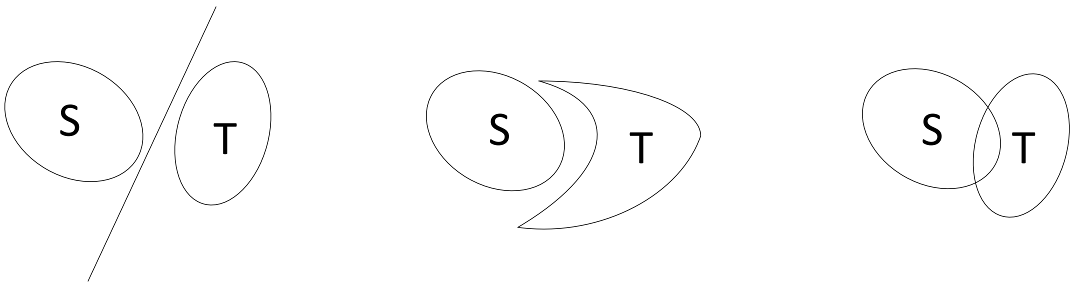
And again, if S is finite, we can give a more concrete description.

Proposition 1.1. *Let* $x_1, ..., x_n$ *be* n *points of* V*. The convex hull of* $\{x_1, ..., x_n\}$ *is the set of all convex combinations of* $x_1, ..., x_n$*:*

$$Co(x_1, ..., x_n) = \{\sum_{i=1}^{n} \lambda_i v_i, \lambda_i \geq 0 \text{ for all } i, \sum_i \lambda_i = 1\}.$$

2 Separating Hyperplane Theorems in $\mathbb{R}^n$

The separating hyperplane theorems are key to the proof of the second welfare theorem. The idea they express is very easy to grasp in the plane $\mathbb{R}^2$. Look at the figures below:



On the left, the two subsets S and T can be separated by a straight line: S is on one side of the line, T on the other. Note that both sets are convex, and that they are disjoint. If one of the set is not convex, then it is easy to think of an example where no straight line can separate the two sets; look for instance at the middle figure. Note also that the two sets are disjoint. It is obvious that if the sets are not disjoints, we can think of an example where no straight line can separate the two sets; look for instance at the figure on the right.

The main (Minkowski's) separating hyperplane theorem states that the two conditions—both convex, and disjoint—are sufficient for a straight line to exist that separate the two sets. This seems fairly obvious in the plane. But the interest of the theorem is that it is expressed not just in the plane, but more generally in $\mathbb{R}^n$ for any n. So, before going to the theorem, how to generalize a straight line to $\mathbb{R}^n$? And what does it mean to separate two sets in $\mathbb{R}^n$?

2.1 Hyperplane, and what it means for an hyperplane to separate two sets

First, what is the equivalent of a straight line in $\mathbb{R}^n$? In the plane a straight line is what divides the plane into two half-planes; note that a straight line is an affine subspace of dimension $1 = 2 - 1$. In $\mathbb{R}^3$, what separates $\mathbb{R}^3$ into two half-spaces is a plane; note that a plane is an affine subspace of dimension $2 = 3 - 1$. So in $\mathbb{R}^n$, we consider affine subspaces of dimension $n - 1$: affine hyperplanes (we call them simply hyperplane from now on).

But we do not really visualize anything after $n = 3$: so in which sense do hyperplanes *"divide* $\mathbb{R}^n$ *into two half-spaces"*? In higher dimensions, drawings do not generalize well, but equations do. Remember that in the plane any straight line (any affine subspace) has an equation $ax + by + c = 0$ for some $a, b, c \in \mathbb{R}^3$ (meaning the straight line is the set $\{(x, y) \in \mathbb{R}^2 / ax + by + c = 0\}$). So we can define the two half-planes as $\{(x, y) \in \mathbb{R}^2 / ax + by + c < 0\}$ and $\{(x, y) \in \mathbb{R}^2 / ax + by + c > 0\}$. It would be nice to have a similar equation for an hyperplane. Well, it turns out that the $n = 2$ case is just a particular case of a more general result:

Proposition 2.1.

- *A subset H of $\mathbb{R}^n$ is a vector hyperplane of $\mathbb{R}^n$ iff there exist $p \neq 0 \in \mathbb{R}^n$ such that:*

$$x = (x_1, ..., x_n)' \in H \Leftrightarrow p'x = \sum_{i=1}^{n} p_i x_i = 0.$$

- *A subset H of $\mathbb{R}^n$ is an affine hyperplane of $\mathbb{R}^n$ iff there exist $p \neq 0 \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that:*

$$x = (x_1, ..., x_n)' \in H \Leftrightarrow p'x = \sum_{i=1}^{n} p_i x_i = c.$$

In both cases, it is possible to normalize p such that $||p||_2 = 1$.

Proof. Let us start with vector hyperplanes. Consider the set $H = \{x/p'x = 0\}$ for some $p \neq 0$. See it as the kernel of the $1 \times n$ matrix p'. As p is rank 1, its kernel has dimension $n - 1$.

Conversely, let H be a vector hyperplane. Let $(h_1, ..., h_{n-1})$ be a basis of H. We can complete the basis with a vector g to get a basis $(h_1, ..., h_{n-1}, g)$ of $\mathbb{R}^n$. A vector $x \in V$ can be written $x = \sum_{i=1}^{n-1} \lambda_i h_i + \lambda_g g$ in the basis; $x \in H$ iff $\lambda_g = 0$. Now, it is easy to check that the function $x \mapsto \lambda_g(x)$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$. But we know that linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ have necessarily the form $x \mapsto Ax$ for some $m \times n$ matrix. So here, $\lambda_g(x) = p'x$ for some $p \in \mathbb{R}^n$, and $x \in V$ iff $p'x = 0$. Finally, $p \neq 0$, otherwise all vectors x satisfy $0x = 0$ and the equation is not one of a hyperplane but of $\mathbb{R}^n$.

Let us now generalize the equation to affine hyperplanes. Consider the set $H = \{x/p'x = c\}$. The image of the matrix p' has dimension 1 (the rank of p'), so it is the entire real line $\mathbb{R}$. So there exists $x^* \in V$ such that $p'x^* = c$. Using the result for vector hyperplanes, we can write $H = \{x/p'(x - x^*) = 0\} = \{y/p'y = 0\} + x^*$: the translation of an vector hyperplane. The converse follows the same logic.

For the normalization, just note that if $p'x = c$ is the equation of H, then $(\lambda p)'x = \lambda c$ is also an equation of H for all $\lambda \neq 0 \in \mathbb{R}$. □

With an equation $p'x = c$ characterizing the hyperspace H, it is possible to talk about the two **half-spaces** $\{x/p'x < c\}$ and $\{x/p'x > c\}$ it divides $\mathbb{R}^n$ into. We then say that H separates two sets if one set is included into one half-space, and the other set into the other half-space. Because it is possible that the sets lie partly on the frontier $\{x/p'x = c\}$—the hyperplane—we distinguish between separation and strict separation:

Definition 2.1. *Let S and T be two subsets of $\mathbb{R}^n$, and H an hyperplane with equation $p'x = c$*

- *We say that the hyperplane $H = \{x/p'x = c\}$* ***separates*** *S and T iff:*

$$\forall s \in S, \forall t \in T, p's \leq c \leq p't$$

- *We say that the hyperplane $H = \{x/p'x = c\}$* ***strictly separates*** *S and T iff:*

$$\forall s \in S, \forall t \in T, p's < c < p't$$

2.2 The theorems

We state 3 separating hyperplane theorems, each one building on the previous one to culminate in Minkowski's separating hyperplane theorem, which deals with 2 convex disjoint sets. The 2 first theorems deal with separating a convex set and a point. We admit the three theorem. You can consult FMEA, section 13.6, for proofs.

Theorem 2.1. *Let S be a non-empty subset of $\mathbb{R}^n$ and t a point of $\mathbb{R}^n$.*

If S is convex and closed and $t \notin S$, then there exists a hyperplane that strictly *separates them.*

If S is not closed, strict separation is not guaranteed. But we can formulate a result with weak separation. When we require only weak separation, we can even let t belong to S, provided it does not belong to $int(S)$:

Theorem 2.2. *Let S be a non-empty subset of $\mathbb{R}^n$ and t a point of $\mathbb{R}^n$.*

If S is convex and $t \notin int(S)$, then there exists a hyperplane that separates them.

Finally:

Theorem 2.3. ***Minkowski's separating hyperplane theorem***

Let S and T be two non-empty subsets of $\mathbb{R}^n$.

If S and T are both convex, and disjoint, then there exists a hyperplane that separates them.

3 Convex and concave real-valued functions

3.1 Definition

We now define convexity (and concavity) for functions, not sets. The notion is defined for real-valued functions—functions that take values in $\mathbb{R}$. Besides, we restrict to functions from $\mathbb{R}^n$ to $\mathbb{R}$ (the definition would work for any vector subspace instead of $\mathbb{R}^n$, but most results and applications concern functions with arguments in $\mathbb{R}^n$, so let us be concrete). Precisely, we consider a function $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$.

A convex function is defined as a function whose area above the curve $\{(x, y)/y \geq f(x)\}$ (called the **epigraph** of f) is a convex set; a concave function is defined as a function whose area below the curve $\{(x, y)/y \leq f(x)\}$ (called the **subgraph** of f) is a convex set. These definitions connect the notion of a convex function to the notion of a convex set, but are not very useful in practice. It is easy to show that the definitions are equivalent to the following ones:

Definition 3.1.

- *f is **convex** iff S is a convex subset of $\mathbb{R}^n$ and:*

$$\forall x, y \in S, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

- *f is **concave** iff S is a convex subset of $\mathbb{R}^n$ and:*

$$\forall x, y \in S, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$$

*We also define **strictly convex** and **strictly concave** functions by requiring in addition that the inequality in the definition be strict whenever $x \neq y$ and $\lambda \in (0, 1)$ (if $x = y$ or $\lambda = 0$ or $\lambda = 1$, then the inequality is $f(x) \leq f(x)$, which is always true but never strict). A strictly convex (concave) function is obviously convex (concave).*

In practice we always forget about the requirement that S be convex because it is obvious: if S is not convex, then $f(\lambda x + (1 - \lambda)y)$ is not even defined! However, be careful never to try to prove that a function defined on a non-convex set is convex or concave. From the definition, a function f is concave iff $-f$ is convex.

A function can be neither convex nor concave. Can a function be both convex and concave? Yes, for instance affine functions (functions of the form $f(x) = a'x + b, a \in \mathbb{R}^n$ and $b \in \mathbb{R}$). (There are the only ones). Some ubiquitous convex/concave functions from $\mathbb{R}$ to $\mathbb{R}$:

- $\exp(x)$ is convex.
- $|x|^p, p \geq 1$ are convex.
- $\log(x)$ is concave.
- $|x|^p, 0 \leq p \leq 1$ are concave (defined on $\mathbb{R}_+$).

3.2 Properties

Simple results on operations on convex and concave functions are much useful in practice. First, sum and multiplication by a scalar. The sum of convex (concave) functions is convex (concave). As for multiplication by a scalar, we know it cannot hold in general, since if f is strictly convex, then $-f$ is strictly concave, hence not convex. However, multiplication by a *positive* scalar maintains convexity/concavity.

Proposition 3.1.

- *If f and g are convex (concave), so is $f + g$.*
- *If f is convex (concave) and $\lambda \geq 0$, then λf is convex (concave).*

The proof simply consists in summing or multiplying the definition inequalities. It follows that positive linear combinations of convex (concave) functions are convex (concave).

Second, composition of functions. The composition of two convex (concave) functions need not be convex (concave). However:

Proposition 3.2.

- *If f and g are convex and g is (weakly) increasing, then $g \circ f$ is convex.*
- *If f and g are concave and g is (weakly) increasing, then $g \circ f$ is concave.*

Proof. Again, the proof merely consists in writing the definition inequalities. If f is convex:

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

Which implies, *because g is increasing*:

$$\begin{aligned} g(f(\lambda x + (1 - \lambda)y)) &\leq g(\lambda f(x) + (1 - \lambda)f(y)) \\ &\leq \lambda g(f(x)) + (1 - \lambda)g(f(y)) \end{aligned}$$

Where the last inequality simply uses the convexity of g. ☐

Third, taking the maximum of convex functions preserves convexity, and taking the minimum of concave functions preserves concavity:

Proposition 3.3.

- *If f and g are convex, then* $\max(f, g)$ *is convex.*
- *If f and g are concave, then* $\min(f, g)$ *is concave.*

Proof. See problem-set. ☐

3.3 Jensen's inequality

The definition of a convex function relies on the value of f at convex combinations of two points x and y. What can we say about the value of f convex at convex combinations of more than two points? The Jensen inequality offers an answer:

Proposition 3.4. ***Jensen's inequality*** *Let $f : S \to \mathbb{R}, S \subseteq \mathbb{R}^n$ be a convex function.*

For any integer n, for any $x_1, ..., x_n \in S$, for any positive reals $\lambda_1, ..., \lambda_n$ such that $\sum \lambda_i = 1$,

$$f(\lambda_1 x_1 + ... + \lambda_n x_n) \leq \lambda_1 f(x_1) + ... + \lambda_n f(x_n)$$

(And correspondingly for concave functions).

Proof. The proof is by induction, starting from the base case $n = 2$ for which the statement is simply the definition of convexity. You are asked to detail it in the problem-set. ☐

3.4 Characterization for differentiable functions

The first characterization is for differentiable functions. A single-variable function is convex if and only if its curve is always above its tangent lines; more generally, a multi-variable function is convex iff its graph is always above its tangent hyperplanes:

Proposition 3.5. *Let $f : S \to \mathbb{R}$, S an open convex subset of $\mathbb{R}^n$. Assume f is differentiable.*

- *f is convex iff:*

$$\forall x, y, f(y) \geq f(x) + f'(x)(y-x)$$

- *f is strictly convex iff:*

$$\forall x \neq y, f(y) > f(x) + f'(x)(y-x)$$

And with the reverse inequalities for concavity.

Proof. We prove the first item. Assume f is convex. The idea is to look at $f'(x)(y-x)$ as the derivative along the vector $y-x$ at x: $f'_{y-x}(x)$. So form the ratio:

$$\frac{f(x+t(y-x))-f(x)}{t} = \frac{f((1-t)x+ty)-f(x)}{t}$$

For $t \in [0,1]$ (only small t matters for limits), we can use the definition of convexity in the right-hand side. So:

$$\frac{f(x+t(y-x))-f(x)}{t} \leq \frac{(1-t)f(x)+tf(y)-f(x)}{t} = f(y)-f(x)$$

At the limit $t \to 0$, we get the desired inequality.

Conversely, assume the characterization holds. Fix x, y, λ, define $z = \lambda x + (1-\lambda)y$. We want to show that $f(z) \leq \lambda f(x) + (1-\lambda)f(y)$. Use the characterization twice, both time picking the derivative in z, but picking alternatively x or y as the second point:

$$f(x) \geq f(z) + f'(z)(x-z) = f(z) + f'(z)(1-\lambda)(x-y)$$
$$f(y) \geq f(z) + f'(z)(y-z) = f(z) + f'(z)\lambda(y-x)$$

Summing λ times the first inequality and $(1-\lambda)$ times the second one eliminates $f'(z)$ and yields the result. □

3.5 Characterizations for $\mathcal{C}^2$ functions

For $\mathcal{C}^2$ functions, we have a characterization of convex functions, although we only have a sufficient condition for strict convexity. A function from $\mathbb{R}$ to $\mathbb{R}$ is convex iff its second-derivative is positive. The result generalizes to functions of $\mathbb{R}^n$ to $\mathbb{R}$ replacing positivity by positive semi-definiteness:

Proposition 3.6. *Let $f : S \to \mathbb{R}$, S an open convex subset of $\mathbb{R}^n$. Assume f is $\mathcal{C}^2$.*

- *f is convex iff its hessian $f''(x)$ is positive semi-definite for all x in S.*
- *f is concave iff its hessian $f''(x)$ is negative semi-definite for all x in S.*

Proof. We *admit* the result. See e.g. Sundaram theorem 7.10 for a proof. □

It seems intuitive that strict convexity should correspond to positive definiteness. The equivalence is false however: we only have sufficient conditions for strict convexity and strict concavity:

Proposition 3.7. *Let $f : S \to \mathbb{R}$, S an open convex subset of $\mathbb{R}^n$. Assume f is $\mathcal{C}^2$.*

- *If the hessian $f''(x)$ is positive definite for all x in S then f is strictly convex.*
- *If the hessian $f''(x)$ is negative definite for all x in S then f is strictly concave.*

Proof. We also *admit* this result. □

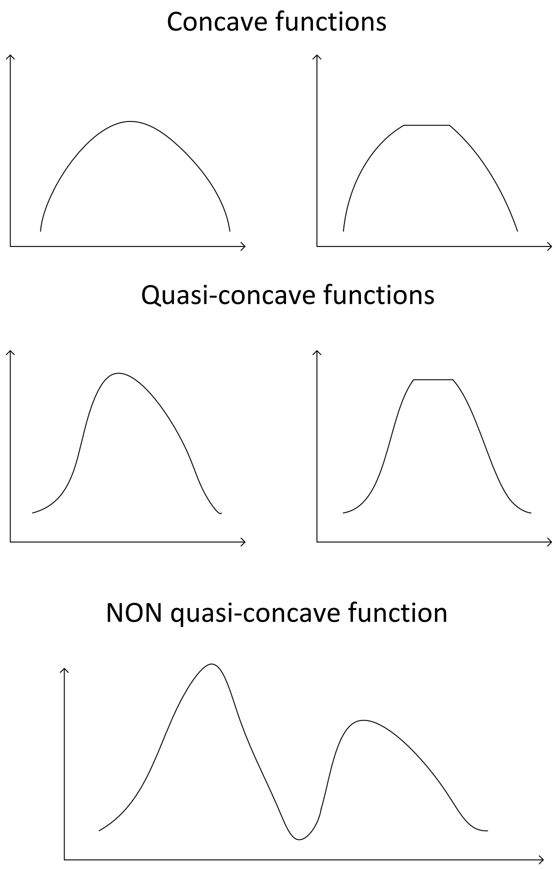
To get convinced that the converse if false, consider an example for $n = 1$. For functions from $\mathbb{R}$ to $\mathbb{R}$, $f''(x)$ positive definite simply means $f''(x) > 0$. The function $x \mapsto x^4$ is strictly convex. However, $f''(0) = 0$.

To show that the Hessian $f''(x)$ is (for instance) positive definite, remember the characterizations of positive definiteness through principal minors and eigenvalues that we have seen in linear algebra.

4 Quasi-convex and quasi-concave real-valued functions

4.1 Definition

One major application of convex and concave functions is optimization. As we will see, any local maximum of a concave function is a global maximum. To get the idea, consider functions of one variable as in the figure below. The two curves at the top represent concave functions. It is intuitive graphically that a concave function cannot have “two peaks”—although it can have a continuum of maxima if the function plateaus at its maximum. But if the only goal is to characterize functions that have at most one peak, then concavity is a unnecessarily strong assumption. Look at the two middle curves. They are not concave, but they still are “single-peaked”. In this section, we define the notion of quasi-concavity (quasi-convexity), which is a weaker requirement than concavity (convexity) and captures this idea of having at most a single peak.



So how to capture this idea? Draw on horizontal line on any of the five graphs in the figure, and consider all the x such that $f(x)$ is above that line. This is called an upper-level set:

Definition 4.1.

- *The* ***upper-level set of*** f ***at level*** a *is the set* $\{x/f(x) \geq a\}$.
- *The* ***lower-level set of*** f ***at level*** a *is the set* $\{x/f(x) \leq a\}$.

In the first 4 graphs, any upper-level set is convex. In the bottom graph, we can find upper-level sets that are not convex because the function has two peaks. We say that the first 4 functions are quasi-concave, but the bottom function is not:

Definition 4.2.

- *A function is* ***quasi-concave*** *if all its upper-level sets are convex.*
- *A function is* ***quasi-convex*** *if all its lower-level sets are convex.*

This is easy to express the definition analytically: if we note U_a the upper-level set at level a of f, then $x \in U_a$ is equivalent to $f(x) \geq a$. So:

$$\begin{aligned}
\text{f quasi-concave iff } & \forall a \in \mathbb{R}, \forall x, y, \forall \lambda \in [0,1], \Big[x, y \in U_a \Rightarrow \lambda x + (1-\lambda)y \in U_a\Big] \\
\text{iff } & \forall a \in \mathbb{R}, \forall x, y, \forall \lambda \in [0,1], \Big[f(x) \geq a, f(y) \geq a \Rightarrow f(\lambda x + (1-\lambda)y) \geq a\Big] \\
\text{iff } & \forall a \in \mathbb{R}, \forall x, y, \forall \lambda \in [0,1], \Big[\min(f(x), f(y)) \geq a \Rightarrow f(\lambda x + (1-\lambda)y) \geq a\Big] \\
\text{iff } & \forall x, y, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \geq \min(f(x), f(y))
\end{aligned}$$

(For the last equivalence, consider the two implications separately; for the first implication, pick $a = \min(f(x), f(y))$). This last line is the most practical definition:

Proposition 4.1.

A function is ***quasi-concave*** *iff:*

$$\forall x, y, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \geq \min(f(x), f(y))$$

*A function is **quasi-convex** iff:*

$$\forall x, y, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leq \max(f(x), f(y))$$

Written this way, it is self-evident that quasi-concavity implies concavity and quasi-convexity implies convexity. We define **strict quasi-concavity** (**strict quasi-convexity**) when the inequalities are strict whenever $x \neq y$ and $\lambda \in (0,1)$.

4.2 Properties

While the sum of two concave functions is concave, it is false for quasi-concave functions (we can easily construct a two-peak function by summing two single-peaked functions). But as a consolation prize: while the composition of a concave function by an increasing function is not in general concave, the composition of a quasi-concave function by an increasing function is:

Proposition 4.2.

- *If f is quasi-concave and g is increasing, then $g \circ f$ is quasi-concave.*
- *If f is quasi-convex and g is increasing, then $g \circ f$ is quasi-convex.*

This is straightforward to see from the characterization.

4.3 Characterization for differentiable functions

Just as concavity, quasi-concavity has a characterization for differentiable functions. Assume f is differentiable.

Proposition 4.3. *Let $f : S \to \mathbb{R}$, S an open convex subset of $\mathbb{R}^n$.*

- *f is quasi-concave iff:*

$$\forall x, y \in S, f(y) \geq f(x) \Rightarrow f'(x)(y-x) \geq 0.$$

- *f is quasi-convex iff:*

$$\forall x, y \in S, f(x) \geq f(y) \Rightarrow f'(x)(y-x) \leq 0.$$

Proof. We do the proof of one implication, for instance for quasi-concavity. Assume f is quasi-concave. Fix $x, y \in S$ and assume $f(y) \geq f(x)$. For any $t \in [0,1]$, $f(x + t(y-x)) \geq \min(f(x), f(y)) = f(x)$. So for all $t \in [0,1]$, $\frac{f(x+t(y-x))-f(x)}{t} \geq 0$. At the limit, $f'_{y-x}(x) \geq 0$. But $f'_{y-x}(x) = f'(x)(y-x)$.

We *admit* the converse. See e.g. Sundaram p.216 for a proof. □

4.4 Characterization for $\mathcal{C}^2$ functions

Just as concavity, quasi-concavity has a characterization for $\mathcal{C}^2$ functions. It relies on bordered matrices (we already encountered bordered matrices in linear algebra). The bordered hessian of a function f is its hessian matrix, bordered by the derivative of f.

Definition 4.3. *Let $f : S \to \mathbb{R}$, S an open subset of $\mathbb{R}^n$.*

The ***bordered hessian of*** *f is the $(n+1) \times (n+1)$ matrix:*

$$B(x) = \begin{pmatrix} 0 & f'(x) \\ \nabla f(x) & H(x) \end{pmatrix}$$

We note $B_k(x)$ the leading principal minor of order k of $B(x)$ (these are determinants, not matrices).

We state the results for quasi-concave functions. First the necessary condition:

Proposition 4.4. *Let $f : S \to \mathbb{R}$, S an open convex subset of $\mathbb{R}^n$. Assume f is $\mathcal{C}^2$.*

If f is quasi-concave, then for all x, $(-1)^k B_{k+1}(x) \geq 0$ for all $k = 1, ..., n$.

Proof. We *admit* the result. □

Just as for concavity, the converse is false, but by requiring strict inequalities, we have some form of a converse:

Proposition 4.5. *Let $f : S \to \mathbb{R}$, S an open convex subset of $\mathbb{R}^n$. Assume f is $\mathcal{C}^2$.*

If for all x, $(-1)^k B_{k+1}(x) > 0$ for all $k = 1, ..., n$, then f is strictly quasi-concave.

Proof. We *admit* the result. □

5 Homogeneous real-valued functions

5.1 Definition

Definition 5.1. *A function $\mathbb{R}^n \to \mathbb{R}^m$ is **homogeneous of degree** k (an integer) iff:*

$$\forall x \in \mathbb{R}^n, \forall \lambda > 0, f(\lambda x) = \lambda^k f(x)$$

When $n = m = 1$, a homogeneous function is simply a power function $x \mapsto Ax^k$, for $x > 0$ and a power function $x \mapsto Bx^k$ for $x < 0$ (the constant A and B may differ). In higher dimensions, consider a half-vector line $\{\lambda x, \lambda > 0\}$, called a **ray through the origin**: homogeneity puts a lot of structure on the values a function can take along the same ray through the origin.

5.2 Two properties if f is differentiable

Proposition 5.1. *Let f be a differentiable function from $\mathbb{R}^n$ to $\mathbb{R}$. If f is homogeneous of degree k, then its derivative is homogeneous of degree $k-1$ (and therefore so are its derivatives along any vectors, in particular all its partial derivatives).*

Proof. Differentiate both side of $f(\lambda x) = \lambda^k f(x)$ with respect to x. □

Taking the transpose, the result can be stated with the gradient: for all $\lambda > 0$, $\nabla f(\lambda x) = \lambda^{k-1} \nabla f(x)$. All the gradients of f at points that belong to the same ray through the origin are proportional. This has a graphical consequence. Consider the **level set of level** c of f, defined as all the points such that $f(x) = c$, $\{x/f(x) = c\}$. The gradient of f at x_0 is orthogonal to the hyperplane tangent to the level set of level $f(x_0)$ at x_0. Hence, if f is homogeneous, the hyperplanes tangent to level sets are parallel along each ray through the origin.

Proposition 5.2. ***Euler's homogeneous function theorem***

Let f be a differentiable function $\mathbb{R}^n \to \mathbb{R}$.

f is homogeneous of degree k iff it satisfies the following identity:

$$\forall x, f'(x)x = \sum_{i=1}^{n} f'_i(x)x_i = kf(x)$$

Proof. Assume f is homogeneous of degree k. Differentiate both sides of $f(\lambda x) = \lambda^k f(x)$ with respect to λ this time. We get: $f'(\lambda x)x = k\lambda^{k-1} f(x)$. Choosing $\lambda = 1$ yields the result.

Conversely, assume f satisfies the identity $f'(x)x = kf(x)$ for all x, and define on $\mathbb{R}_+^*$, $g(\lambda) = \frac{f(\lambda x)-\lambda^k f(x)}{\lambda^k} = \lambda^{-k} f(\lambda x) - f(x)$. We want to show that g is constant to zero. Differentiate g:

$$\begin{aligned} g'(\lambda) &= (-k)\lambda^{-k-1} f(\lambda x) + \lambda^{-k} f'(\lambda x)x \\ &= \lambda^{-k-1}\Big(-kf(\lambda x) + f'(\lambda x)(\lambda x)\Big) \\ &= 0 \end{aligned}$$

Hence g is constant. But $g(1) = f(x) - f(x) = 0$, so g is constant to 0. QED. □

5.3 Economic applications

5.3.1 In producer theory

In producer theory, a **production function** f tells how much of output $f(x)$ one can get with a quantity $x = (x_1, ..., x_n)'$ of inputs. For production functions, homogeneity of degree 1 is called **constant returns to scale**: a production function exhibits constant returns to scale iff when multiplying all inputs by $\lambda > 0$, the output is multiplied by λ.

Euler's theorem has a much meaningful interpretation in producer's theory: it states that a firm with constant returns to scale makes zero profits under perfect competition (meaning that the firm takes prices as given). The profits of the firm are its revenues minus its costs. If p is the price of the good it sells, and w_i the price of input i, then the profits are $\Pi = pf(x) - \sum_{i=1}^{n} w_i x_i$. If the firm maximizes profits, then (anticipating on unconstrained optimization), for all i, $pf_i'(x) = w_i$. Injecting w_i back into the expression of profits, $\Pi = p(f(x) - \sum_{i=1}^{n} f_i'(x)x_i)$. If f exhibits constant returns to scale, this is zero.

5.3.2 In consumer theory

In consumer theory, a consumer has **homothetic preferences** is its preferences can be represented by a homogeneous **utility function**. Because if a utility function U represents the agent's preferences, then any utility function $g \circ U$ with g strictly increasing represents the same preferences, we call a **homothetic function** a function that can be written $g \circ U$, with g strictly increasing and U homogeneous.