Lecture Notes - Convexity

MA Math Camp 2022
Columbia University

Andrea Ciccarone*

This Version : August 28, 2022

Contents

Throughout this lecture, the vector spaces are real vector spaces, unless stated otherwise.

1 Convex Sets

Definition 1.1. *In real vector space V, a set $S \subset V$ is a* ***convex set*** *iff*

$$\lambda x + (1 - \lambda)\, y \in S$$

for any $\lambda \in [0, 1]$ and $x, y \in S$.

Notice that it makes sense to talk about convex sets only in a vector space, since we need to be able to perform vector addition and scalar multiplication. In most applications, the vector space is $\mathbb{R}^n$.

For finitely many vectors $x_1, x_2, \ldots, x_n$ in vector space V, a **convex combination** of $x_1, x_2, \ldots, x_n$ is a vector $\sum_{i=1}^{n} \lambda_i x_i$ for scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}_+$ with $\sum_{i=1}^{n} \lambda_i = 1$. Different from linear combination, convex combination requires that the coefficient λ_i's are nonnegative and that they sum up to 1.

By definition, a set S is convex iff any combination of *two* vectors in S is still in S. However, the next result says that S is convex iff any combination of *finitely many* vectors in S is still in S.

*The present lecture notes were largely based on math camp materials from César Barilla, Palaash Bhargava, Paul Koh, and Xuan Li. All errors in this document are mine. If you find a typo or an error, please send me an email at cesar.barilla@columbia.edu.

Claim 1.2. *In vector space V, the set $S \subset V$ is convex iff any convex combination of $x_1, x_2, \dots, x_n \in S$ is also in S.*

Proof. $\Leftarrow$ is trivial.

$\Rightarrow$:

If $n = 1$, the statement is trivial.

If $n = 2$, then $\lambda_1 x_1 + \lambda_2 x_2 \in S$ by definition of convexity.

Suppose when $n = k$, we have $\sum_{i=1}^{k} \lambda_i x_i \in S$ for any $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}_+$ s.t. $\sum_{i=1}^{k} \lambda_i = 1$. Consider $n = k + 1$. We have

$$\begin{aligned}\sum_{i=1}^{k+1} \lambda_i x_i &= \sum_{i=1}^{k} \lambda_i x_i + \lambda_{k+1} x_{k+1} \\ &= \left(\sum_{j=1}^{k} \lambda_j \right) \left(\sum_{i=1}^{k} \frac{\lambda_i}{\sum_{j=1}^{k} \lambda_j} x_i \right) + \lambda_{k+1} x_{k+1}\end{aligned}$$

Because $\sum_{i=1}^{k} \frac{\lambda_i}{\sum_{j=1}^{k} \lambda_j} = 1$, we have $\sum_{i=1}^{k} \frac{\lambda_i}{\sum_{j=1}^{k} \lambda_j} x_i \in S$ by the induction hypothesis. Because $\left(\sum_{j=1}^{k} \lambda_j \right) + \lambda_{k+1} = 1$, we know that $\sum_{i=1}^{k+1} \lambda_i x_i$ is a convex combination of $\frac{\lambda_i}{\sum_{j=1}^{k} \lambda_j} x_i$ and x_{k+1}, which are both in S. Therefore, $\sum_{i=1}^{k+1} \lambda_i x_i \in S$.

So when $n = k + 1$, we also have $\sum_{i=1}^{n} \lambda_i x_i \in S$ for any $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}_+$ s.t. $\sum_{i=1}^{n} \lambda_i = 1$.

By induction, for any $n \in \mathbb{N}$, we have $\sum_{i=1}^{n} \lambda_i x_i \in S$ for any $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}_+$ s.t. $\sum_{i=1}^{n} \lambda_i = 1$. $\square$

The next proposition states that any arbitrary intersection of convex sets is still convex.

Proposition 1.3. *In vector space V, let $\{S_\alpha\}_{\alpha \in A}$ be a family of convex sets. Then $\bigcap_{\alpha \in A} S_\alpha$ is also convex.*

The proof is straightforward and follows from the definition of convexity.

Definition 1.4. *In vector space V, the **convex hull** of set $S \subset V$ is*

$$Co(S) := \bigcap_{C \in \{X \subset V:\ X \text{ is convex and } X \supset S\}} C$$

Because intersection of convex sets is still convex, we know that $Co(S)$ is convex, and therefore the convex hull can be also interpreted as the smallest convex set that covers S.

The next result says that for a finite set S, the convex hull of S is simply the set of all convex combinations of vectors in S.

Claim 1.5. *In vector space V, let $\{x_1, x_2, \dots, x_n\}$ be a finite set of vectors. Then*

$$\begin{aligned}&Co(\{x_1, x_2, \dots, x_n\}) \\ &= \left\{ \sum_{i=1}^{n} \lambda_i x_i : \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}_+, \text{ and } \sum_{i=1}^{n} \lambda_i = 1 \right\}\end{aligned}$$

Proof. $\subset$: It is sufficient to show that the set on the right-hand side

$$\left\{\sum_{i=1}^{n} \lambda_i x_i : \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}_+, \text{ and } \sum_{i=1}^{n} \lambda_i = 1\right\} =: A$$

covers $\{x_1, \ldots, x_n\}$ and is convex.

Clearly, A covers $\{x_1, \ldots, x_n\}$. To show convexity, take any two vectors $\sum_{i=1}^{n} \lambda_i x_i$ and $\sum_{i=1}^{n} \mu_i x_i$ in A, and consider any $\alpha \in [0, 1]$, we have

$$\alpha \sum_{i=1}^{n} \lambda_i x_i + (1-\alpha) \sum_{i=1}^{n} \mu_i x_i = \sum_{i=1}^{n} \left[\alpha \lambda_i + (1-\alpha)\mu_i\right] x_i$$

Clearly, each coefficient $\alpha\lambda_i + (1-\alpha)\mu_i \geq 0$, and their sum

$$\begin{aligned}\sum_{i=1}^{n} \left[\alpha \lambda_i + (1-\alpha)\mu_i\right] &= \sum_{i=1}^{n} \alpha\lambda_i + \sum_{i=1}^{n} (1-\alpha)\mu_i \\ &= \alpha \sum_{i=1}^{n} \lambda_i + (1-\alpha)\sum_{i=1}^{n} \mu_i \\ &= \alpha + (1-\alpha) = 1\end{aligned}$$

and so $\alpha \sum_{i=1}^{n} \lambda_i x_i + (1-\alpha)\sum_{i=1}^{n} \mu_i x_i \in A$.

$\supset$:

Take any $x = \sum_{i=1}^{n} \lambda_i x_i \in A$. WTS $x \in Co\left(\{x_1, \ldots, x_n\}\right)$.

It is sufficient to show that $\sum_{i=1}^{n} \lambda_i x_i \in C$ for any $C \subset V$ s.t. C is convex and $C \supset \{x_1, \ldots, x_n\}$. This is true because of Claim 1.2. $\square$

1.1 Separating Hyperplane Theorem

In $\mathbb{R}^n$, a **hyperplane** is defined as

$$H(p, c) := \{x \in \mathbb{R}^n : p \cdot x = c\}$$

where $p \in \mathbb{R}^n \setminus \{0\}$, $c \in \mathbb{R}$, and $\cdot$ is the dot product. A hyperplane $H(p, c)$ cuts the whole space $\mathbb{R}^n$ into halves. This is a generalization of a line in $\mathbb{R}^2$ and a plane in $\mathbb{R}^3$.

Theorem 1.6 (Minkowski's Separating Hyperplane)**.** *Let S_1 and S_2 be two disjoint nonempty and convex sets in $\mathbb{R}^n$. Then there exist $p \in \mathbb{R}^n \setminus \{0\}$ and $c \in \mathbb{R}$ s.t. $p \cdot x \geq c$ for any $x \in S_1$ and $p \cdot x \leq c$ for any $x \in S_2$.*

See FMEA Section 13.6 for a proof. Minkowski's separating hyperplane theorem states that for any two disjoint nonempty and convex sets in $\mathbb{R}^n$, we can find a hyperplane $H(p, c)$ that weakly separates them, i.e. one of the two sets is contained in $H_+(p, c) := \{x \in \mathbb{R}^n : p \cdot x \geq c\}$, and the other is contained in $H_-(p, c) := \{x \in \mathbb{R}^n : p \cdot x \leq c\}$.

Minkowski's Separating Hyperplane is used in the proof of Second Welfare Theorem.

1.2 Brouwer's Fixed Point Theorem

Theorem 1.7 (Brouwer's Fixed Point)**.** *Let X be a nonempty, compact, and convex set in $\mathbb{R}^n$, and consider a continuous function $f : X \to X$. Then there exists $x^* \in X$ s.t. $f(x^*) = x^*$.*

The theorem states that a continuous self-map defined on a nonempty, compact, and convex set in $\mathbb{R}^n$ must have a fixed point. Let's admit this result without proof. We will introduce its generalization, Kakutani's fixed point theorem, later when we discuss correspondences.

Brouwer's fixed point theorem and Kakutani's fixed point theorem play an important role in the existence of Walrasian equilibria in the general equilibrium theory and the existence of Nash equilibria in non-cooperative game theory.

2 Convex and Concave Functions

Definition 2.1. *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*

*(1) The function f is a **convex** function iff*

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

for any $x, y \in S$ and $\lambda \in [0,1]$.

*(2) The function f is a **concave** function iff*

$$f\big(\lambda x + (1-\lambda) y\big) \geq \lambda f(x) + (1-\lambda) f(y)$$

for any $x, y \in S$ and $\lambda \in [0,1]$.

*(3) The function f is a **strictly convex** function iff*

$$f\big(\lambda x + (1-\lambda) y\big) < \lambda f(x) + (1-\lambda) f(y)$$

for any $x, y \in S$ with $x \neq y$ and $\lambda \in (0,1)$.

*(4) The function f is a **strictly concave** function iff*

$$f\big(\lambda x + (1-\lambda) y\big) > \lambda f(x) + (1-\lambda) f(y)$$

for any $x, y \in S$ with $x \neq y$ and $\lambda \in (0,1)$.

In the definition above, we only take convex combination of two points in the domain. However, we can also take finitely many points in the domain and still have the function value at a convex combination weakly less/greater than the convex combination of the function values for a convex/concave function. This is known as Jensen's inequality.

Theorem 2.2 (Jensen's Inequality)**.** *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*

(1) f is convex iff

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)$$

for any $x_1, x_2, \ldots, x_n \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}_+$ with $\sum_{i=1}^{n} \lambda_i = 1$.

(2) f is concave iff

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq \sum_{i=1}^{n} \lambda_i f(x_i)$$

for any $x_1, x_2, \ldots, x_n \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}_+$ with $\sum_{i=1}^{n} \lambda_i = 1$.

The proof is by induction on n, similar to the proof of Claim 1.2. This is left as an exercise.

For two vector spaces V and W, a function $f : V \to W$ is **linear** iff $f(x_1 + x_2) = f(x_1) + f(x_2)$ and $f(\lambda x) = \lambda f(x)$. Clearly, a linear function $f : V \to \mathbb{R}$, where V is a vector space, is both convex and concave, but not strictly convex or strictly concave.

Also, notice that f is (strictly) convex iff $-f$ is (strictly) concave. The proof is left as an exercise.

Consider a function $f : S \to \mathbb{R}$. Define the **graph** of f as

$$G(f) := \big\{(x, y) \in S \times \mathbb{R} : y = f(x)\big\}$$

Notice that this is in fact a redundant notation, because $G(f)$ is exactly f. Recall that a relation from S to $\mathbb{R}$ is a subset of $S \times \mathbb{R}$.

Define the **epigraph** of f as

$$G^+(f) := \big\{(x, y) \in S \times \mathbb{R} : y \geq f(x)\big\}$$

and the **subgraph** of f as

$$G^-(f) := \big\{(x, y) \in S \times \mathbb{R} : y \leq f(x)\big\}$$

The next result characterizes a convex/concave function using its epigraph/subgraph.

Proposition 2.3. *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*

(1) f is convex iff its epigraph $G^+(f)$ is a convex set in $V \times \mathbb{R}$.

(2) f is concave iff its subgraph $G^-(f)$ is a convex set in $V \times \mathbb{R}$.

Notice that $V \times \mathbb{R}$ is a vector space, since both V and $\mathbb{R}$ are vector spaces. So it makes sense to talk about convex sets in $V \times \mathbb{R}$. The vector addition and scalar multiplication in $V \times \mathbb{R}$ is defined in a component-by-component fashion. That is, we define $(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$ for any $(x_1, y_1), (x_2, y_2) \in V \times \mathbb{R}$, where $x_1 + x_2$ is vector addition in V, and $y_1 + y_2$ is (vector) addition in $\mathbb{R}$. Also, we define $\lambda(x, y) := (\lambda x, \lambda y)$ for any $(x, y) \in V \times \mathbb{R}$, where λx is the scalar multiplication in V, and λy is the (scalar) multiplication in $\mathbb{R}$.

Proof. (1)

$\Rightarrow$:

Take any $(x_1, y_1), (x_2, y_2) \in G^+(f)$ and any $\lambda \in [0, 1]$. WTS: $\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \in G^+(f)$.

Because $(x_1, y_1), (x_2, y_2) \in G^+(f)$, we have $y_1 \geq f(x_1)$ and $y_2 \geq f(x_2)$. Therefore,

$$f\big(\lambda x_1 + (1 - \lambda) x_2\big) \leq \lambda f(x_1) + (1 - \lambda) f(x_2) \leq \lambda y_1 + (1 - \lambda) y_2$$

where the first inequality is by convexity of f. Therefore,

$$\begin{aligned} &\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \\ &= \big(\lambda x_1 + (1 - \lambda) x_2, \lambda y_1 + (1 - \lambda) y_2\big) \in G^+(f) \end{aligned}$$

$\Leftarrow$:

Take any $x_1, x_2 \in S$ and any $\lambda \in [0, 1]$. WTS:

$$f\big(\lambda x_1 + (1 - \lambda) x_2\big) \leq \lambda f(x_1) + (1 - \lambda) f(x_2)$$

Because $\big(x_1, f(x_1)\big), \big(x_2, f(x_2)\big) \in G^+(f)$, and $G^+(f)$ is convex, we have

$$\lambda\big(x_1, f(x_1)\big) + (1 - \lambda)\big(x_2, f(x_2)\big) \in G^+(f)$$

i.e. $\left(\lambda x_1 + (1-\lambda) x_2, \lambda f(x_1) + (1-\lambda) f(x_2)\right) \in G^+(f)$, which by definition implies

$$f\left(\lambda x_1 + (1-\lambda) x_2\right) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

(2) can be shown symmetrically, which is left as an exercise. $\square$

In the proof above, we can also show (2) using the fact that f is concave iff $-f$ is convex, and the subgraph of f can be mapped to the epigraph of $-f$, in the sense that $(x,y) \in G^-(f)$ iff $(x,-y) \in G^+(-f)$.

The next result states that addition and multiplication by a nonnegative real number preserve convexity/concavity of a function.

Proposition 2.4. *Consider two functions f and g from S to $\mathbb{R}$, where S is a convex set in vector space V. If f and g are both convex/concave functions, then*
(1) $f+g$ is a convex/concave function, and
(2) cf is a convex/concave function, for any $c \in \mathbb{R}_+$.

The proof is straightforward.

The next result says: (1) a weakly increasing convex transformation of a convex function is still convex, and (2) a weakly increasing concave transformation of a concave function is still concave.

Proposition 2.5. *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*
(1) If f is convex and $\phi : \mathbb{R} \to \mathbb{R}$ is weakly increasing and convex, then $\phi \circ f$ is convex.
(2) If f is concave and $\phi : \mathbb{R} \to \mathbb{R}$ is weakly increasing and concave, then $\phi \circ f$ is concave.

The proof is left as an exercise.

A straightforward corollary of the proposition above is: (3) a weakly decreasing concave transformation of a convex function is concave, and (4) a weakly decreasing convex transformation of a concave function is convex. To see (3), suppose f is convex and $\phi : \mathbb{R} \to \mathbb{R}$ is weakly decreasing and concave, then $-\phi$ is weakly increasing and convex. Applying (1), we know that $(-\phi) \circ f$ is convex. Therefore, $\phi \circ f = -(-\phi) \circ f$ is concave. (4) can be shown symmetrically.

Now let's consider a family $\{f_\alpha\}_{\alpha \in A}$ of real-valued functions defined on the same domain S. If the set $\left\{f_\alpha(x) : \alpha \in A\right\}$ is bounded from above for each $x \in S$, we can define the function $\sup \{f_\alpha\}_{\alpha \in A} : S \to \mathbb{R}$ as the pointwise supremum, i.e. for each $x \in S$,

$$\left(\sup \{f_\alpha\}_{\alpha \in A}\right)(x) := \sup \left\{f_\alpha(x) : \alpha \in A\right\}$$

Similarly, if the set $\left\{f_\alpha(x) : \alpha \in A\right\}$ is bounded from below for each $x \in S$, we can define the function $\inf \{f_\alpha\}_{\alpha \in A} : S \to \mathbb{R}$ as the pointwise infimum, i.e. for each $x \in S$,

$$\left(\inf \{f_\alpha\}_{\alpha \in A}\right)(x) := \inf \left\{f_\alpha(x) : \alpha \in A\right\}$$

Now let's state the following result.

Proposition 2.6. *Consider a finite family of functions $\{f_\alpha\}_{\alpha \in A}$ from S to $\mathbb{R}$, where S is a convex set in vector space V.*
(1) If all functions in the family are convex, and the set $\left\{f_\alpha(x) : \alpha \in A\right\}$ is bounded from above for each $x \in S$, then $\sup \{f_\alpha\}_{\alpha \in A}$ is a convex function.
(2) If all functions in the family are concave, and the set $\left\{f_\alpha(x) : \alpha \in A\right\}$ is bounded from below for each $x \in S$, then $\inf \{f_\alpha\}_{\alpha \in A}$ is a concave function.

Shortly put, the proposition states that the sup of convex functions is still convex, and the inf of concave functions is still concave.

Proof. (1) First, I claim that

$$G^{+}\left(\sup\{f_\alpha\}_{\alpha\in A}\right)=\bigcap_{\alpha\in A}G^{+}(f_\alpha)$$

$\subset$:

Take any $(x,y)\in G^{+}\left(\sup\{f_\alpha\}_{\alpha\in A}\right)$. We have $y\geq\left(\sup\{f_\alpha\}_{\alpha\in A}\right)(x)$, and so $y\geq f_\alpha(x)$ for any $\alpha\in A$. Therefore, $(x,y)\in G^{+}(f_\alpha)$ for any $\alpha\in A$, and so $(x,y)\in\bigcap_{\alpha\in A}G^{+}(f_\alpha)$.

$\supset$:

Take any $(x,y)\in\bigcap_{\alpha\in A}G^{+}(f_\alpha)$. We have $(x,y)\in G^{+}(f_\alpha)$ for any $\alpha\in A$, i.e. $y\geq f_\alpha(x)$ for any $\alpha\in A$. Therefore, y is an upper bound of $\left\{f_\alpha(x):\alpha\in A\right\}$, and so $y\geq\sup\left\{f_\alpha(x):\alpha\in A\right\}=\left(\sup\{f_\alpha\}_{\alpha\in A}\right)(x)$. Therefore $(x,y)\in G^{+}\left(\sup\{f_\alpha\}_{\alpha\in A}\right)$.

Then for each $\alpha\in A$, because f_α is a convex function, its epigraph $G^{+}(f_\alpha)$ is a convex set. So $G^{+}\left(\sup\{f_\alpha\}_{\alpha\in A}\right)=\bigcap_{\alpha\in A}G^{+}(f_\alpha)$ is also convex, and therefore $\sup\{f_\alpha\}$ is a convex function.

(2) can be proved symmetrically. □

The proposition implies as a special case that $\max\{f,g\}$ is a convex function if f and g are both convex, and $\min\{f,g\}$ is a concave function if f and g are both concave.

Because linear functions are both convex and concave, another important special case of the proposition above is that the sup of linear functions is convex, and the inf of linear functions is concave.

The next theorem provides a characterization of convexity/concavity of continuously differentiable functions.

Theorem 2.7. *Suppose the function $f:S\to\mathbb{R}$ is a C^1 function, where S is a convex and open set in $\mathbb{R}^n$.*

(1) f is convex iff

$$f\left(x'\right)\geq f(x)+\nabla f(x)\cdot\left(x'-x\right)$$

for any $x',x\in S$.

(2) f is concave iff

$$f\left(x'\right)\leq f(x)+\nabla f(x)\cdot\left(x'-x\right)$$

for any $x',x\in S$.

(3) f is strictly convex iff

$$f\left(x'\right)> f(x)+\nabla f(x)\cdot\left(x'-x\right)$$

for any $x',x\in S$ with $x'\neq x$.

(4) f is strictly concave iff

$$f\left(x'\right)< f(x)+\nabla f(x)\cdot\left(x'-x\right)$$

for any $x',x\in S$ with $x'\neq x$.

The intuition of the theorem above is that the set

$$\left\{\left(x',y\right)\in S\times\mathbb{R}:y=f(x)+\nabla f(x)\cdot\left(x'-x\right)\right\}$$

is the hyperplane that is tangent to the graph of f at x. A convex/concave function should lie above/below this tangent plane. See FMEA Theorem 2.4.1 for a proof.

The next theorem provides a characterization of convexity/concavity for twice continuously differentiable functions using the Hessian matrix.

Theorem 2.8. *Suppose the function $f : S \to \mathbb{R}$ is a C^2 function, where S is a convex and open set in $\mathbb{R}^n$.*

(1) f is convex iff its Hessian matrix $H(x)$ is positive semi-definite for any $x \in S$.

(2) f is concave iff its Hessian matrix $H(x)$ is negative semi-definite for any $x \in S$.

(3) f is strictly convex if its Hessian matrix $H(x)$ is positive definite for any $x \in S$.

(4) f is strictly concave if its Hessian matrix $H(x)$ is negative definite for any $x \in S$.

Let's admit this result without a formal proof. However, we can obtain some intuition of this theorem using Taylor's expansion:

$$f(x+h) = f(x) + \nabla f(x) h + \frac{1}{2} h^T H_f(x) h + o\left(\|h\|^2\right)$$

When h is close to 0, the error term $o\left(\|h\|^2\right)$ is small compared to $\frac{1}{2} h^T H_f(x) h$, and therefore the sign of $f(x+h) - \left(f(x) + \nabla f(x) h\right)$ is determined by the sign of $h^T H_f(x) h$. If $H_f(x)$ is positive/negative definite, then $h^T H_f(x) h$ is strictly positive/negative for any $h \neq 0$, and therefore f is convex/concave.

In the theorem above, notice that (3) and (4) only claim the "if" direction is true. In fact, the "only if" direction is not true. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = x^4$, which is a strictly convex function. However, its Hessian $H(x) = f''(x) = 12x^2$ is not positive definite when $x = 0$.

3 Quasi-convex and Quasi-concave Functions

Definition 3.1. *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*

*(1) The function f is a **quasi-convex** function iff*

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

for any $x, y \in S$ and $\lambda \in [0,1]$.

*(2) The function f is a **quasi-concave** function iff*

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

for any $x, y \in S$ and $\lambda \in [0,1]$.

*(3) The function f is a **strictly quasi-convex** function iff*

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

for any $x, y \in S$ with $x \neq y$ and $\lambda \in (0,1)$.

*(4) The function f is a **strictly quasi-concave** function iff*

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

for any $x, y \in S$ with $x \neq y$ and $\lambda \in (0,1)$.

Compare the definition of quasi-convex functions with that of convex functions in the previous section, clearly a convex function f is also quasi-convex, since

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

Similarly, concavity implies quasi-concavity, strict convexity implies strict quasi-convexity, and strict concavity implies strict quasi-concavity.

Clearly, f is (strictly) quasi-convex iff $-f$ is (strictly) quasi-concave.

For a function $f : S \to \mathbb{R}$, define the **upper contour set** of f with cutoff a as

$$C^{+}(f, a) := \big\{x \in S : f(x) \geq a\big\}$$

and the **lower contour set** of f with cutoff a as

$$C^{-}(f, a) := \big\{x \in S : f(x) \leq a\big\}$$

The next result characterizes a quasi-concave/quasi-convex function using its upper/lower contour set.

Proposition 3.2. *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*

(1) f is quasi-concave iff its upper contour set $C^{+}(f, a)$ is a convex set in V for any $a \in \mathbb{R}$.

(2) f is quasi-convex iff its lower contour set $C^{-}(f, a)$ is a convex set in V for any $a \in \mathbb{R}$.

Notice that the concept of upper/lower contour set is completely different from that of epigraph/subgraph. The upper/lower contour set is in the vector space V, but the epigraph/subgraph is in the vector space $V \times \mathbb{R}$. Using this characterization, it is not difficult to see that quasi-concavity is essentially a single peak condition, and that quasi-convexity is essentially a single trough condition.

Proof. (1)

$\Rightarrow$:

Take any $a \in \mathbb{R}$. WTS: $C^{+}(f, a)$ is convex.

Take any $x_1, x_2 \in C^{+}(f, a)$ and any $\lambda \in [0, 1]$. WTS: $\lambda x_1 + (1 - \lambda) x_2 \in C^{+}(f, a)$.

By definition of $C^{+}(f, a)$, we have $f(x_1) \geq a$ and $f(x_2) \geq a$. Because f is quasi-concave, we have

$$f\big(\lambda x_1 + (1 - \lambda) x_2\big) \geq \min\big\{f(x_1), f(x_2)\big\} \geq a$$

and so $\lambda x_1 + (1 - \lambda) x_2 \in C^{+}(f, a)$.

$\Leftarrow$:

Take any $x_1, , x_2 \in S$ and any $\lambda \in [0, 1]$. WTS: $f\big(\lambda x_1 + (1 - \lambda) x_2\big) \geq \min\big\{f(x_1), f(x_2)\big\}$.

Because $f(x_1) \geq \min\big\{f(x_1), f(x_2)\big\}$ and $f(x_2) \geq \min\big\{f(x_1), f(x_2)\big\}$, we have $x_1, x_2 \in C^{+}\Big(f, \min\big\{f(x_1), f(x_2)\big\}\Big)$. Because $C^{+}\Big(f, \min\big\{f(x_1), f(x_2)\big\}\Big)$ is convex, we have $\lambda x_1 + (1 - \lambda) x_2 \in C^{+}\Big(f, \min\big\{f(x_1), f(x_2)\big\}\Big)$, which by definition implies

$$f\big(\lambda x_1 + (1 - \lambda) x_2\big) \geq \min\big\{f(x_1), f(x_2)\big\}$$

(2) can be proved symmetrically. □

In the proof above, we can also show (2) using the fact that f is quasi-convex iff $-f$ is quasi-concave, and the lower contour set $C^{-}(f, a)$ is the same as the upper contour set $C^{+}(-f, -a)$. This alternative proof is left as an exercise.

The next result states that a weakly increasing transformation of a quasi-convex/quasi-concave function is still quasi-convex/quasi-concave.

Proposition 3.3. *Consider a function $f : S \to \mathbb{R}$, where S is a convex set in vector space V.*

(1) If f is quasi-convex and $\phi : \mathbb{R} \to \mathbb{R}$ is weakly increasing, then $\phi \circ f$ is quasi-convex.

(2) If f is quasi-concave and $\phi : \mathbb{R} \to \mathbb{R}$ is weakly increasing, then $\phi \circ f$ is quasi-concave.

Proof. (1) Take any $x_1, , x_2 \in S$ and any $\lambda \in [0,1]$. WTS:

$$\phi\Big(f\left(\lambda x_1 + (1-\lambda)\,x_2\right)\Big) \geq \min\Big\{\phi\left(f\left(x_1\right)\right), \phi\left(f\left(x_2\right)\right)\Big\}$$

To see this,

$$\begin{aligned}\phi\Big(f\left(\lambda x_1 + (1-\lambda)\,x_2\right)\Big) &\geq \phi\Big(\min\left\{f\left(x_1\right), f\left(x_2\right)\right\}\Big)\\ &= \min\Big\{\phi\left(f\left(x_1\right)\right), \phi\left(f\left(x_2\right)\right)\Big\}\end{aligned}$$

(2) can be shown symmetrically. □

A straightforward corollary of the proposition above is: (3) a weakly decreasing transformation of a quasi-convex function is quasi-concave, and (4) a weakly decreasing transformation of a quasi-concave function is quasi-convex. To see (3), suppose f is quasi-convex and ϕ is weakly decreasing, then $-\phi$ is weakly increasing. By (1), we have $(-\phi) \circ f$ is quasi-convex, and so $\phi \circ f = -(-\phi) \circ f$ is quasi-concave. (4) can be shown symmetrically.

Although quasi-convexity/quasi-concavity is preserved under increasing transformations, the sum of two quasi-convex/quasi-concave functions may no longer be quasi-convex/quasi-concave.