

Convexity and Quasiconvexity

Convex Combinations and Convex Sets.

Definition. Given any finite collection of points $x_1, \dots, x_m \in \mathbb{R}^n$, a point $z \in \mathbb{R}^n$ is said to be a *convex combination* of the points $\{x_1, \dots, x_m\}$ if there is some $\lambda \in \mathbb{R}^m$ satisfying

1. $\lambda_i \geq 0, \quad i = 1, \dots, m,$ and
2. $\sum_{i=1}^m \lambda_i = 1,$

such that $z = \sum_{i=1}^m \lambda_i x_i$. A subset \mathcal{D} of \mathbb{R}^n is *convex* if the convex combination of any two points in \mathcal{D} is also in \mathcal{D} . ▲

- Thus a set is convex if the straight line joining any two points in \mathcal{D} is completely contained in \mathcal{D} i.e. if for all x and y in \mathcal{D} and $\lambda \in (0, 1)$ it is the case that $\lambda x + (1 - \lambda)y$ is a subset of \mathcal{D} .

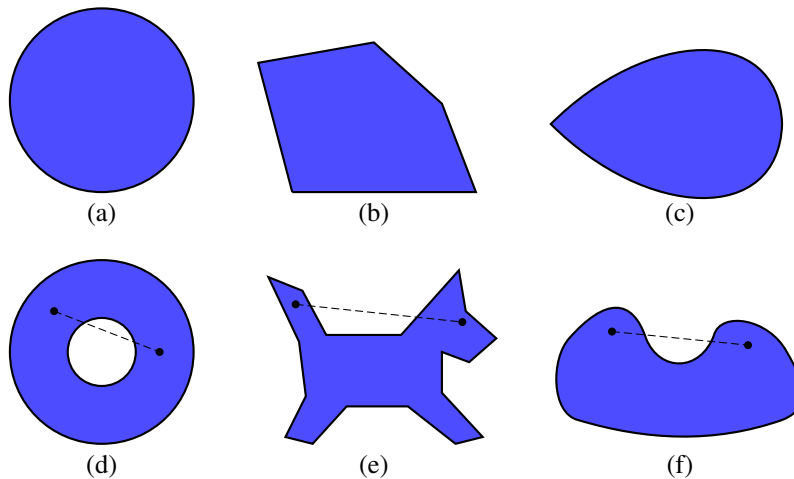


Figure 1: The sets represented by (a), (b) and (c) are convex, while (d), (e) and (f) illustrate nonconvex sets.

Concave and Convex Functions.

Definition. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function.

- The *subgraph* of f , denoted $\text{sub } f$, is the set

$$\text{sub } f = \{(x, y) \in \mathcal{D} \times \mathbb{R} \mid f(x) \geq y\}.$$

- The *epigraph* of f , denoted $\text{epi } f$, is the set

$$\text{epi } f = \{(x, y) \in \mathcal{D} \times \mathbb{R} \mid f(x) \leq y\}$$

▲

- The subgraph of a function is the area lying below the graph of a function.
- On the other hand, the epigraph of a function is the area lying above the graph of the function.

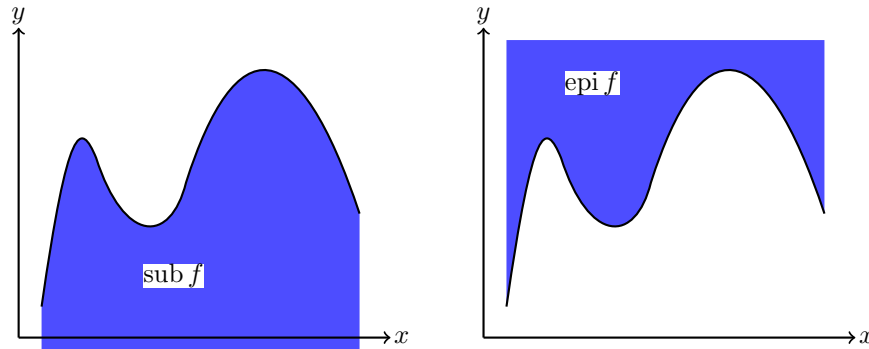


Figure 2: The subgraph and epigraph of f .

Definition. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function.

- We say that f is *concave* on \mathcal{D} if $\text{sub } f$ is a convex set.
- We say that f is *convex* on \mathcal{D} if $\text{epi } f$ is a convex set. ▲
- Note concave and convex functions are required to have convex domains.
- The following theorem provides an alternative definition of concave and convex functions.

Theorem 1. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function. Then

1. f is concave iff for all $x, y \in \mathcal{D}$ and $\lambda \in (0, 1)$, we have

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

2. f is convex iff for all $x, y \in \mathcal{D}$ and $\lambda \in (0, 1)$, we have

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

- So a function is concave iff the function's value at a convex combination of any two points is at least as great as the same convex combination of the function's values at each point.

Definition. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function.

- We say f is *strictly concave* if for all $x, y \in \mathcal{D}$ with $x \neq y$, and all $\lambda \in (0, 1)$, we have

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

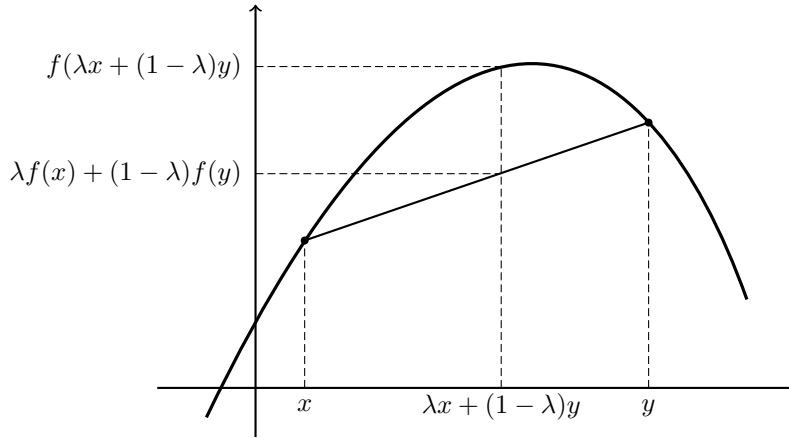


Figure 3: A function f is concave iff the secant line connecting any two points on the graph of f lies below the graph.

- We say f is *strictly convex* if for all $x, y \in \mathcal{D}$ with $x \neq y$, and all $\lambda \in (0, 1)$, we have

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y). \quad \blacktriangle$$

Theorem 2. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function. Then

1. f is concave iff the function $-f$ is convex.
 2. f is strictly concave iff the function $-f$ is strictly convex.
- The previous result allows us to easily apply all results about concave functions to convex functions
 - Another valuable property of concave functions is that they behave well under addition and scalar multiplication by positive numbers.

Theorem 3. Let \mathcal{D} be a convex subset of \mathbb{R}^n . Let $f_i : \mathcal{D} \rightarrow \mathbb{R}$ be concave functions and let a_i be positive numbers $i = 1, \dots, k$. Then

$$a_1 f_1 + \dots + a_k f_k$$

is a concave function.

Proof. Simply apply the definition of a concave function. ■

- An identical result holds for convex functions.
- The assumption of convexity has two important implications.
- First, every concave function must also be continuous except possibly at the boundary points.

- Second, every concave function is differentiable “almost everywhere”.

Theorem 4. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a concave or convex function. Then

1. If \mathcal{D} is open, f is continuous on \mathcal{D} .
 2. If \mathcal{D} is not open, f is continuous on $\text{int } \mathcal{D}$.
 3. If \mathcal{D} is open, f is differentiable “almost everywhere” on \mathcal{D} and the derivative Df of f is continuous at all points where it exists.
- For a discussion of the meaning of “almost everywhere” see Sundaram pp182-183.

Convexity and the Properties of the Derivative.

- We can characterize the concavity or convexity of a differentiable function using the first derivative.

Theorem 5. Let \mathcal{D} be an open convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a C^1 function. Then

1. f is concave iff $Df(x)(y - x) \geq f(y) - f(x)$ for all $x, y \in \mathcal{D}$.
 2. f is convex iff $Df(x)(y - x) \leq f(y) - f(x)$ for all $x, y \in \mathcal{D}$.
- Note that, if we let $z = y - x$, we can rewrite (1) to say f is concave iff $Df(x)z + f(x) \geq f(x + z)$ for all $x, z \in \mathcal{D}$.
 - Thus a function is concave iff the tangent line lies above the graph of the function.

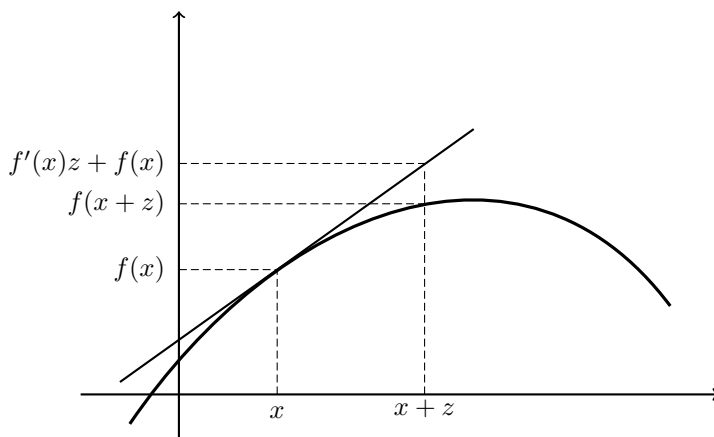


Figure 4: A function is concave iff the tangent line lies above the graph of the function.

- In the next theorem, the concavity or convexity of a C^2 function is characterized using the second derivative.
- The theorem also gives a sufficient condition which can be used to identify strictly concave and strictly convex functions.

Theorem 6. Let \mathcal{D} be an open convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a C^2 . Then

1. f is concave iff $D^2 f(x)$ is a negative semidefinite matrix for all $x \in \mathcal{D}$.
2. f is convex iff $D^2 f(x)$ is a positive semidefinite matrix for all $x \in \mathcal{D}$.
3. If $D^2 f(x)$ is a negative definite matrix for all $x \in \mathcal{D}$, then f is strictly concave.
4. If $D^2 f(x)$ is a positive definite matrix for all $x \in \mathcal{D}$, then f is strictly convex.

- It is important to note that parts (3) and (4) of the theorem are only sufficient conditions. For example, part (3) does not say that if f is strictly concave on \mathcal{D} , then $D^2 f(x)$ is a negative definite matrix for all $x \in \mathcal{D}$.
- The next example illustrates this point.

Example 1. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ be defined by $f(x) = -x^4$ and $g(x) = x^4$ respectively.

- The f is strictly concave on \mathbb{R} , while g is strictly convex on \mathbb{R} .
- However $f''(0) = g''(0)$, so that $f''(0)$ is not negative definite and $g''(0)$ is not positive definite. \blacklozenge
- Our next example illustrates the importance of the theorem for simplifying the identification of concavity in practice.

Example 2. Let $f : \mathbb{R}_{++}^2 \rightarrow \mathbb{R}$ be given by $f(x, y) = x^a y^b$, $a, b > 0$.

- For given a and b , this function is concave if, for any (x, y) and (\hat{x}, \hat{y}) in \mathbb{R}_{++}^2 and any $\lambda \in (0, 1)$, we have

$$[\lambda x + (1 - \lambda)\hat{x}]^a [\lambda y + (1 - \lambda)\hat{y}]^b \geq \lambda x^a y^b + (1 - \lambda)\hat{x}^a \hat{y}^b.$$

- Similarly f is convex, if for all (x, y) and (\hat{x}, \hat{y}) in \mathbb{R}_{++}^2 and any $\lambda \in (0, 1)$, we have

$$[\lambda x + (1 - \lambda)\hat{x}]^a [\lambda y + (1 - \lambda)\hat{y}]^b \leq \lambda x^a y^b + (1 - \lambda)\hat{x}^a \hat{y}^b.$$

- Compare checking for convexity of f using these inequalities to checking using the second derivative test.

- The latter only requires us to identify the definiteness of the following matrix:

$$D^2 f(x, y) = \begin{pmatrix} a(a-1)x^{a-2}y^b & abx^{a-1}y^{b-1} \\ abx^{a-1}y^{b-1} & b(b-1)x^a y^{b-2} \end{pmatrix}.$$

The determinant of this matrix is

$$ab(1-a-b)x^{2(a-1)}y^{2(b-1)}$$

which is positive if $a+b < 1$, zero if $a+b = 1$ and negative if $a+b > 1$.

- Furthermore, if $a, b < 1$ the diagonal terms are negative and so f is a strictly concave function if $a+b < 1$ and concave if $a+b = 1$. If $a+b > 1$, then $D^2 f(x, y)$ is indefinite and f is neither concave nor convex.
- In summary, a Cobb-Douglas production function on \mathbb{R}_{++}^2 is concave iff it exhibits constant or decreasing returns to scale. ♦
- We now present some results which indicate the importance of convexity for optimization theory.
- But first some terminology.

Definition.

- We refer to a maximization problem as a *convex maximization problem* if the constraint set is convex and the objective function is concave.
- Similarly, we refer to a minimization problem as a *convex minimization problem* if the constraint set is convex and the objective function is convex.
- More generally, we refer to an optimization problem as a *convex optimization problem* if it is either of the above. ▲
- The first result establishes that in convex optimization problems, all local optima must also be global optima.
- Thus, to find a global optimum in such problems, it is sufficient to identify a local optimum.

Theorem 7. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be concave. Then

1. Any local maximum of f is a global maximum of f .
 2. The set $\arg \max\{f(x) \mid x \in \mathcal{D}\}$ of maximizers of f on \mathcal{D} is either empty or convex.
- Similar results hold for convex minimization problems.
 - The second part of the result means that we cannot have multiple isolated points as maximizers.

- For example, in the utility maximization problem with two perfect substitutes, either the solution is a unique corner solution or there are infinitely many solutions along the budget constraint.
- The second result shows that if a *strictly* convex optimization problem has a solution, then the solution is unique.

Theorem 8. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be strictly concave. Then the set $\arg \max\{f(x) \mid x \in \mathcal{D}\}$ of maximizers of f on \mathcal{D} is either empty or contains a single point.

- We can combine this result with the Weierstrass theorem to establish the existence of a unique global optimum in a convex optimization problem in which the objective function is continuous and the constraint set is compact.

Quasiconcave and Quasiconvex Functions.

- We have seen that convexity has powerful implications for optimization problems. However, convexity is a very restrictive assumption, which is important when we come to applications.
- For example, we saw that the Cobb-Douglas function production $f(x, y) = x^a y^b$ ($a, b > 0$) is not concave unless $a + b \leq 1$.
- So, we will now look at optimization under a weakening of the condition of convexity, called quasiconvexity.

Definition. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function.

- The *upper contour set* of f at $a \in \mathbb{R}$, denoted $U_f(a)$, is the set

$$U_f(a) = \{x \in \mathcal{D} \mid f(x) \geq a\}.$$

- The *lower contour set* of f at $a \in \mathbb{R}$, denoted $L_f(a)$, is the set

$$L_f(a) = \{x \in \mathcal{D} \mid f(x) \leq a\}. \quad \blacktriangle$$

Definition. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function.

- We say that f is *quasiconcave* on \mathcal{D} if $U_f(a)$ is a convex set for all $a \in \mathbb{R}$.
- We say that f is *quasiconvex* on \mathcal{D} if $L_f(a)$ is a convex set for all $a \in \mathbb{R}$. \blacktriangle
- Thus a function is quasiconcave if its upper contour sets are convex sets.
- Similarly, a function is quasiconvex if its lower contour sets are convex sets.
- As is the case with concave and convex functions, it is also true for quasiconcave and quasiconvex functions that a relationship exists between the value of a function at two points and the value of the function at a convex combination.

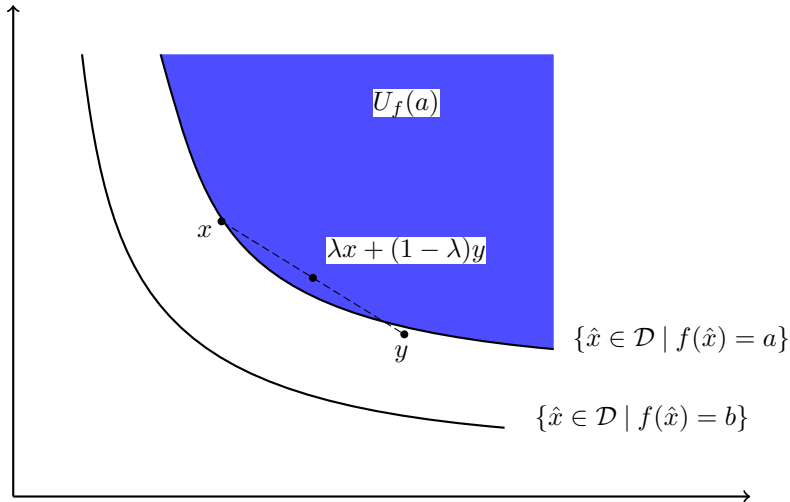


Figure 5: The level sets of a strictly quasiconcave function ($a > b$). The upper contour set of f at a is $U_f(a) = \{\hat{x} \in \mathcal{D} \mid f(\hat{x}) \geq a\}$.

- The following theorem provides two alternative definitions of quasiconcavity.

Theorem 9. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function. Then the following statements are equivalent.

1. f is quasiconcave on \mathcal{D} .
2. For all $x, y \in \mathcal{D}$ and all $\lambda \in (0, 1)$

$$f(x) \geq f(y) \text{ implies } f(\lambda x + (1 - \lambda)y) \geq f(y).$$

3. For all $x, y \in \mathcal{D}$ and all $\lambda \in (0, 1)$

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

- A similar result holds for quasiconvex functions, with the inequalities reversed and “min” replaced with “max”.

Definition. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function.

- We say f is *strictly quasiconcave* if for all $x, y \in \mathcal{D}$ with $x \neq y$, and all $\lambda \in (0, 1)$, we have

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

- We say f is *strictly quasiconvex* if for all $x, y \in \mathcal{D}$ with $x \neq y$, and all $\lambda \in (0, 1)$, we have

$$f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\}. \quad \blacktriangle$$

Theorem 10. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function. Then

1. f is quasiconcave iff the function $-f$ is quasiconvex.
2. f is strictly quasiconcave iff the function $-f$ is strictly quasiconvex.

Quasiconvexity as a Generalization of Convexity.

- It is straightforward to show that the set of all quasiconcave functions contains the set of all concave functions and similarly for quasiconvex functions.

Theorem 11. *Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a function. Then*

1. *If f is concave on \mathcal{D} , then it is also quasiconcave on \mathcal{D} .*
2. *If f is convex on \mathcal{D} , then it is also quasiconvex on \mathcal{D} .*

- The following example demonstrates how to check directly for quasiconvexity and shows the converse of the above result is false.

Example 3. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be any increasing function. Then f is both quasiconcave and quasiconvex.

- To show this, consider any $x, y \in \mathbb{R}$ and any $\lambda \in (0, 1)$. Assume, without loss of generality, that $x > y$. Then

$$x > \lambda x + (1 - \lambda)y > y.$$

- Since f is increasing, we have

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

- Since $f(x) = \max\{f(x), f(y)\}$, the first inequality shows that f is quasiconvex.
- Similarly, since $f(y) = \min\{f(x), f(y)\}$, the second inequality shows that f is quasiconcave.
- Since it is always possible to choose a nondecreasing function f that is neither concave nor convex on \mathbb{R} (say $f(x) = x^3$), we have shown that not every quasiconcave function is concave and not every quasiconvex function is convex. ♦
- The next theorem elaborates on the relationship between concave and quasiconcave functions.

Theorem 12. *Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a quasiconcave function.*

1. *If $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is an increasing function, then the composition $\phi \circ f$ is a quasiconcave function from \mathcal{D} to \mathbb{R} .*
2. *In particular, any increasing transform of a concave function results in a quasiconcave function.*

- The converse of this theorem is not true. That is, we cannot say that every quasiconcave function is an increasing transformation of some concave function. See Sundaram pp207-209 for two concrete examples of quasiconcave functions that are not increasing transformations of any concave function.

Quasiconvexity and the Properties of the Derivative.

- As with concavity we can characterize the quasiconcavity of a differentiable function using the first derivative.

Theorem 13. Let \mathcal{D} be an open convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a C^1 function. Then

1. f is quasiconcave iff $f(y) \geq f(x)$ implies $Df(x)(y - x) \geq 0$ for all $x, y \in \mathcal{D}$.
 2. f is quasiconvex iff $f(y) \leq f(x)$ implies $Df(x)(y - x) \leq 0$ for all $x, y \in \mathcal{D}$.
- The condition (1) is illustrated in the figure. If we think of $Df(x)^T$ as the gradient vector $\nabla f(x)$, then the theorem says that the angle between the gradient and the vector $y - x$ is acute (or right).

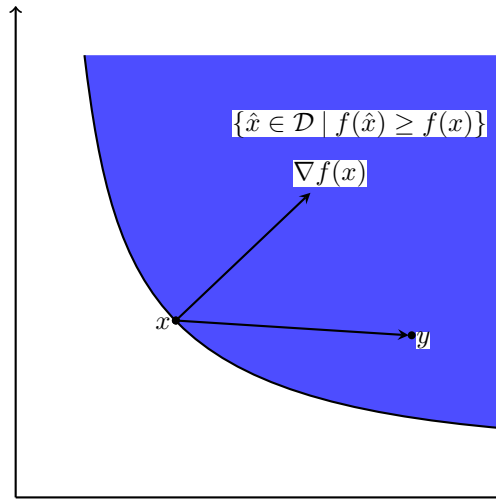


Figure 6: The condition (1) says that the angle between the vector $y - x$ and $\nabla f(x)$ is acute.

- We can also test for quasiconcavity using the second derivative.

Theorem 14. Let \mathcal{D} be an open convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be a C^2 function. Consider the bordered Hessian

$$\overline{H} = \begin{pmatrix} 0 & f_1 & \cdots & f_n \\ f_1 & f_{11} & \cdots & f_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ f_n & f_{n1} & \cdots & f_{nn} \end{pmatrix}$$

Let \overline{H}_k denote the k th order leading principal submatrix of \overline{H} .

1. If f is quasiconcave on \mathcal{D} , then, for all $x \in \mathcal{D}$, $(-1)^{r-1}|\overline{H}_r| \geq 0$ for $r = 2, \dots, n+1$.
 2. If f is quasiconvex on \mathcal{D} , then, for all $x \in \mathcal{D}$, $|\overline{H}_r| \leq 0$ for $r = 2, \dots, n+1$.
 3. If $(-1)^{r-1}|\overline{H}_r| > 0$ for all $r = 2, \dots, n+1$, then f is quasiconcave on \mathcal{D} .
 4. If $|\overline{H}_r| < 0$ for all $r = 2, \dots, n+1$, then f is quasiconvex on \mathcal{D} .
- Part (3) requires the signs of the leading principal minors to alternate, starting with negative for the 2×2 matrix \overline{H}_2 .
 - Compare this theorem with the corresponding theorem on concavity.
 - There are two important differences.
 - In theorem (6), a weak inequality i.e. the negative semidefiniteness of $D^2 f$ was both necessary and sufficient to establish concavity.
 - However, in the result above, the weak inequality is only a *necessary* condition for quasiconcavity. The sufficient condition involves a strict inequality.
 - Second, the theorem does not give a test for *strict* quasiconcavity.

Example 4. Let $f : \mathbb{R}_{++}^2 \rightarrow \mathbb{R}$ be given by $f(x, y) = x^a y^b$, $a, b > 0$.

- We saw that f is strictly concave on if $a + b < 1$, concave if $a + b = 1$, and neither concave nor convex if $a + b > 1$.
- We will show that f is quasiconcave for all $a, b > 0$.
- To show this directly, using the definition of quasiconcavity, requires us to prove that

$$[\lambda x + (1 - \lambda)\hat{x}]^a [\lambda y + (1 - \lambda)\hat{y}]^b \geq \min\{x^a y^b, \hat{x}^a, \hat{y}^b\}$$

holds for all $(x, y) \neq (\hat{x}, \hat{y})$ in \mathbb{R}_{++}^2 and for all $\lambda \in (0, 1)$.

- Compare checking for quasiconcavity f using this inequality to checking using the second derivative test.
- We have to show that $|\overline{H}_2(x, y)| < 0$ and $|\overline{H}_3(x, y)| > 0$ for all $x, y \in \mathbb{R}_{++}^2$, where

$$\overline{H}_2(x, y) = \begin{pmatrix} 0 & ax^{a-1}y^b \\ ax^{a-1}y^b & a(a-1)x^{a-2}y^b \end{pmatrix},$$

$$\overline{H}_3(x, y) = \begin{pmatrix} 0 & ax^{a-1}y^b & bx^a y^{b-1} \\ ax^{a-1}y^b & a(a-1)x^{a-2}y^b & abx^{a-1}y^{b-1} \\ bx^a y^{b-1} & abx^{a-1}y^b & b(b-1)x^a y^{b-2} \end{pmatrix}.$$

- Calculating the determinants we find

$$|\overline{H}_2(x, y)| = -a^2 x^{2(a-1)} y^{2b} < 0$$

$$|\overline{H}_3(x, y)| = ab(a+b)x^{3a-2}y^{3b-2} > 0,$$

for all $(x, y) \in \mathbb{R}_{++}^2$.

- Thus f is quasiconcave on $(x, y) \in \mathbb{R}_{++}^2$. ◆

Quasiconvexity and Optimization.

- Unlike concave and convex functions:
 - Quasiconcave and quasiconvex functions are not necessarily continuous on the interior of their domains.
 - Quasiconcave functions can have local maxima that are not global maxima, and quasiconvex functions can have local minima that are not global minima.
 - First order conditions are not sufficient to identify even local optima under quasiconvexity.
- The following example illustrates these points.

Example 5. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by

$$f(x) = \begin{cases} x^3, & x \leq 1 \\ 1, & x \in (1, 2] \\ x^3, & x > 2 \end{cases}$$

Since f is increasing, it is both quasiconcave and quasiconvex on \mathbb{R} .

- Clearly, f has a discontinuity at $x = 2$.
- Also, f is constant on the open interval $(1, 2)$, so that every point in this interval is a local maximizer and local minimizer of f .
- However, no point in $(1, 2)$ is either a global maximizer or a global minimizer.
- Finally, $f'(0) = 0$, although 0 is not a local maximum or local minimum. ♦
- Another important distinction between convexity and quasiconvexity, is that while a strictly concave function cannot be even weakly convex, a strictly quasiconcave function can also be strictly quasiconvex.
- For example any strictly increasing function on \mathbb{R} is both strictly quasiconvex and strictly quasiconcave. This can be shown by modifying example (3).
- We saw that local maxima of quasiconcave functions need not be global maxima.
- However, when the function is strictly quasiconcave, there is a result identical to that for strictly concave functions.

Theorem 15. Let \mathcal{D} be a convex subset of \mathbb{R}^n and let $f : \mathcal{D} \rightarrow \mathbb{R}$ be strictly quasiconcave. Then

1. Any local maximum of f is a global maximum of f .
2. The set $\arg \max\{f(x) \mid x \in \mathcal{D}\}$ of maximizers of f on \mathcal{D} is either empty or a singleton.

- A similar result holds for strictly quasiconvex functions in minimization problems.
- This is significant because it says that the weaker property of strict quasiconcavity is enough to guarantee uniqueness of the solution (if there is one).