

MS-E2122 - Nonlinear Optimization

Lecture III

Fernando Dias

Department of Mathematics and Systems Analysis

Aalto University
School of Science

September 3, 2023

Outline of this lecture

Convex functions

- Basic properties

- Subgradients of convex functions

- Differentiable functions

- Generalisations of convexity

Last Week ...

- ▶ Convex sets;
- ▶ Weierstrass theorem;
- ▶ Farka's theorem;

Last week...

Outline of this lecture

Convex functions

Basic properties

Subgradients of convex functions

Differentiable functions

Generalisations of convexity

Convexity of functions

Convexity is a key feature in optimisation.

In convex optimisation problems,

local optimality always implies global optimality.

Definition 1 (Convexity of a function)

Let $f : S \rightarrow \mathbb{R}$ where $S \subseteq \mathbb{R}^n$ is a nonempty convex set. The function f is said to be **convex** on S if

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

for each $x_1, x_2 \in S$ and for each $\lambda \in [0, 1]$.

Convexity of functions

Convexity is a key feature in optimisation.

In convex optimisation problems,

local optimality always implies global optimality.

Definition 1 (Convexity of a function)

Let $f : S \rightarrow \mathbb{R}$ where $S \subseteq \mathbb{R}^n$ is a nonempty convex set. The function f is said to be **convex** on S if

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

for each $x_1, x_2 \in S$ and for each $\lambda \in [0, 1]$.

Remarks:

- ▶ f is **concave** if $-f$ is convex;
- ▶ if strict inequality holds, f is strictly convex.

Convexity of functions

Examples of **convex functions**:

▶ $f(x) = a^\top x + b;$

Convexity of functions

Examples of **convex functions**:

- ▶ $f(x) = a^\top x + b$;
- ▶ $f(x) = e^x$;

Convexity of functions

Examples of **convex functions**:

- ▶ $f(x) = a^\top x + b$;
- ▶ $f(x) = e^x$;
- ▶ $f(x) = x^p$ on \mathbb{R}_+ for $p \leq 0$ or $p \geq 1$; concave for $0 \leq p \leq 1$;

Convexity of functions

Examples of **convex functions**:

- ▶ $f(x) = a^\top x + b$;
- ▶ $f(x) = e^x$;
- ▶ $f(x) = x^p$ on \mathbb{R}_+ for $p \leq 0$ or $p \geq 1$; concave for $0 \leq p \leq 1$;
- ▶ $f(x) = \|x\|_p$ (p -norm);

Convexity of functions

Examples of **convex functions**:

- ▶ $f(x) = a^\top x + b$;
- ▶ $f(x) = e^x$;
- ▶ $f(x) = x^p$ on \mathbb{R}_+ for $p \leq 0$ or $p \geq 1$; concave for $0 \leq p \leq 1$;
- ▶ $f(x) = \|x\|_p$ (p -norm);
- ▶ $f(x) = \log x$ and negative entropy $f(x) = -x \log x$ are concave;

Convexity of functions

Examples of **convex functions**:

- ▶ $f(x) = a^\top x + b$;
- ▶ $f(x) = e^x$;
- ▶ $f(x) = x^p$ on \mathbb{R}_+ for $p \leq 0$ or $p \geq 1$; concave for $0 \leq p \leq 1$;
- ▶ $f(x) = \|x\|_p$ (p -norm);
- ▶ $f(x) = \log x$ and negative entropy $f(x) = -x \log x$ are concave;
- ▶ $f(x) = \max \{x_1, \dots, x_n\}$.

Convexity of functions

Convexity **preserving** operations:

- ▶ let $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then these are convex:

Convexity of functions

Convexity **preserving** operations:

- ▶ let $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then these are convex:
 - $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$ where $\alpha_j > 0$ for $j = 1, \dots, k$;

Convexity of functions

Convexity **preserving** operations:

- ▶ let $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then these are convex:
 - $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$ where $\alpha_j > 0$ for $j = 1, \dots, k$;
 - $f(x) = \max \{f_1(x), \dots, f_k(x)\}$;

Convexity of functions

Convexity **preserving** operations:

- ▶ let $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then these are convex:
 - $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$ where $\alpha_j > 0$ for $j = 1, \dots, k$;
 - $f(x) = \max \{f_1(x), \dots, f_k(x)\}$;
- ▶ $f(x) = \frac{1}{g(x)}$ on S , where $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is concave and $S = \{x : g(x) > 0\}$;

Convexity of functions

Convexity **preserving** operations:

- ▶ let $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then these are convex:
 - $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$ where $\alpha_j > 0$ for $j = 1, \dots, k$;
 - $f(x) = \max \{f_1(x), \dots, f_k(x)\}$;
- ▶ $f(x) = \frac{1}{g(x)}$ on S , where $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is concave and $S = \{x : g(x) > 0\}$;
- ▶ $f(x) = g(h(x))$, where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a nondecreasing convex function and $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex.

Convexity of functions

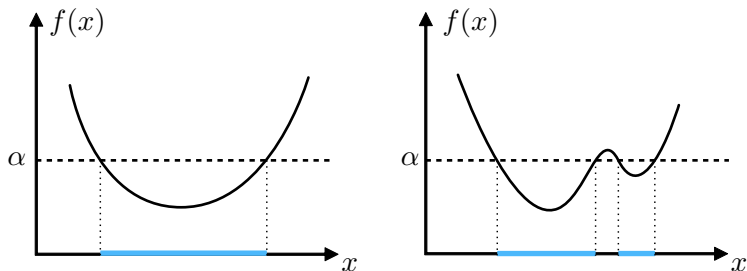
Convexity **preserving** operations:

- ▶ let $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then these are convex:
 - $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$ where $\alpha_j > 0$ for $j = 1, \dots, k$;
 - $f(x) = \max \{f_1(x), \dots, f_k(x)\}$;
- ▶ $f(x) = \frac{1}{g(x)}$ on S , where $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is concave and $S = \{x : g(x) > 0\}$;
- ▶ $f(x) = g(h(x))$, where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a nondecreasing convex function and $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex.
- ▶ $f(x) = g(h(x))$, where $g : \mathbb{R}^m \rightarrow \mathbb{R}$ is convex and $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is affine: $h(x) = Ax + b$ with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Lower level sets

Definition 2 (Lower level set)

Let $S \subseteq \mathbb{R}^n$ be a nonempty set. The lower level set of $f : \mathbb{R}^n \mapsto \mathbb{R}$ for $\alpha \in \mathbb{R}$ is given by $S_\alpha = \{x \in S : f(x) \leq \alpha\}$.



The lower level sets S_α (in blue) of two functions, given a value of α . Notice the nonconvexity of the level set of the nonconvex function (on the right)

Convex functions and lower level sets

Convex level sets are necessary for one to be able to state **global optimality conditions**.

Convex functions **always** present convex lower level sets. The converse is not necessarily true.

Convex functions and lower level sets

Convex level sets are necessary for one to be able to state **global optimality conditions**.

Convex functions **always** present convex lower level sets. The converse is not necessarily true.

Lemma 3

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Then, any level set S_α with $\alpha \in \mathbb{R}$ is convex.

Convex functions and lower level sets

Convex level sets are necessary for one to be able to state **global optimality conditions**.

Convex functions **always** present convex lower level sets. The converse is not necessarily true.

Lemma 3

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Then, any level set S_α with $\alpha \in \mathbb{R}$ is convex.

Proof.

Let $x_1, x_2 \in S_\alpha$. Thus, $x_1, x_2 \in S$ with $f(x_1) \leq \alpha$ and $f(x_2) \leq \alpha$. Let $\lambda \in (0, 1)$ and $x = \lambda x_1 + (1 - \lambda)x_2$. Since S is convex, we have $x \in S$. Now, by the convexity of f , we have

$$f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda\alpha + (1 - \lambda)\alpha = \alpha$$

and thus $x \in S_\alpha$.



Epigraphs and convex functions

Convex functions can be characterised by examining **supporting hyperplanes** of their **epigraphs**.

Definition 4 (Ephigraph)

Let $S \subseteq \mathbb{R}^n$ be a nonempty set and $f : S \rightarrow \mathbb{R}$. The epigraph of f is

$$\mathbf{epi}(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \geq f(x)\} \subseteq \mathbb{R}^{n+1}$$

Epigraphs and convex functions

Convex functions can be characterised by examining **supporting hyperplanes** of their **epigraphs**.

Definition 4 (Ephigraph)

Let $S \subseteq \mathbb{R}^n$ be a nonempty set and $f : S \rightarrow \mathbb{R}$. The epigraph of f is

$$\mathbf{epi}(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \geq f(x)\} \subseteq \mathbb{R}^{n+1}$$

Remark: the hypograph of f is $\{(x, y) : x \in S, y \in \mathbb{R}, y \leq f(x)\}$.

Epigraphs and convex functions

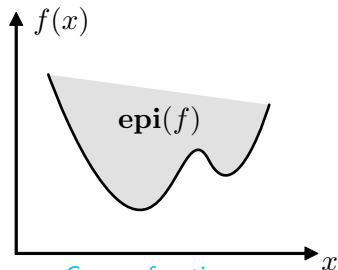
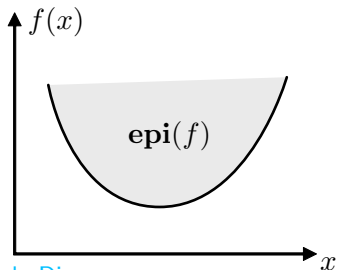
Convex functions can be characterised by examining **supporting hyperplanes** of their **epigraphs**.

Definition 4 (Ephigraph)

Let $S \subseteq \mathbb{R}^n$ be a nonempty set and $f : S \rightarrow \mathbb{R}$. The epigraph of f is

$$\text{epi}(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \geq f(x)\} \subseteq \mathbb{R}^{n+1}$$

Remark: the hypograph of f is $\{(x, y) : x \in S, y \in \mathbb{R}, y \leq f(x)\}$.



Epigraphs and convex functions

Epigraphs can be used to **infer the convexity of functions**.

Theorem 5

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Then f is convex if and only if $\text{epi}(f)$ is a convex set.

Epigraphs and convex functions

Epigraphs can be used to **infer the convexity of functions**.

Theorem 5

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Then f is convex if and only if $\mathbf{epi}(f)$ is a convex set.

Proof.

First, suppose f is convex and let $(x_1, y_1), (x_2, y_2) \in \mathbf{epi}(f)$. Then, for $\lambda \in (0, 1)$ we have

$$\lambda y_1 + (1 - \lambda)y_2 \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2).$$

As $\lambda x_1 + (1 - \lambda)x_2 \in S$, $(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2) \in \mathbf{epi}(f)$.

Epigraphs and convex functions

Epigraphs can be used to **infer the convexity of functions**.

Theorem 5

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Then f is convex if and only if $\mathbf{epi}(f)$ is a convex set.

Proof.

First, suppose f is convex and let $(x_1, y_1), (x_2, y_2) \in \mathbf{epi}(f)$. Then, for $\lambda \in (0, 1)$ we have

$$\lambda y_1 + (1 - \lambda)y_2 \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2).$$

As $\lambda x_1 + (1 - \lambda)x_2 \in S$, $(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2) \in \mathbf{epi}(f)$.

Conversely, suppose $\mathbf{epi}(f)$ is convex. For $x_1, x_2 \in S$. Then

$(x_1, f(x_1)) \in \mathbf{epi}(f)$, $(x_2, f(x_2)) \in \mathbf{epi}(f)$ and thus

$(\lambda x_1 + (1 - \lambda)x_2, \lambda f(x_1) + (1 - \lambda)f(x_2)) \in \mathbf{epi}(f)$ for $\lambda \in (0, 1)$.

These imply that $\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2)$. □

Epigraphs and convex functions

As epigraphs are convex sets, they have **supporting hyperplanes** at their boundary points, which leads to the notion of **subgradients**.

Definition 6 (Subgradients)

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Then $\xi \in \mathbb{R}^n$ is a **subgradient** of f at $\bar{x} \in S$ if

$$f(x) \geq f(\bar{x}) + \xi^\top (x - \bar{x}).$$

Epigraphs and convex functions

As epigraphs are convex sets, they have **supporting hyperplanes** at their boundary points, which leads to the notion of **subgradients**.

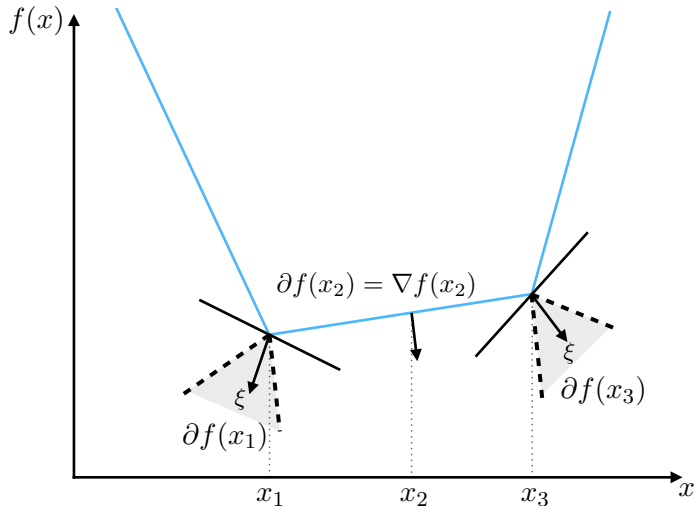
Definition 6 (Subgradients)

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Then $\xi \in \mathbb{R}^n$ is a **subgradient** of f at $\bar{x} \in S$ if

$$f(x) \geq f(\bar{x}) + \xi^\top (x - \bar{x}).$$

Remark: $\partial f(\bar{x}) = \{\xi : f(x) \geq f(\bar{x}) + \xi^\top (x - \bar{x})\}$ is the **subdifferential** of f at \bar{x} , which is convex.

Subgradients



A representation of the subdifferential (in grey) for nondifferentiable (x_1 and x_3) and differentiable (x_2) points

Subgradients and supporting hyperplanes

Every convex function has **at least one subgradient** in the interior of its domain, which can also be used to show convexity of f .

Theorem 7

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Then for all $\bar{x} \in \mathbf{int}(S)$, there exists $\xi \in \mathbb{R}^n$ such that

$$H = \left\{ (x, y) : y = f(\bar{x}) + \xi^\top (x - \bar{x}) \right\}$$

supports $\mathbf{epi}(f)$ at $(\bar{x}, f(\bar{x}))$. In particular,

$$f(x) \geq f(\bar{x}) + \xi^\top (x - \bar{x}), \forall x \in S.$$

Subgradients and supporting hyperplanes

Every convex function has **at least one subgradient** in the interior of its domain, which can also be used to show convexity of f .

Theorem 7

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Then for all $\bar{x} \in \mathbf{int}(S)$, there exists $\xi \in \mathbb{R}^n$ such that

$$H = \left\{ (x, y) : y = f(\bar{x}) + \xi^\top (x - \bar{x}) \right\}$$

supports $\mathbf{epi}(f)$ at $(\bar{x}, f(\bar{x}))$. In particular,

$$f(x) \geq f(\bar{x}) + \xi^\top (x - \bar{x}), \forall x \in S.$$

Proof sketch: use [Theorem 5](#) and the support of convex sets to show that the subgradient inequality holds.

Differentiability and gradients

Let us first define **differentiability**.

Definition 8

Let $S \subseteq \mathbb{R}^n$ be a nonempty set. $f : S \rightarrow \mathbb{R}$ is differentiable at $\bar{x} \in \text{int}(S)$ if there exists a vector $\nabla f(\bar{x})$, called a gradient vector, and a function $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}) + \|x - \bar{x}\| \alpha(\bar{x}; x - \bar{x})$$

where $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}; x - \bar{x}) = 0$

Differentiability and gradients

Let us first define **differentiability**.

Definition 8

Let $S \subseteq \mathbb{R}^n$ be a nonempty set. $f : S \rightarrow \mathbb{R}$ is differentiable at $\bar{x} \in \mathbf{int}(S)$ if there exists a vector $\nabla f(\bar{x})$, called a gradient vector, and a function $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}) + \|x - \bar{x}\| \alpha(\bar{x}; x - \bar{x})$$

where $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}; x - \bar{x}) = 0$

Remarks:

- ▶ This is called the **first-order (Taylor series) expansion** of f . Without the α term, it is the first-order approximation.
- ▶ If this holds for all $\bar{x} \in \mathbf{int}(S)$, then f is differentiable in S .
- ▶ $\nabla f(\bar{x}) = \left(\frac{\partial f(\bar{x})}{\partial x_1}, \dots, \frac{\partial f(\bar{x})}{\partial x_n} \right)$ is unique.

Differentiability and gradients

If f is convex and differentiable, $\partial(x)$ is the singleton $\{\nabla f(x)\}$.

Lemma 9

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Suppose that f is differentiable at $\bar{x} \in \mathbf{int}(S)$. Then $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$ is a singleton.

Differentiability and gradients

If f is convex and differentiable, $\partial f(x)$ is the singleton $\{\nabla f(x)\}$.

Lemma 9

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$ a convex function. Suppose that f is differentiable at $\bar{x} \in \text{int}(S)$. Then $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$ is a singleton.

Proof.

From [Theorem 7](#), $\partial f(\bar{x}) \neq \emptyset$. Moreover, combining the existence of a subgradient ξ and differentiability of f at \bar{x} , we obtain:

$$f(\bar{x} + \lambda d) \geq f(\bar{x}) + \lambda \xi^\top d \quad (1)$$

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})^\top d + \lambda \|d\| \alpha(\bar{x}; \lambda d) \quad (2)$$

Subtracting (2) from (1), we get $0 \geq \lambda(\xi - \nabla f(\bar{x}))^\top d - \lambda \|d\| \alpha(\bar{x}; \lambda d)$.

Dividing by $\lambda > 0$ and letting $\lambda \rightarrow 0^+$, we obtain $(\xi - \nabla f(\bar{x}))^\top d \leq 0$.

Now, by setting $d = \xi - \nabla f(\bar{x})$, it becomes clear that $\xi = \nabla f(\bar{x})$. \square

Differentiability and gradients

Theorem 10 gives an important characterisation of convex differentiable functions.

(Differentiable) functions are convex if and only if they are always underestimated by affine (first-order) approximations at \bar{x} .

Theorem 10

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex open set, and let $f : S \rightarrow \mathbb{R}$ be differentiable on S . f is convex if and only if for any $\bar{x} \in S$, we have

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}), \quad \forall x \in S.$$

Differentiability and gradients

Theorem 10 gives an important **characterisation of convex differentiable functions**.

(Differentiable) functions are **convex** if and only if they are always **underestimated by affine (first-order) approximations** at \bar{x} .

Theorem 10

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex open set, and let $f : S \rightarrow \mathbb{R}$ be differentiable on S . f is convex if and only if for any $\bar{x} \in S$, we have

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}), \quad \forall x \in S.$$

Proof sketch: combine **Theorem 7** and **Lemma 9**.

Remark: this powerful tool, known as **affine bounding**, is a part of many **optimisation algorithms**.

Second-order differentiability

A function is **twice differentiable** if it has a second order expansion.

Definition 11 (second-order differentiability)

Let $S \subseteq \mathbb{R}^n$ be a nonempty set, and let $f : S \rightarrow \mathbb{R}$. Then f is twice differentiable at $\bar{x} \in \text{int}(S)$ if there exists, for each $\bar{x} \in S$, a vector $\nabla f(\bar{x}) \in \mathbb{R}^n$, an $n \times n$ symmetric matrix $H(\bar{x})$ (the Hessian), and a function $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^\top H(\bar{x}) (x - \bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}; x - \bar{x})$$

where $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}; x - \bar{x}) = 0$.

Second-order differentiability

Remarks:

1. Let $f_{ij}(\bar{x}) = \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}$. The Hessian $H(\bar{x})$ at \bar{x} is given by

$$H(\bar{x}) = \begin{bmatrix} f_{11}(\bar{x}) & \dots & f_{1n}(\bar{x}) \\ \vdots & \ddots & \vdots \\ f_{n1}(\bar{x}) & \dots & f_{nn}(\bar{x}) \end{bmatrix}$$

2. The **Hessian** is **positive semi-definite** at \bar{x} if $x^\top H(\bar{x})x \geq 0$ for $x \in \mathbb{R}^n$.

Second-order differentiability

Remarks:

1. Let $f_{ij}(\bar{x}) = \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}$. The Hessian $H(\bar{x})$ at \bar{x} is given by

$$H(\bar{x}) = \begin{bmatrix} f_{11}(\bar{x}) & \dots & f_{1n}(\bar{x}) \\ \vdots & \ddots & \vdots \\ f_{n1}(\bar{x}) & \dots & f_{nn}(\bar{x}) \end{bmatrix}$$

2. The **Hessian** is **positive semi-definite** at \bar{x} if $x^\top H(\bar{x})x \geq 0$ for $x \in \mathbb{R}^n$.

Example: $f(x_1, x_2) = 2x_1 + 6x_2 - 2x_1^2 - 3x_2^2 + 4x_1x_2$.

$$\nabla f(\bar{x}) = \begin{bmatrix} 2 - 4\bar{x}_1 + 4\bar{x}_2 \\ 6 - 6\bar{x}_2 + 4\bar{x}_1 \end{bmatrix} \text{ and } H(\bar{x}) = \begin{bmatrix} -4 & -6 \\ -4 & -6 \end{bmatrix}.$$

Second-order differentiability

Positive semidefinite Hessians are used to **infer convexity**.

Theorem 12

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex open set, and let $f : S \rightarrow \mathbb{R}$ be twice differentiable on S . Then f is convex if and only if the Hessian matrix is positive semidefinite (PSD) at each point in S .

Second-order differentiability

Positive semidefinite Hessians are used to **infer convexity**.

Theorem 12

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex open set, and let $f : S \rightarrow \mathbb{R}$ be twice differentiable on S . Then f is convex if and only if the Hessian matrix is positive semidefinite (PSD) at each point in S .

Remarks:

- ▶ Checking for PSD of Hessians can be done efficiently using Gauss-Jordan eliminations;
- ▶ **Positive definite** Hessians ($x^\top H(\bar{x})x > 0$ for $x \in \mathbb{R}^n$) imply strict convexity.

Quasiconvexity

Some of the results for convex functions can be generalised to **nonconvex** functions that possess **convex lower level sets**.

Definition 13 (quasiconvex functions)

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Function f is quasiconvex if, for each $x_1, x_2 \in S$ and $\lambda \in (0, 1)$, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \max \{f(x_1), f(x_2)\}.$$

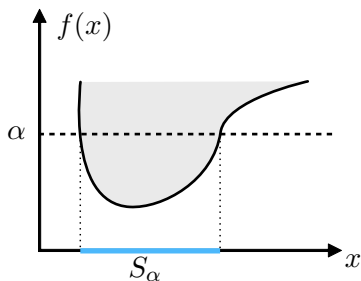
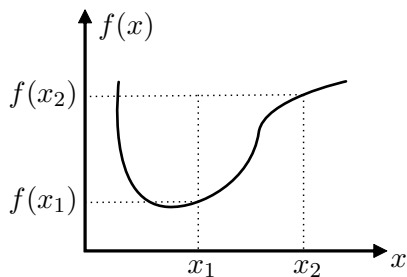
Quasiconvexity

Some of the results for convex functions can be generalised to **nonconvex** functions that possess **convex lower level sets**.

Definition 13 (quasiconvex functions)

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Function f is quasiconvex if, for each $x_1, x_2 \in S$ and $\lambda \in (0, 1)$, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \max \{f(x_1), f(x_2)\}.$$



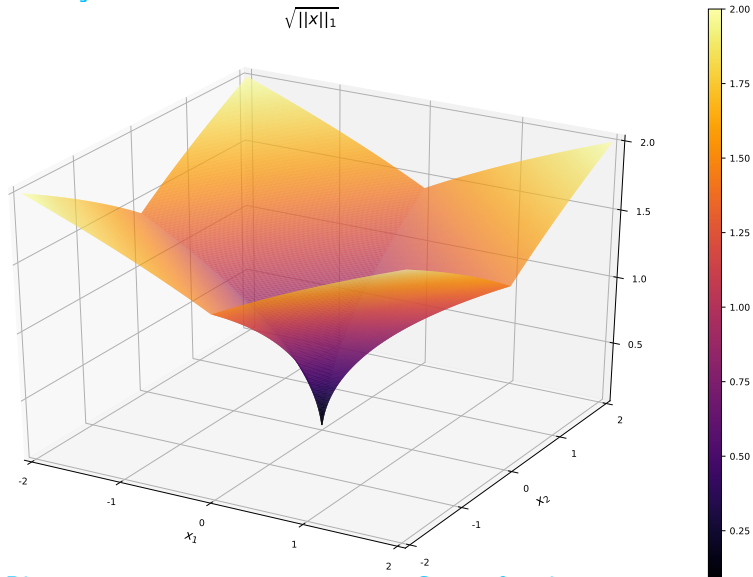
Quasiconvexity

Examples of quasi-convex functions:

- ▶ $f(x) = \sqrt{\|x\|_1}$
- ▶ $f(x) = \log x$ is quasilinear for $x > 0$
- ▶ $f(x) = \inf \{z \in \mathbb{Z} : z \geq x\}$ is quasilinear
- ▶ $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 > 0\}$
- ▶ $f(x_1, x_2) = \log(x_1^2 + x_2^2)$

Remark: f is **quasiconcave** if $-f$ is quasiconvex. Quasilinear functions are both quasiconvex and quasiconcave.

Quasiconvexity



Quasiconvexity

Quasiconvex functions do not necessarily have convex **epigraphs**. However, their lower **level sets** are convex.

Theorem 14

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Function f is quasiconvex if and only if $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ is convex for all $\alpha \in \mathbb{R}$.

Quasiconvexity

Quasiconvex functions do not necessarily have convex **epigraphs**. However, their lower **level sets** are convex.

Theorem 14

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Function f is quasiconvex if and only if $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ is convex for all $\alpha \in \mathbb{R}$.

Proof.

Suppose f is quasiconvex and let $x_1, x_2 \in S_\alpha$. Thus, $x_1, x_2 \in S$ and $\max\{f(x_1), f(x_2)\} \leq \alpha$. Let $x = \lambda x_1 + (1 - \lambda)x_2$ for $\lambda \in (0, 1)$. As S is convex, $x \in S$. By quasiconvexity of f ,

$$f(x) \leq \max\{f(x_1), f(x_2)\} \leq \alpha.$$

Hence, $x \in S_\alpha$ and S_α is convex.

Quasiconvexity

Proof (cont.)

Conversely, assume that S_α is convex for $\alpha \in \mathbb{R}$. Let $x_1, x_2 \in S$, and let $x = \lambda x_1 + (1 - \lambda)x_2$ for $\lambda \in (0, 1)$. Thus, for $\alpha = \max\{f(x_1), f(x_2)\}$, we have $x_1, x_2 \in S_\alpha$. The convexity of S_α implies that $x \in S_\alpha$, and thus $f(x) \leq \alpha = \max\{f(x_1), f(x_2)\}$, which implies that f is quasiconvex.

Quasiconvexity

Proof (cont.)

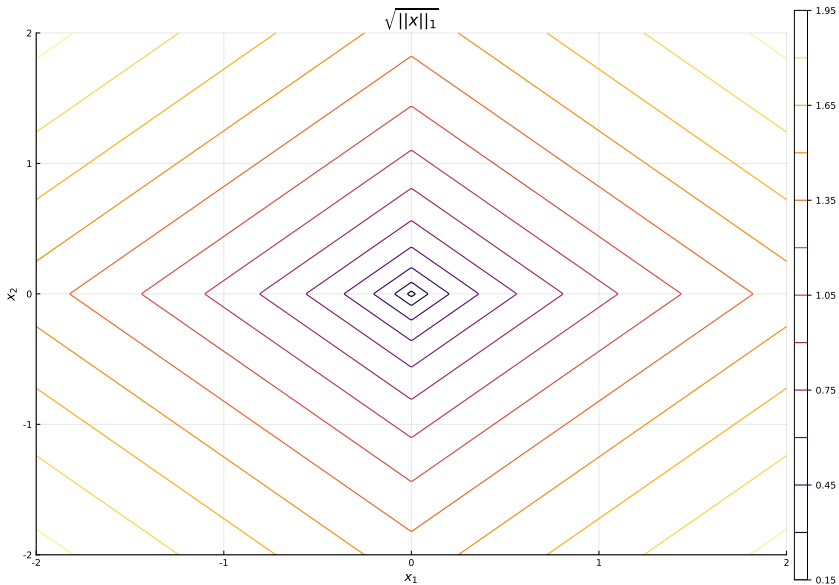
Conversely, assume that S_α is convex for $\alpha \in \mathbb{R}$. Let $x_1, x_2 \in S$, and let $x = \lambda x_1 + (1 - \lambda)x_2$ for $\lambda \in (0, 1)$. Thus, for $\alpha = \max\{f(x_1), f(x_2)\}$, we have $x_1, x_2 \in S_\alpha$. The convexity of S_α implies that $x \in S_\alpha$, and thus $f(x) \leq \alpha = \max\{f(x_1), f(x_2)\}$, which implies that f is quasiconvex.

One important property of continuous quasiconvex functions is the following **first-order condition**.

Theorem 15

Let $S \subseteq \mathbb{R}^n$ be a nonempty open convex set, and let $f : S \rightarrow \mathbb{R}$ be differentiable on S . Then f is quasiconvex if and only if, for $x_1, x_2 \in S$ and $f(x_1) \leq f(x_2)$, $\nabla f(x_2)^\top (x_1 - x_2) \leq 0$.

Remark: geometrically, this means that $\nabla f(x_2)$ defines a **supporting hyperplane** to the (convex) lower level set $S_{f(x_2)}$ at x_2 .



Quasiconvexity and pseudoconvexity

For a **quasiconvex function**, **local optimality does not generalise** to the entire domain, unless the function is **strictly quasiconvex**.

Definition 16 (strict quasiconvexity)

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Function f is strictly quasiconvex if, for each $x_1, x_2 \in S$ with $f(x_1) \neq f(x_2)$ and $\lambda \in (0, 1)$, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) < \max \{f(x_1), f(x_2)\}.$$

Quasiconvexity and pseudoconvexity

For a **quasiconvex function**, **local optimality does not generalise** to the entire domain, unless the function is **strictly quasiconvex**.

Definition 16 (strict quasiconvexity)

Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : S \rightarrow \mathbb{R}$. Function f is strictly quasiconvex if, for each $x_1, x_2 \in S$ with $f(x_1) \neq f(x_2)$ and $\lambda \in (0, 1)$, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) < \max \{f(x_1), f(x_2)\}.$$

Remark:

1. Notice that the definition **precludes the existence of flat spots** anywhere else than at extreme points.
2. **Pseudoconvex** functions are those for which the first-order conditions in **Theorem 15** are **sufficient** for global optimality.



Reference