CO 355
Lecture 13

Convex functions

Vris Cheung

(Substitute class)

October 26, 2010

Outline

Basic notions (that you should know)

- + openness / closedness of a set
- + interior of a set
- + closure of a set
- + (in Euclidean space)

  a set is compact iff it is closed + bounded.
- + Bolzano-Weierstrass theorem:

  A sequence in a compact set has a convergent subsequence.
- + Continuity of functions

Calculus review

+ $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\bar{x} \in \text{int}(S)$ if $\exists\, d \in \mathbb{R}^n$ s.t.

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - d^\top (x - \bar{x})}{\|x - \bar{x}\|} = 0.$$

Here d is called the gradient of f at $\bar{x}$. Notation : $\nabla f(\bar{x})$.

+ Recall:

$$\nabla f(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), \frac{\partial f}{\partial x_2}(\bar{x}), \ldots, \frac{\partial f}{\partial x_n}(\bar{x}) \right).$$

+ If the maps

$$x \mapsto \frac{\partial f}{\partial x_i}(x) \quad (i = 1, \ldots, n)$$

are defined in a nbd. of and are continuous at $\bar{x}$,
then f is continuously differentiable at $\bar{x}$.

Calculus review

If the maps

$$x \mapsto \frac{\partial f}{\partial x_i}(x) \quad (i = 1, \ldots, n)$$

are differentiable at $\bar{x}$, we may define the Hessian of f at $\bar{x}$ as the matrix $\nabla^2 f(\bar{x}) \in \mathbb{R}^{n \times n}$ by

$$\left[\nabla^2 f(\bar{x})\right]_{ij} := \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}).$$

If f is twice continuously differentiable at $\bar{x}$, the Hessian is symmetric.

* In the course notes, the Hessian is denoted by $Hf(\bar{x})$.

Calculus review

Let $S \subseteq \mathbb{R}^n$ be nonempty open,
$\bar{x} \in \text{int}(S)$ and $f : S \to \mathbb{R}$ be given.

+ Gradient of f at $\bar{x}$:

$$\nabla f(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), \frac{\partial f}{\partial x_2}(\bar{x}), \ldots, \frac{\partial f}{\partial x_n}(\bar{x}) \right).$$

+ Hessian of f at $\bar{x}$:
  the matrix $\nabla^2 f(\bar{x}) \in \mathbb{R}^{n \times n}$ by

$$\left[\nabla^2 f(\bar{x})\right]_{ij} := \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}).$$

Outline

Convex functions

Let $S \subseteq \mathbb{R}^n$ be convex (and non-empty).

Definition

A real-valued function $f : S \to \mathbb{R}$ is convex if
for all $x, y \in S, \lambda \in [0,1]$,

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

$f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is concave if $-f$ is convex.

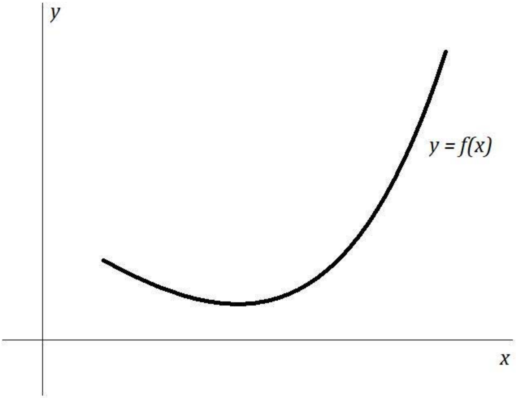
Definition

A function $f : S \to \mathbb{R}$ is strictly convex if
for all distinct $x, y \in S, \lambda \in (0,1)$,

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

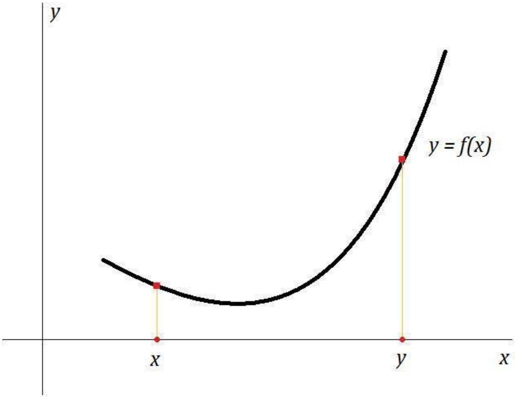
Convex functions

The secant line lies above the curve through two points in S.



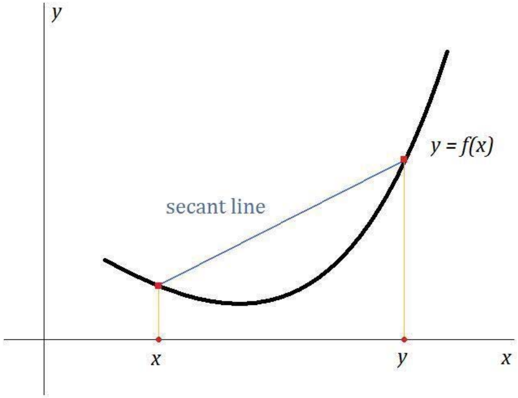
Convex functions

The secant line lies above the curve through two points in S.



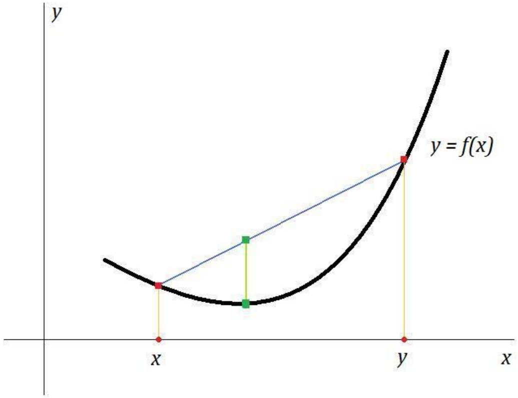
Convex functions

The secant line lies above the curve through two points in S.



Convex functions

The secant line lies above the curve through two points in S.



Convex functions

Examples:

+ affine functions: fox fixed $a \in \mathbb{R}^n$, b,

$$f(x) := a^\top x + b \qquad (x \in \mathbb{R}^n).$$

(If $b = 0$, the function is linear.)

+ norm :

$$f(x) := \|x\|^2 := \sum_{i=1}^{n} x_i^2 \quad \text{for } x \in \mathbb{R}^n.$$

+ some functions involving log:

$$f(t) := -\log t \qquad \text{for } t \in \mathbb{R} \text{ with } x > 0;$$

$$f(x) := -\log\left(\sum_{i=1}^{n} x_i\right) \qquad \text{for } x \in \mathbb{R}^n \text{ with all } x_i > 0;$$

$$f(X) := -\log\det(X) \qquad \text{for positive definite symm. } X \in \mathbb{R}^{n\times n}.$$

Outline

How do we check convexity of a function?

+ by definition... or
+ using calculus.

Equivalent conditions of convexity: (assuming sufficient differentiability)

	$g : \mathbb{R} \to \mathbb{R}$	$f : \mathbb{R}^n \to \mathbb{R}$
epigraph	$\text{epi}(g) \in \mathbb{R} \times \mathbb{R}$ is convex.	$\text{epi}(f) \in \mathbb{R}^n \times \mathbb{R}$ is convex.
1st order cond.	$g(y) \geqslant g(x) + g'(x)(y - x)$ (i.e. g' is non-decreasing)	$f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$
2nd order cond.	$g''(x) \geqslant 0$	$\nabla^2 f(x)$ is positive semidefinite.

Here $\text{epi}(f) := \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leqslant r\}$.

Easy observation (1)

If $\lambda \in [0, 1]$, then

$$x + \lambda(y - x) = (1 - \lambda)x + \lambda y$$

is a conv. combi. of x and y.

Also, if f is convex on a convex set containing x and y, then for any $\lambda \in [0, 1]$

$$\begin{aligned} f(x + \lambda(y - x)) &= f((1 - \lambda)x + \lambda y) \\ &\leqslant (1 - \lambda)f(x) + \lambda f(y) \\ &= f(x) + \lambda [f(y) - f(x)]. \end{aligned}$$

Easy observation (1)

If $\lambda \in [0,1]$, then

$$x + \lambda(y - x) = (1 - \lambda)x + \lambda y$$

is a conv. combi. of x and y.

Also, if f is convex on a convex set containing x and y, then for any $\lambda \in [0,1]$

$$f\left(x + \lambda(y - x)\right) \leqslant f(x) + \lambda\left[f(y) - f(x)\right].$$

Easy observation (2)

Let $S \subseteq \mathbb{R}^n$ be nonempty open,
$\bar{x} \in \operatorname{int}(S)$, non-zero $d \in \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be given.

We can define

$$g(t) := f(\bar{x} + td) \quad (t \in I),$$

the evaluation of f along a line segment through $\bar{x}$ parallel to d.
(Here $I \subseteq \mathbb{R}$ is s.t. $\bar{x} + td \in S$ for all $t \in I$.)

(1) f is convex iff "all such functions g are convex".

(2)

$$\begin{aligned} f \text{ diff. on } S &\implies g \text{ diff. on some nbd. of } 0: \\ & \qquad g'(t) = d^\top \nabla f(\bar{x} + td)\ ; \\ f \text{ twice ctsly diff. on } S &\implies g \text{ twice diff. on some nbd. of } 0: \\ & \qquad g''(t) = d^\top \nabla^2 f(\bar{x} + td) d\ . \end{aligned}$$

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.

Proof.

(1) $\implies$ (2): (Prop. 3.2)
WLOG assume $x \neq y$. Note that if $\lambda \in (0, 1)$, then

$$\lim_{\lambda \searrow 0} \frac{g(x + \lambda(y - x)) - g(x)}{\lambda(y - x)} = g'(x).$$

But $g(x + \lambda(y - x)) \leqslant g(x) + \lambda [g(y) - g(x)]$ by (1). Then, if $y > x$,

$$\frac{\lambda [g(y) - g(x)]}{\lambda(y - x)} \geqslant \frac{g(x + \lambda(y - x)) - g(x)}{\lambda(y - x)} \quad \forall \lambda \in (0, 1).$$

□

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y-x)$ for all $x, y \in I$.

Proof.

(1) $\implies$ (2): (Prop. 3.2)
WLOG assume $x \neq y$. Note that if $\lambda \in (0,1)$, then

$$\lim_{\lambda \searrow 0} \frac{g(x+\lambda(y-x)) - g(x)}{\lambda(y-x)} = g'(x).$$

But $g(x+\lambda(y-x)) \leqslant g(x) + \lambda\,[g(y) - g(x)]$ by (1). Then, if $y > x$,

$$\frac{g(y) - g(x)}{y - x} \geqslant \frac{g(x+\lambda(y-x)) - g(x)}{\lambda} \quad \forall \lambda \in (0,1).$$

□

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y-x)$ for all $x, y \in I$.

Proof.

(1) $\implies$ (2): (Prop. 3.2)
WLOG assume $x \neq y$. Note that if $\lambda \in (0,1)$, then

$$\lim_{\lambda \searrow 0} \frac{g(x+\lambda(y-x)) - g(x)}{\lambda(y-x)} = g'(x).$$

But $g(x+\lambda(y-x)) \leqslant g(x) + \lambda\left[g(y) - g(x)\right]$ by (1). Then, if $y > x$,

$$\frac{g(y)-g(x)}{y-x} \geqslant \lim_{\lambda \searrow 0} \frac{g(x+\lambda(y-x)) - g(x)}{\lambda} = g'(x).$$

□

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.

Proof.

(1) $\implies$ (2): (Prop. 3.2)
WLOG assume $x \neq y$. Note that if $\lambda \in (0,1)$, then

$$\lim_{\lambda \searrow 0} \frac{g(x + \lambda(y-x)) - g(x)}{\lambda(y-x)} = g'(x).$$

But $g(x + \lambda(y-x)) \leqslant g(x) + \lambda\,[g(y) - g(x)]$ by (1). Then, if $y > x$,

$$\frac{g(y) - g(x)}{y - x} \geqslant \lim_{\lambda \searrow 0} \frac{g(x + \lambda(y-x)) - g(x)}{\lambda} = g'(x).$$

The case $y < x$ is similar. □

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.

Proof.

(2) $\implies$ (1): fix any $x, y \in I$ and $\lambda \in [0, 1]$. Let $z := \lambda x + (1 - \lambda)y$.

By (2),

$$g(x) \geqslant g(z) + g'(z)(x - z)$$
$$g(y) \geqslant g(z) + g'(z)(y - z)$$

□

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.

Proof.

(2) $\implies$ (1): fix any $x, y \in I$ and $\lambda \in [0,1]$. Let $z := \lambda x + (1-\lambda)y$.

By (2),

$$\lambda g(x) \geqslant \lambda g(z) + g'(z)\,[\lambda(x - z)]$$
$$(1-\lambda)g(y) \geqslant (1-\lambda)g(z) + g'(z)\,[(1-\lambda)(y - z)]$$

□

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.

Proof.

(2) $\implies$ (1): fix any $x, y \in I$ and $\lambda \in [0, 1]$. Let $z := \lambda x + (1 - \lambda)y$.

By (2),

$$\lambda g(x) \geqslant \lambda g(z) + g'(z)\,[\lambda(x - z)]$$
$$(1 - \lambda)g(y) \geqslant (1 - \lambda)g(z) + g'(z)\,[(1 - \lambda)(y - z)]$$
$$\implies \quad \lambda g(x) + (1 - \lambda)g(y) \geqslant g(z)$$

□

Convexity of functions on $\mathbb{R}$: 1st order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(1) g is convex.
(2) $g(y) \geqslant g(x) + g'(x)(y-x)$ for all $x, y \in I$.

Proof.

(2) $\implies$ (1): fix any $x, y \in I$ and $\lambda \in [0,1]$. Let $z := \lambda x + (1-\lambda)y$.

By (2),

$$
\begin{aligned}
\lambda g(x) &\geqslant \lambda g(z) + g'(z)\,[\lambda(x-z)] \\
(1-\lambda)g(y) &\geqslant (1-\lambda)g(z) + g'(z)\,[(1-\lambda)(y-z)] \\
\implies \quad \lambda g(x) + (1-\lambda)g(y) &\geqslant g(z) \\
&= g(\lambda x + (1-\lambda)y)
\end{aligned}
$$

This is the definition of g being convex. □

Convexity of functions on $\mathbb{R}^n$: 1st order cond.

Theorem (3.5)

Let $S \subseteq \mathbb{R}^n$ be convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(1) f is convex.
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.

Proof.

(1) $\implies$ (2):
The map $g : [0,1] \to \mathbb{R}$ defined by

$$g(t) := f\left(x + t(y - x)\right)$$

is convex. Then by Thm. 3.4,

$$g(1) \geqslant g(0) + g'(0)(1 - 0)$$

□

Convexity of functions on $\mathbb{R}^n$: 1st order cond.

Theorem (3.5)

Let $S \subseteq \mathbb{R}^n$ be convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(1) f is convex.
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.

Proof.

(1) $\implies$ (2):
The map g : [0,1]$\to$ $\mathbb{R}$ defined by

$$g(t) := f\left(x + t(y - x)\right)$$

is convex. Then by Thm. 3.4,

$$\begin{aligned} & g(1) \geqslant g(0) + g'(0)(1 - 0) \\ \implies \quad & f(y) \geqslant f(x) + \nabla f(x)^\top (y - x). \end{aligned}$$

□

Convexity of functions on $\mathbb{R}^n$: 1st order cond.

Theorem (3.5)

Let $S \subseteq \mathbb{R}^n$ be convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(1) f is convex.
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.

Proof.

(1) $\implies$ (2):
The map $g : [0,1] \to \mathbb{R}$ defined by

$$g(t) := f(x + t(y - x))$$

is convex. Then by Thm. 3.4,

$$\begin{aligned} & g(1) \geqslant g(0) + g'(0)(1 - 0) \\ \implies \quad & f(y) \geqslant f(x) + \nabla f(x)^\top (y - x). \end{aligned}$$

The converse is similar to the proof of (2) implying (1) in Thm. 3.4. □

Convexity of functions on $\mathbb{R}$: 2nd order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.
(3) g' is nondecreasing.

Proof.

(2) $\implies$ (3): Let $x < y$. Then

$$g(y) \geqslant g(x) + g'(x)(y - x)$$
$$g(x) \geqslant g(y) + g'(y)(x - y)$$

□

Convexity of functions on $\mathbb{R}$: 2nd order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.
(3) g' is nondecreasing.

Proof.

(2) $\implies$ (3): Let $x < y$. Then

$$\begin{aligned} g(y) &\geqslant g(x) + g'(x)(y - x) \\ g(x) &\geqslant g(y) + g'(y)(x - y) \\ \implies \quad g(y) + g(x) &\geqslant g(x) + g(y) + \left[g'(x) - g'(y)\right](y - x) \end{aligned}$$

□

Convexity of functions on $\mathbb{R}$: 2nd order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.
(3) g' is nondecreasing.

Proof.

(2) $\implies$ (3): Let $x < y$. Then

$$
\begin{aligned}
& g(y) \geqslant g(x) + g'(x)(y - x) \\
& g(x) \geqslant g(y) + g'(y)(x - y) \\
\implies \quad & g(y) + g(x) \geqslant g(x) + g(y) + \left[g'(x) - g'(y)\right](y - x) \\
\implies \quad & 0 \leqslant \left[g'(y) - g'(x)\right](y - x).
\end{aligned}
$$

In particular, $g'(y) \geqslant g'(x)$. □

Convexity of functions on on $\mathbb{R}$: 2nd order cond.

Theorem (3.4)

Let $I \subseteq \mathbb{R}$ and $g : I \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $g(y) \geqslant g(x) + g'(x)(y - x)$ for all $x, y \in I$.
(3) g' is nondecreasing.

Proof.

(3) $\implies$ (2): Let $x < y$ lie in I. By mean-value theorem

$$g(y) - g(x) = g'(z)(y - x)$$

for some $z \in [x, y]$. By (3), $g'(z) \geqslant g'(x)$, so

$$g(y) - g(x) \geqslant g'(x)(y - x).$$

□

Convexity of functions on $\mathbb{R}^n$: 2nd order cond.

Theorem (3.6)

Let $S \subseteq \mathbb{R}^n$ be open convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.
(3) $\nabla^2 f(x)$ is positive semidefinite for all $x \in S$.

Proof.

(2) $\implies$ (3): fix any $x \in S$ and $d \in \mathbb{R}^n$. We show that

$$d^\top \nabla^2 f(x) d \geqslant 0.$$

Then $g(t) := f(x + td)$ is defined on a nbd. I of 0.
We show that for all $t < s \in I$,

$$g(s) \geqslant g(t) + g'(t)(s - t),$$

so that by Thm 3.4, we have that $g'' > 0$ on I. □

Convexity of functions on $\mathbb{R}^n$: 2nd order cond.

Theorem (3.6)

Let $S \subseteq \mathbb{R}^n$ be open convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.
(3) $\nabla^2 f(x)$ is positive semidefinite for all $x \in S$.

Proof.

(2) $\implies$ (3): fix any $x \in S$ and $d \in \mathbb{R}^n$. We show that

$$d^\top \nabla^2 f(x) d \geqslant 0.$$

Then $g(t) := f(x + td)$ is defined on a nbd. I of 0.
For all $t < s \in I$,

$$\begin{aligned} & f(x+sd) \geqslant f(x+td) + \nabla f(x+td)^\top [(s-t)d] \\ \implies \quad & g(s) \geqslant g(t) + g'(t)(s-t) \end{aligned}$$

so by Thm. 3.4,

$$0 \leqslant g''(0) = d^\top \nabla^2 f(x) d.$$

□

Convexity of functions on $\mathbb{R}^n$: 2nd order cond.

Theorem (3.6)

Let $S \subseteq \mathbb{R}^n$ be open convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.
(3) $\nabla^2 f(x)$ is positive semidefinite for all $x \in S$.

Proof.

(3) $\implies$ (2): fix any $x, y \in S$, and let $d := y - x$. Define

$$g : [0,1] \to \mathbb{R} : t \mapsto f(x + td).$$
$$\implies \quad g''(t) = d^\top \nabla^2 f(x + td) d \geqslant 0$$

by positive semidefiniteness of $\nabla^2 f(x + td)$. By Thm 3.4,

$$g(1) \geqslant g(0) + g'(0)$$

□

Convexity of functions on $\mathbb{R}^n$: 2nd order cond.

Theorem (3.6)

Let $S \subseteq \mathbb{R}^n$ be open convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.
(3) $\nabla^2 f(x)$ is positive semidefinite for all $x \in S$.

Proof.

(3) $\implies$ (2): fix any $x, y \in S$, and let $d := y - x$. Define

$$g : [0,1] \to \mathbb{R} : t \mapsto f(x + td).$$
$$\implies \quad g''(t) = d^\top \nabla^2 f(x + td) d \geqslant 0$$

by positive semidefiniteness of $\nabla^2 f(x + td)$. By Thm 3.4,

$$g(1) \geqslant g(0) + g'(0)$$
$$\implies \quad f(x + d) \geqslant f(x) + \nabla f(x)^\top d.$$

□

Convexity of functions on $\mathbb{R}^n$: 2nd order cond.

Theorem (3.6)

Let $S \subseteq \mathbb{R}^n$ be open convex and $f : S \to \mathbb{R}$ be differentiable. Then TFAE:
(2) $f(y) \geqslant f(x) + \nabla f(x)^\top (y - x)$ for all $x, y \in S$.
(3) $\nabla^2 f(x)$ is positive semidefinite for all $x \in S$.

Proof.

(3) $\implies$ (2): fix any $x, y \in S$, and let $d := y - x$. Define

$$\begin{aligned} & g : [0,1] \to \mathbb{R} : t \mapsto f(x + td). \\ \implies \quad & g''(t) = d^\top \nabla^2 f(x + td) d \geqslant 0 \end{aligned}$$

by positive semidefiniteness of $\nabla^2 f(x + td)$. By Thm 3.4,

$$\begin{aligned} & g(1) \geqslant g(0) + g'(0) \\ \implies \quad & f(y) \geqslant f(x) + \nabla f(x)^\top (y - x). \end{aligned}$$

□

Basic properties of convex functions

Suppose $f : S \to \mathbb{R}$ ($S \subseteq \mathbb{R}^n$ convex) is convex.

+ *Jensen's inequality* holds:

$$f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leqslant \sum_{i=1}^{k} \lambda_i f(x_i) \qquad \forall\, x_i \in S,\ \lambda_i \geqslant 0 \text{ with } \sum_{i=1}^{k} \lambda_i = 1.$$

+ f is continuous on $\text{int}(S)$.

+ If $A : \mathbb{R}^m \to \mathbb{R}^n$ is affine, then
  (1) the preimage $A^{-1}(S) \subseteq \mathbb{R}^m$ is convex, and
  (2) $f \circ A$: $A^{-1}(S) \to \mathbb{R}$ is convex.

+ If $g : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing,
  then $g \circ f$ is convex.
  (The non-decreasing condition *cannot* be removed.)