S. Boyd and L. Vandenberghe - Convex Optimization
Chapter 3: Convex Functions

C. Molinari

PhD student at UTFSM
Internship at INRIA

05 December 2016

Indice

We consider functions from the whole space to the *extended real line*:

$$f: \ \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$$

Definition (Effective domain)

$$\mathbf{dom}\,(f) = \{x \in \mathbb{R}^n : \quad f(x) < +\infty\}$$

Example: Indicator function

Let $C \subset \mathbb{R}^n$ be a set and define the indicator function of C as

$$\delta_C(x) = \begin{cases} 0 & x \in C \\ +\infty & \text{else;} \end{cases}$$

then

$$\operatorname*{arg\,min}_{x \in C} \ f(x) = \operatorname*{arg\,min}_{x \in \mathbb{R}^n} \ \{ \ f(x) + \delta_C(x) \ \}$$

Definition (Convex/Concave function: Jensen's inequality)

A function $f:\ \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be *convex* if for every $x, y \in \mathbb{R}^n$ and for every $\theta \in [0,1]$ it holds

$$f\left(\theta x + (1-\theta)\, y\right) \leq \theta f(x) + (1-\theta)\, f(y);$$

f is said to be *concave* if $-f$ is convex

Geometrical interpretation of convexity:

The segment between $(x, f(x))$ and $(y, f(y))$ lies above the graph of f

Example: Affine functions

A function is both *convex* and *concave* if and only if is *affine* (i.e.: $f(x) = \langle a,\ x\rangle + b$, for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$)

Proposition (Restriction to lines)

A function f is convex if and only if the function

$$g:\ t \in \mathbb{R} \longmapsto g(t) = f(x + tv)$$

is convex for every $x, v \in \mathbb{R}^n$ and $t \in \mathbb{R}$ such that $x + tv \in \mathbf{dom}(f)$

Extensions of Jensen's inequality

If f is a convex function, then

- for every $(x_k)_k$ in $\mathbf{dom}\,(f)$ and for every $(\theta_k)_k$ positive such that $\sum_{k=1}^{K} \theta_k = 1$:
$$f\left(\sum_{k=1}^{K} \theta_k x_k\right) \leq \sum_{k=1}^{K} \theta_k f\,(x_k)\,;$$
- for every $S \subseteq \mathbf{dom}\,(f)$ and $p(x) \geq 0$ on S such that $\int_S p(x)\ dx = 1$:
$$f\left(\int_S p(x)\ x\ dx\right) \leq \int_S p(x) f(x)\ dx$$

Proposition

A function f is convex if and only if $f\left(\mathbb{E}\left[x\right]\right) \leq \mathbb{E}\left[f(x)\right]$ for every random variable x such that $x \in \mathbf{dom}\,(f)$ with probability one

Indice

Examples of convex functions

Dimension one:

- $f(x) = e^{ax}$ on $\mathbb{R}$, for any $a \in \mathbb{R}$;
- $f(x) = |x|^p$ on $\mathbb{R}$, for $p \geq 1$;
- $f(x) = -\log(x)$ on $\mathbb{R}_{++}$;
- $f(x) = x\log(x)$ on $\mathbb{R}_{++}$ (called *negative entropy*);

Higher dimension:

- $f(x) = \|x\|$ on $\mathbb{R}^n$, where $\|\cdot\|$ is a generic norm;
- $f(x) = \max(x_1, \ldots, x_n)$ on $\mathbb{R}^n$;
- $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$ on $\mathbb{R}^n$;
- $f(x) = -(x_1 \cdot \ldots \cdot x_n)^{1/n}$ on $\mathbb{R}^n_{++}$ (*geometric mean*)

Proposition (Composition with affine maps)

If f is convex, $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$, then $g(x) = f(Ax + b)$ is convex

Proposition (Non-negative weighted sum and poitwise maximum/supremum)

A. If $(f_k)_k$ are convex functions and $(w_k)_k$ positive weights, then are convex also

$$g(x) = w_1 f_1(x) + \cdots + w_K f_K(x) \quad \text{and}$$
$$h(x) = \max\{f_1(x), \cdots, f_K(x)\};$$

B. if $f(x, y)$ is a convex function for every parameter $y \in \mathcal{A}$ and $w(y) \geq 0$ for every $y \in \mathcal{A}$, then are convex also

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y)\, dy \quad \text{and}$$
$$h(x) = \sup_{y\in\mathcal{A}} f(x, y)$$

Proposition (Composition with non-decreasing functions)

If

- $g_i : \mathbb{R}^n \to \mathbb{R}$ is convex; and
- $h : \mathbb{R}^k \to \overline{\mathbb{R}}$ is convex and non-decreasing in each argument,

then $f = h \circ (g_1, \ldots, g_k)$ is convex

Remark:

If $h : \mathbb{R} \to \overline{\mathbb{R}}$ is convex and non-decreasing, then $\mathbf{dom}\,(h)$ can be only $\mathbb{R}$, $(-\infty, a)$ or $(-\infty, a]$

Proposition (Minimization)

If f is convex in (x, y) and C is a convex non-empty set, then

$$g(x) = \inf_{y \in C} f(x, y)$$

is convex (provided $g(x) > -\infty$ for every $x \in \mathbb{R}^n$)

Example:

The function $\|y - x\|$ is convex in (x, y); then, if $C \subseteq \mathbb{R}^n$ is a convex set, the function

$$\mathbf{dist}_C(x) = \inf_{y \in C} \|y - x\|$$

is also convex

Indice

Convexity

Proposition (First order condition)

Let f be differentiable in $\mathbf{dom}\,(f)$; then the following are equivalent:

- f is a *convex* function: for every $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$
$$f\left(\theta x + (1-\theta)\,y\right) \leq \theta f(x) + (1-\theta)\,f(y);$$
- $\mathbf{dom}\,(f)$ is convex set and for every $x_0, x \in \mathbf{dom}\,(f)$
$$f\,(x) \geq f\,(x_0) + \langle \nabla f(x_0),\ x - x_0 \rangle;$$
- $\mathbf{dom}\,(f)$ is convex set and for every $x_0, x \in \mathbf{dom}\,(f)$
$$\langle \nabla f(x) - \nabla f(x_0),\ x - x_0 \rangle \geq 0$$

Remark: from *local* information to *global* information

If f is *convex* and $\nabla f(x_0) = 0$, then x_0 is a global minimizer of f

Sketch of the proof (1)

- f is convex $\Rightarrow$ $f(x) \geq f(x_0) + \langle \nabla f(x_0),\ x - x_0 \rangle$

  From convexity, we have $f(\theta x + (1-\theta) x_0) \leq \theta f(x) + (1-\theta) f(x_0)$; manipulating, we obtain that for every $\theta \in (0,1)$

  $$f(x) \geq f(x_0) + \frac{f(x_0 + \theta(x - x_0)) - f(x_0)}{\theta};$$

  to conclude, take the limit as $\theta \to 0^+$

Sketch of the proof (2)

- $f(x) \geq f(x_0) + \langle \nabla f(x_0),\ x - x_0 \rangle \ \Rightarrow\ \langle \nabla f(x) - \nabla f(x_0),\ x - x_0 \rangle \geq 0$

  Interchanging the roles of x_0 and x, we have

  $$\begin{cases} f(x) \geq f(x_0) + \langle \nabla f(x_0),\ x - x_0 \rangle & \text{and} \\ f(x_0) \geq f(x) + \langle \nabla f(x),\ x_0 - x \rangle; \end{cases}$$

  summing-up the two inequalities, we obtain the result

Sketch of the proof (3a)

- $\langle \nabla f(x) - \nabla f(x_0),\ x - x_0 \rangle \geq 0 \;\Rightarrow\;$ f is convex

  For $\theta \in [0,1]$, define the function

  $$\phi(\theta) = f(\theta x + (1-\theta) x_0) - \theta f(x) - (1-\theta) f(x_0),$$

  then $\phi(0) = \phi(1) = 0$ and

  $$\phi'(\theta) = \langle \nabla f(\theta x + (1-\theta) x_0),\ x - x_0 \rangle - f(x) + f(x_0);$$

  moreover, for $0 < \theta_1 < \theta_2 < 1$ we have $\phi'(\theta_1) - \phi'(\theta_2) \leq 0$, i.e.: ϕ' is non-decreasing. Then there exists $\bar{\theta} \in (0,1)$ such that $\phi'\left(\bar{\theta}\right) = 0$.

Sketch of the proof (3b)

So ϕ is

- non-increasing in $\left[0, \bar{\theta}\right]$ and
- non-decreasing in $\left[\bar{\theta}, 1\right]$;

then, finally,

$$\phi\left(\theta\right) \leq 0 \qquad \forall \theta \in [0, 1]$$

Proposition (Second order condition)

Let f be twice-differentiable; then f is *convex* if and only if $\mathbf{dom}\,(f)$ is convex and for every $x_0, x \in \mathbf{dom}\,(f)$

$$\langle \nabla^2 f(x_0)\ (x - x_0)\,,\ (x - x_0)\rangle \geq 0;$$

i.e.: $\nabla^2 f(x_0)$ is *positive semi-definite*

Strict convexity

Proposition (First order condition)

Let f be differentiable in $\mathbf{dom}\,(f)$; then the following are equivalent:

- f is *strictly convex*: for every distinct $x, y \in \mathbb{R}^n$ and $\theta \in (0,1)$
$$f\left(\theta x + (1-\theta)\, y\right) < \theta f(x) + (1-\theta)\, f(y);$$
- $\mathbf{dom}\,(f)$ is convex set and for every distinct $x_0, x \in \mathbf{dom}\,(f)$
$$f\,(x) > f\,(x_0) + \langle \nabla f(x_0),\ x - x_0 \rangle;$$
- $\mathbf{dom}\,(f)$ is convex set and for every distinct $x_0, x \in \mathbf{dom}\,(f)$
$$\langle \nabla f(x) - \nabla f(x_0),\ x - x_0 \rangle > 0$$

Remark: from *local* information to *global* information

If f is *strictly convex*, then it has *at most* one (global) minimizer

Proposition (Second order condition)

Let f be twice-differentiable; if $\mathbf{dom}\,(f)$ is convex and $\forall x_0, x \in \mathbf{dom}\,(f)$

$$\langle \nabla^2 f(x_0)\ (x - x_0)\,,\ (x - x_0)\rangle > 0$$

(i.e.: $\nabla^2 f(x_0)$ is *positive definite*), then f is *strictly convex*

Counterexample

$f(x) = x^4$ is strictly convex, but $f''(0) = 0$

Strong convexity

Proposition (First order condition)

Let f be differentiable in $\mathbf{dom}\,(f)$; then the following are equivalent:

- f is α-*strongly convex*: for every $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$
$$f\left(\theta x + (1-\theta)\,y\right) + \frac{\alpha}{2}\theta\,(1-\theta)\,\|y-x\|^2 \leq \theta f(x) + (1-\theta)\,f(y);$$
- $\mathbf{dom}\,(f)$ is convex set and for every $x_0, x \in \mathbf{dom}\,(f)$
$$f\,(x) \geq f\,(x_0) + \langle \nabla f(x_0),\ x - x_0\rangle + \frac{\alpha}{2}\|x - x_0\|^2;$$
- $\mathbf{dom}\,(f)$ is convex set and for every $x_0, x \in \mathbf{dom}\,(f)$
$$\langle \nabla f(x) - \nabla f(x_0),\ x - x_0\rangle \geq \alpha\|x - x_0\|^2$$

Remark:

strong convexity $\Rightarrow$ *strict convexity* $\Rightarrow$ *convexity*

Proposition (Second order condition)

Let f be twice-differentiable; then f is *strongly convex* if and only if $\mathbf{dom}\,(f)$ is convex and for every $x_0, x \in \mathbf{dom}\,(f)$

$$\langle \nabla^2 f(x_0)\ (x - x_0),\ (x - x_0)\rangle \geq \frac{\alpha}{2}\|x - x_0\|^2,$$

i.e.: $\nabla^2 f(x_0)$ is α-*uniformly elliptic*

Indice

Continuity of Convex functions

Theorem

If f is *convex*, then it is continuous at every point of $\textbf{rel int dom}(f)$

Indice

Exercise (a)

- Prove Hölder inequality: for every $x, y \in \mathbb{R}^n$, $p > 1$ and $q \in \mathbb{R}_+$ such that $1/p + 1/q = 1$,

$$\sum_{j=1}^{n} x_j y_j \leq \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} \left(\sum_{j=1}^{n} |y_j|^q \right)^{1/q}$$

  Hint: from convexity of function $f(x) = -log(x)$, it holds that $a^\theta b^{1-\theta} \leq \theta a + (1-\theta)$ for every $a, b \geq 0$ and $\theta \in [0, 1]$

- Show that if $g: \mathbb{R}^n \to \mathbb{R}$ is convex and $h: \mathbb{R} \to \overline{\mathbb{R}}$ is convex and non-decreasing, then $f = h \circ g$ is convex
  - A. in the case that g and h are both twice-differentiable and $\mathbf{dom}\,(g) = \mathbf{dom}\,(h) = \mathbb{R}$;
  - B. in the general case

Exercise (b)

- Let C be a convex set and define the *Minkowski function* as

$$M_C(x) = \inf\left\{t > 0 : \ \frac{x}{t} \in C\right\}.$$

A. What is $\mathbf{dom}\,(M_C)$?

B. Show that M_C is

- homogeneous, i.e. that for every $\alpha \geq 0$ it holds $M_C(\alpha x) = \alpha M_C(x)$; and
- convex

C. Suppose that C is also closed, bounded, symmetric (if $x \in C$, then $-x \in C$) and has non-empty interior; show that M_C is a norm. What is the corresponding unit ball, i.e. $\mathbf{sub}_\gamma\,(M_C)$?

Exercise (c)

- For $x \in \mathbb{R}^n$ we denote by $x_{[i]}$ the i-th largest component of x, i.e.

  $$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]};$$

  show that the sum of the r largest elements of x, i.e.

  $$f(x) = \sum_{i=1}^{r} x_{[i]},$$

  is a convex function.

  Hint: notice that f can be re-written as the maximum of all possible sums of r different components of x, i.e.

  $$f(x) = \max\left\{x_{i_1} + \cdots + x_{i_r} : \ 1 \leq i_1 < \cdots < i_r \leq n\right\}$$