Modern Optimization

Jean-Luc Bouchot

School of Mathematics and Statistics
Beijing Institute of Technology
jlbouchot@bit.edu.cn

Spring 2021

1. Convex functions and analysis– review

Outline

Proposition 2.1 (Cauchy Schwarz)

Remember the Cauchy-Schwarz inequality: let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$*, we have*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\mathbf{v}\|.$$

This allows, if needed, to define an angle between the two vectors (assuming non zero vectors):

$$\cos(\alpha) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

Definition 2.1 (Convex sets)

Let $C \subseteq \mathbb{R}^d$*. We say that* C *is* **convex** *if for any two points* $x, y \in C$*, for any* $\lambda \in [0, 1]$*,*

$$\lambda x + (1 - \lambda) y \in C.$$

Let's doodle some examples.

Remark 2.1

Let $\mathcal{I}$ be a countable index set and $\{C_i\}_{i\in\mathcal{I}}$ a family of convex sets. Then

$$\cap_{i\in\mathcal{I}} C_i$$

is convex.

Definition 2.2 (Graph-Epigraph)

Let $f : \mathrm{dom}(f) \to \mathbb{R}$*.*

- *The* **graph** *of* f *is the set of points* $\{(x, f(x)), x \in \mathrm{dom}(f)\} \subset \mathbb{R}^{d+1}$*.*
- *The* **epigraph** *of* f *is the set of points above the graph:*

$$\mathrm{epi}(f) = \{(x, y), x \in \mathrm{dom}(f), y \geq f(x)\}.$$

Let's see what this looks like...

Definition 2.3 (Convex function)

Let $f : \mathrm{dom}(f) \to \mathbb{R}$. *We say that* f *is* **convex** *if:*

1. $\mathrm{dom}(f)$ *is convex, and*
2. *for all* $x, y \in \mathrm{dom}(f)$, *for all* $\lambda \in [0,1]$, *we have*

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

Example 2.1

Let $f(x) = a^T x + c$ for a given vector $a \in \mathbb{R}^d$ and $c \in \mathbb{R}$. f is convex.

Remark 2.2

The previous example is very specific in that it is convex with, in fact, equality instead of inequality.
If a function has a strict inequality, we will talk about strict convexity (albeit limiting to $\lambda \in (0,1)$). We will come back to this in the near future.

Example 2.2

Let $Q \in \mathbb{R}^{d\times d}$ be a positive definite matrix and define $f(x) = x^T Qx$. f is a convex function.

Proposition 2.2

Let $f : \mathrm{dom}(f) \to \mathbb{R}$. f *is convex* $\Leftrightarrow \mathrm{epi}(f)$ *is convex.*

Proposition 2.3 (Jensen's inequality)

Let $f : \mathrm{dom}(f) \to \mathbb{R}$ *be a convex function. Let* $x_1, \cdots, x_n$ *be* n *points in* $\mathrm{dom}(f) \subset \mathbb{R}^d$ *and let* $\lambda_1, \cdots, \lambda_n$ *be* n *nonnegative numbers such that* $\sum \lambda_i = 1$*. Then*

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

Proposition 2.4

Let f be a convex function on the open set $\mathrm{dom}(f)$. Then f is continuous.

Definition 2.4 (Lipschitz continuity)

A function $f : \mathrm{dom}(f) \to \mathbb{R}$ *is said to be* **Lipschitz continuous** *with Lipschitz constant* L *(sometimes expressed as* L*-Lipschitz or, if context is clear, simply* f *is Lipschitz) if*

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathrm{dom}(f).$$

Theorem 2.1

Let $f : \mathrm{dom}(f) \to \mathbb{R}^m$ be a differentiable function and let $X \subset \mathrm{dom}(f)$ be an open convex set. Let $L \in \mathbb{R}^+$. The following statements are equivalent:

1. *f is L-Lipschitz.*
2. *The differentials of f are bounded by L, i.e.*

$$\|Df(x)\| \leq L, \quad \forall x \in X,$$

where D denotes the differential operator (or Jacobian), defined as the unique operator A such that for all y in a neighbourhood of x, we have

$$f(y) = f(x) + A(y - x) + r(y - x),$$

with

$$\lim_{v \to 0} \frac{\|r(v)\|}{\|v\|} = 0.$$

Example 2.3 (in-class)

Consider all the assumptions from the previous theorem except that X is just convex. What can be said about the conclusion?

Theorem 2.2

Let $f : \mathrm{dom}(f) \to \mathbb{R}$ *with* $\mathrm{dom}(f)$ *open. Assume furthermore that* f *is differentiable on* $\mathrm{dom}(f)$*. Then* f *is convex if and only if*

- $\mathrm{dom}(f)$ *is convex and*
- *the inequality* $f(y) \geq f(x) + \nabla f(x)^T(y - x)$ *holds for all* $x, y \in \mathrm{dom}(f)$*.*

Example 2.4

Use the first order condition to show the convexity of $f : \mathbb{R}^2 \to \mathbb{R}$ defined as $f(x_1, x_2) = x_1^2 + x_2^2$.

Theorem 2.3

Let f be a twice continuously differentiable function. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and for all $x \in \mathrm{dom}(f) \subseteq \mathbb{R}^n$ the Hessian $H_f(x) = \nabla^2 f(x)$ is positive semidefinite.

Lemma 1

A function $f : \mathrm{dom}(f) \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex if and only $g_{x,y}(t) := f(x + ty)$ is (univariate) convex as a function of f.
Note that the domain of g is dependent on the variables $x, y \in \mathrm{dom}(f)$.

Example 2.5

The negative entropy function f defined as

$$f : \begin{cases} \mathbb{R}_{>0} & \to & \mathbb{R} \\ x & \mapsto & x\log(x) \end{cases}$$

is a convex function.

Example 2.6

Let f be the function defined as

$$f : \begin{cases} \mathbb{R} \times \mathbb{R}_{>0} & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & \frac{x^2}{y}. \end{cases}$$

f is convex.

Definition 2.5 (Convex hull)

For a set C, its convex hull is defined as the set of all convex combinations of points in C, i.e.

$$\mathrm{conv}(C) := \{y = \sum_{i=1}^{k} \alpha_i x_i \text{ for some } k \in \mathbb{N}, x_1, \cdots, x_k \in C, \alpha_1, \cdots, \alpha_k \geq 0, \sum \alpha_i = 1$$

Definition 2.6 (Cones)

A set C is called a **cone** *if it is nonnegative homogeneous, i.e.*

$$\forall \alpha \geq 0, x \in C, \alpha x \in C.$$

Definition 2.7 (Convex cones)

A set C is called a **convex cone** *if it is a cone and convex, i.e.*

$$\forall \alpha \geq 0, \beta \geq 0, x, y \in C, \alpha x + \beta y \in C.$$

Example 2.7

Let $\Sigma_s^n := \{x \in \mathbb{R}^n : \|x\|_0 := \#\{i : x_i \neq 0\} \leq s\}$ be the set of s sparse vectors in $\mathbb{R}^n$ with $s \leq n$. Σ_s^n is a cone but not a convex cone.

Proposition 2.5

Let f be a convex function in two variables x and y and let C denote a non empty set. Then,

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

is convex provided $g(x) > \infty$.
Moreover, its domain is defined as the projection along a coordinate of C:

$$\mathrm{dom}(g) = \{x : \exists y \in C, (x, y) \in \mathrm{dom}(f)\}.$$