The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381K: Convex Optimization — Fall 2019

Lecture 10

Aryan Mokhtari Wednesday, October 2, 2019.

Goal: In this lecture we first study operations that preserve the convexity property of a set. Then, we talk about convex functions.

1 Operations that preserve convexity

Several operations preserve convexity. Here are a few:

- **Intersection:** If $\mathcal{C}_1$ and $\mathcal{C}_2$ are two convex sets, then $\mathcal{C}_1 \cap \mathcal{C}_1$ remains convex. In fact, the same holds true for arbitrary intersections of convex sets.

  Proof. Let $\mathbf{x}_1$, $\mathbf{x}_2 \in \mathcal{C}_1 \cap \mathcal{C}_1$, and let $\mathbf{x}$ be a point on the line segment between $\mathbf{x}_1$ and $\mathbf{x}_2$. Then, $\mathbf{x}$ lies in both $\mathcal{C}_1$ & $\mathcal{C}_2$ as they are convex. Consequently, $\mathbf{x} \in \mathcal{C}_1 \cap \mathcal{C}_1$. □

- **Cartesian product:** If $\mathcal{C}_1$ and $\mathcal{C}_2$ are two convex sets, then their cartesian product is convex, i.e.
$$\mathcal{C}_1 \times \mathcal{C}_2 = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} : \mathbf{x} \in \mathcal{C}_1, \mathbf{y} \in \mathcal{C}_2 \right\}$$
  is convex.

  Proof. Let $\mathbf{u}, \mathbf{v} \in \mathcal{C}_1 \times \mathcal{C}_2$ where $\mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}$, and let $\mathbf{w} = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix} = \lambda \mathbf{u} + (1-\lambda)\mathbf{v}$. Now, as $\mathbf{w}_1 \in \mathcal{C}_1$ and $\mathbf{w}_2 \in \mathcal{C}_2$, then $\mathbf{w} \in \mathcal{C}_1 \times \mathcal{C}_2$.

  □

- **Affine function:** Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is an affine function (i.e., can be written as $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ with $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then, the image of a convex set $\mathcal{C}$ under f is a convex set, i.e.,
$$\mathcal{S} \subset \mathbb{R}^n \;\text{ is convex } \Rightarrow \; f(\mathcal{S}) = \{\mathbf{y} = f(\mathbf{x}) \in \mathbb{R}^m | \mathbf{x} \in \mathcal{S}\} \;\text{ is convex}$$

Proof. Let $\mathbf{y}_1, \mathbf{y}_2 \in f(\mathcal{S})$. Then, we know that there exist $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ such that $\mathbf{y}_1 = f(\mathbf{x}_1)$ and $\mathbf{y}_2 = f(\mathbf{x}_2)$. By convexity of the set $\mathcal{S}$ we know that $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in \mathcal{S}$ for $\lambda \in [0, 1]$. Therefore, $\lambda\mathbf{y}_1 + (1-\lambda)\mathbf{y}_2$ also belongs to the set $f(\mathcal{S})$ as we know that it can be written as $\lambda\mathbf{y}_1 + (1-\lambda)\mathbf{y}_2 = f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)$ and $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in \mathcal{S}$. □

2 Convex functions

We give three definitions of convex functions: A basic definition that requires no differentiability of the function; a first-order definition of convexity that uses the gradient; and a second-order condition of convexity. We show that (for smooth functions) these three are equivalent. Later in the course, it will be important to deal with non-differentiable functions.
First, we define the set of points where a function is finite.

Definition 1 (Domain of a function)**.** *The domain of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is denoted* $\mathrm{dom}(f)$, *and is defined as the set of points where a function* f *is finite:*

$$\mathrm{dom}(f) = \{\mathbf{x} \in \mathbb{R}^n \ : \ f(\mathbf{x}) < \infty\}.$$

Definition 2 (Convexity I)**.** *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is convex if* $\mathrm{dom}(f)$ *is convex and* for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathrm{dom}(f) \subseteq \mathbb{R}^n$, $\lambda \in [0,1]$, *we have:*

$$f\left(\lambda \mathbf{x}_1 + (1-\lambda)\,\mathbf{x}_2\right) \ \leq \ \lambda f(\mathbf{x}_1) + (1-\lambda)\, f(\mathbf{x}_2). \tag{1}$$

This inequality is illustrated in Figure 1.

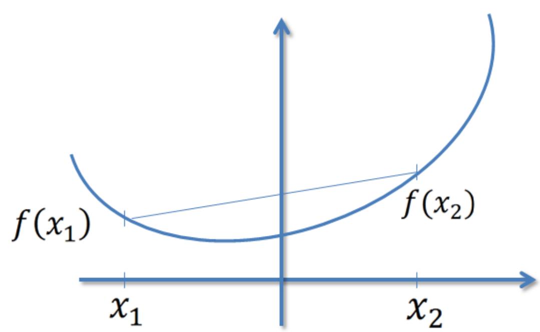

Figure 1: Convex functions

Remark 1. *We say that* f *is concave if* $-f$ *is convex.*

Definition 3 (Strict Convexity)**.** *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is strictly convex if* $\mathrm{dom}(f)$ *is convex and* for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathrm{dom}(f) \subseteq \mathbb{R}^n$ *where* $\mathbf{x}_1 \neq \mathbf{x}_2$ *and* $\lambda \in (0,1)$, *we have:*

$$f\left(\lambda \mathbf{x}_1 + (1-\lambda)\,\mathbf{x}_2\right) \ < \ \lambda f(\mathbf{x}_1) + (1-\lambda)\, f(\mathbf{x}_2). \tag{2}$$

Show that the following functions are convex:

$$f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + b$$

$$f(\mathbf{x}) = \|\mathbf{x}\|_p$$

$$f(\mathbf{X}) = \mathbf{tr}(\mathbf{A}^\top \mathbf{X}) + b = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} X_{ij} + b$$

Theorem 1. *A convex function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(t) = f(\mathbf{x} + t\mathbf{y})$ is convex (as a univariate function) for all $\mathbf{x}$ in domain of f and all $\mathbf{y} \in \mathbb{R}^n$.*

The theorem simplifies many basic proofs in convex analysis.
Allows us to check convexity of a function by checking convexity of functions of one variable.
Example: Consider the function $f(\mathbf{X}) = -\log\det(\mathbf{X})$ where $dom f = \mathbf{S}^n_{++}$. Consider $\mathbf{X} \in \mathbf{S}^n_{++}$ and $\mathbf{V} \in R^{n\times n}$ and the function $g : \mathbb{R} \to \mathbb{R}$ defined as $g(t) = f(\mathbf{X} + t\mathbf{V})$. Then, we have

$$\begin{aligned}
g(t) &= f(\mathbf{X} + t\mathbf{V}) \\
&= -\log\det(\mathbf{X} + t\mathbf{V}) \\
&= -\log\det(\mathbf{X}(\mathbf{I} + t\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2})) \\
&= -\log\det(\mathbf{X}) - \log\det(\mathbf{I} + t\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}) \\
&= -\log\det(\mathbf{X}) - \log\prod_{i=1}^{n}(1 + t\lambda_i) \\
&= -\log\det(\mathbf{X}) - \sum_{i=1}^{n}\log(1 + t\lambda_i),
\end{aligned}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}$. Hence, for any $\mathbf{X} \in \mathbf{S}^n_{++}$ and $\mathbf{V} \in R^{n\times n}$ the function $g(t)$ is convex with respect to t as $-\log$ is a convex function.