W. DAI

EE4.66 LARGE DATA PROCESSING EXERCISE

IMPERIAL COLLEGE LONDON

Contents

1

Convexity

1.1 Convex Set

1.

(a) Let A, B be two convex sets. Define $C := A + B = \{\boldsymbol{a} + \boldsymbol{b} |\ \boldsymbol{a} \in A,\ \boldsymbol{b} \in B\}$. Show that C is convex.

(b) Let A_k, $k = 1, 2, \cdots, K$, be convex sets. Show that $A := \bigcap_{k=1}^{K} A_k$ is convex.

(c) Show that a set is convex if and only if its intersection with any line is convex.

2. This question is about the distance between two parallel hyperplanes $\left\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}^\mathsf{T} \boldsymbol{x} = b_1\right\}$ and $\left\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}^\mathsf{T} \boldsymbol{x} = b_2\right\}$.

(a) Give a scientifically reasonable definition for the distance.

(b) Find the distance in a closed form.

3. For each of the following sets, indicate whether it is convex or not and prove your answer.

(a) A slab, i.e., $\{\boldsymbol{x} \in \mathbb{R}^n | \alpha \leq \boldsymbol{a}^\mathsf{T} \boldsymbol{x} \leq \beta\}$.

(b) A rectangle, i.e., $\{\boldsymbol{x} \in \mathbb{R}^n | \alpha_i \leq x_i \leq \beta_i\}$.

(c) The set of points closer to a given point $\boldsymbol{x}_0$ than to another given point $\boldsymbol{y}$, i.e.,

$$\{\boldsymbol{x} | \|\boldsymbol{x} - \boldsymbol{x}_0\|_2 \leq \|\boldsymbol{x} - \boldsymbol{y}\|_2\}.$$

(d) The set of points closer to a given point than to a given set, i.e.,

$$\{\boldsymbol{x} | \|\boldsymbol{x} - \boldsymbol{x}_0\|_2 \leq \|\boldsymbol{x} - \boldsymbol{y}\|_2,\ \forall \boldsymbol{y} \in S\}$$

where $S \subseteq \mathbb{R}^n$.

(e) The set of points closer to one set than another, i.e.,

$$\{\boldsymbol{x} | \text{dist}(\boldsymbol{x}, S) \le \text{dist}(\boldsymbol{x}, T)\}$$

where

$$\text{dist}(x, S) := \inf\{\|\boldsymbol{x} - \boldsymbol{z}\|_2 \ \mid \ \boldsymbol{z} \in S\}.$$

(f) The set of points whose distance to $\boldsymbol{a}$ does not exceed a fixed fraction θ of the distance to $\boldsymbol{b}$, i.e., the set $\{\boldsymbol{x} \mid \ \|\boldsymbol{x} - \boldsymbol{a}\|_2 \le \theta\|\boldsymbol{x} - \boldsymbol{b}\|_2\}$, where $\boldsymbol{a} \neq \boldsymbol{b}$ and $0 \le \theta \le 1$.

1.2 *Convex Functions*

1. Use the second-order condition of convexity to prove the following functions are convex

   (a) $f(x) = -\log(x)$ where $x \in \mathbb{R}^+$.

   (b) $f(x) = x\log(x)$ where $x \in \mathbb{R}^+$.

   (c) Affine functions $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$.

   (d) Quadratic functions $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\mathsf{T}\boldsymbol{x} + c$ where $\boldsymbol{A} \succeq 0$.

   (e) $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|_2^2$.

2.

   (a) Prove that if $f(\boldsymbol{x})$ is convex, then $g(\boldsymbol{x}) := f(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$ is convex.

   (b) Prove that any norm $\|\cdot\|$ on $\mathbb{R}^m$ is convex.

   (c) Let

$$f(\boldsymbol{x}) := \max_{i=1,\cdots,k} \|\boldsymbol{A}^{(i)}\boldsymbol{x} + \boldsymbol{b}^{(i)}\|,$$

   where $\boldsymbol{A}^{(i)} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b}^{(i)} \in \mathbb{R}^m$, and $\|\cdot\|$ is a norm on $\mathbb{R}^m$. Indicate whether $f(\cdot)$ is convex or not. Prove your claim.

   (d) Let

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

   where $|x|_{[i]}$ is the ith largest component of $|x_1|, \cdots, |x_n|$. Decide whether $f(\cdot)$ is convex or not. Prove your claim.

1.3 *Line Search*

1. Let $\boldsymbol{x} \in \mathbb{R}^2$ and

$$f(\boldsymbol{x}) = 3|x_1| + |x_2|.$$

   Consider the point $\boldsymbol{y} = [0, 1]^\mathsf{T}$.

(a) Show that $\boldsymbol{g} = [3, 1]^\top \in \partial f(\boldsymbol{y})$.

(b) Let $\tau \in (0, 1)$, find the closed form for

$$f(\boldsymbol{y} - \tau \boldsymbol{g}).$$

(c) Comment on whether $-\boldsymbol{g}$ is a descent direction or not.

2. The following is the famous Wolfe's example, which shows that gradient descent method may not converge to a local optimal point.

Let $\boldsymbol{x} \in \mathbb{R}^2$ and

$$f(\boldsymbol{x}) = \begin{cases} 5(9x_1^2 + 16x_2^2)^{1/2} & \text{if } x_1 > |x_2|, \\ 9x_1 + 16|x_2| & \text{if } x_1 \le |x_2|. \end{cases}$$

Suppose that $\boldsymbol{x}^0 = [16/9, 1]^\top$. Consider exact line search where

$$\boldsymbol{x}^{l+1} = \boldsymbol{x}^l - t^{l+1} \nabla f(\boldsymbol{x}^l),$$

where

$$t^{l+1} = \arg\min_t \; f(\boldsymbol{x}^l - t \nabla f(\boldsymbol{x}^l)).$$

(a) Draw the contours of $f(\boldsymbol{x})$ in the region $-2 \le x_1 \le 2$ and $-2 \le x_2 \le 2$.

(b) Is the point $\boldsymbol{x} = [0, 0]^\top$ optimal? Why?

(c) Find the closed form of $\nabla f(\boldsymbol{x})$ in the region where $x_1 > |x_2|$.

(d) Find t^1 and $\boldsymbol{x}^1$.

(e) Find t^2 and $\boldsymbol{x}^2$.

(f) Use mathematical induction, find t^l and $\boldsymbol{x}^l$. It can be concluded that $\boldsymbol{x}^l \to [0, 0]^\top$ as $l \to \infty$.

2
Bibliography