Convex Optimization — DM1

ZHAO Tong

tong.zhao@eleves.enpc.fr

2.12 Which of the following sets are convex?

(a) A slab is a convex set.

Proof: Suppose $x_1, x_2 \in \mathbb{R}^n$ satisfying $\alpha \le a^T x_1 \le \beta$, $\alpha \le a^T x_2 \le \beta$

for $\forall \gamma \in [0,1]$, we have $a^T[\gamma x_1 + (1-\gamma)x_2] = \gamma a^T x_1 + (1-\gamma) a^T x_2$

so we have $\gamma a^T x_1 + (1-\gamma) a^T x_2 \geq \gamma\alpha + (1-\gamma)\alpha = \alpha$

$\gamma a^T x_1 + (1-\gamma) a^T x_2 \le \gamma\beta + (1-\gamma)\beta = \beta$

so we conclude that all the linear combination belongs to the slab set, thus it is convex.

(b) A rectangle is a convex set.

Proof: Suppose $x, y \in \mathbb{R}^n$, $\alpha_i \le x_i \le \beta_i$, $\alpha_i \le y_i \le \beta_i$, $i = 1, \dots, n$

for $\forall \gamma \in [0,1]$, we have $\underbrace{\gamma\alpha_i + (1-\gamma)\alpha_i}_{=\alpha_i} \le \gamma x_i + (1-\gamma) y_i \le \underbrace{\gamma\beta_i + (1-\gamma)\beta_i}_{=\beta_i}$

since all the linear combination belongs to the set, we know that it is convex

(c) A wedge is a convex set

Proof: Suppose $x_1, x_2 \in \mathbb{R}^n$, $a_1^T x_1 \le b_1$, $a_2^T x_1 \le b_2$, $a_1^T x_2 \le b_1$, $a_2^T x_2 \le b_2$

for $\forall \gamma \in [0,1]$, we have $a_1^T(\gamma x_1 + (1-\gamma)x_2) \le \gamma b_1 + (1-\gamma) b_1 = b_1$

$a_2^T(\gamma x_1 + (1-\gamma)x_2) \le \gamma b_2 + (1-\gamma) b_2 = b_2$

so we conclude that it is convex.

(d) The set of points closer to a given point than a given set is a convex set.

Proof: For every $y \in S$, we have $\|x - x_0\|_2 \le \|x - y\|_2$, which is a halfspace, thus it is convex. So our set can be expressed as $\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \le \|x - y\|_2\}$. We know that given any collection of convex sets, their intersection is a convex set. So we conclude that the set is convex.

(e) The set of points closer to one set than another may not be a convex set.

Counter-example: we can construct a set T by taking the midpoint of two points $s_1, s_2 \in S$, then we take $x_1 = s_1$ and $x_2 = s_2$, since $0.5x_1 + 0.5x_2 \in T$, the set can not be convex.

(f) The set $\{x \mid x+S_2 \subseteq S_1\}$ where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 convex is convex.

Proof: Suppose x_1, x_2 satisfying $x_1+S_2 \subseteq S_1$, $x_2+S_2 \subseteq S_1$.

For $\forall s \in S_2$, $\forall \gamma \in [0,1]$, we have $\gamma x_1+(1-\gamma)x_2+s = \gamma\underbrace{(x_1+s)}_{\in S_1} + (1-\gamma)\underbrace{(x_2+s)}_{\in S_1}$

Since S_1 is convex, we know $\gamma x_1+(1-\gamma)x_2+s \in S_1$, so the set $\{x \mid x+s \subseteq S_1\}$ is convex. According to the theorem we mentioned before, we know $\bigcap_{s\in S_2}\{x \mid x+s \subseteq S_1\}$ $= \{x \mid x+S_2 \subseteq S_1\}$ is convex.

(g) The set $\{x \mid \|x-a\|_2 \le \theta\|x-b\|_2\}$ is a convex set.

Proof: The inequality can be expressed as:

$$(x-a)^T(x-a) \le \theta^2(x-b)^T(x-b)$$

$$x^Tx - 2a^Tx + a^Ta \le \theta^2x^Tx - 2\theta^2b^Tx + \theta^2b^Tb$$

$$\underbrace{(1-\theta^2)}_{\ge 0}x^Tx - 2(a^Tx - \theta^2b^Tx) + \underbrace{(a^Ta - \theta^2b^Tb)}_{\text{const}} \le 0$$

If $\theta = 1$, the set is a half space, thus is convex.

If $\theta \in [0,1)$, the set is a ball space, thus is also convex.

3.21 Show that the following functions $f: \mathbb{R}^n \to \mathbb{R}$ are convex.

(a) $f(x) = \max_{i=1,\dots,k}\|A^{(i)}x - b^{(i)}\|$ where $A^{(i)} \in \mathbb{R}^{m\times n}$, $b^{(i)} \in \mathbb{R}^m$

Proof: First we prove that $f_i(x) = A^{(i)}x - b^{(i)}$

Suppose $\alpha \in [0,1]$, we have $f_i(\alpha x_1 + (1-\alpha)x_2) = A^{(i)}(\alpha x_1 + (1-\alpha)x_2) - b^{(i)}$

$= \alpha[A^{(i)}x_1 - b^{(i)}] + (1-\alpha)[A^{(i)}x_2 - b^{(i)}] = \alpha f_i(x_1) + (1-\alpha)f_i(x_2)$,

so it is convex.

Then we prove that the norm of a convex function is convex.

$\|f_i(\alpha x_1 + (1-\alpha)x_2)\| \le \|f_i(\alpha x_1)\| + \|f_i((1-\alpha)x_2)\|$

$= \alpha\|f_i(x_1)\| + (1-\alpha)\|f_i(x_2)\|$

Then we prove that the element-wise function $f = \max_{i=1,\dots,k}(\|f_1(x)\|, \cdots \|f_k(x)\|)$ is convex.

for any i, we have $\|f_i(\alpha x_1 + (1-\alpha)x_2)\| \le \alpha\|f_i(x_1)\| + (1-\alpha)\|f_i(x_2)\|$

$\le \alpha f(x_1) + (1-\alpha)f(x_2)$

so $\max(\|f_1(\overset{\alpha x_1+(1-\alpha)x_2}{\uparrow})\|, \cdots, \|f_k(\overset{\alpha x_1+(1-\alpha)x_2}{\uparrow})\|) \le \alpha f(x_1) + (1-\alpha)f(x_2)$

(b) $f(x) = \sum_{i=1}^{r} |x|_{[i]}$ on $\mathbb{R}^n$, where $|x|$ denotes the vector with $|x|_i = |x_i|$

Proof: Suppose $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, $\alpha \in [0,1]$

$$f(\alpha x_1 + (1-\alpha)x_2) = \sum_{i=1}^{r} |\alpha x + (1-\alpha) y|_{[i]} \quad (1)$$

We take $\{d_1, d_2, \cdots, d_r\}$ indicate the indices of the chosen components.

$$(1) = \sum_{i=1}^{r} |\alpha x_{d_i} + (1-\alpha) y_{d_i}| \le \alpha \sum_{i=1}^{r} |x_{d_i}| + (1-\alpha) \sum_{i=1}^{r} |y_{d_i}|$$

$$\le \alpha f(x) + (1-\alpha) f(y)$$

so the function is convex.

3.32 Prove the following:

(a) If f and g are convex, both nondecreasing, and positive functions on an interval, then fg is convex.

Proof: Suppose $x_1 < x_2$, for $\forall \gamma \in [0,1]$, we have

$$f(\gamma x_1 + (1-\gamma)x_2)\, g(\gamma x_1 + (1-\gamma)x_2) \le (\gamma f(x_1) + (1-\gamma) f(x_2))(\gamma g(x_1) + (1-\gamma) g(x_2))$$

$$= \gamma^2 f(x_1) g(x_1) + \gamma(1-\gamma)(f(x_1) g(x_2) + f(x_2) g(x_1)) + (1-\gamma)^2 g(x_1) g(x_2) \quad (1)$$

we have that

$$\gamma^2 f(x_1) g(x_1) + \gamma(1-\gamma) f(x_1) g(x_2) + (1-\gamma)^2 g(x_1) g(x_2) - \gamma f(x_1) g(x_1) - (1-\gamma) f(x_2) g(x_2)$$

$$= \gamma(1-\gamma)(f(x_1) g(x_2) + f(x_2) g(x_1) - f(x_1) g(x_1) - f(x_2) g(x_2)) \quad (2)$$

since f and g are both nondecreasing. $f(x_2) \ge f(x_1)$ $\quad g(x_2) \geq g(x_1)$

so $(2) \le 0$, then we have $(1) \le \alpha f(x_1) g(x_1) + (1-\alpha) f(x_2) g(x_2)$

which indicates that fg is convex.

(b) If f and g are concave, positive, with one nondecreasing and the other nonincreasing, then fg is concave

Proof: Suppose $x_1 < x_2$, f is nondecreasing and g is nonincreasing, for $\forall \gamma \in [0,1]$,

$$f(\gamma x_1 + (1-\gamma)x_2)\, g(\gamma x_1 + (1-\gamma)x_2) \ge (\gamma f(x_1) + (1-\gamma) f(x_2))(\gamma g(x_1) + (1-\gamma) g(x_2))$$

$$= \gamma^2 f(x_1) g(x_1) + \gamma(1-\gamma) f(x_1) g(x_2) + \gamma(1-\gamma) f(x_2) g(x_1) + (1-\gamma)^2 g(x_1) g(x_2)$$

$$= (-\gamma(1-\gamma) + \gamma) f(x_1) g(x_1) + \gamma(1-\gamma)(f(x_1) g(x_2) + f(x_2) g(x_1)) + ((1-\alpha) - \alpha(1-\alpha)) g(x_2) f(x_2)$$

$$= \gamma f(x_1) g(x_1) + (1-\gamma) f(x_2) g(x_2) + \alpha(1-\alpha)\underline{((f(x_1) - f(x_2))(g(x_2) - g(x_1)))} \ge 0$$

$$\geq \gamma f(x_1) g(x_1) + (1-\gamma) f(x_2) g(x_2)$$

so $f \cdot g$ is concave.

(c) If f is convex, nondecreasing, and positive, and g is concave, nonincreasing, and positive, then f/g is convex.

Proof: Suppose $x_1 < x_2$, for $\forall \gamma \in [0,1]$, we have

$$\frac{f(\gamma x_1+(1-\gamma)x_2)}{g(\gamma x_1+(1-\gamma)x_2)} \le \frac{\gamma f(x_1)+(1-\gamma)f(x_2)}{\gamma g(x_1)+(1-\gamma)g(x_2)} \quad (1)$$

$$(1) - \frac{\gamma f(x_1)}{g(x_1)} - \frac{(1-\alpha)f(x_2)}{g(x_2)} = \frac{\alpha(1-\alpha)}{c}\Big(\underbrace{f(x_1)g(x_1)g(x_2) + g(x_1)g(x_2)f(x_2) - f(x_1)g^2(x_2) - g^2(x_1)f(x_2)}_{(2)}\Big)$$

where $c = (\alpha g(x_1) + (1-\alpha)g(x_2))\,g(x_1)g(x_2) > 0$.

$$(2) = \overset{\le 0}{(g(x_2)-g(x_1))}\cdot\frac{1}{g(x_2)g(x_1)}\cdot\overset{\ge 0}{\left(\frac{f(x_2)}{g(x_2)} - \frac{f(x_1)}{g(x_1)}\right)} \le 0$$

so we have $(1) \le \dfrac{\gamma f(x_1)}{g(x_1)} + \dfrac{(1-\gamma)f(x_2)}{g(x_2)}$

so f/g is convex.

3.36 Derive the conjugates of the following functions.

(a) Max function. $f(x) = \max_{i=1,\cdots,n} x_i$ on $\mathbb{R}^n$

Proof: We have the function $f^*: \mathbb{R}^n \to \mathbb{R}$, defined as

$$f^*(y) = \sup_{x\in\mathbb{R}^n} (y^Tx - \max_{i=1,\cdots,n} x_i)$$

We denote that the maximum element appears at index k, then we have:

$$f^*(y) = \sup_{x\in\mathbb{R}^n} (\sum_{i\neq k}^{n} y_i x_i + y_k x_k - x_k)$$

For $y_k < 0$, we can pick x_k to be $-\infty$, thus the function is unbounded.

If $y \succeq 0$ and $\sum_i y_i > 1$, we can pick x to be $\infty \mathbb{1}_n$ } the function is unbounded

If $y \succeq 0$ and $\sum_i y_i < 1$, we can pick x to be $-\infty \mathbb{1}_n$

If $y \succeq 0$ and $\sum_i y_i = 1$, y^Tx is a linear combination of x, we have $y^Tx \le \max_{i=1,\cdots,n} x_i$, so $f^*(y) = 0$.

We conclude that $f^*(y) = \begin{cases} 0 & y\succeq 0 \text{ and } \sum_i y_i = 1 \\ \infty & \text{otherwise} \end{cases}$

(b) Sum of largest elements. $f(x) = \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$.

Proof: We defined function $f^*: \mathbb{R}^n \to \mathbb{R}$ as

$$f^*(y) = \sup_{x\in\mathbb{R}^n} (y^Tx - \sum_{i=1}^{r} x_{[i]})$$

$(y_i < 0,\ x_i \to -\infty)$

As the above proof, we know that the function is unbounded when $y \preceq 0$ and when $y \succeq 1$. $(y_i > 1,\ x_i \to +\infty)$

If $\mathbb{1}_n^T y \neq r$, $f^*(y) = \sup_{x\in\mathbb{R}^n} (y^Tx - \sum_{i=1}^{r} x_{[i]})$, we take $x = \mathbb{1}_n t$

so $f^*(y) = \sup(t y^T \mathbb{1}_n - rt) = \sup(tc)$ with $c = \mathbb{1}_n^T y - r$

thus the function is unbounded when $c > 0$ and $t \to \infty$ and when $c < 0$ and $t \to -\infty$

when $\mathbb{1}_n^T y = r$ and $y \succeq 0$ and $y \preceq 1$, $f^*(y)$ is bounded by 0

so we conclude that $f^*(y) = \begin{cases} 0 & y \succeq 0 \text{ and } y \preceq 1 \text{ and } \sum y_i = r \\ \infty & \text{otherwise.} \end{cases}$

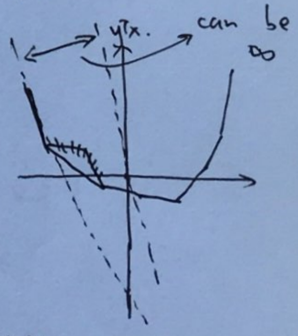
(c) Piecewise-linear function on $\mathbb{R}$. $f(x) = \max_{i=1,\cdots,m} (a_i x + b_i)$

Proof: We define function $f^*: \mathbb{R}^n \to \mathbb{R}$ as

$$f^*(y) = \sup_{x\in\mathbb{R}^n} (y^Tx - \max_{i=1,\cdots,m} (a_i x + b_i))$$

We first sort the slope of the segments in an increasing order, such that

$$a_1 \le a_2 \le \cdots \le a_m$$



We claim that ~~the~~ when $y \le a_1$ or when $y \ge a_m$, $f^*(y) = \infty$

when $a_1 < y < a_m$, $f^*(y)$ is bounded.

For $a_i \le y \le a_{i+1}$, the intersection point is found by

$$a_i x + b_i = a_{i+1} x + b_{i+1} \Rightarrow x = \frac{b_{i+1} - b_i}{a_i - a_{i+1}}$$

So we have $f^*(y) = (y - a_i)\dfrac{b_{i+1} - b_i}{a_i - a_{i+1}} - b_i \quad a_i \le y \le a_{i+1} \quad i \in \{1, \cdots, m-1\}$

and ∞ otherwise.

(14) Power function $f(x) = x^p$ on R_{++}, where $p > 1$. Repeat for $p \leq 0$.

Proof: We define function $f^*(x) = \sup(y^T x - x^p)$ $y \in \mathbb{R}$

$\cancel{f^{*\prime}(y) = x}$ $\quad (y^T x - x^p)' = y - p x^{p-1}$

when $y \leq 0$, $y - p x^{p-1} < 0$ on $\mathbb{R}_{++}$ so, the function has a sup 0.

when $y > 0$, the function is at first increasing and then decreasing.

$$\text{let } y - p x^{p-1} = 0 \text{ we have } x = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$$

$$y^T x - x^p = \frac{y^{\frac{p}{p-1}}}{p^{\frac{1}{p-1}}} - \frac{y^{\frac{p}{p-1}}}{p^{\frac{p}{p-1}}} = (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}$$

$$\text{so } f^*(x) = \begin{cases} 0 & y \leq 0 \\ (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}} & y > 0 \end{cases}$$

(15) Negative geometric mean $f(x) = -(\prod x_i)^{\frac{1}{n}}$ on $\mathbb{R}^n_{++}$

Proof: We define function $f^*(x) = \sup(y^T x + (\prod x_i)^{\frac{1}{n}})$ $\cancel{y \in \mathbb{R}}$

when $y \geq 0$, we take $x_k = \infty$, $f^*(x) \to \infty$ thus it is not bounded.

when $(\prod_i |y_i|)^{\frac{1}{n}} \leq \frac{1}{n}$, and $x_i = -\frac{t}{y_i}$, we have

$$y^T x + (\prod x_i)^{\frac{1}{n}} = -tn + t\left(\prod_i \left(-\frac{1}{y_i}\right)\right)^{\frac{1}{n}} \xrightarrow{t \to \infty} \infty$$

when $(\prod_i |y_i|)^{\frac{1}{n}} > \frac{1}{n}$,

$$\frac{y^T x}{n} \geq \left(\prod_i (-y_i x_i)\right)^{\frac{1}{n}} \geq \frac{1}{n}\left(\prod_i x_i\right)^{\frac{1}{n}}$$

i.e. $-x^T y \geq -f(x)$ with equality for $x \to 0$, Hence $f^*(y) = \begin{cases} 0 & y < 0,\ (\prod_i (y_i))^{\frac{1}{n}} > \frac{1}{n} \\ \infty & \text{otherwise} \end{cases}$

(16) Negative generalized logarithm for second-order cone. $f(x,t) = -\log(t^2 - x^T x)$ on $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\|_2 < t\}$

Proof: ~~We define~~ $f^*(y,u) = -2 + 2\log 2 - \log(u^2 - y^T y)$ $\quad \mathrm{dom} f^* = \{(y,u) \mid \|y\|_2 < -u\}$

we define $f^*(y,u) = ut + y^T x + \log(t^2 - x^T x)$

Suppose $\|y\|_2 \geq -u$, choose $x = sy$, $t = \|x\|_2 + 1 \geq s\|y\|_2 \geq -su$, with $s \geq 0$, Then

$$y^T x + tu \geq s y^T y - s u^2 = s(y^T y - u^2) \geq 0$$

$$\log(t^2 - x^T x) = \log(2s\|y\|_2 + 1)$$

therefore $y^T x + tu + \log(t^2 - x^T x)$ is unbounded above

Assume that $\|y\|_2 < -u$, calculate the gradient respect to x and t

$$x = \frac{2y}{u^2 - y^T y} \qquad t = -\frac{-2u}{u^2 - y^T y}$$

This gives $f^*(y,u) = ut + y^T x + \log(t^2 - x^T x) = -2 + \log 4 - \log(y^T y - u^2)$

when $\mathrm{dom} f^* = \{(y,u) \mid \|y\|_2 < -u\}$, ∞ otherwise.