Convex Optimization
Homework 1

Quang Anh NGUYEN

October 2021

1

1.1

A rectangle is convex as intersection of half spaces, which are themselves convex.

1.2

Since $x \mapsto \frac{1}{x}$ is convex on $\mathbb{R}_++$, its epigraph $\{(x_1, x_2) \in \mathbb{R}^2_{++} : x_2 \geq \frac{1}{x_1}\}$ is convex. Hence, $\{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 x_2 \geq 1\}$ is a convex set.

1.3

For $y \in S$, $\{\|x - x_0\|_2 \leq \|x - y\|_2\}$ is convex as a half space. Therefore, their intersection $\{\|x - x_0\|_2 \leq \sup_{y \in S} \|x - y\|_2\} = \cap_{y \in S}\{\|x - x_0\|_2 \leq \|x - y\|_2\}$ is also convex.

1.4

Consider in $\mathbb{R}^2$ the following sets: $S = \{(-1, 1), (-1, -1)\}$ and $T = \{(1, 0)\}$.

Let $A = \{x \in \mathbb{R}^2 : \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\}$ and $u = (0, 1), v = (0, -1)$.

It can be seen that $u, v \in A$ because $\operatorname{dist}(u, S) = \operatorname{dist}(v, S) = 1$ and $\operatorname{dist}(u, T) = \operatorname{dist}(v, T) = \sqrt{2} \geq 1$. However $0 = \frac{u+v}{2} \notin A$ because $\operatorname{dist}(0, S) = \sqrt{2} > 1 = \operatorname{dist}(0, T)$. Thus A is not convex.

1.5

For $y \in S_2$, $\{x : x + y \in S_1\}$ is the inverse image of S_2 under the affine application $x \mapsto x + y$. Since S_2 is convex, this set is also convex for each y.

Consequently, their intersection $\{x : x + S_1 \subset S_1\} = \cap_{y \in S_2}\{x : x + y \in S_1\}$ is also convex.

2

2.1

The Hessian matrix of f on $\mathbb{R}^2_{++}$ is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is neither semi positive nor semi negative definite. Then f is neither convex nor concave.

f is not quasiconvex because $\{x \in \mathbb{R}^2_{++} : x_1 x_2 \leq 1\}$ is not convex. For instance $(2, 1/2), (1/2, 2)$ is in this set but their midpoint $(5/4, 5/4)$ is not.

Similarly to 1.2, $\{x \in \mathbb{R}^2_{++} : x_1 x_2 \geq \alpha\}$ is convex for all $\alpha > 0$ or equals to $\mathbb{R}^2_{++}$ if $\alpha \leq 0$, which is also convex. Thus, f is quasiconcave.

2.2

The Hessian matrix of f on $\mathbb{R}^2_{++}$ is $\frac{1}{x_1^2 x_2^2} \begin{bmatrix} 2x_2/x_1 & 1 \\ 1 & 2x_1/x_2 \end{bmatrix}$, which is semi positive definite. Then f is convex.

Consider only the non trivial case $\alpha > 0$, since $\frac{1}{x_1 x_2} \leq \alpha \Leftrightarrow x_1 x_2 \geq \frac{1}{\alpha}$, we know that the α-sublevel of f is convex from the previous question, thus f is quasiconvex.

Similarly, f is not quasiconcave because $\{x \in \mathbb{R}^2_{++} : \frac{-1}{x_1 x_2} \leq -1\} = \{x \in \mathbb{R}^2_{++} : x_1 x_2 \leq 1\}$ is not convex.

2.3

The Hessian of f on on $\mathbb{R}^2_{++}$ is $\frac{1}{x_2^2} \begin{bmatrix} 0 & -1 \\ -1 & 2x_1/x_2 \end{bmatrix}$, which is neither semi positive nor semi negative negative definite. Then f is neither convex nor concave.

However, for $\alpha > 0$, the α-sublevel is convex as intersection of two convex sets $\mathbb{R}^2_{++}$ and the half space $\{x : x_1 - \alpha x_2 \leq 0\}$. Thus f is quasiconvex.

Similarly, f is also quasiconcave.

2.4

The Hessian of f on on $\mathbb{R}^2_{++}$ is $\alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \begin{bmatrix} -x_2^2 & x_1 x_2 \\ x_1 x_2 & -x_1^2 \end{bmatrix}$, which is semi negative definite. It can be deduced that f is concave.

The β-sublevel of f, $\{x \in \mathbb{R}^2_{++} : x_1^{\alpha} x_2^{1-\alpha} \leq \beta\} = \{x \in \mathbb{R}^2_{++} : x_2 \leq \beta^{\frac{1}{1-\alpha}} x_1^{\frac{-\alpha}{1-\alpha}}\}$ is not convex since $t \mapsto \beta^{\frac{1}{1-\alpha}} t^{\frac{-\alpha}{1-\alpha}}$ is not concave.

3

3.1

Let $X \in \mathbf{S}^n_{++}$ and $t \in \mathbb{R}, V \in \mathbf{S}^n$ such that $X + tV \in \mathbf{S}^n_{++}$.

Consider

$$\begin{aligned} g(t) = f(X+tV) &= \mathbf{Tr}\left((X+tV)^{-1}\right) = \mathbf{Tr}\left(X^{-1/2}(\mathbf{I}+tX^{-1/2}VX^{-1/2})^{-1}X^{-1/2}\right) \\ &= \mathbf{Tr}\left(X^{-1}(\mathbf{I}+tX^{-1/2}VX^{-1/2})^{-1}\right) = \mathbf{Tr}\left(X^{-1}(\mathbf{I}+tP^TSP)^{-1}\right) \\ &= \mathbf{Tr}\left(X^{-1}P^T(\mathbf{I}+tS)^{-1}P\right) = \mathbf{Tr}\left(PX^{-1}P^T(\mathbf{I}+tS)^{-1}\right) \\ &= \sum_{i=1}^n \frac{(PX^{-1}P^T)_{ii}}{1+tS_{ii}} \end{aligned}$$

Then

$$g''(t) = \sum_{i=1}^n \frac{2S_{ii}^2(PX^{-1}P^T)_{ii}}{(1+tS_{ii})^3}$$

Where P^TSP is the orthogonal diagonalization of $X^{-1/2}VX^{-1/2}$. Note that $PX^{-1}P^T \succeq 0$ and $\mathbf{I} + tS = PX^{-1/2}(X+tV)X^{-1/2}P^T \succeq 0$, which mean their diagonal entries are positive.

It follows that g is convex, thus f is convex.

3.2

Let $g(X, y, z) = -z^TXz + 2y^Tz$ on $\mathbf{S}^n_{++} \times \mathbb{R}^n \times \mathbb{R}^n$. For each y, $g(X, y, z)$ is linear and thus convex in (X, y).

Consequently, $f(X, y) = \sup_z g(X, y, z)$ is also convex.

3.3

Noting $\|\cdot\|$ the operator norm. We will first prove that

$$f(X) = \sup_{\|Y\| \leq 1} \mathbf{Tr}(X^TY) \tag{1}$$

Let $X = USV^T$ be the singular value decomposition of X, where U, V are orthogonal and $S = \mathrm{diag}(\sigma_1(X), ..., \sigma_n(X))$. For Y such that $\|Y\| \leq 1$

$$\mathbf{Tr}(X^TY) = \mathbf{Tr}(VSU^TY) = \mathbf{Tr}(SU^TYV) = \sum_{i=1}^n \sigma_i(X)u_i^TYv_i$$

Where u_i, v_i are respectively the i-th column of U and V. Since $\|Y\| \leq 1$ and u_i, v_i are unit vectors, it follows that

$$\mathbf{Tr}(X^TY) \leq \sum_{i=1}^n \sigma_i(X)\, \|Y\|\, \|u_i\|\, \|v_i\| \leq \sum_{i=1}^n \sigma_i(X)$$

For $Y = UV^T$, we have $U^T Y V = \mathbf{I}$ implies $u_i^T Y v_i = 1$ for all i. In this case the inequality is saturated. We conclude that (1) is correct.

Now, as $X \mapsto \mathbf{Tr}(X^T Y)$ is linear hence convex for all Y such that $\|Y\| \leq 1$, we deduce that f is convex.

4

$$K_{\mathrm{m}+} = \{x \in \mathbb{R}^n : x_1 \geq x_2 \geq ... \geq x_n \geq 0\} \tag{2}$$

4.1

$K_{\mathrm{m}+}$ is closed as inverse image of the closed set $\mathbb{R}_+^n$ under the continous application $x \mapsto (x_1 - x_2, ..., x_{n-1} - x_n, x_n)$.

$K_{\mathrm{m}+}$ is solid since $(n, n-1, ..., 1) \in \mathrm{int} K_{\mathrm{m}+}$.

$K_{\mathrm{m}+}$ contains no lines because $K_{\mathrm{m}+} \in \mathbb{R}_+^n$.

In conclusion, $K_{\mathrm{m}+}$ is a proper cone.

4.2

For $1 \leq k \leq n$, denote $u_k = (1, ..., 1, 0, ..., 0)$, the vector whose first k elements equaling to 1 and others elements equaling to 0. We see that these vectors are in $K_{\mathrm{m}+}$.

It is clear then that

$$K_{\mathrm{m}+}^* \subset \{y \in \mathbb{R}^n : u_k^T y \geq 0, \forall k\} \tag{3}$$

Supposing $u_k^T y \geq 0, \forall k$ and $x \in K_{\mathrm{m}+}$, we have

$$x^T y = \left(x_n u_n + \sum_{k=1}^{n-1} (x_k - x_{k+1}) u_k \right)^T y = x_n u_n^T y + \sum_{k=1}^{n-1} (x_k - x_{k+1}) u_k^T y \geq 0$$

Then $y \in K_{\mathrm{m}+}^*$.

Consequently, we deduce that

$$K_{\mathrm{m}+}^* = \{y \in \mathbb{R}^n : u_k^T y \geq 0, \forall k\} \tag{4}$$

5

We note

$$h(x, y) = y^T x - f(x) \tag{5}$$

And

$$f^*(y) = \sup_x h(x, y) \tag{6}$$

5.1

If $\exists i : y_i < 0$, let $x_j = 0, \forall j \neq i$ and $x_i \to -\infty$ we have $h(x, y) = y_i x_i \to +\infty$

If $\exists i : y_i > 1$, let $x_j = 0, \forall j \neq i$ and $x_i \to -\infty$ we have $h(x, y) = (y_i - 1)x_i \to +\infty$

If $y^T \mathbf{1} \neq 1$ we have $h(y, x_0 \mathbf{1}) = (y^T \mathbf{1} - 1)x_0$, which is linear in x_0 and therefore $f^*(y) = +\infty$.

Let us consider the remaining case $y \in [0, 1]^n$ and $y^T \mathbf{1} = 1$. The partial derivative of h with respect to x_n is $y_n \geq 0$ if $x_n < \max_{1 \leq i \leq n-1} x_i$ and $y_n - 1 \leq 0$ if $x_n > \max_{1 \leq i \leq n-1}$. Thus

$$\sup_x h(x, y) = \sup_x h(x_1, ..., x_{n-1}, \max_{1 \leq i \leq n-1} x_n, y) = (1 - y_n) \sup_{x_1, ..., x_{n-1}} \left(\frac{\sum_{i=1}^{n-1} y_i x_i}{1 - y_n} - \max_{1 \leq i \leq n-1} x_i \right)$$

Note that $\frac{y_i}{1-y_n} \in [0, 1], \forall i \leq n - 1$ and $\sum_{i=1}^{n-1} \frac{y_i}{1-y_n} = 1$. We can then continue the same transformation for remaining variables and finally obtain

$$\sup_{x_1, ..., x_n} h(x, y) = ... = \prod_{i=2}^{n} (1 - y_i) \sup_{x_1} (x_1 - \max(x_1)) = 0$$

We conclude that

$$f^*(y) = \begin{cases} 0 & \text{if } y \in [0, 1]^n, y^T \mathbf{1} = 1 \\ +\infty & \text{otherwise} \end{cases} \tag{7}$$

5.2

Similarly to 5.1, we only consider the case $y \in [0, 1]^n$ and $y^T \mathbf{1} = r$. Otherwise, $f^*(y) = +\infty$.

Note $z_i = x_{[i]}$ (sorted vector of x) and with $\bar{z} = \frac{\sum_{i=1}^{r}(1-y_i)z_i}{\sum_{i=1}^{r}(1-y_i)} \geq z_{r+1} \geq ... \geq z_n$

$$h(x, y) = \sum_{i=1}^{r} (y_i - 1)z_i + \sum_{i=r+1}^{n} y_i z_i \leq -\bar{z} \sum_{i=1}^{r} (1 - y_i) + \bar{z} \sum_{i=r+1}^{n} y_i = \bar{z} \left(\sum_{i=1}^{n} y_i - r \right) = 0$$

And $h(0, y) = 0$. Therefore $f^*(y) = 0$.

We have then We conclude that

$$f^*(y) = \begin{cases} 0 & \text{if } y \in [0, 1]^n, y^T \mathbf{1} = r \\ +\infty & \text{otherwise} \end{cases} \tag{8}$$

5.3

If $y > a_m$ then $h(x, y) \to +\infty$ as $x \to +\infty$.

If $y < a_1$ then $h(x, y) \to +\infty$ as $x \to -\infty$.

Let us assume that $y \in [a_i, a_{i+1}]$ for a certain i.

As f is convex, f' is increasing. Then $f(x)$ must be progressively equal to $a_1 x + b, ..., a_m x + b$ in this order as x increase. The break points of the graph of x are

$$z_1 \leq ... \leq z_m \text{ where } z_j = \frac{b_{j+1} - b_j}{a_{j+1} - a_j}$$

We know that $\frac{\partial h}{\partial x}$ is equal to $y - a_j \geq 0, j \leq i$ if $x < z_i$ and is equal to $y - a_j \leq 0, j > i$ if $x > z_i$. It follows that

$$\sup_x h(x,y) = h(z_i, y) = (y - a_i)\frac{b_{i+1} - b_i}{a_{i+1} - a_i} - b_i$$

Finally

$$f^*(x) = \begin{cases} (y - a_i)\frac{b_{i+1} - b_i}{a_{i+1} - a_i} - b_i & \text{if } \exists i : y \in [a_i, a_{i+1}] \\ +\infty & \text{if } y \notin [a_1, a_m] \end{cases} \tag{9}$$