MVA: Convex Optimization

Homework 1

Borachhun YOU

Exercise 1: Which of the following sets are convex?

① A rectangle, i.e. a set of the form $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i=1,\dots,n\}$

Applying the definition:

Let $R = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i = 1, \dots, n\}$ a rectangle

$x_1, x_2 \in R$

$0 \leq \theta \leq 1$

Then, for x_1 : $\alpha_i \leq x_{1i} \leq \beta_i$, $\forall i$ (1)

for x_2 : $\alpha_i \leq x_{2i} \leq \beta_i$, $\forall i$ (2)

Multiply (1) by θ, and (2) by $(1-\theta)$, we get:

$$\theta \alpha_i \leq \theta x_{1i} \leq \theta \beta_i$$
$$+ \quad (1-\theta)\alpha_i \leq (1-\theta) x_{2i} \leq (1-\theta)\beta_i$$
$$\theta\alpha_i + (1-\theta)\alpha_i \leq \theta x_{1i} + (1-\theta) x_{2i} \leq \theta\beta_i + (1-\theta)\beta_i$$
$$\alpha_i \leq \theta x_{1i} + (1-\theta) x_{2i} \leq \beta_i$$

Then $\theta x_1 + (1-\theta) x_2 \in R$

Thus, a rectangle is a convex set.

② The hyperbolic set $\{x \in \mathbb{R}_+^2 \mid x_1 x_2 \geq 1\}$

Applying the definition:

Let $H = \{x \in \mathbb{R}_+^2 \mid x_1 x_2 \geq 1\}$

$x, y \in H$

$0 \leq \theta \leq 1$

Then, for $x = (x_1, x_2)$: $x_1 x_2 \geq 1$

for $y = (y_1, y_2)$: $y_1 y_2 \geq 1$

- Case $x_1 \geq y_1$ and $x_2 \geq y_2$:

$\theta x_i + (1-\theta) y_i$, $i \in \{1,2\}$

$= \theta(x_i - y_i) + y_i \geq y_i$ since $x_i - y_i \geq 0$

$\Rightarrow [\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2] \geq y_1 y_2 \geq 1$

- Case $x_1 \leq y_1$ and $x_2 \leq y_2$:

$\theta y_i + (1-\theta) x_i$, $i \in \{1,2\}$

$= \theta(y_i - x_i) + x_i \geq x_i$ since $y_i - x_i \geq 0$

$\Rightarrow [\theta y_1 + (1-\theta) x_1][\theta y_2 + (1-\theta) x_2] \geq x_1 x_2 \geq 1$

- Case $x_1 \geq y_1$ and $x_2 \leq y_2$, and case $x_1 \leq y_1$ and $x_2 \geq y_2$:

$$[\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2]$$
$$= \theta^2 x_1 x_2 + \theta(1-\theta) x_1 y_2 + \theta(1-\theta) x_2 y_1 + (1-\theta)^2 y_1 y_2$$
$$= \theta x_1 x_2 + (1-\theta) y_1 y_2 - \theta(1-\theta)(y_1 - x_1)(y_2 - x_2)$$
$$\geq \theta x_1 x_2 + (1-\theta) y_1 y_2 \quad \text{since } (y_1 - x_1)(y_2 - x_2) \leq 0$$
$$\geq 1 \quad \text{since } x_1 x_2 \geq 1 \text{ and } y_1 y_2 \geq 1$$

With the 4 cases above, we get $\theta x + (1-\theta) y \in H$

Thus, the hyperbolic set is a convex set

(3) The set of points closer to a given point than a given set, i.e.
$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$ where $S \subseteq \mathbb{R}^n$

The set can be written as $\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$

$$\|x - x_0\|_2 \leq \|x - y\|_2$$
$$\|x - x_0\|_2^2 \leq \|x - y\|_2^2$$
$$(x - x_0)^T (x - x_0) \leq (x - y)^T (x - y)$$
$$x^T x - 2 x_0^T x + x_0^T x_0 \leq x^T x - 2 y^T x + y^T y$$
$$2 (y - x_0)^T x \leq y^T y - x_0^T x_0$$

The set can be rewritten as $\bigcap_{y \in S} \{x \mid 2(y - x_0)^T x \leq y^T y - x_0^T x_0\}$
which is an intersection of halfspaces, and since halfspaces are convex, the set is also convex.

(4) The set of points closer to one set than another, i.e.
$\{x \mid \text{dist}(x, S) \leq \text{dist}(x, T)\}$ where $S, T \subseteq \mathbb{R}^n$, and
$\text{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$

As an example in $\mathbb{R}$, take $S = \{-2, 3\}$ and $T = \{1\}$

$\{x \mid \text{dist}(x, S) \leq \text{dist}(x, T)\} = \{x \in \mathbb{R} \mid x \leq -\frac{1}{2} \text{ or } x \geq 2\}$

This example set is clearly not convex

Thus, in general, this set is not convex.

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

$$\{x \mid x + S_2 \subseteq S_1\} = \{x \mid x + y \in S_1 \text{ for all } y \in S_2\}$$
$$= \bigcap_{y \in S_2} \{x \mid x + y \in S_1\}$$
$$= \bigcap_{y \in S_2} (S_1 - y)$$

We want to show that $(S_1 - y)$ is a convex set

Let $x_1, x_2 \in S_1$, $0 \leq \theta \leq 1$

S_1 is convex $\Rightarrow \theta x_1 + (1-\theta)x_2 \in S_1$

$\theta x_1 + (1-\theta)x_2 - y \in S_1 - y$

$\theta(x_1 - y) + (1-\theta)(x_2 - y) \in S_1 - y$

Then, $x_1 - y, x_2 - y \in (S_1 - y) \Rightarrow \theta(x_1 - y) + (1-\theta)(x_2 - y) \in (S_1 - y)$

$\Rightarrow S_1 - y$ is convex

Thus, $\{x \mid x + S_2 \subseteq S_1\}$ is an intersection of convex sets and therefore convex.

Exercise 2 : Determine whether a function is convex or concave.

① $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$

dom $f = \mathbb{R}^2_{++}$ is convex

$$\frac{\partial f}{\partial x_1} = x_2 \ , \ \frac{\partial f}{\partial x_2} = x_1$$

$$\frac{\partial^2 f}{\partial x_1^2} = 0 \ , \ \frac{\partial^2 f}{\partial x_2^2} = 0 \ , \ \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 1$$

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Let I be an identity matrix

$$\det(\nabla^2 f - \lambda I) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1$$

$\Rightarrow$ The eigenvalues of $\nabla^2 f$ are -1 and 1

Since the eigenvalues of $\nabla^2 f$ are positive and negative, then $\nabla^2 f$ is neither positive semidefinite nor negative semidefinite

Thus f is neither a convex function nor a concave function.

② $f(x_1, x_2) = \frac{1}{x_1 x_2}$ on $\mathbb{R}^2_{++}$

dom $f = \mathbb{R}^2_{++}$ is convex

$$\frac{\partial f}{\partial x_1} = -\frac{1}{x_1^2 x_2} \ , \ \frac{\partial f}{\partial x_2} = -\frac{1}{x_1 x_2^2}$$

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{2}{x_1^3 x_2} \ , \ \frac{\partial^2 f}{\partial x_2^2} = \frac{2}{x_1 x_2^3} \ , \ \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{1}{x_1^2 x_2^2}$$

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix}$$

$$\det(\nabla^2 f) = \frac{4}{x_1^4 x_2^4} - \frac{1}{x_1^4 x_2^4} = \frac{3}{x_1^4 x_2^4} > 0$$

$\Rightarrow$ We know that the two eigenvalues of $\nabla^2 f$ have the same sign (*)

$$tr(\nabla^2 f) = \frac{2}{x_1^3 x_2} + \frac{2}{x_1 x_2^3} = \frac{2(x_1^2+x_2^2)}{x_1^3 x_2^3} > 0$$

$\Rightarrow$ With $(*)$, the eigenvalues of $\nabla^2 f$ are positive
Then, $\nabla^2 f$ is positive definite
Thus, f is convex.

③ $f(x_1, x_2) = \frac{x_1}{x_2}$ on $\mathbb{R}^2_{++}$
dom $f = \mathbb{R}^2_{++}$ is convex

$$\frac{\partial f}{\partial x_1} = \frac{1}{x_2}, \quad \frac{\partial f}{\partial x_2} = -\frac{x_1}{x_2^2}$$

$$\frac{\partial^2 f}{\partial x_1^2} = 0, \quad \frac{\partial^2 f}{\partial x_2^2} = \frac{2x_1}{x_2^3}, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -\frac{1}{x_2^2}$$

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{pmatrix}$$

$$\det(\nabla^2 f) = -\frac{1}{x_2^4} < 0$$

$\Rightarrow$ We know that the two eigenvalues of $\nabla^2 f$ have different signs
Then, $\nabla^2 f$ is neither positive semidefinite nor negative semidefinite
Thus, f is neither convex nor concave.

④ $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, $0 \leq \alpha \leq 1$, on $\mathbb{R}^2_{++}$
dom $f = \mathbb{R}^2_{++}$ is convex

$$\frac{\partial f}{\partial x_1} = \alpha x_1^{\alpha-1} x_2^{1-\alpha}, \quad \frac{\partial f}{\partial x_2} = (1-\alpha) x_1^\alpha x_2^{-\alpha}$$

$$\frac{\partial^2 f}{\partial x_1^2} = \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha}, \quad \frac{\partial^2 f}{\partial x_2^2} = -\alpha(1-\alpha) x_1^\alpha x_2^{-\alpha-1}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha}$$

$$\nabla^2 f(x_1,x_2) = \begin{pmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & \alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1} \end{pmatrix}$$

$\det(\nabla^2 f) = \alpha^2(\alpha-1)^2 x_1^{2\alpha-2}x_2^{-2\alpha} - \alpha^2(1-\alpha)^2 x_1^{2\alpha-2}x_2^{-2\alpha} = 0$

$\Rightarrow$ The eigenvalues of $\nabla^2 f$ consist of 0 $(*)$

$\mathrm{tr}(\nabla^2 f) = \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} + \alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1}$

$= \alpha(\alpha-1)\left(x_1^{\alpha-2}x_2^{1-\alpha} + x_1^{\alpha}x_2^{-\alpha-1}\right) \leq 0$ since $\alpha-1\leq 0$

$\Rightarrow$ With $(*)$, the eigenvalues ~~ar~~ of $\nabla^2 f$ are 0 and $\lambda \leq 0$

Then, $\nabla^2 f$ is negative semidefinite

Thus, f is concave.

Exercise 3 : Show the following functions are convex

① $f(X) = \mathrm{tr}(X^{-1})$ on $\mathrm{dom}\, f = S^n_{++}$

dom f is convex

Let $S(t) = A + tB$ where $A \in S^n_{++}$ and B is symmetric

The goal is to show that : $\frac{d^2}{dt^2}\mathrm{tr}(S(t)^{-1})\Big|_{t=0} \geq 0$

$S(t)^{-1} = (A+tB)^{-1} = \left[A(I+tA^{-1}B)\right]^{-1} = (I+tA^{-1}B)^{-1}A^{-1}$

We can show that : $I = (I+M)^{-1}(I+M) = (I+M)^{-1} + (I+M)^{-1}M$

$\Rightarrow (I+M)^{-1} = I - (I+M)^{-1}M \quad (*)$

We then recursively apply $(*)$ to $S(t)^{-1}$ 3 times :

$S(t)^{-1} = (I+tA^{-1}B)^{-1}A^{-1}$

$\overset{(*)}{=} \left[I - (I+tA^{-1}B)^{-1}tA^{-1}B\right]A^{-1} = A^{-1} - (I+tA^{-1}B)^{-1}tA^{-1}BA^{-1}$

$\overset{(*)}{=} A^{-1} - \left[I-(I-tA^{-1}B)^{-1}tA^{-1}B\right]tA^{-1}BA^{-1}$

$= A^{-1} - tA^{-1}BA^{-1} + (I-tA^{-1}B)^{-1}t^2A^{-1}BA^{-1}BA^{-1}$

$\overset{(*)}{=} A^{-1} - tA^{-1}BA^{-1} + \left[I-(I-tA^{-1}B)^{-1}tA^{-1}B\right]t^2A^{-1}BA^{-1}BA^{-1}$

$= A^{-1} - tA^{-1}BA^{-1} + t^2A^{-1}BA^{-1}BA^{-1} - (I-tA^{-1}B)^{-1}t^3A^{-1}BA^{-1}BA^{-1}BA^{-1}$

$$tr(S(t)^{-1}) = tr(A^{-1}) - t \cdot tr(A^{-1}BA^{-1}) + t^2 \cdot tr(A^{-1}BA^{-1}BA^{-1})$$
$$- t^3 \cdot tr((I - tA^{-1}B)^{-1} A^{-1}BA^{-1}BA^{-1}BA^{-1})$$

$$\frac{d^2}{dt^2} tr(S(t)^{-1}) = 2tr(A^{-1}BA^{-1}BA^{-1}) - \frac{d^2}{dt^2}\left[t^3 \cdot tr((I - tA^{-1}B)^{-1}A^{-1}BA^{-1}BA^{-1}BA^{-1})\right]$$

$$\left.\frac{d^2}{dt^2} tr(S(t)^{-1})\right|_{t=0} = 2tr(A^{-1}BA^{-1}BA^{-1})$$

Let $C = A^{-1}B \Rightarrow C^T = (A^{-1}B)^T = B^T(A^{-1})^T = B^T(A^T)^{-1} = BA^{-1}$

$$\left.\frac{d^2}{dt^2} tr(S(t)^{-1})\right|_{t=0} = 2tr(CA^{-1}C^T)$$

Let x be a vector. Since A^{-1} is positive definite, we can decompose ~~$x^TCA^{-1}C^Tx$~~ it into: $A^{-1} = Z^TZ$

$$x^TCA^{-1}C^Tx = x^TCZ^TZC^Tx = (ZC^Tx)^T(ZC^Tx) = \|ZC^Tx\|^2 \geq 0$$

$\Rightarrow CA^{-1}C^T$ is positive semidefinite

$\Rightarrow tr(CA^{-1}C^T) \geq 0$

$\Rightarrow \left.\dfrac{d^2}{dt^2} tr(S(t)^{-1})\right|_{t=0} \geq 0$

Thus, f is convex

② $f(X, y) = y^TX^{-1}y$ on $\text{dom } f = S^n_{++} \times \mathbb{R}^n$

Consider $y^Tz - \frac{1}{2}z^TXz$. It attains its maximum at $z = X^{-1}y$

$\Rightarrow \sup\limits_z (y^Tz - \frac{1}{2}z^TXz) = \frac{1}{2}y^TX^{-1}y$

$\sup\limits_z (2y^Tz - z^TXz) = y^TX^{-1}y$

$2y^Tz - z^TXz$ is linear for both X and y

$\Rightarrow$ it is convex in (X, y)

Thus, f is convex because it is a supremum of convex functions

③ $f(X) = \sum_{i=1}^{n} \sigma_i(X)$ on $\operatorname{dom} f = S^n$ where $\sigma_i(X)$ are singular values of a matrix $X \in \mathbb{R}^{n\times n}$

The goal is to show that f is a norm

- $f(X) = 0$
  $\Rightarrow \sum_{i=1}^{n} \sigma_i(X) = 0 \Rightarrow \forall i,\ \sigma_i(X) = 0$
  $\Rightarrow X = 0$
- $f(aX) = \sum_{i=1}^{n} \sigma_i(aX) = \sum_{i=1}^{n} \sqrt{|a|^2 |\lambda_i|^2}$, λ_i eigenvalue of $X \in S^n$
  $= |a| \sum_{i=1}^{n} \sqrt{|\lambda_i|^2} = |a| \sum_{i=1}^{n} \sigma_i(X) = |a| f(X)$
- $f(X) = \sum_{i=1}^{n} \sigma_i(X) \geq 0$ since $\sigma_i(X) \geq 0$
- We want to show that $f(X+Y) \leq f(X) + f(Y)$
  First, we want to show that $\sup_{\sigma_1(Q)\leq 1} \langle Q, X\rangle = \sup_{\sigma_1(Q)\leq 1} \operatorname{tr}(Q^T X) = \sum_{i=1}^{n} \sigma_i(X) = f(X)$

  where we define $Q = UV^T$ and $\sigma_1(Q)$ is the largest singular value of Q which is a norm
  By singular value decomposition : $X = U\Sigma V^T$
  $\langle Q, X\rangle = \langle UV^T, U\Sigma V^T\rangle = \operatorname{tr}(VU^TU\Sigma V^T) = \operatorname{tr}(V^TVU^TU\Sigma)$
  $= \operatorname{tr}(\Sigma) = \sum_{i=1}^{n} \sigma_i(X)$
  $\Rightarrow \sup_{\sigma_1(Q)\leq 1} \langle Q, X\rangle \geq \sum_{i=1}^{n} \sigma_i(X)$ (*)

  $\sup_{\sigma_1(Q)\leq 1} \langle Q, X\rangle = \sup_{\sigma_1(Q)\leq 1} \operatorname{tr}(Q^T U\Sigma V^T) = \sup_{\sigma_1(Q)\leq 1} \operatorname{tr}(V^T Q^T U \Sigma)$

  $= \sup_{\sigma_1(Q)\leq 1} \langle U^T Q V, \Sigma\rangle = \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{n} \sigma_i(X) (U^T Q V)_{ii}$

  $= \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{n} \sigma_i(X)\, u_i^T Q v_i \leq \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{n} \sigma_i(X)\, \sigma_1(Q)$

  $\leq \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{n} \sigma_i(X) = \sum_{i=1}^{n} \sigma_i(X)$ (**)

  (*) and (**) $\Rightarrow \sup_{\sigma_1(Q)\leq 1} \langle Q, X\rangle = \sum_{i=1}^{n} \sigma_i(X) = f(X)$

$$f(x+y) = \sup_{b_1(Q)\leq 1} \langle Q, x+y\rangle \leq \sup_{b_1(Q)\leq 1} \langle Q, x\rangle + \sup_{b_1(Q)\leq 1} \langle Q, y\rangle = f(x)+f(y)$$

From the 4 points above, f is a norm and thus convex.