Oliver Jack

HW 1

Ex. 1

1) Suppose $x, y \in A = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i \in \{1, \dots, n\}\}$ and $\theta \in [0, 1]$. Then $\forall i \in \{1, \dots, n\}$

- $\underbrace{\theta \min\{x_i, y_i\} + (1-\theta)\min\{x_i, y_i\}}_{= \min\{x_i, y_i\}} \leq \theta x_i + (1-\theta) y_i$
- $\theta x_i + (1-\theta) y_i \leq \underbrace{\theta \max\{x_i, y_i\} + (1-\theta)\max\{x_i, y_i\}}_{= \max\{x_i, y_i\}}$

$\Rightarrow \alpha_i \leq \min\{x_i, y_i\} \leq \theta x_i + (1-\theta) y_i \leq \max\{x_i, y_i\} \leq \beta_i$

$\Rightarrow \theta x + (1-\theta) y \in A$

2) Since if $x_1 = 0$ or $x_2 = 0$ the equation never holds, we may consider the set:

$B = \{(x, t) \in \mathbb{R}^2_{++} \mid \frac{1}{x} \leq t\}$

Consider the function

$$f : \mathbb{R}_{++} \longrightarrow \mathbb{R}$$
$$x \longmapsto \frac{1}{x}$$

Since $\forall x, y \in \mathbb{R}_{++}$:

$f(y) \geq f(x) + f'(x)(y - x)$

$\Leftrightarrow \frac{1}{y} \geq \frac{1}{x} - \frac{1}{x^2}(y - x)$

$\Leftrightarrow \frac{(x-y)^2}{y} \geq 0$ True since $y \in \mathbb{R}_{++}$!

So by the 1-st order condition f is convex ($\mathrm{dom}(f) = \mathbb{R}_{++}$ which is convex).

Since f is convex, $\mathrm{epi}\, f = B$ is a convex set.

3) Case 1: $x_0 \in S$

The set would simply be

$C = \{x_0\}$

since every point x in the set would have to verify:

$\|x - x_0\|_2 \leq \|x_0 - \underset{\in S}{x_0}\|_2 = 0$

$\Rightarrow x = x_0$

Case 2: $x_0 \notin S$

Fix $y \in S$. Consider the set

$C_y = \{ x \in \mathbb{R}^n \mid \|x - x_0\|_2^2 \leq \|x - y\|_2^2 \}$

We have:

$\|x - x_0\|_2^2 = x^T x - 2x^T x_0 + x_0^T x_0$

$\|x - y\|_2^2 = x^T x - 2x^T y + y^T y$

Thus the inequality becomes:

$x^T x - 2x^T x_0 + x_0^T x_0 \leq x^T x - 2x^T y + y^T y$

$\Leftrightarrow 2(y - x_0)^T x \leq y^T y - x_0^T x_0$

Hence, we can conclude that C_y is a halfspace, so it is convex.

Since the original set only contains $x \in \mathbb{R}^n$ s.t.:

$\|x - x_0\|_2 \leq \|x - y\|_2 \quad \forall y \in S$

we can write it as:

$C = \bigcap_{y \in S} C_y$

Thus, C is convex since it is the intersection of convex sets.

4) Consider the sets $S = (-\infty, -1] \cup [1, +\infty)$ and $T = \{0\}$ in $\mathbb{R}$. Then our set would be:

$D = \{x \mid \text{dist}(x, S) \leq \text{dist}(x, T)\}$

$= (-\infty, -\frac{1}{2}] \cup [\frac{1}{2}, +\infty)$

Take $x = -\frac{1}{2}$, $y = \frac{1}{2}$ and $\lambda = \frac{1}{2}$. Then:

$\frac{1}{2} \cdot (-\frac{1}{2}) + \frac{1}{2} \cdot \frac{1}{2} = 0 \notin D$

Thus, D is not convex.

5) The set can be rewritten as:

$$E = \{x \mid x + S_2 \subseteq S_1\}$$
$$= \{x \mid \forall y \in S_2 : x + y \in S_1\}$$
$$= \bigcap_{y \in S_2} \underbrace{\{x \mid x + y \in S_1\}}_{= (S_1 - y)}$$
$$= \bigcap_{y \in S_2} (S_1 - y)$$

Since $(S_1 - y)$ is the image of a convex set under an affine function (translation by $-y$), we can say that $(S_1 - y)$ is convex $\forall y \in S_2$. Thus, the intersection $\bigcap_{y \in S_2} (S_1 - y)$ is also convex.

Ex. 2

1) We have that $\mathbb{R}^2_{++}$ is open and

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

$\det(\lambda \cdot I - A) = 0 \iff \lambda = \pm 1$

$\Rightarrow$ A is neither positive / negative semi-definite

Thus, by the second order condition f is neither convex, nor concave.

2) We have that $\mathbb{R}^2_{++}$ is open and

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

By Sylvester's criterion $\forall x_1, x_2 \in \mathbb{R}_{++}$:

- $\det\left(\frac{2}{x_1^3 x_2}\right) = \frac{2}{x_1^3 x_2} > 0$
- $\det(B) = \frac{4}{x_1^4 x_2^4} - \frac{1}{x_1^4 x_2^4} = \frac{3}{x_1^4 x_2^4} > 0$

$\Rightarrow B(x_1, x_2) \succeq 0 \quad \forall (x_1, x_2) \in \mathbb{R}^2_{++}$

$\Rightarrow f$ is convex (and quasiconvex)

3) We have that $\mathbb{R}^2_{++}$ is open and

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

Since $\det(C) = -\frac{1}{x_2^4} < 0$, we can conclude by Sylvester's criterion that C is not semi-positive definite, so f is not convex.

Similarly:

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

Since

$$\det\left(-\frac{2x_1}{x_2^3}\right) = -\frac{2x_1}{x_2^3} < 0$$

we know by Sylvester's criterion that C' is not semi-positive definite, so $-f$ is not convex. Hence, f is not concave.

4) We have that $\mathbb{R}^2_{++}$ is open and for $\alpha \in (0,1)$:

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

$$= D$$

By Sylvester's criterion

- $\det(\alpha(1-\alpha)x_1^{\alpha-2}x_2^{1-\alpha})$

  $= \underbrace{\alpha}_{>0}\underbrace{(1-\alpha)}_{>0}x_1^{\alpha-1}x_2^{1-\alpha} > 0$

- $\det(\alpha(1-\alpha)x_1^{\alpha}x_2^{-1-\alpha})$

  $= \alpha(1-\alpha)x_1^{\alpha}x_2^{-1-\alpha} > 0$

- $\det(D) = \alpha^2(1-\alpha)^2x_1^{2\alpha-2}x_2^{-2\alpha} - \alpha^2(1-\alpha)^2x_1^{2\alpha-2}x_2^{-2\alpha}$

  $= 0$

we can conclude that D is positive semi-definite. Hence, $-f$ is convex, so f is concave.

If $\alpha = 0$ or $\alpha = 1$, then f is linear, so both convex and concave.

(Remark: If f is not an affine function then convexity excludes concavity and vice versa.)

Ex. 3

1) As seen in class, to show that f is convex it is sufficient to show that

$$g : D \to \mathbb{R}$$
$$t \mapsto f(X + tV) \quad \text{for } X \in S^n_{++},\ V \in S^n$$

is convex in t with $D = \{t \mid X + tV \in \text{dom } f\}$

Then:

$$g(t) = \text{Tr}((X + tV)^{-1})$$
$$= \text{Tr}([X^{\frac{1}{2}}(I_n + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})X^{\frac{1}{2}}]^{-1})$$
$$\overset{(*)}{=} \text{Tr}(X^{-\frac{1}{2}}(I_n + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})^{-1}X^{-\frac{1}{2}})$$

Since $\text{Tr}(ABC) = \text{Tr}(CAB) = \text{Tr}(BCA)$ for matrices A, B, C of same size, we get:

$$(*) = \text{Tr}(X^{-1}(I_n + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})^{-1})$$

Since $(X^{-\frac{1}{2}}VX^{-\frac{1}{2}})^T = (X^{-\frac{1}{2}})^T V^T (X^{-\frac{1}{2}})^T = X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$, we know that $X^{-\frac{1}{2}}VX^{-\frac{1}{2}} \in S_n$, so by the eigenvalue decomposition $\exists\, Q$ an $n \times n$ orthogonal matrix and Σ a diagonal matrix with eigenvalues λ_i of the matrix $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$ such that:

$$X^{-\frac{1}{2}}VX^{-\frac{1}{2}} = Q\Sigma Q^T$$

Thus:

$$(*) = \text{Tr}(X^{-1}(I_n + tQ\Sigma Q^T)^{-1})$$
$$= \text{Tr}(X^{-1}(Q^T)^{-1}(I_n + t\Sigma)^{-1}Q^{-1}) \quad // \ Q^{-1} = Q^T$$

$= \mathrm{Tr}\left(X^{-1} Q (I + t\Sigma)^{-1} Q^T\right)$

$= \mathrm{Tr}\left(Q^T X^{-1} Q (I + t\Sigma)^{-1}\right)$ $\quad // \mathrm{Tr}(AB) = \sum_i A_{ii} B_{ii}$

$= \sum_{i=1}^{n} (Q^T X^{-1} Q)_{ii} \underbrace{(I + t\Sigma)^{-1}_{ii}}_{\text{diagonal matrix with entries } 1 + t\lambda_i}$

$= \sum_{i=1}^{n} (Q^T X^{-1} Q)_{ii} \dfrac{1}{1 + t\lambda_i}$

Since $X \in S^n_{++}$, $X^{-1} \in S^n_{++}$.

Thus $Q^T X^{-1} Q \in S^n_{++}$, which implies that the diagonal entries of $Q^T X^{-1} Q$ are positive.

Furthermore, $\forall t \in \mathrm{dom}(g): 1 + t\lambda_i > 0$

So, $\forall i \in \{1, \ldots, n\}$:

$t \mapsto (Q^T X^{-1} Q)_{ii} \dfrac{1}{1 + t\lambda_i}$

is a convex function in t ($\in \mathbb{R}_+$).

Finally, g is convex since it is a non-negative weighted sum.

2) Let us start off by proving:

$y^T X^{-1} y = \sup_{z \in \mathbb{R}^n} \{2y^T z - z^T X z\}$

We have:

$z^T X z - 2y^T z$

$= (z - X^{-1}y)^T X (z - X^{-1}y) - y^T X^{-1} y$

The infimum in terms of z is reached if

$z = X^{-1} y$

Thus:

$\inf_{z \in \mathbb{R}^n} \{z^T X z - 2y^T z\} = -y^T X^{-1} y$

$\Leftrightarrow y^T X^{-1} y = \sup_{z \in \mathbb{R}^n} \{2y^T z - z^T X z\}$

Since the function

$g: S^n_{++} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$

$(X, y, z) \mapsto 2y^T z - z^T X z$

is affine in both X and y, we know that it is convex in (X, y) $\forall z \in \mathbb{R}^n$

Thus:

$$(X, y) \mapsto \sup_{z \in \mathbb{R}^n} \{2y^T z - z^T X z\}$$

is convex.

3) Consider the singular value decomposition of X:

$X = U \Sigma V^T$ where U, V are $n \times n$ orthogonal matrices and Σ is a diagonal matrix with entries $\sigma_1(X), \dots, \sigma_n(X)$

Then:

$$\begin{aligned} f(X) &= \sum_{i=1}^{n} \sigma_i(X) \\ &= \mathrm{Tr}(\Sigma) \\ &= \mathrm{Tr}(U^T U \Sigma V^T V) \quad \| \mathrm{Tr}(ABC) = \mathrm{Tr}(CBA) \\ &= \mathrm{Tr}(V U^T U \Sigma V^T) \\ &= \mathrm{Tr}(V U^T X) \end{aligned}$$

Since U and V^T are orthogonal, we have for $A = U V^T$:

$\|A\| = 1$ where $\|\cdot\|$ is the spectral norm

Thus:

$$\sum_{i=1}^{n} \sigma_i(X) \leq \sup_{\|Y\| \leq 1} (\mathrm{Tr}(Y^T X))$$

On the other hand, consider $\|Y\| \leq 1$. Then:

$$\begin{aligned} \mathrm{Tr}(Y^T X) &= \mathrm{Tr}(Y^T U \Sigma V^T) \\ &= \mathrm{Tr}(V^T Y^T U \Sigma) \end{aligned}$$

Set $Z = V^T Y^T U$.

Since U and V are orthogonal:

$\|Z\| = \|Y\| \leq 1$

Thus, by von Neumann's trace inequality:

$\mathrm{Tr}(Y^T X) = \mathrm{Tr}(Z\Sigma)$

$$\leq \sum_{i=1}^{n} \underbrace{\sigma_i(Z)}_{\leq 1} \underbrace{\sigma_i(\Sigma)}_{= \sigma_i(X)}$$

$$\leq \sum_{i=1}^{n} \sigma_i(X)$$

$\Rightarrow \sup_{\|Y\| \leq 1} (\mathrm{Tr}(Y^T X)) \leq \sum_{i=1}^{n} \sigma_i(X)$

$\Rightarrow \sup_{\|Y\| \leq 1} (\mathrm{Tr}(Y^T X)) = \sum_{i=1}^{n} \sigma_i(X)$

For $Y \in \mathbb{R}^n$ fixed

$X \mapsto \mathrm{Tr}(Y^T X)$

is the composition of two linear functions, so is also linear, thus convex in X.

Then

$f(X) = \sup_{\|Y\| \leq 1} (\mathrm{Tr}(Y^T X))$

is also convex.

* Remark (Ex. 3.1)

To show that $1 + t\lambda_i > 0 \quad \forall t \in \mathrm{dom}(g)$, it suffices to show that

$I_n + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$

is positive definite.

By assumption, we have that $X, X + tV \in S_{++}^n$.

Set $A = \underbrace{X^{\frac{1}{2}}}_{\succ 0}$ and $B = I_n + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$.

Then:

$X + tV = ABA$

Since $X + tV \succ 0$, we have:

$\forall y \in \mathbb{R}^n$:

$y^T ABA y > 0$

$\Leftrightarrow (Ay)^T B (Ay) > 0$

$\forall y' \in \mathbb{R}^n, \exists y \in \mathbb{R}^n$ s.t.: $y = A^{-1} y'$

So $\forall y' \in \mathbb{R}^n$:

$(y')^T B y' > 0$

$\Rightarrow B = I_n + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}} \succ 0$