BARRA
Romane

Convex Optimization
Homework 1

Exercise 1:

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

$$= \left(\bigcap_{i=1}^{n} \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i\}\right) \cap \left(\bigcap_{i=1}^{n} \{x \in \mathbb{R}^n \mid x_i \leq \beta_i\}\right)$$

is the intersection of $2n$ halfspaces (which are convex), hence it is convex.

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

Let $x = (x_1, x_2)$, $y = (y_1, y_2) \in H$ and $0 \leq \theta \leq 1$

We define $z = \theta x + (1-\theta) y$

$$z_1 z_2 = (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2)$$
$$= \theta^2 \underbrace{x_1 x_2}_{\geq 1} + \theta(1-\theta)(x_1 y_2 + y_1 x_2) + (1-\theta)^2 \underbrace{y_1 y_2}_{\geq 1}$$

Since $x_1 x_2 \geq 1$ and $x_1, x_2 > 0$, then $x_2 \geq \frac{1}{x_1}$

and also $y_1 \geq \frac{1}{y_2}$

Thus: $z_1 z_2 \geq \theta^2 + \theta(1-\theta)\left(x_1 y_2 + \frac{1}{x_1 y_2}\right) + (1-\theta)^2$

$g : x \mapsto x + \frac{1}{x}$ is minimized by 2 on $\mathbb{R}_+^*$, hence:

$$z_1 z_2 \geq \theta^2 + 2\theta(1-\theta) + (1-\theta)^2 = 1$$

$z \in H$, thus H is convex.

3) $C = \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$

For a fixed $y \in S$:

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

$\iff \cancel{x^T x} - 2x_0^T x + x_0^T x_0 \leq \cancel{x^T x} - 2y^T x + y^T y$

$$\|x-x_0\|_2 \leq \|x-y\|_2 \iff \underbrace{2(y-x_0)^T}_{=a} x \leq \underbrace{y^Ty - x_0^Tx_0}_{=b}$$

Therefore, $\{x \mid \|x-x_0\|_2 \leq \|x-y\|_2\} = \{x \mid a^Tx \leq b\}$ is a halfspace.

Then, $C = \bigcap_{y \in S} \{x \mid \|x-x_0\|_2 \leq \|x-y\|_2\}$ is the intersection of halfspaces, thus it is convex.

4) $E = \{x \mid \operatorname{dist}(x,S) \leq \operatorname{dist}(x,T)\}$ with $S,T \subseteq \mathbb{R}^n$.

For $n=1$, $S =]-\infty, -2] \cup [2, +\infty[$ and $T = [-1, 1]$:

$$E =]-\infty, -\tfrac{3}{2}] \cup [\tfrac{3}{2}, +\infty[\text{ which is not convex}$$

So, E is not convex in general

5) $C = \{x \mid x + S_2 \subseteq S_1\}$ with $S_1, S_2 \subseteq \mathbb{R}^n$ and S_1 convex

Let $x, y \in C$ and $0 \leq \theta \leq 1$

$$\forall z \in S_2, \quad \theta x + (1-\theta)y + z = \theta\underbrace{(x+z)}_{\in S_1} + (1-\theta)\underbrace{(y+z)}_{\in S_1} \underset{S_1 \text{ convex}}{\in} S_1$$

Thus, $\theta x + (1-\theta)y \in C$ and C is convex.

Exercise 2:

1) $f(x_1, x_2) = x_1 x_2$

$$\nabla f(x_1,x_2) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x_1,x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$\nabla^2 f$ is not positive semidefinite, and $-\nabla^2 f$ neither.

Thus, f is not convex and not concave.

As seen in the previous exercise, question 2, $\{(x_1, x_2) \in \mathbb{R}_+^2 \mid x_1 x_2 \geq 1\}$ is convex, which is still true for $\{(x_1, x_2) \in \mathbb{R}_+^2 \mid x_1 x_2 \geq \alpha\}$ for any α.
Therefore, <u>f is quasiconvex</u>.

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

$$\nabla f(x_1, x_2) = \begin{bmatrix} -\frac{1}{x_1^2 x_2} \\ -\frac{1}{x_1 x_2^2} \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x_1, x_2) = \frac{1}{x_1 x_2} \begin{bmatrix} \frac{2}{x_1^2} & \frac{1}{x_1 x_2} \\ \frac{1}{x_1 x_2} & \frac{2}{x_2^2} \end{bmatrix}$$

Let $y = (y_1, y_2) \in \mathbb{R}^2$

$$^t y \nabla^2 f(x_1, x_2) y = \left(\frac{2y_1^2}{x_1^2} + \frac{2y_1 y_2}{x_1 x_2} + \frac{2y_2^2}{x_2^2} \right) \times \frac{1}{x_1 x_2}$$

$$= \frac{2y_1^2 x_2^2 + 2y_1 y_2 x_1 x_2 + 2y_2^2 x_1^2}{x_1^2 x_2^2} \times \frac{1}{x_1 x_2}$$

$$= \frac{(y_1 x_2 + y_2 x_1)^2 + (y_1 x_2)^2 + (y_2 x_1)^2}{x_1^2 x_2^2} \times \underbrace{\frac{1}{x_1 x_2}}_{\geq 0}$$

$$\geq 0$$

So: $\nabla^2 f \succcurlyeq 0$ and <u>f is convex and quasiconvex</u>.

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

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{1}{x_2} \\ -\frac{x_1}{x_2^2} \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ \frac{-1}{x_2^2} & \frac{2x_1}{x_2^2} \end{bmatrix}$$

For $y = (2x_1, 1)$:

$$y^T \nabla^2 f(x_1, x_2) y = -\frac{2x_1}{x_2^2} < 0$$

For $y = (0, 1)$:

$$y^T \nabla^2 f(x_1, x_2) y = \frac{2x_1}{x_2^2} > 0$$

Therefore, $\nabla^2 f$ and $-\nabla^2 f$ are not positive semidefinite and f is neither convex or concave.

Let $x = (x_1, x_2)$ and $y = (y_1, y_2) \in \mathbb{R}^2_{++}$ such that $f(x) \geq f(y)$.

$$\nabla f(x)^T (y-x) = \frac{y_1 - x_1}{x_2} - \frac{x_1(y_2 - x_2)}{x_2^2}$$

$$= \frac{(y_1 - x_1)x_2 - x_1(y_2 - x_2)}{x_2^2}$$

$$= \frac{y_1 x_2 - x_1 y_2}{x_2^2}$$

$f(x) \geq f(y) \iff \frac{x_1}{x_2} \geq \frac{y_1}{y_2}$

$$\leq 0$$

Thus, f is quasiconvex.

Let $x, y \in \mathbb{R}^2_{++}$ such that $-f(x) \geq -f(y)$

$$\nabla(-f)(x)^T (y-x) = \frac{x_1 y_2 - y_1 x_2}{x_2^2}$$

$-f(x) \geq -f(y) \iff \frac{x_1}{x_2} \leq \frac{y_1}{y_2}$

$$\leq 0$$

f is also quasiconcave, therefore f is quasilinear.

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

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

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

$$= -\alpha(1-\alpha) x_1^{\alpha} x_2^{1-\alpha} \begin{bmatrix} \frac{1}{x_1^2} & \frac{-1}{x_1 x_2} \\ \frac{-1}{x_1 x_2} & \frac{1}{x_2^2} \end{bmatrix}$$

$$\nabla^2 f(x_1, x_2) = \underbrace{-\alpha(1-\alpha)\, x_1^{\alpha} x_2^{1-\alpha}}_{\leq 0} \underbrace{\begin{bmatrix} \frac{-1}{x_1} \\ \frac{1}{x_2} \end{bmatrix} \begin{bmatrix} \frac{-1}{x_1} \\ \frac{1}{x_2} \end{bmatrix}^T}_{\succeq 0} \preceq 0.$$

Hence, f is concave and quasiconcave.

Exercise 3:

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

Let $X, V \in S^n_{++}$

We define $g : \mathbb{R} \to \mathbb{R}$

$$t \mapsto f(X + tV)$$

$$g(t) = \mathrm{Tr}\left((X+tV)^{-1}\right)$$
$$= \mathrm{Tr}\left(X^{-1}\left(Id + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}\right)^{-1}\right)$$

We decompose $X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$ as PDP^T with D a diagonal matrix with the eigenvalues and P orthogonal.

$$g(t) = \mathrm{Tr}\left(X^{-1}\left(Id + tPDP^T\right)^{-1}\right)$$
$$= \mathrm{Tr}\left(X^{-1}P\left(Id + tD\right)^{-1}P^T\right)$$
$$= \mathrm{Tr}\left(P^T X^{-1} P \left(Id + tD\right)^{-1}\right)$$

Since $(Id + tD)^{-1}$ is diagonal:

$$g(t) = \sum_{i=1}^{n} \underbrace{[P^T X^{-1} P]_{ii}}_{\geq 0 \text{ since } X \in S^n_{++}} (1 + t\lambda_i)^{-1}$$ where λ_i are the eigenvalues of $X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$

g is a positive weighted-sum of convex fonctions $t \mapsto \frac{1}{1+t\lambda_i}$ $(\lambda_i > 0)$ therefore g is convex.

Hence f is convex.

2) $f(X, y) = y^T X^{-1} y$ on $S_{++}^n \times \mathbb{R}^n$

Let's define $g(x) = \frac{1}{2} x^T X x$ on $\mathbb{R}^n$

Its conjugate function is:

$$g^*(y) = \sup_{x \in \mathbb{R}^n} \left(y^T x - f(x) \right) = \sup_{x \in \mathbb{R}^n} (\underbrace{y^T x - \frac{1}{2} x^T X x}_{= h(x)})$$

$h(x) = y^T x - \frac{1}{2} x^T X x$

$\nabla h(x) = y^T - x^T X$

$\nabla h(x^*) = 0 \implies x^* = X^{-1} y$

The maximum of h is obtained for $x^* = X^{-1} y$

Then: $g^*(y) = h(x^*) = \frac{1}{2} y^T X^{-1} y = \frac{1}{2} f(X, y)$

g^* is convex (as it is a conjugate function), therefore <u>f is convex.</u>

3) $f(X) = \sum_{i=1}^{n} \sigma_i(X)$ where $\sigma_1(X), \dots, \sigma_n(X)$ are the singular values of X.

Let $X = U \Sigma V^T$ with U, V $n \times n$ unitary matrix and Σ $n \times n$ diagonal matrix be the singular value decomposition of X.

We have that $[\Sigma]_{ii} = \sigma_i(X)$ for $i = 1, \dots, n$.

* On the one hand, let A be a $n \times n$ matrix such that $\sigma_i(A) \leq 1$ for $i = 1, \dots, n$

$$\langle A, X \rangle = \langle A, U \Sigma V^T \rangle = \langle U^T A V, \Sigma \rangle = \sum_{i=1}^{n} \sigma_i [U^T A V]_{ii} = \sum_{i=1}^{n} \sigma_i u_i^T A v_i$$

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

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

Thus: $\boxed{\sup_{A \mid \forall i, \sigma_i(A) \leq 1} \langle A, X \rangle \leq \sum_{i=1}^{n} \sigma_i(X)}$

* On the other hand, let $A = UV^T$. By construction, $\sigma_i(A) = 1$ for $i = 1, \dots, \tilde{n}$.
and: $\langle A, X \rangle = \langle UV^T, U\Sigma V^T \rangle = \mathrm{Tr}(VU^T U \Sigma V^T) = \mathrm{Tr}(V^T V \Sigma) = \sum_{i=1}^{\tilde{n}} \sigma_i(X)$

Thus: $$\boxed{\sup_{A,\ \forall i, \sigma_i(A) \leq 1} \langle A, X \rangle = \sum_{i=1}^{\tilde{n}} \sigma_i(X)}$$

$X \mapsto \langle A, X \rangle$ is convex for all A, so as a pointwise supremum, f is convex.

Exercice 4:

1) $K_{m+} = \{ x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \dots \geq x_n \geq 0 \}$

- K_{m+} is defined by n homogeneous linear inequalities (which are not strict), therefore it is closed.
- $x = (n, n-1, \dots, 2, 1) \in K_{m+}$, therefore K_{m+} is solid.
- All components of $x \in K_{m+}$ are positive, therefore K_{m+} cannot contains any line: it is pointed.

2) By definition, $K_{m+}^* = \{ y \in \mathbb{R}^n \mid \forall x \in K_{m+},\ y^T x \geq 0 \}$

Let $y \in \mathbb{R}^n$ and $x \in K_{m+}$.

$$y^T x = \sum_{i=1}^{\tilde{n}} x_i y_i$$

$$= \overbrace{(x_1 - x_2)}^{\geq 0} y_1 + \overbrace{(x_2 - x_3)}^{\geq 0}(y_1 + y_2) + \overbrace{(x_3 - x_4)}^{\geq 0}(y_1 + y_2 + y_3)$$
$$+ \dots + \underbrace{(x_{n-1} - x_n)}_{\geq 0}(y_1 + \dots + y_{n-1}) + \underbrace{x_n}_{\geq 0}(y_1 + \dots + y_n)$$

Since $x_i - x_{i+1} \geq 0$ for $i = 1, \dots, n-1$, and that it can take any value:

$$\forall x \in K_{m+},\ y^T x \geq 0 \iff y_1 \geq 0,\ y_1 + y_2 \geq 0, \dots, y_1 + \dots + y_n \geq 0$$

and: $$\boxed{K_{m+}^* = \{ y \in \mathbb{R}^n \mid y_1 \geq 0,\ y_1 + y_2 \geq 0, \dots, y_1 + \dots + y_n \geq 0 \}}$$