Convex Optimization Homework 1

Soël Megdoud

October 2024

1 Exercise 1

1) $C = \{x \in R^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, \ldots, n\}$.

Let $x, y \in C$ and $\theta \in [0,1], \forall i \in [1, n]$, we have:

$$\theta x + (1-\theta)y \leq \beta_i \theta + (1-\theta)\beta_i = \beta_i,$$

and

$$\alpha_i \leq \alpha_i \theta + (1-\theta)\alpha_i = \beta_i \leq \theta x + (1-\theta)y.$$

Therefore, C is a convex set.

2) $C = \{x \in R^2_+ \mid x_1 x_2 \geq 1\}$.

Let $\theta \in [0,1]$ we have:

$$(\theta x_1 + (1-\theta)y_1)(\theta x_2 + (1-\theta)y_2) = \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)(x_1 y_2 + y_1 x_2) \geq \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 \geq \theta^2 + (1-\theta)^2 \geq 1$$

because x and $y \in C$, therefore C in convex.

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

Let $y \in S$, $\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$ is convex as it is the halfspace of the median hyperplane between x and y. Indeed,

$$\|x - x_0\| = x \cdot x - 2x \cdot x_0 + x_0 \cdot x_0 \leq \|x - y\| = x \cdot x - 2x \cdot y + x_0 \cdot x_0$$

is equivalent to

$$x \cdot (2(y - x_0)) < y \cdot y - x_0 \cdot x_0$$

This is the equation of a halfspace where $(2(y - x_0))$ is the normal vector. This is true for each $y \in S$ therefore C is convex as the intersection of convex sets.

4) $C = \{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\}$ where $S, T \subseteq R^n$ and dist(x, S) $= \inf\{\|x - z\|_2 \mid z \in S\}$.

We can choose T as the unit ball :

$$T = \{x \mid \|x\| \leq 1\}$$

and S its complement:

$$S = \{x \mid \|x\| > 1\}$$

Let's choose x outside the ball so that $\|x\| > 1$, now let's choose $y = -x$. For $\theta = 1/2$:

$$\theta x + (1-\theta)y = 0$$

However, $\|0\| = 0$ so $\{x \mid \mathrm{dist}(x, S) > \mathrm{dist}(x, T)\}$, therefore C is not convex.

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

We can rewrite it as:

$$C = \bigcap_{s_2 \in S_2} \{x \in R^n \mid x + s_2 \in S_1\}$$

Each set of the form $\{x \in R^n \mid x + s_2 \in S_1\} = S_1 \cap \{x + s_2 \mid x \in R^n\}$ is convex as the intersection of S_1 which is convex, and $\{x + s_2 \ \mid x \in R^n\}$ which is also convex as the image of the convex set R^n by an affine function.

Thus, even if S_2 is not convex, C remains a convex set as the intersection of convex sets.

2 Exercise 2

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

R^2_{++} is a convex set so to determine the convexity of the function, we compute the Hessian matrix:

$$H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The eigen values of H are found from the characteristic polynomial:

$$\det(H - \lambda I) = \lambda^2 - 1 = 0 \quad \Rightarrow \quad \lambda_1 = 1, \quad \lambda_2 = -1.$$

Since one eigenvalue is positive and the other is negative, the function f is neither convex nor concave.

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

R^2_{++} is a convex set, the Hessian is given by

$$H = \begin{pmatrix} \frac{2}{x_1^3 y_1} & \frac{1}{x_2^2 y_1^2} \\ \frac{1}{x_2^2 y_1^2} & \frac{2}{x_1 y_2^3} \end{pmatrix}.$$

Now, fix $x, y \geq 0$. We must show that for $a, b \in R^2_{++}$, we have

$$(ab) \begin{pmatrix} \frac{2}{x_1^3 y_1} & \frac{1}{x_2^2 y_1^2} \\ \frac{1}{x_2^2 y_1^2} & \frac{2}{x_1 y_2^3} \end{pmatrix} (ab) = 2(a^2 x_1^{-3} y_1 {-}^{1} + b^2 x_1 {-}^{1} y_3^{-3} + ab x_2^{-2} y_2^{-2}) \geq 0$$

Let $v = a x_1^{\frac{-3}{2}} y_1^{\frac{-1}{2}}$ and $w = b x_1^{\frac{-1}{2}} y_2^{\frac{-3}{2}}$. Then, we need to show:

$$v^2 + w^2 + vw \geq 0$$

If both $v > 0$ and $w > 0$, or $v < 0$ and $w < 0$, this is true so let's assume $v < 0 < w$. We then have $vw \leq w^2$ so it is always true and the Hessian is positive semi-definite which implies that f is convex.

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

R^2_{++} is a convex set, the Hessian is:

$$H = \begin{pmatrix} 0 & \frac{1}{x_2^2} \\ \frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{pmatrix}$$

We have $\det(H) = \frac{-1}{x_2^4}$ which is always strictly negative therefore H cannot be positive semi-definite and f is not convex.

4) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$, for $(x_1, x_2) \in R^{++}$.

R^2_{++} is a convex set, the Hessian is

$$H = \begin{pmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & -(1-\alpha)\alpha x_1^{\alpha}x_2^{-\alpha-1} \end{pmatrix}.$$

Now let's calculate:

$$\det(H - \lambda I) = \left(\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} - \lambda\right)\left(\alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1} - \lambda\right) - \left(\alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha}\right)^2$$

So we have:

$$\det(H - \lambda I) = \lambda(\lambda - \alpha(\alpha-1)(x_1^{\alpha-2}x_2^{1-\alpha} + x_1^{\alpha}x_2^{-\alpha-1}))$$

Finally, either $\lambda = 0$ or $\lambda = \alpha(\alpha-1)(x_1^{\alpha-2}x_2^{1-\alpha} + x_1^{\alpha}x_2^{-\alpha-1})$. In all cases, λ is negative therefore f is concave.

3 Exercise 3

1) $f(X) = \mathbf{Tr}(X^{-1})$ on S_n^{++}

To show that f is convex, let's study $g(t) = f(Z + tV) = \mathrm{Tr}((Z + tV)^{-1})$, where $Z \in S_n^{++}$ and $V \in S_n$. We are going to check that g is convex for all $V \in S_n$.

We aim to study the inverse of the matrix $(Z + tV)$. We use the fact that for any positive definite matrix $Z \in S_n^{++}$, there exists a square root matrix $Z^{1/2}$ such that $Z = Z^{1/2}Z^{1/2}$ and $Z^{-1} = Z^{-1/2}Z^{-1/2}$.

We rewrite $Z + tV$ by factoring $Z^{1/2}$ as follows:

$$Z + tV = Z^{1/2}\left(I + tZ^{-1/2}VZ^{-1/2}\right)Z^{1/2}.$$

Thus, the inverse of $Z + tV$ can be written as:

$$(Z + tV)^{-1} = Z^{-1/2}\left(I + tZ^{-1/2}VZ^{-1/2}\right)^{-1}Z^{-1/2}.$$

We use the cyclic property of the trace, to obtain:

$$g(t) = \mathrm{Tr}\left(Z^{-1}(I + tZ^{-1/2}VZ^{-1/2})^{-1}\right)$$

Now, we apply the spectral decomposition to the symmetric matrix $Z^{-1/2}VZ^{-1/2}$. Since $Z^{-1/2}VZ^{-1/2}$ is symmetric, it can be decomposed as follows:

$$Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$$

where Q is an orthogonal matrix and Λ is a diagonal matrix containing the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Substituting this decomposition into the expression for $(Z + tV)^{-1}$, we get:

$$g(t) = \mathrm{Tr}\left(Z^{-1}Q(I + t\Lambda)^{-1}Q^T\right) = \mathrm{Tr}\left(Q^TZ^{-1}Q\,(I + t\Lambda)^{-1}\right).$$

Since Λ is a diagonal matrix, $(I + t\Lambda)^{-1}$ is also diagonal, with the i-th diagonal entry being $\frac{1}{1+t\lambda_i}$. This allows us to further simplify:

$$g(t) = \sum_{i=1}^{n} (Q^TZ^{-1}Q)_{ii}\,\frac{1}{1 + t\lambda_i}.$$

where $(Q^TZ^{-1}Q)_{ii}$ are the diagonal elements of the matrix $Q^TZ^{-1}Q$. Each term $\frac{1}{1+t\lambda_i}$ is a convex function of t.

Since this is a positive weighted sum of convex functions, we can conclude that $g(t)$ is convex, which implies that f is convex on S_n^{++}.

2) $f(X, y) = y^T X y$ on $S_n^{++} \times R^n$

According to the lecture:

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

Therefore we have:

$$f(y) = \sup_x \left(2y^T x - x^T X x \right)$$

Let's note $g(X, y, x) = 2y^T x - x^T X x$, g is convex in $X \in S_n^{++}$ and $y \in R^n$ for all x. Therefore, f is a convex function as a supremum of a convex function.

3) $f(X) = \sum_{i=1}^n \sigma_i(X)$, where $\sigma_i(X)$ are the singular values of a matrix X

We are going to show that

$$f(X) = \sup_{\sigma_{\max}(Q) \leq 1} \operatorname{Tr}(Q^T X).$$

1) Let's decompose X according to its singular value decomposition:

$$X = U E V^T,$$

and let's take $Q = UV^T$.

Then we have:

$$\operatorname{Tr}(Q^T X) = \operatorname{Tr}(V U^T U E V^T) = \operatorname{Tr}(V V^T U U^T E) = \operatorname{Tr}(E) = \sum_i \sigma_i(X) = f(X).$$

Therefore,

$$\sup_{\sigma_{\max}(Q) \leq 1} \operatorname{Tr}(Q^T X) \geq f(X).$$

2) Now, consider:

$$\sup_{\sigma_{\max}(Q) \leq 1} \operatorname{Tr}(Q^T X) = \sup_{\sigma_{\max}(Q) \leq 1} \operatorname{Tr}(V^T Q^T U E) = \sup_{\sigma_{\max}(Q) \leq 1} \sum_i \lambda_i \left(U^T Q V \right)_{ii}$$

Where λ_i are the eigen values of X. Since $\sigma_{\max}(Q) \leq 1$, we can write:

$$\sup_{\sigma_{\max}(Q) \leq 1} \operatorname{Tr}(Q^T X) \leq \sup_{\sigma_{\max}(Q) \leq 1} \sum_i \lambda_i \sigma_{\max}(Q) \leq \sup \sum_i \lambda_i = f(X).$$

Therefore, we conclude that

$$f(X) = \sup_{\sigma_{\max}(Q) \leq 1} \operatorname{Tr}(Q^T X).$$

$Tr(Q^T X)$ is convex in X for all Q with $\sigma_{\max}(Q) \leq 1$, as a linear function, therefore f is convex.