# DSA3102: Solutions to Homework 1
# Assigned: 24/08/23, Due: 07/09/2023

Instructions:

- Please do all the four problems below. A (non-empty) subset of them will be graded.
- Write (legibly) on pieces of paper or Latex your answers. Convert your submission to PDF format.
- Submit to the Assignment 1 folder under Assignments in Canvas before 23:59 on 07/09/2023.

1. Consider the set defined by the following inequalities

$$S = \{x \in \mathbb{R}^2 : x_1 \geq x_2 - 1, x_2 \geq 0\} \bigcup \{x \in \mathbb{R}^2 : x_1 \leq x_2 - 1, x_2 \leq 0\}$$

(a) Draw the set $S$. Is it convex?
**Solutions:** See Fig. 1. Clearly the set $S$ is not convex.

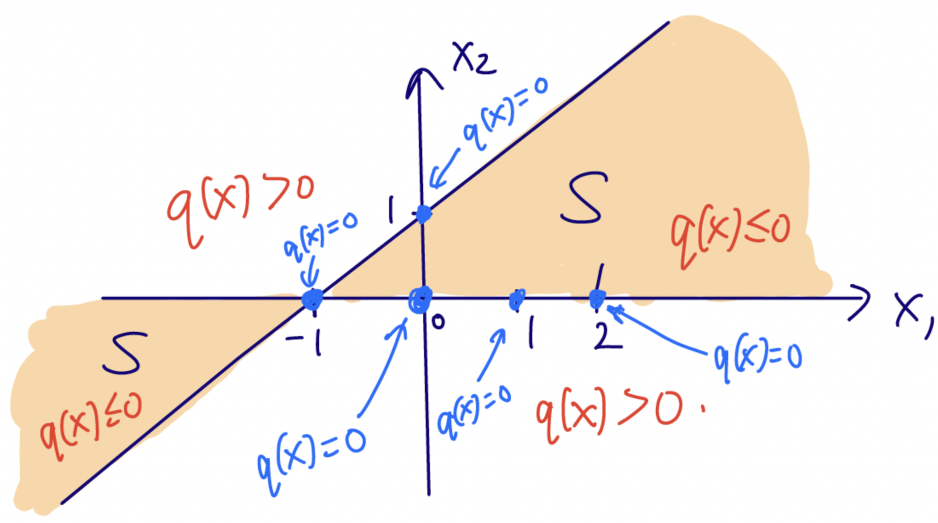


Figure 1: The set $S$ in Question 1

(b) Show that the set $S$ can be described as a single quadratic inequality of the form

$$q(x) = x^T Ax + 2b^T x + c \leq 0$$

for a matrix $A \in \mathbf{S}^2$ (2 by 2 symmetric matrix), $b \in \mathbb{R}^2$ and $c \in \mathbb{R}$ which you should determine.
*Hint: Note that $q$ is continuous, $q \leq 0$ on $S$ and $q > 0$ on $S^c$, hence $q$ must be 0 on the boundary of $S$.*

(c) What is the convex hull of the set $S$?

**Solutions:** We may write

$$q(x) = x^T Ax + 2b^T x + c = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} + 2 \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + c$$

We see from the figure that the points $(0,0)^T, (1,0)^T, (-1,0)^T, (0,1)^T, (1,2)^T$ are on the boundary of $S$. On these points $q(x) = 0$. We will use these points to solve for $a_{11}, a_{12}, a_{22}, b_1, b_2, c$.
First using $x = (0,0)^T$, we see that $x^T Ax + 2b^T x = 0$ and hence, $c = 0$.
Next using $x = (1,0)^T$, we see that $q(x) = a_{11} + 2b_1 = 0$.
Next using $x = (-1,0)^T$, we see that $q(x) = a_{11} - 2b_1 = 0$.
The last two observations imply that $a_11 = b_1 = 0$.
Next using $x = (0,1)^T$, we see that $q(x) = a_{22} - 2b_2 = 0$, which means that $a_{22} = 2b_2$.
Finally, using $x = (1,2)$, we see that $q(x) = 4a_{21} + 4a_{22} + 4b_2 = 4a_{21} + 6a_{22} = 0$. This means that $a_{21} = -\frac{3}{2}a_{22}$.
We may pick $b_2 = t \in \mathbb{R}$. Then $a_{22} = 2t$ and $a_{21} = -3t$ and $c = 0$. This means that

$$A = t \begin{bmatrix} 0 & -3 \\ -3 & 2 \end{bmatrix} \qquad b = t \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Thus,

$$q(x) = t\left[-6x_1x_2 + 2x_2^2 + 2x_2\right].$$

**Solutions:** The convex hull of $S$ is $\mathbb{R}^2$.

2. The following is a very important result in linear programming duality. Let $A \in \mathbb{R}^{m\times n}$ and $y \in \mathbb{R}^m$. Show that one and only one of the following two conditions is satisfied.

- the system of linear equations $Ax = y$ admits a non-negative solution $x \geq 0$;
- there exists $z \in \mathbb{R}^m$ such that $z^T A \geq 0$ and $z^T y < 0$.

You should use the *strict separating hyperplane* theorem.

**Solutions:** It is easy to see that both Statements 1 and 2 cannot hold simultaneously. From Statement 2, we see that there exists $z$ such that $z^T A \geq 0$ and $z^T y < 0$. Let the nonnegative solution of $Ax = y$ be $x$. Then $z^T Ax \geq 0$ since both $z^T A$ and $x$ are nonnegative. But $Ax = y$. Thus, $z^T y \geq 0$ contradicting $z^T y < 0$.

Now suppose Statement 1 does not hold. Then we claim that Statement 2 must hold. Suppose there is no $x \geq 0$ such that $Ax = y$. Thus $y \notin \mathrm{conic}(A)$, where $\mathrm{conic}(A)$ is the set of conic combinations of the columns of $A$. Since $\{y\}$ is convex, closed and bounded and $\mathrm{conic}(A)$ is convex and closed, by the strict separating hyperplane theorem, there exists a hyperplane $\{x : z^T x = q\}$ that strictly separates $\{y\}$ and $\mathrm{conic}(A)$, i.e.,

$$z^T y < q \qquad \text{and} \qquad z^T Av > q \quad \forall\, v \geq 0.$$

Note that $q$ is necessarily negative from the second condition above. (Suppose $q \geq 0$. Then take $v = 0$, which yields $z^T Av = 0$, which contradicts $z^T Av > q$.) This implies that $z^T y < 0$.

Now, we prove that $z^T Av > q$ for all $v \geq 0$ implies that $z^T A \geq 0$, which would conclude the proof (as this is precisely Statement 2). Suppose, to the contrary, that there exists a coordinate of the vector $z^T A$ that is negative. Say the coordinate is $i$, i.e., $(z^T A)_i < 0$. Then we set $v = Le_i \geq 0$ ($e_i$ is the $i$-th unit vector) for some $L > 0$. Then $z^T Av = L(z^T A)_i < 0$. By making $L > 0$ sufficiently large, $z^T Av = L(z^T A)_i < q$. This contradicts $z^T Av > q$. Thus Statement 2 is proved.

Technically, we also have to prove that the following cannot happen. (i) Both Statements 1 and 2 are not satisfied; (ii) Statement 2 does not hold implies Statement 1 holds. However (i) and (ii) are easy to deduce from the fact that "Statements 1 and 2 cannot hold simultaneously" and "Statement 1 does not hold implies Statement 2 must hold". Please convince yourself of this.

3. The *polar* of an arbitrary set $C \subset \mathbb{R}^n$ is defined as the set

$$C^\circ := \{y \in \mathbb{R}^n : y^T x \leq 1 \text{ for all } x \in C\}$$

(a) (3 points) Let $C \subset \mathbb{R}^n$ be any set, not necessarily convex. Is $C^\circ$ convex? Justify your answer carefully.

**Solution:** Yes $C^\circ$ is convex. It can be written as

$$C^\circ = \bigcap_{x \in C} \{y \in \mathbb{R}^n : y^T x \leq 1\}$$

Each set $\{y \in \mathbb{R}^n : y^T x \leq 1\}$ parametrized by $x \in C$ is a halfspace hence convex. Intersection of convex sets is convex.

(b) (5 points) Recall that a *cone* $K$ is such that if $x \in K$ and $\lambda \geq 0$, then $\lambda x \in K$. Recall that the *dual cone* is defined as

$$K^* := \{y \in \mathbb{R}^n : y^T x \geq 0 \text{ for all } x \in K\}$$

It is known that the polar of the cone $K$, denoted as $K^\circ$, and the dual cone, denoted as $K^*$, are related as follows:

$$K^\circ = -cK^*$$

for some *positive* number $c > 0$. Find $c$.

**Solution:** The answer is $c = 1$. We prove this in two parts first noting that

$$-K^* := \{y : y^T x \leq 0 \text{ for all } x \in K\}$$

$K^\circ \subset -K^*$: Take $y \in K^\circ$. This means that $y^T x \leq 1$ for all $x \in K$. Since $K$ is a cone, if $x \in K$, we have $\lambda x \in K$ for all $\lambda \geq 0$. This means that $y^T x \leq 1/\lambda$ for all $x \in K$ and all $\lambda \geq 0$. Now take $\lambda \to \infty$. Then we conclude that $y^T x \leq 0$ for all $x \in K$. This means that $y \in -K^*$.

$-K^* \subset K^\circ$: Take $y \in -K^*$. This means that $y^T x \leq 0$ for all $x \in K$. Clearly $y^T x \leq 1$ for all $x \in K$. This means that $y \in K^\circ$ as desired.

(c) (2 points) Show carefully that the polar of the unit ball $\mathcal{B}(0,1) := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ is the unit ball. You may find the Cauchy-Schwarz inequality (i.e., $x^T y \leq \|x\|_2 \|y\|_2$) useful.

**Solution:** We need to show that $\mathcal{B}(0,1)^\circ = \mathcal{B}(0,1)$.

Take $x \in \mathcal{B}(0,1)^\circ$. This means that $x^T y \leq 1$ for all $y \in \mathcal{B}(0,1)$. Suppose $\|x\|_2 > 1$. Now let $y = x/\|x\|_2$. Then the inner product $y^T x = \|x\|_2 > 1$, which contradicts $x^T y \leq 1$. This means that $\|x\|_2 \leq 1$, i.e., $x \in \mathcal{B}(0,1)$.

Now take $x \in \mathcal{B}(0,1)$. This means that $\|x\|_2 \leq 1$. Fix $y \in \mathcal{B}(0,1)$. We have

$$x^T y \leq \|x\|_2 \|y\|_2 \leq \|y\|_2 \leq 1$$

This means that $x \in \mathcal{B}(0,1)^\circ$.

4. (a) For $x, y$ both positive scalars, show that

$$ye^{x/y} = \sup_{\alpha > 0} \alpha(x+y) - y\alpha \log \alpha.$$

(log here refers to the natural logarithm.) Use the above result to prove that the function $f : \mathbb{R}^n_{++} \to \mathbb{R}$ defined as

$$f(x,y) = ye^{x/y}$$

is convex.

**Solutions:** Let $g(\alpha) = \alpha(x+y) - y\alpha\log\alpha$. Then

$$g'(\alpha) = (x+y) - y\alpha \cdot \frac{1}{\alpha} - y\log\alpha \cdot .$$

Setting this derivative to zero, we obtain

$$x - y\log\alpha = 0$$

or

$$\alpha^* = e^{x/y}$$

Since $\alpha^* > 0$, it satisfies the constraint automatically, and hence

$$g(\alpha^*) = e^{x/y}\,(x+y) - y\,e^{x/y}\,\log(e^{x/y}) = y\,e^{x/y}$$

as desired.

Since $f$ is the pointwise supremum of a bunch of affine, hence convex, functions, it is also convex.

(b) Show that the function $f : \mathbb{R}^2_{++} \to \mathbb{R}$

$$f(x) = \left(\sqrt{x_1} + \sqrt{x_2}\right)^2$$

is concave by examining the definiteness of its Hessian.

**Solution:** We can write $f(x) = x_1 + x_2 + 2\sqrt{x_1 x_2}$. The gradient vector is

$$\nabla f(x) = \begin{bmatrix} 1 + \sqrt{\frac{x_2}{x_1}} & 1 + \sqrt{\frac{x_1}{x_2}} \end{bmatrix}^T.$$

The Hessian is

$$\nabla^2 f(x) = \begin{bmatrix} -\frac{1}{2}\sqrt{\frac{x_2}{x_1^3}} & \frac{1}{2}\frac{1}{\sqrt{x_1 x_2}} \\ \frac{1}{2}\frac{1}{\sqrt{x_1 x_2}} & -\frac{1}{2}\sqrt{\frac{x_1}{x_2^3}} \end{bmatrix}.$$

We check the definiteness of $\nabla^2 f$. We can write $\nabla^2 f$ as

$$\nabla^2 f(x) = -\frac{1}{2\sqrt{x_1 x_2}} v v^T$$

where

$$v = \left[\sqrt{\frac{x_2}{x_1}}, \sqrt{\frac{x_1}{x_2}}\right]^T.$$

Hence $\nabla^2 f(x)$ is negative semidefinite.

Another way to prove that $\nabla^2 f(x)$ is negative semidefinite is to show that $-\nabla^2 f(x)$ is positive semidefinite. That is, we want to show that

$$A = -\nabla^2 f(x) = \begin{bmatrix} \frac{1}{2}\sqrt{\frac{x_2}{x_1^3}} & -\frac{1}{2}\frac{1}{\sqrt{x_1 x_2}} \\ -\frac{1}{2}\frac{1}{\sqrt{x_1 x_2}} & \frac{1}{2}\sqrt{\frac{x_1}{x_2^3}} \end{bmatrix}$$

is PSD. We check all the all the principal minors are nonnegative. The (1,1) and (2,2) entries $\frac{1}{2}\sqrt{\frac{x_2}{x_1^3}}$ and $\frac{1}{2}\sqrt{\frac{x_1}{x_2^3}}$ are clearly positive. The determinant is

$$\det(A) = \frac{1}{2}\sqrt{\frac{x_2}{x_1^3}} \cdot \frac{1}{2}\sqrt{\frac{x_1}{x_2^3}} - \left(-\frac{1}{2}\frac{1}{\sqrt{x_1 x_2}}\right)^2 = 0.$$

Hence $A = -\nabla^2 f(x)$ is positive semidefinite showing that $f$ is concave. Note that you need to check that *all* the principal minors are nonnegative; it is not enough to check that the upper left minors are nonnegative.