Optimization 1 - 098311

Winter 2021 - HW 5

Ido Czerninski 312544596, Asaf Gendler 301727715

November 26, 2020

Problem 1:

let $A \in \mathbb{R}^{mxn}$, and $C \subseteq \mathbb{R}^n$ and $D \subseteq \mathbb{R}^m$ convex sets.

a)

prove that the set $A(C)$ is convex:

$$A(C) = \{Ax \in \mathbb{R}^n : x \in C\}$$

let $z, y \in A(C)$, from $A(C)$ definition:

$$z = Ax_z : x_z \in C$$

$$y = Ax_y : x_y \in C$$

for $\lambda \in [0, 1]$ let's look at:

$$\lambda y + (1-\lambda) z = \lambda A x_y + (1-\lambda) A x_z = A(\lambda x_y + (1-\lambda) x_z)$$

because C is a convex set, and $x_z, x_y \in C$:

$$\lambda x_y + (1-\lambda) x_z \in C$$

thus from $A(C)$ definition:

$$A(\lambda x_y + (1-\lambda) x_z) \in A(C)$$

$\forall \lambda \in [0, 1]$, $z, y \in A(C)$:

$$\lambda y + (1-\lambda) z \in A(C)$$

and therefore $A(C)$ is a convex set by definition.

b)

prove that the set $A^{-1}(D)$ is convex:

$$A^{-1}(D) = \{x \in \mathbb{R}^n : Ax \in D\}$$

let $z, y \in A^{-1}(D)$, from $A^{-1}(D)$ definition:

$$Az \in D$$

$$Ay \in D$$

for $\lambda \in [0,1]$ let's look at:

$$A\left(\lambda y + (1-\lambda) z\right) = \lambda Ay + (1-\lambda) Az$$

because D is a convex set, and $Az, Ay \in D$:

$$\lambda Ay + (1-\lambda) Az \in D$$

we found that:

$$A\left(\lambda y + (1-\lambda) z\right) \in D$$

thus from $A^{-1}(D)$ definition:

$$\lambda y + (1-\lambda) z \in A^{-1}(D)$$

$\forall \lambda \in [0,1]$, $z, y \in A^{-1}(D)$:

$$\lambda y + (1-\lambda) z \in A^{-1}(D)$$

and therefore $A^{-1}(D)$ is a convex set by definition.

Problem 2:

Let $a \neq b \in \mathbb{R}^n$, Find the values of $\mu \in \mathbb{R}$ for which the set S_μ is convex.

$$S_\mu = \{x \in \mathbb{R}^n : \|x-a\| \leq \mu \|x-b\|\}$$

for $0 < \mu \leq 1$**:**

First, let's try to define the set in a different way:

$$\begin{aligned}
& \|x-a\| \leq \mu \|x-b\| \\
\text{all positive} \iff & \|x-a\|^2 \leq \mu^2 \|x-b\|^2 \\
\iff & x^T x - 2a^T x + a^T a \leq \mu^2 x^T x - 2\mu^2 b^T x - \mu^2 b^T b \\
\iff & \underbrace{\left(1-\mu^2\right) x^T x - 2\left(a^T - \mu^2 b^T\right) x + a^T a - \mu^2 b^T b}_{\triangleq f(x)} \leq 0 \\
\iff & f(x) \leq 0
\end{aligned}$$

hence:

$$S_\mu = \{x \in \mathbb{R}^n : \|x-a\| \leq \mu \|x-b\|\} = \{x \in \mathbb{R}^n : f(x) \leq 0\}$$

Easy way

$$\begin{aligned}
f(x) &= \left(1-\mu^2\right) x^T x - 2\left(a^T - \mu^2 b^T\right) x + a^T a - \mu^2 b^T b \\
&= x^T \underbrace{\left(1-\mu^2\right) I}_{\succeq 0 \ (0<\mu\leq 1)} x - 2\left(a^T - \mu^2 b^T\right) x + a^T a - \mu^2 b^T b
\end{aligned}$$

hence S_μ is ellipsoid which is convex (as shown in the lecture).

Hard way

Let $x, y \in S_\mu$ and define:

$$z = tx + (1-t)y, \ t \in (0,1)$$

We need to show that:

$$z \in S_\mu$$

Proof:

$$x, y \in S_\mu \Rightarrow f(x), f(y) \leq 0$$

$$\begin{aligned}
f(z) &= \left(1-\mu^2\right) z^T z - 2\left(a^T - \mu^2 b^T\right) z + a^T a - \mu^2 b^T b \\
&= \left(1-\mu^2\right)(tx + (1-t)y)^T (tx + (1-t)y) - 2\left(a^T - \mu^2 b^T\right)(tx + (1-t)y) + a^T a - \mu^2 b^T b \\
&= \left(1-\mu^2\right)\left[t^2 x^T x + (1-t)^2 y^T y + 2t(1-t)x^T y\right] \\
&- t \cdot 2\left(a^T - \mu^2 b^T\right) x - (1-t) \cdot 2\left(a^T - \mu^2 b^T\right) y + a^T a - \mu^2 b^T b \\
(*) &\leq \left(1-\mu^2\right)\left[t^2 x^T x + (1-t)^2 y^T y + t(1-t)\left(x^T x + y^T y\right)\right] - ... \\
&= \left(1-\mu^2\right)\left[\left(t^2 + t(1-t)\right) x^T x + \left((1-t)^2 + t(1-t)\right) y^T y\right] - ... \\
&= \left(1-\mu^2\right)\left[\left(t^2 + t - t^2\right) x^T x + ((1-t)(1-t+t)) y^T y\right] = ... \\
&= \left(1-\mu^2\right)\left[t x^T x + (1-t) y^T y\right] - t \cdot 2\left(a^T - \mu^2 b^T\right) x - (1-t) \cdot 2\left(a^T - \mu^2 b^T\right) y \\
&+ (t + (1-t))\left(a^T a - \mu^2 b^T b\right) \\
&= t\left[\left(1-\mu^2\right) x^T x - 2\left(a^T - \mu^2 b^T\right) x + a^T a - \mu^2 b^T b\right] \\
&+ (1-t)\left[\left(1-\mu^2\right) y^T y - 2\left(a^T - \mu^2 b^T\right) y + a^T a - \mu^2 b^T b\right] \\
&= t f(x) + (1-t) f(y) \leq 0
\end{aligned}$$

$$f(z) \leq 0 \Rightarrow z \in S_\mu \Rightarrow S_\mu \text{ is convex}$$

Now we need to show why $(*)$ holds :

$$x^T y = \langle x, y\rangle \underbrace{\leq}_{\text{C.S}} \|x\| \cdot \|y\| = \underbrace{\sqrt{x^T x}}_{a} \cdot \underbrace{\sqrt{y^T y}}_{b} \leq \frac{1}{2}\left(\underbrace{x^T x}_{a^2} + \underbrace{y^T y}_{b^2}\right)$$

Moreover:

$$2\underbrace{\left(1-\mu^2\right)}_{>0}\underbrace{t}_{>0}\underbrace{(1-t)}_{>0} x^T y \leq \left(1-\mu^2\right) t(1-t)\left(x^T x + y^T y\right)$$

where $1-\mu^2 \geq 0$ since $0 < \mu < 1$ (and the same for t and $(1-t)$)

for $\mu = 0$:

$$\begin{aligned} S_0 &= \{x \in \mathbb{R}^n : \|x - a\| \leq 0 \, \|x - b\|\} \\ &= \{x \in \mathbb{R}^n : \|x - a\| \leq 0\} \\ \text{(norms are non-negative)} &= \{x \in \mathbb{R}^n : \|x - a\| = 0\} \\ &= \{a\} \end{aligned}$$

which is convex trivially since a convex combination can not be created using only one element.

for $\mu < 0$:

Since all norms are non-negative:

$$S_\mu = \emptyset$$

which is a convex set trivially (same as above).

for $\mu > 1$:

We will show that S_μ is not convex:

define:

$$x = a$$

$$y = \frac{a - \mu b}{1 - \mu}$$

$$\|x - a\| = \|a - a\| = 0 \leq \mu \, \|x - b\|$$

hence:

$$x \in S_\mu$$

moreover:

$$
\begin{aligned}
\|y-a\| &= \left\|\frac{a-\mu b}{1-\mu}-a\right\| = \left\|\frac{a-\mu b-a+\mu a}{1-\mu}\right\| \\
&= \left\|\frac{-\mu b+\mu a}{1-\mu}\right\| = |\mu|\left\|\frac{a-b}{1-\mu}+b-b\right\| \\
&= \mu\left\|\frac{a-b+b-\mu b}{1-\mu}-b\right\| = \mu\left\|\frac{a-\mu b}{1-\mu}-b\right\| \\
&= \mu\|y-b\| \leq \mu\|y-b\|
\end{aligned}
$$

hence:

$$y \in S_\mu$$

however, if we choose $\lambda = \frac{1}{\mu} \in [0,1]$ we get:

$$
\begin{aligned}
z = \lambda x + (1-\lambda)\, y &= \frac{a}{\mu} + \left(1-\frac{1}{\mu}\right)\cdot\left(\frac{a-\mu b}{1-\mu}\right) \\
&= \frac{a}{\mu} - \left(\frac{1-\mu}{\mu}\right)\cdot\left(\frac{a-\mu b}{1-\mu}\right) = \frac{a}{\mu} - \frac{a-\mu b}{\mu} \\
&= \frac{a}{\mu} - \frac{a}{\mu} + b = b
\end{aligned}
$$

z is attained by a convex combination of $x, y \in S_\mu$ but:

$$\|z-a\| = \|b-a\| \underbrace{>}_{a\neq b} 0 = \|b-b\| = \mu\|z-b\|$$

hence:

$$z \notin S_\mu$$

we found $x, y \in S_\mu$ and a scalar $\lambda \in [0,1]$ such that $\lambda x + (1-\lambda)\, y \notin S_\mu$, hence S_μ is not convex by definition. the intuition for that is that we showed earlier that the set is in fact an ellipsoid. for $\mu \in (0,1]$ we got the interior of the ellipsoid, which is a convex set, for $\mu > 1$ we get the exterior of the ellipsoid, which of course is non convex.

To conclude:

$$
\begin{cases}
S_\mu \text{ is convex} & \mu \leq 1 \\
S_\mu \text{ is not convex} & \mu > 1
\end{cases}
$$

Problem 3:

show that the conic hull of the set:

$$S = \left\{x \in \mathbb{R}^2 : (x_1 - 1)^2 + x_2^2 = 1\right\}$$

is the set $\{x \in \mathbb{R}^2 : x_1 > 0\} \cup \{(0,0)\}$

first direction:

let $y \in cone\,(S)$, and let's prove that $y \in \{x \in \mathbb{R}^2 : x_1 > 0\} \cup \{(0,0)\}$

y can be $(0,0)$ as the zero vector always belongs to the conic hull.

in this case it's trivial that $y \in \{x \in R^2 : x_1 > 0\} \cup \{(0,0)\}$

let's assume $y \neq (0,0)$.

such y does exist, for example $y = (2,0) \in cone\,(S)$ because:

$$y = \underbrace{1}_{\geq 0} \cdot \underbrace{(2,0)}_{\in S}$$

because $y \in cone\,(S)$ it can be written as:

$$y = \sum_{i=1}^{k} \lambda_i x_i : x_1, x_2, ..., x_k \in S, \lambda_1, \lambda_2, ..., \lambda_k \in \mathbb{R}_+$$

S is a circle with a center at $(1,0)$ and a radios of 1, thus:

$$\forall x \in S, x_1 \geq 0$$

and equality holds if and only if $x = (0,0)$.

let's look at the first coordinate of y:

$$y_1 = \sum_{i=1}^{k} \underbrace{\lambda_i}_{\geq 0} \underbrace{x_{i_1}}_{\geq 0} \geq 0$$

and equality will hold if and only if:

$$\forall \lambda_i \neq 0 \to x_{i_1} = 0$$

$$\iff \forall \lambda_i \neq 0 \to x_i = (0,0)$$

$$\iff y_1 = (0,0)$$

but we assumed $y_1 \neq (0,0)$, thus the equality is strict and:

$$y_1 > 0$$

hence:

$$y \in \left\{x \in \mathbb{R}^2 : x_1 > 0\right\} \cup \{(0,0)\}$$

we proved that every y that belongs to $cons\,(S)$ also belongs to $\{x \in \mathbb{R}^2 : x_1 > 0\} \cup \{(0,0)\}$ thus:

$$\boxed{cone\,(S) \subseteq \left\{x \in \mathbb{R}^2 : x_1 > 0\right\} \cup \{(0,0)\}}$$

second direction:

now let $y \in \{x \in \mathbb{R}^2 : x_1 > 0\} \cup \{(0,0)\}$, and let's prove that $y \in cone\,(S)$

if $y = (0,0)$ than $y \in cone\,(S)$ as we said before that the zero vector always belongs to the conic hull.

if $y \in \{x \in \mathbb{R}^2 : x_1 > 0\}$, lets show it can be written as a conic combination of vectors from S.

let's assume

$$y = \lambda v \longrightarrow v = \frac{y}{\lambda}$$

for some $v \in S$, $\lambda \in \mathbb{R}_{++}$ then:

$$\begin{aligned} &(v_1 - 1)^2 + v_2^2 = 1 \\ \iff &\left(\frac{y_1}{\lambda} - 1\right)^2 + \left(\frac{y_2}{\lambda}\right)^2 = 1 \\ \iff &\frac{y_1^2}{\lambda^2} - 2\frac{y_1}{\lambda} + 1 + \frac{y_2^2}{\lambda^2} = 1 \\ \iff &\frac{y_1^2}{\lambda^2} - 2\frac{y_1}{\lambda} + \frac{y_2^2}{\lambda^2} = 0 \end{aligned}$$

because $\lambda \in \mathbb{R}_{++}$ we can multiply by λ^2:

$$\iff y_1^2 - 2y_1\lambda + y_2^2 = 0$$

$$\iff 2y_1\lambda = y_1^2 + y_2^2$$

$y \in \{x \in \mathbb{R}^2 : x_1 > 0\}$ thus $y_1 > 0$ and we can divide by it:

$$\iff \lambda = \frac{y_1^2 + y_2^2}{2y_1}$$

looking at what we did in the opposite direction, if we choose:

$$\lambda = \frac{y_1^2 + y_2^2}{2y_1} > 0$$

then the vector $v = \frac{y}{\lambda}$ satisfies:

$$(v_1 - 1)^2 + v_2^2 = 1$$

thus $v \in S$ and for $\lambda \in \mathbb{R}_+$:

$$y = \lambda v$$

therefore y is a conic combination of vectors from S, and by definition:

$$y \in con\,(S)$$

we proved that every y that belongs to $\{x \in \mathbb{R}^2 : x_1 > 0\} \cup \{(0,0)\}$ also belongs to $cone\,(S)$ thus:

$$\boxed{\left\{x \in \mathbb{R}^2 : x_1 > 0\right\} \cup \{(0,0)\} \subseteq cone\,(S)}$$

combining both the directions we have proved, we get:

$$cone\,(S) = \left\{x \in \mathbb{R}^2 : x_1 > 0\right\} \cup \{(0,0)\}$$

Problem 4:

Let $\emptyset \neq S \subseteq \mathbb{R}^n$ and let $\bar{x} \in S$. Consider the set:

$$C_{\bar{x}} = \{y \in \mathbb{R}^n : y = \lambda (x - \bar{x}), \lambda \geq 0, x \in S\}$$

a)

Show that $C_{\bar{x}}$ is a cone and interpret it geometrically.

from a geometrical perspective $C_{\bar{x}}$ is a set of all the points on rays from the origin such that if those rays were taken from $\bar{x}$, they would have an intersection with at least one point in S.

let $z \in C_{\bar{x}}$, than by the set definition:

$$z = \lambda_z (x_z - \bar{x}), \lambda_z \geq 0, x_z \in S$$

let's look at λz for some $\lambda \in \mathbb{R}_+$:

$$\lambda z = \underbrace{\lambda \lambda_z}_{\tilde{\lambda}} (x_z - \bar{x}) = \tilde{\lambda} (x_z - \bar{x})$$

since $\lambda_z, \lambda \in \mathbb{R}_+$ than also $\tilde{\lambda} \in \mathbb{R}_+$

$\tilde{\lambda} \in \mathbb{R}_+$ and $x_z \in S$ thus by $C_{\bar{x}}$ definition:

$$\lambda z = \tilde{\lambda} (x_z - \bar{x}) \in C_{\bar{x}}$$

we proved that $\forall z \in C_{\bar{x}}, \lambda \in \mathbb{R}_+ \rightarrow \lambda z \in C_{\bar{x}}$, thus $C_{\bar{x}}$ is a cone by definition.

b)

Show that if S is convex than $C_{\bar{x}}$ is convex.

since we already showed that $C_{\bar{x}}$ is a cone, using the lemma from the lecture, we just need to show that:

$$z_1, z_2 \in C_{\bar{x}} \Rightarrow z_1 + z_2 \in C_{\bar{x}}$$

let $z_1, z_2 \in C_{\bar{x}}$, than by the $C_{\bar{x}}$ definition:

$$z_1 = \lambda_{z_1} (x_{z_1} - \bar{x}), \lambda_{z_1} \geq 0, x_{z_1} \in S$$

$$z_2 = \lambda_{z_2} (x_{z_2} - \bar{x}), \lambda_{z_2} \geq 0, x_{z_2} \in S$$

now let's look at $z_1 + z_2$:

$$\begin{aligned} z_1 + z_2 &= \lambda_{z_1} (x_{z_1} - \bar{x}) + \lambda_{z_2} (x_{z_2} - \bar{x}) = \\ &= \lambda_{z_1} x_{z_1} - \lambda_{z_1} \bar{x} + \lambda_{z_2} x_{z_2} - \lambda_{z_2} \bar{x} = \\ &= \lambda_{z_1} x_{z_1} + \lambda_{z_2} x_{z_2} - (\lambda_{z_1} + \lambda_{z_2}) \bar{x} = \\ &\overset{\{*\}}{=} (\lambda_{z_1} + \lambda_{z_2}) \left(\frac{\lambda_{z_1} x_{z_1} + \lambda_{z_2} x_{z_2}}{\lambda_{z_1} + \lambda_{z_2}} - \bar{x} \right) = \\ &= (\lambda_{z_1} + \lambda_{z_2}) \left(\frac{\lambda_{z_1}}{\lambda_{z_1} + \lambda_{z_2}} x_{z_1} + \frac{\lambda_{z_2}}{\lambda_{z_1} + \lambda_{z_2}} x_{z_2} - \bar{x} \right) \end{aligned}$$

if we denote:

$$\tilde{\lambda} = \frac{\lambda_{z_1}}{\lambda_{z_1} + \lambda_{z_2}} \in [0, 1]$$

then:

$$1 - \tilde{\lambda} = 1 - \frac{\lambda_{z_1}}{\lambda_{z_1} + \lambda_{z_2}} = \frac{\lambda_{z_1} + \lambda_{z_2} - \lambda_{z_1}}{\lambda_{z_1} + \lambda_{z_2}} = \frac{\lambda_{z_2}}{\lambda_{z_1} + \lambda_{z_2}}$$

hence:

$$z_1 + z_2 = (\lambda_{z_1} + \lambda_{z_2}) \left(\tilde{\lambda} x_{z_1} + \left(1 - \tilde{\lambda}\right) x_{z_2} - \bar{x} \right)$$

$x_{z_1}, x_{z_2} \in S$ and $\tilde{\lambda} \in [0, 1]$, then because S is convex:

$$v = \tilde{\lambda} x_{z_1} + \left(1 - \tilde{\lambda}\right) x_{z_2} \in S$$

in addition $\bar{\lambda} = \lambda_{z_1} + \lambda_{z_2} \in \mathbb{R}_+$ since each one of terms belongs to $\mathbb{R}_+$.

now from $C_{\bar{x}}$ definition:

$$z_1 + z_2 = \bar{\lambda} (v - \bar{x}) \in C_{\bar{x}}$$

$\{*\}$ this equality is true only if $\lambda_{z_1} + \lambda_{z_2} > 0$.

since $\lambda_{z_1}, \lambda_{z_2} \in \mathbb{R}_+$, it always holds that $\lambda_{z_1} + \lambda_{z_2} \geq 0$. if $\lambda_{z_1} + \lambda_{z_2} = 0$, it means $\lambda_{z_1} = \lambda_{z_2} = 0$. in this case $z_1 + z_2 = 0$ and $0 \in C_{\bar{x}}$ and the proof still holds.

$C_{\bar{x}}$ is a cone, and $z_1, z_2 \in C_{\bar{x}} \Rightarrow z_1 + z_2 \in C_{\bar{x}}$, hence $C_{\bar{x}}$ is a convex cone.

c)

Suppose that S is closed, it is not necessarily means that $C_{\bar{x}}$ is closed.

let's take for example the set from pre-lecture 6 quiz in $\mathbb{R}^2$:

$$S = \left\{(x, y) : y = e^{-x}, x \geq 0\right\} \cup \{(0, 0)\}$$

this is a closed set.

let's take $\bar{x} = (0, 0) \in S$.

$$C_{\bar{x}} = \left\{y \in \mathbb{R}^2 : y = \lambda x, \lambda \geq 0, x \in S\right\}$$

$\boxed{\text{in this case } C_{\bar{x}} \text{ is the entire first quadrant but without the positive } x \text{ axis, thus } C_{\bar{x}} \text{ is not closed.}}$

conditions under which $C_{\bar{x}}$ will be closed:

1)

$\boxed{\text{if } S \text{ is a finite set, than } C_{\bar{x}} \text{ is close.}}$

proof:

first every finite set is closed, so the condition holds.

S is finite, thus it has k elements for some $k \in \mathbb{N}$. (k can't be zero because S is non empty)

let $\bar{x} \in S$.

denote the elements of S as $\{a_i\}_{i=1}^k$, and define the set that contains $a_i - \bar{x}$ solely:

$$S_i = \{a_i - \bar{x}\}$$

cone (S_i) is the set $\{y \in \mathbb{R}^n : y = \lambda (a_i - \bar{x}), \lambda \geq 0\}$ and is a closed set using the theorem from the lecture about a conic hull of finite sets.

$C_{\bar{x}}$ can be written as:

$$C_{\bar{x}} = \bigcup_{i=1}^{k} cone\,(S_i)$$

a finite union of closed sets is closed, thus $C_{\bar{x}}$ is closed.

2)

$\boxed{\text{if } \bar{x} \in int\,(S) \text{ then } C_{\bar{x}} \text{ is closed.}}$

proof:

we will show that $C_{\bar{x}} = \mathbb{R}^n$ hence $C_{\bar{x}}$ is closed.

let $y \in C_{\bar{x}}$, then by the definition of $C_{\bar{x}}$, $y \in \mathbb{R}^n \longrightarrow C_{\bar{x}} \subseteq \mathbb{R}^n$

let $y \in \mathbb{R}^n$

if $y = 0_n$ then for $x = \bar{x} \in S$ and $\lambda = \sqrt{\pi} \geq 0$ we get:

$$y = \lambda \left(x - \bar{x}\right)$$

hence $y \in C_{\bar{x}}$

if $y \neq 0_n$ then:

since $\bar{x} \in int\left(S\right)$ then:

$$\exists r > 0 : B\left(\bar{x}, r\right) \subseteq S$$

define:

$$z \overset{\|y\|>0}{=} \bar{x} + \frac{r}{2\left\|y\right\|} y$$

notice that:

$$\left\|z - \bar{x}\right\| = \left\|\bar{x} + \frac{r}{2\left\|y\right\|} y - \bar{x}\right\| = \left\|\frac{r}{2\left\|y\right\|} y\right\| = \frac{r}{2}\left\|\frac{y}{\left\|y\right\|}\right\| = \frac{r}{2} < r$$

hence:

$$z \in B\left(\bar{x}, r\right) \subseteq S \longrightarrow z \in S$$

if we choose:

$$\lambda = \frac{2\left\|y\right\|}{r} \geq 0$$

then:

$$\lambda\left(z - \bar{x}\right) = \frac{2\left\|y\right\|}{r}\left(\bar{x} + \frac{r}{2\left\|y\right\|} y - \bar{x}\right) = \frac{2\left\|y\right\|}{r}\left(\frac{r}{2\left\|y\right\|} y\right) = y$$

therefore $y \in C_{\bar{x}} \longrightarrow \mathbb{R}^n \subseteq C_{\bar{x}}$

to conclude $C_{\bar{x}} = R^n$ which is a closed set

Problem 5:

let $a_1, a_2 \in K$, by K definition:

$$x^* = \arg\min_{x\in S}\left\{a_1^T x\right\} = \arg\min_{x\in S}\left\{a_2^T x\right\}$$

note: we use the terminology arg min in this question, although the minimum doesn't have to be unique, by saying this we mean x^* is one of the arguments that minimizes the expression.

let $\lambda \in [0,1]$, notice that both $\lambda \geq 0$ and $(1-\lambda) \geq 0$ hence:

$$\left(\lambda a_1 + (1-\lambda) a_2\right)^T x = \lambda a_1^T x + (1-\lambda) a_2^T x \geq \lambda a_1^T x^* + (1-\lambda) a_2^T x^* = \left(\lambda a_1 + (1-\lambda) a_2\right)^T x^*$$

and of course for $x = x^* \in S$:

$$\left(\lambda a_1 + (1-\lambda) a_2\right)^T x = \left(\lambda a_1 + (1-\lambda) a_2\right)^T x^*$$

thus:

$$x^* = \arg\min_{x\in S}\left\{\left(\lambda a_1 + (1-\lambda) a_2\right)^T x\right\}$$

meaning $\lambda a_1 + (1-\lambda) a_2 \in K$

we proved that $\forall a_1, a_2 \in K, \lambda \in [0,1] \to \lambda a_1 + (1-\lambda) a_2 \in K$, thus K is convex by definition.

let $\lambda_2 \in \mathbb{R}_+$

if $\lambda_2 = 0$ then $(\lambda_2 a_1)^T x = 0$ and every vector $x \in S$ is an optimal solution for $\min_{x\in S}\left\{(\lambda_2 a_1)^T x\right\}$,specifically x^* is also an optimal solution and thus $\lambda_2 a_1 \in K$.

if $\lambda_2 > 0$:

$$\arg\min_{x\in S}\left\{(\lambda_2 a_1)^T x\right\} = \arg\min_{x\in S}\left\{\lambda_2 a_1^T x\right\} \overset{\lambda_2\geq 0}{=} \arg\min_{x\in S}\left\{a_1^T x\right\} = x^*$$

thus again $\lambda_2 a_1 \in K$

we proved that $\forall a_1 \in K, \lambda_2 \in \mathbb{R}_+ \to \lambda_2 a_1 \in K$, thus K is a cone by definition.

we proved K is both a cone and convex, therefore K is a convex cone.

we could have made things easier, since K is a cone we could have just check if $a_1 + a_2 \in K$ which indeed holds:

$$(a_1 + a_2)^T x = a_1^T x + a_2^T x \geq a_1^T x^* + a_2^T x^* = (a_1 + a_2)^T x^*$$

and for $x = x^*$:

$$(a_1 + a_2)^T x = x^*$$

thus:

$$x^* = \arg\min_{x\in S}\left\{(a_1+a_2)^T x\right\}$$

meaning $a_1 + a_2 \in K$

Problem 6:

a)

Show that the extreme points of $B_\infty[0_n, 1] = \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1\}$ are $\{-1, 1\}^n$

let $x \in \{-1, 1\}^n \subseteq B_\infty[0_n, 1]$

let's assume by contradiction that x is not an extreme point of $B_\infty[0_n, 1]$, thus:

$$\exists z_1 \neq z_2 \in B_\infty[0_n, 1], \lambda \in (0,1) : x = \lambda z_1 + (1-\lambda) z_2$$

$z_1 \neq z_2 \in B_\infty[0_n, 1]$ therefore:

$$\|z_1\|_\infty \leq 1 \rightarrow \forall j : |z_{1_j}| \leq 1$$

$$\|z_2\|_\infty \leq 1 \rightarrow \forall j : |z_{2_j}| \leq 1$$

all the coordinates of x are either 1 or -1 hence :

$$\forall j : \lambda z_{1_j} + (1-\lambda) z_{2_j} = \pm 1$$

since $z_1 \neq z_2$ there exists at least one coordinate which satisfies $z_{1_k} \neq z_{2_k}$

z_{1_k} and z_{2_k} can't be the negative of one another because than:

$$|\lambda z_{1_k} + (1-\lambda) z_{2_k}| = |\lambda z_{1_k} - (1-\lambda) z_{1_k}| = |(\lambda - 1 + \lambda) z_{1_k}| = |(2\lambda - 1) z_{1_k}|$$

$$= |(2\lambda - 1)| |z_{1_k}| \overset{\substack{\lambda \in (0,1) \\ |z_{1_k}| \leq 1}}{<} 1 \cdot 1 = 1$$

since $|z_{1_k}| \leq 1, |z_{2_k}| \leq 1, z_{1_k} \neq \pm z_{2_k}$, at least one of them as to be strictly smaller than one:

$$|z_{1_k}| < 1 \quad or \quad |z_{2_k}| < 1$$

in this case:

$$|\lambda z_{1_k} + (1-\lambda) z_{2_k}| \leq |\lambda z_{1_k}| + |(1-\lambda) z_{2_k}| = |\lambda| |z_{1_k}| + |(1-\lambda)| |z_{2_k}|$$

$$\overset{|z_{1_k}|<1 \quad or \quad |z_{2_k}|<1}{<} |\lambda| + |(1-\lambda)| \overset{\substack{\lambda \in (0,1) \\ 1-\lambda \in (0,1)}}{=} \lambda + 1 - \lambda = 1$$

$$|\lambda z_{1_k} + (1-\lambda) z_{2_k}| < 1$$

and that is a contradiction to the fact that:

$$\lambda z_{1_k} + (1-\lambda) z_{2_k} = \pm 1$$

thus $\{-1,1\}^n$ are extreme points of $B_\infty[0_n,1]$.

this doesn't finish the proof though, as we didn't prove that their aren't any other extreme points.

let $x \in B_\infty[0_n,1]$, $x \notin \{-1,1\}^n$ and let's show it is not an extreme point of $B_\infty[0_n,1]$.

since $x \notin \{-1,1\}^n$ there exist at least one coordinate of x which is not 1 or -1 , let's denote it x_k, and since $x \in B_\infty[0_n,1]$ this coordinate value must be between them, meaning:

$$|x_k| < 1$$

define two new vectors $z_1, z_2 \in \mathbb{R}^n$ such that:

$$\forall j \in \{1,2,...,n\} \setminus \{k\} : z_{1_j} = z_{2_j} = x_j$$

$$z_{1_k} = 1, \quad z_{2_k} = -1$$

first of all since $x \in B_\infty[0_n,1]$ than $\forall j \in \{1,2,...,n\} \setminus \{k\} : \left|z_{1_j}\right| = \left|z_{2_j}\right| = |x_j| \le 1$

in addition:

$$|z_{1_k}| = |z_{2_k}| = 1 \le 1$$

hence $z_1 \ne z_2 \in B_\infty[0_n,1]$.

notice that for $\lambda = \frac{x_k+1}{2}$:

$$\forall j \in \{1,2,...,n\} \setminus \{k\} : \lambda z_{1_j} + (1-\lambda) z_{2_j} = \lambda x_j + (1-\lambda) x_j = x_j$$

$$\lambda z_{1_k} + (1-\lambda) z_{2_k} = \lambda - 1 + \lambda = 2\lambda - 1 = 2 \cdot \frac{x_k+1}{2} - 1 = x_k + 1 - 1 = x_k$$

thus:

$$\lambda z_1 + (1-\lambda) z_2 = x$$

and:

$$\lambda = \frac{x_k+1}{2} \overset{|x_k|<1}{<} \frac{1+1}{2} = \frac{2}{2} = 1$$

$$\lambda = \frac{x_k+1}{2} \overset{|x_k|<1}{>} \frac{-1+1}{2} = \frac{0}{2} = 0$$

we have found two vectors $z_1 \ne z_2 \in B_\infty[0_n,1]$ and a scalar $\lambda \in (0,1)$ such that the vector $x \in B_\infty[0_n,1]$ can be written as:

$$x = \lambda z_1 + (1-\lambda) z_2$$

thus x is not an extreme point of $B_\infty[0_n,1]$.

b)

Let $X_i \subseteq \mathbb{R}^n, i = 1, 2, ..., k$.

Prove that:

$$ext\,(X_1 \times X_2 \times ... \times X_k) = ext\,(X_1) \times ext\,(X_2) \times ... \times ext\,(X_k)$$

let $y \in ext\,(X_1) \times ext\,(X_2) \times ... \times ext\,(X_k)$

first notice that:

$$\forall j : ext\,(X_j) \subseteq X_j$$

thus:

$$ext\,(X_1) \times ext\,(X_2) \times ... \times ext\,(X_k) \subseteq X_1 \times X_2 \times ... \times X_k$$

therefore $y \in X_1 \times X_2 \times ... \times X_k$

let's assume by contradiction that y is not an extreme point of $X_1 \times X_2 \times ... \times X_k$, thus:

$$\exists z_1 \neq z_2 \in X_1 \times X_2 \times ... \times X_k, \lambda \in (0, 1) : y = \lambda z_1 + (1 - \lambda)\, z_2$$

denote $v^j \in \mathbb{R}^n$ as a sub vector consisting of coordinates $(j-1)\,n+1$ to jn of vector v, for example v^1 is the vector consisting of the first n coordinates of v.

because $z_1 \neq z_2$:

$$\exists j \in \{1, 2, .., k\} : z_1^j \neq z_2^j$$

let's look at the vector y^j , from the equation above:

$$y^j = \lambda z_1^j + (1 - \lambda)\, z_2^j$$

since $z_1, z_2 \in X_1 \times X_2 \times ... \times X_k$ than $z_1^j, z_2^j \in X_j$

since $y \in ext\,(X_1) \times ext\,(X_2) \times ... \times ext\,(X_k)$ than $y^j \in ext\,(X_j) \subseteq X_j$

we have found two vectors $z_1^j \neq z_2^j \in X_j$ and $\lambda \in (0, 1)$ such that $y^j \in X_j$ can be written as:

$$y^j = \lambda z_1^j + (1 - \lambda)\, z_2^j$$

this is a contradiction to the fact that $y^j \in ext\,(X_j)$

thus $ext\,(X_1) \times ext\,(X_2) \times ... \times ext\,(X_k)$ are extreme points of $X_1 \times X_2 \times ... \times X_k$.

now we need to prove that their aren't any other extreme points.

let $y \in X_1 \times X_2 \times ... \times X_k$,$y \notin ext\left(X_1\right) \times ext\left(X_2\right) \times ... \times ext\left(X_k\right)$ and let's prove it's not an extreme point.

because $y \notin ext\left(X_1\right) \times ext\left(X_2\right) \times ... \times ext\left(X_k\right)$:

$$\exists j \in \{1, 2, ..., k\} : y^j \notin ext\left(X_j\right)$$

of course $y^j \in X_j$ because $y \in X_1 \times X_2 \times ... \times X_k$ thus:

$$\exists z_1 \neq z_2 \in X_j, \lambda \in (0, 1) : y^j = \lambda z_1 + (1 - \lambda) z_2$$

define two new vectors $v_1, v_2 \in \mathbb{R}^{n^2}$ in the below manner:

$$\forall i \in \{1, 2, ..., k\} \setminus \{j\} : v_1^i = v_2^i = y^i$$

$$v_1^j = z_1 \quad v_2^j = z_2$$

notice that because $y \in X_1 \times X_2 \times ... \times X_k$:

$$\forall i \in \{1, 2, ..., k\} \setminus \{j\} : v_1^i = v_2^i = y^i \in X_i$$

and:

$$v_1^j = z_1 \in X_j \quad v_2^j = z_2 \in X_j$$

thus $v_1, v_2 \in X_1 \times X_2 \times ... \times X_k$ and $v_1 \neq v_2$ because $v_1^j \neq v_2^j$

also notice that:

$$\forall i \in \{1, 2, ..., k\} \setminus \{j\} : \lambda v_1^i + (1 - \lambda) v_2^i = \lambda y^i + (1 - \lambda) y^i = y_i$$

$$\lambda v_1^j + (1 - \lambda) v_2^j = \lambda z_1 + (1 - \lambda) z_2 = y^j$$

hence $\lambda v_1 + (1 - \lambda) v_2 = y$

we have found two vectors $v_1 \neq v_2 \in X_1 \times X_2 \times ... \times X_k$ and a scalar $\lambda \in (0, 1)$ such that $y \in X_1 \times X_2 \times ... \times X_k$ can be written as:

$$y = \lambda v_1 + (1 - \lambda) v_2$$

thus y is not an extreme point of $X_1 \times X_2 \times ... \times X_k$.