Problem 1

Find the optimal solution $t \in \mathbb{R}$ for the following optimization problems. Your solution must be in terms of given matrices and as concise as possible. Assume that the given matrices are of appropriate size including the identity matrix I.

a) Given Symmetric X, find

$$\min\{t \mid tI - X \succeq 0\}.$$

The expression $tI - X \succeq 0$ means that the matrix $tI - X$ is PSD. Therefore all eigenvalues must be non-negative. So the eigenvalues of X must satisfy $t \geq \lambda_i, \forall i$ where λ_i are the eigenvalues of X. Therefore, the optimal t is the maximum of the eigenvalues of X denoted as

$$t = \max\{\lambda_i\}.$$

b) Given Y, find

$$\min\{t \mid tI - Y^T Y \succeq 0\}.$$

The expression $tI - Y^T Y \succeq 0$ means that the matrix is PSD and we know $Y^T Y$ is symmetric.

Let $X = Y^T Y$ then,

$$\min\{t \mid tI - X \succeq 0\}$$

which is equivalent to the previous problem and hence the optimal t is

$$t = \max\{\lambda_i\}.$$

c) Given symmetric W and symmetric positive definite Z, find

$$\min\{t \mid tZ - W \succeq 0\}.$$

If Z is symmetric, then it is diagonizable, its eigenvalues are real, and its eigenvectors are orthogonal. Hence, Z has an eigen decomposition $Z = Q\Lambda Q^T$. If Z is also positive definite then all its eigenvalues are non-negative. So we have,

$$Z = Q\Lambda Q^T, \text{ where } \Lambda = \mathrm{diag}(\lambda_i) \text{ for}$$
non-negative eigenvalues and $Q^TQ = QQ^T = I$.

This decomposition gives us similar form to part (a) and with a symmetric W the optimal t is $\max\{\lambda_i\}$.

Problem 2

For each of the following sets, either prove it is convex, or give a counter example that shows it is not convex.

a) $S = \{(x,t) \in \mathbb{R}^{n+1} \mid 1 \leq \|x\| \leq (1+t), t > 0\}$.

Let $\alpha = 1$, $\beta = (1+t)$ and we know $\beta \geq \alpha$ for $t > 0$.

Then,

$$H_1 = \{x \in \mathbb{R}^n \mid \|x\| \geq \alpha\}$$

and

$$H_2 = \{x \in \mathbb{R}^n \mid \|x\| \leq \beta\}$$

and our set $S = H_1 \cap H_2$ where H_1 and H_2 are halfspaces. Hence our set S is convex.

b) $\{x \mid x + S_2 \subseteq S_1\}$ where $S_1, S_2 \subseteq \mathbb{R}^n$, and S_1 is a convex set.

The set is convex.

proof

we have $C = x + S_2 \subseteq S_1$ where S_1 is convex.

We can write $x + y \in S_1$ for all $y \in S_2$.

so then for all $y \in S_2$ we have

$$\{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\}.$$

For set C, all of the points x such that for every y in S_2, $x+y$ is in S_1.

This allows $\{x \mid x+y \in S_1\} = \bigcap_{y \in S_2} (S_1 - y)$

which is the intersection of convex sets $S_1 - y$
hence, C is convex. $\boxtimes$

c) $\{a = [a_0, a_1, \dots, a_m]^T \in \mathbb{R}^{m+1} \mid \max_{i=1,\dots,k} |p(x_i) - y_i| \leq \varepsilon\}$,
where,

$$p(x) = a_0 x + a_1 x + a_2 x^2 + \dots + a_m x^m$$

is a polynomial in $x \in \mathbb{R}$, and $x_i, y_i \in \mathbb{R}$ are given.

The set is convex

Def. of convex sets: A set is convex if, for any two points a, b in the set, the line segment connecting a and b is also in the set. Let a, b be two arbitrary points such that $\max |p(x_i^a) - y_i| \leq \varepsilon$ and $\max |p(x_i^b) - y_i| \leq \varepsilon$

where x_i^a, x_i^b are given fixed data points.

Consider a convex combination a and b, where $c = \theta a + (1-\theta) b$ for $\theta \in [0,1]$. We want to show

$$\max_{i=1\ldots k} |p(x_i^c) - y_i| \leq \varepsilon .$$

Since $p(x)$ is a polynomial it is linear in the coefficients a_i. Additionally, we know that the $\max(x_1, \ldots x_n)$ is convex

$$\text{for } 0 \leq \theta \leq 1 : f(\theta x + (1-\theta) y) = \max_i (\theta x_i + (1-\theta) y_i)$$

$$\leq \max_i x_i + (1-\theta) \max_i y_i$$

$$= \theta f(x) + (1-\theta) f(y).$$

Therefore $\max_i |p(x_i^c) - y_i|$ is a linear combination of $\max_i |p(x_i^a) - y_i|$ and $\max_i |p(x_i^b - y_i|$ with coefficients θ and $1-\theta$. Hence the maximum value of the linear combination in bound ε is convex. #

d) $\{x \in \mathbb{R}^n \mid \text{dist}(x,a) \leq \text{dist}(x,B)\}$, where $a \in \mathbb{R}^n$, B is a set (not necessarily convex), and for any $x \in \mathbb{R}^n$ and $C \subset \mathbb{R}^n$ we define the distance of a point to a set as: $\text{dist}(x, C) = \inf_{y \in C} \|x - y\|_2$.

This set in general is not convex. The set means the point x is closer to a vector $a \in \mathbb{R}^n$ than a set B, which is not necessarily convex.

If we take B to be non-convex we cannot prove that the infimum of the Euclidean distance over B or for some $y \in B$ will be convex.

e) $S = \{x \in \mathbb{R}^n \mid Tr(B^T A(x)) \geq 1\}$, where $A(x) = A_0 + x_1 A_1 + \ldots + x_n A_n$ and $A_i \in S^n$, $B \in \mathbb{R}^{n \times n}$ are given.

We know Trace is a linear operation, $A(x)$ is the linear combination, A is symmetric, and B is a square matrix.

We want to show that a line segment $x = \theta x_1 + (1-\theta) x_2$ for $0 \leq \theta \leq 1$ is in our set.

Consider: $Tr(B^T A(x)) = Tr(B^T A(\theta x_1 + (1-\theta) x_2))$

Using the linearity of the trace operator:

$$= Tr(\theta B^T A(x_1) + (1-\theta) B^T A(x_2))$$

Now since $Tr(X+Y) = Tr(X) + Tr(Y)$ we can write

$$= \theta Tr(B^T A(X_1)) + (1-\theta) Tr(B^T A(X_2))$$

Since $X_1, X_2 \in S$ we get $Tr(B^T A(X_1)) \geq 1$ and $Tr(B^T A(X_2)) \geq 1$ therefore it is greater or equal to

$$\theta + (1-\theta) = 1.$$

Hence for any $X_1, X_2 \in S$ the line segment $\theta X_1 + (1-\theta) X_2$ is also in S. Hence, S is convex. $\not{\!\#}$

f) $S = \{(x,y) \mid \|x\|_2 \leq \|y\|_2 , x, y \in \mathbb{R}^n\}$.

This set is convex.

proof

Since a norm is always non-negative, we have

$\|x\|_2 \leq \|y\|_2$ iff $^2\|x\|_2 \leq {}^2\|y\|_2$ so

$$\|x\|_2^2 \leq \|y\|_2^2 \Leftrightarrow x_1^2 + x_2^2 + \ldots + x_n^2 \leq y_1^2 + y_2^2 + \ldots + y_n^2$$

Now consider each seperate inequality $y_i^2 - x_i^2$ which can be written as a halfsace $\{(x,y) \mid y_i^2 - x_i^2 \geq\}$ thus we have an intersection of halfspaces

$$\bigcap_{i=1}^{n} \{(x,y) \mid y_i^2 - x_i^2 \geq 0\},$$ hence the set is convex. $\not{\!\#}$

Problem 3

Let x be a real-valued random variable with $\mathrm{Prob}(x=a_i) = p_i$, $i=1,\dots,n$, where $a_1 < a_2 < \dots < a_n$. Note that $p \in \mathbb{R}^n$ lies in the standard probability simplex $P = \{p \mid \mathbf{1}^T p = 1, p \succeq 0\}$. For each case below, either show that given condition defines a convex set (in p), or provide a counter example.

a) $\mathrm{Prob}(x \in C) \leq 0.3$ where $C \subset \mathbb{R}$ is given but not necessarily convex.

$$\mathrm{Prob}(x=a_i) = p_i, i=1,\dots,n = \sum_{i=1}^{n} p_i$$

$$\Rightarrow \mathrm{Prob}(x \in C) = \sum_{i : a_i \in C}^{n} p_i \leq 0.3$$

This constraint is linear in p and therefore the constraint is equivalent to a linear inequality and hence convex.

b) $\mathrm{Prob}(x \in A \mid x \in B) \leq 0.6$, where $A, B \subset \mathbb{R}$ are given but not necessarily convex.

$$\mathrm{Prob}(x \in A \mid x \in B) = \frac{\mathrm{Prob}(x \in A \cap B)}{\mathrm{Prob}(x \in B} \leftarrow \text{conditional probability}$$

$$\Rightarrow \quad \frac{\sum_{i:a_i \in A \cap B} p_i}{\sum_{i:a \in B} p_i} \le 0.6$$

Although they may be linear in p, the quotient is not necessarily convex and thus we cannot conclude that this set of probabilities is convex in p.

c) $\mathbb{E} f(x) \geq 0.2$ where $\mathbb{E} f(x) = \sum_{i=1}^{n} p_i f(a_i)$.

we can write $\sum_{i=1}^{n} p_i f(a_i)$ → not a function of p

↳ linear in p_i

$\mathbb{1}^T p = 1$

↳ hyperplane

So, $\sum_{i=1}^{n} p_i f(a_i) \geq 0.2$ is a halfspace and hence is convex.

d) $\alpha \mathbb{E} x \leq \text{var}(x)$, where α is a given constant and $\text{var}(x) = \mathbb{E}(x - \mathbb{E}x)^2 = \mathbb{E}x^2 - (\mathbb{E}x)^2$

$$= \sum_{i=1}^{n} a_i^2 p_i + \left(\sum_{i=1}^{n} a_i p_i\right)^2$$

And this can be written as

$$\sum_{i=1}^{n} a_i^2 p_i + \left(\sum_{i=1}^{n} a_i p_i\right)^2 = b^T p + p^T A p$$

where $b_i = a_i^2$ and $A = aa^T$. Because aa^T is positive semidefinite this is a convex set.

Lets then write: $\alpha \mathbb{E}x \leq b^T p + p^T A p$ as

$$\alpha \cdot \sum_{i=1}^{n} p_i f(a_i) \leq b^T p + p^T A p$$

$$\Rightarrow \underbrace{\alpha \cdot \sum_{i=1}^{n} p_i f(a_i)}_{\text{linear in } p} - \underbrace{b^T p + p^T A p}_{\text{convex}} \leq 0$$

hence this is a convex set.

Problem 4

Consider the convex cone $K_p = \{(x,t) \mid \|x\|_p \leq t\}$. Is this a proper cone for $p=2$? Is it a proper cone for $p=1$?

A cone $K \subseteq \mathbb{R}^n$ is a proper cone if it satisfies:

1. K is convex
2. K is closed
3. K is solid (non-empty interior)
4. K is pointed

For $p=2$ we have the second-order cone,

$$K_2 = \{(x,t) \mid \|x\|_2 \leq t\}.$$

1) K_2 is convex by definition of the norm cone and given in the problem statement.

2) We can write $\|x\|_2 - t \leq 0$ with a non-strict inequality and therefore K_2 includes its boundary and is closed.

3) By definition the norm cone and thus second-order cone has non-empty interior.

Meaning for any $(x,t) \in K_2$ with $t>0$, we can find an open ball around (x,t) that is entirely contained in K_2.

4) K is pointed: if $x \in K, -x \in K \Rightarrow x = 0$.

consider $(x_1, t_1), (x_2, t_2) \in K_2$ with $t_1, t_2 > 0$.
Show that the line segment connecting these points is in K_2. Let $\theta \in [0,1]$, $\theta \in \mathbb{R}$.

Then the line segment is given by
$$(\theta x_1 + (1-\theta) x_2,\ \theta t_1 + (1-\theta) t_2)$$

we want to show the norm condition holds
$$\|(\theta x_1 + (1-\theta) x_2\|_2 \leq \theta t_1 + (1-\theta) t_2$$

By the triangle inequality:
$$\|(\theta x_1 + (1-\theta) x_2)\|_2 \leq \theta \|x_1\|_2 + (1-\theta)\|x_2\|_2$$

Since $(x_1, t_1), (x_2, t_2) \in K_2$ we have

$\|x_1\|_2 \leq t_1$ and $\|x_2\|_2 \leq t_2$.

Therefore, $\|(\theta x_1 + (1-\theta)x_2)\|_2 \leq \theta t_1 + (1-\theta) t_2$.

This shows the line segment satisfies the norm condition for K_2. So we can say K_2 is pointed.

For $p=1$: $K_1 = \{(x,t) \mid \|x\|_1 \leq t\}$.

1) K_1 is convex given in problem statement and in def. of norm cone

2) We can write $\|x\|_1 - t \leq 0$ with a non-strict inequality and therefore K_1 is closed.

3) By definition K_1 has non-empty interior meaning we can find find an open ball for any $(x,t) \in K_1$ with $t>0$ that is entirely contained in K_1.

4. similar proof to that of l_2-norm.

OR, Take (x,t), $(-x,-t) \in K_1$.

Then, $\|x\|_1 \leq t$ and $\|-x\|_1 \leq -t$.

we know, $\|x\|_1 = |x_1| + \dots + |x_n|$, so
$\|x\|_1 = \|-x\|_1$
$\Rightarrow \|x_1\| \leq t$ and $\|x\|_1 \leq -t$

This implies $t = 0$ and so $\|x\|_1 = 0$
and $x = 0$.
This satisfies the def. of pointedness.

* you can use this proof for K_2 and generalize to any norm cone.