ECE569/AI539 Convex Optimization - Mid-term Exam

Fall 2022
School of Electrical Engineering and Computer Science
Oregon State Univeristy

Due: 11:59 PM, Oct. 31, 2022

Note: please submit through Canvas. Please also sign the attached statement in the last page.

- This cover sheet must be signed and submitted along with the exam answers.
- By submitting the exam answers with my name affixed above,
  - I understand that exam answers submitted at 11:59 PM Oct. 31, 2022, or later will not be accepted,
  - I acknowledge that I am aware of the Oregon State University policy concerning academic misconduct (see `https://studentlife.oregonstate.edu/studentconduct`),
  - I attest that the work I am submitting for this exam is solely my own,
  - I understand that suspiciously similar answers submitted by multiple individuals will be reported to the School of EECS and College of Engineering for investigation, and
  - I understand that the exam materials are **confidential** and will not disclose the materials to any third party.

1a) $S = \{x \in \mathbb{R}^n \mid \max_{i \in \{1 \dots n\}} x_i \leq \alpha\}$

$\alpha \in \mathbb{R}$.

If x_1 and x_2 are in set S.

say $\max x_{1i} \leq \alpha$ & $\max x_{2i} \leq \alpha$

for convexity check $\theta x_1 + (1-\theta) x_2 = x_3$

applying the function max

$\theta x_1 + (1-\theta_2) x_2 = x_3$

applying max on LHS.

$\max(\theta x_1 + (1-\theta) x_2) = \max(\theta x_1) + \max((1-\theta) x_2)$

$\Rightarrow \theta \max x_{1i} + (1-\theta) \max x_{2i}$

$\Rightarrow \theta \alpha + (1-\theta) \alpha = \cancel{\theta\alpha} + 1\alpha - \cancel{\theta\alpha} = \alpha$

$\Rightarrow \alpha = x_3$ which satisfies the condition $x_i \leq \alpha$

$\Rightarrow x_3$ is in set S

$\Rightarrow S$ is a convex set

1b) $S = \left\{x \in \mathbb{R}^n \;\middle|\; \frac{a^T x + b}{c^T x + d} \leq 1,\ c^T x + d > 0\right\}$

$\frac{a^T x + b}{c^T x + d} \leq 1$

$\Rightarrow a^T x + b \leq c^T x + d$

$\Rightarrow a^T x - c^T x \leq d - b$

which is of the form $a^T x <= b$

$\Rightarrow (a^T - c^T) x < (d-b)$ which is of the form of halfspace

Halfspaces are convex.

1c) $f(x) = \|Ax-b\|_2^2 - \gamma \|x\|_2^2$

where $\gamma > 0$

$\|Ax-b\|_2^2 = (Ax-b)^T(Ax-b)$

$\|x\|_2^2 = x^T x$

$\Rightarrow f(x) = (Ax-b)^T(Ax-b) - \gamma(x^T x)$

$x^T A^T A x - x^T A^T b - b^T A x + b^T b - \gamma x^T x$

$\Rightarrow x^T(A^T A - \gamma I)x - x^T A^T b - b^T A x + b^T b$

$\frac{\partial}{\partial x} f(x) = \frac{\partial}{\partial x}\left(\|Ax-b\|_2^2 - \gamma \|x\|_2^2\right)$

$= 2(Ax-b).A - \gamma 2x$

$\frac{\partial^2 f(x)}{\partial x^2} = 2(A) + 2(Ax-b)(0) - \gamma 2$

$\Rightarrow \frac{\partial^2 f(x)}{\partial x^2} = 2A - \gamma.2 = 2(A-\gamma)$

It's convex

1d) $f(x) = \max_P \{\|APx - b\|_4 \mid P \text{ is Permutation matrix}\}$

Suppose maximum is P_{max}

$f(x) = \{\|AP_{max}x - b\|_4\}$

$AP_{max}x - b$ is an affine function

$\Rightarrow$ It's convex as in $\|AP_{max} - b\|_4$

~~Scaling and adding an affine~~

is affine function of x

4 norm is convex.

1e) ~~X~~ Set $S = \{X \in S^n ; rank(X) \leq 1\}$

Let X_1 and X_2 belong to S

$rank(X_1) \leq 1$ $\quad$ $rank(X_2) \leq 1$

$X_3 = \theta X_1 + (1-\theta) X_2$

$\Rightarrow rank(\theta X_1) + rank((1-\theta)X_2)$

$= \theta\, rank(X_1) + (1-\theta)\, rank(X_2)$

$\Rightarrow rank(X_1 + X_2) \leq rank(X_1) + rank(X_2)$

$\Rightarrow rank(X_1) + rank(X_2) - \theta\, rank(X_2)$

$\Rightarrow rank(X_1) + rank(X_2) \leq 1 + 1 = 2$

$rank(X_1) + rank(X_2) \leq 2.$

for convex ~~convex~~ ~~vector~~ rank must ≤ 1 $\forall X \in S^n$

$\therefore$ not convex

2a) given

$X = WH^T$

$X \in R^{m \times n}$

$W \in R_+^{m \times r}$

$H \in R_+^{n \times r}$

Note:
Conic Hull in a set C is set of all conic combinations of points in C
$\sum_{i=1}^{k} \theta_i x_i$, where $\theta_i \geq 0 \ \forall i = 1 \ldots k$ → conic combination of $x_1 \ldots x_k$ — (1)

S.T. $X(:,l) \in \text{cone}(W) \ \forall l = 1, \ldots, n$

$X(:,l) \rightarrow l^{th}$ column of X

$$W = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1r} \\ W_{21} & & & W_{2r} \\ W_{31} & & & \vdots \\ W_{m1} & \cdots & \cdots & W_{mr} \end{bmatrix}_{m \times r} \qquad H = \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1r} \\ \vdots & & & \\ H_{n1} & \cdots & \cdots & H_{nr} \end{bmatrix}_{n \times r}$$

$$WH^T = \begin{bmatrix} W_{11} & \cdots & W_{1r} \\ W_{21} & & \vdots \\ \vdots & & \\ W_{m1} & & W_{mr} \end{bmatrix}_{m \times r} \begin{bmatrix} H_{11} & H_{21} & \cdots & H_{n1} \\ H_{12} & & & \\ & & & \\ H_{1r} & H_{2r} & & H_{nr} \end{bmatrix}_{r \times n}$$

$$= \begin{bmatrix} W_{11}H_{11} + \cdots + W_{1r}H_{1r} & W_{11}H_{21} + \cdots W_{1r}H_{2r} & \cdots \\ \vdots & & \vdots \\ W_{m1}H_{m1} \cdots + W_{mr}H_{mr} & \cdots & W_{m1}H_{n1} \cdots + W_{mr}H_{nr} \end{bmatrix}_{m \times n}$$ which has n columns and m rows

~~$X(:,l) = [\sum W_{ij} H]$~~ Say for an l^{th} column in X where $l \in \{1, \ldots n\}$

$$X(:,l) = \begin{bmatrix} W_{11}H_{l1} + \cdots W_{1r}H_{lr} \\ W_{21}H_{l1} + \cdots \\ W_{31}H_{l1} + \cdots \\ W_{m1}H_{l1} + \cdots W_{mr}H_{lr} \end{bmatrix}_{m \times 1}$$

$\Rightarrow X(:,l) = \left[\sum_{i=1}^{r} H_{li} W_{ji} \right]_{j \times 1} (j = 1 \ldots m)$; $H_{li} > 0$ from question $H \in R_+^{n \times r} \Rightarrow H^T \in R_+^{r \times n}$

$= \left[\sum_{i=1}^{r} W_i H_{il} \right]$

$\Rightarrow X(:,l)$ is a conic combination of $W_{ji} \ \forall j \in \{1, \ldots m\}$, $l \in \{1, \ldots n\}$

$\Rightarrow X(:,l) \in \text{cone}(W) \ \forall l \in \{1 \ldots n\}$
from (1)

2b) $\tilde{X}$ normalized model

$$\tilde{X}(:,l) = \begin{cases} \dfrac{X(:,l)}{\|X(:,l)\|_1}, & \|X(:,l)\|_1 \neq 0 \\ 0 \text{ (all zero vec with len } m\text{)}, & \text{otherwise} \end{cases}$$

By expanding $\tilde{X}(:l)$

$$\tilde{X}(:l) = \frac{\sum_{i=1}^{r} W_i H_{il}}{\|\sum_{i=1}^{r} W_i H_{il}\|} \text{ as } W \in R_+^{m \times r} \text{ and } H \in R_+^{n \times r}$$

$\Rightarrow$ $\|\sum_{i=1}^{r} W_i H_{il}\|_1$ checks if all elements in the vector are non zero

$$\tilde{X}(:,l) = \frac{W_1 H_{1l}}{\|W_1\|_1 H_{1l} + \cdots + \|W_r\|_1 H_{rl}} + \cdots + \frac{W_r H_{rl}}{\|W_1\|_1 H_{1l} + \cdots + \|W_r\|_1 H_{rl}}$$

multiplying and dividing by $\|W_i\|_1$ for each i^{th} term where $i \in \{1..r\}$

$$= \left(\frac{W_1 H_{1l}}{\|W_1\|_1 H_{1l} + \cdots + \|W_r\|_1 H_{rl}} \times \frac{\|W_1\|_1}{\|W_1\|_1}\right) + \cdots +$$

$$= \left(\frac{W_1}{\|W_1\|_1} \cdot \frac{H_{1l}\|W_1\|_1}{\sum_{i=1}^{r}\|W_i\|_1 H_{il}}\right) + \cdots + \left(\frac{W_r}{\|W_r\|_1} \cdot \frac{H_{rl}\|W_r\|_1}{\sum_{i=1}^{r}\|W_i\|_1 H_{il}}\right)$$

which is a convex combination of $\frac{W_i}{\|W_i\|_1}$, $i \in 1,..r$. — ②

which is a convex hull.

$$\left\{\frac{\|W_1\|_1 H_{1l}}{\sum_{i=1}^{r}\|W_i\|H_{il}}, \ldots, \frac{\|W_r\|_1 H_{rl}}{\sum_{i=1}^{r}\|W_i\|H_{rl}}\right\} \in [0,1] \text{ — ③}$$

since ② & ③ are achieved

$$\tilde{X}(:,l) \in \text{conv}\left(\frac{W_i}{\|W_i\|_1}\right) \quad i \in \{1,..r\}$$

2c) $K = \{l_1 \ldots l_r\} \subseteq \{1 \ldots n\}$

ST

$$H(l_i:)^T = \alpha_i e_i \quad \forall i \in \{1, \ldots, r\}$$

where $\alpha_i > 0 \ \forall i = 1 \ldots r$ & $e_i \in \mathbb{R}^r$ is i^{th} unit

$$\hat{l} = \arg\max_{l \in \{1 \ldots n\}} \|\tilde{X}(:l)\|_2$$

ST.

$$X(:\hat{l}) \in \{W(:,1)\alpha_1 \ldots W(:r)\alpha_r\}$$

Hint ST $\hat{l} \in K$.

By expanding $X(:\hat{l})$

$$X(:\hat{l}) = \max \|\tilde{X}(:l)\|_2 \quad l \in \{1, \ldots, n\}$$

$$\Rightarrow X(:\hat{l}) = \max \left\| \frac{\sum_{i=1}^{r} w_i H_{il}}{\|\sum_{i=1}^{r} w_i H_{il}\|_1} \right\|_2 \quad l \in \{1 \ldots n\} \quad -① $$

If the convex hull has max points in $\{x_1 \ldots x_q\}$

$$x_{max} = \max_x \|x\|_2 \quad \forall x \in \{x_1 \ldots x_q\}$$

If we take the comb^n of pts that belong to max class, we have

$$= c_1 x_1 + c_2 x_2 + c_3 x_3 + \ldots + c_q x_q \quad -①$$

where $c_i \in [0,1] \ \forall i \in \{1 \cdots q\}$

$$\sum c_i = 1$$

Using the triangle inequality on $\|①\|_2$

$$\|c_1 x_1 + c_2 x_2 + \cdots c_q x_q\|_2 \le c_1\|x_1\|_2 + c_2\|x_2\|_2 + \cdots$$

$$\le c_1\|x_{max}\|_2 + \cdots c_q\|x_{max}\|_2$$

H^T(:,l) for all l_i is a linear combination and it is sparse, will have only one term.

3a) $a \in \mathbb{R}^n$ $a_1 \geq a_2 \geq a_3 \geq \cdots \geq a_n > 0$

$b \in \mathbb{R}^n$ $b_k = 1/a_k$.

minimize $-\log(a^T x) - \log(b^T x)$

subject to $x \succeq 0$, $1^T x = 1$

(a) $f(x) = -\log(a^T x) - \log(b^T x)$

$\Rightarrow -[\log(a^T x) + \log(b^T x)]$

$= -[\log(a^T x \cdot b^T x)]$

Which can be written as

$-\log(x^T a b^T x)$

we know that $x \succeq 0$ is convex

$1^T x = 1$ is affine.

$ab^T \rightarrow$ is a number of $1 \times 1 \Rightarrow$ PSD

and determinant is

number itself.

$\Rightarrow$ $-\log$ is convex function $\therefore$ the problem is

a convex problem.

3b) KKT of minimize $-\log(a^T x) - \log(b^T x)$

subject to $x \succeq 0$

$1^T x = 1$

$1^T x^* = 1$, $x^* \succeq 0$

$1^T x^* - 1 = 0$, $-x^* \preceq 0$

$\lambda^* \succeq 0$

$\lambda^* x_i^* = 0$

$f_i(x^*) \leq 0$, $i = 1 \dots m$

$h_i(x^*) = 0$, $i = 1 \dots p$

$\lambda_i^* \geq 0$, $i = 1 \dots m$

$\lambda_i^* f_i(x^*) = 0$, $i = 1 \dots m$

$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} v_i^* \nabla h_i(x^*)$

~~$L(x^*, \lambda^*, v^*)$~~

~~$= -\log(a^T x^*) - \log$~~

$L(x^*, \lambda^*, v^*) = -\log(a^T x^*) - \log(b^T x^*) - \lambda x^* + v^* (1^T x^* - 1)$

Taking the gradient

$= -\frac{1}{a^T x^*} \cdot a_i - \frac{1}{b^T x^*} \cdot b_i - \lambda_i^* + v^* 1^T = 0.$

$\Rightarrow \quad \frac{-a_i}{a^T x^*} - \frac{b_i}{b^T x^*} - \lambda_i^* + v^* 1^T = 0$

3c) Verify that $x^* = [1/2, 0 \ldots 0, 1/2]^T$ is an optimal solution

from 3b problem we know that

$$\frac{-a_i}{a^T x^*} - \frac{b_i}{b^T x^*} - \lambda_i^* + v^* = 0$$

from values given in x^*; $k = 1, n$.

$$\lambda_i^* = \frac{-a_i}{a^T x^*} - \frac{b_i}{b^T x^*} + v^*$$

$i = 1, n \rightarrow$ writing a generalised form

$$\lambda_i^* = \frac{-a_i}{a^T x^*} - \frac{b_i}{b^T x^*} + v^*$$

$$= \frac{-a_i}{\frac{a_1 + a_n}{2}} - \frac{b_i}{\frac{b_1 + b_n}{2}} + v^*$$

$$= \frac{-a_i \cdot 2}{a_1 + a_n} - \frac{b_i \cdot 2}{b_1 + b_n} + v^*$$

$$= \frac{-2(a_i(b_1 + b_n) + b_i(a_1 + a_n))}{(a_1 + a_n)(b_1 + b_n)} + v^*$$

$$= \frac{-2\left(a_i \cdot 1 + b_i\left(\frac{a_1 + a_n}{b_1 + b}\right)\right)}{(a_1 + a_n)} + v^*$$

$$= \frac{-2\left(a_i \cdot 1 + b_i\left(\frac{a_1 + a_n}{\frac{a_1 + a_n}{a_1 a_n}}\right)\right)}{(a_1 + a_n)} + v^* ; \quad b_1 = 1/a_1, \; b_n = 1/a_n$$

$$\lambda_i^* = \frac{-2(a_i + b_i a_1 a_n)}{(a_1 + a_n)} + v^*$$

we know that $\lambda_i^* x_i^* = 0 \Rightarrow \lambda_1^* x_1^* = 0 \Rightarrow \lambda_1^* = 0$ as $x_1^* = 1/2$

||rly $\lambda_n^* = 0$ as $x_n^* = 1/2$

by substituting we get

$$0 = \frac{-2(a_1 + 1/a_1 \, a_1 a_n)}{(a_1 + a_n)} + v^* \Rightarrow \underline{\underline{v^* = 2}}$$

$$0 = \frac{-2(a_n + 1/a_n \, a_1 a_n)}{(a_1 + a_n)} + v^* \Rightarrow \underline{v^* = 2}$$

x^* is optimal as $x^* = [1/2, 0 \ldots 0, 1/2]$; $\lambda^* = [0, NA \ldots NA, 0]$

$\boxed{v^* = 2} \Rightarrow x^*, \lambda^*, v^*$ satisfy KKT.

4a) minimize $Tr(CX)$

subject to $Tr(AX) = 1$ — ①

$X \succeq 0$

$rank(X) = 1$

minimize $Tr(CX)$

subject to $Tr(AX) = 1$ — ②

$X \succeq 0$

V^* and V^*_{SDR} are optimal values of ① & ②

From ② it can be observed that the rank $\neq 1$ as it is relaxed which implies that in ② there will only be a lower bound on the optimal objective value

Hence, $V^*_{SDR} \leq V^*$

$f_i(x^*) \preceq_{K_i} 0, \quad i=1\ldots m$

$h_i(x^*) = 0, \quad i=1\ldots p$

$\lambda_i^* \succeq_{K_i^*} 0, \quad i=1\ldots m$

$\lambda^{*T} f_i(x^*) = 0, \quad i=1\ldots m$

$\nabla f_0(x^*) + \sum_{i=1}^{m} Df_i(x^*)^T \lambda_i^* + \sum_{i=1}^{p} v_i^* \nabla h_i(x^*) = 0$

4b) from the question

minimize $Tr(CX)$

Subject to $Tr(AX) = 1$ — ②

$X \succeq 0$

$Tr(AX) = 1 \Rightarrow f(x) >= 0 \Rightarrow -x^* \preceq_{K_i} 0 \quad i=1$

$Tr(AX) - 1 = 0$

$\lambda_i^* \succeq_{K_i} 0 \quad i=1$

$\lambda_i^{*T} x^* = 0 \quad i=1$

Note: $\frac{\partial tr\, AB}{\partial a_{ij}} = b_{ji}$; $\nabla_A\, tr\, AB = B^T$

$\Rightarrow L(x^*, \lambda^*, v^*) = Tr(Cx^*) - \lambda^* x^*$
$+ v^*(Tr(Ax^*) - 1$

$\nabla_A\, tr(CA) = C^T$

1st term $= C^T$

2nd term after gradient $= \lambda^* I$

3rd term $= V^* A^T$

$\Rightarrow \nabla_x L(x^*, \lambda^*, v^*) = C^T - \lambda^* I + v^* A^T = 0$

4c)

$$X^*_{SDR} = \arg\min \ (2)$$

$$\text{rank}(X^*_{SDR}) \leq 1$$

from KKT

$$\lambda_i^{*T} f_i(x^*) = 0$$

$$\lambda^* x^* = 0$$

$$\Rightarrow \lambda^* = 0 \text{ as } x^* > 0$$

Taking $c^T - \lambda^* I + v^* A^T = 0$

2^{nd} term becomes 0 as $\lambda^* I = 0$

$$c^T + v^* A^T = 0$$

$$c^T = -v^* A^T \quad - (3)$$

rank of cx is full rank

rank(cx) is same as rank(x)

multiply x^* on both sides of (3)

$$c^T x^* = -v^* A^T x^*.$$

$$\text{rank}(c^T x^*) = \text{rank}(-v^* A^T x^*)$$

$$\text{rank}(x^*) = \text{rank}(-v^* A^T x^*)$$

$$A = \begin{bmatrix} c \\ \\ \\ \end{bmatrix}_{n\times 1} \begin{bmatrix} & c^T & \end{bmatrix}_{1\times n}$$ which will have a rank 1

$$\Rightarrow \text{rank}(-v^* A^T x^*) = \text{rank}(A^T x^*)$$

$\Rightarrow$ x^* needs to have a rank 1 at the most

$$\Rightarrow \text{rank}(x^*) \leq 1$$