VANDERBILT UNIVERSITY
CE 5999
INTRO TO OPTIMIZATION

HOMEWORK # 5
Junyi Ji
April 17, 2023

It takes about 24 hours to finish this assignment.

1. Solve Lab 1 (available here: `https://ecal.berkeley.edu/files/ce191/Lab1.pdf`) from my colleague's optimization class at UC Berkeley. I was gonna copy the problem herein, but I see no point in doing so. The problem seems too long, but it isn't once you formulate the linear program. Most of the other exercises are basic analysis and changes in parameters for the original problem formulation.

   Solutions:

   (a) The guess is super simple. See Table 1 below for more details. Based on the analysis of the demand and the supply, it is impossible to fulfill the requirements without Source 3, the quality of which is the lowest. I tried to balance Source 3 with other sources to meet the requirements for water quality. I find the mixture of Source 3 and 4 with the proportion of 3:7 can meet the requirements for the water quality.

Town	Appletown	Berrytown	Cherrytown	Grapetown	Mangotown	**Supply Sum**
Source 1	9	0	0	6	0	15
Source 2	8	0	0	2	0	10
Source 3	13	3	15	2	12	45
Source 4	0	7	35	10	28	80
Demand Sum	30	10	50	20	40	± 0
Quality	1125	1180	1180	675	1180	**Not Satisfied**

Table 1: Initial guess for the feasible solution

   However, the common pipe for Source 1 and 2 can only hold 20 ML water. Hence, the guess is not satisfied.

   (b) The problem can be formulated as follows.

      i. The notations of the variables are stated in Table 2.

Notations	Meaning
$f_{i,j}$	water flow from Source i to Town j
s_i	water supply limit for Source i
d_j	water demand of Town j
Q	water quality (hardness) requirement
$c_{i,j}$	cost for the water flow from Source i to Town j
q_i	water quality of Source i
m	number of the water sources
n	number of the towns

Table 2: Notations of the decision variables and constraints of the problem

ii. LP formulation.

$$\min \quad \sum_{i,j=1}^{m,n} c_{i,j} f_{i,j} \tag{1}$$

$$\text{subject to} \sum_{j=1}^{n} f_{i,j} \leq s_i \tag{2}$$

$$\sum_{i=1}^{m} f_{i,j} \geq d_j \tag{3}$$

$$\frac{\sum_{i=1}^{m} q_i f_{i,j}}{\sum_{i=1}^{m} f_{i,j}} \leq Q \tag{4}$$

$$\sum_{i=1,j=1}^{2,n} f_{i,j} \leq 20 \tag{5}$$

$$f_{i,j} \geq 0 \tag{6}$$

In the formulation above, Equation 1 represents the objective function, which is the total cost of the water program. Equation 2 to Equation 5 are the four constraints for the problem:

- Equation 2 represents the supply constraint, meaning that the supply of each source cannot exceed the supply limit.
- Equation 3 represents the demand constraint, meaning the sum of the water delivered to town j should be greater than or equal to the demand.
- Equation 4 represents the water quality constraints, meaning that the water quality of each town should satisfy the lowest standard.
- Equation 5 represents the sum of the water flow from Sources 1 and 2 cannot exceed the limit of the common pipe.
- Equation 6 represents water flow from Source i to Town j is greater than or equal to 0.

iii. Write the $\mathbf{x}$,$\mathbf{c}$,$\boldsymbol{A}$,$\mathbf{b}$.

Write the vector $\mathbf{x}$ first as follows,

$$F_i = \begin{bmatrix} f_{i,1} & \cdots & f_{i,j} & \cdots & f_{i,n} \end{bmatrix} \in \mathbb{R}^n \tag{7}$$

$$\mathbf{x} = \begin{bmatrix} F_1 & \cdots & F_j & \cdots & F_m \end{bmatrix} \in \mathbb{R}^{mn} \tag{8}$$

Then, write the vector $\mathbf{c}$ as follows,

$$C_i = \begin{bmatrix} c_{i,1} & \cdots & c_{i,j} & \cdots & c_{i,n} \end{bmatrix} \in \mathbb{R}^n \tag{9}$$

$$\mathbf{c} = \begin{bmatrix} C_1 & \cdots & C_j & \cdots & C_m \end{bmatrix} \in \mathbb{R}^{mn} \tag{10}$$

$$
A = \begin{bmatrix}
\mathbf{1}_n & \mathbf{0}_n & \cdots & \mathbf{0}_n \\
\mathbf{0}_n & \mathbf{1}_n & \cdots & \mathbf{0}_n \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0}_n & \mathbf{0}_n & \cdots & \mathbf{1}_n \\
(q_1 - Q)D_1 & (q_2 - Q)D_1 & \cdots & (q_m - Q)D_1 \\
\vdots & \vdots & \ddots & \vdots \\
(q_1 - Q)D_i & (q_2 - Q)D_i & \cdots & (q_m - Q)D_i \\
\vdots & \vdots & \ddots & \vdots \\
(q_1 - Q)D_n & (q_2 - Q)D_n & \cdots & (q_m - Q)D_n \\
-D_1 & -D_1 & \cdots & -D_1 \\
\vdots & \vdots & \ddots & \vdots \\
-D_i & -D_i & \cdots & -D_i \\
\vdots & \vdots & \ddots & \vdots \\
-D_n & -D_n & \cdots & -D_n \\
\mathbf{1}_n & \mathbf{1}_n & \cdots & \mathbf{0}_n \\
-1 & \mathbf{0}_{mn-1} & & \\
\vdots & \vdots & \vdots & \\
\mathbf{0}_k & -1 & \mathbf{0}_{mn-k-1} & \\
\vdots & \vdots & \vdots & \\
 & \mathbf{0}_{mn-1} & -1 &
\end{bmatrix}^\top \in \mathbb{R}^{(m+2n+1+mn)\times mn} \quad (11)
$$

In the matrix $\boldsymbol{A}$, A_{qi} can be written as follows:

$$D_1 = \begin{bmatrix} 1 & \mathbf{0}_{n-1} \end{bmatrix} \in \mathbb{R}^n \quad (12)$$

$$D_2 = \begin{bmatrix} 0 & 1 & \mathbf{0}_{n-2} \end{bmatrix} \in \mathbb{R}^n \quad (13)$$

$$\cdots \quad (14)$$

$$D_n = \begin{bmatrix} \mathbf{0}_{n-1} & 1 \end{bmatrix} \in \mathbb{R}^n \quad (15)$$

$$
\mathbf{b} = \begin{bmatrix}
s_1 \\ \vdots \\ s_i \\ \vdots \\ s_m \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ -d_1 \\ \vdots \\ -d_i \\ \vdots \\ -d_n \\ 20 \\ \mathbf{0}_{mn}^\top
\end{bmatrix}^\top \in \mathbb{R}^{(m+2n+1+mn)} \quad (16)
$$

Then the LP can be written in the format as follows:

$$\min \quad \mathbf{c}^\top \mathbf{x} \tag{17}$$
$$\text{subject to} \quad \boldsymbol{A}\mathbf{x} \preceq \mathbf{b} \tag{18}$$

(c) Solution:

Town	Appletown	Berrytown	Cherrytown	Grapetown	Mangotown	**Supply Sum**
Source 1	0	0	0	0	15	15
Source 2	0	0	0	0	5	5
Source 3	9.375	3.125	15.625	6.25	18.281	52.656
Source 4	20.625	6.875	34.375	13.75	1.719	77.344
Demand Sum	30	10	50	20	40	± 0
Quality	1200	1200	1200	1200	1199.99	**Cost: 199,251**

Table 3: Optimal solution for the formulated LP

The MATLAB code can be found in the attached file. Please find Table 3 for the optimal solution. All the constraints are active.

(d) Demand sensitivity analysis.

i. 5% lower than forecast. See Table 4 for the details.
The total cost is 6.01% less than before. Water flow from Source 4 to Mangotown drops to 0.344 ML. Other links keep the same trend (about 5% drop).

Town	Appletown	Berrytown	Cherrytown	Grapetown	Mangotown	**Supply Sum**
Source 1	0	0	0	0	15	15
Source 2	0	0	0	0	5	5
Source 3	8.906	2.969	14.844	5.938	17.656	50.313
Source 4	19.594	6.531	32.656	13.062	0.344	72.187
Demand Sum	28.5	9.5	47.5	19	38	± 0
Quality	1199.99	1199.99	1199.99	1199.99	1199.99	**187,271** (-6.01%)

Table 4: Optimal solution for the scenario of 5% lower than forecast

ii. 5% higher than forecast.
No feasible solution. I tested different scenarios with the increase in demands. 2.9% is the highest feasible increasing range. I guess probably the water quality constraints cannot be satisfied.

iii. the daily supply limits for sources 3 and 4 are 95 and 75 ML respectively.
No feasible solution. I think the reason is quite understandable. The average hardness for Source 3 and 4 is 1594 kg/ML, probably the main contributor to the infeasible issue.

(e) Kiwi Inc. will add water demand to Berrytown. The county of Orchard should **NOT** accept the conditions. Add a decision variable y, representing the water flow sold to Kiwi Inc., to the raw program. Only need to make a slight change to the raw LP and run the LP. The optimal solution keeps the same and the optimal $y = 0$. However, if the price is increased to \$1700/ML, the county of Orchard could accept the conditions. And the maximum capacity to serve Kiwi Inc. is 4.091 ML per day.

Price	y
1500	0
1600	0
1700	3.864
1800	3.864
1900	4.091
2000	4.091
3000+	4.091

Table 5: The relationship between price and the ability to serve Kiwi Inc.

(f) Investment. Assume that the limit of Source 1 has been increased to 17 ML per day. The optimal cost is \$199,232 per day. \$199,232 + \$14 = \$199,246 < \$199,251. Hence, the county should choose the option. It can save \$1825 per year.

```
clear;clc;
f = 1.0 % demand coefficient
c = [360 370 350 355 365 ...
    420 425 430 435 415 ...
    700 715 685 720 725 ...
    2000 2015 1995 1985 2020]
A = [1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1
    -950 0 0 0 0 -1000 0 0 0 0 1100 0 0 0 0 -500 0 0 0 0
    0 -950 0 0 0 0 -1000 0 0 0 0 1100 0 0 0 0 -500 0 0 0
    0 0 -950 0 0 0 0 -1000 0 0 0 0 1100 0 0 0 0 -500 0 0
    0 0 0 -950 0 0 0 0 -1000 0 0 0 0 1100 0 0 0 0 -500 0
    0 0 0 0 -950 0 0 0 0 -1000 0 0 0 0 1100 0 0 0 0 -500
    -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0
    0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0
    0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0
    0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0
    0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1
    1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
    -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 -1 0 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 0 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 0 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 0 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 0 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1]
b = [17 10 60 80 ...
```

```
    0 0 0 0 0 ...
    -30*f -10*f -50*f -20*f -40*f ...
    20 0 0 0 0...
    0 0 0 0 ...
    0 0 0 0...
    0 0 0 0...
    0 0 0 0... ]
n = 20
cvx_begin
    variable x(n)
    minimize(c * x)
    subject to
        A * x<= b.'
cvx_end
```

2. You are given the set of all 2×2 matrices with diagonal elements $(1, 2)$ which we can write as

$$\mathcal{R} = \left\{ \begin{bmatrix} 1 & x \\ y & 2 \end{bmatrix}, \; x, y \in \mathbb{R} \right\}.$$

   (a) Is the set $\mathcal{R}$ convex?
   (b) Are rank constraints on matrices convex or nonconvex?
   (c) Defining a subset of $\mathcal{R}$ as $\mathcal{R}_1 \subset \mathcal{R}$ to be the set of rank-one matrices. Derive conditions on x and y that allow for all matrices $\mathcal{R}_1 \subset \mathcal{R}$ to be rank-one.
   (d) Is the set $\mathcal{R}_1$ convex?
   (e) Write the following optimization problem

$$\text{minimize} \quad ||A||_F^2 \quad \text{subject to} \quad A \in \mathcal{R}_1$$

   as an explicit optimization problem, where $||A||_F$ is the Frobenius norm. Is this problem convex?

Solutions:

(a) According to the definition: For any R_1 and $R_2 \in \mathcal{R}$,

$$R_1 = \begin{bmatrix} 1 & x_1 \\ y_1 & 2 \end{bmatrix}, x_1, y_1 \in \mathbb{R} \tag{19}$$

$$R_2 = \begin{bmatrix} 1 & x_2 \\ y_2 & 2 \end{bmatrix}, x_2, y_2 \in \mathbb{R} \tag{20}$$

$$\theta R_1 + (1-\theta) R_2 = \begin{bmatrix} 1 & \theta x_1 + (1-\theta) x_2 \\ \theta y_1 + (1-\theta) y_2 & 2 \end{bmatrix} \tag{21}$$

$$x_1, y_1, x_2, y_2 \in \mathbb{R} \tag{22}$$

$$\Rightarrow \quad \theta x_1 + (1-\theta) x_2, \theta y_1 + (1-\theta) y_2 \in \mathbb{R} \tag{23}$$

Hence, the set $\mathcal{R}$ is convex.

(b) The rank constraints would be $xy - 2 = 0$ or $xy - 2 \neq 0$.

- $xy - 2 = 0$. It is the rank-1 constraint.

$$(\theta x_1 + (1-\theta) x_2)(\theta y_1 + (1-\theta) y_2) - 2 \tag{24}$$

$$= \quad 2\theta^2 + 2(1-\theta)^2 + \theta(1-\theta)(x_1 y_2 + x_2 y_1) - 2 \neq 0 \tag{25}$$

  It is a nonconvex constraint.

- $xy - 2 \neq 0$. It is the rank-2 constraint.

$$(\theta x_1 + (1-\theta) x_2)(\theta y_1 + (1-\theta) y_2) \tag{26}$$

$$= \quad \theta^2 x_1 y_1 + (1-\theta)^2 x_2 y_2 + \theta(1-\theta)(x_1 y_2 + x_2 y_1) \neq 0 \tag{27}$$

  It is a convex constraint.

(c) $xy - 2 = 0$.

(d) The subset $\mathcal{R}$ is nonconvex. See Equation 25 for the proof.

(e) Write the optimization problem as an explicit one.

$$\min \quad x^2 + y^2 + 5 \tag{28}$$

$$\text{subject to} \quad xy - 2 = 0 \tag{29}$$

Obviously, it is a nonconvex problem (the constraint is nonconvex).

3. (***Reading Section 4.6.2 from the textbook can help in solving this problem***.) Consider the following primal SDP:

$$p^* = \text{minimize } \; x \quad \text{subject to } \begin{bmatrix} 1 & x+y \\ x+y & y \end{bmatrix} \succeq 0, \;\; x, y \in \mathbb{R}$$

(a) Show that this SDP is strictly feasible.

(b) Find the dual of the primal SDP. This will be an SDP in standard form.

(c) Show that the dual SDP is infeasible.

(d) Since the primal SDP is strictly feasible we know that strong duality holds. However, the dual SDP is infeasible. What does this say about the primal value p^*? Can you directly justify this?

Solutions:

(a) According to the constraint, the following inequalities can be satisfied.

$$1 > 0 \tag{30}$$
$$y \geq 0 \tag{31}$$
$$y - (x+y)^2 \geq 0 \tag{32}$$

Hence, x is bounded as follows:

$$-y - \sqrt{y} \leq x \leq -y + \sqrt{y} \tag{33}$$

This SDP is strictly feasible since $p^* = -y - \sqrt{y}$.

(b) SDP in standard form. Write the SDP in inequality form firstly,

$$\min \quad \begin{bmatrix} 1 & 0 \end{bmatrix}^\top \begin{bmatrix} x & y \end{bmatrix} \tag{34}$$
$$\text{subject to} \quad x \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{35}$$

Then write the dual standard form SDP.

$$\max \quad -tr(\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{Z}) \tag{36}$$
$$\text{subject to} \quad 1 + tr(\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \mathbf{Z}) = 0 \tag{37}$$
$$Z \succeq 0 \tag{38}$$

(c) However, $\mathbf{Z}$ is a diagonal matrix, which means that $tr(\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \mathbf{Z}) = 0$. As a result, Equation 37 is not satisfied since $1 \neq 0$. **The dual SDP is infeasible.**

(d) Since the primal SDP is strictly feasible, we know that its optimal value, denoted by p*, is finite. Strong duality then implies that the optimal value of the dual SDP is also equal to p*. However, since the dual SDP is infeasible, it does not have an optimal value. The reason why the primal problem has optimal value while the dual problem **NOT** is that the complementary slackness condition is not satisfied.

4. Consider the well-known constrained, least-squares problem:

$$\text{minimize} \quad ||Ax-b||_2^2 \quad \text{subject to} \quad Gx = h$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, G \in \mathbb{R}^{p \times n}$ and $h \in \mathbb{R}^p$. Assume that $\text{rank}(A) = n$ and $\text{rank}(G) = p$.

(a) Using the KKT conditions, determine the optimal solution of this optimization problem.

Solution:
Let $G = \begin{bmatrix} g_1 & g_2 & \cdots & g_i & \cdots & g_p \end{bmatrix}^\top$, where $g_i \in \mathbb{R}^n$. The Lagrangian is:

$$\begin{aligned}
\mathcal{L}(x,\mu) &= ||Ax-b||_2^2 + \sum_{i=1}^{p} \mu_i (g_i^\top x - h_i) & (39)\\
&= (Ax-b)^\top (Ax-b) + \sum_{i=1}^{p} \mu_i (g_i^\top x - h_i) & (40)\\
&= (Ax-b)^\top (Ax-b) + \mu^\top (Gx-h) & (41)\\
&= x^\top A^\top A x + \left(G^\top \mu - 2A^\top b\right)^\top x - \mu^\top h + b^\top b & (42)
\end{aligned}$$

The KKT conditions are stated as follows:

$$\begin{aligned}
Gx^* = h \ \ &\text{(equality constraint)} & (43)\\
2A^\top A x^* + G^\top \mu^* - 2A^\top b = 0 \ \ &\text{(}1^{st}\text{ order condition)} & (44)
\end{aligned}$$

Since $rank(A) = n$, $A^\top A$ is then invertible.

$$x^* = \frac{1}{2}(A^\top A)^{-1}(-G^\top \mu^* + 2A^\top b) \tag{45}$$

Then put the equation back to Equation 43.

$$G(\frac{1}{2}(A^\top A)^{-1}(-G^\top \mu^* + 2A^\top b)) = h \tag{46}$$

Also, $rank(G) = p$, G is invertible.

$$\mu^* = -2\left(G\left(A^\top A\right)^{-1} G^\top\right)^{-1} \left(h - G\left(A^\top A\right)^{-1} A^\top b\right) \tag{47}$$

Put the optimal μ^* back to Equation 45.

$$x^* = \left(A^\top A\right)^{-1} \left(A^\top b + G^\top \left(G\left(A^\top A\right)^{-1} G^\top\right)^{-1} \left(h - G\left(A^\top A\right)^{-1} A^\top b\right)\right) \tag{48}$$

Here, x^* is the optimal solution for primal problem, μ^* is the optimal solution for the dual problem, and the strong duality holds (See Equation 43).

---

5. Let $p^{(1)}, \ldots, p^{(r)}$ and $q^{(1)}, \ldots, q^{(s)}$ be different vectors in $\mathbb{R}^d$, where $r, s \geq 1$. Let $\mathcal{P}$ denote the polytope defined as the convex hull of the $p^{(i)}$'s, and $\mathcal{Q}$ the polytope defined as the convex hull of of the $q^{(i)}$'s (check the textbook definition on what constitutes a convex hull and polytope).

   Define a new matrix $C \in \mathbb{R}^{d \times n}$ whose i-th column is $p^{(i)}$ for all $1 \leq i \leq r$ and whose $(r+j)$-th column is $-q^{(j)}$ for all $1 \leq j \leq s$.

   (a) Given the above, formulate an optimization problem of computing the minimum $\mathcal{L}_2$ squared distance between points in $\mathcal{P}$ and $\mathcal{Q}$. The problem should be written as a QP with an objective function $||Cx||_2^2$ where $x \in \mathbb{R}^n$.

   (b) Let $y = Cx$. Demonstrate that the previous QP formulated in the previous part of this problem can be written as a QP with the objective function $||y||_2^2$, with two optimization variables x and y of appropriate dimensions.

   (c) Derive the dual of part b) and showcase that the dual objective function maximization can be written as

$$d^* = \underset{\lambda}{\text{maximize}} \left(-\frac{1}{4}\lambda^\top \lambda + \underset{i=1,\ldots,r}{\text{minimize}} \quad \lambda^\top p^{(i)} - \underset{j=1,\ldots,s}{\text{maximize}} \quad \lambda^\top q^{(j)} \right).$$

Solutions:

(a) According to the definition of the convex hull, every point in $\mathcal{P}$ can be written as:

$$\mathcal{P} = \left\{ \sum_{i=1}^{r} x_i p^{(i)}, \ \ s.t. \sum_{i=1}^{r} x_i = 1, \ \ x_i \geq 0 \right\} \tag{49}$$

Similarly, $\mathcal{Q}$ can be defined as follows:

$$\mathcal{Q} = \left\{ \sum_{j=1}^{s} x_{r+j} q^{(j)}, \ \ s.t. \sum_{j=r+1}^{n} x_j = 1, \ \ x_j \geq 0 \right\} \tag{50}$$

Write the $\boldsymbol{C}$ as stated.

$$\boldsymbol{C} = \begin{bmatrix} p^{(1)} \\ \vdots \\ p^{(i)} \\ \vdots \\ p^{(r)} \\ -q^{(1)} \\ \vdots \\ -q^{(j)} \\ \vdots \\ -q^{(s)} \end{bmatrix} \in \mathbb{R}^{d \times n} \tag{51}$$

$$\boldsymbol{x} = \begin{bmatrix} x_1 & \cdots & x_i & \cdots & x_r & x_{r+1} & \cdots & x_j & \cdots & x_n \end{bmatrix} \in \mathbb{R}^n \tag{52}$$

Then the function $\boldsymbol{C}\mathbf{x}$ can be written as follows:

$$\boldsymbol{C}\mathbf{x} = \sum_{i=1}^{r} x_i p^{(i)} - \sum_{j=1}^{s} x_{r+j} q^{(j)} \tag{53}$$

The optimization problem can be written as follows then.

$$\min \quad (\sum_{i=1}^{r} x_i p^{(i)} - \sum_{j=1}^{s} x_{r+j} q^{(j)})^\top (\sum_{i=1}^{r} x_i p^{(i)} - \sum_{j=1}^{s} x_{r+j} q^{(j)}) \tag{54}$$

$$\text{subject to} \quad \sum_{i=1}^{r} x_i = 1 \tag{55}$$

$$\sum_{j=r+1}^{n} x_j = 1 \tag{56}$$

$$x_i \geq 0 \tag{57}$$

(b) QP with decision variables as $(x, y) \in \mathbb{R}^{n+d}$. The QP with (x,y) can be written as follows then.

$$\min \quad y^\top y \tag{58}$$

$$\text{subject to} \quad y = \boldsymbol{C}\mathbf{x} \tag{59}$$

$$\sum_{i=1}^{r} x_i = 1 \tag{60}$$

$$\sum_{j=r+1}^{n} x_j = 1 \tag{61}$$

$$x_i \geq 0 \tag{62}$$

(c) To derive the dual of part 5b, the Lagrangian can be written as:

$$\mathcal{L}(x, y, \mu, \alpha, \beta, \lambda) = y^\top y + \mu^\top (y - \boldsymbol{C}\mathbf{x}) + \alpha(\sum_{i=1}^{r} x_i - 1) + \beta(\sum_{j=r+1}^{n} x_j - 1) - \lambda^\top \mathbf{x} \tag{63}$$

where the variable $\mu \in \mathbb{R}^d, \alpha \in \mathbb{R}, \beta \in \mathbb{R}, \lambda \in \mathbb{R}^n_+$ are the dual variables. The dual problem can then be written as follows:

$$\min \quad \mathcal{L}(x, y, \mu, \alpha, \beta, \lambda) \tag{64}$$

$$\text{subject to} \quad \lambda \succeq 0 \tag{65}$$

The KKT conditions for optimal solution (x^*, y^*) are stated as follows:

$$\partial_{x_i} \mathcal{L}(x, y, \mu, \alpha, \beta, \lambda) = 0 \qquad (1^{st} \text{ order condition for } x, n \text{ equations}) \tag{66}$$

$$\partial_{y_j} \mathcal{L}(x, y, \mu, \alpha, \beta, \lambda) = 0 \qquad (1^{st} \text{ order condition for } y, d \text{ equations}) \tag{67}$$

$$y = Cx \qquad (\text{equality constraints for Equation 59}) \tag{68}$$

$$\sum_{i=1}^{r} x_i = 1 \qquad (\text{equality constraints for Equation 60}) \tag{69}$$

$$\sum_{j=r+1}^{n} x_j = 1 \qquad (\text{equality constraints for Equation 61}) \tag{70}$$

$$\lambda_i x_i = 0 \qquad (\text{complementary slackness}, n \text{ equations}) \tag{71}$$

Solve the Equations 67, which are in the quadratic form without constraints,

$$y_j^* = -\frac{\mu_j}{2} \tag{72}$$

The Lagrangian objective function can then be written as:

$$\mathcal{L} = -\frac{1}{4}\mu^\top \mu - \mu^\top \boldsymbol{Cx} \tag{73}$$

$$= -\frac{1}{4}\mu^\top \mu - \mu^\top (\sum_{i=1}^{r} x_i p^{(i)} - \sum_{j=1}^{s} x_{r+j} q^{(j)}) \tag{74}$$

$$\leq -\frac{1}{4}\mu^\top \mu - \mu^\top \min(p^{(i)}) \sum_{i=1}^{r} x_i + \mu^\top \max(q^{(i)}) \sum_{i=r+1}^{n} x_j \tag{75}$$

$$= -\frac{1}{4}\mu^\top \mu - \mu^\top \min(p^{(i)}) + \mu^\top \max(q^{(j)}) \tag{76}$$

Here, $\min(p^{(i)})$ and $\max(q^{(j)})$ are vectors $\in \mathbb{R}^d$. Take $\min(p^{(i)})$ as an example, assume $p^{(i)} = \begin{bmatrix} p_1^{(i)} & \cdots & p_k^{(i)} & \cdots & p_d^{(i)} \end{bmatrix}$.

$$\min(p^{(i)}) = \begin{bmatrix} \min(p_1^{(1))}, \cdots, p_1^{(i))}, \cdots, p_1^{(r))}) \\ \vdots \\ \min(p_k^{(1))}, \cdots, p_k^{(i))}, \cdots, p_k^{(r))}) \\ \vdots \\ \min(p_d^{(1))}, \cdots, p_d^{(i))}, \cdots, p_d^{(r))}) \end{bmatrix}^\top \in \mathbb{R}^d \tag{77}$$

Same for the notation $\max(q^{(j)})$. Hence, the upper bound of the Lagrangian objective function can be written as follows:

$$\sup \mathcal{L}(\mu) = -\frac{1}{4}\mu^\top \mu - \mu^\top \min(p^{(i)}) + \mu^\top \max(q^{(j)}) \tag{78}$$

Hence, $d^* = \max_\mu(\sup \mathcal{L}(\mu))$.

6. You are given the following optimization problem:

$$\text{minimize } f(x) = x_1^2 + x_2^2 \quad \text{subject to} \quad (x_1 - 1)^2 + (x_2 - 1)^2 \leq 2 \ , \ (x_1 - 1)^2 + (x_2 + 1)^2 \leq 2.$$

(a) Draw the feasible region of this problem on Matlab. Furthermore, on top of the figure sketch the level sets of $f(x)$.

(b) Is this problem convex? Verify that strong duality holds for this problem via Slater's condition.

(c) Solve the problem via deriving the KKT conditions. There are a few cases to consider for the multipliers. List them all. Trivial solutions without justifications are not accepted.

Solutions:

(a) See the attached figure (Figure 1) for details.

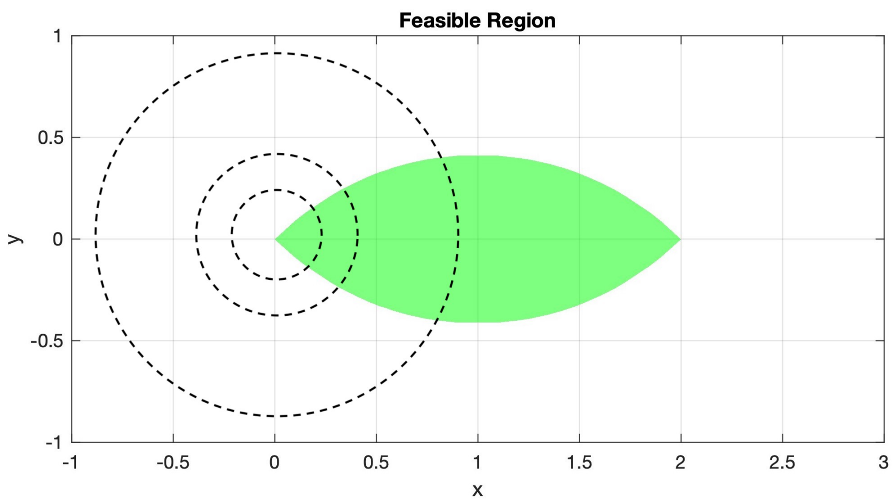

Figure 1: Feasible region of this problem

(b) **Convex**. Write the Lagrangian function as follows.

$$\mathcal{L}(x, \mu, \lambda) = x_1^2 + x_2^2 - \mu(x_1^2 + x_2^2 - 2x_1 - 2x_2) - \lambda(x_1^2 + x_2^2 - 2x_1 + 2x_2) \tag{79}$$

The primal problem is convex, and the dual problem is strictly feasible (seen from Figure 1). Then the strong duality holds for the problem according to Slater's condition.

(c) KKT conditions are stated as follows.

$$\partial \mathcal{L}_{x_1} = 2(1 - \lambda - \mu)x_1 + (2\mu + 2\lambda) = 0 \quad (1^{st} \text{ order condition}) \tag{80}$$
$$\partial \mathcal{L}_{x_2} = 2(1 - \lambda - \mu)x_2 + (2\mu - 2\lambda) = 0 \quad (1^{st} \text{ order condition}) \tag{81}$$
$$\mu(x_1^2 + x_2^2 - 2x_1 - 2x_2) = 0 \quad \text{(complementary slackness)} \tag{82}$$
$$\lambda(x_1^2 + x_2^2 - 2x_1 + 2x_2) = 0 \quad \text{(complementary slackness)} \tag{83}$$
$$x_1^2 + x_2^2 - 2x_1 - 2x_2 \leq 0 \quad \text{(inequality constraint)} \tag{84}$$
$$x_1^2 + x_2^2 - 2x_1 + 2x_2 \leq 0 \quad \text{(inequality constraint)} \tag{85}$$
$$\lambda, \mu \geq 0 \quad \text{(Lagrange multiplier)} \tag{86}$$

- $\lambda = 0, \mu = 0$.
  $\Rightarrow x_1 = 0, \ \ x_2 = 0$. Not optimal.
- $\mu = 0, x_1^2 + x_2^2 - 2x_1 + 2x_2 = 0, \lambda \neq 0$:
  $\Rightarrow x_1 = 0, \ \ x_2 = 0$. Not optimal.
  $\Rightarrow x_1 = 2, \ \ x_2 = -2$. Inequality constraint not satisfied.

- $\lambda = 0, x_1^2 + x_2^2 - 2x_1 - 2x_2 = 0$, $\mu \neq 0$:
  $\Rightarrow x_1 = 0,\ \ x_2 = 0$. Not optimal.
  $\Rightarrow x_1 = 2,\ \ x_2 = 2$. Inequality constraint not satisfied.
- $x_1^2 + x_2^2 - 2x_1 - 2x_2 = 0, x_1^2 + x_2^2 - 2x_1 + 2x_2 = 0$.
  $\Rightarrow x_1 = 0;\ \ x_2 = 0$. Not optimal.
  $\Rightarrow x_1 = 2;\ \ x_2 = 0$. **OPTIMAL!** $\Rightarrow f(x) = 4$

Hence, $x_1^* = 2,\ \ x_2^* = 0,\ \ f(x^*) = 4$.

---

7. For this given SDP

$$\textbf{(LMI-1)} \quad \min \quad \mathrm{trace}(P) \quad \text{subject to} \quad A^\top P + PA - C^\top Y^\top - YC \prec 0, \quad P = P^\top \succ 0,$$

(where matrices P and Y are the optimization variables and matrices A and C are constant) derive the SDP form in the LMI form. This form can be written as $x \in \mathbb{R}^n$

$$\begin{aligned} &\min \quad c^\top x \\ &\text{subject to} \sum_{i=1}^{n} x_i F_i + G \preceq 0 \\ &Dx = b \end{aligned}$$

where $F_1, \ldots, F_n, G \in \mathbb{S}^m$ (i.e., $F_1, \ldots, F_n, G$ are **symmetric** constant matrices). Essentially, I want you to derive these constant matrices for **(LMI-1)** given above.
Solutions:
Rewrite the strict inequalities to "greater than or equal",

$$\begin{bmatrix} -P & 0 \\ 0 & A^\top P + PA - C^\top Y^\top - YC \end{bmatrix} + \begin{bmatrix} \epsilon_1 I & 0 \\ 0 & \epsilon_2 I \end{bmatrix} \preceq 0 \tag{87}$$

where $\epsilon_1, \epsilon_2 > 0$ and are very small.

$$\begin{bmatrix} -P & 0 \\ 0 & A^\top P + PA \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & C^\top Y^\top + YC \end{bmatrix} + \begin{bmatrix} \epsilon_1 I & 0 \\ 0 & \epsilon_2 \end{bmatrix} \preceq 0 \tag{88}$$

$$P = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ x_2 & x_{m+1} & \dots & x_{2m-1} \\ \vdots & \vdots & \ddots & \vdots \\ x_m & x_{2m-1} & \dots & x_{\frac{m(m+1)}{2}} \end{bmatrix} \in \mathbb{R}^{m \times m}, \quad \frac{m(m+1)}{2} \text{ elements.} \tag{89}$$

$$Y = \begin{bmatrix} y_1 & y_2 & \dots & y_m \\ y_{m+1} & y_{m+2} & \dots & y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m^2-m+1} & y_{m^2-m+2} & \dots & y_{m^2} \end{bmatrix} \in \mathbb{R}^{m \times m}, \quad m^2 \text{ elements} \tag{90}$$

$$E_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, \ldots, E_n = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 1 \end{bmatrix}, n = \frac{m(m+1)}{2} \tag{91}$$

$$F_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, F_2 = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, \ldots, F_l = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 1 \end{bmatrix}, l = m^2 \tag{92}$$

The problem is equivalent to the following SDP:

$$\sum_{i=1}^{n} x_i \begin{bmatrix} -E_i & 0 \\ 0 & A^\top E_i + E_i A \end{bmatrix} - \sum_{j=1}^{l} y_j \begin{bmatrix} 0 & 0 \\ 0 & C^\top F_j + F_j C \end{bmatrix} + \begin{bmatrix} \epsilon_1 I & 0 \\ 0 & -\epsilon_2 I \end{bmatrix} \preceq 0 \tag{93}$$

In summary, the SDP problem of **(LMI-1)** can be organized as follows:

$$\min \quad c^\top x \tag{94}$$

$$\text{subject to} \sum_{i=1}^{n} x_i \begin{bmatrix} -E_i & 0 \\ 0 & A^\top E_i + E_i A \end{bmatrix} - \sum_{j=1}^{l} y_j \begin{bmatrix} 0 & 0 \\ 0 & C^\top F_j + F_j C \end{bmatrix} + \begin{bmatrix} \epsilon_1 I & 0 \\ 0 & -\epsilon_2 I \end{bmatrix} \preceq 0 \tag{95}$$

where $c = \begin{bmatrix} 1 & \mathbf{0}_m & 1 & \mathbf{0}_m & \cdots & \mathbf{0}_m & 1 \end{bmatrix} \in \mathbb{R}^n$, which help pick up the diagonal value of the matrix P.

---

8. Solve the previous problem but for a different SDP given as follows

$$\textbf{(LMI-2)} \quad \min \quad ||P||_1 \quad \text{subject to} \quad \begin{bmatrix} A^\top P + PA & PB - C^\top \\ B^\top P - C & D^\top D - I \end{bmatrix} \prec 0, \quad P = P^\top \succ 0,$$

where matrices $B, C,$ and D are all constant. Matrix P is the only matrix variable. Note that for this problem you need to utilize the epigraph trick to model the L_1 norm objective function into a linear objective function with some constraints.

Solutions:

Rewrite the strict inequalities to "greater than or equal",

$$\begin{bmatrix} -P & 0 & 0 \\ 0 & A^\top P + PA & PB - C^\top \\ 0 & B^\top P - C & D^\top D - I \end{bmatrix} + \begin{bmatrix} \epsilon_1 I & 0 & 0 \\ 0 & -\epsilon_2 I & -\epsilon_3 I \\ 0 & -\epsilon_4 I & -\epsilon_5 I \end{bmatrix} \preceq 0 \tag{96}$$

$$E_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, \dots, E_n = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 1 \end{bmatrix}, n = \frac{m(m+1)}{2} \tag{97}$$

The SDP problem of (LMI-2) can be organized as follows:

$$\min \quad t \tag{98}$$

$$\text{subject to} \sum_{i=1}^{n} x_i \begin{bmatrix} -E_i & 0 & 0 \\ 0 & A^\top E_i + E_i A & E_i B - C^\top \\ 0 & B^\top & 0 \end{bmatrix} + \begin{bmatrix} \epsilon_1 I & 0 & 0 \\ 0 & -\epsilon_2 I & -\epsilon_3 I \\ 0 & -\epsilon_4 I & D^\top D - (1+\epsilon_4) I \end{bmatrix} \preceq 0 \tag{99}$$

$$x \preceq t1^\top \tag{100}$$

$$x \succeq -t1^\top \tag{101}$$

$$t \geq 0 \tag{102}$$

9. Show that the following bilinear matrix inequalities (or in general any BMI) are nonconvex.

$$\textbf{(BMI-1)} \quad \text{find} \quad P, \alpha \quad \text{subject to} \quad A^\top P + PA - \alpha P \prec 0, \quad P = P^\top \succ 0, \alpha > 0$$

where α is a scalar.

$$\textbf{(BMI-2)} \quad \text{find} \quad P, Y \quad \text{subject to} \quad A^\top P + PA - Y^\top P - PY \prec 0, \quad P = P^\top \succ 0$$

where P and Y are matrix variables.
Solutions:
BMI-1:

$$A^\top P_1 + P_1 A - \alpha_1 P_1 \prec 0 \tag{103}$$

$$A^\top P_2 + P_2 A - \alpha_2 P_2 \prec 0 \tag{104}$$

Consider $P = \theta P_1 + (1-\theta)P_2$ and $\alpha = \theta\alpha_1 + (1-\theta)\alpha_2$.

$$A^\top[\theta P_1 + (1-\theta)P_2] + [\theta P_1 + (1-\theta)P_2]A - [\theta\alpha_1 + (1-\theta)\alpha_2][\theta P_1 + (1-\theta)P_2] \tag{105}$$

$$= \theta(A^\top P_1 + P_1 A - \alpha_1 P_1) + (1-\theta)(A^\top P_2 + P_2 A - \alpha_2 P_2) - [\theta\alpha_1 + (1-\theta)\alpha_2][\theta P_1 + (1-\theta)P_2] \tag{106}$$

$$\prec \theta\alpha_1 P_1 + (1-\theta)\alpha_2 P_2 - [\theta\alpha_1 + (1-\theta)\alpha_2][\theta P_1 + (1-\theta)P_2] \neq 0 \tag{107}$$

Hence, **BMI-1** is nonconvex.
BMI-2:

$$A^\top P_1 + P_1 A - Y_1^\top P_1 - P_1 Y_1 \prec 0 \tag{108}$$

$$A^\top P_2 + P_2 A - Y_2^\top P_2 - P_2 Y_2 \prec 0 \tag{109}$$

Consider $P = \theta P_1 + (1-\theta)P_2$ and $Y = \theta Y_1 + (1-\theta)Y_2$.

$$A^\top[\theta P_1 + (1-\theta)P_2] + [\theta P_1 + (1-\theta)P_2]A \tag{110}$$

$$- [\theta Y_1^\top + (1-\theta)Y_2^\top][\theta P_1 + (1-\theta)P_2] - [\theta P_1 + (1-\theta)P_2][\theta Y_1 + (1-\theta)Y_2] \tag{111}$$

$$\prec \theta(Y_1^\top P_1 + P_1 Y_1) + (1-\theta)(Y_2^\top P_2 + P_2 Y_2) \tag{112}$$

$$- [\theta Y_1^\top + (1-\theta)Y_2^\top][\theta P_1 + (1-\theta)P_2] - [\theta P_1 + (1-\theta)P_2][\theta Y_1 + (1-\theta)Y_2] \tag{113}$$

$$\neq 0 \tag{114}$$

Hence, **BMI-2** is nonconvex.
