CSE 203B W24 Homework 4

Due Time : 11:50pm, Friday Feb. 23, 2024 Submit to Gradescope

In this homework, we work on exercises from the textbook. Problems 4.1, 4.8, 4.11, and 4.15 are related to LP. Problems 4.21, 4.39, and 4.47 are related to QCQP, and SDP. Problem 5.3 is about the basic definition of duality. Problems 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, and 5.10 are examples and applications of duality. Also, we practice using the convex optimization tools on a linear programming problem, and a quadratic programming problems.

Total points: 50. Exercises are graded by completion, and assignments are graded by correctness.

I. Exercises from textbook chapters 4 & 5 (15 pts, 1pt for each problem)
4.1, 4.8, 4.11, 4.15, 4.21, 4.39, 4.47, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10.

II. Assignments (35 pts)

II.1 Linear Programming: You are free to use any software packages. (15 pts)

Given

\[
A = \begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 \\
2 & 0 & 1 & 0 \\
0 & 3 & -1 & 2 \\
1 & 0 & 2 & 0
\end{bmatrix},
\]

\[
b^T = [1 \ 5 \ 2 \ 3 \ 2],
\]

\[
c^T = [2 \ 1 \ 2 \ 3],
\]

and \( n = 4 \), perform steps A, B, and C for problems II.1.1, II.1.2, II.1.3, and II.1.4.

A. Solve the following linear programming problems twice, once in the primal formulation and once in the dual formulation.

B. Check the feasibility of the solution. If a solution is not found, explain why a solution is not available and suggest how to mitigate the issue if you are the project leader.

C. Compare the primal and dual solutions. If the primal and dual formulation solutions are different, explain the difference.

II.1.1. minimize \( f_0(x) = c^T x \) subject to \( Ax \leq b, x \in \mathbb{R}^n \).

A. Solve in primal formulation.

The minimum value is \(-\infty\), which can be achieved by one feasible point \( x^T = [0 \ t \ 0 \ 0] \), where \( t \leq 0 \).

\[
\lim_{t \to -\infty} c^T x = \lim_{t \to -\infty} t = -\infty
\]

Solve in the dual formulation:

Refers to the formula (5.22) of the textbook Chapter 5.2.1, pp. 225.

The dual problem is equivalent to max \(-b^T \lambda \) s.t. \( A^T \lambda + c = 0, \lambda \geq 0 \). When the constraints are infeasible, we regard \(-\infty\) as the value of this dual problem.
The definition of strictly

∀

However, since λ

II.1.3. minimize

II.1.2. minimize

f

component of b.

p, h

Note: The standard form of the optimization problem is (4.1) in the textbook, which is min

strictly

problem is

∞

value is also unbounded below.

B. The primal problem is feasible but its value is unbounded below. The dual problem is infeasible and its

Thus, the dual problem is infeasible and its value is −∞, which is the same as the primal problem.

B. The primal problem is feasible but its value is unbounded below. The dual problem is infeasible and its

value is also unbounded below.

C. The primal and dual optimal values are the same, which is ∞. They are the same since the primal

problem is strictly feasible: ∃x = [0.1 0.1 0.1 0.1], Ax < b.

Note: The standard form of the optimization problem is (4.1) in the textbook, which is min f0(x) s.t.

∀1 ≤ i ≤ m, f1(x) < 0 and ∀1 ≤ i ≤ p, h1(x) = 0.

The definition of strictly feasible with respect to this form is ∃x, ∀1 ≤ i ≤ m, f1(x) < 0, and ∀1 ≤ i ≤ p, h1(x) = 0.

In this question, m = 5 and p = 0. ∀1 ≤ i ≤ 5, f1(x) = a⊤ i x − bi, where ai is the i-th row of A, bi is the i-th component of b.

II.1.2. minimize f0(x) = c⊤x subject to Ax = b, x ∈ Rn.

Rubric: The solution does not require parts A, B, C, since the primal problem is infeasible. The student

needs to point out that the original constraints are infeasible. Then the student is expected to relax as least

equality constraints to inequality. As shown in Winter 23’s solution, relaxation might be done by changing

= to ≤, such as a5 x − b5 = 0 to a5 x − b5 ≤ 0. However, I think it is also acceptable to change = to ≥. If

just relaxing one constraint does not help, then multiple could be relaxed. There will be a piazza post about

that relaxation.

[A|b] → RREF → Ib, where the last row is [0 0 0 0 | 1], which means [A|b] is inconsistent. Thus, the

primal problem’s constraint (Ax = b, x ∈ Rn) is inconsistent, and the primal problem is infeasible.

The equality constraint a5 x − b5 = 0 could be relaxed to a5 x − b5 ≤ 0, while other constraints remain the

same.

Then x = [2 3 13 3 2 3 2 3 14]⊤ is a feasible point, and the corresponding c⊤x = −7.0

II.1.3. minimize f0(x) = c⊤x subject to Ax ≤ b, x ∈ R+n.
The primal’s optimal value is 0. Considering \( c \geq 0 \), it must be true that \( \forall x \in \mathbb{R}_+^n, c^\top x \geq 0 \). Since \( c^\top x = 0 \) at the feasible point 0, the max possible value of \(-c^\top x\) is 0.

The dual problem’s optimal value is also 0.

The dual problem is max \(-b^\top \lambda\) s.t. \((A')^\top \lambda' + c = 0, \lambda' \geq 0\), where \( \lambda' \in \mathbb{R}^9, \lambda = \begin{bmatrix} \lambda_1' \\ \lambda_2' \\ \lambda_3' \\ \lambda_4' \end{bmatrix}, A' = \begin{bmatrix} A \\ -I_4 \end{bmatrix} \in \mathbb{R}^{9 \times 4} \)

and \( b' = \begin{bmatrix} b \\ 0 \end{bmatrix} \in \mathbb{R}^9 \). When the constraints are infeasible, we regard \(-\infty\) as the value of this dual problem.

\[-b^\top \lambda_* = 0\] at the feasible point \( \lambda_* = \begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 1 & 2 & 3 \end{bmatrix}^\top \).

Considering \( b \geq 0 \), it must be true that \( \forall \lambda \geq 0, -b^\top \lambda \geq 0 \). Thus, the max possible value of \(-b^\top \lambda\) is 0.

\[\text{II.1.4. minimize } f_0(x) = c^\top x \text{ subject to } Ax = b, x \in \mathbb{R}_+^n.\]

Rubric: The solution does not require parts A, B, C, since the primal problem is infeasible. The student needs to point out that the original constraints are infeasible. Then the student is expected to relax at least equality constraints in inequality. As shown in Winter 23’s solution, relaxation might be done by changing = to \( \leq \), such as \( a_{ij}^\top x - b_i = 0 \) to \( a_{ij}^\top x - b_i \leq 0 \). However, I think it is also acceptable to change = to \( \geq \). If just relaxing one constraint does not help, then multiple could be relaxed. There will be a piazza post about that relaxation.

\([A|b] \rightarrow \text{RREF} \rightarrow I_5\), where the last row is \( \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix} \), which means \([A|b]\) is inconsistent. Thus, \((Ax = b, x \in \mathbb{R}^n)\) is infeasible. Thus, the primal problem’s constraint \((Ax = b, x \in \mathbb{R}^n)\) is infeasible.

We can relax the equality constraint \( a_i^\top x - b_i = 0 \) to \( a_i^\top x - b_i \leq 0 \) for \( i = 1, 4 \), and then \( x = \begin{bmatrix} 2/3 \\ 1/9 \\ 2/3 \\ 0 \end{bmatrix} \) is a feasible point, and the corresponding \( c^\top x = \frac{35}{9} \).

\[\text{II.2 Graph embedding (20 pts) Graph embedding is an important problem in machine learning and graph theory. Given an undirected graph } G = (V, E) \text{ with } n \text{ vertices, the problem is to assign coordinates in } \mathbb{R}^m \text{ to each vertex } v \in V. \text{ Typically there are desired qualities or constraints imposed on the embedding—e.g. the coordinates assigned to connected nodes should be close with respect to some notion of distance. For example, the choice of Euclidean distance yields a quadratically constrained quadratic program (QCQP). Let } A \in \{0, 1\}^{n \times n} \text{ be the symmetric adjacency of } G, \text{ i.e. } a_{ij} = a_{ji} = 1 \text{ if edge } e_{ij} \in E \text{ otherwise } a_{ij} = a_{ji} = 0. \text{ Let } D \text{ be the corresponding diagonal degree matrix such that } D_{ii} = \sum_{j} A_{i,j}. \text{ The graph Laplacian is defined to be } L = D - A. \]

One well known way to do graph embedding (Laplacian Eigenmaps) is to solve formulation (I):

\[
\min_{X \in \mathbb{R}^{m \times n}} \langle X, LX \rangle \\
\text{s.t. } 1^\top X = 0, \\
X^\top X = I, \\
\tag{1}
\]

where the inner product \( \langle A, B \rangle \) is defined to be \( \text{tr}(A^\top B) \).

(i) Show that when \( m = 1 \), \( \langle x, LX \rangle = x^\top LX = \sum_{i,j \in E} (x_i - x_j)^2 \).
Similarly, any $x = x^T(D - A)x = x^TDx - x^TAx$

\[
\sum_i D_{ii}x_i^2 - \sum_{i,j \in E} 2x_ix_j
\]
\[
= \sum_i x_i^2 - \sum_{i,j \in E} 2x_ix_j
\]
\[
= \sum_{i,j \in E} (x_i^2 + x_j^2 - 2x_ix_j) = \sum_{i,j \in E} (x_i - x_j)^2
\]

(ii) Let us order the eigenvalues of Laplacian $L$ from small to large, i.e. $\lambda_0 \leq \lambda_1 \leq ... \leq \lambda_{n-1}$. Prove that the Laplacian $L$ is positive semidefinite with eigenvalue $\lambda_0 = 0$ and the corresponding eigenvector $v_0 = 1$.

Since matrices $D$ and $A$ are symmetric, $L$ is also symmetric. As shown in part (i), $x^TLx = \sum_{i,j \in E} (x_i - x_j)^2 \geq 0 \forall x \implies L$ is positive semidefinite.

For any eigenvector $v$,

\[ v^TLv \geq 0 \implies v^T\lambda v \geq 0 \implies \lambda v^Tv \geq 0 \implies \lambda \geq 0 \implies \text{all eigenvalues are non-negative.} \]

For $v = 1$,

\[Lv = (D-A)v = Dv - Av = D1 - A1\]

\[= [D_{11} \quad \sum_j A_{1j} \quad \vdots \quad \sum_j A_{nj}] - [\sum_j A_{2j} \quad \vdots \quad \sum_j A_{nj}] = 0\]

\[\implies L1 = 0 = 01 \implies 0 \text{ is an eigenvalue of } L. \text{ Since all eigenvalues of } L \text{ are non-negative, } 0 \text{ is the smallest eigenvalue i.e. } \lambda_0 = 0 \text{ and the corresponding eigenvector is } 1.\]

(iii) Prove that the optimal solution to formulation (1) is $X = [v_1, v_2, ..., v_m]$, where $v_i$ is the eigenvector corresponding to eigenvalue $\lambda_i$ of Laplacian $L$.

This is an instantiation of the min-max variational theorem in the context of Laplacian matrices. Note that $L$ is symmetric and PSD. It’s eigenvalues $\lambda_i \geq 0$. Additionally, the rows and columns of $L$ sum to zero. The unit vector with constant entries, $\bar{1}$, lies in the null space of $L$ i.e. is an eigenvector associated with eigenvalue 0. First, suppose $(\bar{1}, v_1, ..., v_{n-1})$ is an orthogonal basis for $R^n$ corresponding to eigenvalues $0 \leq \lambda_1 \leq ... \leq \lambda_{n-1}$. Let $X = [x_1, x_2, ..., x_m]$. Then any $x_1 \in R^n$ s.t. $x_1 \neq 0, x_1 \nsubseteq \bar{1}$ can be expressed as a linear combination of $v_1, ..., v_{n-1}; a_1v_1 + ... + a_{n-1}v_{n-1}$. Substituting the Rayleigh quotient, we get

\[
\frac{x_1^TLx_1}{x_1^T} = \frac{(\sum_{i=1}^n a_i v_i) ^T L (\sum_{i=1}^{n-1} a_i v_i)}{(\sum_{i=1}^n a_i v_i) ^T (\sum_{i=1}^{n-1} a_i v_i)} = \frac{\sum_{i=1}^{n-1} \lambda_i a_i^2 || v_i ||^2}{\sum_{i=1}^n a_i^2 || v_i ||^2} \geq \lambda_1
\]

\[\text{min}_{x_1} x_1^TLx_1 = \lambda_1\]

and minimum occurs at $x_1 = v_1$

Similarly, any $x_2 \in R^n$ s.t. $x_2 \neq 0, x_2 \nsubseteq \bar{1}, v_1$ can be expressed as a linear combination of $v_2, ..., v_{n-1}; a_2v_2 + ... + a_{n-1}v_{n-1}$. Substituting the Rayleigh quotient, we get
\[
\frac{x_2^T L x_2}{x_2^T x_2} = \frac{(\sum_{i=2}^{n-1} a_i v_i)^T L (\sum_{i=2}^{n-1} a_i v_i)}{(\sum_{i=2}^{n-1} a_i v_i)^T (\sum_{i=2}^{n-1} a_i v_i)} = \frac{\sum_{i=2}^{n-1} \lambda_i a_i^2 ||v_i||^2}{\sum_{i=2}^{n-1} a_i^2 ||v_i||^2} \geq \lambda_2
\]

\[\min_{x_2} x_2^T L x_2 = \lambda_2\]

and minimum occurs at \(x_2 = v_2\)

Hence, the optimal solution to formulation (1) is \(X = [v_1, v_2, ..., v_m]\), where \(v_i\) is the eigenvector corresponding to eigenvalue \(\lambda_i\) of Laplacian \(L\).

Note that \(1^T X = 0\) for the optimal \(X\) since eigenvectors \(v_1, v_2, ..., v_m\) are orthogonal to \(v_0 = 1\).

(iv) Ignoring the second-order constraint, we have formulation (2):

\[
\begin{align*}
\min_{X \in \mathbb{R}^{n \times m}} & \quad \langle X, LX \rangle \\
\text{s.t.} & \quad 1^T X = 0,
\end{align*}
\]

(2) Derive the dual formulation of formulation (2) and KKT conditions.

Dual formulation:

\[
\begin{align*}
\min_{X \in \mathbb{R}^{n \times m}} & \quad \langle X, LX \rangle \\
\text{s.t.} & \quad 1^T X = 0,
\end{align*}
\]

define the Lagrangian where \(\lambda\) are the Lagrange multipliers.

\[
\mathcal{L}(X, \lambda) = \langle X, LX \rangle + \langle v, 1^T X \rangle
\]

The first-order condition is then

\[
\frac{\partial \mathcal{L}(X, \lambda)}{\partial X} = LX + L^T X + 1v = 0
\]

\[
X = -(L + L^T)^{\dagger}1v
\]

\[
X = -\frac{L^T 1v}{2}
\]

where \(L^\dagger\) represents the pseudo inverse of \(L\). The associated dual problem is

\[
\begin{align*}
\max_v & \quad Tr(X^T LX) + Tr(v^T 1^T X) \\
\Rightarrow \max_v & \quad Tr\left(\frac{-v^T 1^T (L^\dagger)^T L - L^\dagger 1v}{2}\right) + Tr\left(\frac{v^T 1^T L^\dagger 1v}{2}\right) \\
\Rightarrow \max_v & \quad Tr\left(\frac{v^T 1^T L^\dagger 1v}{4}\right) - Tr\left(\frac{v^T 1^T L 1v}{2}\right) \\
\Rightarrow \max_v & \quad -(1v, L^\dagger 1v)/4
\end{align*}
\]

KKT Conditions:

- Primal feasibility: \(1^T X = 0\)
- Stationarity: \(\nabla_X \mathcal{L}(X, \nu) = 0\)
(v) Put back the second order constraint. However, change matrix $I$ into a positive definite matrix $C \in \mathbb{R}^{m \times m}$.

$$
\min_{X \in \mathbb{R}^{n \times m}} \langle X, LX \rangle \\
\text{s.t. } 1^T X = 0, \\
X^T X = C,
$$

Try to find an analytical solution as subproblem (iii). If we could not find an analytical solution, describe your approach to finding a numerical solution.

Consider the constraint $X^T X = C$

$$
X^T X = C^{\frac{1}{2}} C^{\frac{1}{2}} \\
X^T X C^{-\frac{1}{2}} = C^{\frac{1}{2}} C^{\frac{1}{2}} C^{-\frac{1}{2}} \\
X^T X C^{-\frac{1}{2}} = C^{\frac{1}{2}} \\
X^T X C^{-\frac{1}{2}} = (C^{\frac{1}{2}})^T \\
((C^{\frac{1}{2}})^T)^{-1} X^T X C^{-\frac{1}{2}} = ((C^{\frac{1}{2}})^T)^{-1}(C^{\frac{1}{2}})^T \\
((C^{\frac{1}{2}})^T)^{-1} X^T X C^{-\frac{1}{2}} = I \\
(C^{\frac{1}{2}})^T X^T X C^{-\frac{1}{2}} = I \\
(X C^{-\frac{1}{2}})^T X C^{-\frac{1}{2}} = I \\
Y^T Y = I \text{ where } Y = X C^{-\frac{1}{2}}
$$

If $Y = X C^{-\frac{1}{2}}$

$$
\implies X = Y C^{\frac{1}{2}} \\
\implies X^T LX = (Y C^{\frac{1}{2}})^T L Y C^{\frac{1}{2}} \\
= (C^{\frac{1}{2}})^T Y^T L Y C^{\frac{1}{2}} \\
\implies Tr[X^T LX] = Tr[(C^{\frac{1}{2}})^T Y^T L Y C^{\frac{1}{2}}] \\
= Tr[Y^T L Y C]
$$

So the formulation changes to:

$$
\min_{Y \in \mathbb{R}^{n \times m}} \langle Y, L Y C \rangle \\
\text{s.t. } 1^T Y = 0, \\
Y^T Y = I,
$$

$L$ can be decomposed using eigenvalue decomposition into $PDP^T$ where columns of $P$ are eigenvectors of $L$ and $D$ is a diagonal matrix with eigenvalues as the entries:
\[ L = PD^T \]
\[ \Rightarrow Y^TLYC = Y^T(PD^T)YC \]
\[ = (P^TY)^TD(P^TY)C \]

Let \( Z = P^TY \)
\[ \Rightarrow Y = PZ \]
\[ \Rightarrow Y^TLYC = Z^TDZC \]
\[ \Rightarrow Tr[Y^TLYC] = Tr[Z^TDZC] \]
\[ = Tr[ZCZ^TD] \]

For the constraint \( Y^TY = I \)
\[ Y^TY = I \]
\[ \Rightarrow (PZ)^T(PZ) = I \]
\[ \Rightarrow Z^T(P^T P)Z = I \]
\[ \Rightarrow Z^TZ = I \]

For the constraint \( 1^TY = 0 \)
\[ 1^TY = 0 \Rightarrow 1^TPZ = 0 \]

So the formulation becomes:

\[
\min_{Z \in \mathbb{R}^{n \times m}} Tr[ZCZ^TD] \\
\text{s.t. } 1^TPZ = 0, \\
Z^TZ = I, 
\]

This formulation is similar to part (iii). The only difference is that the eigenvalues of \( C \) are multiplied with \( m \) smallest diagonal entries in \( D \) (which are \( m \) smallest eigenvalues of \( L \)). To minimize the objective, we pair eigenvalues of \( C \) and \( L \) such that smallest eigenvalue of \( C \) is multiplied with largest value in \( m \) smallest eigenvalues of \( L \) and vice-versa. Assuming that entries of \( D \) are in ascending order, the solution is given by

\[
Z = \begin{bmatrix}
\leftarrow v_{m}^T \\
\leftarrow v_{m-1}^T \\
\vdots \\
\leftarrow v_1^T \\
\leftarrow 0^T \\
\vdots \\
\leftarrow 0^T \\
\end{bmatrix} \\
Z^T = \begin{bmatrix}
\uparrow v_{m} \\
\uparrow v_1 \\
\downarrow 0 \\
\downarrow 0 \\
\end{bmatrix}
\]

where \( v_i \) is the eigenvector corresponding to eigenvalue \( \lambda_i \) [\( \lambda_m \geq \ldots \geq \lambda_1 \)] of matrix \( C \). So, \( X = YC^{\frac{1}{2}} = PZC^{\frac{1}{2}} \). Note: columns of \( P \) are eigenvectors of \( L \) and \( D \) is a diagonal matrix with eigenvalues as the entries.

For the value of \( Z \) shown above, the minimum value of the objective function is given by:
\[ \text{Tr}[ZCZ^TD] = \sum_{i=0}^{m-1} \left[ \lambda_{m-i} \right] C \left[ v_{m-i}^T v_{m-i} \right] \left( \lambda_{i+1} \right) L \]

\[ = \sum_{i=0}^{m-1} \left[ \lambda_{m-i} \right] C \left[ \lambda_{i+1} \right] L \]

(vi) Let \( X \in \mathbb{R}^{n \times 3} \) represent the coordinates of \( n \) vertices in 3-d. The vertex coordinate assignment can be visualized—a “drawing” of \( G \). Implement your solution to Problem (iii) and show your result for the given graph. We have written a partial framework in Python+CVXPY to get you started:

https://colab.research.google.com/drive/1kSN_oM3gS4nhCjqg2y23OZ0JR_H_uAg?usp=sharing
https://colab.research.google.com/drive/1OTTfZ-vGa0jZTrx7LW13AoCh21zQgiUd?usp=sharing