Numerical Optimization Midterm Exam Solutions

Fan Zhang and Xiangyu Yang

November 5, 2020

1 (8 = 4 + 4 points) Suppose you were restoring a signal from a linear observation operator, i.e., $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\theta}$, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{y} \in \mathbb{R}^m$, and $\boldsymbol{\theta} \in \mathbb{R}^m$ is the random noise.

(1) Construct a linear programming model to find the signal that can best fit the linear model $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$.

(2) Suppose you know that the signal does not have large differences between adjacent elements. Therefore, you also want to add the "total difference of the signal" $\sum_{i=1}^{n-1} |x_{i+1} - x_i|$ to the objective.

Transform this problem to a linear programming model.

Solution:

(1) We could minimize the residual between the true signal and measurements, which essentially makes of the ℓ_1 norm cost criterion. Let $\boldsymbol{e} = (1, \dots, 1)^\top$, we have

$$\begin{aligned} \min_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{r},\boldsymbol{s}\in\mathbb{R}^m}} \quad & \boldsymbol{e}^\top(\boldsymbol{r}+\boldsymbol{s}) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{r} - \boldsymbol{s} \\ & \boldsymbol{r} \geq \boldsymbol{0}, \boldsymbol{s} \geq \boldsymbol{0}. \end{aligned} \tag{0.1}$$

Alternatively, we could also use ℓ_∞ norm to model this problem, and it can be written as follows

$$\begin{aligned} \min_{\substack{u\in\mathbb{R}_+ \\ \boldsymbol{x}\in\mathbb{R}^n}} \quad & u \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} \leq u\boldsymbol{e} \\ & \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} \leq u\boldsymbol{e} \end{aligned} \tag{0.2}$$

(2) $|x_{i+1}-x_i|$ can be transformed to a linear term by using two nonnegative variables $u_i \geq 0, v_i \geq 0$. Specifically, let $x_{i+1}-x_i = u_i - v_i$ and replace every occurrence of $|x_{i+1}-x_i|$ with $u_i + v_i$. We therefore have the following formulation

$$
\begin{aligned}
\min_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{r},\boldsymbol{s}\in\mathbb{R}^m \\ \boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^{n-1}}} \quad & \boldsymbol{e}^\top(\boldsymbol{r}+\boldsymbol{s}) + \sum_{i=1}^{n-1}(u_i+v_i) \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{r} - \boldsymbol{s} \\
& x_{i+1} - x_i = u_i - v_i, \ i = 1, \cdots, n-1 \\
& \boldsymbol{r} \geq \boldsymbol{0}, \boldsymbol{s} \geq \boldsymbol{0}, \boldsymbol{u} \geq \boldsymbol{0}, \boldsymbol{v} \geq \boldsymbol{0}.
\end{aligned}
\tag{0.3}
$$

2 ($8 = 3+2+3$ points) Does set $P = \{\boldsymbol{x} \in \mathbb{R}^2 \mid -1 \le x_1 \le 1\}$ has an extreme point? Write down its standard form. Does this standard form has an extreme point? If it does, then write down an extreme point, and explain why it is an extreme point.

Solution:

(1) No. The polyhedron P contains a line and does not have an extreme point.

(2) The problem transformation involve the following steps:

1) Replace the unrestricted variable x_2 by $x_2^+ - x_2^-$, where $x_2^+ \ge 0$ and $x_2^- \ge 0$.
2) Also note that we can equivalently write $-1 \le x_1 \le 1$ as $|x_1| \le 1$. In this respect, we introduce $x_1^+ \ge 0$ and $x_1^- \ge 0$, and let $x_1 = x_1^+ - x_1^-$, so that we have $|x_1| = x_1^+ + x_1^-$.

Therefore, the standard form of P is

$$P = \{\boldsymbol{x} = (x_1^+, x_1^-, x_2^+, x_2^-, s)^\top \in \mathbb{R}^5 \mid {}_{x_1^+, x_1^-, x_2^+, x_2^-, s \ge 0}^{x_1^+ + x_1^- + s = 1,}\}. \tag{0.4}$$

3) Yes, (0.4) has an extreme point. In particular, $\bar{\boldsymbol{x}} = (1,\ 0,\ 0,\ 0,\ 0)^\top$ is an extreme point of this standard form. This is because

- First of all, $x_1^+ + x_1^- + s = 1$ is active.
- Then, out of the constraints that are active at $\bar{\boldsymbol{x}}$, there are 5 of them that are linearly independent.
- Lastly, $\bar{\boldsymbol{x}} \ge \boldsymbol{0}$ holds.

Therefore, $\bar{\boldsymbol{x}}$ is a basic feasible solution. On the other hand, due to the equivalence of extreme point and basic feasible solution in our context, we have shown that $\bar{\boldsymbol{x}}$ is an extreme point of P.

3 $(4 = 8 \times 0.5$ points) Fill out the following form. Put a "✓" to indicate that this case could happen. Put a "✗" to indicate this case cannot happen. (See the example in the bottom right cell)

Solution:

Primal Problem \ Dual Problem	Infeasible	Unbounded (from above)	Has optimal solution
Infeasible	✓	✓	✗
Unbounded (from below)	✓	✗	✗
Has optimal solutin	✗	✗	✓

4 (6 = 2 + 2 + 2 points) For the standard form

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \ \boldsymbol{c}^\top \boldsymbol{x} \ \text{ s.t. } \ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$$

and consider the following 5 conditions.

$$Ax = b \qquad\qquad x \geq 0$$

$$x^T s = 0$$

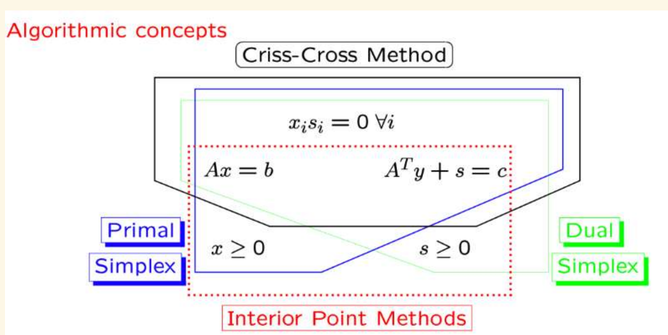
$$A^T y + s = c \qquad\qquad s \geq 0$$

(1) Draw circles on the conditions that the iterates of the primal simplex must satisfy.

(2) Draw squares on the conditions that the iterates of the dual simplex must satisfy.

(3) Draw lines under the conditions that the iterates of the interior point method must satisfy.

Solution:



5 ($12 = 2 + 5 \times 2$ points) For the standard form of linear programming and its dual problem.

(1) Write down what is weak duality and prove.

(2) Write down what is strongly duality and prove.

Solution: Consider the following primal-dual pair with standard form

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \tag{P}$$

$$\begin{aligned} \max \quad & \boldsymbol{y}^\top \boldsymbol{b} \\ \text{s.t.} \quad & \boldsymbol{y}^\top \boldsymbol{A} \leq \boldsymbol{c}^\top \end{aligned} \tag{D}$$

(1) (**Weak duality**) If $\boldsymbol{x}$ is a feasible solution to the primal problem (P) and $\boldsymbol{y}$ is a feasible solution to the dual problem (D), then

$$\boldsymbol{y}^\top \boldsymbol{b} \leq \boldsymbol{c}^\top \boldsymbol{x}. \tag{0.5}$$

Proof. Since $\boldsymbol{x}$ is primal feasible and $\boldsymbol{y}$ is dual feasible, then

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} &\Longrightarrow \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b} \\ \boldsymbol{y}^\top \boldsymbol{A} \leq \boldsymbol{c}^\top, \boldsymbol{x} \geq \boldsymbol{0} &\Longrightarrow \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{c}^\top \boldsymbol{x}. \end{aligned} \tag{0.6}$$

From $\boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b}$ and $\boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{c}^\top \boldsymbol{x}$, we conclude $\boldsymbol{y}^\top \boldsymbol{b} \leq \boldsymbol{c}^\top \boldsymbol{x}$. □

(2) (**Strong duality**) If a linear programming problem has an optimal solution, so does its dual, and the respective optimal costs are equal.

Proof. We prove the strong duality by making use of the simplex tableau. Without loss of generality, we can obtain the primal optimal $\boldsymbol{x}^*$ as well as the optimal basis $\boldsymbol{B}$. Then, we have

$$\begin{aligned} & \boldsymbol{x}_{\boldsymbol{B}}^* = \boldsymbol{B}^{-1}\boldsymbol{b}, \boldsymbol{x}_{\boldsymbol{N}}^* = \boldsymbol{0}, \\ & \boldsymbol{c}_{\boldsymbol{N}}^\top - \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{B}^{-1}\boldsymbol{N} \geq \boldsymbol{0}. \end{aligned}$$

Hence, the dual optimal $\boldsymbol{y}^*$ can be calculated by $(\boldsymbol{y}^*)^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{B}^{-1}$. On the other hand, we have

$$\begin{aligned}
\boldsymbol{y}^\top \boldsymbol{A} &= \begin{bmatrix} \boldsymbol{y}^\top \boldsymbol{B} & \boldsymbol{y}^\top \boldsymbol{N} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{B}^{-1} \boldsymbol{B} & \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{B}^{-1} \boldsymbol{N} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{c}_{\boldsymbol{B}}^\top & \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{B}^{-1} \boldsymbol{N} \end{bmatrix} \\
&\leq \begin{bmatrix} \boldsymbol{c}_{\boldsymbol{B}}^\top & \boldsymbol{c}_{\boldsymbol{N}}^\top \end{bmatrix} \\
&= \boldsymbol{c}^\top
\end{aligned} \tag{0.7}$$

This verifies the dual feasibility of $\boldsymbol{y}^*$. Furthermore, we have

$$(\boldsymbol{y}^*)^T \boldsymbol{b} = \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{A}_{\boldsymbol{B}}^{-1} \boldsymbol{b} = \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{x}_{\boldsymbol{B}}^* + \boldsymbol{c}_{\boldsymbol{B}}^\top \boldsymbol{x}_{\boldsymbol{N}}^* = \boldsymbol{c}^\top \boldsymbol{x}^*.$$

This completes the proof. $\square$

6 (35 = 5 × 7 points) Consider the linear optimization problem

$$\begin{aligned} \min \quad & -4x_1 - x_2 - 3x_3 \\ \text{s.t.} \quad & 2x_1 + 2x_2 + x_3 = 4 \\ & x_1 + 2x_2 = 2 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{aligned} \tag{0.8}$$

(1) Write down its first-phase problem, and use it to find a basic feasible solution.

(2) Turn to the second phase and use primal-simplex method to find the optimal solution.

(3) Define the dual variables and write down the Lagrangian function.

(4) Write down the dual problem. Point out which constraints in the dual problem are active and inactive at the optimal dual solution.

(5) What is $\boldsymbol{B}^{-1}$ in your final tableau?

(6) If the right-hand-side 2 is changed to $2 + \delta$, determine an interval for δ, so that the optimal basis are still the same.

(7) If the second constraint is removed from this problem, is the optimal solution you have found still optimal? If not, find the optimal solution in this new case.

Solution:

(1) To find a feasible solution to (0.8), we form the auxiliary problem as follows

$$\begin{aligned} \min \quad & y_1 + y_2 \\ \text{s.t.} \quad & 2x_1 + 2x_2 + x_3 + y_1 = 4 \\ & x_1 + 2x_2 + y_2 = 2 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ y_1 \geq 0,\ y_2 \geq 0 \end{aligned} \tag{0.9}$$

A basic feasible solution to (0.9) is therefore obtained by letting $(y_1, y_2) = (4, 2)$.

(2) • We first form the initial tableau of Phase I

Table 1: The first tableau.

	x_1	x_2	x_3	y_1	y_2	
	2	2	1	1	0	4
	1	2	0	0	1	2
$\boldsymbol{r}^T$	-3	-4	-1	0	0	-6

Table 2: The second tableau.

	x_1	x_2	x_3	y_1	y_2	
	2	2	1	1	0	4
	1	2	0	0	1	2
$\boldsymbol{r}^T$	-1	-2	0	1	0	-2

- We bring x_3 into the basis and have y_1 exit the basis. The basis matrix $\boldsymbol{B}$ is still the identity and only the last row of the tableau changes. We obtain
- We now bring x_2 into the basis and have y_2 exit the basis. The new tableau is

Table 3: The third tableau.

	x_1	x_2	x_3	y_1	y_2	
	1	0	1	1	-1	2
	0.5	1	0	0	0.5	1
$\boldsymbol{r}^T$	0	0	0	1	1	0

Note that the cost in (0.9) has dropped to zero, indicating that we have a feasible solution to the original problem (0.8).

We now start executing the primal simplex method on the original problem. We bring x_1 into the basis and have x_3 exit the basis. The

Table 4: The first tableau of (0.8)

	x_1	x_2	x_3	
	1	0	1	2
	0.5	1	0	1
$\boldsymbol{r}^T$	-0.5	0	0	7

basis matrix $\boldsymbol{B}$ is still the identity. We obtain

After two pivots, we observe that the all reduced costs are nonnegative.

Table 5: The last tableau of (0.8).

	x_1	x_2	x_3	
	1	0	1	2
	0	1	-0.5	0
$\boldsymbol{r}^T$	0	0	0.5	8

This indicates that the optimality is satisfied. Hence, $(x_1, x_2, x_3)^* = (2, 0, 0)$ is the optimal solution to (0.8), and moreover, we have $z^* = -8$.

(3) Let $\boldsymbol{\lambda} \in \mathbb{R}^2$ and $\boldsymbol{\mu} \in \mathbb{R}^3_+$ be the dual variables associated with those two inequality constraints of (0.8), respectively. Define the Lagrangian $\mathcal{L} : \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}$ associated with the problem (0.8) as

$$\begin{aligned}\mathcal{L}(\boldsymbol{x}; \boldsymbol{\lambda}, \boldsymbol{\mu}) := & (-2\lambda_1 - \lambda_2 - 4 - \mu_1)x_1 + (-2\lambda_1 - 2\lambda_2 - 1 - \mu_2)x_2 \\ & + (-\lambda_1 - 3 - \mu_3)x_3 + 4\lambda_1 + 2\lambda_2.\end{aligned} \tag{0.10}$$

(4) From Lagrangian, the dual objective reads

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}; \boldsymbol{\lambda}, \boldsymbol{\mu}). \tag{0.11}$$

Since we merely have interests in the case that $g(\boldsymbol{\lambda}, \boldsymbol{\mu}) > -\infty$, this implies all the coefficients in front of each primal variable, i.e., $x_i,\ i = 1, 2, 3$, should be set as 0. Hence, the dual problem is written as follows

$$\begin{aligned} \max \quad & 4\lambda_1 + 2\lambda_2 \\ \text{s.t.} \quad & 2\lambda_1 + \lambda_2 \leq -4 \\ & 2\lambda_1 + 2\lambda_2 \leq -1 \\ & \lambda_1 \leq -3, \lambda_2 \text{ free }. \end{aligned} \tag{0.12}$$

It should be noticed that μ_1, μ_2 and μ_3 are eliminated in the dual problem since they are all nonnegative.

For checking the active constraints, we calculate the optimal dual variables by $(\lambda^*)^T = \boldsymbol{c_B B}^{-1} = [-4, -1] \begin{bmatrix} 1 & -1 \\ -0.5 & 1 \end{bmatrix} = (-3.5, 3)$. By checking all constraints by λ^*, the first two constraints are active and the last constraint is inactive.

(5) Since the optimal basic variables are (x_1, x_2), the corresponding basis is $\boldsymbol{B} = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}$, yielding $\boldsymbol{B}^{-1} = \begin{bmatrix} 1 & -1 \\ -0.5 & 1 \end{bmatrix}$.

(6) Consider the optimal tableau table 5. We look at the first column of $\boldsymbol{B}^{-1}$ which is $(-1, 1)^\top$. The feasibility condition is then $(2, 0)^T + \delta(-1, 1)^\top \geq \mathbf{0}$, yielding $0 \leq \delta \leq 2$.

(7) No. $\boldsymbol{x} = (0, 0, 3)^T$ is the optimal solution.

7 ($15 = 4+4+3+4$ points) A large manufacturing company produces liquid nitrogen in 5 plants spread out in Jiangsu Province. Each plant has monthly production capacity.

Plant i	1	2	3	4	5
Capacity p_i	120	95	150	120	140

It has 7 retailers in the same area. Each retailer has a monthly demand to be satisfied.

Retailer j	1	2	3	4	5	6	7
Demand d_j	55	72	80	110	85	30	78

Transportation between any plant i and any retailer j has a cost of c_{ij} dollars per volume unit of nitrogen.

(1) Build a model to decide each retailer should be served by which plant. Your objective is to minimize the total transportation cost.

(2) Since the monthly rental fee is rising up for each plant. The company is considering to shut down some of the plant while still keep all the retailers' demands are satisfied. Suppose the rental fee for each plant is now r_i, build a model to minimize the total cost.

(3) Suppose you want to use the Lagrangian relaxation method. Which constraint(s) do you want to relax? Give your reason(s) for your choice.

(4) Show that for any multipliers in your relaxation, the optimal value of your relaxed problem is always a lower bound for the optimal value of the original problem.

Solution:

(1) Define the following notation:
Sets
$I = \{i \mid 1, 2, \dots, 5\}$ (set of plants)
$J = \{j \mid 1, 2, \dots, 7\}$ (set of retailers)
Parameters
p_i = capcity of plant $i \in I$
d_j = each unit demand of retailer $j \in J$
c_{ij} = cost to transport one volume unit of nitrogen from plant $i \in I$ to retailer $j \in J$

Decision Variables
x_{ij} = the fraction of nitrogen's quantity from plant $i \in I$ to retailer $j \in J$

Therefore, our problem is formulated as follows

$$\begin{aligned}
\min \quad & \sum_{i=1}^{5}\sum_{j=1}^{7} d_j x_{ij} c_{ij} & \\
\text{s.t.} \quad & \sum_{i=1}^{5} x_{ij} = 1, & \forall j \in J \\
& \sum_{j=1}^{7} d_j x_{ij} \le p_i, & \forall i \in I \\
& x_{ij} \ge 0, & \forall i \in I, \forall j \in J
\end{aligned} \tag{0.13}$$

(2) **Decision Variables**
$y_i = 1$ if plant i is opened, 0 otherwise
x_{ij} = the fraction of nitrogen's quantity from plant $i \in I$ to retailer $j \in J$

Therefore, our problem is formulated as follows

$$\begin{aligned}
\min \quad & \sum_{i=1}^{5} r_i y_i + \sum_{i=1}^{5}\sum_{j=1}^{7} d_j x_{ij} c_{ij} & \\
\text{s.t.} \quad & \sum_{i=1}^{5} x_{ij} = 1, & \forall j \in J \\
& \sum_{j=1}^{7} d_j x_{ij} \le p_i, & \forall i \in I \\
& y_i \in (0,1), & \forall i \in I \\
& x_{ij} \le y_i, & \forall i \in I, \forall j \in J \\
& x_{ij} \ge 0, & \forall i \in I, \forall j \in J
\end{aligned} \tag{0.14}$$

(3) Relax the constraint $\sum_{i=1}^{5} x_{ij} = 1, \ \forall j \in J$.

1. Introduce fewer Lagrangian multipliers so that the relaxed problem is a tighter lower bound of the original one, and moreover, it

generally makes it easier to find good multipliers using subgradient optimization.

2. As for the relaxation of $\sum_{j=1}^{7} d_j x_{ij} \leq p_i$, $\forall i \in I$ or $x_{ij} \leq y_i$, $\forall i \in I, \forall j \in J$, it can be easily observed that x_{ij} will be 0 for many i that are open, there will be many constraints such that $\sum_{j=1}^{7} d_j x_{ij} \leq p_i$ and $x_{ij} \leq y_i$. It is often difficult to get good results when relaxing inequality constraints many of which are slack.

(4) For ease of presentation, we use LR_λ to denote the Lagrangian subproblem that is obtaine by relaxing constraint $\sum_{i=1}^{5} x_{ij} = 1\ \forall j \in J$.

Proof. Let (x, y) be a feasible solution to (0.14). It is obvious to see that (x, y) is also feasible for LR_λ, and it has the same objective value in both problems since $\sum_{j=1}^{7} \lambda_i (1 - \sum_{i=1}^{5} x_{ij}) = 0$. Hence, the optimal objective value for LR_λ is no greater than that of (0.14). $\square$

8 (12 = 5 + 2 + 5 points) Suppose in the optimality condition for standard form you relax the complementarity condition to have $x_i s_i = \tau > 0$. Suppose your interior point method has the iteration:

$$(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1}, \boldsymbol{s}^{k+1}) \leftarrow (\boldsymbol{x}^{k}, \boldsymbol{y}^{k}, \boldsymbol{s}^{k}) + (\Delta\boldsymbol{x}, \Delta\boldsymbol{y}, \Delta\boldsymbol{s}).$$

(1) For any $(\boldsymbol{x}^k, \boldsymbol{y}^k, \boldsymbol{s}^k)$ with $\boldsymbol{x}^k \geq \boldsymbol{0}$, $\boldsymbol{s}^k \geq \boldsymbol{0}$, derive the system of linear equations that is used for computing $(\Delta\boldsymbol{x}, \Delta\boldsymbol{y}, \Delta\boldsymbol{s})$.

(2) For any $(\boldsymbol{x}^k, \boldsymbol{y}^k, \boldsymbol{s}^k)$ with $\boldsymbol{x}^k \geq \boldsymbol{0}$, $\boldsymbol{s}^k \geq \boldsymbol{0}$, show the linear system matrix you have in (1) is nonsigular.

(3) Suppose for any symmetric matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$, you have a very eficient subroutine to compute its Jordan Canonical form decomposition: $\boldsymbol{Q} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top$, where $\boldsymbol{P}^\top = \boldsymbol{P}^{-1}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix consisting of the eigenvalues of $\boldsymbol{Q}$. Now use this decomposition subroutine to design an eficient algorithm for solving the linear equations you have in (1).

Solution: We consider the linear programming problem in standard form; that is

$$\begin{array}{ll} \min & \boldsymbol{c}^\top\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{array} \tag{0.15}$$

where $\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, and $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $\text{rank}(\boldsymbol{A}) = m$. The dual problem for (0.15) is

$$\begin{array}{ll} \min & \boldsymbol{b}^\top\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{A}^\top\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ & \boldsymbol{s} \geq \boldsymbol{0}, \end{array} \tag{0.16}$$

(1) The optimality conditions for the primal-dual pair is a mapping $F : \mathbb{R}^{2n+m} \to \mathbb{R}^{2n+m}$:

$$F(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = \begin{bmatrix} \boldsymbol{A}^\top\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{c} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} \end{bmatrix} = \boldsymbol{0}, (\boldsymbol{x}, \boldsymbol{s} \geq \boldsymbol{0}) \tag{0.17}$$

where $\boldsymbol{X}, \boldsymbol{S} \in \mathbb{R}^{n\times n}$ are diagonal matrices with diagonal $\boldsymbol{x}$ and $\boldsymbol{s}$, respectively, and $\boldsymbol{e} \in \mathbb{R}^n$ is a vector of ones. Since the complementary condition is relaxed, we can apply Newton's method to solve the following equation for a fixed relaxation parameter $\tau > 0$; that is

$$F_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = \begin{bmatrix} \boldsymbol{A}^\top\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{c} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \\ \boldsymbol{X}\boldsymbol{s} - \tau\boldsymbol{e} \end{bmatrix} = \boldsymbol{0}, (\boldsymbol{x}, \boldsymbol{s} \geq \boldsymbol{0}) \tag{0.18}$$

At the kth iterate, the Newton direction $(\Delta \boldsymbol{x}, \Delta \boldsymbol{y}, \Delta \boldsymbol{s})$ is computed via solving the following linear system:

$$\begin{bmatrix} \mathbf{0}_{n\times n} & \boldsymbol{A}_{n\times m}^\top & \boldsymbol{I}_{n\times n} \\ \boldsymbol{A}_{m\times n} & \mathbf{0}_{m\times m} & \mathbf{0}_{m\times n} \\ \boldsymbol{S}_{n\times n}^k & \mathbf{0}_{n\times m} & \boldsymbol{X}_{n\times n}^k \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \\ \Delta \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \tau \boldsymbol{e} - \boldsymbol{X}^k \boldsymbol{s}^k \end{bmatrix} \tag{0.19}$$

(2) *Proof.* By performing a block Gaussian elimination, we have

$$\begin{aligned} \begin{bmatrix} \mathbf{0} & \boldsymbol{A}^\top & \boldsymbol{I} \\ \boldsymbol{A} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{S}^k & \mathbf{0} & \boldsymbol{X}^k \end{bmatrix} & \longrightarrow \begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{X}^k & \mathbf{0} & \boldsymbol{S}^k \end{bmatrix} \\ \longrightarrow \begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}^\top & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{X}^k\boldsymbol{A}^\top & \boldsymbol{S}^k \\ \mathbf{0} & \mathbf{0} & \boldsymbol{A} \end{bmatrix} & \longrightarrow \begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}^\top & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{X}^k\boldsymbol{A}^\top & \boldsymbol{S}^k \\ \mathbf{0} & \boldsymbol{A}(\boldsymbol{S}^k)^{-1}\boldsymbol{X}^k\boldsymbol{A}^\top & \mathbf{0} \end{bmatrix} \\ \longrightarrow \begin{bmatrix} \boldsymbol{I} & \mathbf{0} & \boldsymbol{A}^\top \\ \mathbf{0} & \boldsymbol{S}^k & -\boldsymbol{X}^k\boldsymbol{A}^\top \\ \mathbf{0} & \mathbf{0} & \boldsymbol{A}(\boldsymbol{S}^k)^{-1}\boldsymbol{X}^k\boldsymbol{A}^\top \end{bmatrix}, & \end{aligned}$$

which is a block diagonal matrix with full rank diagonals. This completes the proof.

□

(3) The above system (0.19) is of size $2n + m$, we can symmetrize the Jacobian by multiplying the last block row from the left by $(\boldsymbol{S}^k)^{-1}$, yielding

$$\begin{bmatrix} \mathbf{0}_{n\times n} & \boldsymbol{A}_{n\times m}^\top & \boldsymbol{I}_{n\times n} \\ \boldsymbol{A}_{m\times n} & \mathbf{0}_{m\times m} & \mathbf{0}_{m\times n} \\ \boldsymbol{I}_{n\times n} & \mathbf{0}_{n\times m} & \boldsymbol{D}_{n\times n}^k \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \\ \Delta \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \tau(\boldsymbol{s}^k)^{-1} - \boldsymbol{X}^k \boldsymbol{e}^k \end{bmatrix},$$

where $\boldsymbol{D}^k = (\boldsymbol{S}^k)^{-1}\boldsymbol{X}^k$. Now, the Jordan Canonical form decomposition can be applied to solve the linear equations (0.19).