

Midterm

MS&E 310

Course Instructor : Yinyu Ye

Question 1 (30pts)

This problem is concerned with using an optimality criterion for linear programming to decide whether the vector $\mathbf{x}^* = (0, 2, 0, 7, 0)$ is optimal for the linear program:

$$(P) \max \quad 8x_1 - 9x_2 + 12x_3 + 4x_4 + 11x_5,$$

$$\text{subject to} \quad 2x_1 - 3x_2 + 4x_3 + x_4 + 3x_5 \leq 1,$$

$$x_1 + 7x_2 + 3x_3 - 2x_4 + x_5 \leq 1$$

$$5x_1 + 4x_2 - 6x_3 + 2x_4 + 3x_5 \leq 22,$$

$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

(6 pts each)

- Write down the dual (D) of the problem (P)
- Noting that the constraints of (P) have the form $\mathbf{Ax} \leq \mathbf{b}$, $\mathbf{x} \geq 0$, write down the support of the vectors \mathbf{x}^* and $\mathbf{b} - \mathbf{Ax}^*$.
- For the moment, assume that \mathbf{x}^* is an optimal solution to (P). Write down a system of equations in the dual variables that must hold as a consequence of this assumption.
- Solve the system in part (c).

Question 2 (25 pts)

In this question, we will develop linear programming approximation for an integer optimization problem. Consider a function

$$\phi(\mathbf{X}) = \boldsymbol{\alpha}^T \mathbf{X} + \mathbf{X}^T \mathbf{W} \mathbf{X} = \sum_i \alpha_i X_i + \sum_{i \neq j} w_{ij} X_i X_j,$$

where $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a binary vector ($X_i \in \{0, 1\}$) and $\mathbf{W}_{ii} = 0$. Our goal is to

$$\begin{aligned} \text{(D-OPT)} \quad & \max_{\mathbf{X}} \phi(\mathbf{X}) \\ & \text{subject to } X_i \in \{0, 1\} \end{aligned}$$

This is a simplified version of the Ising model's potential function widely used statistical physics and computer vision. You may think of the function value $\phi(\mathbf{X})$ proportional to the log likelihood of a configuration of atomic spinings or image pixel values. To maximize the objective function is actually the procedure to find the most probable configuration. Since this integer programming problem is pretty hard to solve, people always try to find an LP approximation.

- (a) (15 pts) Formulate an LP approximation for the above D-OPT problem by using the technique of relaxation. (Hint. Consider relaxing $X_i \in \{0, 1\}$ into $X_i \in [0, 1]$. And also you should notice that there are quadratic terms that need to be removed. You may want to introduce $Y_{ij} = X_i X_j$ and investigate the relation between X_i, X_j and Y_{ij} .)
- (b) (10 pts) What is the relation of the optimal values of the D-OPT, your LP, and its dual? We don't require you to write out the dual explicitly here but you may do that if you are interested.

Question 3 (25 pts)

Quantile regression is a type of regression model in statistics and econometrics. While the least square regression models the expectation of target variable Y , the quantile regression models its quantile. In this question, we will investigate the computational aspect of the quantile regression. The model assumes the τ -quantile of the target random variable $Y \in R$ is a linear function of the response variable

$$Q_Y(\tau) = \boldsymbol{\beta}^T \mathbf{x},$$

where $\tau \in (0, 1)$, $\boldsymbol{\beta} \in R^p$ and $\mathbf{x} \in R^p$. Given the response data matrix $\mathbf{X} \in R^{n \times p}$ ($n > p$) and the target variable $\mathbf{Y} \in R^n$, the way we estimate the parameter $\boldsymbol{\beta}$ is to minimize the following objective function:

$$\min_{\beta} \sum_{i=1}^n l_{\tau}(Y_i - \beta^T \mathbf{X}_{(i)}).$$

$\mathbf{X}_{(i)}$ denotes the i -th row of \mathbf{X} . And the function $l_{\tau}(\cdot)$ is defined by

$$l_{\tau}(t) = \begin{cases} \tau t, & t \geq 0, \\ (\tau - 1)t, & t < 0. \end{cases}$$

The quantile τ is determined based on the interest of the concrete problems. For example, we set $\tau = 0.99$ to estimate water level when we plan to build a dam that could prevent flood once in a hundred years. The minimizer of the above objective function would be our estimate of β .

- (a) (15 pts) Formulate the optimization problem into a linear programming.
- (b) (10 pts) Show that there exists an optimal β^* with a line passing at least p data points. In other words, $l_{\tau}(Y_i - \beta^{*T} \mathbf{X}_{(i)}) = 0$ for at least p different indices.

Question 4 (20 pts)

In this question, we consider the problem of robust linear programming problem.

$$\begin{aligned} \min_x \sup_{c \in \mathcal{C}} c^T x, \\ \text{subject to } Ax \geq b, \end{aligned}$$

with variable $x \in R^n$ and the set $\mathcal{C} = \{c \mid Fc \leq g, Ec = h\}$. We use the set \mathcal{C} to characterize the uncertainty for the parameters in the real world settings due to error of measurement and randomness. The objective function is to minimize the superior value over all the choice of c in \mathcal{C} , that is, to obtain a worst case optimality.

- (a) (10 pts) Find the dual of the problem

$$\begin{aligned} \max_c c^T x \\ \text{subject to } Fc \leq g, Ec = h \end{aligned}$$

with decision variable c . (You can regard x as a known vector here.)

- (b) (10 pts) Use the result above to obtain a single LP formulation equivalent to the original robust LP.

Bonus (10 pts)

During the the simplex method, if the current basic feasible solution is non-degenerate and there is only one negative reduced cost coefficient, then the entering variable will remain as a basic variable for the rest of the simplex method procedures.