Problem set # 1

Numerical Methods for Data Science 2019/20

UC3M — *Master on Statistics for Data Science*

Due date: September 30. Value: 50% of the final grade.

Note: This is an individual assignment. Evidence of plagiarism will be penalized. Hand in the assignment in paper or as a pdf file, with Gurobi-Python code printouts, but no electronic files.

Problem 1 (35 points). Consider the linear optimization problem

$$\begin{aligned}
&\text{maximize } 2x_1 + 3x_2 - 9x_3 \\
&\text{subject to:} \\
&5x_1 + x_2 + 5x_3 = 3 \\
&x_1 + 2x_2 - x_3 = 6 \\
&x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0.
\end{aligned}$$

(a, 7 points) Formulate the dual problem, and find all its solutions using the graphical method.

(b, 7 points) Formulate the optimality conditions that must be satisfied by any optimal primal solution in relation with a dual optimal solution $\boldsymbol{\pi}^*$. Apply them, along with part (a), to find all solutions of the primal problem.

(c, 7 points) Obtain all possible values for the reduced cost of each primal variable. Are reduced costs unique? Interpret the reduced costs obtained.

(d, 7 points) Carry out a sensitivity analysis with respect to simultaneous changes of contraint right-hand sides for the primal problem. Contrast the results with those obtained through Gurobi-Python.

(e, 7 points) Carry out a sensitivity analysis with respect to simultaneous changes of objective coefficients for the primal problem. Contrast the results with those obtained through Gurobi-Python.

Answers.

(a) The dual problem is:

$$\begin{aligned}
&\text{minimize } 3\pi_1 + 6\pi_2 \\
&\text{subject to:} \\
x_1:\ &5\pi_1 + \pi_2 \geqslant 2 \\
x_2:\ &\pi_1 + 2\pi_2 \geqslant 3 \\
x_3:\ &5\pi_1 - \pi_2 \geqslant -9.
\end{aligned}$$

See Figure 1. All feasible solutions in the feasible's region edge $\pi_1 + 2\pi_2 = 3$, which is a semi-infinite line segment with vertex $(1/9, 13/9)$, are optimal. Such a line segment has parametric equations

$$(\pi_1^*, \pi_2^*) = (1/9, 13/9) + \lambda(2, -1), \quad \lambda \geqslant 0.$$

(b) The optimality conditions that must be satisfied by any primal optimal solution $\mathbf{x}^*$ in relation with a dual optimal solution $\boldsymbol{\pi}^*$ are:

(OC1) : $\mathbf{x}^*$ is primal feasible.

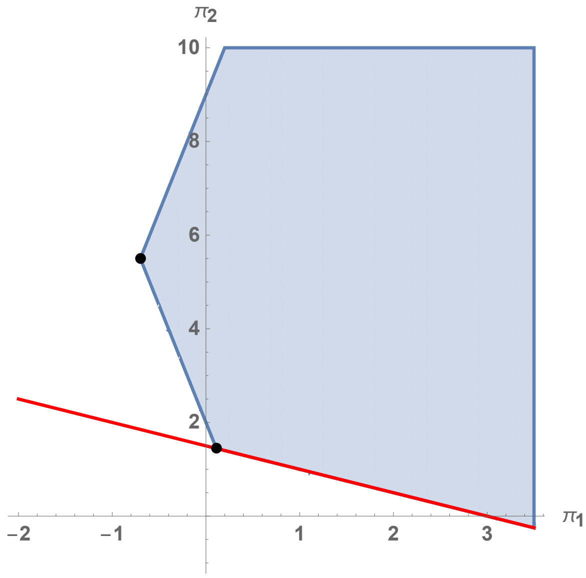

Figure 1: Graphical solution of Problem 1(a).

(OC3') : Complementary slackness:

$$5\pi_1^* - \pi_2^* > -9 \text{ (note: this always holds)} \Longrightarrow x_3^* = 0$$

$$5\pi_1^* + \pi_2^* > 2 \text{ (note: this holds except for } \boldsymbol{\pi}^* = (1/9, 13/9)) \Longrightarrow x_1^* = 0$$

If $5\pi_1^* + \pi_2^* > 2$, we have $\mathbf{x}^* = (0, 3, 0)$. If $5\pi_1^* + \pi_2^* = 2$, i.e., $\boldsymbol{\pi}^* = (1/9, 13/9)$, we have

$$\begin{aligned}
&5x_1 + x_2 = 3\\
&x_1 + 2x_2 = 6\\
&x_1 \geqslant 0, x_2 \geqslant 0,
\end{aligned}$$

whose unique solution is, again, $\mathbf{x}^* = (0, 3, 0)$.

(c) The reduced costs of the primal variables are

$$\begin{aligned}
\bar{c}_1 &= 5\pi_1^* + \pi_2^* - 2 = 9\lambda\\
\bar{c}_2 &= \pi_1^* + 2\pi_2^* - 3 = 0\\
\bar{c}_3 &= 5\pi_1^* - \pi_2^* + 9 = 73/9 + 11\lambda,
\end{aligned}$$

where $\boldsymbol{\pi}^*$ is a dual optimal solution, i.e., for any $\lambda \geqslant 0$ (see the solution of part (a)). Therefore, reduced costs are not unique.

(d) Carry out a sensitivity analysis with respect to simultaneous changes of contraint right-hand sides for the primal problem.

$$\begin{aligned}
&\text{maximize}\, 2x_1 + 3x_2 - 9x_3\\
&\text{subject to:}\\
&5x_1 + x_2 + 5x_3 = 3 + \Delta b_1\\
&x_1 + 2x_2 - x_3 = 6 + \Delta b_2\\
&x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0.
\end{aligned}$$

Consider $\boldsymbol{\pi}^* = (1/9, 13/9)$. We have $\hat{x}_3^* = 0$ and

$$\begin{aligned}
5x_1 + x_2 &= 3 + \Delta b_1\\
x_1 + 2x_2 &= 6 + \Delta b_2,
\end{aligned}$$

so

$$\hat{\mathbf{x}}^* = (\frac{1}{9}\left(2\Delta b_1 - \Delta b_2\right), \frac{1}{9}\left(-\Delta b_1 + 5\Delta b_2 + 27\right), 0).$$

Imposing primal feasibility gives

$$\begin{aligned} 2\Delta b_1 - \Delta b_2 &\geqslant 0 \\ -\Delta b_1 + 5\Delta b_2 &\geqslant -27. \end{aligned}$$

Further, if $\Delta b_2 = 0$ then $0 \leqslant \Delta b_1 \leqslant 27$. And, if $\Delta b_1 = 0$ then $-27/5 \leqslant \Delta b_2 \leqslant 0$.

(e) The primal problem has a single feasible solution, therefore the optimal solution $\mathbf{x}^*$ will remain optimal for the modified problem

$$\begin{aligned} &\text{maximize}\,(2+\Delta r_1)x_1 + (3+\Delta r_2)x_2 - (9+\Delta r_3)x_3 \\ &\text{subject to:} \\ &5x_1 + x_2 + 5x_3 = 3 \\ &x_1 + 2x_2 - x_3 = 6 \\ &x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, \end{aligned}$$

regardless of the values of the increments Δr_j.

Such a result can also be recovered by a standard analysis, as follows. Consider $\mathbf{x}^* = (0,3,0)$. By CS we can consider the system of equations

$$\begin{aligned} 5\pi_1 + \pi_2 &= 2 + \Delta r_1 \\ \pi_1 + 2\pi_2 &= 3 + \Delta r_2, \end{aligned}$$

which gives $\hat{\boldsymbol{\pi}}^* = \frac{1}{9}(1 + 2\Delta r_1 - \Delta r_2, 13 - \Delta r_1 + 5\Delta r_2)$. Dual feasibility means that

$$5\hat{\pi}_1^* - \hat{\pi}_2^* + 9 - \Delta r_3 = \frac{1}{9}\left(73 + 11\Delta r_1 - 10\Delta r_2 - 9\Delta r_3\right) \geqslant 0,$$

i.e.,

$$11\Delta r_1 - 10\Delta r_2 - 9\Delta r_3 \geqslant -73.$$

So, if $\Delta r_2 = \Delta r_3 = 0$, then $\Delta r_1 \geqslant -73/11$.

And, if $\Delta r_1 = \Delta r_3 = 0$, then $\Delta r_2 \leqslant 73/10$.

And, if $\Delta r_1 = \Delta r_2 = 0$, then $\Delta r_3 \leqslant 73/9$.

We can also consider alternatively the the system of equations

$$\begin{aligned} \pi_1 + 2\pi_2 &= 3 + \Delta r_2 \\ 5\hat{\pi}_1 - \hat{\pi}_2 &= -9 + \Delta r_3, \end{aligned}$$

which gives $\hat{\boldsymbol{\pi}}^* = \frac{1}{11}(-15 + \Delta r_2 + 2\Delta r_3, 24 + 5\Delta r_2 - \Delta r_3)$. Dual feasibility means that

$$5\hat{\pi}_1^* + \hat{\pi}_2^* - 2 - \Delta r_1 = \frac{1}{11}\left(-73 - 11\Delta r_1 + 10\Delta r_2 + 9\Delta r_3\right) \geqslant 0,$$

i.e.,

$$-11\Delta r_1 + 10\Delta r_2 + 9\Delta r_3 \geqslant 73.$$

So, if $\Delta r_2 = \Delta r_3 = 0$, then $\Delta r_1 \leqslant -73/11$.

And, if $\Delta r_1 = \Delta r_3 = 0$, then $\Delta r_2 \geqslant 73/10$.

And, if $\Delta r_1 = \Delta r_2 = 0$, then $\Delta r_3 \geqslant 73/9$.

Combining both cases we obtain that $\mathbf{x}^*$ is optimal for any $\Delta r_j \in \mathbb{R}$.

Problem 2 (35 points). In a CSI investigation, the crime suspect left his/her shoe imprints in the crime scene. From that evidence the investigators want to infer the suspect's height. For that purpose, they plan to obtain a prediction equation for height based on shoe size based on the following data:

Shoe size (cm)	Height (cm)
29.7	175.3
29.7	177.8
31.4	185.4
31.8	175.3
27.6	172.7

(a, 20 points) Formulate the Linear Optimization moof seen in class for estimating the best prediction equation under the Mean Absolute Error (MAE) criterion, and implement it in Gurobi-Python.

(b, 7 points) Solve the model and give the optimal solution (prediction equation). Is it unique?

(c, 8 points) Obtain the optimal dual solution and discuss its possible interpretation.

Answers.

(a) The primal formulation is, as seen in class,

$$
\begin{aligned}
(P) \qquad & \text{minimize } z = \sum_{i=1}^{n} (e_i^+ + e_i^-) \\
& \text{subject to} \\
\pi_i\colon\ & e_i^+ - e_i^- + b_0 + x_i b_1 = y_i, \quad i = 1, \ldots, n \\
& e_i^+, e_i^- \geqslant 0, \quad i = 1, \ldots, n
\end{aligned}
$$

(b) The optimal solution (prediction equation) obtained is $\hat{y} = 138.53 + 1.24x$. In general, an optimal prediction equation under the MAE criterion need not be unique. In this case, the variables that must take the value zero in any optimal solution because they have positive reduced costs are:

$$e_1^+ = e_4^+ = e_5^+ = e_1^- = e_2^- = e_3^- = e_5^- = 0.$$

We thus have that the following system of equations characterizes the optimal solutions:

$$
\begin{aligned}
b_0 + x_1 b_1 &= y_1 \\
e_2^+ + b_0 + x_2 b_1 &= y_2 \\
e_3^+ + b_0 + x_3 b_1 &= y_3 \\
-e_4^- + b_0 + x_4 b_1 &= y_4 \\
b_0 + x_5 b_1 &= y_5.
\end{aligned}
$$

Since this system of equations has a unique solution, there is a unique optimal solution.

(c) The dual problem is

$$
\begin{aligned}
(D) \qquad & \text{maximize } d = \sum_{i=1}^{n} y_i \pi_i \\
& \text{subject to} \\
b_0\colon\ & \sum_{i=1}^{n} \pi_i = 0 \\
b_1\colon\ & \sum_{i=1}^{n} x_i \pi_i = 0 \\
e_i^+\colon\ & \pi_i \leqslant 1, \quad i = 1, \ldots, n \\
e_i^-\colon\ & -\pi_i \leqslant 1, \quad i = 1, \ldots, n
\end{aligned}
$$

By strong duality, writing the MAE as $z^*(\mathbf{y})$, we have

$$\pi_i^* = \frac{\partial}{\partial y_i} z^*(\mathbf{y}),$$

whenever the derivative exists. Furthermore, $\pi_i^* = 1$ for observations where the prediction equation underestimates the response ($e_i^+ > 0$), $\pi_i^* = -1$ for observations where the prediction equation overestimates the response ($e_i^- > 0$).

The optimal dual solution obtained is:

$$\pi_1^* = -0.81, \pi_2^* = 1, \pi_3^* = 1, \pi_4^* = -1, \pi_5^* = -0.19.$$

Problem 3 (30 points). Consider the following optimal production planning model,

$$\begin{aligned}
&\text{minimize} \sum_{t=1}^{L} \sum_{i=1}^{N} (p_{it} x_{it} + h_{it} I_{it}) \\
&\text{subject to:} \\
&x_{it} + I_{i,t-1} - I_{it} = D_{it}, \quad i = 1, \ldots, N,\ t = 1, \ldots, L \\
&\sum_{i=1}^{N} m_i x_{it} \leqslant K_t, \quad t = 1, \ldots, L \\
&\sum_{i=1}^{N} I_{it} \leqslant \text{INV}_t, \quad t = 1, \ldots, L \\
&x_{it}, I_{it} \geqslant 0, \quad i = 1, \ldots, N,\ t = 1, \ldots, L,
\end{aligned}$$

where the following notation is used (note that the term “resource” refers to a generic productive resource, and costs are in euros):

- N: number of different products.
- L: number of planning periods.
- x_{it}: quantity made of product i in period t.
- I_{it}: inventory of product i at the end of period t.
- D_{it}: demand of product i in period t.
- p_{it}: production cost per unit of product i in period t.
- h_{it}: inventory holding cost per unit of product i in period t.
- m_i: resource consumption per unit of product i.
- K_t: available resource capacity in period t.
- INV_t: inventory storage capacity in period t.

The goal is to determine the optimal production quantities x_{it} and the inventory levels I_{it} in each planning period. The optimal plan minimizes the production and inventory holding costs, while respectinv production and inventory capacities.

Consider an instance with $N = L = 2$, $m_1 = m_2 = 1$, and the following additional data:

Parameters		Period 1	Period 2
Demand (D_{it})	Prod 1	0	200
	Prod 2	100	100
Production cost (p_{it})	Prod 1	10	25
	Prod 2	10	20
Inventory holding cost (h_{it})	Prod 1	5	5
	Prod 2	5	5
Production capacity (K_t)		300	200
Inventory (storage) capacity (INV_t)		200	200

(a, 12 puntos) Implement the model in Gurobi-Python, solve it, and interpret the optimal solution. Obtain also the optimal dual solution. From the sensitivity analysis reported by Gurobi-Python, respond to the following questions.

(b, 3 puntos) What is the marginal cost of an additional unit of demand for product 1 in the first period?

(c, 3 puntos) What are the marginal costs of increasing and of decreasing the demand of product 1 in the second period?

(d, 3 puntos) How much does the optimal cost change if the production capacity in the first period increases by one unit?

(e, 3 puntos) How much does the optimal cost change if the production capacity in the first period decreases by one unit?

(f, 3 puntos) How much does the optimal cost change if the inventory storage capacity in the first period decreases by one unit?

(g, 3 puntos) If the production cost of product 2 in the second period decreases, does the optimal production plan need to be changed?

Answers.

(a) The optimal primal solution is

$$x_{11}^* = 200, x_{12}^* = 0, x_{21}^* = 100, x_{22}^* = 100,$$

and

$$I_{11}^* = 200, I_{12}^* = 0, I_{21}^* = 0, I_{22}^* = 0,$$

with an optimal cost of 6000 €.

Furthermore, denoting the dual variables as indicated next,

$$
\begin{aligned}
& \text{minimize} \sum_{t=1}^{L} \sum_{i=1}^{N} (p_{it} x_{it} + h_{it} I_{it}) \\
& \text{subject to:} \\
\pi_{it}\colon\ & x_{it} + I_{i,t-1} - I_{it} = D_{it}, \quad i = 1, \ldots, N,\ t = 1, \ldots, L \\
u_t\colon\ & \sum_{i=1}^{N} m_i x_{it} \leqslant K_t, \quad t = 1, \ldots, L \\
v_t\colon\ & \sum_{i=1}^{N} I_{it} \leqslant \mathrm{INV}_t, \quad t = 1, \ldots, L \\
& x_{it}, I_{it} \geqslant 0, \quad i = 1, \ldots, N,\ t = 1, \ldots, L,
\end{aligned}
$$

their optimal values as obtained by Gurobi are

$$\pi_{11}^* = 15, \pi_{12}^* = 20, \pi_{21}^* = 15, \pi_{22}^* = 20,$$

and

$$u_1^* = -5, u_2^* = 0, v_1^* = 0, v_2^* = 0.$$

(b, 3 puntos) What is the marginal cost of an additional unit of demand for product 1 in the first period? 20 €.

(c, 3 puntos) What are the marginal costs of increasing and of decreasing the demand of product 1 in the second period? Of increasing: 25 €. Of decreasing: -20 €.

(d, 3 puntos) How much does the optimal cost change if the production capacity in the first period increases by one unit? Nothing.

(e, 3 puntos) How much does the optimal cost change if the production capacity in the first period decreases by one unit? It increases by 10 €.

(f, 3 puntos) How much does the optimal cost change if the inventory storage capacity in the first period decreases by one unit? It increases by 10 €.

(g, 3 puntos) If the production cost of product 2 in the second period decreases, does the optimal production plan need to be changed? The current optimal solution remains optimal as long as that production cost satisfies $15 \leqslant p_{22} \leqslant 25$. Since the original value is $p_{22} = 20$, it can decrease by 5 €.