KONSTANTINOS PAGANOPOULOS 01769789

OPTIMISATION AND DECISION MODELS

ASSIGNMENT 3 29/11/2019

INTRODUCTION

In this assignment solutions for homework 3 and 4 of "Optimisation and Decision Models MSc Business Analytics 2019/20" Exercises 3 are provided. The solutions include hand written calculations, as well as using Microsoft Excel and AMPL program for solving discrete optimisation problems.

QUESTION 3

(a)

The discrete problem is the following:

<u>Decision Variables:</u>

We introduce the integer variables $x_1, x_2, x_3, x_4 \geq 0$, where

$x_i = number\ of\ production\ units\ of\ product\ i,\ \ where\ i = 1,2,3,4$

We also introduce the binary variables Y_1, Y_2, Y_3, Y_4, where $Y_i = \begin{cases} 1, & if\ product\ i\ is\ produced \\ 0, & otherwise \end{cases}$, $where\ i = 1,2,3,4$

<u>Objective Function:</u>

$$maximise\ 70x_1 + 60x_2 + 90x_3 + 80x_4 - 50,000Y_1 - 40,000Y_2 - 70,000Y_3 - 60,000Y_4$$

<u>Constraints:</u>

Due to consideration 1: $Y_1 + Y_2 + Y_3 + Y_4 \leq 2$

Due to consideration 2: $Y_3 \leq Y_1 + Y_2$

Due to consideration 3: either $x_1 + x_2 \leq 20,000$

or

(or both) $x_3 + x_4 \leq 20,000$

We now introduce a binary variable Y ε {0,1} and we make use of the "Big -M- trick", where M equals a very large number (let's say 1,000,000) as follows:

$$x_1 + x_2 \leq 20,000 + MY$$
$$x_3 + x_4 \leq 20,000 + M(1 - Y)$$

Due to demand: $x_1 \leq 10,000Y_1$

$x_2 \leq 15,000Y_2$

$x_3 \leq 12,500Y_3$

$x_4 \leq 9,000Y_4$

$Y \epsilon \{0,1\}$

(b)

The Excel model could look as follow:

1	**Product**	**Net revenues or Start-up costs**	
2	**1**	£70	
3	**2**	£60	
4	**3**	£90	
5	**4**	£80	
6	**1**	£50,000	
7	**2**	£40,000	
8	**3**	£70,000	
9	**4**	£60,000	
10			
11	**Decision variables**		
12	**X1**	0	
13	**X2**	15000	
14	**X3**	12500	
15	**X4**	0	
16	**Y1**	0	
17	**Y2**	1	
18	**Y3**	1	
19	**Y4**	0	
20	**Y**	0	
21			
22	**Objective function**	£1,915,000	
23			
24	**Constraints**		
25	**Consideration 1**	2	2
26	**Consideration 2**	1	1
27	**Consideration 3**	15000	20000
28	**Consideration 4**	12500	1020000
29	**demand 1**	0	0
30	**demand 2**	15000	15,000
31	**demand 3**	12500	12,500
32	**demand 4**	0	0

Thus, the profit-maximising production mix is to produce 15,000 production quantities of x_2 and 12,500 of x_3, leading to revenues of £1,915,000.

Please find attached the solution in the Excel Spreadsheet ("Konstantinos_Paganopoulos_Assingment_3.xlsx") for more details.

QUESTION 4

(a)

We follow the same procedure with previous assignment to convert the 1-norm regression problem to a linear one. In the end, we are going to add some new constraints that transform the problem to a mixed binary linear one.

The 1-norm regression problem is the following:

$$Chance = \beta_0 + \beta_1 GRE + \beta_2 TOEFL + \beta_3 Univ + \beta_4 SOP + \beta_5 LOR + \beta_6 CGPA + \beta_7 Res$$

We also know that k = 7 = number of independent variables and n = 100 = number of records (the value of the parameter k is chosen by the decision maker to be the maximum possible for more accurate results).

As a result, we can rewrite the 1-norm regression problem as:

$$Chance_i = \beta_0 + \beta_1 GRE_{i1} + \ldots + \beta_7 Res_{i7}$$

We would like to solve:

$$minimise \parallel Y - X\beta \parallel_1$$

$$where \quad Y = \begin{bmatrix} Chance_1 \\ Chance_2 \\ . \\ . \\ . \\ Chance_{100} \end{bmatrix}, \quad X = \begin{bmatrix} 1 & GRE_1 & TOEFL_1 & Univ_1 & SOP_1 & LOR_1 & CGPA_1 & Res_1 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 1 & GRE_{100} & TOEFL_{100} & Univ_{100} & SOP_{100} & LOR_{100} & CGPA_{100} & Res_{100} \end{bmatrix}, \quad \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_7 \end{bmatrix}$$

Which is equal to

$$\begin{aligned} & minimise \parallel Y - X\beta \parallel_1 \\ & s.t.\ \beta \in R^8 \end{aligned}$$

As a result, we solve the following:

$$\begin{aligned} minimise & \parallel Chance_1 - (\beta_0 + \beta_1 GRE_1 + \beta_2 TOEFL_1 + \ldots + \beta_7 Res_1) \parallel + \\ & . \\ & . \\ & . \\ & \parallel Chance_{100} - (\beta_0 + \beta_1 GRE_{100} + \beta_2 TOEFL_{100} + \ldots + \beta_7 Res_{100}) \parallel \\ & s.t.\ \beta_0, \beta_1, \beta_2, \ldots, \beta_7 \in \mathbb{R} \end{aligned}$$

We now move the non-linearities to the constraints:

$$minimise\ \theta_1 + \theta_2 + \ldots + \theta_{100}$$

$$s.t.\ \theta_1 = \| Chance_1 - (\beta_0 + \beta_1 GRE_1 + \beta_2 TOEFL_1 + \ldots + \beta_7 Res_1) \|$$

$$\vdots$$

$$\theta_{100} = \| Chance_{100} - (\beta_0 + \beta_1 GRE_{100} + \beta_2 TOEFL_{100} + \ldots + \beta_7 Res_{100}) \|$$
$$\beta_0, \beta_1, \beta_2, \ldots \beta_7, \theta_1, \theta_2 \ldots, \theta_{100} \in \mathbb{R}$$

We now relax the equalities to inequalities:

$$minimise\ \theta_1 + \theta_2 + \ldots + \theta_{100}$$

$$s.t.\ \theta_1 \geq \| Chance_1 - (\beta_0 + \beta_1 GRE_1 + \beta_2 TOEFL_1 + \ldots + \beta_7 Res_1) \|$$

$$\vdots$$

$$\theta_{100} \geq \| Chance_{100} - (\beta_0 + \beta_1 GRE_{100} + \beta_2 TOEFL_{100} + \ldots + \beta_7 Res_{100}) \|$$

$$\beta_0, \beta_1, \beta_2, \ldots \beta_7, \theta_1, \theta_2 \ldots, \theta_{100} \in \mathbb{R}$$

We now rewrite the $\| \quad \|$ - terms:

It is a mathematical fact that if we take any $a \in \mathbb{R}$ then $\| a \| = max\{a, -a\}$

In other words, $x \geq \| y \|$ is the same as writing $x \geq max\{y, -y\}$, which in turn holds exactly when $x \geq y\ \ and\ \ x \geq -y$.

Hence, we want to solve the following:

$$minimise\ \theta_1 + \theta_2 + \ldots + \theta_{100}$$

$$s.t.\ \theta_1 \geq Chance_1 - (\beta_0 + \beta_1 GRE_1 + \beta_2 TOEFL_1 + \ldots + \beta_7 Res_1)$$
$$\theta_1 \geq (\beta_0 + \beta_1 GRE_1 + \beta_2 TOEFL_1 + \ldots + \beta_7 Res_1) - Chance_1$$

$$\vdots$$

$$\theta_{100} \geq Chance_{100} - (\beta_0 + \beta_1 GRE_{100} + \beta_2 TOEFL_{100} + \ldots + \beta_7 Res_{100})$$
$$\theta_{100} \geq (\beta_0 + \beta_1 GRE_{100} + \beta_2 TOEFL_{100} + \ldots + \beta_7 Res_{100}) - Chance_{100}$$

$$\beta_0, \beta_1, \beta_2, \ldots \beta_7, \theta_1, \theta_2 \ldots, \theta_{100} \in \mathbb{R}$$

However, now we also have some extra constraints for the problem to be constructed in such a way that at most “k” of the slopes corresponding to the independent variables can be nonzero. First, we have to specify that by giving the value 7 to “k” we mean that for all 7 “betas” none of their value is equal to zero. We also know that the intercept of the regression line should always be present and not count towards the quota “k”. In other words, $\beta_0 \neq 0$. As a result, we have the following extra constraints:

$\beta_1, \ldots, \beta_k \neq 0$, where k = 7

Hence the previous constraint is equal to: $\beta_1, \ldots, \beta_7 \neq 0$ which is equal to the following:
$\beta_1, \ldots, \beta_7 \geq 0 \; and \; \beta_1, \ldots, \beta_7 \leq 0$

We now introduce the binary variables $k_1, \cdots, k_7$ ε {0,1} and we use the Big-M-trick for a large number M = 100 as follows:

$$\begin{cases} \beta_1 \geq 0 \\ \beta_1 \leq 0 \end{cases} \qquad \begin{cases} \beta_1 \geq 0 - Mk_1 \\ \beta_1 \leq 0 + Mk_1 \end{cases} \qquad \begin{cases} \beta_1 \geq -Mk_1 \\ \beta_1 \leq Mk_1 \end{cases}$$

$$\vdots \qquad \Rightarrow \qquad \vdots \qquad \Rightarrow \qquad \vdots$$

$$\begin{cases} \beta_7 \geq 0 \\ \beta_7 \leq 0 \end{cases} \qquad \begin{cases} \beta_7 \geq 0 - Mk_7 \\ \beta_7 \leq 0 + Mk_7 \end{cases} \qquad \begin{cases} \beta_7 \geq -Mk_7 \\ \beta_7 \leq Mk_7 \end{cases}$$

We also must consider that at most “k” of the slopes corresponding to the independent variables are allowed to be nonzero. As a result, our last constraint is the following:

$$k_1 + k_2 + k_3 + k_4 + k_5 + k_6 + k_7 \leq k$$

Of course, if we had chosen another value for “k”, for example 3, then we would have only three “betas” that were nonzero. As previously written, we choose the maximum possible value of “k” to construct a more accurate model.

(b)

We now construct an AMPL file that solves the mixed binary linear problem using “CPLEX” as a solver (“option solver ‘cplex’;”). The AMPL model could look like this:

```
set NUM ordered; # candidate number

param GRE {NUM}; # GRE Score
param TOEFL {NUM}; # TOEFL Score
param Univ {NUM}; # University Rating
param SOP {NUM}; # Statement of Purpose Strength
param LOR {NUM}; # Letter of Recommend. Strength
param CGPA {NUM}; # Undergraduate GPA
param Res {NUM}; # Research Experience
param Chance {NUM}; # Chance of Admission
```

```
# import the data set
data Graduate_Admissions.dat;

# define the rest variables
var x {NUM};

var b0;
var b1;
var b2;
var b3;
var b4;
var b5;
var b6;
var b7;

var k1 binary;
var k2 binary;
var k3 binary;
var k4 binary;
var k5 binary;
var k6 binary;
var k7 binary;

param k = 7;
param M = 100;

# define the objective function
minimize error: sum {i in NUM} x[i];

# define the constraints
subject to constr1 {i in NUM}: x[i] >= Chance[i] - (b0 + b1*GRE[i] + b2*TOEFL[i] +
b3*Univ[i] + b4*SOP[i] + b5*LOR[i] + b6*CGPA[i] + b7*Res[i]);
subject to constr2 {i in NUM}: x[i] >= (b0 + b1*GRE[i] + b2*TOEFL[i] + b3*Univ[i] +
b4*SOP[i] + b5*LOR[i] + b6*CGPA[i] + b7*Res[i]) - Chance[i];

subject to b1g: b1 >= -k1*M;
subject to b2g: b2 >= -k2*M;
subject to b3g: b3 >= -k3*M;
subject to b4g: b4 >= -k4*M;
subject to b5g: b5 >= -k5*M;
subject to b6g: b6 >= -k6*M;
subject to b7g: b7 >= -k7*M;

subject to b1l: b1 <= k1*M;
subject to b2l: b2 <= k2*M;
subject to b3l: b3 <= k3*M;
subject to b4l: b4 <= k4*M;
subject to b5l: b5 <= k5*M;
subject to b6l: b6 <= k6*M;
subject to b7l: b7 <= k7*M;

subject to nonzerosk: k1 + k2 + k3 + k4 + k5 + k6 + k7 <= k;
```

The optimal value and the optimal coefficients and binary variables for “k” = 7 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 5.840174093
197 MIP simplex iterations
0 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -1.33337
b1 = 0.00332792
b2 = 0.00359076
b3 = 0.0377193
b4 = 0.00475836
b5 = 0.0418593
b6 = 0.0349495
b7 = 0.0102293
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 1
k2 = 1
k3 = 1
k4 = 1
k5 = 1
k6 = 1
k7 = 1
```

The optimal value and the optimal coefficients and binary variables for "k" = 6 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 5.844671944
211 MIP simplex iterations
12 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -1.38411
b1 = 0.00343736
b2 = 0.0038768
b3 = 0.0372375
b4 = 0.0100766
b5 = 0.0435541
b6 = 0.0308012
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 1
k2 = 1
k3 = 1
k4 = 1
k5 = 1
k6 = 1
k7 = 0
```

The optimal value and the optimal coefficients and binary variables for “k” = 5 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 5.867923725
240 MIP simplex iterations
19 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -1.26681
b1 = 0.00333198
b2 = 0.00332792
b3 = 0.046741
b4 = 0
b5 = 0.04241
b6 = 0.0292119
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 1
k2 = 1
k3 = 1
k4 = 0
k5 = 1
k6 = 1
k7 = 0
```

The optimal value and the optimal coefficients and binary variables for “k” = 4 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 5.96751606
450 MIP simplex iterations
37 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -1.0604
b1 = 0.00273019
b2 = 0.00481799
b3 = 0.0538651
b4 = 0
b5 = 0.0562955
b6 = 0
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 1
k2 = 1
k3 = 1
k4 = 0
k5 = 1
k6 = 0
k7 = 0
```

The optimal value and the optimal coefficients and binary variables for “k” = 3 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 6.134
404 MIP simplex iterations
23 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -0.6056
b1 = 0
b2 = 0.0084
b3 = 0.0592
b4 = 0
b5 = 0.0592
b6 = 0
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 0
k2 = 1
k3 = 1
k4 = 0
k5 = 1
k6 = 0
k7 = 0
```

The optimal value and the optimal coefficients and binary variables for “k” = 2 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 6.637182847
424 MIP simplex iterations
27 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -2.44545
b1 = 0.00687314
b2 = 0
b3 = 0
b4 = 0
b5 = 0
b6 = 0.112865
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 1
k2 = 0
k3 = 0
k4 = 0
k5 = 0
k6 = 1
k7 = 0
```

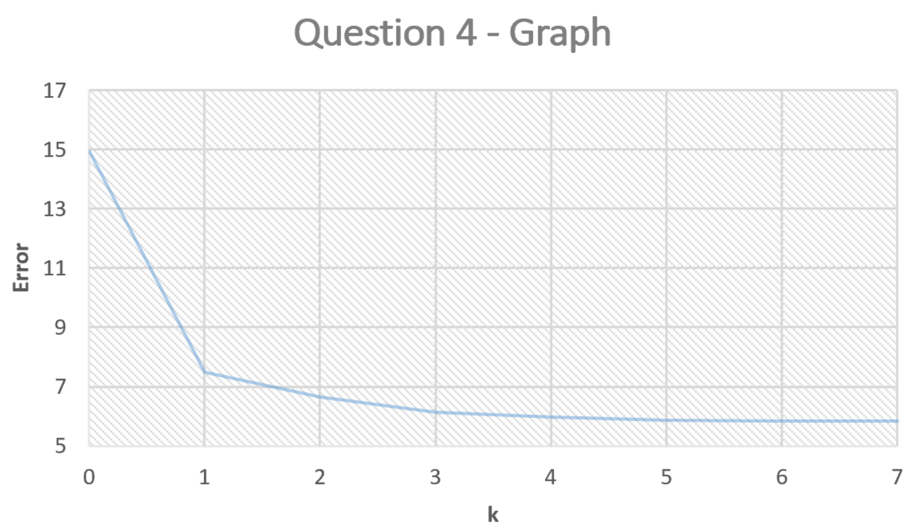
The optimal value and the optimal coefficients and binary variables for “k” = 1 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal integer solution; objective 7.480512821
195 MIP simplex iterations
0 branch-and-bound nodes
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = -1.0812
b1 = 0
b2 = 0
b3 = 0
b4 = 0
b5 = 0
b6 = 0.209402
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 0
k2 = 0
k3 = 0
k4 = 0
k5 = 0
k6 = 1
k7 = 0
```

The optimal value and the optimal coefficients and binary variables for “k” = 0 are the following:

```
ampl: model Konstantinos_Paganopoulos_Assignment_3.mod;
ampl: option solver cplex;
ampl: solve;
CPLEX 12.9.0.0: optimal solution; objective 14.93
67 dual simplex iterations (48 in phase I)
ampl: display b0, b1, b2, b3, b4, b5, b6, b7;
b0 = 0.68
b1 = 0
b2 = 0
b3 = 0
b4 = 0
b5 = 0
b6 = 0
b7 = 0
ampl: display k1, k2, k3, k4, k5, k6, k7;
k1 = 0
k2 = 0
k3 = 0
k4 = 0
k5 = 0
k6 = 0
k7 = 0
```

Now we plot a graph with "k" on the x-axis and the estimation error (i.e., the optimal objective value from the regression problem) on the y-axis. The graph could be the following:



As we can see from the above graph, the error decreases as we increase "k", or in other words the error decreases as the number of the slopes corresponding to the independent variables increases. That is expected, since more variables result in a better modelling of the problem.

Please find attached the solution in the AMPL file ("Konstantinos_Paganopoulos_Assignment_3.mod") for more details.

CONCLUSION

In the given coursework assignment, solutions to 2 discrete optimisation problems through various ways are provided. The Excel spreadsheet and the AMPL file are attached separately.