Solutions 2

Solution to (1) (a): Using the decision variables suggested in the question, we obtain the following linear program:

$$
\begin{array}{ll}
\text{minimise} & 17 * (AO + AT + AH + AD) + 3AO + 2AT + 5AH + 7AD + \\
& 20 * (BO + BT + BH + BD) + 6BO + 4BT + 8BH + 3BD + \\
& 24 * (CO + CT + CH + CD) + 9CO + 1CT + 5CH + 4CD \\
\text{subject to} & AO + AT + AH + AD \leq 800 \\
& BO + BT + BH + BD \leq 700 \\
& CO + CT + CH + CD \leq 700 \\
& AO + BO + CO = 300 \\
& AT + BT + CT = 500 \\
& AH + BH + CH = 400 \\
& AD + BD + CD = 600 \\
& AO, AT, AH, AD, BO, BT, BH, BD, CO, CT, CH, CD \geq 0
\end{array}
$$

Solution to (1) (b): The AMPL file could look as follows:

```
var AO >= 0;
var AT >= 0;
var AH >= 0;
var AD >= 0;

var BO >= 0;
var BT >= 0;
var BH >= 0;
var BD >= 0;

var CO >= 0;
var CT >= 0;
var CH >= 0;
var CD >= 0;

minimize obj:       17 * (AO + AT + AH + AD) + 3 * AO + 2 * AT + 5 * AH + 7 * AD +
                    20 * (BO + BT + BH + BD) + 6 * BO + 4 * BT + 8 * BH + 3 * BD +
                    24 * (CO + CT + CH + CD) + 9 * CO + 1 * CT + 5 * CH + 4 * CD;

subject to Cap_A: AO + AT + AH + AD <= 800;
subject to Cap_B: BO + BT + BH + BD <= 700;
subject to Cap_C: CO + CT + CH + CD <= 700;
subject to Demand_O: AO + BO + CO = 300;
subject to Demand_T: AT + BT + CT = 500;
subject to Demand_H: AH + BH + CH = 400;
subject to Demand_D: AD + BD + CD = 600;
```

The optimal objective is 40,400, and the transshipment plan is as follows:

```
display AO, AT, AH, AD;
AO = 300
AT = 100
AH = 400
AD = 0

display BO, BT, BH, BD;
BO = 0
BT = 100
BH = 0
BD = 600

display CO, CT, CH, CD;
CO = 0
CT = 300
CH = 0
CD = 0
```

Apart from the optimal production and transhipment plan itself, here are some possible conclusions that can be drawn from this information:

- In the optimal solution, Birmingham only serves two customers and Cardiff only serves one customer; this may enable us to generate further process efficiencies.

- In the optimal solution, the Cardiff factory only operates at ~43% of its capacity; one may therefore consider scaling down the operations (at least temporarily) at this plant (*e.g.*, in terms of the actively used machines, as well as temporary staff).

- Increasing the capacity of the Aberdeen factory (Cap_A has a shadow price of -6) would be six times as beneficial (in terms of cost savings) as increasing the capacity of the Birmingham factory (Cap_B has a shadow price of -1); this should be considered in future capacity expansion plans.

- Based on the shadow prices of the customer demand constraints for the optimal solution, one could consider charging customer-specific prices (*e.g.*, it is more expensive to serve Hilton Appliances than the other customers) — this may have an impact on the customer demands, however, and thus needs to be evaluated carefully. Also, we would need to check the valid ranges for the shadow prices first!

Solution to (2) (a): We bring the problem into the standard form of the dual problem. To this end, we first replace the equality with two inequalties and subsequently transform all inequalities into greater-or-equal-to inequalities:

$$\begin{array}{ll} \text{minimise} & 3\,x_1 + 5\,x_2 - x_3 \\ \text{subject to} & x_1 + x_3 \geq 4 \\ & -x_1 - x_3 \geq -4 \\ & -x_2 + 2\,x_3 \geq -2 \\ & x_1,\, x_2 \geq 0,\, x_3 \text{ unrestricted} \end{array}$$

We now replace the unrestricted variable with the difference of two nonnegative ones:

$$\begin{array}{lll} \text{minimise} & 3\,x_1 + 5\,x_2 - x_3^+ + x_3^- & \\ \text{subject to} & x_1 + x_3^+ - x_3^- \geq 4 & [y_1] \\ & -x_1 - x_3^+ + x_3^- \geq -4 & [y_2] \\ & -x_2 + 2\,x_3^+ - 2\,x_3^- \geq -2 & [y_3] \\ & x_1,\, x_2,\, x_3^+,\, x_3^- \geq 0 & \end{array}$$

We can now use the definition of the primal-dual pair to obtain:

$$\begin{array}{ll} \text{maximise} & 4\,y_1 - 4\,y_2 - 2\,y_3 \\ \text{subject to} & y_1 - y_2 \leq 3 \\ & -y_3 \leq 5 \\ & y_1 - y_2 + 2\,y_3 \leq -1 \\ & -y_1 + y_2 - 2\,y_3 \leq 1 \\ & y_1,\, y_2,\, y_3 \geq 0 \end{array}$$

Note that this problem can be further simplified to (not necessary, though):

$$\begin{array}{ll} \text{maximise} & 4\,y_1 - 4\,y_2 - 2\,y_3 \\ \text{subject to} & y_1 - y_2 \leq 3 \\ & -y_3 \leq 5 \\ & y_1 - y_2 + 2\,y_3 = -1 \\ & y_1,\, y_2,\, y_3 \geq 0 \end{array}$$

and

$$\begin{array}{ll} \text{maximise} & 4\,y_1 - 2\,y_3 \\ \text{subject to} & y_1 \leq 3 \\ & -y_3 \leq 5 \\ & y_1 + 2\,y_3 = -1 \\ & y_1 \text{ unrestricted},\, y_3 \geq 0 \end{array}$$

Solution to (2) (b): We create two variables y_1 and y_2 for the two constraints, as well as three constraints corresponding to the three variables of the primal:

???	$4 y_1 + 2 y_2$
subject to	y_1 ??? 1
	y_1 ??? 0
	y_2 ??? -1
	y_1, y_2 need to satisfy ???

We can now fill in the rest of the details, given that the primal is a maximisation problem:

minimise	$4 y_1 + 2 y_2$	
subject to	$y_1 \leq 1$	[bizarre]
	$y_1 \leq 0$	[bizarre]
	$y_2 = -1$	[odd]
	y_1 unrestricted	[odd]
	$y_2 \geq 0$	[sensible]

Note: We can see that the dual problem is infeasible (due to the restrictions placed on y_2) — hence the primal must be unbounded or infeasible. Indeed, one readily checks that the primal problem is infeasible as well (since there are no nonpositive x_1 and x_2 that add up to 4), so this is an example of a primal-dual pair where both problems are infeasible.

Solution to (3): We bring the dual problem into the standard form by multiplying the objective function and the constraints with "-1":

$$
\begin{array}{ll}
\text{maximise} & (-\mathbf{b})^T\mathbf{x} \\
\text{subject to} & (-\mathbf{A}^T)\mathbf{x} \leq (-\mathbf{c}) \\
 & \mathbf{x} \geq \mathbf{0}
\end{array}
$$

We need to keep in mind that whatever the optimal value of this problem is, we have to take the negative of that value (due to the change in the objective function). Also, we have renamed the variables in order not to get confused. We can now apply the definition of the dual to obtain:

$$
\begin{array}{ll}
\text{minimise} & (-\mathbf{c})^T\mathbf{y} \\
\text{subject to} & (-\mathbf{A}^T)^T\mathbf{y} \geq (-\mathbf{b}) \\
 & \mathbf{y} \geq \mathbf{0}
\end{array}
$$

Multiplying the constraints with "-1", we obtain:

$$
\begin{array}{ll}
\text{minimise} & (-\mathbf{c})^T\mathbf{y} \\
\text{subject to} & \mathbf{A}\mathbf{y} \leq \mathbf{b} \\
 & \mathbf{y} \geq \mathbf{0}
\end{array}
$$

Also, we can turn the minimisation objective into a maximisation objective by multiplying the objective function with "-1", keeping in mind that whatever the optimal value of this problem is, we have to take the negative of that value *twice* now (because we already took the negative before):

$$
\begin{array}{ll}
\text{maximise} & \mathbf{c}^T\mathbf{y} \\
\text{subject to} & \mathbf{A}\mathbf{y} \leq \mathbf{b} \\
 & \mathbf{y} \geq \mathbf{0}
\end{array}
$$

This is the primal problem. Note that taking the negative of the optimal value twice just means that we take the optimal value of the problem, since (-1) * (-1) = 1.

Solution to (4) (a): The linear program is

$$
\begin{array}{lll}
\text{minimise} & \text{theta}_1 + \ldots + \text{theta}_{100} & \\
\text{subject to} & \text{theta}_i \geq \text{Chance}_i - \text{GRE}_i * \text{beta\_GRE} - \ldots - \text{Res}_i * \text{beta\_Res} & \\
& \qquad - \text{intercept} & i = 1,\ldots,100 \\
& \text{theta}_i \geq \text{GRE}_i * \text{beta\_GRE} + \ldots + \text{Res}_i * \text{beta\_Res} + \text{intercept} & \\
& \qquad - \text{Chance}_i & i = 1,\ldots,100
\end{array}
$$

Solution to (4) (b): The AMPL file could look as follows:

```
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

data Graduate_Admissions.dat;

var theta {NUM};
var beta_GRE;
var beta_TOEFL;
var beta_Univ;
var beta_SOP;
var beta_LOR;
var beta_CGPA;
var beta_Res;
var intercept;

minimize objective: sum {i in NUM} theta[i];

subject to abs1 {i in NUM}: theta[i] >= Chance[i] - GRE[i] * beta_GRE
                                        - TOEFL[i] * beta_TOEFL
                                        - Univ[i] * beta_Univ
                                        - SOP[i] * beta_SOP
                                        - LOR[i] * beta_LOR
                                        - CGPA[i] * beta_CGPA
                                        - Res[i] * beta_Res
                                        - intercept;


subject to abs2 {i in NUM}: theta[i] >= -Chance[i] + GRE[i] * beta_GRE
                                         + TOEFL[i] * beta_TOEFL
                                         + Univ[i] * beta_Univ
```

```
+ SOP[i] * beta_SOP
+ LOR[i] * beta_LOR
+ CGPA[i] * beta_CGPA
+ Res[i] * beta_Res
+ intercept;
```

I obtained the objective value `5.84017` as well as the solution

```
beta_GRE = 0.00332792
beta_TOEFL = 0.00359076
beta_Univ = 0.0377193
beta_SOP = 0.00475836
beta_LOR = 0.0418593
beta_CGPA = 0.0349495
beta_Res = 0.0102293
intercept = -1.33337
```

Solution to (4) (c): The idea of the reformulation is very similar to the reformulation of the 1-norm: We can write the model as

minimise $\quad \max \{ \text{theta}_i : i = 1, \dots, 100 \}$

subject to $\quad \text{theta}_i \geq \text{Chance}_i - \text{GRE}_i * \text{beta\_GRE} - \dots - \text{Res}_i * \text{beta\_Res} - \text{intercept} \qquad i = 1,\dots,100$

$\qquad \text{theta}_i \geq \text{GRE}_i * \text{beta\_GRE} + \dots + \text{Res}_i * \text{beta\_Res} + \text{intercept} - \text{Chance}_i \qquad i = 1,\dots,100$

We now introduce a new variable tau and set it equal to the objective function:

minimise $\quad \text{tau}$

subject to $\quad \text{tau} = \max \{ \text{theta}_i : i = 1, \dots, 100 \}$

$\qquad \text{theta}_i \geq \text{Chance}_i - \text{GRE}_i * \text{beta\_GRE} - \dots - \text{Res}_i * \text{beta\_Res} - \text{intercept} \qquad i = 1,\dots,100$

$\qquad \text{theta}_i \geq \text{GRE}_i * \text{beta\_GRE} + \dots + \text{Res}_i * \text{beta\_Res} + \text{intercept} - \text{Chance}_i \qquad i = 1,\dots,100$

The inequality can be relaxed to a greater-than-or-equal-to inequality since that inequality will be binding at optimality:

minimise $\quad \text{tau}$

subject to $\quad \text{tau} \geq \max \{ \text{theta}_i : i = 1, \dots, 100 \}$

$\qquad \text{theta}_i \geq \text{Chance}_i - \text{GRE}_i * \text{beta\_GRE} - \dots - \text{Res}_i * \text{beta\_Res} - \text{intercept} \qquad i = 1,\dots,100$

$\qquad \text{theta}_i \geq \text{GRE}_i * \text{beta\_GRE} + \dots + \text{Res}_i * \text{beta\_Res} + \text{intercept} - \text{Chance}_i \qquad i = 1,\dots,100$

We finally observe that tau is greater than or equal to the maximum over all theta_i if and only if it is greater than or equal to all theta_i. We thus get the following linear program:

$$
\begin{array}{lll}
\text{minimise} & \text{tau} & \\
\text{subject to} & \text{tau} \geq \text{theta}_i & i = 1,\dots,100 \\
 & \text{theta}_i \geq \text{Chance}_i - \text{GRE}_i * \text{beta\_GRE} - \dots - \text{Res}_i * \text{beta\_Res} & \\
 & \qquad - \text{intercept} & i = 1,\dots,100 \\
 & \text{theta}_i \geq \text{GRE}_i * \text{beta\_GRE} + \dots + \text{Res}_i * \text{beta\_Res} + \text{intercept} & \\
 & \qquad - \text{Chance}_i & i = 1,\dots,100
\end{array}
$$

Solution to (4) (d): The AMPL file could look as follows:

```
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

data Graduate_Admissions.dat;

var tau;
var theta {NUM};
var beta_GRE;
var beta_TOEFL;
var beta_Univ;
var beta_SOP;
var beta_LOR;
var beta_CGPA;
var beta_Res;
var intercept;

minimize objective: tau;

subject to epi {i in NUM}: tau >= theta[i];

subject to abs1 {i in NUM}: theta[i] >= Chance[i] - GRE[i] * beta_GRE
                                        - TOEFL[i] * beta_TOEFL
                                        - Univ[i] * beta_Univ
                                        - SOP[i] * beta_SOP
                                        - LOR[i] * beta_LOR
                                        - CGPA[i] * beta_CGPA
                                        - Res[i] * beta_Res
```

```
                                                - intercept;


subject to abs2 {i in NUM}: theta[i] >= -Chance[i] + GRE[i] * beta_GRE
                                                 + TOEFL[i] * beta_TOEFL
                                                 + Univ[i] * beta_Univ
                                                 + SOP[i] * beta_SOP
                                                 + LOR[i] * beta_LOR
                                                 + CGPA[i] * beta_CGPA
                                                 + Res[i] * beta_Res
                                                 + intercept;
```

I obtained the objective value `0.149249` as well as the solution

```
beta_GRE = -0.000741659
beta_TOEFL = 0.00233573
beta_Univ = 0.0360209
beta_SOP = -0.0184988
beta_LOR = 0.037078
beta_CGPA = 0.105112
beta_Res = 0.0572708
intercept = -0.471969
```

To interpret these quantities, we should first scale them so that all explanatory variables live on the same range — let's say [0, 10]. To this end, we need to multiply `beta_GRE` by 340 / 10, `beta_TOEFL` by 120/10, `beta_Univ` by 5/10 and so on. The rescaled values are:

```
beta_GRE = -0.025216406
beta_TOEFL = 0.02802876
beta_Univ = 0.01801045
beta_SOP = -0.0092494
beta_LOR = 0.018539
beta_CGPA = 0.105112
beta_Res = 0.5727080
```

So we see that research experience appears to be key for explaining the chances of admissions success! Beyond that, a high undergraduate GPA is important as well.

Solution to (4) (e): The larger the "*p*" in the $\|\,.\,\|_p$-norm regression, the more sensitive the regression becomes to outliers (*i.e.*, individual samples are seem to contradict the general trend). To see this, let's consider a regression problem of the form

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n,$$

where the true values y_i of 4 samples are $y_1 = 2$, $y_2 = 3$, $y_3 = 4$ and $y_4 = 5$, whereas the estimated values of y'_i as predicted by our regression are $y'_1 = 1$, $y'_2 = 4$, $y'_3 = 3$ and $y'_4 = 8$. Let's calculate the 1-, 2- and infinity-norms of the errors $y_i - y'_i$:

$$\|(1, -1, 1, -3)\|_1 = |1| + |-1| + |1| + |-3| = 6$$
$$\|(1, -1, 1, -3)\|_2 = \text{sqrt}(1^2 + [-1]^2 + [1]^2 + [-3]^2) = \text{sqrt}(12) = 3.464$$

$$\| (1, -1, 1, -3) \|_{inf} = \max \{|1|, |-1|, |1|, |-3|\} = 3$$

We can make two observations here:

1. The higher the "*p*" in the norm, the smaller the value. This is the case in general (but not very relevant for the purposes of our regression problem).
2. In the 1-norm, each error counts "equally", whereas in the infinity-norm, only the maximum error counts. In fact, we have

   $$\| (1, -1, 1, -3) \|_{inf} = \| (2, -2, 2, -3) \|_{inf} = \ldots = 3,$$

   that is, the value of the norm would not change if the errors in the first three samples were larger (as long as they do not exceed 3 in magnitude). The 2-norm is between these two extremes; since it squares the errors inside the square root, larger deviations receive more attention — but the 2-norm does not solely focus on the largest deviation as the inf-norm does.

Thus, the choice of the norm expresses our attitude towards outliers: A high norm (such as the infinity-norm) places all emphasis on the largest error and hence focuses on performing well on outliers, whereas a small norm accounts for all errors equally and thus puts less emphasis on outliers. Which norm to choose would therefore depend on whether we feel that the outliers are crucial in our estimation or whether they are less interesting (*e.g.*, because they could be caused by incorrect input data).