Lösningar till tentan i SF1861 Optimeringslära, 3 Juni, 2016

Uppgift 1.(a)

We note that each column in the matrix $\mathbf{A}$ contains one "+1" and one "−1", while all the other elements in the column are zeros. We also note that the sum of the elements in the vector $\mathbf{b}$ is zero. These observations imply that the LP problem is in fact a balanced network flow problem with 4 nodes (one for each row in $\mathbf{A}$) and 4 directed arcs (one for each column in $\mathbf{A}$). The network corresponding to the given $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{c}$ in this exercise can be illustrated by FIGURE 1 below, where the supply at the nodes (i.e. the components in the vector $\mathbf{b}$), and the unit costs of the arcs (i.e. the components in the vector $\mathbf{c}$) are written in the figure. Arcs from Node1 are directed from left to right, while arcs from Node3 are directed from right to left. Negative supply means demand. If the flow in the arc from node i to node j is denoted x_{ij}, the variable vector is $\mathbf{x} = (x_{12}, x_{32}, x_{34}, x_{14})^{\mathsf{T}}$.

The suggested solutions $\hat{\mathbf{x}} = (50, 0, 30, 10)^{\mathsf{T}}$ and $\tilde{\mathbf{x}} = (20, 30, 0, 40)^{\mathsf{T}}$ can be illustrated in FIGURE 2 and FIGURE 3, respectively, with the arc-flows written on the arcs.

```
            Node2
              O -50                      O -50                O -50
             / \                        /                    / \
            /   \                      /                    /   \
           /     \                    /                    /     \
          2       4                 50                   20       30
         /         \                /                    /         \
        /           \              /                    /           \
       /             \            /                    /             \
Node1 O +60       +30 O Node3    O +60       +30 O    O +60       +30 O
       \             /            \             /      \
        \           /              \           /        \
         \         /                \         /          \
          1       3                 10      30          40
           \     /                    \     /             \
            \   /                      \   /                \
             \ /                        \ /                  \
              O -40                      O -40                O -40
            Node4

           FIGURE 1                   FIGURE 2             FIGURE 3
```

Both the suggested solutions are clearly feasible, since there is flow balance in each of the four nodes and all arc-flows are non-negative. Moreover, each of the two solutions corresponds to a spanning tree in the network. Thus, both solutions are basic feasible solutions to the LP problem. It remains to show that they are optimal.

For each of the two solutions, the simplex multipliers y_i for the four nodes are calculated by $y_4 = 0$ and $y_i - y_j = c_{ij}$ for all arcs (i, j) in the corresponding spanning tree.
In FIGURE 4, the y_i for $\hat{\mathbf{x}}$ are calculated in the order $y_4 = 0$, $y_3 = 3$, $y_1 = 1$, $y_2 = -1$.
In FIGURE 5, the y_i for $\tilde{\mathbf{x}}$ are calculated in the order $y_4 = 0$, $y_1 = 1$, $y_2 = -1$, $y_3 = 3$.

Then the reduced cost for the single non-basic variable x_{32} in FIGURE 4 is calculated by $r_{32} = c_{32} - y_3 + y_2 = 4 - 3 + (-1) = 0$,
while the reduced cost for the single non-basic variable x_{34} in FIGURE 5 is calculated by $r_{34} = c_{34} - y_3 + y_4 = 3 - 3 + 0 = 0$.

```
             O y2=-1                       O y2=-1
            /                             / \
           /                             /   \
          /                             /     \
         2     r32=4-3+(-1)=0          2       4
        /                             /         \
       /                             /           \
      /                             /             \
y1=1 O               O y3=3   y1=1 O               O y3=3
      \             /               \
       \           /                 \
        \         /                   \
         1       3                     1     r34=3-3+0=0
          \     /                       \
           \   /                         \
            \ /                           \
             O y4=0                        O y4=0

          FIGURE 4                      FIGURE 5
```

Since $r_{32} \geq 0$ in FIGURE 4, $\hat{\mathbf{x}}$ is an optimal solution to the LP problem.
Since $r_{34} \geq 0$ in FIGURE 5, $\tilde{\mathbf{x}}$ is an optimal solution to the LP problem.

As a final check, we see that $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = 2 \cdot 50 + 4 \cdot 0 + 3 \cdot 30 + 1 \cdot 10 = 200$,
and $\mathbf{c}^\mathsf{T}\tilde{\mathbf{x}} = 2 \cdot 20 + 4 \cdot 30 + 3 \cdot 0 + 1 \cdot 40 = 200$, so that $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{c}^\mathsf{T}\tilde{\mathbf{x}}$.

Uppgift 1.(b)

Först använder vi Gauss–Jordans metod på matrisen

$$[\,\mathbf{A}\ \ \mathbf{b}\,] = \begin{bmatrix} 1 & 0 & 0 & 1 & 60 \\ -1 & -1 & 0 & 0 & -50 \\ 0 & 1 & 1 & 0 & 30 \\ 0 & 0 & -1 & -1 & -40 \end{bmatrix}.$$

Efter några "enkla radoperationer" erhålls matrisen

$$[\,\mathbf{T}\ \ \mathbf{p}\,] = \begin{bmatrix} 1 & 0 & 0 & 1 & 60 \\ 0 & 1 & 0 & -1 & -10 \\ 0 & 0 & 1 & 1 & 40 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Nu är $\mathbf{A}$ överförd till trappstegsform med *tre trappstegsettor*, medan den ursprungliga högerledsvektorn $\mathbf{b}$ samtidigt överförts till den nya högerledsvektorn $\mathbf{p}$.

Observera att de bägge ekvationssystemet $\mathbf{Ax} = \mathbf{b}$ och $\mathbf{Tx} = \mathbf{p}$ har samma lösningsmängd, eftersom de enkla radoperationerna inte ändrar denna.

Men den allmänna lösningen till $\mathbf{Tx} = \mathbf{p}$ erhålls i detta fall genom att först sätta den variabel som inte svarar mot trappstegsettor till ett godtyckligt tal v, dvs $x_4 = v$, och därefter bestämma hur "trappstegsvariablerna" x_1, x_2 och x_3 måste bero av v för att $\mathbf{Tx} = \mathbf{p}$ ska bli uppfyllt.
Detta ger att den allmänna lösningen till $\mathbf{Tx} = \mathbf{p}$, och därmed även till $\mathbf{Ax} = \mathbf{b}$, är
$x_1 = 60 - v,\ x_2 = -10 + v,\ x_3 = 40 - v$ och $x_4 = v$, som kan skrivas

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 60 \\ -10 \\ 40 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \cdot v\,,$$ för $v \in I\!R$, ur vilket framgår att vektorn

$$\bar{\mathbf{x}} = \begin{pmatrix} 60 \\ -10 \\ 40 \\ 0 \end{pmatrix}$$ är en lösning till $\mathbf{Ax} = \mathbf{b}$, medan $$\mathbf{Z} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$ är en matris vars (i detta fall enda) kolonn utgör en bas till $\mathcal{N}(\mathbf{A})$.

Om vi nu kräver att $\mathbf{Ax} = \mathbf{b}$ och $x_j \geq 15$ för alla j, så innebär detta att v måste uppfylla:

$$\begin{aligned} 60 - v &\geq 15, \quad \text{dvs } v \leq 45, \\ -10 + v &\geq 15, \quad \text{dvs } v \geq 25, \\ 40 - v &\geq 15, \quad \text{dvs } v \leq 25, \ \text{och} \\ 0 + v &\geq 15, \quad \text{dvs } v \geq 15. \end{aligned}$$

Vi ser att enda möjligheten är att $v = 25$, vilket svarar mot lösningen $\mathbf{x} = (35, 15, 15, 25)^{\mathsf{T}}$.

Uppgift 2.(a)

Vi har ett LP-problem på standardformen

$$\begin{aligned} \text{minimera} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{då} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

där $\mathbf{A} = \begin{bmatrix} 2 & 2 & 1 & 1 & 0 \\ 1 & 2 & 2 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 60 \\ 90 \end{pmatrix}$ och $\mathbf{c}^{\mathsf{T}} = (-5, -8, -5, 0, 0)$.

Startlösningen ska ha basvariablerna x_4 och x_5, så att $\beta = (4, 5)$ och $\nu = (1, 2, 3)$.

Motsvarande basmatris är $\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, medan $\mathbf{A}_\nu = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$.

Motsvarande baslösning har $\mathbf{x}_\beta = \bar{\mathbf{b}}$ och $\mathbf{x}_\nu = \mathbf{0}$, där $\bar{\mathbf{b}}$ ges av systemet $\mathbf{A}_\beta\bar{\mathbf{b}} = \mathbf{b}$,

dvs $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 60 \\ 90 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 60 \\ 90 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna erhålls ur systemet $\mathbf{A}_\beta^{\mathsf{T}}\mathbf{y} = \mathbf{c}_\beta$,

dvs $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av

$\mathbf{r}_\nu^{\mathsf{T}} = \mathbf{c}_\nu^{\mathsf{T}} - \mathbf{y}^{\mathsf{T}}\mathbf{A}_\nu = (-5, -8, -5) - (0,\, 0) \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix} = (-5, -8, -5)$.

Eftersom $r_{\nu_2} = r_2 = -8$ är minst, och < 0, ska vi låta x_2 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_2$ ur systemet $\mathbf{A}_\beta\,\bar{\mathbf{a}}_2 = \mathbf{a}_2$,

dvs $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_2 = \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_2 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i2}} \mid \bar{a}_{i2} > 0 \right\} = \min \left\{ \frac{60}{2}\,,\ \frac{90}{2} \right\} = \frac{60}{2} = \frac{\bar{b}_1}{\bar{a}_{12}}.$$

Minimerande index är $i = 1$, varför $x_{\beta_1} = x_4$ inte längre ska vara basvariabel.
Dess plats tas av x_2.

Nu är alltså $\beta = (2, 5)$ och $\nu = (1, 4, 3)$.

Motsvarande basmatris är $\mathbf{A}_\beta = \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix}$, medan $\mathbf{A}_\nu = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$.

Motsvarande baslösning har $\mathbf{x}_\beta = \bar{\mathbf{b}}$ och $\mathbf{x}_\nu = \mathbf{0}$, där $\bar{\mathbf{b}}$ ges av systemet $\mathbf{A}_\beta\bar{\mathbf{b}} = \mathbf{b}$,

dvs $\begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 60 \\ 90 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 30 \\ 30 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna erhålls ur systemet $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,
dvs $\begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -8 \\ 0 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av
$\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu = (-5,\, 0,\, -5) - (-4,\, 0) \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix} = (3,\, 4,\, -1)$.

Eftersom $r_{\nu_3} = r_3 = -1$ är minst, och < 0, ska vi låta x_3 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_3$ ur systemet $\mathbf{A}_\beta\, \bar{\mathbf{a}}_3 = \mathbf{a}_3$,
dvs $\begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_3 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{30}{0.5}\,,\, \frac{30}{1} \right\} = \frac{30}{1} = \frac{\bar{b}_2}{\bar{a}_{23}}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_5$ inte längre ska vara basvariabel.
Dess plats tas av x_3.

Nu är alltså $\beta = (2,3)$ och $\nu = (1,4,5)$.

Motsvarande basmatris är $\mathbf{A}_\beta = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}$, medan $\mathbf{A}_\nu = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

Motsvarande baslösning har $\mathbf{x}_\beta = \bar{\mathbf{b}}$ och $\mathbf{x}_\nu = \mathbf{0}$, där $\bar{\mathbf{b}}$ ges av systemet $\mathbf{A}_\beta\bar{\mathbf{b}} = \mathbf{b}$,
dvs $\begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 60 \\ 90 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 15 \\ 30 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna erhålls ur systemet $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,
dvs $\begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -8 \\ -5 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av
$\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu = (-5,\, 0,\, 0) - (-3, -1) \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = (2,\, 3,\, 1)$.

Eftersom $\mathbf{r}_\nu \geq \mathbf{0}$ så är den aktuella tillåtna baslösningen, dvs $\mathbf{x} = (0, 15, 30, 0, 0)^\mathsf{T}$,
optimal lösning till P. Optimalvärdet är $\mathbf{c}^\mathsf{T}\mathbf{x} = -8 \cdot 15 - 5 \cdot 30 = -270$.

Uppgift 2.(b)

Det primala problemet P är på formen

$$\begin{aligned} \text{minimera} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

Det motsvarande duala problemet D är då på formen

$$\begin{aligned} \text{maximera} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{då} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}, \end{aligned}$$

som utskrivet blir:

$$\begin{array}{rrcrrr} \text{maximera} & 60y_1 & + & 90y_2 & & \\ \text{då} & 2y_1 & + & y_2 & \leq & -5, \\ & 2y_1 & + & 2y_2 & \leq & -8, \\ & y_1 & + & 2y_2 & \leq & -5, \\ & y_1 & & & \leq & 0, \\ & & & y_2 & \leq & 0. \end{array}$$

En noggrann figur visar *dels* att det tillåtna området till D har de fyra hörnpunkterna

$$\mathbf{y}^{(1)} = \begin{pmatrix} -5 \\ 0 \end{pmatrix}, \ \mathbf{y}^{(2)} = \begin{pmatrix} -3 \\ -1 \end{pmatrix}, \ \mathbf{y}^{(3)} = \begin{pmatrix} -1 \\ -3 \end{pmatrix} \ \text{och} \ \mathbf{y}^{(4)} = \begin{pmatrix} 0 \\ -5 \end{pmatrix},$$

dels att det tillåtna området till D är obegränsat och består av alla punkter *på* och *sydväst* om de fem linjesegmenten:

1. y_1−axeln *väster* om punkten $\mathbf{y}^{(1)}$,
2. linjestycket mellan punkterna $\mathbf{y}^{(1)}$ och $\mathbf{y}^{(2)}$,
3. linjestycket mellan punkterna $\mathbf{y}^{(2)}$ och $\mathbf{y}^{(3)}$,
4. linjestycket mellan punkterna $\mathbf{y}^{(3)}$ och $\mathbf{y}^{(4)}$,
5. y_2−axeln *söder* om punkten $\mathbf{y}^{(4)}$.

Nivåkurvorna till målfunktionen $60y_1 + 90y_2$ är parallella linjer vinkelräta mot vektorn $(60, 90)^\mathsf{T}$, med växande värden åt *nordnordost* i figuren.

Den nivåkurva som svarar mot det maximala värdet på målfunktionen inom det tillåtna området ges ur figuren av den nivåkurva som går genom hörnpunkten $\mathbf{y}^{(2)}$.

Denna hörnpunkt, med $y_1 = -3$ och $y_2 = -1$ är alltså en optimal lösning till D.

Optimalvärdet är $\mathbf{b}^\mathsf{T}\mathbf{y} = 60 \cdot (-3) + 90 \cdot (-1) = -270$.

Uppgift 2.(c) Komplementaritetsvillkoren blir i detta fall:

$\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$, $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$ och $x_j(\mathbf{c} - \mathbf{A}^\mathsf{T}\mathbf{y})_j = 0$, $j = 1, \ldots, 5$, dvs

$$\begin{aligned}
2x_1 + 2x_2 + x_3 + x_4 &= 60, \\
x_1 + 2x_2 + 2x_3 + x_5 &= 90,
\end{aligned}$$

$$\begin{aligned}
&x_1 \geq 0, \\
&x_2 \geq 0, \\
&x_3 \geq 0, \\
&x_4 \geq 0, \\
&x_5 \geq 0,
\end{aligned}$$

$$\begin{aligned}
&2y_1 + y_2 + 5 \leq 0, \\
&2y_1 + 2y_2 + 8 \leq 0, \\
&y_1 + 2y_2 + 5 \leq 0, \\
&y_1 \leq 0, \\
&y_2 \leq 0,
\end{aligned}$$

$$\begin{aligned}
&x_1\,(2y_1 + y_2 + 5) = 0, \\
&x_2\,(2y_1 + 2y_2 + 8) = 0, \\
&x_3\,(y_1 + 2y_2 + 5) = 0, \\
&x_4\,y_1 = 0, \\
&x_5\,y_2 = 0.
\end{aligned}$$

Insättning av våra erhållna lösningar $\mathbf{x} = (0, 15, 30, 0, 0)^\mathsf{T}$ och $\mathbf{y} = (-3, -1)^\mathsf{T}$ ger att

$$\begin{aligned}
&2 \cdot 0 + 2 \cdot 15 + 1 \cdot 30 + 0 = 60, \\
&1 \cdot 0 + 2 \cdot 15 + 2 \cdot 30 + 0 = 90,
\end{aligned}$$

$$\begin{aligned}
0 &\geq 0, \\
15 &\geq 0, \\
30 &\geq 0, \\
0 &\geq 0, \\
0 &\geq 0,
\end{aligned}$$

$$\begin{aligned}
&2 \cdot (-3) + 1 \cdot (-1) + 5 \leq 0, \\
&2 \cdot (-3) + 2 \cdot (-1) + 8 \leq 0, \\
&1 \cdot (-3) + 2 \cdot (-1) + 5 \leq 0, \\
&-3 \leq 0 \\
&-1 \leq 0
\end{aligned}$$

$$\begin{aligned}
&0 \cdot (\,2 \cdot (-3) + 1 \cdot (-1) + 5\,) = 0 \cdot (-2) = 0, \\
&15 \cdot (\,2 \cdot (-3) + 2 \cdot (-1) + 8\,) = 15 \cdot 0 = 0, \\
&30 \cdot (\,1 \cdot (-3) + 2 \cdot (-1) + 5\,) = 30 \cdot 0 = 0, \\
&0 \cdot (-3) = 0, \\
&0 \cdot (-1) = 0.
\end{aligned}$$

Samtliga villkor är alltså uppfyllda, vilket verifierar att $\mathbf{x} = (0, 15, 30, 0, 0)^\mathsf{T}$ och $\mathbf{y} = (-3, -1)^\mathsf{T}$ är optimala lösningar till P respektive D.

Uppgift 3.(a)

Med $\mathbf{x}^{(1)} = (0,\,0)^{\mathsf{T}}$ är $\mathbf{h}(\mathbf{x}^{(1)}) = (3,\,3,\,-5,\,-5)^{\mathsf{T}}$ och $f(\mathbf{x}^{(1)}) = \frac{1}{2}\,\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\mathbf{h}(\mathbf{x}^{(1)}) = 34$.

Vidare är $\nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 2(x_1-2) & 2(x_2-0) \\ 2(x_1-0) & 2(x_2-2) \\ 2(x_1+2) & 2(x_2-0) \\ 2(x_1-0) & 2(x_2+2) \end{bmatrix}$, så att $\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} -4 & 0 \\ 0 & -4 \\ 4 & 0 \\ 0 & 4 \end{bmatrix}$.

I Gauss-Newtons metod ska man lösa ekvationssystemet

$\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\mathbf{h}(\mathbf{x}^{(1)})\,.$

I vårt fall är $\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 32 & 0 \\ 0 & 32 \end{bmatrix}$ och $\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -32 \\ -32 \end{pmatrix}$,

så ekvationssystemet blir $\begin{bmatrix} 32 & 0 \\ 0 & 32 \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 32 \\ 32 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Vi prövar $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Då blir $\mathbf{h}(\mathbf{x}^{(2)}) = (1,\,1,\,1,\,1)^{\mathsf{T}}$ och $f(\mathbf{x}^{(2)}) = 2 < 34 = f(\mathbf{x}^{(1)})$ (med bred marginal) så steget $t_1 = 1$ accepteras. Därmed har vi utfört en fullständig iteration med Gauss-Newtons metod, och erhållit $\mathbf{x}^{(2)} = (1,\,1)^{\mathsf{T}}$.

Uppgift 3.(b)

Eftersom f är godtyckligt många gånger deriverbar i alla punkter $\mathbf{x} \in I\!R^2$ så är f en konvex funktion på $I\!R^2$ om och endast om Hessianen $\mathbf{F}(\mathbf{x})$ (dvs matrisen med andraderivator) är positivt semidefinit i alla punkter $\mathbf{x} \in I\!R^2$.

När $f(x) = \frac{1}{2}\,\mathbf{h}(\mathbf{x})^{\mathsf{T}}\mathbf{h}(\mathbf{x})$ så ges Hessianen av $\mathbf{F}(\mathbf{x}) = \nabla\mathbf{h}(\mathbf{x})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}) + \sum_{i=1}^{4} h_i(\mathbf{x})\mathbf{H}_i(\mathbf{x})$.

I vårt fall är $\nabla\mathbf{h}(\mathbf{x})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 16x_1^2+32 & 16x_1x_2 \\ 16x_1x_2 & 16x_2^2+32 \end{bmatrix}$.

Vidare är $\mathbf{H}_i(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, oberoende av i och $\mathbf{x}$, samt $\sum_{i=1}^{4} h_i(\mathbf{x}) = 4x_1^2 + 4x_2^2 - 4$,

så att $\mathbf{F}(\mathbf{x}) = \nabla\mathbf{h}(\mathbf{x})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}) + \sum_{i=1}^{4} h_i(\mathbf{x})\mathbf{H}_i(\mathbf{x}) \;=\; \begin{bmatrix} 24x_1^2+8x_2^2+24 & 16x_1x_2 \\ 16x_1x_2 & 8x_1^2+24x_2^2+24 \end{bmatrix} =$

$= 8\cdot\begin{bmatrix} a & b \\ b & c \end{bmatrix}$, där $a = 3x_1^2 + x_2^2 + 3$, $\;b = 2x_1x_2\;$ och $\;c = x_1^2 + 3x_2^2 + 3$.

Här är både $a \geq 3 > 0$ och $c \geq 3 > 0$ för alla $\mathbf{x}$.
Vidare är $ac - b^2 = 3x_1^4 + 3x_2^4 + 6x_1^2x_2^2 + 12x_1^2 + 12x_1^2 + 9 \geq 9 > 0$ för alla $\mathbf{x}$.

Därmed är $\mathbf{F}(\mathbf{x})$ positivt definit för alla $\mathbf{x} \in I\!R^2$,
vilket medför att f är en strikt konvex funktion på hela $I\!R^2$.

Uppgift 4.(a)

Vi har i denna deluppgift ett QP med likhetsbivillkor, dvs ett problem på formen

minimera $\frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$,

där $\mathbf{H} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\mathbf{A} = [\,1\;\;1\;\;1\,]$ och $\mathbf{b} = 1$.

Allmänna lösningen till $\mathbf{A}\mathbf{x} = \mathbf{b}$, dvs till $x_1 + x_2 + x_3 = 1$, ges av

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1-v_1-v_2 \\ v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -v_1 \\ v_1 \\ 0 \end{pmatrix} + \begin{pmatrix} -v_2 \\ 0 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} =$$

$= \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$, där $\bar{\mathbf{x}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ är en lösning till $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{Z} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ är en matris vars kolonner utgör en bas till $\mathcal{N}(\mathbf{A})$, och $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, där v_1 och v_2 är godtyckliga reella tal.

Variabelbytet $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$ leder till att vi ska lösa systemet $(\mathbf{Z}^{\mathsf{T}}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, förutsatt att $\mathbf{Z}^{\mathsf{T}}\mathbf{H}\mathbf{Z}$ är åtminstone positivt semidefinit.

Vi får att $\mathbf{Z}^{\mathsf{T}}\mathbf{H}\mathbf{Z} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, som är positivt definit (ty $2 > 0$, $2 > 0$ och $2 \cdot 2 - 1 \cdot 1 > 0$)

och $\mathbf{Z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Systemet $(\mathbf{Z}^{\mathsf{T}}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ har då den unika lösningen $\hat{\mathbf{v}} = (1/3, -2/3)^{\mathsf{T}}$, vilket ger att $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}} = (4/3,\, 1/3, -2/3)^{\mathsf{T}}$ är unik optimal lösning till QP-problemet.

Vår erhållna vektor $\hat{\mathbf{x}} = (4/3,\, 1/3, -2/3)^{\mathsf{T}}$ och skalären u uppfyller Lagrangevillkoren för QP-problemet om $\mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{A}^{\mathsf{T}}u$ och $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$, dvs om

$$\begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{pmatrix} 4/3 \\ 1/3 \\ -2/3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} u \\ u \\ u \end{pmatrix} \quad \text{och} \quad 4/3 + 1/3 + (-2/3) = 1.$$

Vi ser att allt detta är uppfyllt om $u = 4/3$.

Uppgift 4.(b)

Nu ska vi istället försöka *maximera* $f(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$,

med $\mathbf{H}$, $\mathbf{c}$, $\mathbf{A}$ och $\mathbf{b}$ som i (a)-uppgiften.

För att få ett ekvivalent minimeringsproblem (som vi är väl hemmastadda med)

byter vi tecken på målfunktionen och studera följande ekvivalenta minimeringsproblem:

$\tilde{\mathrm{P}}$: minimera $\frac{1}{2}\,\mathbf{x}^\mathsf{T}(-\mathbf{H})\mathbf{x} + (-\mathbf{c})^\mathsf{T}\mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Vi tillämpar nollrumsmetoden och erhåller samma $\bar{\mathbf{x}}$ och $\mathbf{Z}$ som (a)-uppgiften.

Men nu blir $\mathbf{Z}^\mathsf{T}(-\mathbf{H})\mathbf{Z} = \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix}$, som ej är positivt semidefinit,

vilket medför att det inte finns någon optimal lösning till QP-problemet $\tilde{\mathrm{P}}$.

Nu ska vi försöka hitta tillåtna lösningar $\mathbf{x}$ med godtyckligt stora värden på $f(\mathbf{x})$.
Variabelbytet $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$, som garanterar att $\mathbf{A}\mathbf{x} = \mathbf{b}$, följt av en Taylorutveckling av andra graden, som är exakt för kvadratiska funktioner, leder till att

$$f(\mathbf{x}) = f(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}) = f(\bar{\mathbf{x}}) + (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})^\mathsf{T}\mathbf{Z}\mathbf{v} + \tfrac{1}{2}\,(\mathbf{Z}\mathbf{v})^\mathsf{T}\mathbf{H}\,(\mathbf{Z}\mathbf{v}) =$$

$$= f(\bar{\mathbf{x}}) + (\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}))^\mathsf{T}\mathbf{v} + \tfrac{1}{2}\,\mathbf{v}^\mathsf{T}(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = \text{ se (a)-uppgiften}$$

$$= 1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \tfrac{1}{2} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}^\mathsf{T} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 1 + v_2 + v_1^2 + v_1 v_2 + v_2^2.$$

Genom att välja t ex $\tilde{\mathbf{v}} = (10^6,\, 0)^\mathsf{T}$ så blir $f(\tilde{\mathbf{x}}) = f(\bar{\mathbf{x}} + \mathbf{Z}\tilde{\mathbf{v}}) = 1 + 10^{12} > 10^{12}$,

där alltså $\tilde{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\tilde{\mathbf{v}} = (1-10^6,\, 10^6,\, 0)^\mathsf{T}$.

Uppgift 5.

Problemet kan skrivas på formen: minimera $f(\mathbf{x})$ då $g_i(\mathbf{x}) \leq 0$, $i = 1, 2, 3$, där

$f(\mathbf{x}) = x_1^4 + x_2^4 + cx_3^4 + 4x_1 + 4x_2 + 4x_3$,
$g_1(\mathbf{x}) = 4 - x_1 - x_2$,
$g_2(\mathbf{x}) = 4 - x_1 - x_3$,
$g_3(\mathbf{x}) = 4 - x_2 - x_3$.

Bivillkorsfunktionerna är linjära, och därmed även konvexa.

Målfunktionen har Hessianen $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 12x_1^2 & 0 & 0 \\ 0 & 12x_2^2 & 0 \\ 0 & 0 & 12cx_3^2 \end{bmatrix}$,

som är positivt semidefinit for alla $\mathbf{x} \in I\!\!R^3$ om och endast om $c \geq 0$.

För icke-negativa värden på konstanten c har vi alltså ett konvext optimeringsproblem. Vidare uppfyller exempelvis $\mathbf{x} = (3, 3, 3)^\mathsf{T}$ samtliga bivillkor med strikt olikhet, så det betraktade problemet är ett *regulärt* konvext problem om $c \geq 0$, vilket medför att $\hat{\mathbf{x}}$ är en globalt optimal lösning till problemet om och endast om $\hat{\mathbf{x}}$ är en KKT-punkt.
Om däremot $c < 0$ så har problemet ingen optimal lösning.
(Låt exempelvis $\mathbf{x}(t) = (4, 4, t)^\mathsf{T}$. Detta är en tillåten lösning för varje värde på $t \geq 0$, och om $c < 0$ så gäller att $f(\mathbf{x}(t)) \to -\infty$ då $t \to \infty$.)

Uppgift 5.(a)

KKT-villkoren kan delas upp i fyra grupper enligt följande.

(KKT–1) $\dfrac{\partial f}{\partial x_j}(\mathbf{x}) + \displaystyle\sum_{i=1}^{3} y_i \frac{\partial g_i}{\partial x_j}(\mathbf{x}) = 0$, för $j = 1, 2, 3$:
$4x_1^3 + 4 - y_1 - y_2 = 0$,
$4x_2^3 + 4 - y_1 - y_3 = 0$,
$4cx_3^3 + 4 - y_2 - y_3 = 0$.

(KKT–2) Tillåten punkt, dvs $g_i(\mathbf{x}) \leq 0$ för $i = 1, 2, 3$:
$4 - x_1 - x_2 \leq 0$,
$4 - x_1 - x_3 \leq 0$,
$4 - x_2 - x_3 \leq 0$.

(KKT–3) Lagrangemultiplikatorerna icke-negativa:
$y_1 \geq 0$,
$y_2 \geq 0$,
$y_3 \geq 0$.

(KKT–4) Komplementaritetsvillkor, dvs $y_i g_i(\mathbf{x}) = 0$ för $i = 1, 2, 3$:
$y_1(4 - x_1 - x_2) = 0$,
$y_1(4 - x_1 - x_3) = 0$,
$y_1(4 - x_2 - x_3) = 0$.

Uppgift 5.(b) Antag först att $\mathbf{x} = (3,\, 3,\, 1)^{\mathsf{T}}$.

Då är $4 - x_1 - x_2 < 0$, $4 - x_1 - x_3 = 0$ och $4 - x_2 - x_3 = 0$, så KKT-2 är uppfyllt.

Vidare är KKT-4 uppfyllt om och endast om $y_1 = 0$.

När $y_1 = 0$ så är KKT-1 uppfyllt om och endast om
$112 - y_2 = 0$, $112 - y_3 = 0$, och $4c + 4 - y_2 - y_3 = 0$,
vilket är uppfyllt om och endast om $y_2 = 112$, $y_3 = 112$ och $c = 55$.
Då är även KKT-3 uppfyllt, och eftersom $c = 55 \geq 0$ så är $\mathbf{x} = (3,\, 3,\, 1)^{\mathsf{T}}$
inte bara en KKT-punkt utan även en globalt optimal lösning när $c = 55$.

Uppgift 5.(c) Antag nu att $\mathbf{x} = (1,\, 3,\, 3)^{\mathsf{T}}$.

Då är $4 - x_1 - x_2 = 0$, $4 - x_1 - x_3 = 0$ och $4 - x_2 - x_3 < 0$, så KKT-2 är uppfyllt.

Vidare är KKT-4 uppfyllt om och endast om $y_3 = 0$.

När $y_3 = 0$ så är KKT-1 uppfyllt om och endast om
$8 - y_1 - y_2 = 0$, $112 - y_1 = 0$, och $108c + 4 - y_2 = 0$.
Men de två första av dessa villkor är uppfyllda endast
om $y_1 = 112$ och $y_2 = -104$, vilket bryter mot KKT-3.

Det går alltså inte att välja något värde på c som gör $\mathbf{x} = (1,\, 3,\, 3)^{\mathsf{T}}$
till en globalt optimal lösning.

Uppgift 5.(d) Antag slutligen att $\mathbf{x} = (2,\, 2,\, 2)^{\mathsf{T}}$.

Då är $4 - x_1 - x_2 = 0$, $4 - x_1 - x_3 = 0$ och $4 - x_2 - x_3 = 0$, så KKT-2 och KKT-4 är uppfyllda.

Villkoren KKT–1 kan nu skrivas

$$\begin{aligned} y_1 + y_2 &= 36, \\ y_1 + y_3 &= 36, \\ y_2 + y_3 &= 32\,c + 4. \end{aligned}$$

Lösningen till detta ekvationssystem är, med utnyttjande av den givna räknehjälpen,

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} 36 \\ 36 \\ 32\,c+4 \end{pmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{pmatrix} 36 \\ 36 \\ 32\,c+4 \end{pmatrix} =$$

$$= \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{pmatrix} 18 \\ 18 \\ 16\,c+2 \end{pmatrix} = \begin{pmatrix} 34 - 16\,c \\ 16\,c+2 \\ 16\,c+2 \end{pmatrix}.$$

Vi ser att KKT-3 blir uppfyllda om och endast om $c \leq 34/16 = 17/8$ och $c \geq -2/16$.
Men vi kräver enligt ovan även att $c \geq 0$ för att ett globalt optimum ska finnas.

Alltså: $\mathbf{x} = (2,\, 2,\, 2)^{\mathsf{T}}$ en global optimallösning om och endast om $0 \leq c \leq 17/8$.