Lösningar till tentan i SF1861/51 Optimeringslära, 3 juni, 2015

Uppgift 1.(a) Första delen:

The network is illustrated in the following figure, where all the links are directed from left to right.

```
      1       3
500 O------O -600
     \     /
      \   /
       \/
       /\
      /  \
     /    \
400 O------O -300
    2       4
```

Let $\mathbf{x} = (\ x_{13},\ x_{14},\ x_{23},\ x_{24}\ )^{\mathsf{T}}$, where the variable x_{ij} stands for the flow in the link from node i to node j, and let $\mathbf{c} = (\ c_{13},\ c_{14},\ c_{23},\ c_{24}\ )^{\mathsf{T}}$. Then the total cost for the flow is given by $\mathbf{c}^{\mathsf{T}}\mathbf{x}$.
The flow balance conditions in the nodes can be written $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 500 \\ 400 \\ -600 \\ -300 \end{pmatrix}.$$

Finally, the given directions of the links imply the constraints $\mathbf{x} \geq \mathbf{0}$.

Then the minimum cost network flow problem can be formulated as:

minimize $\mathbf{c}^{\mathsf{T}}\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

(It is possible, and often recommended, to remove the last row in $\mathbf{A}$ and the corresponding last component in $\mathbf{b}$ to get a system without redundant equations.)

Uppgift 1.(a) Andra delen:

The four different spanning trees are shown in the following figure,
together wih the unique link flows which satisfy $\mathbf{Ax} = \mathbf{b}$.

```
1   200    3    1   600    3    1   500    3    1          3
 O------O        O------O        O------O        O        O
  \    /          \                     /          \     /
300\  /            \                   /        500\  /
    \/          -100\                 /              \/
    /\                \            100/               /\
400/  \                \             /            600/  \
  /    \                \           /               /    \
 O      O        O------O        O------O        O------O
2        4    2   400    4    2   300    4    2  -200   4

 Spanning        Spanning        Spanning        Spanning
 tree 1          tree 2          tree 3          tree 4
```

The link flows in the first spanning tree are calculated as follows:
The only way to satisfy the supply constraint in node 2 is to let $x_{23} = 400$.
Then the only way to satisfy the demand constraint in node 3 is to let $x_{13} = 200$.
Then the only way to satisfy the supply constraint in node 1 is to let $x_{14} = 300$.
Then the demand constraint in node 4 is also satisfied.

The link flows in the second spanning tree are calculated as follows:
The only way to satisfy the demand constraint in node 3 is to let $x_{13} = 600$.
Then the only way to satisfy the supply constraint in node 1 is to let $x_{14} = -100$.
The only way to satisfy the supply constraint in node 2 is to let $x_{24} = 400$.
Then the demand constraint in node 4 is also satisfied.

The link flows for the third and fourth spanning trees are calculated in a similar way.

We see that in the spanning trees 1 and 3, all the link flows are non-negative,
so these two spanning trees correspond to feasible basic solutions, $\mathbf{x} = (200,\ 300,\ 400,\ 0)^{\mathsf{T}}$ and $\mathbf{x} = (500,\ 0,\ 100,\ 300)^{\mathsf{T}}$, respectively. which satisfy both $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

But in the spanning trees 2 and 4, the flows are not non-negative in all links,
so these two spanning trees correspond to *infeasible* basic solutions which satisfy $\mathbf{Ax} = \mathbf{b}$ but not $\mathbf{x} \geq \mathbf{0}$.

Uppgift 1.(b)

Först använder vi Gauss–Jordans metod på den givna matrisen $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$.

Efter några "enkla radoperationer" erhålls matrisen $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{T}$.

Nu är $\mathbf{A}$ överförd till trappstegsform med *två trappstegsettor.*
En bas till $\mathcal{R}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^\mathsf{T})^\perp$ erhålls genom att som basvektorer välja de kolonner i $\mathbf{A}$ som svarar mot trappstegsettor i $\mathbf{T}$, dvs kolonnerna nummer 1 och 2 i $\mathbf{A}$.

De två vektorerna $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ och $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ utgör alltså en bas till $\mathcal{N}(\mathbf{A}^\mathsf{T})^\perp$

Observera att de bägge ekvationssystemet $\mathbf{Ax} = \mathbf{0}$ och $\mathbf{Tx} = \mathbf{0}$ har exakt samma lösningsmängd, eftersom de enkla radoperationerna inte ändrar denna.

Men den allmänna lösningen till $\mathbf{Tx} = \mathbf{0}$ erhålls genom att först sätta den variabel som inte svarar mot trappstegsettor, dvs x_3, till ett godtyckligt tal t, dvs $x_3 = t$, och därefter bestämma hur "trappstegsvariablerna" x_1 och x_2 måste bero av t för att $\mathbf{Tx} = \mathbf{0}$ ska bli uppfyllt. Detta ger att den allmänna lösningen till $\mathbf{Tx} = \mathbf{0}$, och därmed även till $\mathbf{Ax} = \mathbf{0}$, är

$x_1 = -2t,\ y_2 = t,\ y_3 = t$, som kan skrivas $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot t\,,$ för $t \in I\!R$.

Av detta framgår att vektorn $\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ utgör en bas till $\mathcal{N}(\mathbf{A})$.

Vi upprepar nu ovanstående metodik på matrisen $\mathbf{A}^\mathsf{T} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$.

Efter några "enkla radoperationer" erhålls matrisen $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \tilde{\mathbf{T}}$.

Nu är $\mathbf{A}^\mathsf{T}$ överförd till trappstegsform med *två trappstegsettor*, i kolonnerna 1 och 2.

De två vektorerna $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ och $\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$ utgör alltså en bas till $\mathcal{R}(\mathbf{A}^\mathsf{T}) = \mathcal{N}(\mathbf{A})^\perp$.

Slutligen ges den allmänna lösningen till $\tilde{\mathbf{T}}\mathbf{y} = \mathbf{0}$, och därmed även till $\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0}$, av

$y_1 = -t,\ y_2 = -t,\ y_3 = t$, så att vektorn $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ utgör en bas till $\mathcal{N}(\mathbf{A}^\mathsf{T})$.

Kontroll: $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = 0.$

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 $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & -2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (5,\, 6,\, 7,\, 8)$.

Den förslagna lösningen har basvariablerna x_1 och x_4, dvs $\beta = (1,4)$ och $\delta = (2,3)$.

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$, medan $\mathbf{A}_\delta = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

Basvariablernas värden i baslösningen ges av $\mathbf{x}_\beta = \bar{\mathbf{b}}$, där vektorn $\bar{\mathbf{b}}$ beräknas ur ekvationssystemet $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

dvs $\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$,

dvs $x_1 = 3$ och $x_4 = 1$, vilket stämmer med förslaget.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna värden erhålls ur systemet $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,
dvs $\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 6 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av
$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (6,\, 7) - (-1,\, 6) \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = (6,\, 7) - (4,\, 5) = (2,\, 2)$.

Eftersom reducerade kostnaderna är icke-negativa så är den föreslagna lösningen optimal.

$x_1 = 3$, $x_4 = 1$, övriga $x_j = 0$, med optimalvärdet $= 23$.

Uppgift 2.(b)

Efter införande av slackvariabler x_5 och x_6 får vi ett nytt LP-problem på standardformen

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

där nu $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & -2 & -1 & 0 \\ 1 & 1 & 1 & 1 & 0 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ och $\mathbf{c}^{\mathsf{T}} = (5,\, 6,\, 7,\, 8,\, 0,\, 0)$.

Vi startar från baslösningen ovan, med $\beta = (1, 4)$ och $\delta = (2, 3, 5, 6)$,
som är en tillåten baslösning även till detta nya problem.

Vektorerna $\bar{\mathbf{b}}$ och $\mathbf{y}$ blir desamma som ovan.

Reducerade kostnaderna för icke-basvariablerna ges nu av
$\mathbf{r}_\delta^{\mathsf{T}} = \mathbf{c}_\delta^{\mathsf{T}} - \mathbf{y}^{\mathsf{T}}\mathbf{A}_\delta = (6,\, 7,\, 0,\, 0) - (-1,\, 6) \begin{bmatrix} 2 & 1 & -1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix} = (2,\, 2, -1,\, 6)$.

Eftersom $r_{\delta_3} = r_5 = -1$ är minst, och < 0, ska vi låta x_5 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_5$ ur systemet $\mathbf{A}_\beta \bar{\mathbf{a}}_5 = \mathbf{a}_5$,

dvs $\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_1 = \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} -1/3 \\ 1/3 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_1 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i5}} \mid \bar{a}_{i5} > 0 \right\} = \frac{1}{1/3} = \frac{\bar{b}_2}{\bar{a}_{25}}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_4$ inte längre får vara kvar som basvariabel.
Dess plats tas av x_5.

Nu är alltså $\beta = (1, 5)$ och $\delta = (2, 3, 4, 6)$.

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, medan $\mathbf{A}_\delta = \begin{bmatrix} 2 & 1 & -2 & 0 \\ 1 & 1 & 1 & -1 \end{bmatrix}$.

Basvariablernas värden i baslösningen ges av $\mathbf{x}_\beta = \bar{\mathbf{b}}$, där vektorn $\bar{\mathbf{b}}$ beräknas ur ekvationssystemet $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

dvs $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna värden erhålls ur systemet $\mathbf{A}_\beta^{\mathsf{T}}\mathbf{y} = \mathbf{c}_\beta$,
dvs $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges nu av
$\mathbf{r}_\delta^{\mathsf{T}} = \mathbf{c}_\delta^{\mathsf{T}} - \mathbf{y}^{\mathsf{T}}\mathbf{A}_\delta = (6,\, 7,\, 8,\, 0) - (0,\, 5) \begin{bmatrix} 2 & 1 & -2 & 0 \\ 1 & 1 & 1 & -1 \end{bmatrix} = (1,\, 2,\, 3,\, 5)$.

Eftersom reducerade kostnaderna är icke-negativa så är den aktuella lösningen optimal.

$x_1 = 4$, $x_5 = 3$, övriga $x_j = 0$, med optimalvärdet $= 20$.

Uppgift 2.(c)

Det primala problemet P2 är på formen

$$\text{minimera} \;\; \mathbf{c}^\mathsf{T}\mathbf{x} \;\; \text{då} \;\; \mathbf{A}\mathbf{x} \geq \mathbf{b} \;\; \text{och} \;\; \mathbf{x} \geq \mathbf{0}.$$

Det motsvarande duala problemet D2 är då på formen

$$\text{maximera} \;\; \mathbf{b}^\mathsf{T}\mathbf{y} \;\; \text{då} \;\; \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c} \;\; \text{och} \;\; \mathbf{y} \geq \mathbf{0},$$

som utskrivet blir:

$$\begin{array}{rrcl}
\text{maximera} & y_1 + 4y_2 & & \\
\text{då} & y_1 + y_2 & \leq & 5, \\
 & 2y_1 + y_2 & \leq & 6, \\
 & y_1 + y_2 & \leq & 7, \\
 & -2y_1 + y_2 & \leq & 8, \\
 & y_i & \geq & 0, \;\; i = 1, 2.
\end{array}$$

En noggrann figur ger att det tillåtna området är en fyrhörning med hörnen i pukterna $(3, 0)$, $(1, 4)$, $(0, 5)$ och $(0, 0)$.

Nivåkurvorna till målfunktionen $y_1 + 4y_2$ är parallella linjer vinkelräta mot vektorn $(1, 4)$, med växande värden "uppåt" i figuren.

Den nivåkurva som svarar mot det maximala värdet på målfunktionen i det tillåtna området ges ur figuren av linjen $y_1 + 4y_2 = 20$ som går genom hörnpunkten $(y_1, y_2) = (0, 5)$. Denna hörnpunkt är alltså den optimala lösningen till det duala problemet D2.
Duala målfunktionsvärdet är 20, vilket överensstämmer med optimalvärdet till P2.

Uppgift 3.(a)

The objective function is $f(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$, with $\mathbf{H} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$.

All the three given problems are QP problem with linear equality constraints, i.e. problems of the form: minimize $f(\mathbf{x})$ subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$, with $f(\mathbf{x})$ as above.

In the first problem, QP1, $\mathbf{A} = [\,1\;\;1\;\;1\,]$ and $\mathbf{b} = 3$.
Then the general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, i.e. to $x_1 + x_2 + x_3 = 3$, is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}, \text{ for arbitrary values on } v_1 \text{ and } v_2.$$

After the variable change $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$, we should solve the system

$(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, provided that $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z}$ is positive semidefinite.

$\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite ($2 > 0$, $2 > 0$, $2{\cdot}2 - 1{\cdot}1 > 0$), and the system

$(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}}+\mathbf{c})$ becomes $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, with unique solution $\hat{\mathbf{v}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

This implies that $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is the unique global optimal solution to QP1.

In the second problem, QP2, $\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$, and $\mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

The general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is now (after some simple calculations)

$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot v = \bar{\mathbf{x}} + \mathbf{z}\,v$, for arbitrary value on v. After the variable change $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{z}\,v$, we should solve the system $(\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z})\,v = -\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, provided that $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}$ is positive semidefinite. But $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z} = -6 < 0$, so $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}$ is *not* positive semidefinite.
Thus, there is no global optimal solution to QP2.

In the third problem, QP3, $f(\mathbf{x})$ should be *maximized*, which is equivalent to *minimize* $-f(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^\mathsf{T}(-\mathbf{H})\mathbf{x} + (-\mathbf{c})^\mathsf{T}\mathbf{x}$. Further, $\mathbf{A} = [\,0\;\;0\;\;1\,]$ and $\mathbf{b} = 1$.
The general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, i.e. to $x_3 = 1$, is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}, \text{ for arbitrary values on } v_1 \text{ and } v_2.$$

After the variable change $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$, we should solve the system

$(\mathbf{Z}^\mathsf{T}(-\mathbf{H})\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}((-\mathbf{H})\bar{\mathbf{x}} + (-\mathbf{c}))$, provided that $\mathbf{Z}^\mathsf{T}(-\mathbf{H})\mathbf{Z}$ is positive semidefinite.

But $\mathbf{Z}^\mathsf{T}(-\mathbf{H})\mathbf{Z} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is *not* positive semidefinite, since $0{\cdot}0 - 1{\cdot}1 < 0$.

Thus, there is no global optimal solution to QP3.

Uppgift 3.(b)

Låt, i QP2, $\mathbf{x}(v) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot v = \begin{pmatrix} v \\ v \\ v \end{pmatrix}$ Speciellt är $\mathbf{x}(1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \hat{\mathbf{x}}$ (= den givna punkten).

Vidare är $\mathbf{Ax}(v) = \mathbf{b}$ för alla $v \in I\!R$, och $f(\mathbf{x}(v)) = -3v^2 + 6v$, som har en *unik maxpunkt* i $v = 1$. Det medför att varje punkt $\mathbf{x}(v)$ med $v \neq 1$ är en bättre tillåten lösning än $\hat{\mathbf{x}}$.

Välj exempelvis $\tilde{\mathbf{x}} = \mathbf{x}(0) = \mathbf{0}$. Då är $\tilde{\mathbf{x}}$ en tillåten lösning till QP2 med $f(\tilde{\mathbf{x}}) = 0 < 3 = f(\hat{\mathbf{x}})$.

Låt, i QP3, $\mathbf{x}(v_1, v_2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ 1 \end{pmatrix}$. Speciellt är $\mathbf{x}(1, 1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \hat{\mathbf{x}}$.

Vidare är $\mathbf{Ax}(v_1, v_2) = \mathbf{b}$ för alla v_1 och $v_2 \in I\!R$, och $f(\mathbf{x}(v_1, v_2)) = 2 + v_1 + v_2 - v_1 v_2$.

Speciellt är $f(\hat{\mathbf{x}}) = f(\mathbf{x}(1, 1)) = 3$, men detta är inte maxvärdet.

Låt exempelvis $\tilde{\mathbf{x}} = \mathbf{x}(2, 0) = (2,\ 0,\ 1)^{\mathsf{T}}$.

Då är $\tilde{\mathbf{x}}$ en tillåten lösning till QP3 med $f(\tilde{\mathbf{x}}) = f(\mathbf{x}(2, 0)) = 4 > 3 = f(\mathbf{x}(1, 1)) = f(\hat{\mathbf{x}})$.

Uppgift 3.(c)

Några välkända fakta:

1.) Den kvadratiska funktionen $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{Hx} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ är *konvex* om och endast om matrisen $\mathbf{H}$ är *positivt semidefinit*.

2.) Den symmetriska matrisen $\mathbf{H}$ är *positivt semidefinit* om och endast om $\mathbf{y}^{\mathsf{T}}\mathbf{Hy} \geq 0$ för alla vektorer $\mathbf{y}$.

3.) Den kvadratiska funktionen $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{Hx} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ är *konkav* om och endast om den kvadratiska funktionen $-f(\mathbf{x})$ är *konvex*, vilket enligt ovan är fallet om och endast om matrisen $-\mathbf{H}$ är positivt semidefinit, dvs om och endast om $\mathbf{y}^{\mathsf{T}}\mathbf{Hy} \leq 0$ för alla vektorer $\mathbf{y}$.

I vårt fall är $\mathbf{y}^{\mathsf{T}}\mathbf{Hy} = -2y_1y_2 - 2y_2y_3 - 2y_3y_1$.

Inspirerade av QP2 i (a)-uppgiften kan vi välja exempelvis $\mathbf{y} = \mathbf{z} = (1, 1, 1)^{\mathsf{T}}$.
Då blir $\mathbf{y}^{\mathsf{T}}\mathbf{Hy} = \mathbf{z}^{\mathsf{T}}\mathbf{Hz} = -6 < 0$, vilket visar att f *ej* är en konvex funktion.

Inspirerade av QP1 i (a)-uppgiften kan vi välja exempelvis $\mathbf{y} = (1, -1, 0)^{\mathsf{T}}$ = den första kolonnen i matrisen $\mathbf{Z}$. Då blir $\mathbf{y}^{\mathsf{T}}\mathbf{Hy} = 2 > 0$, vilket visar att f *ej* är en konkav funktion.

Slutsatsen är att den givna funktionen f varken är konvex eller konkav på $I\!R^3$.

Uppgift 4.(a)

Ändra beteckningarna och låt variabelvektorn heta $\mathbf{x}$, dvs

$\mathbf{x} = (x_1, x_2, x_3)^\mathsf{T} = (x, y, r)^\mathsf{T}$.

Då är $f(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{m} h_i(\mathbf{x})^2 = \frac{1}{2}\mathbf{h}(\mathbf{x})^\mathsf{T}\mathbf{h}(\mathbf{x})$, där

$h_i(\mathbf{x}) = \sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2} - x_3$ och $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \ldots, h_m(\mathbf{x}))^\mathsf{T}$.

Gradienten av h_i ges av

$$\nabla h_i(\mathbf{x}) = \left(\frac{x_1 - a_i}{\sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2}}, \frac{x_2 - b_i}{\sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2}}, -1 \right),$$

och $\nabla\mathbf{h}(\mathbf{x})$ betecknar $m \times 3$ matrisen med ovanstående gradienter till rader.

Med givna data erhålls att $f(\mathbf{x}) = \frac{1}{2}(h_1(\mathbf{x})^2 + h_2(\mathbf{x})^2 + h_3(\mathbf{x})^2 + h_4(\mathbf{x})^2)$, där

$$\begin{aligned} h_1(\mathbf{x}) &= \sqrt{(x_1 - 12)^2 + x_2^2} - x_3, \\ h_2(\mathbf{x}) &= \sqrt{x_1^2 + (x_2 - 10)^2} - x_3, \\ h_3(\mathbf{x}) &= \sqrt{(x_1 + 10)^2 + x_2^2} - x_3, \\ h_4(\mathbf{x}) &= \sqrt{x_1^2 + (x_2 + 8)^2} - x_3. \end{aligned}$$

Startpunkten ska vara $\mathbf{x}^{(1)} = (0,\ 0,\ 10)^\mathsf{T}$. Då blir

$$\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}, \ f(\mathbf{x}^{(1)}) = 4 \text{ och } \nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

I Gauss-Newtons metod ska vi lösa systemet $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)})$,

som här blir $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Vi testar först $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = (1,\ 1,\ 10)^\mathsf{T}$. Då blir

$\mathbf{h}(\mathbf{x}^{(2)}) = (\sqrt{121^2 + 1^2} - 10,\ \sqrt{1^2 + 9^2} - 10,\ \sqrt{121^2 + 1^2} - 10,\ \sqrt{1^2 + 9^2} - 10)^\mathsf{T} =$

$(\sqrt{122} - 10,\ \sqrt{82} - 10,\ \sqrt{122} - 10,\ \sqrt{82} - 10)^\mathsf{T} \approx (1, -1, 1, -1)^\mathsf{T}$, så att

$f(\mathbf{x}^{(2)}) = \frac{1}{2}\mathbf{h}(\mathbf{x}^{(2)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(2)}) \approx \frac{1}{2}(1 + 1 + 1 + 1) = 2 < 4 = f(\mathbf{x}^{(1)})$.

Steglängden $t_1 = 1$ accepteras alltså, och $\mathbf{x}^{(2)} = (1,\ 1,\ 10)^\mathsf{T}$ blir nästa iterationspunkt.

Uppgift 4.(b)

Låt
(x_1, x_2) = koordinaterna för den gemensamma mittpunkten till C_1 och C_2,
z_1 = kvadraten på radien för den mindre cirkeln C_1, och
z_2 = kvadraten på radien för den större cirkeln C_2.

Då kan problemet formuleras enligt följande i variablerna x_1, x_2, z_1 och z_2:

$$\begin{aligned} \text{minimera} \quad & \pi z_2 - \pi z_1 \\ \text{då} \quad & (x_1 - a_i)^2 + (x_2 - b_i)^2 - z_1 \ \geq \ 0, \quad i = 1, \ldots, m, \\ & (x_1 - a_i)^2 + (x_2 - b_i)^2 - z_2 \ \leq \ 0, \quad i = 1, \ldots, m. \end{aligned}$$

Uppgift 5.
Lagrangefunktionen kan skrivas $L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) + y_1 g_1(\mathbf{x}) + y_2 g_2(\mathbf{x}) =$
$= \frac{1}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2) + y_1(4 - x_1 - x_2 - x_3 - x_4) + y_2(k - x_1 - 2x_2 - 3x_3 - 4x_4)$.

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

(KKT–1) $\nabla_{\mathbf{x}} L(\mathbf{x},\mathbf{y}) = \mathbf{0}^{\mathsf{T}}$:

$\mathbf{x}^{\mathsf{T}} - y_1(1,1,1,1) - y_2(1,2,3,4) = (0,0,0,0)$.

(KKT–2) Tillåten punkt, dvs $g_i(\mathbf{x}) \leq 0$ för $i = 1,2$:

$4 - (1,1,1,1)^{\mathsf{T}}\mathbf{x} \leq 0$ och $k - (1,2,3,4)^{\mathsf{T}}\mathbf{x} \leq 0$.

(KKT–3) Lagrangemultiplikatorerna icke-negativa:

$y_1 \geq 0$ och $y_2 \geq 0$.

(KKT–4) Komplementaritetsvillkor, dvs $y_i g_i(\mathbf{x}) = 0$ för $i = 1,2$:

$y_1(4 - (1,1,1,1)^{\mathsf{T}}\mathbf{x}) = 0$ och $y_2(k - (1,2,3,4)^{\mathsf{T}}\mathbf{x}) = 0$.

(KKT-3) medför att vi kan dela upp problemlösningen i fyra olika fall.

Fall 1: $y_1 = 0$ och $y_2 = 0$.
Här medför (KKT-1) att $\mathbf{x} = (0,0,0,0)^{\mathsf{T}}$, vilket strider mot (KKT-2).
Det finns alltså ingen KKT-punkt under Fall 1.

Fall 2: $y_1 > 0$ och $y_2 = 0$.
Här medför (KKT-1) att $\mathbf{x} = y_1(1,1,1,1)^{\mathsf{T}}$,
medan (KKT-4) medför att $(1,1,1,1)^{\mathsf{T}}\mathbf{x} = 4$.
Tillsammans ger detta att $y_1 = 1$ och $\mathbf{x} = (1,1,1,1)^{\mathsf{T}}$,
som uppfyller (KKT-2) om och endast om $k \leq 10$.
Om $k = 11$ eller $k = 15$ så finns alltså ingen KKT-punkt under Fall 2.
Men om $k = 9$ så uppfyller $\mathbf{x} = (1,1,1,1)^{\mathsf{T}}$,
tillsammans med $\mathbf{y} = (1,0)^{\mathsf{T}}$, samtliga KKT-villkor.

Fall 3: $y_1 = 0$ och $y_2 > 0$.
Här medför (KKT-1) att $\mathbf{x} = y_2(1,2,3,4)^{\mathsf{T}}$,
medan (KKT-4) medför att $(1,2,3,4)^{\mathsf{T}}\mathbf{x} = k$.
Tillsammans ger detta att $y_2 = k/30$ och $\mathbf{x} = (k/30)(1,2,3,4)^{\mathsf{T}}$,
som uppfyller (KKT-2) om och endast om $k \geq 12$.
Om $k = 9$ eller $k = 11$ så finns alltså ingen KKT-punkt under Fall 3.
Men om $k = 15$ så uppfyller $\mathbf{x} = (0.5,\ 1.0,\ 1.5,\ 2.0)^{\mathsf{T}}$,
tillsammans med $\mathbf{y} = (0,\ 0.5)^{\mathsf{T}}$, samtliga KKT-villkor.

Fall 4: $y_1 > 0$ och $y_2 > 0$.
Här medför (KKT-1) att $\mathbf{x} = y_1(1,1,1,1)^{\mathsf{T}} + y_2(1,2,3,4)^{\mathsf{T}}$,
medan (KKT-4) medför att $(1,1,1,1)^{\mathsf{T}}\mathbf{x} = 4$ och $(1,2,3,4)^{\mathsf{T}}\mathbf{x} = k$.
Tillsammans ger detta att $y_1 = 6 - k/2$ och $y_2 = -2 + k/5$,
som uppfyller $y_1 > 0$ och $y_2 > 0$ om och endast om $10 < k < 12$.
Om $k = 9$ eller $k = 15$ så finns alltså ingen KKT-punkt under Fall 3.
Men om $k = 11$ så uppfyller $\mathbf{x} = (0.7,\ 0.9,\ 1.1,\ 1.3)^{\mathsf{T}}$,
tillsammans med $\mathbf{y} = (0.5,\ 0.2)^{\mathsf{T}}$, samtliga KKT-villkor.

Målfunktionen är en kvadratisk funktion med Hessianen $\mathbf{H} = \mathbf{I} =$ enhetsmatrisen som är positivt definit. Därmed är målfunktionen (strikt) konvex. Bivillkorsfunktionerna är linjära, och därmed konvexa. Alltså har vi ett konvext optimeringsproblem, och då utgör varje KKT-punkt en globalt optimal lösning till problemet.

Således gäller följande:

(a): Om $k = 9$ så är $\mathbf{x} = (1, 1, 1, 1)^\mathsf{T}$ en globalt optimal lösning.

(b): Om $k = 11$ så är $\mathbf{x} = (0.7,\ 0.9,\ 1.1,\ 1.3)^\mathsf{T}$ en globalt optimal lösning.

(c): Om $k = 15$ så är $\mathbf{x} = (0.5,\ 1.0,\ 1.5,\ 2.0)^\mathsf{T}$ en globalt optimal lösning.