Lösningar till tentan i SF1861 Optimeringslära, 14 aug 2017

Uppgift 1.(a)

Som vid varje formulering av ett LP-problem behöver vi specificera
– variabler,
– målfunktion (som ska vara linjär i variablerna), och
– bivillkor (som ska vara linjära i variablerna).

Variabler
Vi inför följande variabler x_{ij} (i sorten ton), för $i = 1, 2, 3$ och $j = 1, 2, 3, 4$:
$x_{ij} =$ kvantitet bränsle som flygbolaget beställer från Lev i för leverans till Fp j.
Då har vi tolv variabler x_{ij}, och variabelvektorn $\mathbf{x}$ kan skrivas
$\mathbf{x} = (x_{11}, x_{12}, x_{13}, x_{14}, x_{21}, x_{22}, x_{23}, x_{24}, x_{31}, x_{32}, x_{33}, x_{34})^{\mathsf{T}}$.

Målfunktion
Vi inför följande konstanter c_{ij} (i sorten kkr/ton) för $i = 1, 2, 3$ och $j = 1, 2, 3, 4$:
c_{ij} = priset (per ton) för leverans från Lev i till Fp j.
Dessa konstanter kan samlas i en vektor $\mathbf{c}$:
$\mathbf{c} = (c_{11}, c_{12}, c_{13}, c_{14}, c_{21}, c_{22}, c_{23}, c_{24}, c_{31}, c_{32}, c_{33}, c_{34})^{\mathsf{T}}$.
Den givna tabellen ger att $\mathbf{c} = (5, 4, 4, 5, 7, 4, 4, 6, 8, 6, 5, 7)^{\mathsf{T}}$.
Då kan målfunktionen, som står för flygbolagets kostnader för bränsle, skrivas:
$\mathbf{c}^{\mathsf{T}}\mathbf{x} = c_{11}x_{11} + c_{12}x_{12} + \cdots + c_{34}x_{34} = 5x_{11} + 4x_{12} + \cdots + 7x_{34}$.

Bivillkor
Den begränsade totala kapaciteten hos respektive leverantör ger bivillkoren
$x_{11} + x_{12} + x_{13} + x_{14} \leq s_1$,
$x_{21} + x_{22} + x_{23} + x_{24} \leq s_2$,
$x_{31} + x_{32} + x_{33} + x_{34} \leq s_3$,
där $s_1 = 600$, $s_2 = 350$ och $s_3 = 450$.
Bränslebehovet vid respektive flygplats ger bivillkoren
$x_{11} + x_{21} + x_{31} = d_1$,
$x_{12} + x_{22} + x_{32} = d_2$,
$x_{13} + x_{23} + x_{33} = d_3$,
$x_{14} + x_{24} + x_{34} = d_4$,
där $d_1 = 200$, $d_2 = 300$, $d_3 = 400$ och $d_4 = 500$.
Slutligen måste alla variabler vara icke-negativa, dvs
$x_{ij} \geq 0$, för $i = 1, 2, 3$ och $j = 1, 2, 3, 4$.

Kompakt LP-formulering:

$$\begin{aligned} \text{minimera} \quad & \textstyle\sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} x_{ij} \\ \text{då} \quad & \textstyle\sum_{j=1}^{4} x_{ij} \leq s_i, \text{ för } i = 1, 2, 3, \\ & \textstyle\sum_{i=1}^{3} x_{ij} = d_j, \text{ för } j = 1, 2, 3, 4, \\ & x_{ij} \geq 0, \text{ för } i = 1, 2, 3 \text{ och } j = 1, 2, 3, 4. \end{aligned}$$

Korrekt formulering utan användning av symbolen "$\sum$" kan också ge full poäng.

Uppgift 1.(b)

Först använder vi Gauss–Jordans metod på den givna matrisen $\mathbf{A} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$.

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

Nu är $\mathbf{A}$ överförd till trappstegsform med *två trappstegsettor.*
En bas till $\mathcal{R}(\mathbf{A})$ 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 $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ och $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ utgör alltså en bas till $\mathcal{R}(\mathbf{A})$.

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

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

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

Vi upprepar nu ovanstående metodik på matrisen $\mathbf{A}^\mathsf{T} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

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

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

De två vektorerna $\begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}$ och $\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix}$ utgör alltså en bas till $\mathcal{R}(\mathbf{A}^\mathsf{T})$.

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

$y_1 = t,\ y_2 = -2t,\ y_3 = t$, så att vektorn $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ utgör en bas till $\mathcal{N}(\mathbf{A}^\mathsf{T})$.

Man kan notera att för den aktuella matrisen $\mathbf{A}$ gäller att $\mathcal{N}(\mathbf{A}^\mathsf{T}) = \mathcal{N}(\mathbf{A})$, och därmed även $\mathcal{R}(\mathbf{A}^\mathsf{T}) = \mathcal{N}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^\mathsf{T})^\perp = \mathcal{R}(\mathbf{A})$. De två erhållna basvektorerna för $\mathcal{R}(\mathbf{A})$ utgör alltså en bas även för $\mathcal{R}(\mathbf{A}^\mathsf{T})$, och vice versa.

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 $\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 & 1 \\ 1 & 2 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (\,5,\ 4,\ 3,\ 2\,)$.

Den förslagna lösningen har basvariablerna x_1 och x_2, dvs $\beta = (1,2)$ och $\delta = (3,4)$.

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

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 $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$,

dvs $x_1 = 2$ och $x_2 = 3$, vilket stämmer med förslaget.

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

dvs $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (3,2) - (2,1) \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = (3,2) - (-1,1) = (4,1)$.

Samtliga reducerade kostnader är icke-negativa, så den aktuella baslösningen (= den i tentauppgiften föreslagna lösningen) $\mathbf{x} = (2,3,0,0)^\mathsf{T}$ är optimal, med optimalvärdet $\mathbf{c}^\mathsf{T}\mathbf{x} = 5 \cdot 2 + 4 \cdot 3 = 22$.

Eftersom det primala problemet är: minimera $\mathbf{c}^\mathsf{T}\mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$ och $\mathbf{x} \geq \mathbf{0}$,

så är det motsvarande duala problemet: maximera $\mathbf{b}^\mathsf{T}\mathbf{y}$ då $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$,
som utskrivet blir:

$$\begin{aligned} \text{maximera} \quad & 7y_1 + 8y_2 \\ \text{då} \quad & 2y_1 + y_2 \leq 5, \\ & y_1 + 2y_2 \leq 4, \\ & -y_1 + y_2 \leq 3, \\ & y_1 - y_2 \leq 2. \end{aligned}$$

Vektorn $\mathbf{y} = (2,1)^\mathsf{T}$ med simplexmultiplikatorer för den optimala baslösningen till det primala problemet uppfyller alla bivillkor i det duala problemet, och har duala målfunktionsvärdet $\mathbf{b}^\mathsf{T}\mathbf{y} = 7 \cdot 2 + 8 \cdot 1 = 22 = \mathbf{c}^\mathsf{T}\mathbf{x}$, vilket visar att $\mathbf{y} = (2,1)^\mathsf{T}$ är en optimal lösning till det duala problemet.

Uppgift 2.(b)

Efter införande av slackvariabler x_5 och x_6 får vi ett nytt 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 nu $\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 & 1 & 1 & 0 \\ 1 & 2 & 1 & -1 & 0 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (\,5,\, 4,\, 3,\, 2,\, 0,\, 0\,)$.

Vi startar från baslösningen ovan, med $\beta = (1, 2)$ och $\delta = (3, 4, 5, 6)$,
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
$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (3,\, 2,\, 0,\, 0) - (2,\, 1)\begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & -1 & 0 & -1 \end{bmatrix} = (4,\, 1, -2,\, 1).$

Eftersom $r_{\delta_3} = r_5 = -2$ 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 $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_5 = \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_5 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i5}} \mid \bar{a}_{i5} > 0 \right\} = \frac{2}{2/3} = \frac{\bar{b}_1}{\bar{a}_{15}}.$$

Minimerande index är $i = 1$, varför $x_{\beta_1} = x_1$ inte längre får vara kvar som basvariabel. Dess plats tas av x_5.

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

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

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 $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

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

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

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

$x_5 = 3$, $x_2 = 4$, övriga $x_j = 0$, med optimalvärdet $= 16$.

Uppgift 3.(a)

Problemet kan skrivas 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

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},\ \mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{bmatrix},\ \mathbf{b} = \begin{pmatrix} 6 \\ 2 \\ -2 \end{pmatrix}.$$

$\mathbf{H}$ är en diagonalmatris med alla diagonalelement > 0, så är $\mathbf{H}$ positivt definit.
Enligt en känd sats är då $\hat{\mathbf{x}} \in I\!R^6$ en optimal lösning till problemet om och endast om det finns en vektor $\hat{\mathbf{u}} \in I\!R^3$ sådan att $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ är en lösning till följande *Lagrangevillkor*:

$$\begin{aligned} \mathbf{H}\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{u} &= -\mathbf{c} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

I det aktuella problemet är $\mathbf{H} = \mathbf{I}$ och $\mathbf{c} = \mathbf{0}$, så ovanstående Lagrangevillkor förenklas till

$$\begin{aligned} \mathbf{x} &= \mathbf{A}^{\mathsf{T}}\mathbf{u} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

Vi ska nu, för var och en av de tre givna punkterna $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$ och $\mathbf{x}^{(3)}$, undersöka om det finns en motsvarande vektor $\mathbf{u}$ så att Lagrangevillkoren blir uppfyllda.

Vi undersöker först vilka av de tre punkterna som uppfyller $\mathbf{A}\mathbf{x} = \mathbf{b}$.
Vi får att $\mathbf{A}\mathbf{x}^{(1)} = \mathbf{b}$ och $\mathbf{A}\mathbf{x}^{(2)} = \mathbf{b}$, men $\mathbf{A}\mathbf{x}^{(3)} \neq \mathbf{b}$.
Alltså kan $\mathbf{x}^{(3)}$ ej vara en optimal lösning.

Det återstår att undersöka om $\mathbf{x}^{(1)}$ eller $\mathbf{x}^{(2)}$ uppfyller $\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{u}$ för någon vektor $\mathbf{u} \in I\!R^3$.

$$\mathbf{x}^{(1)} = \mathbf{A}^{\mathsf{T}}\mathbf{u} \iff \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} u_1 - u_2 \\ u_1 - u_3 \\ u_1 \\ u_2 - u_3 \\ u_2 \\ u_3 \end{pmatrix}.$$

Ekvationerna 3, 5 och 6 ger att $u_1 = u_2 = u_3 = 2$, men då uppfylls inte (t ex) ekvation 1.
Alltså kan $\mathbf{x}^{(1)}$ ej vara en optimal lösning.

$$\mathbf{x}^{(2)} = \mathbf{A}^{\mathsf{T}}\mathbf{u} \iff \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} u_1 - u_2 \\ u_1 - u_3 \\ u_1 \\ u_2 - u_3 \\ u_2 \\ u_3 \end{pmatrix}.$$

Ekvationerna 3, 5 och 6 ger att $u_1 = 3$, $u_2 = 2$ och $u_3 = 1$.
Med dessa värden på u_i uppfylls även ekvationerna 1, 2 och 4!

Alltså är $\mathbf{x}^{(2)}$ en optimal lösning till problemet.

Uppgift 3.(b)

Även problemet P2 kan skrivas 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 nu

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ 2 \\ -2 \end{pmatrix}.$$

$\mathbf{H}$ är nu nollmatrisen som är positivt semidefinit (ty $\mathbf{v}^{\mathsf{T}}\mathbf{H}\mathbf{v} = 0 \geq 0$ för alla $\mathbf{v} \in I\!R^6$).
Därmed är $\hat{\mathbf{x}} \in I\!R^6$ en optimal lösning till P2 om och endast om det finns en vektor $\hat{\mathbf{u}} \in I\!R^3$ sådan att $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ är en lösning till Lagrangevillkoren

$$\begin{aligned} \mathbf{H}\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{u} &= -\mathbf{c} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

I det aktuella problemet är $\mathbf{H}$ = nollmatrisen, så ovanstående Lagrangevillkor förenklas till

$$\begin{aligned} \mathbf{A}^{\mathsf{T}}\mathbf{u} &= \mathbf{c} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

Eftersom $\mathbf{A}\mathbf{x}^{(3)} \neq \mathbf{b}$ kan $\mathbf{x}^{(3)}$ aldrig vara en optimal lösning.
Men eftersom $\mathbf{A}\mathbf{x}^{(1)} = \mathbf{b}$ och $\mathbf{A}\mathbf{x}^{(2)} = \mathbf{b}$ så är både $\mathbf{x}^{(1)}$ och $\mathbf{x}^{(2)}$ optimala lösningar om vektorn $\mathbf{c}$ är sådan att systemet $\mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{c}$ har någon lösning $\mathbf{u}$, dvs om $\mathbf{c}$ ligger i bildrummet till matrisen $\mathbf{A}^{\mathsf{T}}$.

Ett exempel på en sådan vektor får vi om vi sätter $\mathbf{u} = (3, 2, 1)^{\mathsf{T}}$ varvid $\mathbf{A}^{\mathsf{T}}\mathbf{u} = (1, 2, 3, 1, 2, 1)^{\mathsf{T}}$.
Om alltså $\mathbf{c} = (1, 2, 3, 1, 2, 1)^{\mathsf{T}}$ så är både $\mathbf{x}^{(1)}$ och $\mathbf{x}^{(2)}$ optimala lösningar till P2.
(Men det finns oändligt många andra vektorer $\mathbf{c}$ i bildrummet till $\mathbf{A}^{\mathsf{T}}$.)

Uppgift 4.(a) Vi använder oss av att en symmetrisk 2×2–matris $\mathbf{H} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

* är positivt definit (p.d.) om och endast om $a > 0$, $c > 0$ och $ac - b^2 > 0$,
* är positivt semidefinit (p.s.d.) om och endast om $a \geq 0$, $c \geq 0$ och $ac - b^2 \geq 0$,

vilket kan verifieras med exempelvis en LDL^T–faktorisering.

Målfunktionen är $f(\mathbf{x}) = (x_1{+}x_2)^3 + (x_1{+}x_2)^2 + (x_1{-}2)^2 + (x_2{-}4)^2$.

Gradient till f ges då av $\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 3(x_1{+}x_2)^2 + 2(x_1{+}x_2) + 2(x_1{-}2) \\ 3(x_1{+}x_2)^2 + 2(x_1{+}x_2) + 2(x_2{-}4) \end{pmatrix}$

och Hessianen till f ges av $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 6(x_1{+}x_2) + 4 & 6(x_1{+}x_2) + 2 \\ 6(x_1{+}x_2) + 2 & 6(x_1{+}x_2) + 4 \end{bmatrix}$.

Startpunkten ges av $\mathbf{x}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, med $\nabla f(\mathbf{x}^{(1)})^\mathsf{T} = \begin{pmatrix} -4 \\ -8 \end{pmatrix}$ och $\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$.

Eftersom $\mathbf{F}(\mathbf{x}^{(1)})$ är positivt definit ($4 > 0$, $4 > 0$ och $4{\cdot}4 - 2^2 > 0$) så bestäms Newtonriktningen $\mathbf{d}^{(1)}$ ur ekvationssystemet $\mathbf{F}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\mathsf{T}$, som här blir

$\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} \mathbf{d} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$.

Vi prövar först steget $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1 \mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$.

Då blir $f(\mathbf{x}^{(2)}) = 20 = f(\mathbf{x}^{(1)})$, så steget $t_1 = 1$ accepteras ej.

Vi prövar då i stället steget $t_1 = 0.5$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1 \mathbf{d}^{(1)} = \mathbf{x}^{(1)} + 0.5\,\mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Då blir $f(\mathbf{x}^{(2)}) = 15 \,<\, 20 = f(\mathbf{x}^{(1)})$, så steget $t_1 = 0.5$ accepteras.

Därmed har vi utfört en fullständig iteration med Newtons metod och erhållit iterationspunkten $\mathbf{x}^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ med $f(\mathbf{x}^{(2)}) = 15$.

Uppgift 4.(b)

Eftersom bivillkor saknas måste varje lokal minpunkt till $f(\mathbf{x})$ uppfylla att

$$\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 3(x_1{+}x_2)^2 + 2(x_1{+}x_2) + 2(x_1{-}2) \\ 3(x_1{+}x_2)^2 + 2(x_1{+}x_2) + 2(x_2{-}4) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Genom att subtrahera den första ekvationen från den andra så erhålls att $x_2 = x_1 + 2$.
Om man sätter in detta i någon av de två ekvationerna så erhålls att
$3(2x_1{+}2)^2 + 2(2x_1{+}2) + 2(x_1{-}2) = 0$, dvs $12x_1^2 + 30x_1 + 12 = 0$,
dvs $x_1^2 + 2.5x_1 + 1 = 0$, med lösningarna $x_1 = -0.5$ och $x_1 = -2$.
Motsvarande värden på $x_2 = x_1 + 2$ är $x_2 = 1.5$ och $x_2 = 0$.

De enda lösningarna till $\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ är alltså $\hat{\mathbf{x}} = \begin{pmatrix} -0.5 \\ 1.5 \end{pmatrix}$ och $\bar{\mathbf{x}} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$.

$\mathbf{F}(\hat{\mathbf{x}}) = \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix}$ är positivt definit, så $\hat{\mathbf{x}} = \begin{pmatrix} -0.5 \\ 1.5 \end{pmatrix}$ är en lokal minpunkt, medan

$\mathbf{F}(\bar{\mathbf{x}}) = \begin{bmatrix} -8 & -10 \\ -10 & -8 \end{bmatrix}$ ej är positivt semidefinit, så $\bar{\mathbf{x}} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$ är *ej* en lokal minpunkt.

Uppgift 4.(c)

Eftersom mängden $C_a = \{\,(x_1,x_2)^\mathsf{T} \in \mathbb{R}^2 \mid x_1+x_2 \geq a\}$ är en konvex mängd med inre punkter, t ex $\mathbf{x} = (a,1)^\mathsf{T}$, så är funktionen f konvex på C_a om och endast om Hessianen $\mathbf{F}(\mathbf{x})$ är positivt semidefinit för alla $\mathbf{x} \in C_a$.

Men $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 6x_1+6x_2+4 & 6x_1+6x_2+2 \\ 6x_1+6x_2+2 & 6x_1+6x_2+4 \end{bmatrix}$ är positivt semidefinit om och endast om

(i) $6x_1+6x_2+4 \geq 0$,
(ii) $6x_1+6x_2+4 \geq 0$, och
(iii) $(6x_1+6x_2+4)(6x_1+6x_2+4) - (6x_1+6x_2+2)^2 \geq 0$.

(i) och (ii) är uppfyllda om och endast om $x_1+x_2 \geq -2/3$.

(iii) är uppfylld om och endast om $(6x_1+6x_2+4)^2 - (6x_1+6x_2+2)^2 =$ (konjugatregeln) $=$ $((6x_1+6x_2+4) - (6x_1+6x_2+2))((6x_1+6x_2+4) + (6x_1+6x_2+2)) = 2\,(12x_1+12x_2+6) \geq 0$, dvs om och endast om $x_1+x_2 \geq -1/2$.

Slutsatsen blir att $\mathbf{F}(\mathbf{x})$ är positivt semidefinit om och endast om $x_1+x_2 \geq -1/2$.

Det mista värde på a för vilket f är konvex på C_a är därmed $a = -1/2$.

Uppgift 5.(a) Problemet kan skrivas på formen:

$$\begin{aligned} \text{minimera} \quad & f(\mathbf{x}) \\ \text{då} \quad & g_1(\mathbf{x}) \le 0, \\ & g_2(\mathbf{x}) \le 0. \end{aligned}$$

$$\begin{aligned} \text{där} \quad f(\mathbf{x}) &= 4x_1^2 - 3x_2^2 + 5x_3^2 - 2x_4^2, \\ g_1(\mathbf{x}) &= x_1^2 + x_2^2 + x_3^2 + x_4^2 - 4, \\ g_2(\mathbf{x}) &= 1 - x_1^2 - x_2^2 - x_3^2 - x_4^2. \end{aligned}$$

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

(KKT1) $\dfrac{\partial f}{\partial x_j}(\mathbf{x}) + y_1 \dfrac{\partial g_1}{\partial x_j}(\mathbf{x}) + y_2 \dfrac{\partial g_2}{\partial x_j}(\mathbf{x}) = 0$, för $j = 1, 2, 3, 4$:
$2x_1\,(4 + y_1 - y_2) = 0$,
$2x_2\,(-3 + y_1 - y_2) = 0$,
$2x_3\,(5 + y_1 - y_2) = 0$,
$2x_4\,(-2 + y_1 - y_2) = 0$.

(KKT2) Tillåten punkt, dvs $g_i(\mathbf{x}) \le 0$ för $i = 1, 2$:
$x_1^2 + x_2^2 + x_3^2 + x_4^2 - 4 \le 0$,
$1 - x_1^2 - x_2^2 - x_3^2 - x_4^2 \le 0$.

(KKT3) Lagrangemultiplikatorerna icke-negativa:
$y_1 \ge 0$,
$y_2 \ge 0$.

(KKT4) Komplementaritetsvillkor, dvs $y_i g_i(\mathbf{x}) = 0$ för $i = 1, 2$:
$y_1\,(x_1^2 + x_2^2 + x_3^2 + x_4^2 - 4) = 0$,
$y_2\,(1 - x_1^2 - x_2^2 - x_3^2 - x_4^2) = 0$.

Uppgift 5.(b)

Ur (KKT4) följer att minst en av y_1 och y_2 måste vara $= 0$. .... **(*)**

Ur (KKT1) följer att minst tre av de fyra variablerna x_j måste vara $= 0$.
Å andra sidan kan inte alla fyra vara $= 0$ samtidigt, ty då uppfylls inte (KKT2).
Alltså är, i varje KKT-punkt, exakt tre av de fyra variablerna $x_j = 0$.

Antag först att $x_1 \neq 0$ medan $x_2 = x_3 = x_4 = 0$.
Då följer ur (KKT1), **(*)** och (KKT3) att $y_1 = 0$ och $y_2 = 4$,
varefter (KKT4) ger att $x_1 = 1$ eller -1.

Antag sedan att $x_2 \neq 0$ medan $x_3 = x_4 = x_1 = 0$.
Då följer ur (KKT1), **(*)** och (KKT3) att $y_1 = 3$ och $y_2 = 0$,
varefter (KKT4) ger att $x_2 = 2$ eller -2.

Antag nu att $x_3 \neq 0$ medan $x_4 = x_1 = x_2 = 0$.
Då följer ur (KKT1), **(*)** och (KKT3) att $y_1 = 0$ och $y_2 = 5$,
varefter (KKT4) ger att $x_3 = 1$ eller -1.

Antag slutligen att $x_4 \neq 0$ medan $x_1 = x_2 = x_3 = 0$.
Då följer ur (KKT1), **(*)** och (KKT3) att $y_1 = 2$ och $y_2 = 0$,
varefter (KKT4) ger att $x_4 = 2$ eller -2.

Problemet har alltså 8 KKT-punkter $\mathbf{x}^{(k)}$, med motsvarande målfunktionsvärden $f(\mathbf{x}^{(k)})$:

$$\begin{aligned}
\mathbf{x}^{(1)} &= (\,1\,,\,0\,,\,0\,,\,0\,)^\mathsf{T}, & f(\mathbf{x}^{(1)}) &= 4,\\
\mathbf{x}^{(2)} &= (-1\,,\,0\,,\,0\,,\,0\,)^\mathsf{T}, & f(\mathbf{x}^{(2)}) &= 4,\\
\mathbf{x}^{(3)} &= (\,0\,,\,2\,,\,0\,,\,0\,)^\mathsf{T}, & f(\mathbf{x}^{(3)}) &= -12,\\
\mathbf{x}^{(4)} &= (\,0\,,-2\,,\,0\,,\,0\,)^\mathsf{T}, & f(\mathbf{x}^{(4)}) &= -12,\\
\mathbf{x}^{(5)} &= (\,0\,,\,0\,,\,1\,,\,0\,)^\mathsf{T}, & f(\mathbf{x}^{(5)}) &= 5,\\
\mathbf{x}^{(6)} &= (\,0\,,\,0\,,-1\,,\,0\,)^\mathsf{T}, & f(\mathbf{x}^{(6)}) &= 5,\\
\mathbf{x}^{(7)} &= (\,0\,,\,0\,,\,0\,,\,2\,)^\mathsf{T}, & f(\mathbf{x}^{(7)}) &= -8,\\
\mathbf{x}^{(8)} &= (\,0\,,\,0\,,\,0\,,-2\,)^\mathsf{T}, & f(\mathbf{x}^{(8)}) &= -8.
\end{aligned}$$

Uppgift 5.(c)

Samtliga KKT-punkter ovan är tillåtna lösningar till problemet, pga (KKT2), och det följer från de beräknade målfunktionsvärdena att de enda av KKT-punkterna som kan vara globalt optimala lösningar är $\mathbf{x}^{(3)}$ och $\mathbf{x}^{(4)}$, båda med målfunktionsvärdet -12.

Att bevisa att $\mathbf{x}^{(3)}$ och $\mathbf{x}^{(4)}$ verkligen är globalt optimala lösningar är ekvivalent med att bevisa att det inte finns någon tillåten lösning $\mathbf{x}$ till problemet med $f(\mathbf{x}) < -12$, och det kan göras på exempelvis följande sätt:

Antag att $\mathbf{x} = (x_1, x_2, x_3, x_4)^\mathsf{T}$ är en tillåten lösning till problemet, dvs att

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 - 4 \leq 0 \quad \text{och} \quad 1 - x_1^2 - x_2^2 - x_3^2 - x_4^2 \leq 0.$$

Då är

$$\begin{aligned}
f(\mathbf{x}) &= 4x_1^2 - 3x_2^2 + 5x_3^2 - 2x_4^2\\
&\geq 4x_1^2 - 3x_2^2 + 5x_3^2 - 2x_4^2 + 3\,(x_1^2 + x_2^2 + x_3^2 + x_4^2 - 4)\\
&= 4x_1^2 - 3x_2^2 + 5x_3^2 - 2x_4^2 + 3x_1^2 + 3x_2^2 + 3x_3^2 + 3x_4^2 - 12\\
&= -12 + 7x_1^2 + 0x_2^2 + 8x_3^2 + x_4^2\\
&\geq -12.
\end{aligned}$$

Slutsatsen är att de båda KKT-punkterna $\mathbf{x}^{(3)} = (\,0\,,\,2\,,\,0\,,\,0\,)^\mathsf{T}$ och $\mathbf{x}^{(4)} = (\,0\,,-2\,,\,0\,,\,0\,)^\mathsf{T}$, bägge med målfunktionsvärdet -12, är globalt optimala lösningar till problemet.

Alternativt bevis av att $\mathbf{x}^{(3)}$ och $\mathbf{x}^{(4)}$ är globalt optimala lösningar:

Det tillåtna området är begränsat och dessutom slutet, eftersom bivillkorsfunktionerna är kontinuerliga. Vidare är målfunktionen kontinuerlig. Då vet vi att det existerar minst en globalt optimal lösning till problemet.

Varje punkt i tillåtna området är en regulär punkt, ty högst ett bivillkor är aktivt och gradienten till motsvarande bivillkorsfunktion är nollskild (så att gradienterna till de aktiva bivillkorsfunktionerna är linjärt oberoende). Varje globalt (eller lokalt) optimal lösning måste därmed vara en KKT-punkt.

Eftersom det endast finns ändligt många (åtta stycken) KKT-punkter så är de KKT-punkter, i detta fall $\mathbf{x}^{(3)}$ och $\mathbf{x}^{(4)}$, som har lägst målfunktionsvärde globalt optimala lösningar.