Lösningar till tentan i SF1861 Optimeringslära, 15 aug 2016

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:

X_{HF} = kvantitet (kg) hasselnötter som används för Familjeblandning under perioden,
X_{HL} = kvantitet (kg) hasselnötter som används för Lyxblandning under perioden,
X_{MF} = kvantitet (kg) mandel som används för Familjeblandning under perioden,
X_{ML} = kvantitet (kg) hasselnötter som används för Lyxblandning under perioden.

Målfunktion:

Försäljningsintäkten (kr) under perioden ges av $C_F\,(X_{HF}+X_{MF}) + C_L(X_{HL}+X_{ML})$.
Inköpskostnaden (kr) under perioden ges av $C_H\,(X_{HF}+X_{HL}) + C_M\,(X_{MF}+X_{ML})$.
Målfunktionen, som enligt problemlydelsen ska definieras som differensen mellan föräljningsintäkten och inköpskostnaden, ges då av
$C_F\,(X_{HF}+X_{MF}) + C_L(X_{HL}+X_{ML}) - C_H\,(X_{HF}+X_{HL}) - C_M\,(X_{MF}+X_{ML})$,
som ekvivalent kan skrivas
$(C_F-C_H)X_{HF} + (C_L-C_H)X_{HL} + (C_F-C_M)X_{MF} + (C_L-C_M)X_{ML}$.

Bivillkoren:

Givna inköpsbegränsningar från underleverantören:
$X_{HF}+X_{HL} \le 250$ och $X_{MF}+X_{ML} \le 300$.

Garanterat hasselnötsinnehåll i Familjeblandningen:
$0.3\,(X_{HF}+X_{MF}) \le X_{HF} \le 0.4\,(X_{HF}+X_{MF})$,
som ekvivalent kan skrivas som de två bivillkoren
$0.3\,X_{MF} - 0.7\,X_{HF} \le 0$ och $0.6\,X_{HF} - 0.4\,X_{MF} \le 0$.

Garanterat hasselnötsinnehåll i Lyxblandningen:
$0.5\,(X_{HL}+X_{ML}) \le X_{HL} \le 0.6\,(X_{HL}+X_{ML})$,
som ekvivalent kan skrivas som de två bivillkoren
$0.5\,X_{ML} - 0.5\,X_{HL} \le 0$ och $0.4\,X_{HL} - 0.6\,X_{ML} \le 0$.

Slutligen måste samtliga variabler vara icke-negativa.

Problemformulering:

$$
\begin{array}{rl}
\text{maximera} & (C_F-C_H)X_{HF} + (C_L-C_H)X_{HL} + (C_F-C_M)X_{MF} + (C_L-C_M)X_{ML} \\
\text{då} & X_{HF}+X_{HL} \le 250, \\
& X_{MF}+X_{ML} \le 300, \\
& 0.3\,X_{MF} - 0.7\,X_{HF} \le 0, \\
& 0.6\,X_{HF} - 0.4\,X_{MF} \le 0, \\
& 0.5\,X_{ML} - 0.5\,X_{HL} \le 0, \\
& 0.4\,X_{HL} - 0.6\,X_{ML} \le 0, \\
& X_{HF} \ge 0,\; X_{HL} \ge 0,\; X_{MF} \ge 0,\; X_{ML} \ge 0.
\end{array}
$$

Uppgift 1.(b)

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

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

Nu är $\mathbf{A}$ överförd till trappstegsform med *tre 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, 2 och 3 i $\mathbf{A}$.

De tre vektorerna $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ och $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ 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 sätta den variabel som inte svarar mot någon trappstegsetta, dvs x_4, till ett godtyckligt tal $t$och därefter bestämma hur "trappstegsvariablerna" 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
$x_1 = t,\ x_2 = -t,\ x_3 = -t$ och $x_4 = t$, som kan skrivas

$\mathbf{x} = \begin{pmatrix} x_1 \\ x_3 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \cdot t\,,$ för $t \in I\!R$. Vektorn $\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$ utgör alltså en bas till $\mathcal{N}(\mathbf{A})$.

Nu använder vi Gauss–Jordans metod på den transponerade matrisen $\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$.

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

Analogt med ovan följer att de tre första kolonnerna i $\mathbf{A}^{\mathsf{T}}$, dvs vektorerna

$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ och $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ utgör en bas till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$, medan $\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$ utgör en bas till $\mathcal{N}(\mathbf{A}^{\mathsf{T}})$.

Vi ser att var och en av basvektorerna till $\mathcal{R}(\mathbf{A})$ är ortogonal mot basvektorn till $\mathcal{N}(\mathbf{A}^{\mathsf{T}})$, och att var och en av basvektorerna till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$ är ortogonal mot basvektorn till $\mathcal{N}(\mathbf{A})$.
Detta är helt i överensstämmelse med de välkända relationerna
$\mathcal{R}(\mathbf{A})^{\perp} = \mathcal{N}(\mathbf{A}^{\mathsf{T}}),\ \mathcal{N}(\mathbf{A}^{\mathsf{T}})^{\perp} = \mathcal{R}(\mathbf{A}).\ \mathcal{R}(\mathbf{A}^{\mathsf{T}})^{\perp} = \mathcal{N}(\mathbf{A})$ och $\mathcal{N}(\mathbf{A})^{\perp} = \mathcal{R}(\mathbf{A}^{\mathsf{T}})$.

Uppgift 2.(a) Vi har ett LP-problem på standardformen

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

där $\mathbf{A} = \begin{bmatrix} 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 40 \\ 80 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (8,7,6,5,4,5,6,7,8)$.

Startlösningen ska ha basvariablerna x_1 och x_9, så att $\beta = (1,9)$ och $\nu = (2,3,4,5,6,7,8)$.

Motsvarande basmatris är $\mathbf{A}_\beta = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$, medan $\mathbf{A}_\nu = \begin{bmatrix} 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{bmatrix}$.

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 $\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 40 \\ 80 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av $\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu =$

$(7,6,5,4,5,6,7) - (1,1) \begin{bmatrix} 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{bmatrix} = (-1,-2,-3,-4,-3,-2,-1)$.

Eftersom $r_{\nu_4} = r_5 = -4$ 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 $\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_5 = \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$.

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

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i5}} \mid \bar{a}_{i5} > 0 \right\} = \min \left\{ \frac{5}{0.5}\,,\ \frac{10}{0.5} \right\} = \frac{5}{0.5} = \frac{\bar{b}_1}{\bar{a}_{15}}.$$

Minimerande index är $i = 1$, varför $x_{\beta_1} = x_1$ inte längre ska vara basvariabel.
Dess plats tas av x_5, så att $\beta = (5,9)$ och $\nu = (1,2,3,4,6,7,8)$.

Motsvarande basmatris är $\mathbf{A}_\beta = \begin{bmatrix} 4 & 0 \\ 4 & 8 \end{bmatrix}$, medan $\mathbf{A}_\nu = \begin{bmatrix} 8 & 7 & 6 & 5 & 3 & 2 & 1 \\ 0 & 1 & 2 & 3 & 5 & 6 & 7 \end{bmatrix}$.

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 $\begin{bmatrix} 4 & 0 \\ 4 & 8 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 40 \\ 80 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 4 & 4 \\ 0 & 8 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av $\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu =$

$(8,7,6,5,5,6,7) - (0,1) \begin{bmatrix} 8 & 7 & 6 & 5 & 3 & 2 & 1 \\ 0 & 1 & 2 & 3 & 5 & 6 & 7 \end{bmatrix} = (8,6,4,2,0,0,0)$.

Eftersom $\mathbf{r}_\nu \geq \mathbf{0}$ så är den aktuella tillåtna baslösningen, dvs $\mathbf{x} = (0,0,0,0,10,0,0,0,5)^\mathsf{T}$, en optimal lösning till P. Optimalvärdet är $\mathbf{c}^\mathsf{T}\mathbf{x} = 4 \cdot 10 + 8 \cdot 5 = 80$.

Uppgift 2.(b) När primala problemet P är på standardformen

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

så är det motsvarande duala problemet D på formen

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

som utskrivet blir:

$$\begin{array}{rrcrcl} \text{maximera} & 40y_1 & + & 80y_2 & & \\ \text{då} & 8y_1 & + & 0y_2 & \leq & 8, \\ & 7y_1 & + & 1y_2 & \leq & 7, \\ & 6y_1 & + & 2y_2 & \leq & 6, \\ & 5y_1 & + & 3y_2 & \leq & 5, \\ & 4y_1 & + & 4y_2 & \leq & 4, \\ & 3y_1 & + & 5y_2 & \leq & 5, \\ & 2y_1 & + & 6y_2 & \leq & 6, \\ & 1y_1 & + & 7y_2 & \leq & 7, \\ & 0y_1 & + & 8y_2 & \leq & 8. \end{array}$$

En noggrann figur i ett koordinatsystem med y_1 längs horisontella axeln och y_2 längs vertikala axeln visar *dels* att det tillåtna området till D enbart har två hörnpunkter, nämligen

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

dels att det tillåtna området till D är obegränsat och består av alla punkter *på* och *sydväst* om följande tre linjestycken:

1. Det vertikala linjestycket *rakt söder* om $\mathbf{y}^{(1)}$, (dvs $y_1 = 1$ och $y_2 \leq 0$),
2. Linjestycket mellan punkterna $\mathbf{y}^{(1)}$ och $\mathbf{y}^{(2)}$, (dvs $y_1 = 1-t$ och $y_2 = t$ för $t \in [\,0\,,1\,]$),
3. Det horisontella linjestycket *rakt väster* om $\mathbf{y}^{(2)}$, (dvs $y_2 = 1$ och $y_1 \leq 0$).

(Man kan notera att de fyra första bivillkoren samtliga är uppfyllda med likhet i $\mathbf{y}^{(1)}$ och med strikt olikhet i $\mathbf{y}^{(2)}$, att det femte bivillkoret är uppfyllt med likhet i både $\mathbf{y}^{(1)}$ och $\mathbf{y}^{(2)}$, och att de fyra sista bivillkoren samtliga är uppfyllda med likhet i $\mathbf{y}^{(2)}$ och med strikt olikhet i $\mathbf{y}^{(1)}$.)

Nivåkurvorna till målfunktionen $40y_1 + 80y_2$ är parallella linjer vinkelräta mot vektorn $(40, 80)^{\mathsf{T}}$, 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 $y_1 = 0$ och $y_2 = 1$, är alltså enligt figuren en optimal lösning till D.

Kontroll: Punkten $\mathbf{y}^{(2)}$ uppfyller alla bivillkoren i D och har det duala målfunktionsvärdet $\mathbf{b}^{\mathsf{T}}\mathbf{y}^{(2)} = 40 \cdot 0 + 80 \cdot 1 = 80 =$ det optimala primala målfunktionsvärdet i (a)-uppgiften. Alltså är $\mathbf{y}^{(2)}$ en optimal lösning till D.

Uppgift 2.(c):

Antag att x_5 och x_k ska vara basvariabler, dvs $\beta = (5\,,k)$, där $k \neq 5$.

Då ges basmatrisen av $\mathbf{A}_\beta = \begin{bmatrix} 4 & 9-k \\ 4 & k-1 \end{bmatrix}$.

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 $\begin{bmatrix} 4 & 9-k \\ 4 & k-1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 40 \\ 80 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \dfrac{1}{k-5} \begin{pmatrix} 15k-95 \\ 20 \end{pmatrix}$.

Baslösningen svarande mot $\beta = (5\,,k)$ ges alltså av

$x_5 = \dfrac{15k-95}{k-5}$, $x_k = \dfrac{20}{k-5}$ och $x_j = 0$ för $j \notin \{\,5\,,k\,\}$.

Detta är en *tillåten* baslösning om och endast om $x_5 \geq 0$ och $x_k \geq 0$,
vilket är uppfyllt om och endast om $k > 5$ och $15k \geq 95$,
vilket i sin tur är uppfyllt om och endast om $k \in \{\,7\,,8\,,9\,\}$.

Det finns alltså tre olika tillåtna baslösningar med x_5 som en av basvariablerna:

$\beta = (5,7)$ med $\mathbf{x}_\beta = (x_5\,,x_7)^\mathsf{T} = (5\,,10)^\mathsf{T}$,

$\beta = (5,8)$ med $\mathbf{x}_\beta = (x_5\,,x_8)^\mathsf{T} = (25/3\,,20/3)^\mathsf{T}$,

$\beta = (5,9)$ med $\mathbf{x}_\beta = (x_5\,,x_9)^\mathsf{T} = (10\,,5)^\mathsf{T}$.

Målfunktionsvärdena för dessa tre tillåtna baslösningar ges av

$c_5x_5 + c_kx_k = 4x_5 + (k-1)x_k = \dfrac{60k-380}{k-5} + \dfrac{20k-20}{k-5} = 80$,

vilket enligt (a)-uppgiften är optimalvärdet till P.

Alla de tre tillåtna baslösningarna ovan är alltså optimala baslösningar till P.

Uppgift 3.(a)

Problemet kan skrivas på formen:

$$\begin{aligned} \text{minimera} \quad & f(\mathbf{x}) \\ \text{då} \quad & g_i(\mathbf{x}) \le 0, \quad i = 1, 2, \end{aligned}$$

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

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 $j = 1, 2, 3, 4$:

$x_1 - 4y_1 - 1y_2 = 0,$
$x_2 - 3y_1 - 2y_2 = 0,$
$x_3 - 2y_1 - 3y_2 = 0,$
$x_4 - 1y_1 - 4y_2 = 0.$

(KKT2) Tillåten punkt, dvs $g_i(\mathbf{x}) \le 0$ för $i = 1, 2$:

$p - 4x_1 - 3x_2 - 2x_3 - 1x_4 \le 0,$
$q - 1x_1 - 2x_2 - 3x_3 - 4x_4 \le 0.$

(KKT3) Lagrangemultiplikatorerna icke-negativa:

$y_1 \ge 0,$
$y_2 \ge 0.$

(KKT4) Komplementaritetsvillkor, dvs $y_i g_i(\mathbf{x}) = 0$ för $i = 1, 2$:

$y_1(p - 4x_1 - 3x_2 - 2x_3 - 1x_4) = 0,$
$y_2(q - 1x_1 - 2x_2 - 3x_3 - 4x_4) = 0.$

Uppgift 3.(b)

Vi söker nu en lösning till (KKT1)–(KKT4) sådan att de båda olikhetsvillkoren i KKT2 är uppfyllda med likhet i $\mathbf{x}$. Då kommer de bägge ekvationerna i (KKT4) att bli automatiskt uppfyllda, så det räcker att söka en lösning till följande åtta villkor:

$$\begin{aligned} &x_1 - 4y_1 - 1y_2 = 0, \\ &x_2 - 3y_1 - 2y_2 = 0, \\ &x_3 - 2y_1 - 3y_2 = 0, \\ &x_4 - 1y_1 - 4y_2 = 0, \\ &p - 4x_1 - 3x_2 - 2x_3 - 1x_4 = 0, \\ &q - 1x_1 - 2x_2 - 3x_3 - 4x_4 = 0, \\ &y_1 \ge 0, \\ &y_2 \ge 0. \end{aligned}$$

Inför matrisen $\mathbf{A} = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}$ och vektorn $\mathbf{b} = \begin{pmatrix} p \\ q \end{pmatrix}$.

Då kan ovanstående åtta villkor skrivas

$$\begin{aligned} &\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{y} = \mathbf{0}, \\ &\mathbf{b} - \mathbf{A}\mathbf{x} = \mathbf{0}, \\ &\mathbf{y} \geq \mathbf{0}. \end{aligned}$$

Insättning av $\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{y}$ i $\mathbf{A}\mathbf{x} = \mathbf{b}$ ger ekvationssystemet $\mathbf{A}\mathbf{A}^{\mathsf{T}}\mathbf{y} = \mathbf{b}$, som i vårt fall blir

$\begin{bmatrix} 30 & 20 \\ 20 & 30 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$, med lösningen $\hat{\mathbf{y}} = \dfrac{1}{50} \begin{pmatrix} 3p-2q \\ 3q-2p \end{pmatrix}$.

Denna lösning uppfyller $\hat{\mathbf{y}} \geq \mathbf{0}$ om och endast om $3p \geq 2q$ och $3q \geq 2p$,

dvs om och endast om $\dfrac{2q}{3} \leq p \leq \dfrac{3q}{2}$.

För dessa värden på p och q ges motsvarande lösning $\hat{\mathbf{x}}$ av

$$\hat{\mathbf{x}} = \mathbf{A}^{\mathsf{T}}\hat{\mathbf{y}} = \begin{pmatrix} 4\hat{y}_1 + 1\hat{y}_2 \\ 3\hat{y}_1 + 2\hat{y}_2 \\ 2\hat{y}_1 + 3\hat{y}_2 \\ 1\hat{y}_1 + 4\hat{y}_2 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 2p-q \\ p \\ q \\ 2q-p \end{pmatrix}$$

Uppgift 3.(c)

Bivillkorsfunktionerna g_i är linjära, och därmed även konvexa.
Målfunktionen f är en kvadratisk funktion med Hessianen $\mathbf{H} = \mathbf{I} =$ enhetsmatrisen (4×4) som är positivt definit, vilket medför att f är en konvex funktion på $\mathbb{R}^4$.
P är alltså ett konvext optimeringsproblem. Därmed är (enligt en känd sats) varje KKT-punkt till P även en globalt optimal lösning till P.

Slutsats: Om $\dfrac{2q}{3} \leq p \leq \dfrac{3q}{2}$ så är $\hat{\mathbf{x}} = \dfrac{1}{10} \begin{pmatrix} 2p-q \\ p \\ q \\ 2q-p \end{pmatrix}$ en globalt optimal lösning till P.

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 $f(\mathbf{x}) = \frac{1}{4}\,x_1^4 - x_1 - x_1x_2 + x_1^2x_2^2$.

Gradienten till f ges av $\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} x_1^3 - 1 - x_2 + 2x_1x_2^2 \\ -x_1 + 2x_1^2x_2, \end{pmatrix}$,

och Hessianen till f ges av $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 3x_1^2 + 2x_2^2 & 4x_1x_2 - 1 \\ 4x_1x_2 - 1 & 2x_1^2 \end{bmatrix}$.

Startpunkten ges av $\mathbf{x}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, med $\nabla f(\mathbf{x}^{(1)})^\mathsf{T} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ och $\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$.

Eftersom $\mathbf{F}(\mathbf{x}^{(1)})$ är positivt definit ($3 > 0$, $2 > 0$ och $3 \cdot 2 - (-1)^2 > 0$) 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

$\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} \mathbf{d} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 1/5 \\ 3/5 \end{pmatrix}$.

Vi prövar först steget $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 6/5 \\ 3/5 \end{pmatrix}$.

Då blir $f(\mathbf{x}^{(2)}) = \cdots = -552/625 < -3/4 = f(\mathbf{x}^{(1)})$, så steget $t_1 = 1$ accepteras.

Därmed har vi utfört en fullständig iteration med Newtons metod och erhållit iterationspunkten $\mathbf{x}^{(2)} = \begin{pmatrix} 6/5 \\ 3/5 \end{pmatrix}$ med $f(\mathbf{x}^{(2)}) = -552/625$.

Uppgift 4.(b)
Först ska vi avgöra om $\tilde{\mathbf{x}} = (0\,, -1)^\mathsf{T}$ är en lokal minpunkt till f. Vi har att
$\nabla f(\tilde{\mathbf{x}})^\mathsf{T} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ och $\mathbf{F}(\tilde{\mathbf{x}}) = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}$ som *ej* är p.s.d. (ty $2 \cdot 0 - (-1)^2 < 0$).
Nödvändiga villkor för att $\tilde{\mathbf{x}}$ ska kunna vara en lokal minpunkt till f är att
$\nabla f(\tilde{\mathbf{x}})^\mathsf{T}$ är nollvektorn och $\mathbf{F}(\tilde{\mathbf{x}})$ är positivt semidefinit (p.s.d.).
Men eftersom $\mathbf{F}(\tilde{\mathbf{x}})$ *inte* är p.s.d. så kan $\tilde{\mathbf{x}}$ *inte* vara en lokal minpunkt till f,
och därmed inte heller en global minpunkt till f.

Sedan ska vi avgöra om $\hat{\mathbf{x}} = (1\,, 0.5)^\mathsf{T}$ är en lokal minpunkt till f. Vi har att
$\nabla f(\hat{\mathbf{x}})^\mathsf{T} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ och $\mathbf{F}(\hat{\mathbf{x}}) = \begin{bmatrix} 3.5 & 1 \\ 1 & 2 \end{bmatrix}$ som är p.d. (ty $3.5 > 0,\ 2 > 0,\ 3.5 \cdot 2 - 1^2 > 0$).
Tillräckliga villkor för att $\hat{\mathbf{x}}$ ska vara en lokal minpunkt till f är att
$\nabla f(\hat{\mathbf{x}})^\mathsf{T}$ är nollvektorn och $\mathbf{F}(\hat{\mathbf{x}})$ är positivt definit.
Eftersom båda dessa villkor är uppfyllda så *är* $\hat{\mathbf{x}}$ en lokal minpunkt till f.

Återstår frågan om $\hat{\mathbf{x}}$ är en global minpunkt till f, dvs om $f(\hat{\mathbf{x}}) \leq f(\mathbf{x})$ för alla $\mathbf{x} \in \mathbb{R}^2$.
Notera först att $\frac{1}{4}\,x_1^4 - x_1$ är en konvex funktion i x_1 med globalt minimum för $x_1 = 1$ (där derivatan är noll), vilket ger att $\frac{1}{4}\,x_1^4 - x_1 \geq \frac{1}{4} - 1 = -0.75$ för alla $\mathbf{x} \in \mathbb{R}^2$.
Notera sedan att $-x_1x_2 + x_1^2x_2^2 = (x_1x_2 - 0.5)^2 - 0.25 \geq -0.25$ för alla $\mathbf{x} \in \mathbb{R}^2$.
Tillsammans ger detta att $f(\mathbf{x}) \geq -0.75 - 0.25 = -1$ för alla $\mathbf{x} \in \mathbb{R}^2$.
Men med $\hat{\mathbf{x}} = (1\,, 0.5)^\mathsf{T}$ är $f(\hat{\mathbf{x}}) = -1$, så $\hat{\mathbf{x}}$ *är* en global minpunkt till f.

5.(a) Problemet

$$\text{P:}\quad \text{minimera}\ \tfrac{1}{2}\,\|\,\mathbf{x}-\mathbf{q}\,\|^2 = \tfrac{1}{2}\,(\mathbf{x}-\mathbf{q})^{\mathsf{T}}(\mathbf{x}-\mathbf{q})\ \ \text{då}\ \ \mathbf{a}^{\mathsf{T}}\mathbf{x} = b,$$

är ekvivalent med ett QP-problem på formen

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

med $\mathbf{H} = \mathbf{I}$ (enhetsmatrisen), $\mathbf{c} = -\mathbf{q}$, $\mathbf{A} = \mathbf{a}^{\mathsf{T}}$ och $\mathbf{b} = b$.

Lagrangevillkoren för P blir därför följande: $\mathbf{I}\,\mathbf{x} - \mathbf{a}\,u = \mathbf{q}$ och $\mathbf{a}^{\mathsf{T}}\mathbf{x} = b$.

Eftersom $\mathbf{I}$ är en diagonalmatris med strikt positiva diagonalelement så är $\mathbf{I}$ positivt definit, vilket medför att Lagrangevillkoren är såväl nödvändiga som tillräckliga för en globalt optimal lösning till P.

Insättning av $\mathbf{x} = \mathbf{q}+\mathbf{a}\,u$ i $\mathbf{a}^{\mathsf{T}}\mathbf{x} = b$ ger ekvationen $\mathbf{a}^{\mathsf{T}}\mathbf{a}\,u = b - \mathbf{a}^{\mathsf{T}}\mathbf{q}$, med lösningen $\hat{u} = \dfrac{b - \mathbf{a}^{\mathsf{T}}\mathbf{q}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}}$, varefter den optimala lösningen $\hat{\mathbf{x}}$ till P ges av $\hat{\mathbf{x}} = \mathbf{q} + \left(\dfrac{b - \mathbf{a}^{\mathsf{T}}\mathbf{q}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}}\right)\mathbf{a}$.

Då blir $\|\,\hat{\mathbf{x}}-\mathbf{q}\,\|^2 = (\hat{\mathbf{x}}-\mathbf{q})^{\mathsf{T}}(\hat{\mathbf{x}}-\mathbf{q}) = \dfrac{(b - \mathbf{a}^{\mathsf{T}}\mathbf{q})^2}{(\mathbf{a}^{\mathsf{T}}\mathbf{a})^2}\,\mathbf{a}^{\mathsf{T}}\mathbf{a} = \dfrac{(\mathbf{a}^{\mathsf{T}}\mathbf{q} - b)^2}{\mathbf{a}^{\mathsf{T}}\mathbf{a}}$, vilket skulle visas.

5.(b) Användning av formeln ovan, med $\mathbf{x}$ i stället för $\mathbf{q}$, ger att

$$d_1(\mathbf{x})^2 = x_1^2,\quad d_2(\mathbf{x})^2 = x_2^2,\quad d_3(\mathbf{x})^2 = x_3^2\ \ \text{och}$$

$$d_4(\mathbf{x})^2 = \frac{(2x_1 + x_2 + 2x_3 - 9)^2}{2\cdot 2 + 1\cdot 1 + 2\cdot 2} = (\,\tfrac{2}{3}\,x_1 + \tfrac{1}{3}\,x_2 + \tfrac{2}{3}\,x_3 - 3\,)^2.$$

Om vi inför $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$ och $\mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}$, så att $\mathbf{A}\mathbf{x}-\mathbf{b} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \frac{2}{3}\,x_1+\frac{1}{3}\,x_2+\frac{2}{3}\,x_3-3 \end{pmatrix}$,

så blir alltså $d_1(\mathbf{x})^2 + d_2(\mathbf{x})^2 + d_3(\mathbf{x})^2 + d_4(\mathbf{x})^2 = \|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2$,

och vårt problem blir följande minsta-kvadratproblem:

$$\text{minimera}\ \tfrac{1}{2}\,\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2\ \ \text{då}\ \mathbf{x} \in I\!R^3,$$

vilket är ekvivalent med normalekvationerna $\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{b}$, som i detta fall blir

$$\frac{1}{9}\begin{bmatrix} 13 & 2 & 4 \\ 2 & 10 & 2 \\ 4 & 2 & 13 \end{bmatrix}\begin{pmatrix} x_1 \\ x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

Med den föreslagna ansatsen $\begin{pmatrix} x_1 \\ x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2t \\ t \\ 2t \end{pmatrix}$ blir vänsterledet ovan $= \begin{pmatrix} 4t \\ 2t \\ 4t \end{pmatrix}$, som överensstämmer med högerledet ovan om $t = 0.5$. Optimal lösning till P är alltså $\hat{\mathbf{x}} = (1,\ 0.5,\ 1)^{\mathsf{T}}$.