Lösningar till 5B1752 Optimeringslära för E, 10/1–07

Uppgift 1.(a)

Inför följande beslutsvariabler:

x_j = antal kg av produkten P_j som blandas till per vecka.
y_i = antal kg av råvaran R_i som köps in per vecka.

Då kan företagets frågeställning formuleras som följande LP-problem i variablerna x_j och y_i:

$$\begin{array}{rl} \text{maximera} & \displaystyle\sum_{j=1}^{n} c_j x_j - \sum_{i=1}^{m} b_i y_i \\ \text{då} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j - y_i = 0\,, \quad i = 1, \dots, m \\ & 0 \le x_j \le d_j\,, \quad j = 1, \dots, n \\ & 0 \le y_i \le e_i\,, \quad i = 1, \dots, m. \end{array}$$

Om man vill så kan man eliminera variablerna y_i ur problemet, med hjälp av bivillkoren $y_i = \sum_{j=1}^{n} a_{ij} x_j$, varvid erhålls följande LP-problem i enbart variablerna x_j:

$$\begin{array}{rl} \text{maximera} & \displaystyle\sum_{j=1}^{n} k_j x_j \\ \text{då} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \le e_i\,, \quad i = 1, \dots, m \\ & 0 \le x_j \le d_j\,, \quad j = 1, \dots, n, \end{array}$$

där konstanterna k_j i målfunktionen ges av $k_j = c_j - \sum_{i=1}^{m} b_i a_{ij}$.

Den första formuleringen är nog att föredra. Det finns ingen anledning att välja en formulering med så få variabler som möjligt.

Uppgift 1.(b)

Att problemet är ett minkostnadsflödesproblem följer av att varje kolonn i $\mathbf{A}$ består av ett element $+1$, ett element -1, och resten 0 :or. Varje rad i $\mathbf{A}$ svarar då mot en nod i nätverket och varje kolonn i $\mathbf{A}$ svarar mot en båge i nätverket, nämligen en båge *från* den nod som svarar mot raden med $+1$ *till* den nod som svarar mot raden med -1.
Nätverket består alltså av 7 stycken noder samt bågmängden
$\mathcal{B} = \{(1,3),\ (1,4),\ (1,5),\ (1,6),\ (1,7), (2,3),\ (2,4),\ (2,5),\ (2,6),\ (2,7)\}$.
Noderna nr 1 och 2 är källnoder, var och en med påfyllningen 5 flödesenheter.
Noderna 3, 4, 5, 6 och 7 är sänknoder, var och en med avtappningen 2 flödesenheter.

Vi kallar fortsättningsvis variablerna för x_{ij} och motsvarande kostnader för c_{ij}, dvs
$\mathbf{x} = (x_{13}, x_{14}, x_{15}, x_{16}, x_{17}, x_{23}, x_{24}, x_{25}, x_{26}, x_{27})^\mathsf{T}$ och
$\mathbf{c} = (c_{13},\ c_{14},\ c_{15},\ c_{16},\ c_{17},\ c_{23},\ c_{24},\ c_{25},\ c_{26},\ c_{27})^\mathsf{T}$.

Den föreslagna lösningen uppfyller $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ och $\hat{\mathbf{x}} \geq \mathbf{0}$.
Vidare svarar den mot ett uppspännande träd i nätverket med basbågarna
$\mathcal{B}_\beta = \{(1,3),\ (1,6),\ (1,7),\ (2,4),\ (2,5),\ (2,6)\}$.
Den föreslagna lösningen $\hat{\mathbf{x}}$ är alltså en tillåten baslösning.
De reducerade kostnaderna ges nu ur formeln

$$r_{ij} = c_{ij} - y_i + y_j \quad \text{för alla icke-basbågar,}$$

där skalärerna (simplexmultiplikatorerna) y_i ges ur formeln

$$y_i - y_j = c_{ij} \quad \text{för alla basbågar samt } y_7 = 0.$$

Skalärerna y_i beräknas i exempelvis följande ordning:
Först sätts $y_7 = 0$, vilket gäller per definition.
Basbågen $(1,7)$ ger sedan att $y_1 - y_7 = c_{17}$, dvs $y_1 = c_{17} = 4$.
Basbågen $(1,6)$ ger sedan att $y_1 - y_6 = c_{16}$, dvs $y_6 = y_1 - c_{16} = 4 - 3 = 1$.
Basbågen $(1,3)$ ger sedan att $y_1 - y_3 = c_{13}$, dvs $y_3 = y_1 - c_{13} = 4 - 3 = 1$.
Basbågen $(2,6)$ ger sedan att $y_2 - y_6 = c_{26}$, dvs $y_2 = y_6 + c_{26} = 1 + 6 = 7$.
Basbågen $(2,5)$ ger sedan att $y_2 - y_5 = c_{25}$, dvs $y_5 = y_2 - c_{25} = 7 - 5 = 2$.
Basbågen $(2,4)$ ger slutligen att $y_2 - y_4 = c_{24}$, dvs $y_4 = y_2 - c_{24} = 7 - 4 = 3$.

Nästa steg är att beräkna reducerade kostnaderna för icke-basvariablerna, vilket ger
$r_{14} = c_{14} - y_1 + y_4 = 2 - 4 + 3 = 1$,
$r_{15} = c_{15} - y_1 + y_5 = 3 - 4 + 2 = 1$,
$r_{23} = c_{23} - y_2 + y_3 = 7 - 7 + 1 = 1$,
$r_{27} = c_{27} - y_2 + y_7 = 8 - 7 + 0 = 1$.

Eftersom alla $r_{ij} \geq 0$ så är den givna tillåtna baslösningen optimal.

Anmärkning: Problemet kan också lösas som ett transportproblem.

Uppgift 2.(a)

Inför slackvariabler x_4, x_5 och x_6 så att problemet blir 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 $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$ och $\mathbf{c} = (1,\ 2,\ 5,\ 0,\ 0,\ 0)^\mathsf{T}$.

I startlösningen ska enligt uppgiftslydelsen x_1, x_2 och x_3 vara basvariabler, dvs $\beta = (1,2,3)$ och $\delta = (4,5,6)$.

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

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 $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$.

Här använde vi den givna räknehjälpen.

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

dvs $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$.

Här använde vi igen den givna räknehjälpen.

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

$= (0,\ 0,\ 0) - (-1,\ 2,\ 3) \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = (-1,\ 2,\ 3).$

Eftersom $r_{\delta_1} = r_4 = -1$ är minst, och < 0, ska vi låta x_4 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_4$ ur systemet $\mathbf{A}_\beta \bar{\mathbf{a}}_4 = \mathbf{a}_4$,

dvs $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \\ \bar{a}_{34} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \\ \bar{a}_{34} \end{pmatrix} = \begin{pmatrix} -0.5 \\ -0.5 \\ 0.5 \end{pmatrix}$.

Här använde vi igen den givna räknehjälpen.

Det största värde som den nya basvariabeln x_4 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i4}} \mid \bar{a}_{i4} > 0 \right\} = \frac{\bar{b}_3}{\bar{a}_{34}} = \frac{2}{0.5}\ .$$

Minimerande index är $i = 3$, varför $x_{\beta_3} = x_3$ inte längre får vara kvar som basvariabel.

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

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

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 $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

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

$= (5,\ 0,\ 0) - (0,\ 1,\ 2) \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = (2,\ 1,\ 2)$.

Eftersom $\mathbf{r}_\delta \geq \mathbf{0}$ så är den aktuella baslösningen optimal.
Alltså är punkten $x_1 = 4$, $x_2 = 4$, $x_3 = 0$, $x_4 = 4$, $x_5 = 0$, $x_6 = 0$ optimal.
Optimalvärdet är $\mathbf{c}^\mathsf{T} \mathbf{x} = 12$.

Uppgift 2.(b)
Om primala problemet är på standardformen

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

så är det duala problemet

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

som utskrivet blir:

$$\begin{aligned} \text{maximera} \quad & 4y_1 + 4y_2 + 4y_3 \\ \text{då} \quad & y_1 + y_2 \leq 1, \\ & y_1 + y_3 \leq 2, \\ & y_2 + y_3 \leq 5, \\ & -y_1 \leq 0, \\ & -y_2 \leq 0, \\ & -y_3 \leq 0. \end{aligned}$$

Det är välkänt att en optimal lösning till detta duala problem ges av vektorn $\mathbf{y}$ med "simplexmultiplikatorerna" i optimala baslösningen i (a)-uppgiften, dvs $\mathbf{y} = (0,\ 1,\ 2)^\mathsf{T}$.
Man kontrollerar snabbt att detta är en tillåten lösning till det duala problemet.
Vidare är optimalvärdet $\mathbf{b}^\mathsf{T} \mathbf{y} = 12 =$ optimalvärdet för det primala problemet ovan.

Uppgift 3.(a)

Vi har att $\mathbf{h}(\mathbf{x}) = \begin{pmatrix} x_1^2 - x_2 - \delta_1 \\ x_1^2 + x_2 - \delta_2 \\ x_2^2 - x_1 - \delta_3 \\ x_2^2 + x_1 - \delta_4 \end{pmatrix}$ och $f(\mathbf{x}) = \frac{1}{2}\mathbf{h}(\mathbf{x})^\mathsf{T}\mathbf{h}(\mathbf{x}) \geq 0$ för alla $\mathbf{x} \in I\!R^2$.

Speciellt om alla $\delta_i = 0$ och $\hat{\mathbf{x}} = (0,\, 0)^\mathsf{T}$ så är $\mathbf{h}(\hat{\mathbf{x}}) = (0,\, 0,\, 0,\, 0)^\mathsf{T}$ och $f(\hat{\mathbf{x}}) = 0$.

Då är $f(\hat{\mathbf{x}}) \leq f(\mathbf{x})$ för alla $\mathbf{x} \in I\!R^2$, vilket innebär att $\hat{\mathbf{x}}$ är en global minpunkt till $f(\mathbf{x})$.

Uppgift 3.(b)

Nu är $\delta_1 = -0.1,\ \delta_2 = 0.1,\ \delta_3 = -0.2,\ \delta_4 = 0.2$ och $\mathbf{x}^{(1)} = (0, 0)^\mathsf{T}$.

Deriveringar ger att

$\nabla h_1(\mathbf{x}) = (2x_1,\, -1),\ \ \nabla h_2(\mathbf{x}) = (2x_1,\, 1),\ \ \nabla h_3(\mathbf{x}) = (-1,\, 2x_2),\ \ \nabla h_4(\mathbf{x}) = (\,1,\, 2x_2)$.

Därmed är $\nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 2x_1 & -1 \\ 2x_1 & 1 \\ -1 & 2x_2 \\ 1 & 2x_2 \end{bmatrix}$, så att $\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ -1 & 0 \\ 1 & 0 \end{bmatrix}$ och $\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} 0.1 \\ -0.1 \\ 0.2 \\ -0.2 \end{pmatrix}$.

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 $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ och $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -0.4 \\ -0.2 \end{pmatrix}$,

så ekvationssystemet blir $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.2 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$.

Vi prövar $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$.

Då blir
$h_1(\mathbf{x}^{(2)}) = 0.04 - 0.1 + 0.1 = 0.04$,
$h_2(\mathbf{x}^{(2)}) = 0.04 + 0.1 - 0.1 = 0.04$,
$h_3(\mathbf{x}^{(2)}) = 0.01 - 0.2 + 0.2 = 0.01$,
$h_4(\mathbf{x}^{(2)}) = 0.01 + 0.2 - 0.2 = 0.01$,

så att $f(\mathbf{x}^{(2)}) = 0.0017 < 0.05 = f(\mathbf{x}^{(1)})$. Steget $t_1 = 1$ gick alltså bra.

Därmed har vi utfört en fullständig iteration med Gauss-Newtons metod.

Uppgift 3.(c)

Nu är $\delta_1 = 0.1,\ \delta_2 = 0.1,\ \delta_3 = 0.2,\ \delta_4 = 0.2$ och $\mathbf{x}^{(1)} = (0,0)^\mathsf{T}$.

Då blir $\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ -1 & 0 \\ 1 & 0 \end{bmatrix}$ och $\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -0.1 \\ -0.1 \\ -0.2 \\ -0.2 \end{pmatrix}$ så att $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Ekvationssystemet i Gauss-Newtons metod blir då

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

Oavsett steglängd blir därmed $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ och algoritmen stannar.

Frågan är nu om $\mathbf{x}^{(2)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ är en lokal minpunkt till $f(\mathbf{x})$.

Vi har att gradienten $\nabla f(\mathbf{x}^{(2)})^\mathsf{T} = \nabla\mathbf{h}(\mathbf{x}^{(2)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(2)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

och hessianen $\mathbf{F}(\mathbf{x}^{(2)}) = \nabla\mathbf{h}(\mathbf{x}^{(2)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(2)}) + \sum_{i=1}^{4} h_i(\mathbf{x}^{(2)})\,\mathbf{H}_i(\mathbf{x}^{(2)}) =$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - 0.1\cdot\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} - 0.1\cdot\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} - 0.2\cdot\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} - 0.2\cdot\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1.6 & 0 \\ 0 & 1.2 \end{bmatrix}.$$

En symmetrisk 2×2-matris $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ är positivt definit om och endast om

$a > 0$, $c > 0$ och $ac - b^2 > 0$, vilket är uppfyllt för ovanstående hessian $\mathbf{F}(\mathbf{x}^{(2)})$.

Eftersom dessutom $\nabla f(\mathbf{x}^{(2)}) = (0,0)$ så är $\mathbf{x}^{(2)}$ (åtminstone) en lokal minpunkt till $f(\mathbf{x})$.

Uppgift 4.
Vi har ett kvadratiska optimeringsproblem med linjära olikhetsbivillkor på formen

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

där $\mathbf{H} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ och $\mathbf{b} = \begin{pmatrix} 3 \\ 6 \\ 7 \end{pmatrix}$.

Första iterationen: I den givna startpunkten är alla tre bivillkoret uppfyllda med likhet. Därför startar vi med $\alpha = (1, 2, 3)$ och γ tom.

Då är $\bar{\mathbf{x}} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 2 \\ 4 \\ 10 \end{pmatrix}$ och $\mathbf{A}_{\alpha} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

Vi får svaret "JA" i Steg 1, ty med hjälp av den givna räknehjälpen erhålls att

$\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_{\alpha}^{\mathsf{T}}\bar{\mathbf{u}}$ med $\bar{\mathbf{u}} = (-2,\ 4,\ 6)^{\mathsf{T}}$, så vi går till Steg 2.

Här konstateras att $\bar{u}_1 < 0$ (och minst), varför $\alpha_1 = 1$ flyttas över till γ–vektorn.

Sedan går vi till Steg 3 med $\alpha = (2, 3)$, $\gamma = (1)$ och $\mathbf{A}_{\alpha} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

I Steg 3 ska vi minimera $\tfrac{1}{2}\mathbf{d}^{\mathsf{T}}\mathbf{H}\mathbf{d} + (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})^{\mathsf{T}}\mathbf{d}$ under bivillkoret $\mathbf{A}_{\alpha}\mathbf{d} = \mathbf{0}$,

Optimalitetsvillkoren för detta konvexa QP-problem med likhetsbivillkor ges av

$\mathbf{H}\mathbf{d} - \mathbf{A}_{\alpha}^{\mathsf{T}}\mathbf{u} = -(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ och $\mathbf{A}_{\alpha}\mathbf{d} = \mathbf{0}$.

Eftersom $\mathbf{H} = 2\,\mathbf{I}$ och $\mathbf{c} = \mathbf{0}$ så ger de första ekvationerna att $\mathbf{d} = \tfrac{1}{2}\mathbf{A}_{\alpha}^{\mathsf{T}}\mathbf{u} - \bar{\mathbf{x}}$, som

insatt i $\mathbf{A}_{\alpha}\mathbf{d} = \mathbf{0}$ ger systemet $\tfrac{1}{2}\mathbf{A}_{\alpha}\mathbf{A}_{\alpha}^{\mathsf{T}}\mathbf{u} = \mathbf{A}_{\alpha}\bar{\mathbf{x}}$, dvs $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 7 \end{pmatrix}$.

Lösningen till detta system är $\mathbf{u} = \begin{pmatrix} 10/3 \\ 16/3 \end{pmatrix}$, varefter $\hat{\mathbf{d}} = \tfrac{1}{2}\mathbf{A}_{\alpha}^{\mathsf{T}}\mathbf{u} - \bar{\mathbf{x}} = \begin{pmatrix} 2/3 \\ 2/3 \\ -2/3 \end{pmatrix}$.

Eftersom $\bar{\mathbf{x}} + \hat{\mathbf{d}} = (5/3,\ 8/3,\ 13/3)^{\mathsf{T}}$ uppfyller alla bivillkor låter vi denna punkt bli nästa iterationspunkt $\bar{\mathbf{x}}$ och går till Steg 1.

Nu är $\alpha = (2, 3)$, $\gamma = (1)$, $\bar{\mathbf{x}} = \begin{pmatrix} 5/3 \\ 8/3 \\ 13/3 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 10/3 \\ 16/3 \\ 26/3 \end{pmatrix}$ och $\mathbf{A}_{\alpha} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

Vi får svaret "JA" i Steg 1, ty $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_{\alpha}^{\mathsf{T}}\bar{\mathbf{u}}$ med $\bar{\mathbf{u}} = \begin{pmatrix} 10/3 \\ 16/3 \end{pmatrix}$, så vi går till Steg 2.

Här konstateras att $\bar{\mathbf{u}} \geq \mathbf{0}$, varför den aktuella iterationspunkten

$\bar{\mathbf{x}} = \begin{pmatrix} 5/3 \\ 8/3 \\ 13/3 \end{pmatrix}$, tillsammans med vektorn $\hat{\mathbf{y}} = \begin{pmatrix} 0 \\ 10/3 \\ 16/3 \end{pmatrix}$, uppfyller optimalitetsvillkoren,

dvs $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}^{\mathsf{T}}\hat{\mathbf{y}}$, $\mathbf{A}\bar{\mathbf{x}} \geq \mathbf{b}$, $\hat{\mathbf{y}} \geq \mathbf{0}$ och $\hat{\mathbf{y}}^{\mathsf{T}}(\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}) = 0$. Vi stannar därför här.

Uppgift 5.

Problemet kan skrivas på formen: minimera $f(\mathbf{x})$ då $g_i(\mathbf{x}) \leq 0$, $i = 1, \ldots, 8$.

$f(\mathbf{x}) = (x_1 - q_1)^2 + (x_2 - q_2)^2 + (x_3 - q_3)^2$,
$g_1(\mathbf{x}) = -1 + x_1 + x_2 + x_3,\;\; g_2(\mathbf{x}) = -1 + x_1 + x_2 - x_3$,
$g_3(\mathbf{x}) = -1 + x_1 - x_2 + x_3,\;\; g_4(\mathbf{x}) = -1 + x_1 - x_2 - x_3$,
$g_5(\mathbf{x}) = -1 - x_1 + x_2 + x_3,\;\; g_6(\mathbf{x}) = -1 - x_1 + x_2 - x_3$,
$g_7(\mathbf{x}) = -1 - x_1 - x_2 + x_3,\;\; g_8(\mathbf{x}) = -1 - x_1 - x_2 - x_3$.

Problemet har tydligen en konvex kvadratisk målfunktion och 8 st linjära olikhetsbivillkor. Därmed är KKT-villkoren både nödvändiga och tillräckliga för en global optimallösning.

Lagrangefunktionen kan skrivas $L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \sum_{i=1}^{8} y_i g_i(\mathbf{x}) =$

$$
\begin{aligned}
&= (x_1 - q_1)^2 + (x_2 - q_2)^2 + (x_3 - q_3)^2 + \\
&+ y_1(-1 + x_1 + x_2 + x_3) + y_2(-1 + x_1 + x_2 - x_3) + \\
&+ y_3(-1 + x_1 - x_2 + x_3) + y_4(-1 + x_1 - x_2 - x_3) + \\
&+ y_5(-1 - x_1 + x_2 + x_3) + y_6(-1 - x_1 + x_2 - x_3) + \\
&+ y_7(-1 - x_1 - x_2 + x_3) + y_8(-1 - x_1 - x_2 - x_3).
\end{aligned}
$$

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

(KKT–1) $\partial L/\partial x_j = 0$ för $j = 1, 2, 3$:
$$
\begin{aligned}
&2(x_1 - q_1) + y_1 + y_2 + y_3 + y_4 - y_5 - y_6 - y_7 - y_8 = 0,\\
&2(x_2 - q_2) + y_1 + y_2 - y_3 - y_4 + y_5 + y_6 - y_7 - y_8 = 0,\\
&2(x_3 - q_3) + y_1 - y_2 + y_3 - y_4 + y_5 - y_6 + y_7 - y_8 = 0.
\end{aligned}
$$

(KKT–2) Tillåten punkt: $g_i(\mathbf{x}) \leq 0$, för $i = 1, \ldots, 8$.

(KKT–3) Lagrangemultiplikatorerna icke-negativa: $y_i \geq 0$, för $i = 1, \ldots, 8$.

(KKT–4) Komplementaritetsvillkor: $y_i g_i(\mathbf{x}) = 0$, för $i = 1, \ldots, 8$.

5.(a)

Här är $\mathbf{q} = (-0.5,\ 0.4,\ -0.4)^\mathsf{T}$.
I punkten $\hat{\mathbf{x}} = (-0.4,\ 0.3,\ -0.3)^\mathsf{T}$ är enbart det 6:e bivillkoret aktivt, medan övriga 7 st bivillkor är uppfyllda med strikt olikhet.
Samtliga y_i utom y_6 måste alltså vara $= 0$ i KKT-villkoren.
De återstående KKT-villkoren som måste uppfyllas i $\hat{\mathbf{x}}$ är då

$$
\begin{aligned}
2(-0.4 + 0.5) - y_6 &= 0 \qquad (\partial L/\partial x_1 = 0)\\
2(0.3 - 0.4) + y_6 &= 0 \qquad (\partial L/\partial x_2 = 0)\\
2(-0.3 + 0.4) - y_6 &= 0 \qquad (\partial L/\partial x_3 = 0)\\
y_6 &\geq 0 \qquad (\text{dual tillåtenhet})
\end{aligned}
$$

Samtliga dessa villkor uppfylls om $y_6 = 0.2$.
KKT-villkoren är alltså uppfyllda i $\hat{\mathbf{x}}$ som därmed är en global optimallösning.

5.(b)
Här är $\mathbf{q} = (-0.8,\ 0.6, -0.1)^{\mathsf{T}}$.
I punkten $\hat{\mathbf{x}} = (-0.6,\ 0.4,\ 0)^{\mathsf{T}}$ är enbart det 5:e och det 6:e bivillkoret aktiva,
medan övriga 6 st bivillkor är uppfyllda med strikt olikhet.
Samtliga y_i utom y_5 och y_6 måste alltså vara $= 0$ i KKT-villkoren.
De återstående KKT-villkoren som måste uppfyllas i $\hat{\mathbf{x}}$ är då

$$\begin{aligned}
2(-0.6 + 0.8) - y_5 - y_6 &= 0 \qquad && (\partial L/\partial x_1 = 0)\\
2(0.4 - 0.6) + y_5 + y_6 &= 0 && (\partial L/\partial x_2 = 0)\\
2(0 + 0.1) + y_5 - y_6 &= 0 && (\partial L/\partial x_3 = 0)\\
y_5 &\geq 0 && (\text{dual tillåtenhet})\\
y_6 &\geq 0 && (\text{dual tillåtenhet})
\end{aligned}$$

Samtliga dessa villkor uppfylls om $y_5 = 0.1$ och $y_6 = 0.3$.
KKT-villkoren är alltså uppfyllda i $\hat{\mathbf{x}}$ som därmed är en global optimallösning.

5.(c)
Här är $\mathbf{q} = (-0.7,\ 1.9, -0.5)^{\mathsf{T}}$.
I punkten $\hat{\mathbf{x}} = (0,\ 1,\ 0)^{\mathsf{T}}$ är 1:a, 2:a, 5:e och 6:e bivillkoren aktiva,
medan övriga 4 st bivillkor är uppfyllda med strikt olikhet.
Därmed måste $y_3 = y_4 = y_7 = y_8 = 0$ i KKT-villkoren.
De återstående KKT-villkoren som måste uppfyllas i $\hat{\mathbf{x}}$ är då

$$\begin{aligned}
2(0 + 0.7) + y_1 + y_2 - y_5 - y_6 &= 0 \qquad && (\partial L/\partial x_1 = 0)\\
2(1 - 1.9) + y_1 + y_2 + y_5 + y_6 &= 0 && (\partial L/\partial x_2 = 0)\\
2(0 + 0.5) + y_1 - y_2 + y_5 - y_6 &= 0 && (\partial L/\partial x_3 = 0)\\
y_1 &\geq 0 && (\text{dual tillåtenhet})\\
y_2 &\geq 0 && (\text{dual tillåtenhet})\\
y_5 &\geq 0 && (\text{dual tillåtenhet})\\
y_6 &\geq 0 && (\text{dual tillåtenhet})
\end{aligned}$$

Gauss-Jordans metod tillämpad på ekvationssystemet

$$\begin{aligned}
y_1 + y_2 - y_5 - y_6 &= -1.4\\
y_1 + y_2 + y_5 + y_6 &= 1.8\\
y_1 - y_2 + y_5 - y_6 &= -1.0
\end{aligned}$$

ger följande lösningar: $\begin{pmatrix} y_1 \\ y_2 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} -1.2 \\ 1.4 \\ 1.6 \\ 0.0 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$ för varje $t \in I\!R$.

Vi ser att om $1.2 \leq t \leq 1.4$ (t ex $t = 1.3$) så är samtliga $y_i \geq 0$, och då är
KKT-villkoren uppfyllda i $\hat{\mathbf{x}}$ som därmed är en global optimallösning.