Lösningar till 5B1712 Optimeringslära, 7 mars 2007

Uppgift 1.(a) Låt:

X_A = antal hektoliter Äppelmust som produceras per vecka.
X_P = antal hektoliter Päronmust som produceras per vecka.
X_B = antal hektoliter Blandmust som produceras per vecka.
X_C = antal hektoliter Cidermust som produceras per vecka.

Vi får då problemformuleringen

$$\begin{array}{rl} \text{maximera} & 196X_A + 210X_P + 280X_B + 442X_C \\ \text{då} & 1.6X_A + 1.8X_P + 3.2X_B + 5.4X_C \leq 80 \\ & 1.2X_A + 1.2X_P + 1.2X_B + 1.8X_C \leq 40 \\ & -0.8X_A + 0.2X_P + 0.2X_B + 0.2X_C \leq 0 \\ & -0.3X_A + 0.7X_P - 0.3X_B - 0.3X_C \leq 0 \\ & X_A \geq 0,\ X_P \geq 0,\ X_B \geq 0,\ X_C \geq 0. \end{array}$$

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 6 stycken noder samt bågmängden
$\mathcal{B} = \{(1,4),\ (1,5),\ (1,6),\ (2,4),\ (2,5),\ (2,6),\ (3,4),\ (3,5),\ (3,6)\}$.
Noderna nr 1, 2 och 3 är källnoder medan noderna nr 4, 5, 6 är sänknoder.

Vi kallar fortsättningsvis variablerna för x_{ij} och motsvarande kostnader för c_{ij},
dvs $\mathbf{x} = (x_{14}, x_{15}, x_{16}, x_{24}, x_{25}, x_{26}, x_{34}, x_{35}, x_{36})^\mathsf{T}$
och $\mathbf{c} = (c_{14},\ c_{15},\ c_{16},\ c_{24},\ c_{25},\ c_{26},\ c_{34},\ c_{35},\ c_{36})^\mathsf{T}$.

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
$\mathcal{B}_\beta = \{(1,4),\ (1,6),\ (2,6),\ (3,5),\ (3,6)\}$.
Den föreslagna lösningen $\hat{\mathbf{x}}$ är alltså en tillåten baslösning.
De reducerade kostnaderna ges nu ur formeln

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

där skalärerna (simplexmultiplikatorerna) y_i ges ur formeln

$$y_i - y_j = c_{ij} \ \text{ för alla basbågar samt } y_6 = 0.$$

Skalärerna y_i beräknas i exempelvis följande ordning:
Först sätts $y_6 = 0$, vilket gäller per definition.
Basbågen $(3,6)$ ger sedan att $y_3 - y_6 = c_{36}$, dvs $y_3 = c_{36} = 4$.
Basbågen $(2,6)$ ger sedan att $y_2 - y_6 = c_{26}$, dvs $y_2 = c_{26} = 4$.
Basbågen $(1,6)$ ger sedan att $y_1 - y_6 = c_{16}$, dvs $y_1 = c_{16} = 4$.
Basbågen $(1,4)$ ger sedan att $y_1 - y_4 = c_{14}$, dvs $y_4 = y_1 - c_{14} = 4 - 2 = 2$.
Basbågen $(3,5)$ ger sedan att $y_3 - y_5 = c_{35}$, dvs $y_5 = y_3 - c_{35} = 4 - 2 = 2$.

Nästa steg är att beräkna reducerade kostnaderna för icke-basvariablerna, vilket ger
$r_{15} = c_{15} - y_1 + y_5 = 3 - 4 + 2 = 1$,
$r_{24} = c_{24} - y_2 + y_4 = 3 - 4 + 2 = 1$,
$r_{25} = c_{25} - y_2 + y_5 = 3 - 4 + 2 = 1$,
$r_{34} = c_{34} - y_3 + y_4 = 3 - 4 + 2 = 1$.
Eftersom alla $r_{ij} \geq 0$ så är den givna tillåtna baslösningen optimal.

Anmärkning: Problemet kan också lösas som ett transportproblem!

Uppgift 2.(a)

Det aktuella LP-problemet är på standardformen

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

där $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 1 & 0 \\ 2 & 3 & 4 & 5 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (0,\, 0,\, 0,\, 0,\, 1,\, 1)$.

Den naturliga startbaslösningen har basvariablerna x_5 och x_6, vilket betyder att $\beta = (5, 6)$ och $\delta = (1, 2, 3, 4)$.

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, medan $\mathbf{A}_\delta = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix}$.

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 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna värden erhålls ur systemet $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,
dvs $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, 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
$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (0,\ 0,\ 0,\ 0) - (1,\ 1) \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix} = (-3,\ -5,\ -7,\ -9)$.

Eftersom $r_{\delta_4} = r_4 = -9$ är minst, och < 0, låter vi 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 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$.

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

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i4}} \mid \bar{a}_{i4} > 0 \right\} = \min \left\{ \frac{10}{4},\ \frac{12}{5} \right\} = \frac{12}{5} = \frac{\bar{b}_2}{\bar{a}_{24}}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_6$ inte längre får vara kvar som basvariabel. Dess plats tas av x_4.

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

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 1 & 4 \\ 0 & 5 \end{bmatrix}$, medan $\mathbf{A}_\delta = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 1 \end{bmatrix}$.

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 $\begin{bmatrix} 1 & 4 \\ 0 & 5 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 2.4 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 1 & 0 \\ 4 & 5 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -0.8 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (0,\ 0,\ 0,\ 1) - (1,\ -0.8) \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 1 \end{bmatrix} = (0.6,\ 0.4,\ 0.2,\ 0.8)$.

Eftersom $\mathbf{r}_\delta \geq \mathbf{0}$ så är den aktuella baslösningen optimal.

Därmed är punkten $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_4 = 2.4$, $x_5 = 0.4$, $x_6 = 0$ optimal.

Optimalvärdet ges av $x_5 + x_6 = 0.4 > 0$.

Om det finns en tillåten lösning $(\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)^\mathsf{T}$ till det ursprungliga systemet, så är $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4,\ 0,\ 0)^\mathsf{T}$ en tillåten lösning till LP-problemet med målfunktionsvärdet $= 0$. Eftersom målfunktionsvärdet till LP-problemet aldrig kan bli negativt så är då $\hat{\mathbf{x}}$ en optimal lösning till LP-problemet.

Alltså: Om det finns en tillåten lösning till det ursprungliga systemet, så är optimalvärdet till LP-problemet $= 0$. Men vi fann ovan att optimalvärdet till LP-problemet är 0.4, vilket medför att det ursprungliga systemet inte kan ha någon lösning.

Uppgift 2.(b)

Antag att $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \hat{y}_3)^\mathsf{T}$ är en optimal lösning till det duala problemet.

Från dualitetssatsen följer då att $\hat{y}_3 = \hat{x}_3 = -0.2$.

Vidare följer från komplementaritetssatsen att eftersom $\hat{x}_1 > 0$ så måste $-\hat{y}_1 + 3\hat{y}_2 + \hat{y}_3 = 0$ och eftersom $\hat{x}_2 > 0$ så måste $2\hat{y}_1 - 4\hat{y}_2 + \hat{y}_3 = 0$.

Tillsammans ger detta att $\hat{y}_1 = 0.7$ och $\hat{y}_2 = 0.3$, som också mycket riktigt uppfyller de övriga bivillkoren i det duala problemet.

Alltså: $\hat{y}_1 = 0.7$, $\hat{y}_2 = 0.3$ och $\hat{y}_3 = -0.2$ är en optimal lösning till det duala problemet (i själva verket den unika optimallösningen).

Uppgift 3

$f(\mathbf{x}) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2 = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 - 2x_3x_1 =$
$= \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}$ med $\mathbf{H} = \begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix}$.

Eftersom $f(\mathbf{x})$ är en summa av tre kvadrater så är $f(\mathbf{x}) \geq 0$ för alla $\mathbf{x} \in I\!R^3$, vilket medför att $\mathbf{H}$ är positivt semidefinit. (Men inte positivt definit eftersom $f(\mathbf{x}) = 0$ om $x_1 = x_2 = x_3$.)

Gauss-Jordan ger efter några enkla radoperationer att systemet $\mathbf{Ax} = \mathbf{b}$ är ekvivalent med systemet $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$. Ur detta följer att *en* tillåten lösning är $\bar{\mathbf{x}} = (2, 4, 0)^\mathsf{T}$.

Genom att sätta högerleden till 0 ovan så erhålls vidare att systemet $\mathbf{Az} = \mathbf{0}$ är ekvivalent med systemet $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \mathbf{z} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Ur detta följer att en bas till $\mathcal{N}(\mathbf{A})$ ges av den ensamma vektorn $\mathbf{z} = (1, -2, 1)^\mathsf{T}$. ($z_3 = 1 \Rightarrow z_1 = 1$ och $z_2 = -2$.)

Vi söker nu en optimal lösning till QP-problemet

$$\begin{array}{rl} \text{minimera} & \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} \\ \text{då} & \mathbf{Ax} = \mathbf{b}\,, \end{array}$$

Vi vet att $\mathbf{Ax} = \mathbf{b}$ är ekvivalent med att $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{z}\,v$ för $v \in I\!R$.

Insättning av detta uttryck i målfunktionen leder till följande optimeringsproblen i den ensamma variabeln v:

minimera $\frac{1}{2}\,v\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}v + \mathbf{z}^\mathsf{T}\mathbf{H}\bar{\mathbf{x}}v + \frac{1}{2}\,\bar{\mathbf{x}}^\mathsf{T}\mathbf{H}\bar{\mathbf{x}} = 18v^2 - 36v + 24$.

Detta är en konvex kvadratisk funktion som minimeras av $\hat{v} = 1$.

Optimal lösning till QP-problemet är därmed $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = (3,\, 2,\, 1)^\mathsf{T}$.

Eftersom $\mathbf{H}$ enligt ovan är positivt semidefinit så är $\hat{\mathbf{x}}$ en minpunkt till $\frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ om och endast om $\mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{0}$. Det finns alltså minst en minpunkt till $\frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ om och endast om ekvationssystemet $\mathbf{Hx} = -\mathbf{c}$ har minst en lösning.
Men detta system har minst en lösning om och endast om $\mathbf{c} \in \mathcal{R}(\mathbf{H}) = \mathcal{N}(\mathbf{H}^\mathsf{T})^\perp$.

Gauss-Jordan ger efter några enkla radoperationer att systemet $\mathbf{H}^\mathsf{T}\mathbf{z} = \mathbf{0}$ är ekvivalent med systemet $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{z} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Ur detta följer att en bas till $\mathcal{N}(\mathbf{H}^\mathsf{T})$ ges av $\mathbf{z} = (1,\, 1,\, 1)^\mathsf{T}$.

Låt oss kalla denna ensamma basvektor för $\mathbf{a}$, dvs $\mathbf{a} = (1,\, 1,\, 1)^\mathsf{T}$.

Då gäller att $\mathbf{c} \in \mathcal{N}(\mathbf{H}^\mathsf{T})^\perp$ om och endast om $\mathbf{c}$ är ortogonal mot denna basvektor $\mathbf{a}$.

Alltså: $\frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ har minst en minpunkt om och endast om $\mathbf{a}^\mathsf{T}\mathbf{c} = 0$.

Uppgift 4.(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 4.(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 och hamnat i punkten $\mathbf{x}^{(2)} = (0.2,\, 0.1)^\mathsf{T}$.

Gradienten av målfunktionen i denna punkt $\mathbf{x}^{(2)}$ ges av

$$\nabla f(\mathbf{x}^{(2)})^\mathsf{T} = \nabla\mathbf{h}(\mathbf{x}^{(2)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(2)}) = \begin{bmatrix} 0.4 & 0.4 & -1 & 1 \\ -1 & 1 & 0.2 & 0.2 \end{bmatrix}\begin{pmatrix} 0.04 \\ 0.04 \\ 0.01 \\ 0.01 \end{pmatrix} = \begin{pmatrix} 0.032 \\ 0.004 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eftersom gradienten inte är nollvektorn så kan $\mathbf{x}^{(2)}$ inte vara en lokal minpunkt.

Uppgift 5.(a) Vi har ett QP-problem på formen

$$\begin{array}{rl} \text{minimera} & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \end{array}$$

där $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 4 \\ 2 \\ 6 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ och $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Första iterationen: I den givna startpunkten är villkoren nr 1, 3 och 4 uppfyllda med likhet. Därför startar vi med $\alpha = (1,3,4)$ och $\gamma = (2)$.

Då är $\bar{\mathbf{x}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 5 \\ 2 \\ 6 \end{pmatrix}$, $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ och $\mathbf{A}_\alpha^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

Vi får svaret "JA" i Steg 1, ty $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_\alpha^\mathsf{T}\bar{\mathbf{u}}$ med $\bar{\mathbf{u}} = (5,\ -3,\ 1)^\mathsf{T}$, så vi går till Steg 2.

Här konstateras att $\bar{u}_2 < 0$ (och minst), varför $\alpha_2 = 3$ flyttas över till γ–vektorn.

Sedan går vi till Steg 3 med $\alpha = (1,4)$, $\gamma = (2,3)$, $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{A}_\gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

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 $\mathbf{H} = \mathbf{I}$ så ger de första ekvationerna att $\mathbf{d} = \mathbf{A}_\alpha^\mathsf{T}\mathbf{u} - \bar{\mathbf{x}} - \mathbf{c}$,

som insatt i $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$ ger ekvationssystemet $\mathbf{A}_\alpha\mathbf{A}_\alpha^\mathsf{T}\mathbf{u} = \mathbf{A}_\alpha(\bar{\mathbf{x}} + \mathbf{c})$;, dvs

$\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 6 \end{pmatrix}$, med lösningen $\mathbf{u} = \begin{pmatrix} 3.5 \\ 2.5 \end{pmatrix}$, varefter $\hat{\mathbf{d}} = \begin{pmatrix} -1.5 \\ 1.5 \\ 0 \end{pmatrix}$.

Eftersom $\bar{\mathbf{x}} + \hat{\mathbf{d}} = (-0.5,\ 1.5,\ 0)^\mathsf{T}$ inte uppfyller alla bivillkor beräknas
$\mathbf{s} = \mathbf{A}_\gamma\bar{\mathbf{x}} - \mathbf{b}_\gamma = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{g} = \mathbf{A}_\gamma\hat{\mathbf{d}} = \begin{pmatrix} -1.5 \\ 1.5 \end{pmatrix}$ och $\hat{t} = \min\limits_i \left\{ \dfrac{s_i}{-g_i} \mid g_i < 0 \right\} = \dfrac{1}{1.5} = \dfrac{s_2}{-g_2}$.

Sedan ändras $\bar{\mathbf{x}}$ till $\bar{\mathbf{x}} + \hat{t}\cdot\hat{\mathbf{d}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \dfrac{1}{1.5}\cdot\begin{pmatrix} -1.5 \\ 1.5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, medan γ_2 flyttas över till α–vektorn.

Ny iteration. Nu är $\alpha = (1,2,4)$, $\gamma = (3)$. Vidare är

$\bar{\mathbf{x}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 4 \\ 3 \\ 6 \end{pmatrix}$, $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ och $\mathbf{A}_\alpha^\mathsf{T} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

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

Här konstateras att $\bar{\mathbf{u}} \geq \mathbf{0}$, vilket betyder att den aktuella iterationspunkten är optimal, varvid algoritmen stannar. En optinal lösning till problemet är alltså $\hat{\mathbf{x}} = (0,1,0)^\mathsf{T}$.

Uppgift 5.(b)

Problemet är ett konvext QP-problem, och därmed är KKT-villkoren både nödvändiga och tillräckliga villkor för en global optimallösning.

Lagrangefunktionen till problemet kan skrivas:

$L(\mathbf{x}, \mathbf{y}) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2 + 4x_1 + 2x_2 + 6x_3 + y_1(1 - x_1 - x_2 - x_3) - y_2x_1 - y_3x_2 - y_4x_3.$

KKT-villkoren blir då:

$$
\begin{array}{rll}
x_1 - y_1 - y_2 = -4 & \partial L/\partial x_1 = 0 & \text{(KKT1)} \\
x_2 - y_1 - y_3 = -2 & \partial L/\partial x_2 = 0 & \text{(KKT2)} \\
x_3 - y_1 - y_4 = -6 & \partial L/\partial x_3 = 0 & \text{(KKT3)} \\
x_1 + x_2 + x_3 \geq 1 & \text{primal tillåtenhet} & \text{(KKT4)} \\
x_1 \geq 0 & \text{primal tillåtenhet} & \text{(KKT5)} \\
x_2 \geq 0 & \text{primal tillåtenhet} & \text{(KKT6)} \\
x_3 \geq 0 & \text{primal tillåtenhet} & \text{(KKT7)} \\
y_1 \geq 0 & \text{dual tillåtenhet} & \text{(KKT8)} \\
y_2 \geq 0 & \text{dual tillåtenhet} & \text{(KKT9)} \\
y_3 \geq 0 & \text{dual tillåtenhet} & \text{(KKT10)} \\
y_4 \geq 0 & \text{dual tillåtenhet} & \text{(KKT11)} \\
y_1(1 - x_1 - x_2 - x_3) = 0 & \text{komplementaritet} & \text{(KKT12)} \\
y_2x_1 = 0 & \text{komplementaritet} & \text{(KKT13)} \\
y_3x_2 = 0 & \text{komplementaritet} & \text{(KKT14)} \\
y_4x_3 = 0 & \text{komplementaritet} & \text{(KKT15)}
\end{array}
$$

Antag först att $\mathbf{x} = (1, 0, 0)^\mathsf{T}$. Då ger (KKT13) att $y_2 = 0$, varefter (KKT1)-(KKT3) ger att $y_1 = 5$, $y_3 = -3$ och $y_4 = 1$. Men detta strider mot (KKT10).
KKT-villkoren kan alltså inte uppfyllas med $\mathbf{x} = (1, 0, 0)^\mathsf{T}$.

Antag nu att $\mathbf{x} = \hat{\mathbf{x}} = (0, 1, 0)^\mathsf{T}$. Då ger (KKT13) att $y_3 = 0$, varefter (KKT1)-(KKT3) ger att $y_1 = 3$, $y_2 = 1$ och $y_4 = 3$. En snabb kontroll visar att alla KKT-villkoren nu är uppfyllda. KKT-villkoren uppfylls alltså av $\hat{\mathbf{x}} = (0, 1, 0)^\mathsf{T}$ och $\hat{\mathbf{y}} = (3, 1, 0, 3)^\mathsf{T}$.

Uppgift 5.(c)

Nu betraktar vi problemet att minimera $f(\mathbf{x})$ då $g_1(\mathbf{x}) \leq 0$ och $\mathbf{x} \in X$,

där $g_1(\mathbf{x}) = 1 - x_1 - x_2 - x_3$ och $X = \{\mathbf{x} \in I\!R^3 \mid \mathbf{x} \geq \mathbf{0}\}$.

Lagrangefunktionen till problemet kan nu skrivas:

$$L(\mathbf{x}, y_1) = \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 + \tfrac{1}{2}x_3^2 + 4x_1 + 2x_2 + 6x_3 + y_1(1 - x_1 - x_2 - x_3) =$$
$$= y_1 + \tfrac{1}{2}x_1^2 + (4 - y_1)x_1 + \tfrac{1}{2}x_2^2 + (2 - y_1)x_2 + \tfrac{1}{2}x_3^2 + (6 - y_1)x_3.$$

För att erhålla den duala målfunktionen $\varphi(y_1)$ ska man minimera $L(\mathbf{x}, y_1)$ med avseende på $\mathbf{x} \in X$. Som synes kan denna minimering utföras map varje enskild variabel x_j för sig.

Notera att om uttrycket $\frac{1}{2}x_j^2 + (c_j - y_1)x_j$ ska minimeras under kravet att $x_j \geq 0$ så ges det minimerande värdet på x_j av:

$x_j(y_1) = y_1 - c_j$ om $y_1 > c_j$ och $x_j(y_1) = 0$ om $y_1 \leq c_j$.

Detta kan kortfattat skrivas $x_j(y_1) = \max\{\, 0\,,\; y_1 - c_j \,\} = (y_1 - c_j)_+$

där den sista likheten utgör definitionen av $(y_1 - c_j)_+$.

Den duala målfunktionen ges nu av

$$\varphi(y_1) = \min_{\mathbf{x} \in X} L(\mathbf{x}, y_1) = L(\mathbf{x}(y_1), y_1) = y_1 +$$
$$+ \tfrac{1}{2}(y_1 - 4)_+^2 + (4 - y_1)(y_1 - 4)_+ + \tfrac{1}{2}(y_1 - 2)_+^2 + (2 - y_1)(y_1 - 2)_+ + \tfrac{1}{2}(y_1 - 6)_+^2 + (6 - y_1)(y_1 - 6)_+ =$$
$$= y_1 - \tfrac{1}{2}(y_1 - 4)_+^2 - \tfrac{1}{2}(y_1 - 2)_+^2 - \tfrac{1}{2}(y_1 - 6)_+^2 \,.$$

Från (a) och (b) gissar vi att $\hat{y}_1 = 3$.

Insättning ovan ger att $\varphi(3) = 3 - \frac{1}{2}(3 - 4)_+^2 - \frac{1}{2}(3 - 2)_+^2 - \frac{1}{2}(3 - 6)_+^2 = 2.5$.

Enligt (a) och (b) är $\hat{\mathbf{x}} = (0, 1, 0)^{\mathsf{T}}$ med $f(\hat{\mathbf{x}}) = 2.5$, dvs $f(\hat{\mathbf{x}}) = \varphi(\hat{y}_1)$.

Detta visar att gissningen $\hat{y}_1 = 3$ var korrekt.