

Lösningar till SF1852 Optimeringslära för E, 22/10–07

Uppgift 1.(a)

Inför följande variabler:

X_{HL} = antal kg hasselnöt som används till Lyxblandning.

X_{HF} = antal kg hasselnöt som används till Familjebländning.

X_{JL} = antal kg jordnöt som används till Lyxblandning.

X_{JF} = antal kg jordnöt som används till Familjebländning.

Nötmans vinstdå kan tecknas:

$$35(X_{HL} + X_{JL}) + 22(X_{HF} + X_{JF}) - 20(X_{HL} + X_{HF}) - 10(X_{JL} + X_{JF}) = \\ (35 - 20)X_{HL} + (35 - 10)X_{JL} + (22 - 20)X_{HF} + (22 - 10)X_{JF}.$$

Kravet att Lyxblandningen ska innehålla minst 65% hasselnötter kan skrivas
 $X_{HL} \geq 0.65(X_{HL} + X_{JL})$ eller, ekvivalent, $0.65X_{JL} - 0.35X_{HL} \leq 0$.

Kravet att Lyxblandningen ska innehålla högst 80% hasselnötter kan skrivas
 $X_{HL} \leq 0.80(X_{HL} + X_{JL})$ eller, ekvivalent, $0.20X_{HL} - 0.80X_{JL} \leq 0$.

Kravet att Familjebländningen ska innehålla minst 25% hasselnötter kan skrivas
 $X_{HF} \geq 0.25(X_{HF} + X_{JF})$ eller, ekvivalent, $0.25X_{JF} - 0.75X_{HF} \leq 0$.

Kravet att Familjebländningen ska innehålla högst 40% hasselnötter kan skrivas
 $X_{HF} \leq 0.40(X_{HF} + X_{JF})$ eller, ekvivalent, $0.60X_{HF} - 0.40X_{JF} \leq 0$.

Begränsningen av tillgången på hassel- resp jordnötter kan skrivas
 $X_{HL} + X_{HF} \leq 600$ och $X_{JL} + X_{JF} \leq 500$.

Slutligen måste alla fyra variablerna vara ≥ 0 .

Vi får alltså LP-problemet:

$$\begin{aligned} & \text{maximera } 15X_{HL} + 25X_{JL} + 2X_{HF} + 12X_{JF} \\ & \text{då } 0.65X_{JL} - 0.35X_{HL} \leq 0, \\ & \quad 0.20X_{HL} - 0.80X_{JL} \leq 0, \\ & \quad 0.25X_{JF} - 0.75X_{HF} \leq 0, \\ & \quad 0.60X_{HF} - 0.40X_{JF} \leq 0, \\ & \quad X_{HL} + X_{HF} \leq 600, \\ & \quad X_{JL} + X_{JF} \leq 500, \\ & \quad X_{HL} \geq 0, X_{HF} \geq 0, X_{JL} \geq 0, X_{JF} \geq 0. \end{aligned}$$

Uppgift 1.(b)

Först använder vi Gauss–Jordans metod på den givna matrisen

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Addition av -1 gånger första raden till andra raden ger till resultat matrisen

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Addition av 1 gånger andra raden till första raden, samt addition av -1 gånger andra raden till tredje raden, ger till resultat matrisen

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{T}.$$

Nu är \mathbf{A} överförd till trappstegsform med *två trappstegsettör*.

En bas till $\mathcal{R}(\mathbf{A})$ erhålls genom att som basvektorer välja de kolonner i \mathbf{A} som svarar mot trappstegsettör i \mathbf{T} , dvs kolonnerna 1 och 2 i \mathbf{A} .

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

En bas till $\mathcal{N}(\mathbf{A})$ kan bestämmas enligt följande:

Sätt $x_3 = 1$ (den enda variabel som inte svarar mot en trappstegssetta) och bestäm sedan x_1 och x_2 (variablerna svarande mot trappstegsettör)

så att $\mathbf{T}\mathbf{x} = \mathbf{0}$. Det ger den första och enda basvektorn $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ till $\mathcal{N}(\mathbf{A})$.

Nu använder vi Gauss–Jordans metod på matrisen

$$\mathbf{A}^T = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}.$$

Addition av 1 gånger första raden till andra raden ger till resultat matrisen

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}.$$

Addition av -1 gånger andra raden till första raden, samt addition av 1 gånger andra raden till tredje raden, ger till resultat matrisen

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \tilde{\mathbf{T}}.$$

Nu är \mathbf{A}^T överförd till trappstegsform med *två trappstegsettör*.

En bas till $\mathcal{R}(\mathbf{A}^T)$ erhålls genom att som basvektorer välja de kolonner i \mathbf{A}^T som svarar mot trappstegsettör i $\tilde{\mathbf{T}}$, dvs kolonnerna 1 och 2 i \mathbf{A}^T .

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

En bas till $\mathcal{N}(\mathbf{A}^T)$ kan bestämmas enligt följande:

Sätt $y_3 = 1$ (den enda variabel som inte svarar mot en trappstegsetta) och bestäm sedan y_1 och y_2 (variablerna svarande mot trappstegsettör)

så att $\tilde{\mathbf{T}}\mathbf{y} = \mathbf{0}$. Det ger den första och enda basvektorn $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ till $\mathcal{N}(\mathbf{A}^T)$.

Uppgift 2.(a)

Den i uppgiften föreslagna lösningen uppfyller balansekvationerna i alla noder och inga bågföden är negativa. Vidare svarar den mot ett uppspänande träd i nätverket, med basbågarna $B_\beta = \{(1, 3), (1, 4), (2, 4), (2, 5)\}$. Den föreslagna lösningen är alltså en tillåten baslösning.

De reducerade kostnaderna ges ur formeln

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

där skalärerna (simplexmultiplikatorerna) y_i ges ur formeln

$$y_i - y_j = c_{ij} \text{ för alla basbågar samt } y_5 = 0.$$

Skalärerna y_i beräknas i exempelvis följande ordning:

Först sätts $y_5 = 0$, vilket gäller per definition.

Basbågen $(2, 5)$ ger sedan att $y_2 - y_5 = c_{25}$, dvs $y_2 = 4$.

Basbågen $(2, 4)$ ger sedan att $y_2 - y_4 = c_{24}$, dvs $y_4 = 2$.

Basbågen $(1, 4)$ ger sedan att $y_1 - y_4 = c_{14}$, dvs $y_1 = 5$.

Basbågen $(1, 3)$ ger sedan att $y_1 - y_3 = c_{13}$, dvs $y_3 = 0$.

Nästa steg är att beräkna reducerade kostnaderna för icke-basvariablerna, vilket ger

$$r_{15} = c_{15} - y_1 + y_5 = 7 - 5 + 0 = 2,$$

$$r_{23} = c_{23} - y_2 + y_3 = 3 - 4 + 0 = -1.$$

Eftersom $r_{23} = -1$ ska vi låta x_{23} bli ny basvariabel.

Motvarande båge $(2, 3)$ bildar en slinga i nätverket tillsammans med träd-bågarna $(1, 3)$ (baklänges), $(1, 4)$ (framlänges) och $(4, 2)$ (baklänges).

Flödet i bågen $(2, 3)$, dvs x_{23} , kan öka till 200 innan en av baklängesbågarnas flöde har gått ner till 0, nämligen x_{24} . x_{23} blir alltså ny basvariabel i stället för x_{24} .

Den nya tillåtna baslösningen blir följande:

$$x_{13} = 200, \quad x_{14} = 500, \quad x_{23} = 200, \quad x_{25} = 600, \quad \text{övriga } x_{ij} = 0.$$

Nu beräknas skalärerna y_i svarande mot denna nya baslösning:

Först sätts $y_5 = 0$, vilket gäller per definition.

Basbågen $(2, 5)$ ger sedan att $y_2 - y_5 = c_{25}$, dvs $y_2 = 4$.

Basbågen $(2, 3)$ ger sedan att $y_2 - y_3 = c_{23}$, dvs $y_3 = 1$.

Basbågen $(1, 3)$ ger sedan att $y_1 - y_3 = c_{13}$, dvs $y_1 = 6$.

Basbågen $(1, 4)$ ger sedan att $y_1 - y_4 = c_{14}$, dvs $y_4 = 3$.

Nu beräknas reducerade kostnaderna för icke-basvariablerna, vilket ger

$$r_{15} = c_{15} - y_1 + y_5 = 7 - 6 + 0 = 1,$$

$$r_{24} = c_{24} - y_2 + y_4 = 2 - 4 + 3 = 1.$$

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

Optimalvärdet är $5 \cdot 200 + 3 \cdot 500 + 3 \cdot 200 + 4 \cdot 600 = 5500$.

Anmärkning: Problemet kan också lösas som ett transportproblem.

Uppgift 2.(b)

Flödesbalansvillkoren i de fem noderna kan skrivas $\mathbf{Ax} = \mathbf{b}$, där

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 700 \\ 800 \\ -400 \\ -500 \\ -600 \end{pmatrix} \text{ och } \mathbf{x} = \begin{pmatrix} x_{13} \\ x_{14} \\ x_{15} \\ x_{23} \\ x_{24} \\ x_{25} \end{pmatrix}.$$

Vektorn med kostnadskoefficienter ges av

$$\mathbf{c} = (c_{13}, c_{14}, c_{15}, c_{23}, c_{24}, c_{25})^T = (5, 3, 7, 3, 2, 4)^T.$$

Att bågarna är (enkel-)riktade ger kravet att $\mathbf{x} \geq \mathbf{0}$.

Uppgift 2.(c)

Om primala problemet är på standardformen

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

så är det duala problemet på formen: maximera $\mathbf{b}^T \mathbf{y}$ då $\mathbf{A}^T \mathbf{y} \leq \mathbf{c}$, som här blir:

$$\begin{aligned} &\text{maximera } 700y_1 + 800y_2 - 400y_3 - 500y_4 - 600y_5 \\ &\text{då } y_1 - y_3 \leq 5, \\ &\quad y_1 - y_4 \leq 3, \\ &\quad y_1 - y_5 \leq 7, \\ &\quad y_2 - y_3 \leq 3, \\ &\quad y_2 - y_4 \leq 2, \\ &\quad y_2 - y_5 \leq 4. \end{aligned}$$

Optimal lösning ges av (a)-uppgiften, dvs $y_1 = 6, y_2 = 4, y_3 = 1, y_4 = 3, y_5 = 0$.

Denna lösning uppfyller bivillkoren ovan och har duala målfunktionsvärdet

$$700 \cdot 6 + 800 \cdot 4 - 400 \cdot 1 - 500 \cdot 3 - 600 \cdot 0 = 5500 = \text{primala målfunktionsvärdet.}$$

Uppgift 3.

Kalla v_1 och v_2 för x_5 och x_6 . Då är Fas1-problemet ett LP-problem på standardformen

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

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

Startlösningen ska ha startbasvariablerna x_5 och x_6 , dvs $\beta = (5, 6)$ och $\delta = (1, 2, 3, 4)$.

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

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

$$\text{dvs } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}, \text{ med lösningen } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}.$$

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

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

Reducerade kostnaderna för icke-basvariablerna ges av

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

Eftersom $r_{\delta_2} = r_2 = -4$ är minst, och < 0 , ska vi låta x_2 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_2$ ur systemet $\mathbf{A}_\beta \bar{\mathbf{a}}_2 = \mathbf{a}_2$,

$$\text{dvs } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \text{ med lösningen } \bar{\mathbf{a}}_2 = \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Det största värde som den nya basvariabeln x_2 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i2}} \mid \bar{a}_{i2} > 0 \right\} = \min \left\{ \frac{6}{2}, \frac{8}{2} \right\} = \frac{6}{2} = \frac{\bar{b}_1}{\bar{a}_{12}}.$$

Minimerande index är $i = 1$, varför $x_{\beta_1} = x_5$ inte längre får vara kvar som basvariabel. Dess plats tas av x_2 .

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

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

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

$$\text{dvs } \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}, \text{ med lösningen } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

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

$$\mathbf{r}_\delta^\top = \mathbf{c}_\delta^\top - \mathbf{y}^\top \mathbf{A}_\delta = (0, 1, 0, 0) - (-1, 1) \begin{bmatrix} 2 & 1 & 1 & -1 \\ 1 & 0 & -1 & -2 \end{bmatrix} = (1, 2, 2, 1).$$

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

Därmed är punkten $x_1 = 0, x_2 = 3, x_3 = 0, x_4 = 0, x_5 = v_1 = 0, x_6 = v_2 = 2$ optimal.
Optimalvärdet är $v_1 + v_2 = 2 > 0$.

Vi kan nu dra slutsatsen att det inte finns någon lösning till det ursprungliga systemet av linjära ekvationer och olikheter. Argumentet är följande: Antag att det funnes en lösning (x_1, x_2, x_3, x_4) till det ursprungliga systemet. Denna lösning skulle då, tillsammans med $(v_1, v_2) = (0, 0)$, utgöra en tillåten lösning med målfunktionsvärdet $v_1 + v_2 = 0$ till Fas1-problemet. Men detta leder till en motsägelse, eftersom vi redan har konstaterat att optimalvärdet till Fas1-problemet är = 2.

Uppgift 4.(a)

Vi har ett problem på formen: minimera $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{H}\mathbf{x} + \mathbf{c}^T \mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$,

$$\text{där } \mathbf{A} = [1 \ 1 \ 1 \ 1], \ \mathbf{b} = 1, \ \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}, \ \mathbf{c} = \begin{pmatrix} -1 \\ -2 \\ -4 \\ -8 \end{pmatrix}.$$

Optimalitetsvillkoren ges av $\mathbf{H}\mathbf{x} - \mathbf{A}^T \mathbf{u} = -\mathbf{c}$ och $\mathbf{A}\mathbf{x} = \mathbf{b}$, vilket ger följande ekvationssystem i $\mathbf{x} \in \mathbb{R}^4$ och $\mathbf{u} \in \mathbb{R}$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 2 & 0 & 0 & -1 \\ 0 & 0 & 4 & 0 & -1 \\ 0 & 0 & 0 & 8 & -1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ u \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 8 \\ 1 \end{pmatrix}.$$

Efter några mycket enkla radoperationer erhålls det ekvivalenta systemet

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1/2 \\ 0 & 0 & 1 & 0 & -1/4 \\ 0 & 0 & 0 & 1 & -1/8 \\ 0 & 0 & 0 & 0 & 15/8 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ u \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -3 \end{pmatrix}$$

med den unika lösningen $u = -1.6$ och $\hat{\mathbf{x}} = (-0.6, 0.2, 0.6, 0.8)^T$.

Uppgift 4.(b)

Det är välkänt att om \mathbf{H} är *positivt semidefinit* så är de ovan använda optimalitetsvillkoren $\mathbf{H}\mathbf{x} - \mathbf{A}^T \mathbf{u} = -\mathbf{c}$ och $\mathbf{A}\mathbf{x} = \mathbf{b}$ båda nödvändiga och tillräckliga för en global optimallösning till minimeringsproblemet. Här är \mathbf{H} är en diagonalmatris med positiva diagonalelement, vilket medför att \mathbf{H} positivt definit och därmed även positivt semidefinit.

$\hat{\mathbf{x}} = (-0.6, 0.2, 0.6, 0.8)^T$ är alltså en global optimallösning till problemet.

Uppgift 5.(a)

Vi har att $\mathbf{h}(\mathbf{x}) = \begin{pmatrix} (x_1 + 2)^2 + x_2^2 - 5 \\ (x_1 - 2)^2 + x_2^2 - 3 \\ x_1^2 + (x_2 - 2)^2 - 6 \\ x_1^2 + (x_2 + 2)^2 - 2. \end{pmatrix}$ och $\nabla \mathbf{h}(\mathbf{x}) = \begin{bmatrix} 2x_1 + 4 & -2x_2 \\ 2x_1 - 4 & 2x_2 \\ 2x_1 & 2x_2 - 4 \\ 2x_1 & 2x_2 + 4 \end{bmatrix}.$

Vi ska starta i $\mathbf{x}^{(1)} = (0, 0)^\top$. Då är $\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -1 \\ 1 \\ -2 \\ 2 \end{pmatrix}$ och $\nabla \mathbf{h}(\mathbf{x}) = \begin{bmatrix} 4 & 0 \\ -4 & 0 \\ 0 & -4 \\ 0 & 4 \end{bmatrix}.$

I Gauss-Newtonens metod ska man lösa systemet $\nabla \mathbf{h}(\mathbf{x}^{(1)})^\top \nabla \mathbf{h}(\mathbf{x}^{(1)}) \mathbf{d} = -\nabla \mathbf{h}(\mathbf{x}^{(1)})^\top \mathbf{h}(\mathbf{x}^{(1)})$,

dvs $\begin{bmatrix} 32 & 0 \\ 0 & 32 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 8 \\ -16 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix}$.

Vi prövar $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1 \mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix}$. Då blir

$$\mathbf{h}(\mathbf{x}^{(2)}) = \begin{pmatrix} (9/4)^2 + (1/2)^2 - 5 \\ (7/4)^2 + (1/2)^2 - 3 \\ (1/4)^2 + (5/2)^2 - 6 \\ (1/4)^2 + (3/2)^2 - 2. \end{pmatrix} = \begin{pmatrix} 5/16 \\ 5/16 \\ 5/16 \\ 5/16 \end{pmatrix} \text{ och } f(\mathbf{x}^{(2)}) = 25/128 < 5 = f(\mathbf{x}^{(1)}),$$

så steget $t_1 = 1$ gick bra. Därmed har vi utfört en iteration med Gauss-Newtonens metod.

Uppgift 5.(b)

Vid Newtons metod, utgående från samma startpunkt $\mathbf{x}^{(1)} = (0, 0)^\top$ som ovan, ska man i stället lösa ekvationssystemet $\mathbf{F}(\mathbf{x}^{(1)}) \mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\top$, där $\mathbf{F}(\mathbf{x}^{(1)})$ är Hessianen av f i $\mathbf{x}^{(1)}$.

För icke linjära minsta-kvadratproblem kan detta system skrivas

$$\left(\nabla \mathbf{h}(\mathbf{x}^{(1)})^\top \nabla \mathbf{h}(\mathbf{x}^{(1)}) + \sum_{i=1}^m h_i(\mathbf{x}^{(1)}) \mathbf{H}_i(\mathbf{x}^{(1)}) \right) \mathbf{d} = -\nabla \mathbf{h}(\mathbf{x}^{(1)})^\top \mathbf{h}(\mathbf{x}^{(1)}),$$

där $\mathbf{H}_i(\mathbf{x}^{(1)})$ är Hessianen av h_i i $\mathbf{x}^{(1)}$.

I vårt fall är $\mathbf{H}_i(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ för $i = 1, 2, 3, 4$, oberoende av \mathbf{x} .

$$\text{Då blir } \sum_{i=1}^m h_i(\mathbf{x}^{(1)}) \mathbf{H}_i(\mathbf{x}^{(1)}) = \left(\sum_{i=1}^m h_i(\mathbf{x}^{(1)}) \right) \mathbf{H}_i(\mathbf{x}^{(1)}) = 0 \cdot \mathbf{H}_i(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Därmed är Newton-systemet $\mathbf{F}(\mathbf{x}^{(1)}) \mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\top$ identiskt med

Gauss-Newton-systemet $\nabla \mathbf{h}(\mathbf{x}^{(1)})^\top \nabla \mathbf{h}(\mathbf{x}^{(1)}) \mathbf{d} = -\nabla \mathbf{h}(\mathbf{x}^{(1)})^\top \mathbf{h}(\mathbf{x}^{(1)})$.

Därför blir nästa iterationspunkt $\mathbf{x}^{(2)}$ densamma för bågge metoderna.

Uppgift 5.(c) Det gäller här att avgöra om $\mathbf{F}(\mathbf{x})$ är positivt semidefinit för alla $\mathbf{x} \in \mathbb{R}^n$. En del kalkyler visar att $\mathbf{F}(\mathbf{x})$ faktiskt är positivt definit för alla $\mathbf{x} \in \mathbb{R}^n$, vilket betyder att f inte bara är konvex utan till och med strikt konvex på hela \mathbb{R}^n .

Uppgift 6.(a)

Problemet kan skrivas på formen: minimera $f(\mathbf{x})$ då $g_i(\mathbf{x}) \leq 0$ för $i = 1, 2$, där

$$f(\mathbf{x}) = x_1^2 + x_2^2 + (x_3 - 2)^2, \quad g_1(\mathbf{x}) = (x_1 - 1)^2 + x_2^2 + x_3^2 - 1, \quad g_2(\mathbf{x}) = x_1^2 + (x_2 - 1)^2 + x_3^2 - 1.$$

Motsvarande Lagrange-funktion ges då av $L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + y_1 g_1(\mathbf{x}) + y_2 g_2(\mathbf{x}) =$

$$x_1^2 + x_2^2 + (x_3 - 2)^2 + y_1((x_1 - 1)^2 + x_2^2 + x_3^2 - 1) + y_2(x_1^2 + (x_2 - 1)^2 + x_3^2 - 1).$$

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

(KKT-1) $\partial L / \partial x_j = 0$ för $j = 1, 2, 3$:

$$\begin{aligned} 2x_1 + 2y_1(x_1 - 1) + 2y_2x_1 &= 0, \\ 2x_2 + 2y_1x_2 + 2y_2(x_2 - 1) &= 0, \\ 2(x_3 - 2) + 2y_1x_3 + 2y_2x_3 &= 0, \end{aligned}$$

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

$$\begin{aligned} (x_1 - 1)^2 + x_2^2 + x_3^2 - 1 &\leq 0, \\ x_1^2 + (x_2 - 1)^2 + x_3^2 - 1 &\leq 0, \end{aligned}$$

(KKT-3) Lagrangemultiplikatorerna icke-negativa:

$$\begin{aligned} y_1 &\geq 0, \\ y_2 &\geq 0, \end{aligned}$$

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

$$\begin{aligned} y_1((x_1 - 1)^2 + x_2^2 + x_3^2 - 1) &= 0, \\ y_2(x_1^2 + (x_2 - 1)^2 + x_3^2 - 1) &= 0. \end{aligned}$$

Uppgift 6.(b) Antag först att $\mathbf{x} = (0, 0, 2)^\top$.

Då övergår KKT-1 till

$$\begin{aligned} -2y_1 &= 0, \\ -2y_2 &= 0, \\ 0 &= 0, \end{aligned}$$

som är uppfyllt för $y_1 = y_2 = 0$, medan KKT-2 övergår till

$$\begin{aligned} 1 + 0 + 4 - 1 &\leq 0, \\ 0 + 1 + 4 - 1 &\leq 0, \end{aligned}$$

som inte stämmer! $\mathbf{x} = (0, 0, 2)^\top$ är alltså ej en KKT-punkt.

Uppgift 6.(c) Antag nu att $\mathbf{x} = \left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right)^\top$.

Då övergår KKT-1 till

$$\begin{aligned} 2/3 - 4y_1/3 + 2y_2/3 &= 0, \\ 2/3 + 2y_1/3 - 4y_2/3 &= 0, \\ -8/3 + 4y_1/3 + 4y_2/3 &= 0, \end{aligned}$$

som är uppfyllt om och endast om $y_1 = 1$ och $y_2 = 1$.

KKT-2 övergår till

$$\begin{aligned} 4/9 + 1/9 + 4/9 - 1 &\leq 0, \\ 1/9 + 4/9 + 4/9 - 1 &\leq 0, \end{aligned}$$

som är uppfyllt eftersom bågge vänsterleden = 1.

Enligt ovan är $y_1 = y_2 = 1$, vilket gör att KKT-3 uppfyllt.

Slutligen är KKT-4 uppfyllt eftersom båda vänsterleden i KKT-2 är = 0.

$\mathbf{x} = \left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right)^T$ är alltså en KKT-punkt.

Uppgift 6.(d)

Såväl den kvadratiska målfunktionen f som de bågge kvadratiska bivillkorsfunktionerna g_i har andraderivatsmatrisen $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, som är en positivt definit matris.

Därmed är både målfunktionen och bivillkorsfunktionerna konvexa, varför det betraktade problemet är ett konvext optimeringsproblem. Vidare uppfyller exempelvis $\mathbf{x} = (0.5, 0.5, 0)^T$ samtliga bivillkor med strikt olikhet, så det betraktade problemet är ett *regulärt* konvext problem. Det betyder att en punkt \mathbf{x} är en globalt optimal lösning till problemet om och endast om \mathbf{x} är en KKT-punkt.

Men eftersom enligt ovan $\mathbf{x} = \left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right)^T$ är en KKT-punkt, så är därmed denna punkt en global optimallösning, vilket visar att det verkligen finns en global optimallösning.