Lösningar till SF1861/SF1851 Optimeringslära, 24/5 2013

Uppgift 1.(a) Inför de 13 variablerna

x_{ij} = kvantitet (i sorten ton) som fabrik nr i åläggs att tillverka av produkt nr j, samt

t = tiden (i sorten timmar) tills alla fabrikerna är färdiga med sina åläganden.

Kravet att uppfylla kundens beställning kan då formuleras $\sum_{i=1}^{3} x_{ij} = b_j$, för $j = 1, 2, 3, 4$.

De tre fabrikerna drivs förstås parallellt, så om t_i betecknar tiden då fabrik nr i är färdig med sitt åläggande så är $t = \max_i \{t_i\}$ = det största av de tre talen t_1, t_2 och t_3.

Återstår att uttrycka t_i i variablerna x_{ij}. Det finns två tänkbara modeller:

Modell I: Om det, vid varje fabrik, förhåller sig på det viset att de fyra produkterna tillverkas en i taget, dvs efter varandra, så är $t_i = \sum_{j=1}^{4} T_{ij} x_{ij}$, för $i = 1, 2, 3$, och då kan det aktuella planeringsproblemet formuleras som följande LP-problem:

$$
\begin{aligned}
\text{minimera} \quad & t \\
\text{då} \quad & t - \sum_{j=1}^{4} T_{ij} x_{ij} \geq 0, \quad \text{för } i = 1, 2, 3, \\
& \sum_{i=1}^{3} x_{ij} = b_j, \quad \text{för } j = 1, 2, 3, 4, \\
& x_{ij} \geq 0, \quad \text{för } i = 1, 2, 3 \text{ och } j = 1, 2, 3, 4.
\end{aligned}
$$

Modell II: Om det i stället, vid varje fabrik, förhåller sig på det viset att de fyra produkterna tillverkar parallellt, dvs oberoende av varandra, så är $t_i = \max_j \{T_{ij} x_{ij}\}$, varvid $t = \max_{i,j} \{T_{ij} x_{ij}\}$ = det största av de 12 talen $T_{11}x_{11}, \ldots, T_{34}x_{34}$, och då kan det aktuella planeringsproblemet formuleras som följande LP-problem:

$$
\begin{aligned}
\text{minimera} \quad & t \\
\text{då} \quad & t - T_{ij} x_{ij} \geq 0, \quad \text{för } i = 1, 2, 3 \text{ och } j = 1, 2, 3, 4. \\
& \sum_{i=1}^{3} x_{ij} = b_j, \quad \text{för } j = 1, 2, 3, 4, \\
& x_{ij} \geq 0, \quad \text{för } i = 1, 2, 3 \text{ och } j = 1, 2, 3, 4.
\end{aligned}
$$

Det framgår inte av problemformuleringen vilken av dessa båda modeller som bäst avbildar verkligheten, så båda ger full poäng på uppgiften.

Uppgift 1.(b) För att avgöra om $\mathbf{H}$ är positivt definit, eller positivt semidefinit, eller ingetdera, försöker vi $\mathrm{LDL}^{\mathsf{T}}$-faktorisera $\mathbf{H}$.

$$\mathbf{H} = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -3 \\ 2 & -3 & a \end{bmatrix}.$$

Addera först $+1$ gånger rad 1 till rad 2 och -2 gånger rad 1 till rad 3.
Addera sedan $+1$ gånger kolonn 1 till kolonn 2 och -2 gånger kolonn 1 till kolonn 3.
Det ger

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \text{ och } \mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & a-4 \end{bmatrix}.$$

Addera nu $+1$ gånger rad 2 till rad 3 och addera sedan $+1$ gånger kolonn 2 till kolonn 3.
Det ger

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ och } \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}\mathbf{E}_2^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a-5 \end{bmatrix}.$$

Därmed är $\mathrm{LDL}^{\mathsf{T}}$-faktoriseringen klar, och man har att

$$\mathbf{H} = \mathbf{L}\mathbf{D}\mathbf{L}^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a-5 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrisen $\mathbf{H}$ är positivt definit om och endast om diagonalmatrisen $\mathbf{D}$ är positivt definit, dvs om och endast om samtliga diagonalelement i $\mathbf{D}$ är > 0, dvs om och endast om $a > 5$.
Om $a = 5$ så är $\mathbf{D}$ positivt semidefinit, men inte positivt definit, vilket då gäller även för $\mathbf{H}$.
Om $a < 5$ så är $\mathbf{D}$ indefinit, vilket då gäller även för $\mathbf{H}$.

Alltså: $\mathbf{H}$ är positivt definit för $a > 5$ och positivt semidefinit för $a \geq 5$.

Uppgift 2.(a) Det givna 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 $\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 & -2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ och $\mathbf{c} = (4,\ 3,\ 6,\ 1)^{\mathsf{T}}$.

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

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

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

dvs $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 3.5 \end{pmatrix}$.

Reducerade kostnaderna icke-basvariablerna ges av

$r_3^{\mathsf{T}} = c_3 - \mathbf{y}^{\mathsf{T}} \mathbf{a}_3 = 6 - (0.5,\ 3.5) \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 1.5.$

$r_4^{\mathsf{T}} = c_4 - \mathbf{y}^{\mathsf{T}} \mathbf{a}_4 = 1 - (0.5,\ 3.5) \begin{pmatrix} -2 \\ 1 \end{pmatrix} = -1.5.$

Eftersom $r_4 = -1.5$ 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 $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} -0.5 \\ 1.5 \end{pmatrix}$.

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

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i4}} \mid \bar{a}_{i4} > 0 \right\} = \min \left\{ -\ ,\ \frac{0.5}{1.5} \right\}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_2$ ska lämna basen.

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

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

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

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

dvs $\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

Reducerade kostnaderna icke-basvariablerna ges nu av

$r_3 = c_3 - \mathbf{y}^\mathsf{T}\mathbf{a}_3 = 6 - (1,\ 3)\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 1$ och $r_2 = c_2 - \mathbf{y}^\mathsf{T}\mathbf{a}_2 = 3 - (1,\ 3)\begin{pmatrix} -1 \\ 1 \end{pmatrix} = 1$.

Eftersom både $r_3 \geq 0$ och $r_2 \geq 0$ är den aktuella baslösningen optimal.
Alltså är punkten $\hat{\mathbf{x}} = (2/3,\ 0,\ 0,\ 1/3)^\mathsf{T}$ en optimal lösning till problemet.
Optimalvärdet ges av $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = 8/3 + 1/3 = 3$.

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:

$$\begin{array}{lrcl} \text{maximera} & y_2 & & \\ \text{då} & y_1 + y_2 & \leq & 4, \\ & -y_1 + y_2 & \leq & 3, \\ & 2y_1 + y_2 & \leq & 6, \\ & -2y_1 + y_2 & \leq & 1. \end{array}$$

För att underlätta figurritandet kan man skriva om detta duala problem på formen

$$\begin{array}{rlcl} \text{maximera} & y_2 & & \\ \text{då} & y_2 & \leq & 4 - y_1, \\ & y_2 & \leq & 3 + y_1, \\ & y_2 & \leq & 6 - 2y_1, \\ & y_2 & \leq & 1 + 2y_1. \end{array}$$

Uppgift 2.(c)

Ur sin figur ser man att optimal lösning till duala problemet ges av skärningspunkten mellan de bägge linjerna $y_2 = 4 - y_1$ och $y_2 = 1 + 2y_1$, dvs $\hat{y}_1 = 1$ och $\hat{y}_2 = 3$.

Det är också välkänt att en optimal lösning till det duala problem ges av vektorn $\mathbf{y}$ med "simplexmultiplikatorerna" till den optimala baslösningen i (a)-uppgiften, vilket stämmer utmärkt med ovanstående $\hat{\mathbf{y}} = (1,\ 3)^\mathsf{T}$.
Man kontrollerar snabbt att detta är en tillåten lösning till det duala problemet.
Vidare ges optimalvärdet av $\mathbf{b}^\mathsf{T}\hat{\mathbf{y}} = \hat{y}_2 = 3 =$ optimalvärdet för det primala problemet ovan.

Uppgift 3.(a)

För att utreda huruvida det linjära ekvationssystemet $\mathbf{A}\mathbf{x} = \mathbf{b}$ har någon lösning eller ej utför vi radoperationer på matrisen $\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix}$.

Addition av (+1) gånger första raden till andra raden ger $\begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & -1 & 1 & -1 \end{bmatrix}$.

Addition av (+1) gånger andra raden till tredje raden ger $\begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Tredje raden motsvarar nu ekvationen $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 1$ som saknar lösning.

Alltså saknar det ursprungliga systemet $\mathbf{A}\mathbf{x} = \mathbf{b}$ lösning.

Uppgift 3.(b)

Att minimera $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$ är ekvivalent med att lösa "normalekvationerna" $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b}$.

Vi utför därför radoperationer på matrisen $\begin{bmatrix} \mathbf{A}^\mathsf{T}\mathbf{A} & \mathbf{A}^\mathsf{T}\mathbf{b} \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 2 \\ -1 & -1 & 2 & -2 \end{bmatrix}$.

Addition av (+0.5) gånger första raden till andra raden, samt
addition av (+0.5) gånger första raden till tredje raden, ger $\begin{bmatrix} 2 & -1 & -1 & 0 \\ 0 & 1.5 & -1.5 & 2 \\ 0 & -1.5 & 1.5 & -2 \end{bmatrix}$.

Addition av (+1) gånger andra raden till tredje raden, samt
addition av (+2/3) gånger andra raden till första raden, ger $\begin{bmatrix} 2 & 0 & -2 & 4/3 \\ 0 & 1.5 & -1.5 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Multiplikation av första raden med $\frac{1}{2}$ och av andra raden med $\frac{2}{3}$ ger $\begin{bmatrix} 1 & 0 & -1 & 2/3 \\ 0 & 1 & -1 & 4/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Den allmänna lösningen till normalekvationerna ges alltså av $\mathbf{x}(t) = (2/3 + t,\ 4/3 + t,\ t)^\mathsf{T}$, där t är ett godtyckligt reellt tal. För samtliga dessa är $\mathbf{A}\mathbf{x}(t) = (2/3,\ 2/3,\ -4/3)^\mathsf{T}$.

En lösning till normalekvationerna är alltså $\bar{\mathbf{x}} = (2/3,\ 4/3,\ 0)^\mathsf{T}$, med $\mathbf{A}\bar{\mathbf{x}} = (2/3,\ 2/3,\ -4/3)^\mathsf{T}$.
Då gäller ekvivalensen $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b} \;\Leftrightarrow\; \mathbf{A}\mathbf{x} = \mathbf{A}\bar{\mathbf{x}}$.

Vi ska nu minimera $\frac{1}{2}\|\mathbf{x}\|^2$ då $\mathbf{A}\mathbf{x} = \mathbf{A}\bar{\mathbf{x}}$.

Optimal lösning till detta ges av $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\bar{\mathbf{u}}$, där $\mathbf{A}\mathbf{A}^\mathsf{T}\bar{\mathbf{u}} = \mathbf{A}\bar{\mathbf{x}}$.

(Ty enligt Lagrangemetoden är $\hat{\mathbf{x}} \in I\!R^n$ en optimal lösning om och endast om $\hat{\mathbf{x}}$, tillsammans med en vektor $\bar{\mathbf{u}}$, uppfyller $\mathbf{I}\,\hat{\mathbf{x}} - \mathbf{A}^\mathsf{T}\bar{\mathbf{u}} = \mathbf{0}$ och $\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}\bar{\mathbf{x}}$. De första ekvationerna ger att $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\bar{\mathbf{u}}$, som insatt i de senare ekvationerna ger att $\mathbf{A}\mathbf{A}^\mathsf{T}\bar{\mathbf{u}} = \mathbf{A}\bar{\mathbf{x}}$.)

Vi utför därför radoperationer på matrisen $\left[\begin{array}{cc} \mathbf{A}\mathbf{A}^\mathsf{T} & \mathbf{A}\bar{\mathbf{x}} \end{array}\right] = \begin{bmatrix} 2 & -1 & -1 & 2/3 \\ -1 & 2 & -1 & 2/3 \\ -1 & -1 & 2 & -4/3 \end{bmatrix}$.

Addition av (+0.5) gånger första raden till andra raden, samt
addition av (+0.5) gånger första raden till tredje raden, ger $\begin{bmatrix} 2 & -1 & -1 & 2/3 \\ 0 & 1.5 & -1.5 & 1 \\ 0 & -1.5 & 1.5 & -1 \end{bmatrix}$.

Addition av (+1) gånger andra raden till tredje raden, samt
addition av (+2/3) gånger andra raden till första raden, ger $\begin{bmatrix} 2 & 0 & -2 & 4/3 \\ 0 & 1.5 & -1.5 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Multiplikation av första raden med $\frac{1}{2}$ och av andra raden med $\frac{2}{3}$ ger $\begin{bmatrix} 1 & 0 & -1 & 2/3 \\ 0 & 1 & -1 & 2/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Den allmänna lösningen till detta system ges alltså av $\mathbf{u}(t) = (2/3 + t,\ 2/3 + t,\ t)^\mathsf{T}$,
där t är ett godtyckligt reellt tal. För samtliga dessa är $\mathbf{A}^\mathsf{T}\mathbf{u}(t) = (0,\ 2/3,\ -2/3)^\mathsf{T}$.

En lösning (av flera) till systemet $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{A}\bar{\mathbf{x}}$ är alltså $\bar{\mathbf{u}} = (2/3,\ 2/3,\ 0)^\mathsf{T}$.

Då är $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\bar{\mathbf{u}} = (0,\ 2/3,\ -2/3)^\mathsf{T}$ minsta-normlösningen till minsta-kvadratproblemet.

Uppgift 3.(c)

Om $\bar{\mathbf{x}}$ är en lösning till normalekvationerna så gäller

dels att $\mathbf{A}\bar{\mathbf{x}}$ tillhör bildrummet för $\mathbf{A}$ (förstås!),

dels att "felvektorn" $\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}$ tillhör nollrummet för $\mathbf{A}^\mathsf{T}$
(ty normalekvationerna säger just att $\mathbf{A}^\mathsf{T}(\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = \mathbf{0}$),

dels att $\mathbf{A}\bar{\mathbf{x}} + (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = \mathbf{b}$.

Låt alltså $\mathbf{v} = \mathbf{A}\bar{\mathbf{x}} = (2/3,\ 2/3,\ -4/3)^\mathsf{T}$ och $\mathbf{w} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}} = (1/3,\ 1/3,\ 1/3)^\mathsf{T}$.

Uppgift 4.(a)

Betrakta ekvationssystemet

$$\begin{aligned} x_2 - x_1^2 &= 0.1\,, \\ x_2 + x_1^2 &= 0.1\,, \\ x_1 - x_2^2 &= 0.2\,, \\ x_1 + x_2^2 &= 0.2\,. \end{aligned}$$

Addition av de två första ekvationerna ger att $2x_2 = 0.2$, dvs $x_2 = 0.1$.
Addition av de två sista ekvationerna ger att $2x_1 = 0.4$, dvs $x_1 = 0.2$.
Men $(x_1, x_2) = (0.2,\, 0.1)$ är inte en lösning till ekvationssystemet, som därmed saknar lösning.

Uppgift 4.(b)

Vi har här att $\mathbf{h}(\mathbf{x}) = \begin{pmatrix} x_2 - x_1^2 - 0.1 \\ x_2 + x_1^2 - 0.1 \\ x_1 - x_2^2 - 0.2 \\ x_1 + x_2^2 - 0.2 \end{pmatrix}$. 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

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

så att $f(\mathbf{x}^{(2)}) = \frac{1}{2}(0.0016 + 0.0016 + 0.0001 + 0.0001) = 0.0017$,
medan $f(\mathbf{x}^{(1)}) = \frac{1}{2}(0.01 + 0.01 + 0.04 + 0.04) = 0.05$.
Steget $t_1 = 1$ gick alltså bra. Därmed har vi utfört en fullständig iteration med Gauss-Newtons metod och hamnat i punkten $\mathbf{x}^{(2)} = (0.2,\, 0.1)^\mathsf{T}$.

Uppgift 4.(c)
Vid Newtons metod ska man i stället lösa ekvationssystemet $\mathbf{F}(\mathbf{x}^{(1)})\,\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\mathsf{T}$.

För ickelinjära minsta-kvadratproblem kan detta system skrivas

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

I vårt fall är funktionerna h_i kvadratiska, så andraderivatsmatriserna är oberoende av $\mathbf{x}$ och ges av

$$\mathbf{H}_1(\mathbf{x}) = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}, \ \mathbf{H}_2(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \ \mathbf{H}_3(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}, \ \mathbf{H}_4(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

Med $\mathbf{h}(\mathbf{x}^{(1)})$ enligt ovan blir $\displaystyle\sum_{i=1}^{m} h_i(\mathbf{x}^{(1)}) \mathbf{H}_i(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, så vi råkar få exakt samma ekvationssystem, och därmed exakt samma nya iterationspunkt $\mathbf{x}^{(2)} = (0.2,\ 0.1)^\mathsf{T}$, som vi fick med Gauss-Newtons metod i förra deluppgiften.

Uppgift 4.(d) Andraderivatsmatrisen till målfunktionen $f(\mathbf{x})$ ges av

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

som efter insättning och förenklingar ger att

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 4x_1^2 + 2 & 0 \\ 0 & 4x_2^2 + 2 \end{bmatrix} + \begin{bmatrix} 4x_1^2 & 0 \\ 0 & 4x_2^2 \end{bmatrix} = \begin{bmatrix} 8x_1^2 + 2 & 0 \\ 0 & 8x_2^2 + 2 \end{bmatrix},$$

som är positivt definit för alla $\mathbf{x} \in I\!R^2$.

Därmed är $f(\mathbf{x})$ en (strikt) konvex funktion på hela $I\!R^2$.

(Detta är dock inte alltid fallet för ickelinjära minsta-kvadratproblem.)

Uppgift 5.(a)

Lagrangefunktionen kan skrivas $L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) + y_1 g_1(\mathbf{x}) + y_2 g_2(\mathbf{x}) =$

$= (x_1 - a_1)^4 + (x_2 - a_2)^4 + y_1((x_1-3)^4 + (x_2-1)^4 - 16) + y_2((x_1-1)^4 + (x_2-3)^4 - 16).$

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

(KKT–1) $\partial L/\partial x_j = 0$ för $j = 1,2$:
$4(x_1-a_1)^3 + 4y_1(x_1-3)^3 + 4y_2(x_1-1)^3 = 0,$
$4(x_2-a_2)^3 + 4y_1(x_2-1)^3 + 4y_2(x_2-3)^3 = 0.$

(KKT–2) Tillåten punkt, dvs $g_i(\mathbf{x}) \le 0$ för $i = 1,2$:
$(x_1-3)^4 + (x_2-1)^4 - 16 \le 0,$
$(x_1-1)^4 + (x_2-3)^4 - 16 \le 0.$

(KKT–3) Lagrangemultiplikatorerna icke-negativa:
$y_1 \ge 0,$
$y_2 \ge 0.$

(KKT–4) Komplementaritetsvillkor, dvs $y_i g_i(\mathbf{x}) = 0$ för $i = 1,2$:
$y_1((x_1-3)^4 + (x_2-1)^4 - 16) = 0,$
$y_2((x_1-1)^4 + (x_2-3)^4 - 16) = 0.$

Uppgift 5.(b) Andraderivatsmatriserna till målfunktionen och bivillkorsfunktionerna är

$$\begin{bmatrix} 12(x_1-a_1)^2 & 0 \\ 0 & 12(x_2-a_2)^2 \end{bmatrix}, \begin{bmatrix} 12(x_1-3)^2 & 0 \\ 0 & 12(x_2-1)^2 \end{bmatrix} \text{ och } \begin{bmatrix} 12(x_1-1)^2 & 0 \\ 0 & 12(x_2-3)^2 \end{bmatrix}.$$

Eftersom samtliga diagonalelement är ≥ 0 i dessa tre diagonalmatriser så är de tre matriserna positivt semidefinita för alla $\mathbf{x} \in I\!R^2$. Därmed är såväl målfunktion som bivillkorsfunktioner konvexa, vilket enligt en känd sats medför att varje KKT-punkt är en globalt optimal lösning till problemet

Uppgift 5.(c) Antag att $\mathbf{x} = (3,\, 3)^{\mathsf{T}}$.

Då är både $g_1(\mathbf{x}) = 0$ och $g_2(\mathbf{x}) = 0$, så (KKT–2) och (KKT–4) är uppfyllda.
(KKT–1) förenklas till $4(3-a_1)^3 + 4y_2(3-1)^3 = 0$ och $4(3-a_2)^3 + 4y_1(3-1)^3 = 0$,
som är uppfyllt om och endast om $y_1 = ((a_2-3)^3)/8$ och $y_2 = ((a_1-3)^3)/8$.
Då blir (KKT–3) uppfyllt om och endast om både $a_1 \ge 3$ och $a_2 \ge 3$.

Alltså: $\mathbf{x} = (3,\, 3)^{\mathsf{T}}$ är en KKT-punkt om och endast om både $a_1 \ge 3$ och $a_2 \ge 3$.

Uppgift 5.(d) Antag nu att $\mathbf{x} = (3-2^{3/4},\, 1+2^{3/4})^{\mathsf{T}}$.

Då är $g_1(\mathbf{x}) = 0$ och $g_2(\mathbf{x}) < 0$, så (KKT–4) medför att $y_2 = 0$, varvid både (KKT–2) och (KKT–4) blir uppfyllda. (KKT–1) förenklas nu till

$4(3-2^{3/4}-a_1)^3 + 4y_1(-2^{3/4})^3 = 0$ och $4(1+2^{3/4}-a_2)^3 + 4y_1(2^{3/4})^3 = 0,$

Addition av dessa villkor ger att $4(3-2^{3/4}-a_1)^3 = -4(1+2^{3/4}-a_2)^3$, vilket är uppfyllt om och endast om $a_1 + a_2 = 4$. Vi sätter därför $a_1 = 2-t$ och $a_2 = 2+t$.

Då blir $y_1 = (1-2^{3/4}+t)^3/(2^{3/4})^3$, så (KKT–3) blir uppfyllt om och endast om $t \ge 2^{3/4}-1$.

Alltså: $\mathbf{x} = (3-2^{3/4},\, 1+2^{3/4})^{\mathsf{T}}$ är en KKT-punkt om och endast om $a_1 = 2-t$ och $a_2 = 2+t$, där $t \ge 2^{3/4}-1$.