Lösningar till SF1861 Optimeringslära för T, 22/5-08

Uppgift 1.(a)

I ett koordinatsystem med x_1 och x_2 på axlarna blir det tillåtna området en fyrhörning med hörnen i punkterna $(4,0)$, $(2,1)$, $(1,2)$ och $(0,4)$.

Nivåkurvorna till målfunktionen $3x_1 + 2x_2$ är parallella linjer vinkelräta mot vektorn $(3,2)^\mathsf{T}$. Ju längre åt "sydväst" en nivåkurva ligger, desto lägre målfunktionsvärde svarar den mot. Det följer ur figuren att den unika optimala lösningen ges av hörnpunkten $(1,2)$.

Nivåkurvorna till målfunktionen $(x_1 - 3)^2 + (x_2 - 1)^2$ är cirklar med mittpunkten i $(3,1)$. Ju mindre cirkel, desto lägre målfunktionsvärde svarar den mot.
Speciell är punkten $(3,1)$ en minpunkt till målfunktionen om man bortser från bivillkoren. Men ur figuren (eller analytiskt) ser vi att $(3,1)$ faktiskt ligger i det tillåtna området, och därmed är denna punkt den unika optimala lösningen även till problemet med bivillkor.

Uppgift 1.(b)

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

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 \end{bmatrix}, \; \mathbf{b} = \begin{pmatrix} 10 \\ 30 \\ 0 \\ -15 \\ -25 \end{pmatrix} \text{ och } \mathbf{x} = \begin{pmatrix} x_{12} \\ x_{13} \\ x_{14} \\ x_{23} \\ x_{25} \\ x_{34} \\ x_{35} \\ x_{45} \end{pmatrix}.$$

Här betecknar variabeln x_{ij} flödet i bågen som går från nod i till nod j.

Vektorn med kostnadskoefficienter ges av

$\mathbf{c} = (c_{12},\, c_{13},\, c_{14},\, c_{23},\, c_{25},\, c_{34},\, c_{35},\, c_{45})^\mathsf{T} = (2,\, 2,\, 2,\, 1,\, 3,\, 1,\, 1,\, 2)^\mathsf{T}$.

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

Uppgift 2.(a)

Vi har ett LP-problem 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 $\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (1,\, 3,\, 2,\, 1)$.

Den förslagna lösningen har basvariablerna x_1 och x_2, dvs $\beta = (1, 2)$ och $\delta = (3, 4)$.

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

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 $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix}$,

dvs $x_1 = 1.5$ och $x_2 = 0.5$, vilket stämmer med förslaget.

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

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

Reducerade kostnaderna för icke-basvariablerna ges av

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (2,\, 1) - (-1,\, 2) \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} = (5,\, 2)$.

Eftersom reducerade kostnaderna är icke-negativa så är den föreslagna lösningen optimal.

$x_1 = 1.5$, $x_2 = 0.5$, övriga $x_j = 0$. Optimalvärdet $= 3$.

Uppgift 2.(b)

Efter införande av slackvariabler x_5 och x_6 får vi ett nytt LP-problem 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 nu $\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & -1 & -1 & 0 \\ 1 & 1 & -1 & -1 & 0 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (1,\, 3,\, 2,\, 1,\, 0,\, 0)$.

Vi startar från baslösningen ovan, med $\beta = (1, 2)$, som är en tillåten baslösning även till detta nya problem.

Vektorerna $\bar{\mathbf{b}}$ och $\mathbf{y}$ blir desamma som ovan.

Reducerade kostnaderna för icke-basvariablerna ges nu av

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (2,\, 1,\, 0,\, 0) - (-1,\, 2) \begin{bmatrix} 1 & -1 & -1 & 0 \\ -1 & -1 & 0 & -1 \end{bmatrix} = (5,\, 2, -1,\, 2)$.

Eftersom $r_{\delta_3} = r_5 = -1$ 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 $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_1 = \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_1 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i5}} \mid \bar{a}_{i5} > 0 \right\} = \frac{0.5}{0.5} = \frac{\bar{b}_2}{\bar{a}_{25}}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_2$ inte längre får vara kvar som basvariabel. Dess plats tas av x_5.

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

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

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

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

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

Reducerade kostnaderna för icke-basvariablerna ges nu av

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

Eftersom reducerade kostnaderna är icke-negativa så är den aktuella lösningen optimal.

$x_1 = 2$, $x_5 = 1$, övriga $x_j = 0$.

Optimalvärdet $= 2$.

Uppgift 3.(a)

Målfunktionen är $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$, där $\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ och $\mathbf{c} = \begin{pmatrix} -3 \\ -5 \\ -4 \\ -3 \end{pmatrix}$.

Punkten $\bar{\mathbf{x}}$ är en global minpunkt till $f(\mathbf{x})$ om och endast om $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{0}$ och $\mathbf{H}$ är positivt semidefinit. Om $\mathbf{H}$ dessutom är positivt definit så är $\bar{\mathbf{x}}$ den unika minpunkten.
Insättning av $\bar{\mathbf{x}} = (2, 1, 1, 1)^{\mathsf{T}}$ visar att $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{0}$.
För att avgöra om $\mathbf{H}$ är positivt definit, eller positivt semidefinit, eller ingetdera, kan man exemplvis använda vanlig kvadratkomplettering.
Här väljer vi i stället att försöka $\mathrm{LDL}^{\mathsf{T}}$-faktorisera $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Addera först -1 gånger rad 1 till rad 2 och addera sedan -1 gånger kolonn 1 till kolonn 2. Det ger

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

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

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ och } \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}\mathbf{E}_2^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Addera nu -1 gånger rad 3 till rad 4 och addera sedan -1 gånger kolonn 3 till kolonn 4. Det ger

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \text{ och } \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}\mathbf{E}_2^{\mathsf{T}}\mathbf{E}_3^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

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

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

Alla diagonalelement i $\mathbf{D}$ är strikt positiva, vilket medför att $\mathbf{H}$ är positivt definit.

Därmed är $\bar{\mathbf{x}} = (2, 1, 1, 1)^{\mathsf{T}}$ den unika globala minpunkten till $f(\mathbf{x})$.

Uppgift 3.(b)

Vi ska här minimera $\frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ under bivillkoret att $(-\mathbf{c})^{\mathsf{T}}\mathbf{x} = 1$.

Eftersom $\mathbf{H}$ är positivt definit så har vi ett konvext QP-problem med ett linjärt bivillkor.

Då gäller att $\hat{\mathbf{x}}$ är en globalt optimal lösning om och endast om $\hat{\mathbf{x}}$ tillsammans med en skalär $\hat{u}$ uppfyller optimalitetsvillkoren: $\mathbf{H}\hat{\mathbf{x}} - (-\mathbf{c})\hat{u} = -\mathbf{c}$ och $(-\mathbf{c})^{\mathsf{T}}\hat{\mathbf{x}} = 1$.

De första av dessa kan skrivas $\mathbf{H}\hat{\mathbf{x}} = -(\hat{u}+1)\cdot\mathbf{c}$ ur vilket följer att $\hat{\mathbf{x}} = (\hat{u}+1)\cdot\bar{\mathbf{x}}$, där $\bar{\mathbf{x}} = (2,1,1,1)^{\mathsf{T}}$ från (a)-uppgiften. (Ty $\mathbf{H}\bar{\mathbf{x}} = -\mathbf{c}$.)

Kravet att $(-\mathbf{c})^{\mathsf{T}}\hat{\mathbf{x}} = 1$ ger då att $\hat{u}+1 = 1/18$, ty $(-\mathbf{c})^{\mathsf{T}}\bar{\mathbf{x}} = 18$.

Därmed är $\hat{\mathbf{x}} = (2/18,\ 1/18,\ 1/18,\ 1/18)^{\mathsf{T}}$ optimal lösning till problemet.

Uppgift 4.(a)

Om både $x_1 > 0$ och $x_2 > 0$ så ges gradienten till f av

$$\nabla f(\mathbf{x}) = \left(\frac{-16}{(x_1+1)^2(x_2+1)} - 1\,,\ \frac{-16}{(x_1+1)(x_2+1)^2} - 3 \right),$$

medan Hessianen till f ges av

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \dfrac{32}{(x_1+1)^3(x_2+1)} & \dfrac{16}{(x_1+1)^2(x_2+1)^2} \\ \dfrac{16}{(x_1+1)^2(x_2+1)^2} & \dfrac{32}{(x_1+1)(x_2+1)^3} \end{bmatrix}.$$

Speciellt med $\mathbf{x}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ är $\nabla f(\mathbf{x}^{(1)}) = (3,\,3)$ och $\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

Matrisen $\mathbf{F}(\mathbf{x}^{(1)})$ är positivt definit, ty en matris av typen $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ är positivt definit om och endast om $a > 0$, $c > 0$ och $ac - b^2 > 0$, vilket är uppfyllt för $\mathbf{F}(\mathbf{x}^{(1)})$.

Därmed bestäms Newtonriktningen $\mathbf{d}^{(1)}$ ur ekvationssystemet $\mathbf{F}(\mathbf{x}^{(1)})\mathbf{d}^{(1)} = -\nabla f(\mathbf{x}^{(1)})^{\mathsf{T}}$,

dvs systemet $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

Vi prövar steget med $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Då blir $f(\mathbf{x}^{(2)}) = 16$, medan $f(\mathbf{x}^{(1)}) = 4 + 5 + 5 = 14$.

Eftersom $f(\mathbf{x}^{(2)}) > f(\mathbf{x}^{(1)})$ prövar vi istället steget $t = 0.5$,

så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + 0.5\mathbf{d}^{(1)} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$.

Då blir $f(\mathbf{x}^{(2)}) = 16/2.25 + 2.5 + 2.5 < 14 = f(\mathbf{x}^{(1)})$.

Steget med $t_1 = 0.5$ gick alltså bra, och vi accepterar $\mathbf{x}^{(2)} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$ som nästa iterationspunkt. Därmed har vi utfört en iteration med Newtons metod.

Uppgift 4.(b)

Funktionen f är konvex om och endast om

$f((1-t)\mathbf{u} + t\mathbf{v}) \leq (1-t)f(\mathbf{u}) + tf(\mathbf{v})$ för alla $\mathbf{u}$ och $\mathbf{v}$ i $I\!R^2$ och alla $t \in (0,1)$.

Inspirerade av räkningarna i (a)-uppgiften testar vi

$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ och $t = 0.5$, så att $(1-t)\mathbf{u} + t\mathbf{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Då är $f((1-t)\mathbf{u} + t\mathbf{v}) = 16$ medan $(1-t)f(\mathbf{u}) + tf(\mathbf{v}) = 0.5 \cdot 14 + 0.5 \cdot (-6) = 4$,

så f är *inte konvex*.

Uppgift 5.

Målfunktionen är strikt konvex, ty Hessianen ges av $\mathbf{F}(\mathbf{x}) = 2\,\mathbf{I}$,
där $\mathbf{I}$ är enhetsmatrisen (4×4).
Bivillkoren är linjära, och därmed konvexa.
Vidare finns det uppenbarligen punkter som uppfyller alla bivillkor med strikt olikhet, t ex $\mathbf{x} = (-1,-1,-1,-1)^{\mathsf{T}}$.

Därmed är problemet ett regulärt konvext problem, vilket betyder att KKT-villkoren är både nödvändiga och tillräckliga villkor för en global optimallösning.

Lagrangefunktionen till problemet kan skrivas:

$$L(\mathbf{x},\mathbf{y}) = (x_1-1)^2 + (x_2-1)^2 + (x_3-4)^2 + (x_4-3)^2 +$$
$$+\ y_1(x_1+x_2+x_4) + y_2(x_1+x_3+x_4) + y_3(x_2+x_3+x_4).$$

KKT-villkoren blir då:

$$\begin{array}{rll}
2x_1 - 2 + y_1 + y_2 = 0 & \partial L/\partial x_1 = 0 & \text{(KKT1)}\\
2x_2 - 2 + y_1 + y_3 = 0 & \partial L/\partial x_2 = 0 & \text{(KKT2)}\\
2x_3 - 8 + y_2 + y_3 = 0 & \partial L/\partial x_3 = 0 & \text{(KKT3)}\\
2x_4 - 6 + y_1 + y_2 + y_3 = 0 & \partial L/\partial x_4 = 0 & \text{(KKT4)}\\
x_1 + x_2 + x_4 \le 0 & \text{primal tillåtenhet} & \text{(KKT5)}\\
x_1 + x_3 + x_4 \le 0 & \text{primal tillåtenhet} & \text{(KKT6)}\\
x_2 + x_3 + x_4 \le 0 & \text{primal tillåtenhet} & \text{(KKT7)}\\
y_1 \ge 0 & \text{dual tillåtenhet} & \text{(KKT8)}\\
y_2 \ge 0 & \text{dual tillåtenhet} & \text{(KKT9)}\\
y_3 \ge 0 & \text{dual tillåtenhet} & \text{(KKT10)}\\
y_1(x_1+x_2+x_4) = 0 & \text{komplementaritet} & \text{(KKT11)}\\
y_2(x_1+x_3+x_4) = 0 & \text{komplementaritet} & \text{(KKT12)}\\
y_3(x_2+x_3+x_4) = 0 & \text{komplementaritet} & \text{(KKT13)}
\end{array}$$

(a). Med $\mathbf{x} = (0,0,0,0)^T$ blir (KKT1)–(KKT4) uppfyllda om och endast om $(y_1,y_2,y_3) = (-2,4,4)$, men detta strider mot (KKT8).
KKT-villkoren kan alltså *inte* uppfyllas med $x = (0,0,0,0)^T$.

(b). Med $\mathbf{x} = (-0.6,\ -0.6,\ 0.8,\ -0.2)^T$ uppfylls (KKT6)–(KKT7) med likhet och (KKT5) med strikt olikhet. Därmed är (KKT12)–(KKT13) uppfyllda, medan (KKT11) kräver att $y_1 = 0$. (KKT1)–(KKT2) ger då att $y_2 = y_3 = 3.2$, och då blir (faktiskt) även (KKT3)–(KKT4) uppfyllda.
Eftersom alla $y_i \ge 0$ så är slutligen även (KKT8)–(KKT10) uppfyllda.
Därmed är samtliga KKT-villkor uppfyllda!

(c). Eftersom $\hat{\mathbf{x}} = (-0.6,\ -0.6,\ 0.8,\ -0.2)^T$ uppfyller KKT-villkoren och problemet är konvext (enligt ovan) så är $\hat{\mathbf{x}}$ en globalt optimal lösning till problemet.

(d). Eftersom problemet är konvext och målfunktionen är *strikt* konvex,
så kan det inte finnas flera olika globala optimallösningar, ty då skulle mittpunkten mellan varje par av optimallösningar vara ännu bättre vilket är en motsägelse.
Alltså är $\hat{\mathbf{x}} = (-0.6,\ -0.6,\ 0.8,\ -0.2)^T$ den unika globala optimallösningen.

Uppgift 6. Endast kortfattade lösningar ges här!

(a).
Vi söker här den punkt bland alla punkter i $\mathcal{R}(\mathbf{A})$ som ligger närmast punkten $\mathbf{b}$.
Varje punkt i $\mathcal{R}(\mathbf{A})$ kan skrivas på formen $\mathbf{Av}$, så den sökta punkten ges av $\mathbf{A\hat{v}}$, där $\hat{\mathbf{v}}$ är en optimal lösning till problemet att minimera $|\mathbf{Av}-\mathbf{b}|^2$.
Detta löses som bekant med normalekvationerna $\mathbf{A^\mathsf{T}Av}=\mathbf{A^\mathsf{T}b}$ som ger att

$$\hat{\mathbf{v}} = \begin{pmatrix} (b_1-b_2+b_3)/3 \\ (-b_1+b_2+2b_3)/6 \end{pmatrix}, \text{ varefter } \mathbf{A\hat{v}} = \begin{pmatrix} (b_1-b_2)/2 \\ (b_2-b_1)/2 \\ b_3 \end{pmatrix}$$

(b).
Vi söker här den punkt bland alla punkter i $\mathcal{N}(\mathbf{A}^\mathsf{T})$ som ligger närmast punkten $\mathbf{b}$.
Denna sökta punkt ges av optimala lösningen $\hat{\mathbf{y}}$ till problemet att minimera $\frac{1}{2}|\,\mathbf{y}-\mathbf{b}\,|^2$ under bivillkoren $\mathbf{A^\mathsf{T}y}=\mathbf{0}$. Detta problem kan ekvivalent skrivas:

minimera $\frac{1}{2}\mathbf{y}^\mathsf{T}\mathbf{I}\,\mathbf{y}-\mathbf{b}^\mathsf{T}\mathbf{y}+\frac{1}{2}|\,\mathbf{b}\,|^2$ då $\mathbf{A^\mathsf{T}y}=\mathbf{0}$.

De nödvändiga och tillräckliga optimalitetsvillkoren för detta QP-problem med linjära likhetsbivillkor blir: $\mathbf{y}-\mathbf{Au}=\mathbf{b}$ och $\mathbf{A^\mathsf{T}y}=\mathbf{0}$.

Insättning av $\mathbf{y}=\mathbf{Au}+\mathbf{b}$ i $\mathbf{A^\mathsf{T}y}=\mathbf{0}$ ger att $\mathbf{A^\mathsf{T}Au}=-\mathbf{A^\mathsf{T}b}$ (nästan normalekvationerna).

$$\text{Vi får att } \hat{\mathbf{u}} = -\begin{pmatrix} (b_1-b_2+b_3)/3 \\ (-b_1+b_2+2b_3)/6 \end{pmatrix}, \text{ varefter } \hat{\mathbf{y}}=\mathbf{A\hat{u}}+\mathbf{b} = \begin{pmatrix} (b_1+b_2)/2 \\ (b_1+b_2)/2 \\ 0 \end{pmatrix}.$$

(c).
Vi söker här den punkt $\hat{\mathbf{w}}$ bland alla punkter i $\mathcal{R}(\mathbf{A}^\mathsf{T})$ som ligger närmast punkten $\mathbf{c}$.
Man konstaterar snabbt (med exempelvis Gauss-Jordan) att första och tredje kolonnerna i $\mathbf{A}^\mathsf{T}$ utgör en bas till $\mathbb{R}^2$. Därmed är $\mathcal{R}(\mathbf{A}^\mathsf{T})=\mathbb{R}^2$, och vår sökta punkt är förstås $\hat{\mathbf{w}}=\mathbf{c}$.

(d).
Vi söker här den punkt $\hat{\mathbf{z}}$ bland alla punkter i $\mathcal{N}(\mathbf{A})$ som ligger närmast punkten $\mathbf{c}$.
Man konstaterar snabbt (med exempelvis Gauss-Jordan) att enda lösningen till $\mathbf{Az}=\mathbf{0}$ är $\mathbf{z}=0$. Därmed är $\mathcal{N}(\mathbf{A})=\{\,\mathbf{0}\,\}$, och vår sökta punkt är förstås $\hat{\mathbf{z}}=\mathbf{0}$.