Lösningar till SF1861 Optimeringslära för T, 26/8-08

Uppgift 1.(a)

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

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 25 \\ 10 \\ 0 \\ 0 \\ -15 \\ -20 \end{pmatrix}.$$

Variabelvektorn ges av $\mathbf{x} = \begin{pmatrix} x_{13} \\ x_{14} \\ x_{23} \\ x_{24} \\ x_{35} \\ x_{36} \\ x_{45} \\ x_{46} \end{pmatrix}$, där variabeln x_{ij} betecknar flödet i bågen

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

Flödets kostnad (som ska minimeras) ges av $\mathbf{c}^\mathsf{T}\mathbf{x}$, där $\mathbf{c} = \begin{pmatrix} c_{13} \\ c_{14} \\ c_{23} \\ c_{24} \\ c_{35} \\ c_{36} \\ c_{45} \\ c_{46} \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 3 \\ 5 \\ 9 \\ 8 \\ 7 \\ 6 \end{pmatrix}$.

Om primala problemet är på standardformen

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

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

$$\begin{aligned} \text{maximera} \quad & 25y_1 + 10y_2 - 15y_5 - 20y_6 \\ \text{då} \quad & y_1 - y_3 \leq 2, \\ & y_1 - y_4 \leq 4, \\ & y_2 - y_3 \leq 3, \\ & y_2 - y_4 \leq 5, \\ & y_3 - y_5 \leq 9, \\ & y_3 - y_6 \leq 8, \\ & y_4 - y_5 \leq 7, \\ & y_4 - y_6 \leq 6. \end{aligned}$$

Uppgift 1.(b)

Först använder vi Gauss–Jordans metod på den givna matrisen $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

Efter några "enkla radoperationer" erhålls matrisen $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{T}$.

Nu är $\mathbf{A}$ överförd till trappstegsform med *två trappstegsettor*.
En bas till $\mathcal{R}(\mathbf{A})$ erhålls genom att som basvektorer välja de kolonner i $\mathbf{A}$ som svarar mot trappstegsettor i $\mathbf{T}$, dvs kolonnerna 1 och 2 i $\mathbf{A}$.

De två vektorerna $\begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}$ och $\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix}$ 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 trappstegsetta)
och bestäm sedan x_1 och x_2 (variablerna svarande mot trappstegsettor)

så att $\mathbf{Tx} = \mathbf{0}$. Det ger den första och enda basvektorn $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ till $\mathcal{N}(\mathbf{A})$.

Nu använder vi Gauss–Jordans metod på den transponerade matrisen $\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$.

Efter några "enkla radoperationer" erhålls matrisen $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \tilde{\mathbf{T}}$.

Nu är $\mathbf{A}^{\mathsf{T}}$ överförd till trappstegsform med *två trappstegsettor*.
En bas till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$ erhålls genom att som basvektorer välja de kolonner i $\mathbf{A}^{\mathsf{T}}$ som svarar mot trappstegsettor i $\tilde{\mathbf{T}}$, dvs kolonnerna 1 och 2 i $\mathbf{A}^{\mathsf{T}}$.

De två vektorerna $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ och $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ utgör alltså en bas till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$

En bas till $\mathcal{N}(\mathbf{A}^{\mathsf{T}})$ 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 trappstegsettor)

så att $\tilde{\mathbf{T}}\mathbf{y} = \mathbf{0}$. Det ger den första och enda basvektorn $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ till $\mathcal{N}(\mathbf{A}^{\mathsf{T}})$.

Notera att vi erhållit att $\mathcal{N}(\mathbf{A}^{\mathsf{T}}) = \mathcal{N}(\mathbf{A})$, vilket medför att motsvarande ortogonala komplement också är lika, dvs $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{A}^{\mathsf{T}})$. De bägge basvektorer som vi ovan beräknade till $\mathcal{R}(\mathbf{A})$ duger alltså även som basvektorer till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$, och de bägge basvektorer som vi ovan beräknade till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$ duger även som basvektorer till $\mathcal{R}(\mathbf{A})$.

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 $\mathbf{A} = \begin{bmatrix} 7 & 3 & 1 & -1 \\ 2 & 1 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 10 \\ 3 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (5,\, 2,\, 1,\, 1)$.

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

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$, 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 $\begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,

dvs $x_1 = 1$ och $x_2 = 1$, vilket stämmer med förslaget.

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

dvs $\begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$, 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

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

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

$x_1 = 1$, $x_2 = 1$, övriga $x_j = 0$. Optimalvärdet $= 5 + 2 = 7$.

Uppgift 2.(b)

Efter införande av slackvariabler x_5 och x_6 får vi ett nytt LP-problem på standardform, dvs minimera $\mathbf{c}^\mathsf{T}\mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$ och $\mathbf{x} \geq \mathbf{0}$, där nu

$\mathbf{A} = \begin{bmatrix} 7 & 3 & 1 & -1 & -1 & 0 \\ 2 & 1 & 1 & -1 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 10 \\ 3 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (5,\, 2,\, 1,\, 1,\, 0,\, 0)$.

Vi startar från baslösningen ovan, med $\beta = (1, 2)$, som ju 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

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

Eftersom reducerade kostnaderna är icke-negativa så är den aktuella lösningen optimal även till detta modifierade problem.

Optimal lösning: $x_1 = 1$, $x_2 = 1$, övriga $x_j = 0$. Optimalvärdet $= 7$.

Uppgift 3.(a)

Målfunktionen är $q(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$, där $\mathbf{H} = \begin{bmatrix} 2 & -2 & 0 & 0 \\ -2 & 6 & -4 & 0 \\ 0 & -4 & 6 & -2 \\ 0 & 0 & -2 & 6 \end{bmatrix}$ och $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Punkten $\hat{\mathbf{x}}$ är en global minpunkt till $q(\mathbf{x})$ om och endast om $\mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{0}$ och $\mathbf{H}$ är positivt semidefinit. Om $\mathbf{H}$ dessutom är positivt definit så är $\hat{\mathbf{x}}$ den unika minpunkten.

För att avgöra om $\mathbf{H}$ är positivt definit, eller positivt semidefinit, eller ingetdera, försöker vi LDL^T-faktorisera $\mathbf{H}$.

$$\mathbf{H} = \begin{bmatrix} 2 & -2 & 0 & 0 \\ -2 & 6 & -4 & 0 \\ 0 & -4 & 6 & -2 \\ 0 & 0 & -2 & 6 \end{bmatrix}.$$

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

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & -4 & 0 \\ 0 & -4 & 6 & -2 \\ 0 & 0 & -2 & 6 \end{bmatrix}.$$

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

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -2 & 6 \end{bmatrix}.$$

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

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T}\mathbf{E}_3^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Därmed är LDL^T-faktoriseringen klar, och man har att

$$\mathbf{H} = \mathbf{L}\mathbf{D}\mathbf{L}^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

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

Ekvationssystemet $\mathbf{H}\mathbf{x} + \mathbf{c} = \mathbf{0}$ blir i detta fall $\mathbf{H}\mathbf{x} = \mathbf{0}$ (ty $\mathbf{c} = \mathbf{0}$), och eftersom $\mathbf{H}$ är positivt definit (och således ickesingulär) så är $\hat{\mathbf{x}} = (0,0,0,0)^\mathsf{T}$ den unika lösningen till detta ekvationssystem. Därmed är $\hat{\mathbf{x}} = (0,0,0,0)^\mathsf{T}$ även den unika minpunkten till $q(\mathbf{x})$.

Uppgift 3.(b)

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

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

Då gäller att $\mathbf{x}$ är en optimal lösning om och endast om $\mathbf{x}$ tillsammans med en skalär u uppfyller de välkända optimalitetsvillkoren, som i detta fall kan skrivas:

$$\mathbf{H}\mathbf{x} = \mathbf{e}\,u \quad \text{och} \quad \mathbf{e}^\mathsf{T}\mathbf{x} = 2.$$

För den första föreslagna punkten $\mathbf{x} = (0.7,\ 0.6,\ 0.5,\ 0.2)^\mathsf{T}$ gäller dels att $\mathbf{e}^\mathsf{T}\mathbf{x} = 2$,
dels att $\mathbf{H}\mathbf{x} = (0.2,\ 0.2,\ 0.2,\ 0.2)^\mathsf{T} = \mathbf{e}\,u$, med $u = 0.2$.
Därmed är $\mathbf{x} = (0.7,\ 0.6,\ 0.5,\ 0.2)^\mathsf{T}$ en optimal lösning.

Eftersom $\mathbf{H}$ är positivt definit så är $q(\mathbf{x})$ strikt konvex, och då kan det inte finnas flera optimala lösningar. Som en kontroll testar vi dock även den andra föreslagna punkten, dvs $\mathbf{x} = (0.8,\ 0.5,\ 0.4,\ 0.3)^\mathsf{T}$. Även för denna gäller att $\mathbf{e}^\mathsf{T}\mathbf{x} = 2$, men nu blir $\mathbf{H}\mathbf{x} = (0.6,\ -0.2,\ -0.2,\ 1.0)^\mathsf{T}$ som förvisso inte kan skrivas på formen $\mathbf{e}\,u$ för något u.
Därmed är $\mathbf{x} = (0.8,\ 0.5,\ 0.4,\ 0.3)^\mathsf{T}$ inte en optimal lösning.

Uppgift 4.(a)

Här är $f(\mathbf{x}) = e^{x_1} + e^{x_2} + x_1^2 + x_2^2 - x_1 x_2 - 3x_1 - 3x_2$.

Gradienten till f blir då $\nabla f(\mathbf{x}) = (\ e^{x_1} + 2x_1 - x_2 - 3,\ \ e^{x_2} - x_1 + 2x_2 - 3\ )$,

medan Hessianen till f ges av $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} e^{x_1}+2 & -1 \\ -1 & e^{x_2}+2 \end{bmatrix}$.

Speciellt med $\mathbf{x}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ är $\nabla f(\mathbf{x}^{(1)}) = (-2,\ -2)$ och $\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$.

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

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

Då blir $f(\mathbf{x}^{(2)}) = 2e - 5 < 1$, medan $f(\mathbf{x}^{(1)}) = 2$.

Eftersom $f(\mathbf{x}^{(2)}) < f(\mathbf{x}^{(1)})$ accepterar vi $\mathbf{x}^{(2)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ som nästa iterationspunkt.
Därmed har vi utfört en iteration med Newtons metod.

Uppgift 4.(b)

Funktionen f är konvex på hela $\mathbb{R}^2$ om och endast om Hessianen

$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} e^{x_1}+2 & -1 \\ -1 & e^{x_2}+2 \end{bmatrix}$ är positivt semidefinit för alla $\mathbf{x} \in \mathbb{R}^2$.

Men $\mathbf{F}(\mathbf{x})$ är på formen $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ med $b = -1$, $a = e^{x_1} + 2 > 2$ och $c = e^{x_2} + 2 > 2$.
Således är $a > 0$, $c > 0$ och $ac - b^2 > 0$, vilket medför att $\mathbf{F}(\mathbf{x})$ är positivt definit, och därmed även positivt semidefinit, för alla $\mathbf{x} \in \mathbb{R}^2$.

Alltså är funktionen f konvex på hela $\mathbb{R}^2$.

Uppgift 5.

Målfunktionen kan skrivas

$$f(\mathbf{x}) = \tfrac{1}{2}\,(\mathbf{x}-\mathbf{c})^{\mathsf{T}}(\mathbf{x}-\mathbf{c}) = \tfrac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{I}\,\mathbf{x} - \mathbf{c}^{\mathsf{T}}\mathbf{x} + \tfrac{1}{2}\,\mathbf{c}^{\mathsf{T}}\mathbf{c}\,,$$

där $\mathbf{I}$ betecknar enhetsmatrisen.

Denna målfunktion är strikt konvex eftersom Hessianen ges av $\mathbf{F}(\mathbf{x}) = \mathbf{I}$, som är positivt definit. Bivillkoren är linjära, och därmed konvexa. Vidare finns det uppenbarligen punkter som uppfyller alla bivillkor med strikt olikhet, t ex $\mathbf{x} = (0,0,0,0)^{\mathsf{T}}$.

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

Skriv bivillkoren på formen $x_j - 1 \le 0$ och $-x_j - 1 \le 0$, samt låt motsvarande lagrangemultiplikatorer heta λ_j resp μ_j, så att lagrangemultiplikatorvektorn blir $\mathbf{y} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu_1, \mu_2, \mu_3, \mu_4)^{\mathsf{T}}$.

Då kan Lagrangefunktonen skrivas

$$L(\mathbf{x},\mathbf{y}) = \tfrac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{I}\,\mathbf{x} - \mathbf{c}^{\mathsf{T}}\mathbf{x} + \tfrac{1}{2}\,\mathbf{c}^{\mathsf{T}}\mathbf{c} + \sum_{j=1}^{4}\lambda_j(x_j-1) + \sum_{j=1}^{4}\mu_j(-x_j-1) =$$

$$= \sum_{j=1}^{4}\Big(\tfrac{1}{2}\,x_j^2 - c_j x_j + \tfrac{1}{2}\,c_j^2 + \lambda_j(x_j-1) + \mu_j(-x_j-1)\Big),$$

och KKT-villkoren blir följande:

(1)	$x_j - c_j + \lambda_j - \mu_j = 0,$	för $j = 1,2,3,4$	($\partial L/\partial x_j = 0$)
(2)	$-1 \le x_j \le 1,$	för $j = 1,2,3,4$	(Tillåten punkt)
(3)	$\lambda_j \ge 0$ och $\mu_j \ge 0,$	för $j = 1,2,3,4$	(Ickeneg. lagrangemult.)
(4)	$\lambda_j(1-x_j) = 0$ och $\mu_j(1+x_j) = 0,$	för $j = 1,2,3,4$	(Kompl.villkor)

Att hitta en lösning till KKT-villkoren förenklas väsentligt av att man för detta speciella ("separabla") problem kan betrakta ett index j i taget:

Om $-1 \le c_j \le 1$ så uppfylls KKT-villkoren av $x_j = c_j,\ \lambda_j = 0$ och $\mu_j = 0$.
Om $c_j > 1$ så uppfylls KKT-villkoren av $x_j = 1,\ \lambda_j = c_j - 1\ (> 0)$ och $\mu_j = 0$.
Om $c_j < -1$ så uppfylls KKT-villkoren av $x_j = -1,\ \lambda_j = 0$ och $\mu_j = -1 - c_j\ (> 0)$.

Om speciellt $\mathbf{c} = (-1.4,\ -0.4,\ 0.6,\ 1.6)^{\mathsf{T}}$ så får vi följande lösning till KKT-villkoren:

$\hat{\mathbf{x}} = (-1,\ -0.4,\ 0.6,\ 1)^{\mathsf{T}},\ \hat{\lambda} = (0,\ 0,\ 0,\ 0.6)^{\mathsf{T}}$ och $\hat{\mu} = (0.4,\ 0,\ 0,\ 0)^{\mathsf{T}}$.

Eftersom målfunktionen är strikt konvex och tillåtna området är konvext, så är $\hat{\mathbf{x}}$ den unika optimallösningen.

Uppgift 6.

Vi inför först några beteckningar.

Rummet som vår tvetjärt befinner sig i består av alla punkter (x, y, z) som uppfyller $0 \leq x \leq a,\ \ 0 \leq y \leq b$ och $0 \leq z \leq c$.

Rummet har sex stycken "begränsningsytor", nämligen

Golvet	(G)	där $z = 0$
Taket	(T)	där $z = c$
Söderväggen	(S)	där $y = 0$
Norrväggen	(N)	där $y = b$
Västerväggen	(V)	där $x = 0$
Österväggen	(Ö)	där $x = a$

och det är endast på dessa ytor som vår tvestjärt kan färdas.
Starten går i skärningspunkten av G, S och V, medan målet ligger i skärningspunkten av T, N och Ö.

Det finns tolv stycken kantlinjer, som var och en utgör skärningen av två begränsningsytor. Sex av dessa kantlinjer är speciellt intressanta, nämligen

Norrväggen/Golvet	(NG)	där $y = b$ och $z = 0$
Golvet/Österväggen	(GÖ)	där $z = 0$ och $x = a$
Öster/Söderväggen	(ÖS)	där $x = a$ och $y = 0$
Söderväggen/Taket	(ST)	där $y = 0$ och $z = c$
Taket/Västerväggen	(TV)	där $z = c$ och $x = 0$
Väster/Norrväggen	(VN)	där $x = 0$ och $y = b$

Tillsammans bildar dessa sex kantlinjer en sluten kurva K som vår tvestjärt måste passera på sin väg från $(0, 0, 0)$ till (a, b, c). (Om $(0, 0, 0)$ är "sydpolen" och (a, b, c) är "nordpolen" så är kurvan K "ekvatorn".)

Kortaste vägen mellan två punkter på en begränsningsyta är en rät linje.

Varje punkt (x, y, z) på kurvan K kan nås från hörnet $(0, 0, 0)$ längs en rät linje i någon av ytorna G, S eller V, och från varje punkt (x, y, z) på kurvan K kan hörnet (a, b, c) nås längs en rät linje i någon av ytorna T, N eller Ö.

Därför måste kortaste vägen från $(0, 0, 0)$ till (a, b, c) finnas bland något (eller några) av följande alternativ, där $\longrightarrow$ betyder "längs en rät linje":

1. $(0, 0, 0) \longrightarrow (x, b, 0) \longrightarrow (a, b, c)$, där $0 \leq x \leq a$.
2. $(0, 0, 0) \longrightarrow (a, y, 0) \longrightarrow (a, b, c)$, där $0 \leq y \leq b$.
3. $(0, 0, 0) \longrightarrow (a, 0, z) \longrightarrow (a, b, c)$, där $0 \leq z \leq c$.
4. $(0, 0, 0) \longrightarrow (x, 0, c) \longrightarrow (a, b, c)$, där $0 \leq x \leq a$.
5. $(0, 0, 0) \longrightarrow (0, y, c) \longrightarrow (a, b, c)$, där $0 \leq y \leq b$.
6. $(0, 0, 0) \longrightarrow (0, b, z) \longrightarrow (a, b, c)$, där $0 \leq z \leq c$.

Låt oss analysera fall 1.
Här korsar tvetjärtens väg "ekvatorn" i punkten $(x, b, 0)$ på kantlinjen NG.

Totala längden på denna väg är $f_1(x) = \sqrt{x^2 + b^2} + \sqrt{(a-x)^2 + c^2}$,
där x ska väljas så att $f_1(x)$ minimeras.

Kalkyler (deriveringar) visar att $f_1(x)$ är en konvex funktion som minimeras unikt av $\hat{x} = \dfrac{ab}{b+c}$, varvid $f_1(\hat{x}) = \sqrt{a^2 + (b+c)^2} = \sqrt{a^2 + b^2 + c^2 + 2bc}$.

(Detta sista resultat kan alternativt erhållas genom att "vika ner" Norrväggen så att den utgör en fortsättning på Golvet. Hörnet (a, b, c) flyttas då ner till punkten $(a, b+c, 0)$, och kortaste vägen från $(0, 0, 0)$ till denna nerflyttade punkt är längs diagonalen i en rektangel med sidorna a och $b + c$. Längden på denna diagonal är mycket riktigt $\sqrt{a^2 + (b+c)^2} = \sqrt{a^2 + b^2 + c^2 + 2bc}$.)

Motsvarande analyser ger följande vägar och väglängder för de sex fallen:

1. $(0,0,0) \longrightarrow (\dfrac{ab}{b+c}, b, 0) \longrightarrow (a,b,c)$. Väglängd $= \sqrt{a^2+b^2+c^2+2bc}$.

2. $(0,0,0) \longrightarrow (a, \dfrac{ba}{a+c}, 0) \longrightarrow (a,b,c)$. Väglängd $= \sqrt{a^2+b^2+c^2+2ac}$.

3. $(0,0,0) \longrightarrow (a, 0, \dfrac{ca}{a+b}) \longrightarrow (a,b,c)$. Väglängd $= \sqrt{a^2+b^2+c^2+2ab}$.

4. $(0,0,0) \longrightarrow (\dfrac{ac}{b+c}, 0, c) \longrightarrow (a,b,c)$. Väglängd $= \sqrt{a^2+b^2+c^2+2bc}$.

5. $(0,0,0) \longrightarrow (0, \dfrac{bc}{a+c}, c) \longrightarrow (a,b,c)$. Väglängd $= \sqrt{a^2+b^2+c^2+2ac}$.

6. $(0,0,0) \longrightarrow (0, b, \dfrac{cb}{a+b}) \longrightarrow (a,b,c)$. Väglängd $= \sqrt{a^2+b^2+c^2+2ab}$.

Om $a > b > c$ så är $2bc < 2ac < 2ab$ och då ger 1 och 4 de kortaste vägarna.

Om $c > a > b$ så är $2ab < 2bc < 2ac$ och då ger 3 och 6 de kortaste vägarna.