Lösningar till SF1852 Optimeringslära för E, 16/1–08

Uppgift 1.(a)

Problemet är ett transportproblem, dvs ett specialfall av minkostnadsflödesproblem.
Nätverket består av 7 st noder A,B,C,P,Q,R,S, alternativt kallade 1,2,...,7,
samt 12 st bågar (A,P),...,(C,S), alternativt kallade (1,4),...,(3,7).

Den av transportchefen föreslagna lösningen svarar mot ett uppspännande träd i nätverket, dvs mot en baslösning till problemet. Det är vidare en tillåten baslösning, ty balansekvationerna är uppfyllda i alla noder och inga variabler är negativa.

Man beräknar simplexmultiplikatorer y_i ur villkoren $y_i - y_j = c_{ij}$ för basvariabler
(dvs trädbågar), samt $y_7 = 0$. Det ger att $y = (4, 6, 5, -2, -1, -1, 0)$.
Därefter beräknas reducerade kostnader för icke-basvariabler ur $r_{ij} = c_{ij} - y_i + y_j$.
Det ger att $r_{ij} = 1$ för alla icke-basvariabler.
Den förslagna lösningen är alltså optimal, eftersom $r_{ij} \geq 0$ för alla icke-basvariabler.

Om man inför variabelvektorn

$\mathbf{x} = (\ x_{14}\ \ x_{15}\ \ x_{16}\ \ x_{17}\ \ x_{24}\ \ x_{25}\ \ x_{26}\ \ x_{27}\ \ x_{34}\ \ x_{35}\ \ x_{36}\ \ x_{37}\ )^{\mathsf{T}}$

och kostnadsvektorn

$\mathbf{c} = (\ c_{14}\ \ c_{15}\ \ c_{16}\ \ c_{17}\ \ c_{24}\ \ c_{25}\ \ c_{26}\ \ c_{27}\ \ c_{34}\ \ c_{35}\ \ c_{36}\ \ c_{37}\ )^{\mathsf{T}} =$
$\quad = (\ 6\ \ 5\ \ 6\ \ 5\ \ 9\ \ 7\ \ 7\ \ 7\ \ 8\ \ 7\ \ 6\ \ 5\ )^{\mathsf{T}},$

så kan det betraktade transportproblemet skrivas på formen

$$\begin{array}{lrl} \text{TP}: & \text{minimera} & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ & \text{då} & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & & \mathbf{x} \geq \mathbf{0}, \end{array}$$

där

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \end{bmatrix} \quad \text{och} \quad \mathbf{b} = \begin{pmatrix} 70 \\ 90 \\ 60 \\ -50 \\ -60 \\ -70 \\ -40 \end{pmatrix}.$$

En optimal lösning är, enligt ovan, $\hat{\mathbf{x}} = (\ 50\ \ 20\ \ 0\ \ 0\ \ 0\ \ 40\ \ 50\ \ 0\ \ 0\ \ 0\ \ 20\ \ 40\ )^{\mathsf{T}},$

Uppgift 1.(b)

Först kan konstateras att $\mathbf{A}$ är symmetrisk, dvs $\mathbf{A}^{\mathsf{T}} = \mathbf{A}$, vilket medför att $\mathcal{R}(\mathbf{A}^{\mathsf{T}}) = \mathcal{R}(\mathbf{A})$ och $\mathcal{N}(\mathbf{A}^{\mathsf{T}}) = \mathcal{N}(\mathbf{A})$.

Nu använder vi Gauss–Jordans metod på den givna matrisen

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Multiplikation av första raden med 0.5 ger till resultat matrisen

$$\begin{bmatrix} 1 & -0.5 & -0.5 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Addition av (+1) gånger första raden till andra raden, samt addition av (+1) gånger första raden till tredje raden, ger till resultat matrisen

$$\begin{bmatrix} 1 & -0.5 & -0.5 \\ 0 & 1.5 & -1.5 \\ 0 & -1.5 & 1.5 \end{bmatrix}.$$

Multiplikation av andra raden med 1/1.5 ger till resultat matrisen

$$\begin{bmatrix} 1 & -0.5 & -0.5 \\ 0 & 1 & -1 \\ 0 & -1.5 & 1.5 \end{bmatrix}.$$

Addition av (+0.5) gånger andra raden till första raden, samt addition av (+1.5) 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 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 nr 1 och 2 i $\mathbf{A}$.

De två vektorerna $\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$ och $\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$ utgör alltså en bas till såväl $\mathcal{R}(\mathbf{A})$ som till $\mathcal{R}(\mathbf{A}^{\mathsf{T}})$.

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 $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ till såväl $\mathcal{N}(\mathbf{A})$ som till $\mathcal{N}(\mathbf{A}^{\mathsf{T}})$.

Uppgift 2.(a)

Om vi inför slackvariabler x_5 och x_6, för att överföra olikhetsbivillkoren till likhetsbivillkor, samt byter tecken på målfunktionen, för att överföra maximeringsproblemet till ett minimeringsproblem, så får vi 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 $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 1 & 1 & 0 \\ 2 & 2 & 1 & 1 & 0 & 1 \end{bmatrix}$, $\quad \mathbf{b} = \begin{pmatrix} 90 \\ 60 \end{pmatrix}$ och $\mathbf{c}^{\mathsf{T}} = (-3, -4, -2, -1, 0, 0)$.

Startlösningen ska ha basvariablerna x_5 och x_6, vilket innebär 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 $\mathbf{A}_\delta = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 2 & 1 & 1 \end{bmatrix}$.

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 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 90 \\ 60 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 90 \\ 60 \end{pmatrix}$.

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

Reducerade kostnaderna för icke-basvariablerna ges av

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

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 $\mathbf{A}_\beta \bar{\mathbf{a}}_2 = \mathbf{a}_2$,

dvs $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, 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

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i2}} \mid \bar{a}_{i2} > 0 \right\} = \min \left\{ \frac{90}{2}, \frac{60}{2} \right\} = \frac{60}{2} = \frac{\bar{b}_2}{\bar{a}_{22}}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_6$ inte längre får vara kvar som basvariabel. Dess plats tas av x_2.

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

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

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 $\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 90 \\ 60 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 30 \\ 30 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -4 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av

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

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

Därmed är punkten $x_1 = 0$, $x_2 = 30$, $x_3 = 0$, $x_4 = 0$ optimal till det ursprungliga problemet. Optimalvärdet är -120 till minimeringsproblemet och $+120$ till maximeringsproblemet.

Uppgift 2.(b)

Eftersom $r_{\delta_3} = r_3 = 0$ så påverkas inte målfunktionen om vi låter ickebasvariabeln x_3 öka från sitt aktuella värde 0. Vi kan därför låta x_3 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_3$ ur systemet $\mathbf{A}_\beta\bar{\mathbf{a}}_3 = \mathbf{a}_3$,

dvs $\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_3 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{30}{1}, \frac{30}{0.5} \right\} = \frac{30}{1} = \frac{\bar{b}_1}{\bar{a}_{13}}.$$

Minimerande index är $i = 1$, varför $x_{\beta_1} = x_5$ inte längre får vara kvar som basvariabel. Dess plats tas av x_3.

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

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

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 $\begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 90 \\ 60 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 30 \\ 15 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av

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

Eftersom $\mathbf{r}_\delta \geq \mathbf{0}$ så är även denna baslösningen optimal (vilket vi redan visste).

Därmed är punkten $x_1 = 0$, $x_2 = 15$, $x_3 = 30$, $x_4 = 0$ optimal till det ursprungliga problemet. Optimalvärdet är -120 till minimeringsproblemet och $+120$ till maximeringsproblemet.

Uppgift 3.(a)

Vi utför radoperationer på matrisen $\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 6 \\ -1 & 0 & 1 & -3 \\ 0 & -1 & -1 & 0 \end{bmatrix}$.

Addition av (+1) gånger första raden till andra raden ger $\begin{bmatrix} 1 & 1 & 0 & 6 \\ 0 & 1 & 1 & 3 \\ 0 & -1 & -1 & 0 \end{bmatrix}$.

Addition av (+1) gånger andra raden till tredje raden ger $\begin{bmatrix} 1 & 1 & 0 & 6 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

Tredje raden motsvarar nu ekvationen $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 9$ som saknar lösning.

Uppgift 3.(b) Att minimera $|\mathbf{Ax} - \mathbf{b}|^2$ är ekvivalent med att lösa $\mathbf{A}^\mathsf{T}\mathbf{Ax} = \mathbf{A}^\mathsf{T}\mathbf{b}$.

Vi utför därför radoperationer på matrisen $\begin{bmatrix} \mathbf{A}^\mathsf{T}\mathbf{A} & \mathbf{A}^\mathsf{T}\mathbf{b} \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 9 \\ 1 & 2 & 1 & 6 \\ -1 & 1 & 2 & -3 \end{bmatrix}$.

Addition av (−0.5) gånger första raden till andra raden, samt
addition av (+0.5) gånger första raden till tredje raden, ger $\begin{bmatrix} 2 & 1 & -1 & 9 \\ 0 & 1.5 & 1.5 & 1.5 \\ 0 & 1.5 & 1.5 & 1.5 \end{bmatrix}$.

Addition av (−1) gånger andra raden till tredje raden, samt
addition av (−2/3) gånger andra raden till tredje raden, ger $\begin{bmatrix} 2 & 0 & -2 & 8 \\ 0 & 1.5 & 1.5 & 1.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

En lösning (av flera) till detta systemet är $\bar{\mathbf{x}} = (4,\ 1,\ 0)^\mathsf{T}$.

Uppgift 3.(c) En lösning till normalekvationerna är enligt ovan $\bar{\mathbf{x}} = (4,\ 1,\ 0)^\mathsf{T}$,
med $\mathbf{A}\bar{\mathbf{x}} = (5,\ -4,\ -1)^\mathsf{T}$. Då gäller ekvivalensen $\mathbf{A}^\mathsf{T}\mathbf{Ax} = \mathbf{A}^\mathsf{T}\mathbf{b} \Leftrightarrow \mathbf{Ax} = \mathbf{A}\bar{\mathbf{x}}$.

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

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

Vi utför därför radoperationer på matrisen $\begin{bmatrix} \mathbf{AA}^\mathsf{T} & \mathbf{A}\bar{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 & 5 \\ -1 & 2 & -1 & -4 \\ -1 & -1 & 2 & -1 \end{bmatrix}$.

Addition av (+0.5) gånger första raden till andra raden, samt
addition av (+0.5) gånger första raden till tredje raden, ger $\begin{bmatrix} 2 & -1 & -1 & 5 \\ 0 & 1.5 & -1.5 & -1.5 \\ 0 & -1.5 & 1.5 & 1.5 \end{bmatrix}$.

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

En lösning (av flera) till detta system är $\hat{\mathbf{u}} = (2,\ -1,\ 0)^\mathsf{T}$.

Då är $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{u}} = (3,\ \ 2,\ -1)^\mathsf{T}$ minimum-normlösningen till minsta-kvadratproblemet.

Uppgift 4.(a)

Låt $f(\mathbf{x}) = 8x_1^2 - 6x_2^2$, $g_1(\mathbf{x}) = 2x_1^2 + 3x_2^2 - 12$ och $g_2(\mathbf{x}) = 8 - 2x_1^2 - 3x_2^2$.

Då kan vårt givna problem skrivas: minimera $f(\mathbf{x})$ då $g_1(\mathbf{x}) \leq 0$ och $g_2(\mathbf{x}) \leq 0$.

Gradienterna ges av
$\nabla f(\mathbf{x}) = (16x_1,\ -12x_2)$, $\nabla g_1(\mathbf{x}) = (4x_1,\ 6x_2)$ och $\nabla g_2(\mathbf{x}) = (-4x_1,\ -6x_2)$.

KKT-villkoren för detta problem är:

KKT-1: $\nabla f(\mathbf{x}) + y_1 \nabla g_1(\mathbf{x}) + y_2 \nabla g_2(\mathbf{x}) = \mathbf{0}^\mathsf{T}$,

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

KKT-3: $y_i \geq 0$ för $i = 1, 2$,

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

som i vårt speciella fall kan skrivas

KKT-1: $4x_1(4 + y_1 - y_2) = 0$ och $6x_2(-2 + y_1 - y_2) = 0$.

KKT-2: $2x_1^2 + 3x_2^2 - 12 \leq 0$ och $8 - 2x_1^2 - 3x_2^2 \leq 0$.

KKT-3: $y_1 \geq 0$ och $y_2 \geq 0$.

KKT-4: $y_1(2x_1^2 + 3x_2^2 - 12) = 0$ och $y_2(8 - 2x_1^2 - 3x_2^2) = 0$.

Uppgift 4.(b)

Ur KKT-4 följer att minst en av y_1 och y_2 måste vara lika med noll.

Ur KKT-1 följer att minst en av x_1 och x_2 måste vara lika med noll.
Å andra sidan kan inte både x_1 och x_2 vara noll samtidigt, ty då uppfylls inte KKT-2.
Alltså är exakt en av x_1 och x_2 lika med noll i varje KKT-punkt.

Antag först att $x_1 \neq 0$ medan $x_2 = 0$.
Då följer ur KKT-1 att $y_1 = 0$ och $y_2 = 4$, varefter KKT-4 ger att $x_1 = 2$ eller -2.

Antag sedan att $x_2 \neq 0$ medan $x_1 = 0$.
Då följer ur KKT-1 att $y_1 = 2$ och $y_2 = 0$, varefter KKT-4 ger att $x_2 = 2$ eller -2.

Vi har alltså följande fyra KKT-punkter och tillhörande Lagrangemultiplikatorer:

$$\mathbf{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \text{ och } \mathbf{y} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \text{ och } \mathbf{y} = \begin{pmatrix} 0 \\ 4 \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \text{ och } \mathbf{y} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} \text{ och } \mathbf{y} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Uppgift 5.(a)

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

$$\nabla f(\mathbf{x}) = \left(\frac{-8}{(x_1+1)^2(x_2+1)} - 1 \,,\; \frac{-8}{(x_1+1)(x_2+1)^2} - 3 \right),$$

medan Hessianen till f ges av

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \dfrac{16}{(x_1+1)^3(x_2+1)} & \dfrac{8}{(x_1+1)^2(x_2+1)^2} \\ \dfrac{8}{(x_1+1)^2(x_2+1)^2} & \dfrac{16}{(x_1+1)(x_2+1)^3} \end{bmatrix}.$$

Speciellt är $\nabla f(\mathbf{x}^{(1)}) = (-2,\, -4)$ och $\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$.

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 $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$.

Vi prövar steget med $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$.

Då blir $f(\mathbf{x}^{(2)}) = 8/12 - 1 - 15 < -15$, medan $f(\mathbf{x}^{(1)}) = 8/4 - 1 - 3 = -2$

Eftersom $f(\mathbf{x}^{(2)}) < f(\mathbf{x}^{(1)})$ gick steget med $t_1 = 1$ bra, och vi accepterar $\mathbf{x}^{(2)} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$ som nästa iterationspunkt. Därmed har vi utfört en iteration med Newtons metod.

Uppgift 5.(b)

Vi ser att för alla värden på x_1 och x_2 är $f(\mathbf{x}) \leq 8 - x_1 - 3x_2$,
så om exempelvis $x_1 = 1$ och $x_2 \to +\infty$ så får vi att $f(\mathbf{x}) \to -\infty$.

Det betyder att det inte finns någon global minpunkt till $f(\mathbf{x})$, och därmed är det förstås omöjligt för Newtons metod att konvergera mot en sådan.

Uppgift 6

Kalla de sökta matriselementen för x_{ij} i stället för a_{ij} och inför variabelvektorn

$\mathbf{x} = (\ x_{11}\ \ x_{12}\ \ x_{13}\ \ x_{21}\ \ x_{22}\ \ x_{23}\ \ x_{31}\ \ x_{32}\ \ x_{33}\ \ )^{\mathsf{T}}$.

Då kan det betraktade problemet skrivas på formen

$$\begin{aligned} \text{minimera} \quad & \tfrac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{I}\,\mathbf{x} \\ \text{då} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned}$$

där

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ k_1 \\ k_2 \\ k_3 \end{pmatrix}$$

och $\mathbf{I}$ är en 9×9 enhetsmatris.

Detta är ett konvext QP-problem med linjära likhetsbivillkor, så $\hat{\mathbf{x}}$ är optimal om och endast om det finns en vektor $\hat{\mathbf{u}}$ så att $\hat{\mathbf{x}}$ och $\hat{\mathbf{u}}$ tillsammans uppfyller $\mathbf{I}\,\hat{\mathbf{x}} - \mathbf{A}^{\mathsf{T}}\hat{\mathbf{u}} = \mathbf{0}$ och $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$.

Antag att $\hat{\mathbf{x}} = (\ \hat{x}_{11}\ \ \hat{x}_{12}\ \ \hat{x}_{13}\ \ \hat{x}_{21}\ \ \hat{x}_{22}\ \ \hat{x}_{23}\ \ \hat{x}_{31}\ \ \hat{x}_{32}\ \ \hat{x}_{33}\ \ )^{\mathsf{T}}$ definieras av att

$\hat{x}_{ij} = \dfrac{r_i + k_j}{3}$, för $i = 1,2,3$ och $j = 1,2,3$.

Då är $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ uppfyllt (pga att $r_1 + r_2 + r_3 = 0$ och $k_1 + k_2 + k_3 = 0$).

Låt $\hat{\mathbf{u}} = (v_1, v_2, v_3, w_1, w_2, w_3)^{\mathsf{T}}$. Då kan villkoren $\mathbf{I}\,\hat{\mathbf{x}} - \mathbf{A}^{\mathsf{T}}\hat{\mathbf{u}} = \mathbf{0}$ skrivas

$\hat{x}_{ij} = v_i + w_j$ för $i = 1,2,3$ och $j = 1,2,3$, dvs

$\dfrac{r_i + k_j}{3} = v_i + w_j$ för $i = 1,2,3$ och $j = 1,2,3$,

som är uppfyllda för $v_i = \dfrac{r_i}{3}$ och $w_j = \dfrac{k_j}{3}$ för $i = 1,2,3$ och $j = 1,2,3$.

Alltså är $\hat{\mathbf{x}}$ optimal.