KTH Matematik

Lösningar till SF1852 Optimeringslära för E, 9/1–09

1. (**a**) Den tidsdiskreta modellen ger uphov till följande bivillkor

$$\mathbf{z}(1) = \begin{bmatrix} e^{-h} & 0 \\ 1-e^{-h} & 1 \end{bmatrix} \mathbf{z}(0) + \begin{bmatrix} 1-e^{-h} \\ h-1+e^{-h} \end{bmatrix} u(0)$$
$$\mathbf{z}(2) = \begin{bmatrix} e^{-h} & 0 \\ 1-e^{-h} & 1 \end{bmatrix} \mathbf{z}(1) + \begin{bmatrix} 1-e^{-h} \\ h-1+e^{-h} \end{bmatrix} u(1)$$

Om vi använder att $\mathbf{z}(0) = \begin{bmatrix} 0 & \theta \end{bmatrix}^\mathsf{T}$ samt $\mathbf{z}(2) = \mathbf{0}$ samt definition för $\mathbf{x}$ så får vi bivillkoret

$$\underbrace{\begin{bmatrix} -1 & 0 & 1-e^{-h} & 0 \\ 0 & -1 & h-1+e^{-h} & 0 \\ e^{-h} & 0 & 0 & 1-e^{-h} \\ 1-e^{-h} & 1 & 0 & h-1+e^{-h} \end{bmatrix}}_{\mathbf{A}} \mathbf{x} = \underbrace{\begin{bmatrix} 0 \\ -\theta \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{b}}$$

Målfunktionen kan skrivas

$$z_1(1)^2 + z_2(1)^2 + u(0)^2 + u(1)^2 = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}$$

med $\mathbf{H} = 2 \cdot \mathbf{I}_4$ (4×4 identitetsmatris).

Kommentar: Det kvadratiska styrproblemet i denna uppgift är för alla rimliga värden på samplingsintervalet h (dvs $h \neq 0$) ointressant eftersom bivillkoret har en unik lösning (matrisen $\mathbf{A}$ är inverterbar). Det finns således inget att optimera.

Om man däremot optimerar över fler tidssteg så blir problemet praktiskt intressant.

(**b**) Vi använder Gauss-Jordans metod på den givna matrisen. Elementära radoperationer ger upphov till följande sekvens av ekvivalenta matriser

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0.1 & 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0.1 & 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{T}$$

Nu är $\mathbf{A}$ överförd till trappstegsform med tre 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, 2 och 3 i $\mathbf{A}$. Vi får alltså

$$\mathcal{R}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0.1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

En bas till $\mathcal{N}(\mathbf{A})$ kan bestämmas enligt följande. Sätt $x_4 = 1$ (den enda variabel som inte svarar mot en trappstegsetta) och bestäm sedan x_1, x_2, och x_3 (variablerna svarande mot trappstegsettor) så att $\mathbf{Tx} = \mathbf{0}$. Detta ger att

$$\mathcal{N}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

2. (**a**) Problemet kan skrivas

$$\text{minimera} \quad \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$$

där

$$\mathbf{H} = \begin{bmatrix} 10 & -6 \\ -6 & 10 \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} 44 \\ -20 \end{bmatrix}$$

Problemet är konvext eftersom $\mathbf{H}$ är positivt definit och därmed gäller att

$$\hat{\mathbf{x}} = -\mathbf{H}^{-1}\mathbf{c} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$$

är en unik optimallösning.

(**b**) Problemet kan skrivas

$$\begin{array}{rl} \text{minimera} & \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} & \mathbf{Ax} = \mathbf{b} \end{array}$$

där

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Problemet kan lösas med Lagrangemetoden eller nollrumsmetoden. Eftersom $\mathbf{H} = I$ så är Lagrangemetoden enkel att använda. Här använder vi dock nollrumsmetoden. Efter några elementära radoperationerpå bivillkoret $\mathbf{Ax} = \mathbf{b}$ får vi ekvationssytemet

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Vi ser att $\bar{\mathbf{x}} = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}^{\mathsf{T}}$ är en lösning till bivillkoret och

$$\mathcal{N}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Med nollrumsmatrisen $\mathbf{Z} = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}^{\mathsf{T}}$ får vi

$$\mathbf{Z}^{\mathsf{T}}\mathbf{H}\mathbf{Z} = 3 \quad \text{(Positivt definit, dvs det existerar en unik optimallösning)}$$
$$\mathbf{Z}^{\mathsf{T}}\mathbf{H}\mathbf{Z}\hat{\mathbf{v}} = -\mathbf{Z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = 0 \quad \Rightarrow \quad \hat{\mathbf{v}} = 0$$

och därmed får vi att $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}} = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}^{\mathsf{T}}$ är en unik optimallösning.

3. (**a**) Från första ekvationen får vi $x_5 = x_1 + 2x_2 + x_3 + 2x_4 - 7$. Om vi sätter in detta i målfunktionen samt de två sista bivillkoren så får vi

$$f(\mathbf{x}) = 4x_1 + 7x_2 + x_3 + 5x_4 - 2(x_1 + 2x_2 + x_3 + 2x_4 - 7) = 2x_1 + 3x_2 - x_3 + x_4 + 14$$
$$g_1(\mathbf{x}) = 2x_1 + 3x_2 + x_3 + x_4 - (x_1 + 2x_2 + x_3 + 2x_4 - 7) - 6 = x_1 + x_2 - x_4 + 1$$
$$g_2(\mathbf{x}) = x_1 + x_2 + 2x_3 + x_4 - (x_1 + 2x_2 + x_3 + 2x_4 - 7) - 4 = -x_2 + x_3 - x_4 + 3$$

Det reducerade LP problemet blir på standard form

$$\begin{array}{rl} \text{minimera} & 2x_1 + 3x_2 - x_3 + x_4 + 14 \\ \text{då} & -x_1 - x_2 + x_4 = 1 \\ & x_2 - x_3 + x_4 = 3 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq,\ x_4 \geq 0 \end{array}$$

(**b**) Enligt förslag skall vi starta med $\beta = (1, 4)$. Vi har då

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

Vi beräknar motsvarande baslösning, simplexmultiplikatorer samt reducerade kostnader

$$\bar{\mathbf{b}} = \mathbf{x}_\beta = \mathbf{A}_\beta^{-1}\mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$
$$\mathbf{y} = (\mathbf{A}_\beta^{\mathsf{T}})^{-1}\mathbf{c}_\beta = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$
$$\mathbf{r}_\delta = \mathbf{c}_\delta - \mathbf{A}_\delta^{\mathsf{T}}\mathbf{y} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

Eftersom $r_2 < 0$ så skall x_2 in i basen. Utför kvottest för att avgöra vilken variabel som skall ut ur basen

$$\bar{\mathbf{a}}_2 = \mathbf{A}_\beta^{-1}\mathbf{a}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$t^{\max} = \min_i \left\{ \frac{\bar{\mathbf{b}}_i}{\bar{\mathbf{a}}_{i2}} | \bar{\mathbf{a}}_{i2} > 0 \right\} = \min_i \left\{ \frac{2}{2}, \frac{3}{2} \right\} = 1$$

Det följer att x_1 skall ut ur basen. För nästa simplexiteration har vi därmed

$$\beta = (2,4) \qquad \delta = (1,3)$$

$$\mathbf{A}_\beta = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \qquad \mathbf{A}_\delta = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\mathbf{c}_\beta^\mathsf{T} = \begin{bmatrix} 3 & 1 \end{bmatrix} \qquad \mathbf{c}_\delta^\mathsf{T} = \begin{bmatrix} 2 & -1 \end{bmatrix}$$

Vi beräknar motsvarande baslösning, simplex multiplikatorer samt reducerade kostnader

$$\begin{aligned} \bar{\mathbf{b}} = \mathbf{x}_\beta = \mathbf{A}_\beta^{-1}\mathbf{b} &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ \mathbf{y} = (\mathbf{A}_\beta^\mathsf{T})^{-1}\mathbf{c}_\beta &= \begin{bmatrix} -1 \\ 2 \end{bmatrix} \\ \mathbf{r}_\delta = \mathbf{c}_\delta - \mathbf{A}_\delta^\mathsf{T}\mathbf{y} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

Eftersom $r_\delta \geq 0$ så har vi hittat en optimal baslösning, $\hat{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 0 & 2 \end{bmatrix}^\mathsf{T}$.

(c) Den optimala lösningen till det ursprungliga problemet blir därmed

$$\begin{aligned} \hat{x}_1 &= 0 \\ \hat{x}_2 &= 1 \\ \hat{x}_3 &= 0 \\ \hat{x}_4 &= 2 \\ \hat{x}_5 &= \hat{x}_1 + 2\hat{x}_2 + \hat{x}_3 + 2\hat{x}_4 - 7 = -1 \end{aligned}$$

4. **(a)** Optimeringsproblemet kan skrivas

$$\begin{aligned} \text{minimera} \quad & f(\mathbf{x}) \\ \text{då} \quad & g_1(\mathbf{x}) \leq 0 \\ & g_2(\mathbf{x}) \leq 0 \\ & g_3(\mathbf{x}) \leq 0 \end{aligned}$$

där

$$\begin{aligned} f(\mathbf{x}) &= (x_1+1)^2 + (x_2+2)^2 \\ g_1(\mathbf{x}) &= x_1 - 2x_2 - 2 \\ g_2(\mathbf{x}) &= 4 - x_1^2 - x_2^2 \\ g_3(\mathbf{x}) &= -x_1 \end{aligned}$$

Gradienterna är

$$\begin{aligned} \nabla f(\mathbf{x}) &= \begin{bmatrix} 2(x_1+2) & 2(x_2+2) \end{bmatrix} \\ \nabla g_1(\mathbf{x}) &= \begin{bmatrix} 1 & -2 \end{bmatrix} \\ \nabla g_2(\mathbf{x}) &= \begin{bmatrix} -2x_1 & -2x_2 \end{bmatrix} \\ \nabla g_3(\mathbf{x}) &= \begin{bmatrix} -1 & 0 \end{bmatrix} \end{aligned}$$

Vi provar punkten $\bar{x} = \begin{bmatrix} 0 & 2 \end{bmatrix}^\mathsf{T}$. Vi får

$$\begin{aligned} \nabla f(\bar{\mathbf{x}}) &= \begin{bmatrix} 4 & 8 \end{bmatrix} \\ \nabla g_1(\bar{\mathbf{x}}) &= \begin{bmatrix} 1 & -2 \end{bmatrix} \\ \nabla g_2(\bar{\mathbf{x}}) &= \begin{bmatrix} 0 & -4 \end{bmatrix} \\ \nabla g_3(\bar{\mathbf{x}}) &= \begin{bmatrix} -1 & 0 \end{bmatrix} \end{aligned}$$

KKT villkoren är

$$\begin{array}{ll} \text{KKT1} & \begin{bmatrix} 4 & 8 \end{bmatrix} + \begin{bmatrix} 1 & -2 \end{bmatrix} y_1 + \begin{bmatrix} 0 & -4 \end{bmatrix} y_2 + \begin{bmatrix} -1 & 0 \end{bmatrix} y_3 = \begin{bmatrix} 0 & 0 \end{bmatrix} \\ & g_1((0,2)) = -6 < 0 \quad \text{ok} \\ \text{KKT2} & g_2((0,2)) = 0 = 0 \quad \text{ok} \\ & g_3((0,2)) = 0 = 0 \quad \text{ok} \\ & y_1 \geq 0 \\ \text{KKT3} & y_2 \geq 0 \\ & y_3 \geq 0 \\ & y_1 \cdot (-6) = 0 \\ \text{KKT4} & y_2 \cdot 0 = 0 \\ & y_3 \cdot 0 = 0 \end{array}$$

Från KKT4 får vi att $y_1 = 0$. Vi ser då att KKT1 och KKT3 är uppfyllda om $y_1 = 0$, $y_2 = 2$ och $y_3 = 4$. $\bar{\mathbf{x}} = \begin{bmatrix} 0 & 2 \end{bmatrix}^\mathsf{T}$ är således en KKT-punkt.

Vi provar nu punkten $\hat{x} = \begin{bmatrix} 2 & 0 \end{bmatrix}^\mathsf{T}$. Vi får

$$\begin{aligned} \nabla f(\hat{\mathbf{x}}) &= \begin{bmatrix} 8 & 4 \end{bmatrix} \\ \nabla g_1(\hat{\mathbf{x}}) &= \begin{bmatrix} 1 & -2 \end{bmatrix} \\ \nabla g_2(\hat{\mathbf{x}}) &= \begin{bmatrix} -4 & 0 \end{bmatrix} \\ \nabla g_3(\hat{\mathbf{x}}) &= \begin{bmatrix} -1 & 0 \end{bmatrix} \end{aligned}$$

KKT villkoren är

$$\begin{array}{ll} \text{KKT1} & \begin{bmatrix} 8 & 4 \end{bmatrix} + \begin{bmatrix} 1 & -2 \end{bmatrix} y_1 + \begin{bmatrix} -4 & 0 \end{bmatrix} y_2 + \begin{bmatrix} -1 & 0 \end{bmatrix} y_3 = \begin{bmatrix} 0 & 0 \end{bmatrix} \\ & g_1((2,0)) = 0 = 0 \quad \text{ok} \\ \text{KKT2} & g_2((2,0)) = 0 = 0 \quad \text{ok} \\ & g_3((2,0)) = -2 < 0 \quad \text{ok} \\ & y_1 \geq 0 \\ \text{KKT3} & y_2 \geq 0 \\ & y_3 \geq 0 \\ & y_1 \cdot 0 = 0 \\ \text{KKT4} & y_2 \cdot 0 = 0 \\ & y_3 \cdot (-4) = 0 \end{array}$$

Från KKT4 får vi att $y_3 = 0$. Vi ser då att KKT1 och KKT3 är uppfyllda om $y_1 = 2$, $y_2 = 2.5$ och $y_3 = 0$. $\hat{\mathbf{x}} = \begin{bmatrix} 2 & 0 \end{bmatrix}^\mathsf{T}$ är således en KKT-punkt.

(b) I en KKT-punkt kan målfunktionens gradient skrivas som en negativ linjärkombination (negativa koefficienter) av de bindande (aktiva) bivillkorens gradienter. I Figur 1

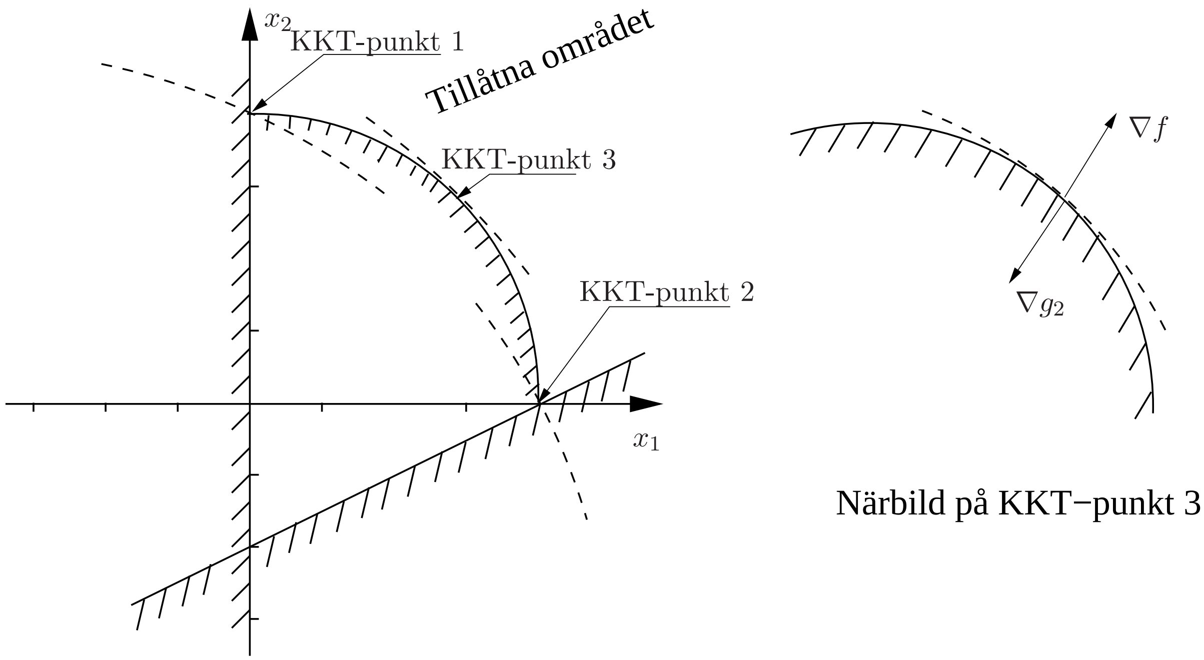

Figure 1: Den vänstra figuren illustrerar de tre bivillkoren samt tre isokostlinjer för målfunktionen som passerar genom de tre KKT-punkterna. I den högra figuren visas en närbild på den tredje KKT-punkten.

ser vi att det finns tre KKT-punkter. KKT-punkt 1 och KKT-punkt 2 är de två punkterna som studerades i problem **a**. I KKT-punkt 3, $\check{x} = (\sqrt{2}, \sqrt{2})$, är endast det andra bivillkoret bindande.

5. (**a**) Bivillkoret kan skrivas om enligt

$$
\begin{aligned}
& \frac{g_{k,k}x_k}{1 + \sum_{l=1,l\neq k}^{n} g_{k,l}x_l} \geq \gamma_k,\ k = 1, \dots, n \\
\Leftrightarrow \quad & g_{k,k}x_k - \gamma_k \sum_{l=1,l\neq k}^{n} g_{k,l}x_l \geq \gamma_k,\ k = 1, \dots, n \\
\Leftrightarrow \quad & \frac{g_{k,k}}{\gamma_k}x_k - \sum_{l=1,l\neq k}^{n} g_{k,l}x_l \geq 1,\ k = 1, \dots, n \\
\Leftrightarrow \quad & \mathbf{Gx} \geq \mathbf{1}
\end{aligned}
$$

Nedlänksproblemet är således ekvivalent med LP-problemet

$$
\begin{aligned}
\text{minimera} \quad & \mathbf{1}^{\mathsf{T}}\mathbf{x} \\
\text{då} \quad & \mathbf{Gx} \geq \mathbf{1} \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned} \tag{0.1}
$$

(**b**) En liknande omskrivning som i problem (**a**) ger upplänkproblemet kan skrivas

$$\begin{array}{rl} \text{maximera} & \mathbf{1}^{\mathsf{T}}\mathbf{y} \\ \text{då} & \mathbf{G}^{\mathsf{T}}\mathbf{y} \leq \mathbf{1} \\ & \mathbf{y} \geq \mathbf{0} \end{array} \tag{0.2}$$

Optimeringsproblemen (0.1) och (0.2) är ett primal-dualt par. Då det antagits att en optimal lösning finns så gäller att $\mathbf{1}^{\mathsf{T}}\hat{\mathbf{x}} = \mathbf{1}^{\mathsf{T}}\hat{\mathbf{y}}$.