KTH Matematik

Lösningar för Tentamen i SF1851 Optimeringslära för E.
fredag Oktober 23, 2009, tid. 08.00–13.00

Examinator: Per Enqvist, tel. 790 62 98

Det kan finnas alternativa lösningar.

---

1. (a) Vi har tre ekvationer som borde gälla om mätningarna vore exakta.

$$U_1 = I_1 R_1, \quad \text{d.v.s } 12 = 0.3 R_1$$

$$U_2 = I_2 R_2, \quad \text{d.v.s } 12 = 0.4 R_2$$

$$U_3 = I_3 R_3, \quad \text{d.v.s } 12 = 0.6/(1/R_1 + 1/R_2)$$

och med ledningsförmågan som variabler

$$12 L_1 = 0.3, \quad 12 L_2 = 0.4, \quad 12(L_1 + L_2) = 0.6.$$

Minsta kvadratproblemet blir att minimera kvadratfelen i dessa ekvationer, d.v.s. vi ska minimera

$$f(L_1, L_2) = (12L_1 - 0.3)^2 + (12L_2 - 0.4)^2 + (12(L_1 + L_2) - 0.6)^2.$$

Detta kan skrivas som $f(L) = \|AL - b\|^2$ där $L = [L_1,\ L_2]^T$ och

$$A = \begin{bmatrix} 12 & 0 \\ 0 & 12 \\ 12 & 12 \end{bmatrix}, \quad b = \begin{bmatrix} 0.3 \\ 0.4 \\ 0.6 \end{bmatrix}.$$

Optimal lösning ges då av $L = (A^T A)^{-1} A^T b = (1/45, 11/360)$ och motsvarande resistanser blir $R = (45, 360/11)$.

(b) Vi börjar med att överföra A på trappstegsform genom att göra elementära radoperationer:

$$A = \begin{bmatrix} 2 & 4 & 1 & 2 \\ 4 & 4 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1/2 & 1 \\ 0 & -4 & -2 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1/2 & -1/2 \\ 0 & 1 & 1/2 & 3/4 \end{bmatrix}$$

Eftersom de två första kolumnerna är enhetsvektorer så spänns bildrummet av de två första kolumnerna i A, dvs bildrummet spänns av vektorerna $[2,\ 4]^T$ och $[4,\ 4]^T$.

Nollrummet spänns av vektorerna

$$\begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1/2 \\ -3/4 \\ 0 \\ 1 \end{bmatrix}.$$

2. (a) Nodbalans ges av ekvationerna $Ax = b$, där

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 10 \\ 0 \\ 0 \\ 0 \\ -3 \\ -2 \\ -5 \end{bmatrix},$$

(b) Baslösningen ges genom att bara betrakta flöde genom det spännande trädet och upprätta nodbalans, basvariablerna ges då av

$$x_{12} = 0, \quad x_{13} = 8, \quad x_{14} = 2, \quad x_{35} = 8, \quad x_{46} = 2, \quad x_{57} = 5.$$

(c) Låt nodpotentialen y_7 vid nod 7 vara 0. $y_5 - y_7 = c_{57} = 3$ ger $y_5 = 3$.
$y_3 - y_5 = c_{35} = 2$ ger $y_3 = 5$.
$y_1 - y_3 = c_{13} = 3$ ger $y_1 = 8$.
$y_1 - y_2 = c_{12} = 2$ ger $y_2 = 6$.
$y_1 - y_4 = c_{14} = 1$ ger $y_4 = 7$.
$y_4 - y_6 = c_{46} = 1$ ger $y_6 = 6$.
De reducerade kostnaderna är nu $r_{25} = c_{25} - y_2 + y_5 = 4 - 6 + 3 = 1$, $r_{36} = c_{36} - y_3 + y_6 = 3 - 5 + 6 = 4$. and $r_{67} = c_{67} - y_6 + y_7 = 2 - 6 + 0 = -4$. Eftersom den reducerade kostnaden r_{67} är negativ så ska den in i basen. Vi lägger till motsvarande båge i grafen och då bildas en cykel. Variabeln x_{67} kan ökas tills den blir 5 varvid x_{57} blir noll och går ur basen. Ny baslösning ges då av

$$x_{12} = 0, \quad x_{13} = 3, \quad x_{14} = 7, \quad x_{35} = 3, \quad x_{46} = 7, \quad x_{67} = 5.$$

Låt nodpotentialen y_7 vid nod 7 vara 0. $y_6 - y_7 = c_{67} = 2$ ger $y_6 = 2$.
$y_4 - y_6 = c_{46} = 1$ ger $y_4 = 3$.
$y_1 - y_4 = c_{14} = 1$ ger $y_1 = 4$.
$y_1 - y_2 = c_{12} = 2$ ger $y_2 = 2$.
$y_1 - y_3 = c_{13} = 3$ ger $y_3 = 1$.
$y_3 - y_5 = c_{35} = 2$ ger $y_5 = -1$.
De reducerade kostnaderna är nu $r_{25} = c_{25} - y_2 + y_5 = 4 - 2 + 1 = 1$, $r_{36} = c_{36} - y_3 + y_6 = 3 - 1 + 2 = 4$. and $r_{57} = c_{57} - y_5 + y_7 = 3 - (-1) + 0 = 4$. Eftersom alla reducerade kostnader är positiva så är den nuvarande baslösningen det unika optimat.

3. (a) Vi börjar med x_1 och x_2 som basvariabler. D.v.s bas och icke-basvariabel index ges av $\beta = \{1, 2\}$ och $\eta = \{3, 4\}$, så

$$B = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}, N = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

och $\bar{b} = B^{-1}b = [1\ 1]^T$ vilket ger startlösningen $x = (1, 1, 0, 0)$.

Nu ger ekvationerna $B^T y = c_B$ och $\hat{c}_N^T = c_N^T - y^T N$ att

$$y = \begin{bmatrix} -2 \\ 5/2 \end{bmatrix}, \quad r_N^T = \begin{bmatrix} -3/2 & 1/2 \end{bmatrix}.$$

Låt x_3 gå in i basen. Vilken ska gå ut ?

Från $B\hat{a}_4 = a_4$, får vi att $\hat{a}_4 = (3/2, -1/2)^T$, och eftersom det första elementet är det enda positiva så går x_1 ur basen.

Uppdatera bas och ickebasmatriserna: Bas och icke-basvariablernas index ges av $\beta = \{2, 3\}$ och $\eta = \{1, 4\}$, och

$$B = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}, N = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Ekvationerna $B^T y = c_B$ och $\hat{c}_N^T = c_N^T - y^T N$ ger

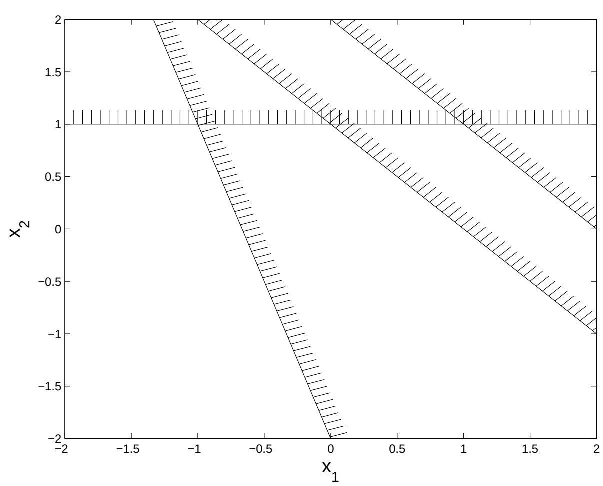
$$y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad r_N^T = \begin{bmatrix} 2 & 1 \end{bmatrix}.$$

Eftersom alla reducerade kostnader är ickenegativa, så är $\hat{x} = (0, 4/3, 2/3, 0)^T$ optimal.

(b) Dualen blir

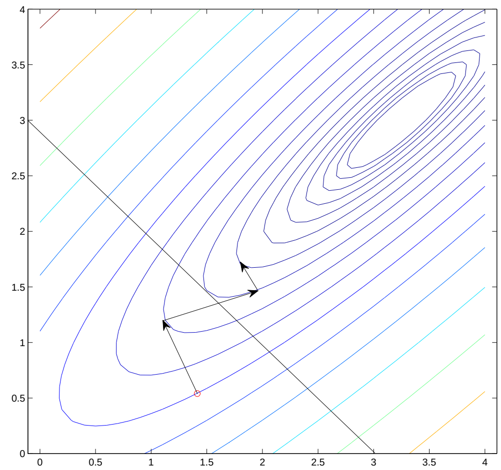
$$(D) \quad \begin{bmatrix} \min\limits_{y} & 4y_1 + 2y_2 \\ \text{då} & y_1 + y_2 \leq 2 \\ & 3y_1 + y_2 \leq -2 \\ y_2 \leq 1 & \\ & y_1 + y_2 \leq 1 \end{bmatrix}.$$

(c) Tillåtna området för dualen ges av



(d) Optimum för dualen ges för $y = (-1, 1)$, ses t.ex. från figur. Vi kontrollerar att den är tillåten för dualen, dvs. $A^T y \leq c$. Dessutom gäller att målfunktionen för dualen $b^T y = -4 + 2 = -2$, och målfunktionen för primalen $c^T x = 4/3 * (-2) + 2/3 = -6/3 = -2$, och enligt svag dualitet så måste alltså vårat y vara optimal till dualen.

4. (a) Iterationerna för gradientmetoden kan ses i figuren nedan.



Det som är viktigast att notera är att sökriktningarna är ortogonala mot den nivåkurva som man startar vid, och att man går så långt i den riktningen att man precis kommer att tangera en ny nivåkurva i nästa iterationspunkt. Det leder till att nästa sökriktning blir ortogonal mot den föregående.

(b) Matrisen H kan faktoriseras som $H = LDL^T$ där

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

Eftersom diagonalelementen i D är positiva så är matrisen H positivt definit och den kvadratiska målfunktionen är konvex.

(c) Lösningen till problemet är unik eftersom H är positivt definit, och den ges av det x som löser $Hx = -c$, dvs av $x = (13/4,\ 3)$.

(d) Först bestämmer vi nollrummet till A, det ges av

$$Z = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

En lösning $\bar{x}$ till bivillkoret ges av

$$\bar{x} = \begin{bmatrix} 0 \\ 3 \end{bmatrix},$$

Ekvationssystemet $(ZHZ^T)v = -Z^T(H\bar{x}+c)$, i.e. $17v = -(-26)$ ger $v = 26/17$ Slutligen, $\hat{x} = \bar{x} + Zv = [26/17\ 25/17]^T$ är ett globalt min eftersom problemet är konvext.

5. (a) Målfunktionen $f(x)$ är summan av de konvexa funktionerna $e^{x_1+x_2}$ och $-x_1 - 3x_2$, så den är också konvex.

Det tillåtna området är ett snitt av två mängder. Eftersom snittet inte är sammanhängande, utan består av två disjunkta mängder, så är det inte konvext.

Eftersom det tillåtna området inte är konvext så är inte optimeringsproblemet konvext.

(b) Om vi vet att mängderna $\mathcal{C}_1$ och $\mathcal{C}_2$ båda är konvexa, så vet vi enligt en sats att deras snitt också är konvext. Detta villkor är dock bara tillräckligt. Om tex $\mathcal{C}_1$ bara består av en punkt, och $\mathcal{C}_1$ är en godtycklig mängd så är deras snitt också konvext. Dvs, det är inte nödvändigt att båda mängderna är konvexa för att snittet ska vara konvex.

(c) Låt

$$g_1(x) = -e^{x_1} + x_2, \quad g_2(x) = e \cdot x_1^2 - x_2.$$

vilket ger

$$\nabla f(x) = \begin{bmatrix} e^{x_1+x_2} - 1 & e^{x_1+x_2} - 3 \end{bmatrix},$$

$$\nabla g_1(x) = \begin{bmatrix} -e^{x_1} & 1 \end{bmatrix}, \quad \nabla g_2(x) = \begin{bmatrix} 2e \cdot x_1 & -1 \end{bmatrix}.$$

I punkten $\mathbf{x}^{(c)} = (0,0)$ gäller att $g_1(\mathbf{x}^{(c)}) = -1$ och $g_2(\mathbf{x}^{(c)}) = 0$. Bivillkor 1 är inte aktivt och alltså måste $\hat{y}_1 = 0$.

Men då har inte ekvationen

$$\nabla f(\mathbf{x}^{(c)}) + \hat{y}_1 \nabla g_1(\mathbf{x}^{(c)}) = \begin{bmatrix} e^0 - 1 & e^0 - 3 \end{bmatrix} + \hat{y}_2 \begin{bmatrix} 0 & -1 \end{bmatrix} = 0$$

någon lösning och KKT-villkoren är inte uppfyllda. Alltså kan inte $\mathbf{x}^{(c)} = (0,0)$ vara en lokal minpunkt.

(d) I punkten $\mathbf{x}^{(d)} = (1, e)$ gäller att $g_1(\mathbf{x}^{(c)}) = 0$ och $g_2(\mathbf{x}^{(c)}) = 0$. Båda bivillkoren är alltså aktiva och alltså kan både $\hat{y}_1$ och $\hat{y}_2$ vara nollskilda.

Då har ekvationen

$$\nabla f(\mathbf{x}^{(d)}) + \hat{y}_1 \nabla g_1(\mathbf{x}^{(d)}) + \hat{y}_2 \nabla g_2(\mathbf{x}^{(d)}) =$$

$$= \begin{bmatrix} e^{1+e} - 1 & e^{1+e} - 3 \end{bmatrix} + \hat{y}_1 \begin{bmatrix} -e^1 & 1 \end{bmatrix} + \hat{y}_2 \begin{bmatrix} 2e^1 & -1 \end{bmatrix} = 0$$

en lösning

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} -e & 2e \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} e^{1+e} - 1 \\ e^{1+e} - 3 \end{bmatrix} = \frac{1}{3e} \begin{bmatrix} 1 & 2e \\ 1 & e \end{bmatrix} \begin{bmatrix} e^{1+e} - 1 \\ e^{1+e} - 3 \end{bmatrix}$$

som är positiv och KKT-villkoren är uppfyllda.

Alltså kan $\mathbf{x}^{(c)} = (0,0)$ vara en lokal minpunkt, men eftersom problemet inte är konvext så är KKT-villkoren bara nödvändiga men inte tillräckliga, så vi kan inte säga att det är en lokal minpunkt.

(e) I en inre punkt så är inga bivillkor aktiva och alltså måste både $\hat{y}_1$ och $\hat{y}_2$ vara lika med noll. Eftersom $\nabla f(x) = \begin{bmatrix} e^{x_1+x_2} - 1 & e^{x_1+x_2} - 3 \end{bmatrix}$ så finns det inga punkter sådana att gradienten är noll, och alltså kan inte KKT-villkoren uppfyllas i en inre punkt.