KTH Mathematics

Lösningsförslag till tentamen i SF1861 Optimeringslära för T.
Onsdag 25 augusti 2010 kl. 14.00–19.00

Examinator: Per Enqvist, tel. 790 62 98

---

1. (a) Vi har ett nätverksflödesproblem med 5 noder. Låt $x = (x_{12}, x_{13}, x_{24}, x_{25}, x_{34}, x_{45})^T$ vara variabelvektorn och vi noterar att den angivna punkten är en tillåten lösning till problemet och en baslösning eftersom den svarar mot ett spännande träd. Börja nätverksversionen av simplexalgoritmen i den angivna punkten.

Låt nodpotentialen i nod 5 vara noll, dvs låt $y_5 = 0$. Utnyttja att $r_{jk} = c_{jk} - y_j + y_k = 0$ för basvariabler, vilket ger att $y_2 = 3$, $y_1 = 7$, $y_3 = 5$ och $y_4 = 2$. För de två ickebasvariablerna har vi då att de reducerade kostnaderna ges av $r_{34} = 4 - 5 + 2 = 1$ och $r_{45} = 3 - 2 + 0 = 1$ båda är positiva och den givna punkten är optimal.

(b) Utför rad- och kolonnoperation på H tills vi erhåller matrisen

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -2/3 \end{bmatrix}$$

och vi noterar att matrisen inte är positiv på diagonal och alltså varken är positivt definit eller positivt semidefinit. Problemet att minimera $f(x)$ är därför inte begränsat underifrån och det finns ingen optimal punkt för problemet.

2. (a) Problemet (P) kan skrivas på standardform

$$(P) \quad \begin{bmatrix} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{bmatrix}$$

där

$$A = \begin{bmatrix} 1 & 0 & 3 & 1 & 0 \\ 1 & 1 & -1 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad c = \begin{bmatrix} -1 & -2 & -1 & 0 & 0 \end{bmatrix}^T.$$

Vi börjar med slackvariablerna x_4 and x_5 i basen, d.v.s. $\beta = \{4, 5\}$ och $\delta = \{1, 2, 3\}$.

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, N = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & -1 \end{bmatrix}$$

Då ger ekvationerna $A_\beta^T y = c_B$ och $r_\delta^T = c_\delta^T - y^T A_\delta$ att

$$y = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad r_\delta^T = \begin{bmatrix} -1 & -2 & -1 \end{bmatrix}.$$

Låt x_2 gå in i basen. Vilken variabel ska lämna basen ?
Från $A_\beta \bar{a}_4 = a_4$, får vi $\bar{a}_4 = (0,1)^T$, och eftersom det första elementet är noll, så måste den andra basvariabeln, nämligen x_5, lämna basen.
Nu är $\beta = \{4,2\}$ och $\delta = \{1,5,3\}$, och vi uppdaterar bas- och ickebasmatriser:

$$A_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_\delta = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & -1 \end{bmatrix}$$

Nu ger ekvationerna $A_\beta^T y = c_\beta$ och $r_\delta^T = c_\delta^T - y^T A_\delta$ att

$$y = \begin{bmatrix} 0 \\ -2 \end{bmatrix}, \quad r_\delta^T = \begin{bmatrix} 1 & 2 & -3 \end{bmatrix}.$$

Låt x_3 gå in i basen. Vilken variabel ska lämna basen ?
Från $A_\beta \bar{a}_3 = a_3$, får vi $\bar{a}_3 = (3,-1)^T$, och eftersom det andra elementet är negativt, så måste den första basvariabeln, nämligen x_4, lämna basen.
Nu är $\beta = \{3,2\}$ och $\delta = \{1,5,4\}$, och vi uppdaterar bas- och ickebasmatriser:

$$A_\beta = \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}, A_\delta = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Nu ger ekvationerna $A_\beta^T y = c_\beta$ och $r_\delta^T = c_\delta^T - y^T A_\delta$ att

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

Eftersom alla reducerade kostnader är positiva så är den aktuella baslösningen $\hat{x} = (0, \frac{4}{3}, \frac{1}{3}, 0, 0)$ optimal.

(b) Det duala problemet blir

$$(D) \quad \begin{bmatrix} \max\limits_{y} & 2y_1 + y_2 \\ \text{då} & y_1 - y_2 \leq 1 \\ & y_1 + y_2 \leq 2 \\ & y_1 - y_2 \leq -2 \\ & 2y_1 + y_2 \leq 5 \end{bmatrix}$$

Komplementaritet ger att bivillkor 2 och 3 är uppfyllda med likhet och vi ser att $y_1 = 0$ och $y_2 = 2$ uppfyller detta ekvationssystem. Dessutom uppfylls övriga bivillkor med strikt olikhet.

(c) Först och främst uppfyller x_b alla bivillkor, så det är en tillåten punkt.
Vi kollar att $c^T x = b^T y$, dvs $2 * 3/2 - 2 * 1/2 = 2 = 2 * 0 + 1 * 2$, alltså är x optimal för primalen och y optimal för dualen enligt dualitetssatsen.

3. (a) Genom LDL^T-faktorisering ser vi att:
Matris H_1 är inte positivt semidefinit, så motsvarande optimeringsproblem har ingen ändlig lösning.

Matris H_2 är positivt definit, så motsvarande optimeringsproblem har en unik ändlig lösning.

Matriserna H_3 och H_4 är positivt semidefinita, men inte positivt definita, så motsvarande optimeringsproblem har en ändlig lösning om och endast om $c_j \in \mathcal{R}(H_j)$.

Eftersom $c_3 \in \mathcal{R}(H_3)$, så har motsvarande optimeringsproblem en ändlig lösning. (t.o.m. oändligt många lösningar som ligger på en linje i $I\!R^2$)

Eftersom $c_4 \notin \mathcal{R}(H_4)$, så har motsvarande optimeringsproblem ingen ändlig lösning.

(b) Bestäm nollrummet till A

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

Det spänns av kolumnvektorn

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

Den reducerade hessianen $Z^T H_1 Z = 6$ är positivt definit, så motsvarande optimeringsproblem har en unik ändlig lösning.

Den reducerade hessianen $Z^T H_2 Z = 5$ är positivt definit, så motsvarande optimeringsproblem har en unik ändlig lösning.

Den reducerade hessianen $Z^T H_3 Z = 8$ är positivt definit, så motsvarande optimeringsproblem har en unik ändlig lösning.

Den reducerade hessianen $Z^T H_4 Z = 8$ är positivt definit, så motsvarande optimeringsproblem har en unik ändlig lösning.

4. Gradienten till f är $\nabla f(x) = [2x_1 + 2x_3 + 1,\ 4x_2 + 4x_3,\ 10x_3 + 2x_1 + 4x_2 + 4x_3^3]$, och hessianen ges av

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 4 & 4 \\ 2 & 4 & 10 + 12x_3^2 \end{bmatrix}.$$

(a) Vi LDL^T-faktoriserar, efter första steget får vi

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 4 \\ 0 & 4 & 8 + 12x_3^2 \end{bmatrix},$$

efter andra steget får vi

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 + 12x_3^2 \end{bmatrix},$$

och denna diagonalmatris är positivt definit för alla x och alltså är f strikt konvex.

(b) I punkten $x^{(0)} = 0$ har vi då $d^{(0)} = -\nabla f(x^{(0)})$, dvs $d^{(0)} = -(1, 0, 0)$.
Nästa punkt väljs med hjälp av exakt linjesökning, minimera $g(\alpha) = f(x^{(0)} + \alpha d^{(0)}) = f(-\alpha, 0, 0) = \alpha^2 - \alpha$ för positiva α.
f är en kvadratisk funktion med positiv hessian, och därför konvex, och har alltså ett unikt minimum som ants där derivatan är noll, dvs i punkten $\hat{\alpha} = 1/2$ (vilket är i det tillåtna området).
Nästa punkt ges då av $x^{(1)} = x^{(0)} + \hat{\alpha} d^{(0)} = (-1/2, 0, 0)$.

(c) Det tillåtna området ges av $0 \le x_j \le 1$ för $j = 1, 2, 3$. Det kan beskrivas med hjälp av $g_i(x) \le 0$ för $i = 1, 2, 3, 4, 5, 6$ där

$$g_1(x) = -x_1, \quad g_2(x) = -x_2, \quad g_3(x) = -x_3,$$
$$g_4(x) = x_1 - 1, \quad g_5(x) = x_2 - 1, \quad g_6(x) = x_3 - 1.$$

Då kan problemet skrivas min $f(x)$ då $g(x) \le 0$.
I punkten $x^* = (0, 0, 0)$, är bivillkoren 1,2,3 aktiva och 4,5,6 inaktiva (och alltså måste $\lambda_4, \lambda_5, \lambda_6$ alla vara noll).

$$\nabla f(x^*) + \lambda \nabla g(x^*) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = 0.$$

Detta är sant för $\lambda_1 = 1 > 0$, $\lambda_2 = 0$, $\lambda_3 = 0$. Alltså är KKT1 och KKT3 uppfyllda. Eftersom x^* är tillåten så är även KKT2 uppfyllt och i och med att bivillkor 1,2,3 är aktiva och 4,5,6 är uppfyllda med strikt olikhet så är KKT4 uppfyllt.
Problemet är konvext, så KKT-villkoren är ekvivalenta med de globala optimalitetsvillkoren och vi kan alltså säga att x^* är ett globalt minimum.

5. (a) Området är inte konvext. Punkterna $x^{(1)} = (1, 0)$ och $x^{(2)} = (-1, 0)$ är tillåtna, men punkten $1/2x^{(1)} + 1/2x^{(2)} = 0$ är inte tillåten.
Tillåtna området bestäms av att $h(x) = x_1^2 + x_2^2 - 1 = 0$. Gradienten till denna funktion ges av $\nabla h(x) = (2x_1,\ 2x_2\ 0)$.
Målfunktionen ges av $f(x) = Ax_1 + Bx_2$. Gradienten till denna funktion ges av $\nabla f(x) = (A,\ B)$.
Vi har att $\nabla f + \lambda \nabla g = (A + \lambda 2x_1,\ B + \lambda 2x_2\ 0) = 0$, om och endast om $x_1 = -A/(2\lambda)$ och $x_2 = -B/(2\lambda)$.
Värdet på λ bestäms genom att sätta in dessa punkter i bivillkoret,

$$x_1^2 + x_2^2 = \frac{A^2}{4\lambda^2} + \frac{B^2}{4\lambda^2} = 1$$

vilket ger $\lambda = \pm\frac{1}{2}\sqrt{A^2 + B^2}$.
Funktionerna f och h är differentierbara oändlig många gånger i hela $\mathbb{R}^2$, gradienten för h är nollskild för alla tillåtna punkter, så problemet är reguljärt. Minimum måste då antas i en punkt som uppfyller KKT-villkoren och alltså i en av punkterna $\left(\frac{A}{\sqrt{A^2+B^2}}, \frac{B}{\sqrt{A^2+B^2}}\right)$ och $-\left(\frac{A}{\sqrt{A^2+B^2}}, \frac{B}{\sqrt{A^2+B^2}}\right)$. Genom att jämföra målfunktionsvärdena i dessa två punkter ser vi att minimum antas i den senare av punkterna. (maximum antas i den andra)

(b) Det nya problemet har målfunktionen

$$g(t) = f(x(t)) = A\cos(t) + B\sin(t).$$

Det tillåtna området är nu konvext, men istället så är målfunktionen inte konvex eftersom $g''(t) = -A\cos(t) - B\sin(t)$ kommer att vara negativ för något $t \in [-\pi, \pi]$ om inte både A och B är noll. Så problemet är inte konvext nu heller.

Låt $t^{(0)} = 0$. Newtons metod ger nu att vi bestämmer riktningen till nästa punkt från $g''(0)d^{(0)} = -g'(0)$, dvs $-Ad^{(0)} = -B$ så $d^{(0)} = B/A$.

Med ett enhetssteg blir nu nästa iterationspunkt $t^{(1)} = t^{(0)} + d^{(0)} = 0 + B/A = B/A$.

Newtonmetoden ger en descentriktning om hessianen är positiv i iterationspunkten, dvs om $A < 0$.