KTH Mathematics

Lösningsförslag till tentamen i SF1861 Optimeringslära för T.
Tisdag 24 Maj 2011 kl. 14.00–19.00

Examinator: Per Enqvist, tel. 790 62 98

---

1. (a) Utför radoperationer på A tills vi erhåller matrisen

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Då ser vi att bildrummet till A kommer att spännas upp av kolumn 1 och 3, dvs

$$\mathcal{R}(A) = \text{span}(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix})$$

Utför radoperationer på A^T tills vi erhåller matrisen

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

Då ser vi att nollrummet till A^T spänns av

$$\mathcal{N}(A^T) = \text{span}(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix})$$

Eftersom

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = 0, \quad \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}^T \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = 0$$

så är underrummen ortogonala.

(b) Ett spännande träd är en delmängd av bågarna i grafen som innehåller ett minimalt antal bågar sådana att alla noder förbinds av en väg, dvs grafen är sammanhängande, och det inte bildas några cykler i delgrafen.

Ett exempel är om man tar bågarna $(1,4)$,$(2,4)$,$(3,4)$,$(4,5)$,$(4,7)$ och $(4,8)$, som svarar mot baslösningen $x_{14} = 10$, $x_{24} = 5$,$x_{34} = 5$,$x_{45} = 4$,$x_{47} = 8$, $x_{48} = 8$ och övriga flöden lika med noll. Eftersom alla flödena är icke-negativa och uppfyller flödesbalansen i varje nod så är denna baslösning tillåten.

2. (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.

3. (a) Utför radoperationer på matrisen

$$[A\ b] = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 \\ 0 & -1 & -1 & 0 \end{array}\right],$$

så erhåller vi

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{array}\right].$$

och sista ekvationen kan aldrig uppfyllas.

(b) Lös $A^TAx = A^Tb$, vi utför därför radoperationer på matrisen

$$[A^TA\ A^Tb] = \left[\begin{array}{rrr|r} 2 & 1 & -1 & 0 \\ 1 & 2 & 1 & 1 \\ -1 & 1 & 2 & 1 \end{array}\right],$$

så erhåller vi

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & -1/3 \\ 0 & 1 & 1 & 2/3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

En vektor som löser detta är

$$x^* = \begin{bmatrix} -1/3 \\ 2/3 \\ 0 \end{bmatrix}.$$

(c) Vi vill nu lösa ekvationssystemet $AA^Tu = Ax^*$ och utför därför radoperationer på matrisen

$$[AA^T\ Ax^*] = \left[\begin{array}{rrr|r} 2 & -1 & -1 & 1/3 \\ -1 & 2 & -1 & 1/3 \\ -1 & -1 & 2 & -2/3 \end{array}\right],$$

så erhåller vi

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & 1/3 \\ 0 & 1 & -1 & 1/3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

En vektor som löser detta är

$$\hat{u} = \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \end{bmatrix}.$$

Då är $\hat{x} = A^T\hat{u} = (0, 1/3, 1/3)^T$.

(d) Vi har att $A\hat{x} - b = (-2/3, -2/3, -2/3)^T$, och kolumnerna i A är ortogonala mot denna vektor, dvs

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -2/3 \\ -2/3 \\ -2/3 \end{bmatrix} = 0, \quad \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} -2/3 \\ -2/3 \\ -2/3 \end{bmatrix} = 0, \quad \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}^T \begin{bmatrix} -2/3 \\ -2/3 \\ -2/3 \end{bmatrix} = 0$$

4. (a) Gradienten till f är $\nabla f(x) = [2x_1 + 2x_2 + 2x_3, \quad 4x_2 + 2x_1 + 3x_3, \quad 6x_3 + 2x_1 + 3x_2]$, och Hessianen ges av

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 2 & 3 \\ 2 & 4 & 3 \\ 3 & 3 & 6 \end{bmatrix} = H,$$

Genom LDL^T-faktorisering erhålls en D-matris med diagonalelementen $2, 2, 3/2$ och därmed är Hessianen positivt definit och f är strikt konvex på hela rummet.

(b) I startpunkten är $\nabla f(x^{(0)}) = (2, 4, 3)^T$, och ekvationssystemet $\nabla^2 f(x^{(0)})d^{(0)} = -\nabla f(x^{(0)})$ har lösningen $d^{(0)} = (0, -1, 0)$ eftersom $\nabla f(x^{(0)})$ är lika som andra kolumnen i H.

Nu blir $x^{(1)} = x^{(0)} + d^{(0)} = (0, 0, 0)^T$, eftersom $f(x^{(1)}) = 0 \; < \; f(x^{(0)}) = 2$.

I nästa iteration är $\nabla f(x^{(0)}) = (0, 0, 0)^T$, så då blir lösningen till $\nabla^2 f(x^{(1)})d^{(1)} = -\nabla f(x^{(1)})$ vektorn $d^{(1)} = (0, 0, 0)^T$ och $x^{(2)} = x^{(1)} + d^{(1)} = (0, 0, 0)^T = x^{(1)}$ och algoritmen har konvergerat.

Eftersom vi har ett konvext kvadratiskt problem så konvergerar Newtons algoritm i ett steg. (Den kvadratiska approximationen som används i Newtons metod är då exakt.)

(c) Bivillkoren som har lagts till definierar ett konvext område, eftersom de kan skrivas som $g_i(x) \le 0$ med konvexa funktioner g_i (här linjära).

(d) Låt $g_1(x) = 1 - x_1$, $g_2(x) = 1 - x_2$, $g_3(x) = -x_3$ och $\hat{x} = (1, 1, 1)^T$.

Villkoret KKT1: $\nabla f + \sum_{i=1}^{3} y_i \nabla g_i = 0$ ger då att

$$\begin{bmatrix} 6 \\ 9 \\ 11 \end{bmatrix} + y_1 \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + y_2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = 0$$

och följdaktligen blir $y_1 = 6$, $y_2 = 9$ och $y_3 = 11$. Eftersom de är positiva så är även KKT3 uppfyllt.

Men $g_3(\hat{x}) = -1$ och $y_3 = 11$, så $y_3 g_3(\hat{x}) = 0$ är inte uppfyllt.

EFtersom det tillåtna området har en inre punkt, och problemet är konvext så är det reguljärt och då är KKT villkoren både nödvändiga och tillräckliga så punkten $\hat{x}$ kan varken vara en lokal eller global minpunkt.

5. (a) Problemet kan skrivas min $f(x)$ då $h(x) = 0$, där bivillkoret $h(x) = \{x : Ax = b\}$ definierar ett konvext område som även kan beskrivas som $\{x = x_0 + Zv, v \in I\!R\}$ där $x_0 = (0, 1, 0)$ och $Z = [1, -1, 1]^T$.

För att (P_1) ska vara ett konvext optimeringsproblem krävs det att den reducerade hessianen ska vara positivt semidefinit. Här:

$$Z^T H Z = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 2$$

Så problemet är konvext.

Vi försöker lösa följande ekvationssystem

$$\nabla f(x^*) + u \nabla h(x^*) \;\; = \;\; \begin{pmatrix} x_3 & 0 & x_1 \end{pmatrix} + u_1 \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} = 0$$

Detta är sant för $u_1 = u_2 = 0$ och $x_1 = x_3 = 0$ vilket innebär att $x_2 = 1$ måste gälla för att få en tillåten punkt. Eftersom problemet är konvext och reguljärt så är detta det globala optimat.

(b) Bivillkoren är samma som i (a) så de bildar även här ett konvext område.

För att (P_2) ska vara ett konvext optimeringsproblem krävs det att den reducerade hessianen ska vara positivt semidefinit. Här:

$$Z^T H Z = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & 2x_2x_3 & x_2^2 \\ 2x_2x_3 & 2x_1x_3 & 2x_1x_2 \\ x_2^2 & 2x_1x_2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 2(x_2^2 - x_1x_3)$$

Om vi väljer, t.ex. $x = (1, 0, 1)$, så blir detta uttryck negativt och den reducerade hessianen är inte positivt semidefinit. Så problemet är inte konvext.

Vi försöker lösa följande ekvationssystem

$$\nabla g(x^*) + u\nabla h(x^*) \quad = \quad \begin{pmatrix} x_2^2x_3 & 2x_1x_2x_3 & x_1x_2^2 \end{pmatrix} + u_1 \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} = 0$$

Detta är sant för $u_1 = u_2 = -1/8$ i punkten $x^{(*)}$. Eftersom problemet inte är konvext så kan vi inte säga att det är en global minpunkt. Man kan t.ex. se att $g(x^{(*)}) = 1/2^8 > g(0, 1, 0) = 0$.