KTH Mathematics

Lösningsförslag till tentamen i SF1861 Optimeringslära för T.
Torsdag 28 maj 2010 kl. 14.00–19.00

Examinator: Per Enqvist, tel. 790 62 98

---

1. (a) Inför variablerna

$$x = (x_{sR}, x_{sM}, x_{sP}, x_{sA}, x_{sD}, x_{gR}, x_{gM}, x_{gP}, x_{gA}, x_{gD}, x_{mR}, x_{mM}, x_{mP}, x_{mA}, x_{mD},)^T$$

för antalet bussar som ska transporteras från städerna Stockholm (s), Göteborg (g) och Malmö (m), till städerna Rom (R), Madrid,(M) Paris (P), Aten (A) och till en Dummy stad (D). Vi inför dummy staden för att få ett balanserat problem.

Enligt antagandet känner vi till kostnaderna för transporterna och vi ordnar dessa i en vektor c i samma ordning som i x. Kostnaderna för transporter till dummynoden sätter vi till noll.

Låt s_i beteckna hur många bussar som finns i stad i, och d_j beteckna hur många busslaster som finns i stad j. Antalet busslaster i Dummynoden sätts till 10 så att problemet blir balanserat.

Optimeringsproblemet kan nu skrivas

$$(\text{TP}) \quad \left[\begin{array}{lll} \min\limits_x & c^T x & \\ \text{då} & \sum_i x_{ij} = d_j, & j = R, M, P, A, D \\ & \sum_j x_{ij} = s_i, & i = s, g, m. \\ & x \geq 0 & \end{array} \right].$$

En tillåten startbaslösning ges av Northwest- corner rule

x_{ij}	R	M	P	A	D	s_i
s	10	5	10	0	0	25
g	0	0	5	15	0	20
m	0	0	0	5	10	15
d_j	10	5	15	20	10	

(b) Utför radoperation på A

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \rightarrow U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

där de två första kolumnerna är identiteten, dvs $\beta_1 = 1$, $\beta_2 = 2$ och $\gamma_1 = 3$.

Nollrummet till A spänns av vektorn $n = (0,\ -1,\ 1)^T$ och bildrummet till A^T spänns av vektorerna i U^T, dvs $r_1 = (1,\ 0,\ 0)^T$ och $r_2 = (0,\ 1,\ 1)^T$

Nu gäller att

$$z = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = v + u = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \alpha_1 + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \eta_1 + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \eta_2$$

Med $\alpha_1 = -1$ och $\eta_1 = \eta_2 = 0$ har vi $z = u + v$ där

$$v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \in \mathcal{N}(A), \quad u = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in \mathcal{R}(A^T),$$

2. (a) Problemet (P) är på standardform

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

där

$$A = \begin{bmatrix} 1 & 1 & 0 & 2 & 4 \\ 0 & 2 & -1 & 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad c = \begin{bmatrix} 1 & 3 & -1 & 2 & 4 \end{bmatrix}^T.$$

Vi börjar med variablerna x_1 and x_2 i basen, d.v.s. $\beta = \{1, 2\}$ och $\delta = \{3, 4, 5\}$.

$$A_\beta = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, A_\delta = \begin{bmatrix} 0 & 2 & 4 \\ -1 & 1 & 1 \end{bmatrix}$$

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

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

Man kan välja mellan att låta 4 och 5 bli nya basvariabelindex. Låt x_4 gå in i basen. Vilken variabel ska lämna basen ? $\bar{b} = (1/2,\ 1/2)$ bestäms från $A_\beta \bar{b} = b$. Från $A_\beta \bar{a}_4 = a_4$, får vi $\bar{a}_4 = (3/2, 1/2)^T$. Från $x_\beta = \bar{b} - \bar{a}_4 t$ ser vi att det första elementet i x_β blir noll först (eftersom den kvoten är minst), och då måste den första basvariabeln, nämligen x_1, lämna basen.
Nu är $\beta = \{4, 2\}$ och $\delta = \{3, 1, 5\}$, och vi uppdaterar bas- och ickebasmatriser:

$$A_\beta = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, A_\delta = \begin{bmatrix} 0 & 1 & 4 \\ -1 & 0 & 1 \end{bmatrix}$$

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

$$y = \begin{bmatrix} 1/3 \\ 4/3 \end{bmatrix}, \quad r_\delta^T = \begin{bmatrix} 1/3 & 2/3 & 4/3 \end{bmatrix}.$$

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

(b) Det duala problemet blir

$$
(D)\quad \left[\begin{array}{ll}
\max\limits_{y} & y_1 + y_2 \\
\text{då} & y_1 \leq 1 \\
 & y_1 + 2y_2 \leq 3 \\
 & -y_2 \leq -1 \\
 & 2y_1 + y_2 \leq 2 \\
 & 4y_1 + y_2 \leq 4
\end{array}\right]
$$

Komplementaritet ger att bivillkor 2 och 4 är uppfyllda med likhet och vi ser att $y_1 = 1/3$ och $y_2 = 4/3$ från simplexmetoden uppfyller dessa villkor. Dessutom uppfylls övriga bivillkor med strikt olikhet och variablerna är icke-negativa. Som extra koll kan man se att $c^T x = b^T y$, dvs $1/3 * 3 + 2 * 1/3 = 1/3 + 4/3$.

(c) Om vi ändrar koefficienten i högerledet så kommer samma baslösning vara optimal så länge som den är tillåten. (reducerade kostnaderna ändras ju inte av denna förändring) Vi har att

$$
x_\beta = A_\beta^{-1} b^{nya} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix},
$$

vilken är icke-negativ. I så fall gäller att

$$
c_\beta^T x_\beta = \begin{bmatrix} 2 \\ 3 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} = 3/2
$$

3. (a) Givet

$$
H = \begin{bmatrix} 2 & -6 & 8 \\ -6 & 20 & -20 \\ 8 & -20 & 41 \end{bmatrix}.
$$

Diagonalmatrisen i LDL^T-faktoriseringen av matrisen H bestäms av radoperationer med faktorer 3 och -4, och motsvarande kolumnoperationer:

$$
\begin{bmatrix} 2 & -6 & 8 \\ 0 & 2 & 4 \\ 0 & 4 & 9 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 9 \end{bmatrix}
$$

Nu använder vi rad 2 för en radoperation med faktorn -2, och motsvarande kolumnoperation:

$$
\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix} \quad \Rightarrow \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

Eftersom diagonalelementen är positiva så är H positivt definit. L-matrisen bestäms av radoperationerna ovan

$$
L = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}
$$

(b) Det tillåtna området är konvext eftersom det bestäms av ett linjärt ekvationssystem. Målfunktionen är konvex på hela $I\!R^3$ enligt a-uppgiften, och är därför även konvex på det tillåtna området. Alltså är optimeringsproblemet konvext. Bestäm nollrummet till A

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

Det spänns av kolumnvektorn

$$Z = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

En lösning till ekvationen $Ax = b$ ges av $\bar{x} = [1\ 0\ 0]^T$. Med

$$\bar{c} = Z^T c = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix} = -7$$

ges en lösning v till ekvationssystemet $Z^T H Z v = -Z^T(H\bar{x} + c)$, d.v.s.

$$34v = -1$$

av $v = -1/34$.

Eftersom problemet var strikt konvext så är $\hat{x} = \bar{x} + vZ = (33/34, 1/34, 0)^T$ det unika minimat till (P).

(c) Vi kollar riktningsderivatan i punkten x i riktningen d

$$\nabla f(x)d = (Hx + c)d = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \\ 8 \end{bmatrix} = 1 > 0.$$

Eftersom denna är positiv så är d ej en avtaganderiktning.

4. (a) f är summan av två konvexa funktioner och alltså konvex. Den första termen är konvex eftersom det är en konvex och växande funktion (exponentialfunktionen) av en konvex funktion.

(b) Gradienten till f är $\nabla f(x) = [2x_1\ \exp\{x_2 + x_3^2\}\ 2x_3 \exp\{x_2 + x_3^2\}]$, och hessianen ges av

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \exp\{x_2 + x_3^2\} & 2x_3 \exp\{x_2 + x_3^2\} \\ 0 & 2x_3 \exp\{x_2 + x_3^2\} & 2(1 + 2x_3^2)\exp\{x_2 + x_3^2\} \end{bmatrix}.$$

I punkten $x^{(0)} = 0$ har vi då $\nabla^2 f(x^{(0)}) d^{(0)} = -\nabla f(x^{(0)})$, dvs

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} d^{(0)} = -\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \quad \Rightarrow d^{(0)} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}.$$

Med ett enhetssteg blir nu nästa iterationspunkt $x^{(1)} = x^{(0)}+d^{(0)} = (0,\ -1,\ 0)^T$. Om Newtons algorithm konvergerar så kommer den gå till ett lokalt min och vi vet att eftersom f är oändligt differentierbar så måste dess gradient vara lika med noll där, men gradienten till noll är inte noll någonstans så algoritmen kommer inte att konvergera. (Eftersom det inte finns några lokala minpunkter)

(c) Låt

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

Då kan problemet skrivas min $f(x)$ då $g(x) \leq 0$.
I punkten $x^* = (0, 0, 0)$, är alla bivillkoren aktiva.

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

Detta är sant för $\lambda_1 = 0$, $\lambda_2 = 1 > 0$, $\lambda_3 = 0$. Alltså är KKT1 och KKT3 uppfyllda. Eftersom x^* är tillåten så är även KKT2 uppfyllt och i och med att alla bivillkor är aktiva 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.

(d) Låt

$$g(x) = 1 - (x_1 - 1)^2 - x_2^2$$

Då kan problemet skrivas min $f(x)$ då $g(x) \leq 0$.
I punkten x^*, är bivillkoret aktivt, dvs $g(x^*) = 0$.

$$\nabla f(x^*) + \lambda \nabla g(x^*) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = 0.$$

Detta är aldrig sant, så KKT1 är ej uppfyllt.
Eftersom punkten x^* är reguljär kan detta inte vara ett lokalt minimum.

5. (a) Detta är ett linjärt minstakvadrat problem.
Problemet kan skrivas som $\min_x f(x)$, där $f(x) = \|\mathcal{A}x - b\|^2$, $x = [A\ B\ C]^T$ och

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & \pi/4 & -1/\sqrt{2} \\ 1 & \pi/3 & -1/2 \\ 1\pi/2 & 0 & \end{pmatrix}, \qquad = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 4 \end{pmatrix}.$$

Eftersom kolumnerna i $\mathcal{A}$ är oberoende finns det ett unikt optimalt x och det bestäms av $\mathcal{A}^T\mathcal{A}x = \mathcal{A}^T b$.

(b) Området är inte konvext. Punkterna $x^{(1)} = (1, 0, 0)$ och $x^{(2)} = (-1, 0, 0)$ är tillåtna, men punkten $1/2x^{(1)} + 1/2x^{(2)} = 0$ är inte tillåten.
Tillåtna området bestäms av att $h_1(x) = x_1^2 + x_2^2 - 1 = 0$ och $h_2(x) = x_1^2 + x_3^2 - 1 = 0$. Gradienterna till dessa ges av $\nabla h_1(x) = (2x_1,\ 2x_2\ 0)$ och $\nabla h_2(x) = (2x_1,\ 0\ 2x_3)$. Dessa är linjärt beroende om $x_2 = x_3 = 0$ Då följer av bivillkoren att $x_1^2 = 1$. De punkter som inte är reguljära ges alltså av $x^{(1)}$ och $x^{(2)}$ ovan.

(c) Problemet har en ändlig lösning om och endast om dualen har en tillåten lösning.

$$(D) \quad \left[\begin{array}{ll} \max\limits_{y} & 2y_1 + y_2 \\ \text{då} & -y_1 + y_2 \leq c_1 \\ & y_1 + y_2 \leq c_2 \\ & y_1 \geq 0, y_2 \geq 0 \end{array}\right]$$

Den andra ekvationen har icke-negativa lösningar om och endast om $c_2 \geq 0$. Dessa ligger i så fall i en triangel i positiva kvadranten. Det minsta värdet för c_1 då det första bivillkoret har en tillåten punkt tillsammans med det andra är då $c_2 = -c_1$. (rita en figur)

Det primala problemet är då ändligt om $c_2 \geq 0$ och $c_1 \geq -c_2$.