KTH Mathematics

Lösningsförslag till tentamen i SF1861 Optimeringslära för T.
Tisdag 19 maj 2009 kl. 08.00–13.00

Examinator: Per Enqvist, tel. 790 62 98

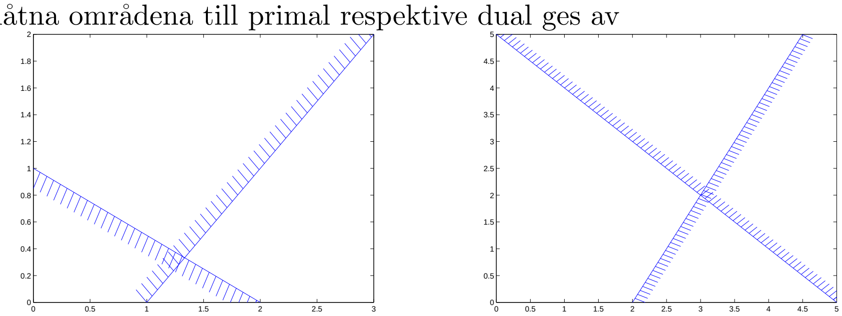
---

1. (a) Inför artificiella variabler x_4 och x_5 och definiera fas 1 problemet

$$\text{(Fas 1)} \quad \left[\begin{array}{ll} \min\limits_{x} & x_4 + x_5 \\ \text{då} & x_1 - x_2 + 2x_3 + x_4 = 1 \\ & 2x_1 + 3x_2 + x_3 + x_5 = 7 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0, \end{array} \right].$$

En tillåten startbaslösning ges av $x_4 = 1$ och $x_5 = 7$, och $x_1 = x_2 = x_3 = 0$. Om fas 1 problemet har ett optimalvärde som är strikt större än noll så har (P) ingen tillåten baslösning, och är optimalvärdet lika med noll så ges en tillåten baslösning av de optimala x_1, x_2 och x_3.

(b) De tillåtna områdena till primal respektive dual ges av



Vi ser att optimum för primalen ges av skärningspunkten till de två linjerna $x_1 + 2x_2 = 2$ och $x_1 - x_2 = 1$, d.v.s. $x_1 = 4/3$ och $x_2 = 1/3$.

Vi ser att optimum för dualen ges av skärningspunkten till de två linjerna $y_1 + y_2 = 5$ och $2y_1 - y_2 = 4$, d.v.s. $x_1 = 3$ och $x_2 = 2$.

De två punkterna ovan är tillåtna till primalen respektive dualen och dessutom gäller att

$$5x_1 + 4x_2 = 5\frac{4}{3} + 4\frac{1}{3} = 8 = 2 \cdot 3 + 2 = 2y_1 + y_2$$

och enligt svag dualitet så måste dessa punkter vara optimala till respektive problem.

2. (a) Inför slackvariabler för att föra över (P) på standardform

$$(P_s) \quad \left[\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \right]$$

där

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

Vi börjar med slackvariablerna x_3 and x_4 i basen, d.v.s. $\beta = \{3,4\}$ och $\delta = \{1,2\}$.

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

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

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

Låt x_2 gå in i basen. Vilken variabel ska lämna basen ?

Från $A_\beta \bar{a}_2 = a_2$, får vi $\bar{a}_2 = (3, -1)^T$, och eftersom det andra elementet är negativt, så måste den första basvariabeln, nämligen x_1, lämna basen.

Nu är $\beta = \{2,4\}$ och $\delta = \{1,3\}$, och vi uppdaterar bas- och ickebasmatriser:

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

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

$$y = \begin{bmatrix} -2/3 \\ 0 \end{bmatrix}, \quad r_\delta^T = \begin{bmatrix} 5/3 & 2/3 \end{bmatrix}.$$

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

(b) 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

$$x_\beta = A_\beta^{-1} b = \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1+\delta \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1+\delta \\ 4+\delta \end{bmatrix},$$

vilken är icke-negativ om $\delta \geq -1$. I så fall gäller att

$$c_\beta^T x_\beta = \begin{bmatrix} -2 \\ 0 \end{bmatrix}^T \frac{1}{3} \begin{bmatrix} 1+\delta \\ 4+\delta \end{bmatrix} = -2/3(1+\delta).$$

Koefficienten $-2/3$ svarar också mot simplexmultiplikatorn y_1 och den optimala dual variabeln.

(c) Om vi ändrar koefficienten i målfunktionen så kommer samma baslösning vara optimal så länge som motsvarande reducerade kostnad fortfarande är icke-negativ, d.v.s. så länge som $r_1 = c_1 - y^T a_1 = (1+\epsilon) - [-2/3\ 0][1\ 1]^T = \epsilon + 5/3 \geq 0$. Om $\epsilon \geq -5/3$ så kommer målfunktionens värde alltså ändras med $x_1 \epsilon$, d.v.s. inte alls eftersom x_1 inte är en basvariabel och därför noll.

3. (a) Först bildrummet till A^T:

$$A^T = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Trappstegskolumnerna i denna matris är nr. 1 och 2 och därför spänns bildrummet till A^T av kolumn 1 och 2 i A^T.

Nollrummet till A

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

spänns av kolumnerna i

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

Att dessa underrum är ortogonala följer av att $x \in \mathcal{R}(A^T)$ kan skrivas $x = A_{12}^T v$ och $y \in \mathcal{N}(A)$ kan skrivas $y = Zw$ och

$$x^T y = v^T A_{12} Z w = v^T \begin{bmatrix} 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} w = v^T \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} w = 0.$$

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

$$Z^T H Z = \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T H = \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 13 & -2 \\ -2 & 1 \end{bmatrix}.$$

Diagonalmatrisen i LDL^T-faktoriseringen av denna reducerade hessian bestäms av

$$\begin{bmatrix} 13 & -2 \\ -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 13 & -2 \\ 0 & 1 + 4/13 \end{bmatrix} \sim \begin{bmatrix} 13 & 0 \\ 0 & 17/13 \end{bmatrix},$$

och eftersom diagonalelementen är positiva så är den reducerade hessianen positivt definit.

(c) En lösning till ekvationen $Ax = b$ ges av $\bar{x} = 0$. Med

$$\bar{c} = Z^T c = \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 7 \\ 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -13 \\ 2 \end{bmatrix}$$

ges en lösning v till ekvationssystemet $Z^T HZv = -Z^T c$, d.v.s.

$$\begin{bmatrix} 13 & -2 \\ -2 & 1 \end{bmatrix} v = - \begin{bmatrix} -13 \\ 2 \end{bmatrix}$$

av $v = [1\ 0]^T$.
Eftersom den reducerade Hessianen $Z^T HZ$ är positivt definit, så är $\hat{x} = \bar{x} + vZ = (-2, 1, 1, 0)^T$ det unika minimat till (P).

4. (a) Gradienten till f är $\nabla f(x) = [\frac{x_1}{\sqrt{1+x_1^2}} \quad 2x_2]$, och hessianen ges av

$$\nabla^2 f(x) = \begin{bmatrix} \frac{1}{\sqrt{1+x_1^2}^3} & 0 \\ 0 & 2 \end{bmatrix}.$$

Eftersom hessianen är positivt definit för alla $x \in I\!R^2$ så är f konvex.

(b) Vi startar i punkten $x^{(0)} = (2, 1)$. Då gäller

$$\nabla f(x^{(0)}) = [\frac{2}{\sqrt{5}} \quad 2], \quad \text{och} \quad \nabla^2 f(x^{(0)}) = \begin{bmatrix} \frac{1}{\sqrt{5}^3} & 0 \\ 0 & 2 \end{bmatrix},$$

och ekvationssystemet $\nabla^2 f(x^{(0)}) d^{(0)} = -\nabla f(x^{(0)})$ ger $d^{(0)} = [-10 \quad -1]^T$.
Vi har att $f(x^{(0)}) = \sqrt{5} + 1$.
Steglängdsval:
Full steglängd $f(x^{(0)} + d^{(0)}) = f(-8, 0) = \sqrt{1 + 64} > \sqrt{5} + 1$.
Halv steglängd $f(x^{(0)} + 1/2d^{(0)}) = f(-3, 1/2) = \sqrt{10} + 1/4 > \sqrt{5} + 1$.
Kvarts steglängd $f(x^{(0)} + 1/4d^{(0)}) = f(-0.5, 0.75) = \sqrt{(5/4)} + 9/16 < \sqrt{5} + 1$.

Nästa punkt blir $x^{(1)} = x^{(0)} + 1/4d^{(0)} = (-0.5, 0.75)$.

(c) Låt

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

Då kan problemet skrivas min $f(x)$ då $g(x) \leq 0$.
I punkten $x^{(1)} = (1, 0)$, är bivillkoret aktivt.

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

Detta är sant för $\lambda = 1/2^{3/2} > 0$.
Problemet är inte konvext, (Tillåtna området innehåller ett hål), så KKT-villkoren är bara nödvändiga och vi kan inte säga att x^* är ett lokalt eller globalt minimum. (Första ordningens) KKT-villkor kan bara användas till att utesluta punkter för dessa icke-konvexa problem.

(d) Låt

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

Då kan problemet skrivas min $f(x)$ då $g(x) \leq 0$.
I punkten $x^{(1)} = (1, 1)$, är bivillkoret aktivt.

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

Detta är sant för $\lambda = 1/2^{3/2} > 0$.
Problemet är konvext, (Tillåtna området är en disk), så KKT-villkoren är ekvivalenta med de globala optimalitetsvillkoren och vi kan alltså säga att x^* är ett globalt minimum.

5. (a) Cirkeln är den som ska bestämmas och våra variabler ska beskriva denna. Låt (α, β) beteckna centrum för cirkeln och låt r beteckna dess radie. Problemet kan skrivas som $\min_x f(x)$, där $f(x) = \sum_{k=1}^{m} h_k(x)^2 = h(x)^T h(x)$, $x = [\alpha,\ \beta,\ r]^T$ och

$$h_k(x) = (x(1) - p_k)^2 + (x(2) - q_k)^2 - r^2, \quad k = 1, \cdots, m.$$

Alternativt kan man även välja

$$\tilde{h}_k(x) = \sqrt{(x(1) - p_k)^2 + (x(2) - q_k)^2} - |r|, \quad k = 1, \cdots, m.$$

(b) Första ordningens nödvändiga villkor säger att $\nabla f(x) = 0$. Här är $\nabla f = h^T \nabla h$, där

$$h(x) = \begin{bmatrix} (x(1) - p_1)^2 + (x(2) - q_1)^2 - r^2, \\ (x(1) - p_2)^2 + (x(2) - q_2)^2 - r^2, \\ (x(1) - p_3)^2 + (x(2) - q_3)^2 - r^2, \\ (x(1) - p_4)^2 + (x(2) - q_4)^2 - r^2, \end{bmatrix},$$

och

$$\nabla h(x) = \begin{bmatrix} 2(x(1) - p_1) & 2(x(2) - q_1) & -2r, \\ 2(x(1) - p_2) & 2(x(2) - q_2) & -2r, \\ 2(x(1) - p_3) & 2(x(2) - q_3) & -2r, \\ 2(x(1) - p_4) & 2(x(2) - q_4) & -2r, \end{bmatrix}.$$

För givna data får vi att

$$h(x^{(0)}) = \begin{bmatrix} (-0.6)^2 + (-0.8)^2 - 1^2, \\ (-0.8)^2 + (-0.6)^2 - 1^2, \\ (1.1)^2 + (0)^2 - 1^2, \\ (0)^2 + (0.9)^2 - 1^2, \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0.21 \\ -0.19 \end{bmatrix}$$

och

$$\nabla h(x^{(0)}) = \begin{bmatrix} 2(-0.6) & 2(-0.8) & -2, \\ 2(-0.8) & 2(-0.6) & -2, \\ 2(1.1) & 2(0) & -2, \\ 2(0) & 2(0.9) & -2, \end{bmatrix} = \begin{bmatrix} -1.2 & -1.6 & -2 \\ -1.6 & -1.2 & -2 \\ 2.2 & 0 & -2 \\ 0 & 1.8 & -2 \end{bmatrix}.$$

vilket ger

$$\nabla f(x^{(0)}) = \begin{bmatrix} 0.462 & -0.342 & -0.04 \end{bmatrix}.$$

där $x^{(0)} = [0,\ 0,\ 1]^T$.

(c) Sökriktningen d bestäms av ekvationssystemet

$$\nabla h(x^{(0)})^T \nabla h(x^{(0)}) d = -\nabla h(x^{(0)})^T h(x^{(0)}).$$

Nu gäller

$$\nabla h(x^{(0)})^T \nabla h(x^{(0)}) = \begin{bmatrix} 8.84 & 3.84 & 1.2 \\ 3.84 & 7.24 & 2 \\ 1.2 & 2 & 16 \end{bmatrix},$$

så systemet som bestämmer d ges av

$$\begin{bmatrix} 8.84 & 3.84 & 1.2 \\ 3.84 & 7.24 & 2 \\ 1.2 & 2 & 16 \end{bmatrix} d = -\begin{bmatrix} 0.462 \\ -0.342 \\ -0.04 \end{bmatrix}.$$

När vi har bestämt d så gör vi en approximativ linjesökning i d's riktning. T.ex. kan vi bestämma t_k som det första i en avtagande följd av tal som går mot noll sådant att $f(x^{(0)} + t_k d) < f(x^{(0)})$ och sedan låta den nya iterationspunkten ges av $x^{(1)} = x^{(0)} + t_k d$.