KTH Matematik

Tentamen i 5B1817 Tillämpad ickelinjär optimering
Fredagen den 8 juni 2007 kl. 8.00–13.00
Kortfattade lösningsförslag

---

1. (a) Vi kan partitionera A enligt $A = (B\ N)$, där B består av den första kolumnen i A. En matris Z vars kolumner formar en bas för nullrummet till A kan då fås enligt

$$Z = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Målfunktionen kan nu uttryckas som

$$\tfrac{1}{2}(x_0 + Zv)^T H (x_0 + Zv).$$

Minimerande v, som då är en global minpunkt, ges alltså av $Z^T H(x_0 + Zv) = 0$ under förutsättning att sådant v finns samt att dessutom $Z^T H Z$ är positivt semidefinit. Insatta värden ger

$$Z^T H Z = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},$$

vilket är en positivt definit matrix. Alltså finns unik global minpunkt. Minimerande v ges av

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}.$$

Lösningen är $v = (-1\ 3)^T$. Insättning ger

$$x = x_0 + Zv = \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}.$$

(b) Analogt får vi här

$$Z = \begin{pmatrix} 0 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{och} \quad Z^T H Z = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Nu är $Z^T H Z$ inte positivt semidefinit, varför global minpunkt saknas.

2. (Se kursmaterialet.)

3. Vi har

$$f(x) = (x_1+2)^2 + (x_2+3)^2, \quad \nabla f(x) = \begin{pmatrix} 2(x_1+2) \\ 2(x_2+3) \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

$$g_1(x) = 1 - \frac{1}{2}x_1^2 - \frac{1}{2}x_2^2, \quad \nabla g_1(x) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}, \quad \nabla^2 g_1(x) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$g_2(x) = x_2, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla^2 g_2(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Insättning ger första QP-problemet enligt

$$\begin{array}{ll} \text{min} & \frac{3}{2}p_1^2 + \frac{3}{2}p_2^2 + 6p_1 + 8p_2 \\ \text{då} & -p_1 - p_2 \geq 0, \\ & p_2 \geq -1. \end{array}$$

Lösning, exempelvis grafiskt, ges av $p = (-2\ -1)^T$, $\lambda = (0\ 5)^T$. (Man kan också betrakta problemet och inse att bivillkoret $-p_1 - p_2 \geq 0$ inte kommer att vara aktivt. Om det bivillkoret tas bort får man två separata envariabelproblem.) I uppgiften krävs ingen linjesökning, så vi väljer steglängden ett. Därmed får vi $x^{(1)} = (-1\ 0)^T$ och $\lambda^{(1)} = (0\ 5)^T$.

4. (a) Derivatorna finns uträknade i förra uppgiften. Vi får

$$\nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)}) = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}, \quad \nabla_x \mathcal{L}(x^{(0)}, \lambda^{(0)}) = \begin{pmatrix} 4 \\ 9 \end{pmatrix},$$

$$g(x^{(0)}) = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}, \quad A(x^{(0)}) = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}$$

Då $g(x^{(0)}) > 0$ behöver vi inte inför slackvariabler. (Alternativt kan vi införa slackvariabler och exempelvis låta $s^{(0)} = g(x^{(0)}) > 0$.) Newtonsteget $\Delta x^{(0)}$, $\Delta\lambda^{(0)}$ ges av

$$\begin{pmatrix} \nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)}) & A(x^{(0)})^T \\ \Lambda^{(0)} A(x^{(0)}) & -G(x^{(0)}) \end{pmatrix} \begin{pmatrix} \Delta x^{(0)} \\ -\Delta\lambda^{(0)} \end{pmatrix} = -\begin{pmatrix} \nabla_x \mathcal{L}(x^{(0)}, \lambda^{(0)}) \\ G(x^{(0)})\lambda^{(0)} - \mu e \end{pmatrix},$$

där $\Lambda^{(0)} = \text{diag}(\lambda^{(0)})$.

Insättning av numeriska värden ger

$$\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & -1 & 1 \\ 0 & -2 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \Delta x_1^{(0)} \\ \Delta x_2^{(0)} \\ -\Delta\lambda_1^{(0)} \\ -\Delta\lambda_2^{(0)} \end{pmatrix} = \begin{pmatrix} -4 \\ -9 \\ 0 \\ 0 \end{pmatrix}.$$

(b) Nya iterationspunkter ges av $x^{(1)} = x^{(0)} + \alpha^{(0)}\Delta x^{(0)}$ och $\lambda^{(1)} = \lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)}$, där steglängden $\alpha^{(0)}$ väljs till 1 i en ren Newtoniteration. Vi måste dock begränsa steglängden så att $g(x^{(0)} + \alpha^{(0)}\Delta x^{(0)}) > 0$ och $\lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)} > 0$

bibehålls. Alternativt, om slackvariabler införs med ekvationen $g(x) - s = 0$, bibehåller vi $s^{(0)} + \alpha^{(0)}\Delta s^{(0)} > 0$ och $\lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)} > 0$. För att garantera konvergens, bör vi dessutom utföra linjesökning med lämplig meritfunktion.

5. (a) Matrisen

$$\begin{pmatrix} Ix & -A^T \\ -A & Ix \end{pmatrix}$$

har egenvärden $x \pm \sigma_i$, $i = 1, \dots, n$. Därmed ges optimalt x av största singulärvärdet hos A, dvs σ_1.

(b) Det duala problemet blir

$$(DSDP) \qquad \begin{array}{ll} \max & \operatorname{trace}(AY_{12}) + \operatorname{trace}(A^T Y_{21}) \\ \text{då} & \operatorname{trace}(Y_{11}) + \operatorname{trace}(Y_{22}) = 1, \\ & \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} = \begin{pmatrix} Y_{11}^T & Y_{21}^T \\ Y_{12}^T & Y_{22}^T \end{pmatrix} \succeq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{array}$$

Egenskaper hos spåroperatorn ger

$$\operatorname{trace}(AY_{12}) = \operatorname{trace}(AY_{21}^T) = \operatorname{trace}(Y_{21}A^T) = \operatorname{trace}(A^T Y_{21}),$$

varför vi kan skriva problemet som

$$(DSDP) \qquad \begin{array}{ll} \max & 2\operatorname{trace}(AY_{12}) \\ \text{då} & \operatorname{trace}(Y_{11}) + \operatorname{trace}(Y_{22}) = 1, \\ & \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} = \begin{pmatrix} Y_{11}^T & Y_{21}^T \\ Y_{12}^T & Y_{22}^T \end{pmatrix} \succeq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{array}$$

(c) Då vi vet att optimalvärdet är uppåt begränsat av σ_1 räcker det att hitta tillåtet Y så att $\operatorname{trace}(AY_{12}) = \sigma_1/2$. Om vi sätter in $A = U\Sigma V^T$ betyder det att $\operatorname{trace}(U\Sigma V^T Y_{12}) = \sigma_1/2$. Om vi nu enligt uppgiftslydelsen låter $Y_{12} = (1/2)v_1 u_1^T$ får vi

$$\operatorname{trace}(U\Sigma V^T \frac{1}{2} v_1 u_1^T) = \frac{1}{2}\operatorname{trace}(\Sigma V^T v_1 u_1^T U) = \frac{1}{2}\operatorname{trace}(\sigma_1 e_1 e_1^T) = \frac{\sigma_1}{2},$$

eftersom U och V är ortonormala matriser. Därmed kan vi låta $Y_{11} = (1/2)v_1 v_1^T$ och $Y_{22} = (1/2)u_1 u_1^T$, vilket ger tillåten lösning då $\operatorname{trace}(Y_{11}) + \operatorname{trace}(Y_{22}) = 1/2 + 1/2 = 1$. Alltså har vi en optimallösning, som ges av

$$\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} v_1 \\ u_1 \end{pmatrix}\begin{pmatrix} v_1^T & u_1^T \end{pmatrix}.$$