Kungl Tekniska Högskolan

Institutionen för Matematik
Avdelningen för Optimeringslära och systemteori

Tentamen i 5B1816 Tillämpad matematisk programmering—ickelinjära problem.
Torsdagen den 7 mars 2002 kl. 14.00–19.00.

Kortfattade lösningsförslag.

---

1. (a) Ja.
(b) Nej.
(c) Ja.
(d) Ja.
(e) Nej.

2. (a) Bivillkor 1 och 2 är bindande i $\widehat{x}$. Dessutom är $\widehat{x}$ en tillåten reguljär punkt. Derivering ger

$$\nabla f(x) = \begin{pmatrix} -4x_1+2 \\ -x_2-2 \\ -8 \end{pmatrix}, \quad A_A(x) = \begin{pmatrix} 2x_1 & 2x_2 & 2x_3 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \begin{pmatrix} -4-2\lambda_1 & 0 & 0 \\ 0 & -1-2\lambda_1 & 0 \\ 0 & 0 & -2\lambda_1 \end{pmatrix}.$$

Första ordningens optimalitetsvillkor kräver då att det finns

$$\hat{\lambda}_A = \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix},$$

där $\hat{\lambda}_2 \geq 0$, så att $\nabla f(\widehat{x}) = A_A(\widehat{x})^T\hat{\lambda}_A$, dvs

$$\begin{pmatrix} 2 \\ -4 \\ -8 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 4 & 0 \\ 8 & 0 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix}.$$

Lösning finns då $\hat{\lambda}_1 = -1$ och $\hat{\lambda}_2 = 2$. Därmed upfyller $\widehat{x}$ tillsammans med $\hat{\lambda}$ första ordningens nödvändiga optimalitetsvillkor.

För att studera reducerade Hessianen till $\mathcal{L}(x,\lambda)$, låt $Z_A(\widehat{x}) = (0\ 2\ -1)^T$, vilket ger $Z_A(\widehat{x})^T\nabla^2_{xx}\mathcal{L}(\widehat{x},\hat{\lambda})Z_A(\widehat{x}) = 6$. Eftersom vi därmed också har positivt definit reducerad Hessian samt strikt komplementaritet, uppfyller $\widehat{x}$ tillsammans med $\hat{\lambda}$ andra ordningens tillräckliga optimalitetsvillkor. Därmed är $\widehat{x}$ en lokal optimalpunkt till (NLP).

(b) Första ordningens optimalitetsvillkor är oförändrade. Däremot får vi

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \nabla^2 f(x) = \begin{pmatrix} -4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

vilket ger $Z_A(\widehat{x})^T\nabla^2_{xx}\mathcal{L}(\widehat{x},\hat{\lambda})Z_A(\widehat{x}) = -4 < 0$. Därmed är andra ordningens nödvändiga optimalitetsvillkor till (NLP') inte uppfyllda i $\widehat{x}$, och $\widehat{x}$ är alltså inte en lokal minpunkt till (NLP').

Anmärkning: Bivillkorens krökning kan alltså vara avgörande för lokal optimalitet.

3. (Se kursmaterialet.)

4. Ett sätt att ta fram dualen är att notera att vi för varje fixerat $v \geq 0$ får en överskattning till optimalvärdet till (SDP) ur

$$(SDP(v)) \qquad \begin{array}{ll} \max & b^Ty + v^Ty \\ \text{då} & \sum_{i=1}^m A_i y_i + S = C, \\ & S = S^T \succeq 0. \end{array}$$

(Detta är lagrangerelaxering av bivillkoret $y \geq 0$.) Men $(SDP(v))$ har samma form som $(DSDP)$, vilket medför att vi för varje fixerat $v \geq 0$ får en överskattning av optimalvärdet ur problemet

$$(PSDP(v)) \qquad \begin{array}{ll} \min & \operatorname{trace}(CX) \\ \text{då} & \operatorname{trace}(A_iX) = b_i + v_i, \quad i = 1,\dots,m, \\ & X = X^T \succeq 0. \end{array}$$

Problemet $(PSDP(v))$ ger en överskattning av optimalvärdet till (SDP) för varje $v \geq 0$. Speciellt kan vi välja v så att överskattningen blir så stark som möjligt. Då kan v kan elimineras ur $(PSDP(v))$, vilket ger

$$\begin{array}{ll} \min & \operatorname{trace}(CX) \\ \text{då} & \operatorname{trace}(A_iX) \geq b_i, \quad i = 1,\dots,m, \\ & X = X^T \succeq 0. \end{array}$$

Därmed har vi fått vårt duala problem.

5. (a) Derivering ger

$$f(x) = e^{x_1+x_2} + x_1^2 + x_1x_2 + 2x_2^2 + x_1,$$

$$\nabla f(x) = \begin{pmatrix} e^{x_1+x_2} + 2x_1 + x_2 + 1 \\ e^{x_1+x_2} + x_1 + 4x_2 \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} e^{x_1+x_2}+2 & e^{x_1+x_2}+1 \\ e^{x_1+x_2}+1 & e^{x_1+x_2}+4 \end{pmatrix},$$

$$g_1(x) = 10 - x_1^2 - x_2^2,\ \nabla g_1(x) = -2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ \nabla^2 g_1(x) = -2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

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

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \begin{pmatrix} e^{x_1+x_2} + 2 + 2\lambda_1 & e^{x_1+x_2} + 1 \\ e^{x_1+x_2} + 1 & e^{x_1+x_2} + 4 + 2\lambda_1 \end{pmatrix},$$

SQP-subproblemet blir

$$(QP) \quad \begin{array}{ll} \min & \frac{5}{2}p_1^2 + 2p_1p_2 + \frac{7}{2}p_2^2 + 7p_2 \\ \text{då} & 4p_1 - 4p_2 \geq -2, \\ & p_2 \geq -1. \end{array}$$

Ursprungsproblemet är konvext med strikt konvex målfunktion, så (QP) är väldefinierat.

(b) Vi kan skriva (QP) på formen

$$(QP) \quad \begin{array}{ll} \min & \frac{1}{2}p^T H p + c^T p, \\ \text{då} & Ap \geq b, \end{array}$$

där $H = \nabla^2_{xx}\mathcal{L}(x,\lambda)$, $A = A(x)$, $c = \nabla f(x)$ och $b = -g(x)$. En iteration i en primal-dual inrepunktsmetod för (QP) tar formen

$$\begin{pmatrix} H & -A^T \\ \Lambda A & \operatorname{diag}(Ap - b) \end{pmatrix} \begin{pmatrix} \Delta p \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} Hp + c - A^T\lambda \\ \Lambda(Ap - b) - \mu e \end{pmatrix},$$

där $e = (1\ 1\ \ldots 1)^T$ och $\Lambda = \operatorname{diag}(\lambda)$. Med insatta numeriska värden får vi

$$\begin{pmatrix} 5 & 2 & -4 & 0 \\ 2 & 7 & 4 & -1 \\ 4 & -4 & 2 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta p_1 \\ \Delta p_2 \\ \Delta\lambda_1 \\ \Delta\lambda_2 \end{pmatrix} = \begin{pmatrix} 4 \\ -9 \\ -1 \\ -1 \end{pmatrix}.$$

(c) En iteration i en primal-dual inrepunktsmetod för (P) tar formen

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

I (5b) har vi $H = \nabla^2_{xx}\mathcal{L}(x,\lambda)$, $A = A(x)$, $c = \nabla f(x)$ och $b = -g(x)$. Om vi låter $p = 0$ i (5b) och använder samma λ i (5b) och här, blir ekvationssystemen identiska.

Anmärkning: Denna likhet gäller normalt endast det initiala ekvationssystemet. En inrepunktslösare för att lösa (QP) behåller ju samma värden på H, A, c och b i alla iterationer, medan en inrepunktslösare för att lösa (P) evaluerar $\nabla^2_{xx}\mathcal{L}(x,\lambda)$, $A(x)$, $\nabla f(x)$ och $g(x)$ i varje iteration.