KTH Matematik

5B1817 Tillämpad ickelinjär optimering 2006/2007

Teorifrågor

En av nedanstående åtta teoriuppgifter kommer att ges på tentan. Den valda uppgiften kommer att vara värd 10 poäng. Totalpoängen på tentan är 50.
Observera att du på tentamen måste motivera dina slutsatser ordentligt. Det ska framgå att du förstår vad du gör. Full poäng kräver ett korrekt resonemang utan logiska eller formella brister.
Eventuella satser som behövs som hjälpresultat i bevisen får användas utan att dessa bevisas.

1. Betrakta problemet (P) definierat av

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{då} & x \in I\!R^n, \end{array}$$

där f är två gånger kontinuerligt deriverbar. (Vi har alltså inga bivillkor i problemet.)

(a) Hur lyder definitionen av att x^* är en lokal optimalpunkt till (P)?

(b) Hur lyder definitionen av att x^* är en global optimalpunkt till (P)?

(c) Antag att f är konvex på $I\!R^n$, och att x^* är en lokal optimalpunkt till (P). Visa att x^* då är en global optimalpunkt till (P).

(Nash och Sofer, avsnitt 10.2 och 2.3.)

2. Betrakta problemet (P) definierat av

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{då} & x \in I\!R^n, \end{array}$$

där f är två gånger kontinuerligt deriverbar. (Vi har alltså inga bivillkor i problemet.)

(a) Formulera och bevisa andra ordningens nödvändiga optimalitetsvillkor till (P).

(b) Formulera och bevisa andra ordningens tillräckliga optimalitetsvillkor till (P).

(Nash och Sofer, avsnitt 10.2.)

3. Betrakta NLP-problemet (P) definierat av

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) = 0, \quad i = 1, \dots, m, \\ & x \in I\!R^n, \end{array}$$

där f och g är två gånger kontinuerligt deriverbara.

En reguljär punkt x^* till bivillkoren är en punkt x^* sådan att $\nabla g_i(x^*)$, $i = 1, \dots, m$, är linjärt oberoende. I en sådan punkt x^* kan tangentplanet $T(x^*)$ uttryckas som

$$T(x^*) = \mathcal{N}(A(x^*)) = \{p : A(x^*)p = 0\}.$$

(a) Formulera och bevisa första ordningens nödvändiga villkor för att en reguljär punkt x^* ska vara en lokal optimalpunkt till (P).

(b) Formulera andra ordningens nödvändiga villkor för att en reguljär punkt x^* ska vara en lokal optimalpunkt till (P).

(Nash och Sofer, avsnitt 14.5 och 14.7.)

4. Betrakta NLP-problemet (P) definierat av

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \ge 0, \quad i = 1, \dots, m, \\ & x \in I\!R^n, \end{array}$$

där f och g är två gånger kontinuerligt deriverbara.

En reguljär punkt x^* till bivillkoren är, som bekant, en punkt x^* sådan att $\nabla g_i(x^*)$, $i \in \{l : g_l(x^*) = 0\}$, är linjärt oberoende.

(a) Formulera andra ordningens nödvändiga villkor för att en reguljär punkt x^* ska vara en lokal optimalpunkt till (P).

(b) För det specialfall då $g(x) = Ax - b$, bevisa första ordningens nödvändiga villkor för att en reguljär punkt x^* ska vara en lokal optimalpunkt till (P).

Ledning (som ej kommer att finnas på tentan): I uppgift 4b kan man betrakta de primal-duala paren LP-problem (P) och (D) definierade av

$$(P) \quad \begin{array}{ll} \min & \nabla f(x^*)^T p \\ \text{då} & A_A p \ge 0, \end{array} \qquad \text{och} \qquad (D) \quad \begin{array}{ll} \max & 0^T \lambda_A \\ \text{då} & A_A^T \lambda_A = \nabla f(x^*), \\ & \lambda_A \ge 0, \end{array}$$

där A_A består av de rader i A som svarar mot bindande bivillkor i x^*. Detta ger ett alternativt sätt till bokens bevis av resultatet, och kräver inte att x^* är reguljär.
(Nash och Sofer, avsnitt 14.4, 14.5 och 14.7.)

5. Härled uttrycket för den symmetriska rang-1 uppdatering, C_k, i kvasi-Newton uppdateringen $B_{k+1} = B_k + C_k$.
(Nash och Sofer, avsnitt 11.3.)

6. Betrakta problemet (P) definierat av

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{då} & g(x) = 0, \\ & x \in I\!R^n, \end{array}$$

där $x \in I\!R^n$ och $g(x) \in I\!R^m$. Första ordningens nödvändiga optimalitetsvillkor för (P) i en reguljär punkt är (som bekant) följande:

$$\begin{aligned} \nabla f(x) - \nabla g(x)\lambda &= 0, \\ g(x) &= 0. \end{aligned}$$

Antag att Newtons metod används för att lösa detta ickelinjära ekvationssystem i $x \in I\!R^n$ och $\lambda \in I\!R^m$. Härled det linjära ekvationssystem som behöver lösas, givet den aktuella iterationspunkten (x_k, λ_k), vid en Newtoniteration. Visa att detta linjära ekvationssystem under lämpliga förutsättningar är ekvivalent med ett visst QP-problem. Ange förutsättningarna och ställ upp QP-problemet.
(Nocedal och Wright i kompletteringsbunten.)

7. Betrakta det ickelinjära matematiska programmeringsproblemet

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{då} & g(x) \geq 0, \end{array}$$

där $f : I\!R^n \to I\!R$ och $g : I\!R^n \to I\!R^m$ är kontinuerligt differentierbara.

En barriärtransformation av (P) ger för en fix positiv barriärparameter μ problemet

$$(P_\mu) \qquad \min \quad f(x) - \mu \sum_{i=1}^{m} \ln(g_i(x)).$$

Visa att första ordningens nödvändiga optimalitetsvillkor för (P_μ) är ekvivalenta med ekvationssystemet

$$\begin{aligned} \nabla f(x) - \nabla g(x)\lambda &= 0, \\ g_i(x)\lambda_i - \mu &= 0, \quad i = 1, \ldots, m, \end{aligned}$$

förutsatt att $g(x) > 0$ och $\lambda > 0$ hålls implicit.
(Nash och Sofer, avsnitt 16.2.)

8. Betrakta det semidefinita programmeringsproblemet (P) givet av

$$(P) \qquad \begin{array}{ll} \min & c^T x \\ \text{då} & G(x) \succeq 0, \end{array}$$

där $G(x) = \sum_{j=1}^{n} A_j x_j - B$ för B och A_j, $j = 1, \dots, n$, symmetriska $m \times m$-matriser. Det tillhörande duala problemet ges av

$$(D) \quad \begin{array}{ll} \max & \operatorname{trace}(BY) \\ \text{då} & \operatorname{trace}(A_j Y) = c_j, \quad j = 1, \dots, n, \\ & Y = Y^T \succeq 0. \end{array}$$

En barriärtransformation av (P) ger för en fix positiv barriärparameter μ problemet

$$(P_\mu) \qquad \min \quad c^T x - \mu \ln(\det(G(x))).$$

(a) Visa att första ordningens nödvändiga optimalitetsvillkor för (P_μ) är ekvivalenta med ekvationssystemet

$$\begin{aligned} c_j - \operatorname{trace}(A_j Y) &= 0, \quad j = 1, \dots, n, \\ G(x)Y - \mu I &= 0, \end{aligned}$$

förutsatt att $G(x) \succ 0$ och $Y \succ 0$ hålls implicit.

(b) Visa att en lösning $x(\mu)$ och $Y(\mu)$ till ekvationssystemet, sådan att $G(x(\mu)) \succ 0$ och $Y(\mu) \succ 0$, är tillåten till (P) respektive (D) med dualitetsgap $m\mu$.

(c) Till skillnad från linjärprogrammering spelar det roll hur ekvationen $G(x)Y - \mu I = 0$ skrivs för de ekvationssystem som uppstår då ekvationssystemet ovan ska lösas med Newtons metod. Exempelvis ger $G(x)Y - \mu I = 0$ respektive $YG(x) - \mu I = 0$ i allmänhet olika Newtonriktningar. Varför?

Anmärkning: För en symmetrisk matris M används ovan $M \succ 0$ respektive $M \succeq 0$ för att beteckna att M är positivt definit respektive positivt semidefinit. Du kan utan bevis använda dig av relationerna

$$\frac{\partial \ln(\det(G(x)))}{\partial x_j} = \operatorname{trace}(A_j G(x)^{-1}) \quad \text{för} \quad j = 1, \dots, n.$$

(Vandenberghe och Boyd i kompletteringsbunten samt utdelat papper.)