Tentamen i TMA945 Tillämpad optimeringslära 980819

1. Matematisk modell:

$$x_j = \text{antal enheter som körs i process } j, j = 1, 2;$$
$$y = \text{antal halvtimmar som fotomodellen anlitas}$$

$$\begin{aligned} \text{maximera } f(x, y) :=& 50(3x_1 + tx_2) - 3(x_1 + 2x_2) \\ & - 2(2x_1 + 3x_2) - 5000y, \end{aligned}$$

$$\begin{aligned} \text{då} \quad x_1 + 2x_2 &\le 20,000, \\ 2x_1 + 3x_2 &\le 35,000, \\ 3x_1 + 5x_2 &\le 1,000 + 200y \\ x_1 \ge 0, x_2 \ge 0, 0 \le y &\le \end{aligned}$$

---

2. Antag att $x_1, \dots, x_m$ är startbas. Inför ett extra bivillkor

$$\sum_{j=m+1}^{n} x_j \ge M, \quad M \ll 0.$$

Låt s vara slackvariabel i villkoret. En baslösning är $(x_1, \dots, x_m, s)$. En dualt tillåten baslösning skapas nu som följer: välj som inkommande basvariabel x_{j^*}, där $j_{j^*} \in \arg\min\{\bar{c}_j | j = m+1, \dots, n\}$. Om $c_{j^*} \ge 0$ är den nuvarande basen redan dualt tillåten, annars väljs s som utgående basvariabel. Nya reducerade kostnader är 0 för $x_1, \dots, x_m, \bar{c}_j - \bar{c}_{j^*} (\le 0)$ för $x_{m+1}, \dots, x_n$ och $-\bar{c}_{j^*}$ för s, och alltså gäller att den nya basen är dualt tillåten.

Detta är ekvivalent med stora M-metoden. Det framgår till exempel genom att studera dualen till det nya problemet.

$[p]$ minimera $f(x) := c^T x$

$$\begin{aligned} \text{då} \quad Ax &= b \qquad \mathbf{1} = (1, 1, \dots, 1)^T \\ \mathbf{1}^T x &\le M \\ x &\ge 0 \end{aligned}$$

har ett dualt problem som är ekvivalent med LP-problemet.

$[D'']$ maximera $g(y, a) := b^T y - Ma$,

$$\begin{aligned} \text{då} \quad A^T y - \mathbf{1} a &\le c, \\ a &\ge 0 \end{aligned}$$

och de två pivoteringarna är dualt ekvivalenta, ty att prirotera in a i $[D'']$ så att högerledet $(\bar{c})$ blir positivt är precis detsamma som att pivotera ut s i $[p']$.

---

3. Fall I: $\{\nabla f(x^T)\} \to 0^n, \{x^t\}$ och $\{f(x^t)\}$ divergerar.

   Exempel: $f(x) = -\log x; \{x^t\} \to \infty, \{f(x^t)\} \to -\infty, \{f'(x^t)\} \to 0$

   Fall II: $\{\nabla f(x^t)\} \to 0^n, \{x^t\}$ divergerar , $\{f(x^t)\}$ konvergerar.

   Exempel: $f(x) = 1/x; \{x^t\} \to \infty, \{f(x^t)\} \to 0, \{f(x^t)\} \to 0$

   Fall III: $\{\nabla f(x^t)\} \to 0^n, \{x^t\}$ är begränsad, $\{f(x^t)\}$ är begränsad.

   Exempel: $f(x) = \frac{1}{3}x^3 - x; x^t = \begin{cases} 1 + 1/t, & t \text{ udda} \\ -1 - 1/t & t \text{ jämna} \end{cases}$

   $\{x^t\}$ har två hopningspunkter, ± 1, och $\{f(x^t)\}$ har två hopningspunkter $\pm 2/3$.

   Fall IV: $\{\nabla f(x^t)\} \to 0^n, \{x^t\}$ är begränsad, $\{f(x^t)\}$ konvergerar.

   Exempel: $f(x) = x^2 - 1; x^t$ som ovan; $\{f(x^t)\} \to 0$.

   Fall V: $\{\nabla f(x^t)\} \to 0^n, \{x^t\}$ och $\{f(x^t)\}$ konvergerar.

   Exempel: f som ovan, $x^2 - 1 + 1/t$

4 a) I en baslösning (x_0, x_N) gäller att

$$BX_B + NX_N = b \Leftrightarrow X_B = B^{-1}b - B^{-1}NX_N. \quad \text{Om } x_j (j \in N)$$

ökar från noll fås därför att

$$X^{NY} = \begin{bmatrix} B^{-1} \\ \mathbf{0} \end{bmatrix} + \ell \begin{bmatrix} -B^{-1}a_j \\ e_j \end{bmatrix},$$

där a_j är kolumn j i N och e_j ör den j:te enhetsvektorn $\quad \because p_j = \begin{bmatrix} -B^{-1}a_j \\ e_j \end{bmatrix}$

I baslösningen gäller också att den reducerade kostnaden är

$$\bar{C}_N^T = C_N^T - C_B^T B^{-1} N.$$

Vi har nu att

$$\begin{aligned} C^T \rho_j = (C_B, C_N)^T \rho_j &= -C_B^T B^{-1} a_j + C_N^T e_j \\ &= C_j - C_B^T B^{-1} a_j = \bar{C}_j. \end{aligned}$$

Att välja minsta värdet på $\bar{C}_j$ är alltså detsamma som att välja minsta värdet av ρ_j.

b) Inkommandekriteriet är skalningsberoende, och beror till exempel på längden hos varje halvlinjes riktningsvektor ρ_j. Ett skalningsoberoende mått är i stället

$$\min_{j \in N} \left\{ \frac{C^T \rho_j}{\|\rho_j\|} \right\}.$$

---

5. Antag att $\hat{x} \in \hat{X}$. Konvexiteten hos f ger att

$$f(x^*) \geq f(\hat{x}) + \nabla f(\hat{x})^T (x^* - \hat{x}) = f(\hat{x}) + \nabla f(x^*)^T (x^* - \hat{x})$$
$$= f(\hat{x}). \text{ Alltså är } \hat{x} \in x^*, \text{ och } \hat{X} \leq X^* \text{ gäller.}$$

Antag att $\hat{x} \in X^*$. Konvexiteten hos f ger att

$$\left. \begin{array}{l} f(x^*) \geq f(\hat{x}) + \nabla f(\hat{x})^T (x^* - \hat{x}) \\ f(\hat{x}) \geq f(x^*) + \nabla f(x^*)^T (\hat{x} - x^*) \end{array} \right\} \Rightarrow [\nabla f(\hat{x}) = \nabla f(x^*)]^T (\hat{x} - x^*) \geq 0.$$

Men den omvända olikheten gäller också, eftersom

$$\begin{array}{l} \hat{x} \in X^* \Rightarrow \nabla f(\hat{x})^T (x^* - \hat{x}) \geq 0 \\ x^* \in X^* \Rightarrow \nabla f(x^*)^T (\hat{x} - x^*) \geq 0 \end{array} \quad \text{enligt optimalitetsvillkorer.}$$

Alltså gäller att $[\nabla f(x^*) - \nabla f(\hat{x})]^T (x^* - \hat{x}) = 0$. Vi får då ur konvexitetsolikheten

$$f(\hat{x}) \geq f(x^*) + \nabla f(x^*)^T (\hat{x} - x^*)$$

att $\nabla f(x^*)^T (\hat{x} - x^*) = 0$ måste gälla, ty $f(\hat{x}) = f(x^*)$ gäller och $\nabla f(x^*)^T (\hat{x} - x^*) \geq 0$. Från konvexiteten hos f fås nu att $f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*) = f(\hat{x}) + \nabla f(x^*)^T (x - \hat{x}), \forall x \in X$, ty $f(\hat{x}) = f(x^*)$. Alltså är $\nabla f(x^*) = \nabla f(\hat{x})$.

Alltså är $\hat{x} \in \hat{X}$, och $x^* = \hat{X}$ gäller. ■

---

6 a) Lös LP-problemet med målfunktionskoefficient $c^2 + \lambda_p$, där $p = C^1 - C^2$, och lös först för $\lambda = 0$.

Givet optimalbasen B bestäms det högsta värdet för λ för vilket optimalitet bibehålls: vi skall ha

$$(C_N^2 + \lambda p_N)^T - (C_B^1 + \lambda p_B)^T B^{-1} N \leq \mathbf{0} \Leftrightarrow$$
$$\lambda (p_N^T - p_B^T B^{-1} N) \leq -(C_N^{2T} - C_B^{T_2} B^{-1} N).$$

Av intresse är de i för vilka $(p_N^T - p_B^T B^{-1} N)_i > 0$. Vi får:

$$\bar{\lambda} = \min_i \left\{ -\frac{(C_N^{2^T} - C_B^{2^T} B^{-1} N)_i}{(p_N^T - p_B^T B^{-1} N)_i} \middle| (p_N^T - p_B^T B^{-1} N)_i > 0 \right\}.$$

Det index i som ger min motsvarar den inkommande basvariabeln vid besbytet som sker då $\lambda > \bar{\lambda}$. Om inget index i finns betyder det att basen är optimal för alla $\lambda \geq 0$, annars pivoteras till en ny optimal bas är nådd, varefter förfarandet upprepas från nästa aktuella värde av $\lambda(\bar{\lambda})$.

Kompromisslösningarna består av de optimala lösningar som passerats då λ har ändrats från 0 till 1.

b) Geometrisk lösning ger att

$$x^* = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \text{ för } \lambda \in [0, 1/2]$$
$$x^* = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \text{ för } \lambda \in [1/2, 1]$$

---

7 a) Brantaste lutningsmetodens sökriktning ges till exempel av $\min \nabla f(x^t)^T p + 1/(2\gamma)\|\rho\|^2$, ty första ordningens villkor är att $\nabla f(x^t) + (1/\gamma)\rho = 0^n$, d.v.s. $\rho = -\gamma \nabla f(x^t)$. Motsvarande p från $\min \nabla f(x^t)^T \rho$ då $\|p\| \leq \Delta$ fås genom att betrakta det ekvivalenta problemet $\min \nabla f(x^t)^T p$. Eftersom målfunktionen då $\|p\|^2 \leq \Delta^2$ är linjär är bivillkoret med nödvändighet bindande, varför vi får att, för något $\lambda > 0$,

$$\nabla f(x^t) + 2\lambda \cdot \rho = 0^n \text{ eller } \rho = -1/(2\lambda)\nabla f(x^t),$$

dvs samma riktning som minus gradienten av f i x^t i båda fallen.

b) I. Ett ekvivalent problem är

$$\begin{aligned} &\min \psi_t(p) \\ &\text{då } \|p\|^2 \leq \Delta_t^2. \end{aligned}$$

första ordningens nödvändiga krav ger att

$$\begin{aligned} \nabla \psi_t(\rho) + \lambda \rho &= \mathbf{0}^n, \text{ dvs} \\ \nabla f(x^t) + \nabla^2 f(x^t)\rho + \lambda\rho &= \mathbf{0}^n, \text{ eller} \\ [\nabla^2 f(x^t) + \lambda I^n]\rho &= -\nabla f(x^t). \end{aligned}$$

Andra ordningen nödvändiga krav ger att

$$\nabla^2 f(x^t) + \lambda I^n \text{ är positivt semidefinition 7.}$$

II. Om $\lambda = 0$ fås att p är Newtonriktningen. Då λ ökar mot ∞ fås i stället en sökriktning som i limes är brantaste lutningsriktningen. (Det förra händer om Δ_t är start nog medan det senare inträffar när Δ_t blir allt mindre. För att se det kan man använda relationen $\Delta_t \geq \|p\| = \|(\nabla^2 f(x^t) + \lambda I)^{-1}\nabla f(x^t)\|$.)

Fördelen med en trustregion är att medan en sökmetod fixerar sökningen så får den här variera (med Δ_t) så att metoden bättre anpassas till problemets egenskaper.