Lösning till Tentamen TMA946/MAN280 Tillämpad Optimeringslära 020311

1a. Vi har problemet

$$\begin{aligned} \min \quad & 3x_1 + x_2 \\ \text{Då} \quad & 3x_1 + 2x_2 \geq 1 \\ & 2x_1 + x_2 \leq 2 \\ & x_1, \ x_2 \geq 0 \end{aligned}$$

till för slackvariabler:

$$\begin{aligned} \min \quad & 3x_1 + x_2 \\ \text{Då} \quad & 3x_1 + 2x_2 - s_1 = 1 \\ & 2x_1 + x_2 + s_2 = 2 \\ & x_1, x_2, s_1, s_2 \geq 0 \end{aligned}$$

Ingen uppenbar bas: inför fas I målfunktion och artificiell variabel

$$\begin{aligned} \min \quad & a_1 \\ \text{Då} \quad & 3x_1 + 2x_2 - s_1 + a_1 = 1 \\ & 2x_1 + x_2 + s_2 = 2 \\ & x_1, x_2, s_1, s_2, a_1 \geq 0 \end{aligned} \quad \Leftrightarrow \quad \begin{array}{c} \min[0 \quad 0 \quad 0 \quad 0 \quad 1] \begin{array}{c} \alpha_1 \\ \vdots \\ a_1 \end{array} \\ \begin{bmatrix} 3 & 2 & -1 & 0 & 1 \\ 2 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{array}{c} \alpha_1 \\ \vdots \\ a_1 \end{array} = \begin{array}{c} 1 \\ 2 \end{array} \end{array}$$

Vi har bas a_1, s_2 and $B = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$, $B^{-1} = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$

Reducerad kostnad blir

$$\bar{c} = c_N^T - c_B^T B^{-1} N = [0 \quad 0 \quad 0] - [1 \quad 0] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & -1 \\ 2 & 1 & 0 \end{bmatrix} = [-3, -2, 1]$$

Vi väljer första icke-basvariabel som inkommande, ty den har lägst reducerad kostnad.

↙ kolonn ur A för x_1

Sätt $y = B^{-1}_{\text{ink}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{matrix} 3 \\ 2 \end{matrix}$

Vidare har vi $B^{-1}b = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{matrix} 1 \\ 2 \end{matrix} = \begin{matrix} 1 \\ 2 \end{matrix}$

Vi får nu utgående som

$$u_t = \operatorname{argmin}_{i, y_i > 0} \frac{(B^{-1}b)_i}{y_i} = 1 \text{ ty } \frac{1}{3} < \frac{2}{2}$$

↑ Det index

↑ som minimerar detta

Därmed blir första basvariabeln (a_1) utgående.

Då a_1 ej är med i basen vet vi att $a_1 = 0$, och vi har en inåten bas till originalproblemet.

Fas II: Vi har bas $x_1 s_2$ och problemet

$$\min[3 \quad 1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ \vdots \\ s_2 \end{bmatrix}$$

Då $\begin{bmatrix} 23 & 2 & -1 & 0 \\ 2 & 1 & 1 & \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ s_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$B = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}, B^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix}$

Vi får reducerad kostnad (Bas=$x_1 s_2$), (icke bas (x_2, s_1))

$$\bar{c} = c_N^T - c_B^T B^{-1} N = [1 \quad 0] - [3 \quad 0] \frac{1}{3} \begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} =$$

$$= [1 \quad 0] - [2 \quad -1] = [-2 \quad 1]$$

Vi får att första icke-basvariabeln blir inkommande (d.v.s. x_2).

Bilda

$$y = B^{-1}A_{\text{ink}} = \frac{1}{3}\begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \end{bmatrix}, B^{-1}b = \frac{1}{3}\begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix}\begin{matrix} 1 \\ 2 \end{matrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$ut = \text{argmin}_{i,y_i>0}\frac{(B^{-1}b)_i}{y_i} \Rightarrow ut = 1 \Rightarrow x_1$ blir utgående. (första basrum)

Ny bas: $x_2, s_2, B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \quad B^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$

$$\bar{c} = c_N^T - C_B^T B^{-1}N = [3 \quad 0] - [1 \quad 0]\frac{1}{2}\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} =$$

$$= [3 \quad 0] - \frac{1}{2}[3 \quad -1] = \frac{3}{2} \quad 1$$

$\bar{c} \geq 0 \Rightarrow x_2, s_2$ är optimal bas.

Vi har $\begin{bmatrix} x_2 \\ s_2 \end{bmatrix} = B^{-1}b = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{matrix} 1 \\ 2 \end{matrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \end{bmatrix}$

1b. Ritar vi en bild får vi

Om vi minskar d får vi $x_1 = x_2 = 0$ då $d = 0$

Vidare blir problemet olösligt för $d > 4$.

Titta analytiskt: om vi förändrar hl får vi

$$\begin{matrix} x_2 \\ s_2 \end{matrix} = B^{-1} \begin{bmatrix} d \\ 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} d \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{d}{2} \\ 2 - \frac{d}{2} \end{bmatrix}$$

basen är tillåten för $0 \leqslant d \leqslant 4$

$$z(d) = c_B^T \cdot x_B(d) = [1 \quad 0] \begin{bmatrix} \frac{d}{2} \\ 2 - \frac{d}{2} \end{bmatrix} = \frac{d}{2}$$

för $d < 0$ har vi $x_1 = x_2 = 0 \Rightarrow z = 0$

Vi får därmed $z(d) = \begin{cases} 0\,, & d < 0 \\ \frac{d}{2}\,, & 0 \leqslant d \leqslant 4 \\ \infty\,, & d > 4 \\ \uparrow & \\ \text{problemet} & \\ \text{otillåtet} & \end{cases}$

2a. Vi kallar Kalles lösning för $\bar{x}$ och motsvarande duala lösning för y.

Vi har problemet

$$\begin{array}{rlcr} \min & c^T x & & \max b^T y \\ \text{Då} & Ax \geq b & \text{med dual} & \text{Då } A^T y \leq c \\ x \geq 0 & & & y \geq 0 \end{array}$$

Optimalitetskriterierna för LP ger oss att (komplementaritet)

$$y^T(A\bar{x} - b) = 0 \quad \bar{x}^T(A^T y - c) = 0$$

Låt $\bar{A}$ vara de kolonner som svarar mot $\bar{x} > 0$. P.S.S. för $\bar{c}$. Då är $\bar{A}$ en $m \times m + k$ matris för att vi skall ha

$$\bar{A}^T y = \bar{c}.$$

Detta överbestämda ekvationssystem har exakt en lösning, den unika duala lösningen.

2b. Titta på $A\bar{x} \geq b$.

Vi lägger till slackvariabler och får $[AI] \begin{matrix} \bar{x} \\ \\ \bar{s} \end{matrix} = b$, $\bar{x}$ och $\bar{s}$ är Kalles optimala lösning.

Välj nu ut de kolonner som svarar mot $\bar{x}$ eller $\bar{s} > 0$ och vi får ekvationssystemet

$$\tilde{A}^0 = b \quad \text{med} \tilde{x}^0 > 0$$

Hitta nu en vektor p. $\tilde{A}^0$:s nollrum.

Då $\tilde{x} > 0$ har vi att $\tilde{x}^0 + \alpha p^0 > 0$ för små α.

Välj nu α s..a. $\tilde{x}^0 + \alpha p^0 \geq 0$ med minst ett element $= 0$ (notera att vi vet att $\tilde{c}^T p^0 = 0$ ty annars hade $\bar{x}$ aldrig varit optimal.)

Vi kan nu ta bort de kolonner i $\tilde{A}^0$ som svarar mot $\tilde{x} + \alpha p = 0$. Vi får då $\tilde{A}^2$ med färre kolonner.

Vi itererar tills dess $\tilde{A}$:s nollrum endast innehåller 0-vektorn. Vi har då hittat ett hörn av polyedern och därmed en baslösning.

3. $\min \frac{1}{2}(x_1^2 + |x_2|^\rho) + 2x_2$

   s.t. ${}_2 = 0$.

   1) KKT-conditions ($\rho > 1$):

$$\begin{cases} x_2 = 0 \\ \begin{pmatrix} x_1 \\ \frac{\rho}{2}|x_2|^{\rho-1} + 2 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases} \overset{\text{primal feasibility}}{\Leftrightarrow} \begin{cases} x_1 = 0 \\ x_2 = 0 \\ \lambda = -2 \end{cases}$$

   KKT are sufficient for global optimality because the problem is convex.

   2) $\rho = 2 \qquad \sigma > 0$

$$\min \frac{1}{2}(x_1^2 + |x_2|^2) + 2x_2 + \sigma x_2^2 = \min \mathcal{F}(x_1, x_2)$$

$$\nabla \tilde{f} = \begin{pmatrix} x_1 \\ x_2 + 2 + 2\sigma x_2 \end{pmatrix} = 0 \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = -\frac{2}{1+2\sigma} \end{cases}$$

   $\tilde{f}$ is convex $\Rightarrow \nabla \tilde{f} = 0$ is a sufficient for otpimality condition

$$x_\sigma^* = \begin{pmatrix} 0 \\ -\frac{2}{1+2\sigma} \end{pmatrix} \xrightarrow[\sigma \to \infty]{} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = x^*$$

3) $\rho = 1 \quad \sigma > 0 \min \frac{1}{2}(x_1^2 + |x_2|) + 2x_2 + \sigma x_2^2 = \min \bar{f}(x_1, x_2)$
3 cases:

(1) $x_2 > 0 \,\&\, \nabla \bar{f} \begin{pmatrix} x_1 \\ -\frac{1}{2} + 2 + 2\sigma x_2 \end{pmatrix} = 0$ – contradiction

(2) $x_2 = 0 \to \min \frac{1}{2}x_1^2 \Rightarrow x_1 = 0$

$\bar{f}(0,0) = 0$

(3) $x_2 > 0 \quad \nabla \bar{f} = \begin{pmatrix} x_1 \\ -\frac{1}{2} + 2 + 2\sigma x_2 \end{pmatrix} = 0 \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = -\frac{3}{4\sigma} \end{cases}$

$\bar{f}(0, -\frac{3}{4\sigma}) = -\frac{3}{16\sigma} < 0$ – golobal mh! ($\bar{f}$ is convex)

$$x_\sigma^* = \begin{pmatrix} 0 \\ -\frac{3}{4\sigma} \end{pmatrix} \xrightarrow[\sigma > \infty]{\qquad\qquad} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = x^*$$

We cannot use groelient-based method, because $\bar{f}$ is not differentiable!

4a. Definition: en funktion $f : \mathbb{R}^n \mapsto \mathbb{R}$ är konvex om, for varje val av $x^1, x^2 \in \mathbb{R}^n$ och $\lambda \in [0,1]$ gäller att

$$(\circledast) \qquad f(\lambda x^1 + (1-\lambda)x^2) \le \lambda f(x^1) + (1-\lambda) f(x^2).$$

C^1: För varje val av $x^2, x^2 \in \mathbb{R}^n$ gäller att

$$(\circledast\circledast) \qquad f(x^2) \ge f(x^1) + \nabla f(x^1)^T (x^2 - x^1).$$

C^2: För varje val av $x \in \mathbb{R}^n$ och $p \in \mathbb{R}^n$ gäller att

$$p^T \nabla^2 f(x) p \ge 0.$$

C^1 igen: ($\Leftarrow$) Vi har: $\circledast\circledast$ ger:
$f(x^1) \ge f(\lambda x^1 + (1-\lambda)x^2) + (1-\lambda)\nabla f(\lambda x^1 + (1-\lambda)x^2)^T (x_1 - x_2)$
$f(x^2) \ge f(\lambda x^1 + (1-\lambda)x^2) + \lambda \nabla f(\lambda x^1 + (1-\lambda)x^2)^T (x_2 - x_1)$
Addera: $\lambda \cdot$ (rad 1) $+ (1-\lambda) \cdot$ (rad 2) $\Rightarrow \circledast$

($\Rightarrow$) $\circledast \Rightarrow (1-\lambda) f(x^2) \ge (1-\lambda) f(x^1) + [f(x^1 + (1-\lambda)(x^2 - x^1)) - f(x^1)]$.
Dividera med $1-\lambda$ och låt $\lambda \to 1$. Då fås $\circledast\circledast$

b) Det mesta talar för en quasi-Newton algoritm: vi har bara tillgång till ∇f, inte $\nabla^2 f$; problemet är konvext, så approximationen av Newton-riktningen bör vara bra; problemet är ganska stort men varje rang-2-uppdatering fordrar bara ytterligare $2 * n$ i minnesutrymme.

c) (D) $\max (b^k)^T y$ Lös dualen! Återanvänd den optimala basen för nästa k – reoptimera!

$$\begin{aligned} \max\ & (b^k)^T y \\ \text{då } & A^T y \leq c \\ & y \geq \mathbf{0} \end{aligned}$$

5. Variabler: d : godstjocklek (m)

ℓ : längd hos cylinder delen (m)

r : radie har kloten (m)

Modell:

minimera $(2\pi_r \ell + 4\pi r^2)d$

då

$$\begin{aligned} 2r + \ell &\leq 5.87 - 2d && \text{(längd)} \\ 2r &\leq 2.34 - 2d && \text{(bredd)} \\ prc_2 &\leq d && \text{(kapacitet)} \\ p(\pi r^2 + 3/4\pi r^3) &= c_1 && \text{(tryck-volym)} \\ d, r, \ell &\geq 0 \end{aligned}$$

6a) Lagrange funktion: $L(x, \lambda) := f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$.

Lagrangedual funktion: $L_*(\lambda) := \min_{x \in X} L(x, = lambda), \lambda \geq \mathbf{0}^m$.

Det Lagrangeduala problemet är:

(LD) $$\text{maximera}_{\lambda \geq \mathbf{0}^m} L_*(\lambda).$$

Svag dualitet: Antag att $\bar{x}$ är tillåten i ursprungsproblemet (1)–(3), och att $\bar{\lambda} \geq \mathbf{0}^m$, dvs, tillåten i (LD). Då gäller att $L_*(\bar{\lambda}) \leq f(\bar{x})$.

Bevis:

$$\begin{aligned} f(\bar{x}) &\geq f(\bar{x}) - \sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{x}) && /\bar{\lambda} \geq \mathbf{0}^m, g(\bar{x}) \geq \mathbf{0}^m/ \\ &= L(\bar{x}, \bar{\lambda}) \\ &\geq L_*(\bar{\lambda}). \quad \blacksquare && /\bar{x} \text{ ej säkert i } \operatorname{argmin}_{x \in X} L(x, \bar{\lambda})/ \end{aligned}$$

b) Antag att $\bar{\lambda} \geq \mathbf{0}^m$ (tillåten i (LD)). Om följande mängd är icke-tom innehåller den samtliga optimala lösningar till (1)–(2), $\bar{\lambda}$ är optimal i (LD), och $L_*(\bar{\lambda}) = f(x^*)$:

$$X^* = \operatorname{argmin}_{x \in X} L(x, \bar{\lambda}) \cap \{x \in \nwarrow^n | g(x) \geq \mathbf{0}^m\} \cap \{x \in \nwarrow^n | \bar{\lambda}^T g(x) = 0\}$$

Bevis: Låt $x^* \in X^*$. Då är x^* tillåten (1)–(3).

Dessutom: $L_*(\bar{\lambda}) = f(x^*) - \bar{\lambda}^T f(x^*) = f(x^*)$. Eftersom svag dualitet gäller enligt a) måste x^* vara optimal i (1)–(3) och $\bar{\lambda}$ i (LD).

c) (P) $\quad \begin{array}{ll} \min & z = c^T x \\ \text{då} & Ax \geq b \\ & x \geq \mathbf{0} \end{array} \qquad (D) \quad \begin{array}{ll} \max & r = b^T y \\ \text{då} & A^T y \leq c \\ & y \geq 0 \end{array}$

För (P) är då $L(x, \lambda) := c^T x - \lambda^T (Ax - b) = (c - A^T \lambda)^T x + b^T \lambda$, och med $X = \{x \in \mathbb{R}^n | x \geq \mathbf{0}^n\}$ får att

$$L_x(\lambda) := \begin{cases} b^T \lambda & \text{om } c \geq A^T \lambda \\ -a & \text{annars.} \end{cases}$$

Alltså är (Ld) ekvivalent med (D).

Optimalitetsvillkoren är: x^* tillåten i (P), y^* tillåten i (D), x^* och y^* är komplementära.

$$\begin{array}{ccccc} Ax^* \geq b & & A^T y^* \leq c & & y^{*T}(Ax^* - b) = 0 \\ & ; & & ; & \\ x^* \geq \mathbf{0} & & y^* \geq \mathbf{0} & & x^{*T}(A^T y^* - c) = 0 \\ (a) & & (b) & & (c) \end{array}$$

Från $g(x^*) \geq \mathbf{0}$ och $x^* \in X$ finner vi (a).

Från $\bar{\lambda}^T g(x^*) = 0$ finner vi $y^{*T}(Tx^* - b) = 0$.

Kvar: x^* löser $\min_{x \in X} L(x, \bar{\lambda})$. I fallet LP fås: x^* minimerar $(c - A^T y^*)^T x$ över $x \geq \mathbf{0}$. Detta problem separerar över x_j: $\min(c_j - A_j^T b^*) x_j$ över $x_j \geq 0$.

Optimum: Antingen är $\quad x_j^* > 0$ och $c_j = A_j^T y^*$ eller

$\qquad\qquad\qquad\qquad\qquad\; x_j^* = 0$ och $c_j \geq A_j^T y^*$ eller

$\because x^* \geq \mathbf{0}; c \geq A^T y^*; x^{*T}(A^T y^* - c)$, dvs resten av (a)–(c). Klart.

7. Refined optimality conditions

   1) Necessary conditions Suppose x_0 is a local min, but

$$f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \text{ and } f^{(n)}(x_0) \neq 0$$

   for some *odd* integer n.

Taylor expansion:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \ldots + \frac{1}{n!} f^{(n)}(\tilde{x})(x - x_0)^n =$$
$$= f(x_0) + \frac{1}{n!} f^{(n)}(\tilde{x})(x - x_0)^n,$$

where $\tilde{x}$ is a point between x and x_0.
For all sufficiently close to x_0 points x both $f^{(n)}(\tilde{x})$ and $f^{(n)}(x_0)$ have the same sign. Therefore, we can get $f(x) < f(x_0)$ by approximating x_0 from left (if $f^{(n)}(x_0) > 0$) or right (if $f^{(n)}(x_0) < 0$), which contradicts local minimality of x_0.
The other possibility is to have

$$f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \,\&\, f^{(n)}(x_0) < 0$$

for some *even* integer n.
Taylor expansion: $f(x) = f(x_0) + \frac{1}{n!} f^{(n)}(\tilde{x})(x - x_0)^n < f(x_0)$ for all sufficiently close to x_0 points x!

2) <u>Sufficient conditions</u>
Taylor expansion

$$f(x) - f(x_0) = \frac{1}{n!} f^{(n)}(\tilde{x})(x - x_0)^n > 0$$

for all x, sufficiently close to x_0 (since $\tilde{x}$ is a point between x and x_0, $f^{(n)}(x_0) > 0$, and $f^{(n)}(\cdot)$ is a continuous function).

3) "Gap"

$$\begin{array}{lll} f_1(x) < 0 & \text{for} & x < 0 \\ f_1(x) > 0 & \text{for} & x > 0 \end{array} \Rightarrow \begin{array}{l} f(0) = 0 \\ x_0 = 0 \end{array} \quad \text{is not a local min}$$

$$\begin{array}{lll} f_2(x) > 0 & \text{for} & x \neq 0 \\ f_2(x) = 0 & \text{for} & x = 0 \end{array} \Rightarrow x_0 = 0 \text{ is a strict local min}$$

However, $f_i^{(n)}(0) = \lim_{x \to 0} f_i^{(n)}(x) = \lim_{x \to 0} P_i^n(\frac{1}{x}) e^{-1/x^2} = 0$, where P_n^i is a some polynomial function. Therefore, necessary conditions hold for f_1 at $x_0 = 0$, but it is not a local minimum!
Sufficient conditions are violated for f_2 of $x_0 = 0$, but it is a strict local minimum!