Tentamen TMA946/MAN280 tillämpad optimeringslära 010821

1. Uppgift: $\min z = 3x_1 + x_2$

   $$\text{Då} \left[\begin{array}{rl} 2x_1 + x_2 & \geq 6 \\ -x_1 + x_2 & \geq -2 \\ x_1, x_2 & \geq 0 \end{array} \right]$$

   Skriv på standardform m.h.aślackvariabler

   $\min z = 3x_1 + x_2$

   $$\text{Då} \left[\begin{array}{r} 2x_1 + x_2 - s_1 = 6 \\ x_1 - x_2 + s_2 = 2 \\ x_1, x_2, s, s_2 \geq 0 \end{array} \right]$$

   Vi ser ingen uppenbar bas $\Rightarrow$ skapa FAS I problem

   $\min w = a$

   $$\text{Då} \left[\begin{array}{r} 2x_1 + x_2 - s_1 + a = 6 \\ x_1 - x_2 + s_2 = 2 \end{array} \right]$$

   Välj a, s_2 som bas $\Rightarrow B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

   Titta på reducerad kostnad:

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

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

   $\Rightarrow$ välj x_1 som inkommande.

   Vi gör min-ratio test för att hitta utgående

   $$Y = B^{-1} A_{ink} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

   $$\bar{b} = B^{-1} b = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

Utgående $= \begin{bmatrix} \text{argmin} \\ i, Y_i > 0 \end{bmatrix} \dfrac{\bar{b}}{Y_i} \Rightarrow$ bas variabel 2 utgår.

Ny bas blir a, x_1

$$\Rightarrow B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} B^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} B^{-1}b = \begin{bmatrix} 2 \\ 2 \end{bmatrix} c_N^T = [0 \quad 0 \quad 0]$$
$$c_B^T = [1 \quad 0] N = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Hitta reducerad kostnad $C_N^T - C_B^T B^{-1} N =$

$$[0 \quad 0 \quad 0] - [1 \quad 0] \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} = [-1 \quad 2] \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} = [-3 \quad 1 \quad 2]$$

icke basvar 1, d.v.s. x_2 blir inkommande.

Gör min ratio.

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

utgående $= \begin{bmatrix} \text{argmin} \\ i, Y_i > 0 \end{bmatrix} \frac{\bar{b}_i}{Y_i} \Rightarrow$ basvar 1 utgår.

Ny bas blir x_2, x_1.

Vi har ej länge någon artificiell variabel i basen $\Rightarrow$ Fas I är klar.

Starta fas II. Vi har bas x_1, x_2

$$\Rightarrow B = \begin{bmatrix} 2 & 1 \\ 1 & = 1 \end{bmatrix} \Rightarrow B^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} B^1 b = \frac{1}{3} \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$
$$c_B^T = [3 \quad 1] c_N^T = [0 \quad 0] N = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hitta reducerad kostnad. $\bar{c} = c_N^T - c_B^T - c_B^T B^{-1} N = [0 \quad 0] - [3 \quad 1] \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
$= -[4/3 \quad 1/3] \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = [4/3 \quad -1/3]$ icke basvar $2 = s_2$ är inkommen.

Vi gör min ratio test. $Y = B^{-1} Aink = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \end{bmatrix}$

Utgående $= \begin{bmatrix} \text{argmin} \\ i, y_i > 0 \end{bmatrix} \frac{\bar{b}_i}{y_i} \Rightarrow$ basvar 1 är utgående.

Ny bas x_2, s_2.

Hitta reducerad kostnad:

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

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

$$\Rightarrow i = [3 \quad 0] - [1 \quad 0] \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = 1 \quad 1 \geq 0$$

$\Rightarrow$ vi har en optimal bas, x_1, s_2

med värden $B^{-1}b = \begin{bmatrix} 1 & 0 & 6 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix} \Rightarrow Z^* = 3x_1 + x_2 = 6$

Vi har $x_1 = 0, x_2 = 6, s_1 = 0, s_2 = 8$.

Kolla! $\begin{bmatrix} \\ 2x_1 + x_2 = 6 \geq 6 \quad \text{OK} \\ -x_1 + x_2 = 6 \geq -2 \quad \text{OK} \end{bmatrix}$

1b För att avgöra för vilka c mängden påverkar bildar vi problemet. $\min z = x_1 + x_2$

Då $\begin{bmatrix} \\ \\ 2x_1 + x_2 \geq 6 \\ -x_1 + x_2 \geq -2 \\ x_1, x_2 \geq 0 \end{bmatrix}$

Antag att det optimala värdet på detta problem är z^*. Då påverkas mängden för $c > z^*$.

2a) Ett exakt straff för "$g_i(x) \geq 0$" är $\min\{0, g_i(x)\}$. En straffunktionsmetod baserad på denna är följande:

0. Välj $\mu_0 \geq 0$. Sätt $t = 0$.
1. Lös $\min_{x \in \mathbb{R}^n} f(x) - \mu^t \cdot \min\{0, g_i(x)\} \to x^t$.
2. Sätt $\mu^{t+1} > \mu^t$ (så att $\lim_{t\to\infty} \mu^t = +\infty$), $t := t + 1$, gå till 1.

b) Låt $x := \{x = \mathbb{R}^n | a_i^T x \geq b_i, i = 1, \ldots, m\}$. Frank-Wolfe algoritmens iteration är följande:

0. Välj $x^0 = X$. Sätt $t = 0$.
1. Lös $\min_{y \in X} \nabla f(x^t)^T y \to y^t$.
2. Om $\nabla f(x^t)^T (b^t - x^t) = 0 \Rightarrow$ Stopp! x^t är en KKT-punkt.
3. Lös $\min_{\ell \in [0,1]} f(x^t + \ell(y^t - t^t)) \to \ell_t$.

4. Sätt $x^{t+1} = x^t + \ell_t(y^t - x^t), t := t+1$, gå till 1.

För att en iteration skall kunna genomföras krävs att stegen 1 och 3 (de två optimeringsproblemen) är genomförbara. Steg 1 är genomförbart om LP-problemet har en lösning, d.v.s. om $\nabla f(x^t)^T y$ är nedåt begränsad på X. Steg 2 är genomförbart om f har ett minimum på linjesegmentet $[x^t, y^t]$. För detta räcker det med att f är kontinuerligt differentiarbar eftersom den då är kontinuerlig och $[x^t, y^t]$ är en kompakt mängd (Weierstrass). For steg 1 fordras i allmänhet att x *är begränsad.*

c) Eftersom f bara är differentierbar en gång har vi inte tillgång till rena Newtonmetoder. Problemet är konvext och inte särdeles stort ($n = 500$ är att betrakta som ganska litet för ett konvext obegräansat problem). Om $\nabla f(x)$ är någorlunda lätt att beräkna rekommenderas Qvari-Neewton/konjugerade gradientmetoder. (se kurslitteraturen för beskrivningar.)

3 a) Modell: Variabler:

x_{ij} = andel av rxxxx j:s efterfrågan som tillgodoses av central i $\forall i, \forall j$.

$$y_i = \left\{ \begin{array}{ll} 1, & \text{om central i byggs, } \forall i. \\ 0, & \text{annars.} \end{array} \right]$$

$$\text{Ny konstant: } a_{ij} = \left\{ \begin{array}{l} 1 \text{ om } d_{ij} \leq D \\ 0 \text{ annars} \end{array} \right] \forall i, \forall j \text{ (uppnåelighetsmatris)}$$

$$\text{Minimera} \sum^{10} i = j c_i b_i$$

Då

$$\begin{aligned}
x_i j &\leq a_{ij} y_i, i = 1, \ldots, 10; j = 1, \ldots, 30 \quad \text{(uppnåelighet)} \\
\sum_{j=1}^{30} e_j x_{ij} &\leq k_i y_i, i = 1, \ldots, 10 \quad \text{(tillgång)} \\
\sum_{i=1}^{10} k_{ij} &= 1, j = 1, \ldots, 30 \quad \text{(efterfrågan)} \\
x_{ij} &\geq 0, \quad i = 1, \ldots, 10, j = 1, \ldots, 30 \\
y_i &\in \{0, 1\}, i = 1, \ldots, 10.
\end{aligned}$$

b) tillägg: $x_{ij} \in \{0, 1\}, i = 1, \ldots, 10; j = 1, \ldots, 30$.

4 a) x^* är ett lokalt minimum $\Leftrightarrow f(x^*) \leq f(x), \forall x \in B(x^*)$, där $B(x^*) = \{x \in \mathbb{R}^n | \|x - x^*\| \leq \epsilon\}$ för ett tillräckligt litet $\epsilon > 0$.

x^* är ett lokalt minimum $\Rightarrow \nabla f(x^*) = \mathbf{0}^n$.

b) x^* är ett lokalt minimum $\Leftrightarrow f(x^*) \leq f(x), \forall x \in B(x^*)nS$.

x^* är ett lokalt minimum $\Rightarrow \exists \lambda_* \geq \mathbf{0}^m$ så att

- $\nabla f(x^*) = A^T \lambda_*$.
- $\lambda_*^T (Ax^* - b) = 0$
- $Ax^* \geq b$.

5 a) Sätt $\phi(x) = -\sum_{i=1}^{m} \log(-g_i)(x)), \beta(x, \mu) = f(x) + \mu\phi(x)$. Låt $\mu_1, \mu_e, \ldots$ vara en positiv och monotont avtagande följd om tal med gränsvärde 0. Sekvensen $x_1, x_2, \ldots,$ ges av

$$x_k \in \arg\min_{x \in \mathbb{R}^n} \beta(x, \mu_k).$$

b) (Iterationsindex k struket här.) Från optimalitetsvillkoret $\nabla_x \beta(x, \mu) = o^n$ fås att

$$\nabla f(x) = \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) = 0^n. \text{ Eftersom } g_i(x) > 0, \forall i.$$

kan vi skriva detta som: $(\lambda_i = \mu / g_i(x), i = 1, \ldots, m)$

$$\begin{aligned} \nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = o^n, \\ \lambda_i \cdot g_i(x) = \mu, i = 1, \ldots, m \\ g_i(x) \geq 0, i = 1, \ldots, m. \end{aligned}$$

Skillnaden mellan detta och KKT för problemet (1)-(2) är att högerledet 0 i komplementariteten ersatts av $\mu > 0$.

Multiplikator estimat: $\lambda_i = \mu / g_i(x), i = 1, \ldots, m$..

c) Låt $I(x^*) = \{i | g_i(x^*) = 0\}$. x^* är reguljär betyder att $\nabla g_i(x^*), i \in I(x^*)$ är linjärt oberoende. Pga att f och g_i är kontinuerligt differentierbara så följer att $\{x_k\} \to x^* \Rightarrow \{\nabla f(x_k)\} \to \nabla f(x^*); \{\nabla g_i(x_k)\} \to \nabla g_i(x^*), \forall i$. Multiplikator estimatet ger att för $i \notin I(x^*) : \{\lambda_{ik}\} = \{\mu_k / g_i(x_k)\} \to 0/g_i(x^*) = 0$. För $i \in I(x^*)$, notera att systemet $\nabla f(x^*) - \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0^n$ har en unik lösning $\lambda_i^*, i \in I(x^*)$, vilken måste vara gränsvärdet för $\{\lambda_{ik}\}, i \in I(x^*)$. Ty antag att $\{\lambda_{ik}\}, i \in (x^*)$, konvergerar mot $\bar{\lambda}$, där $\bar{\lambda}_i \neq \lambda_i^*$ för något $i \in I(x^*)$. Då följer att

$$0 = \nabla f(x_k) = \sum_{i \notin j(x^*)} \lambda_{ik} \nabla g_i(x_k) - \sum_{i \in I(x^*)} \lambda_i^* \nabla g_i(x_k) - \sum_{i \in I(x_*)} [\lambda_{ik} - \lambda_i^*] \nabla g_i(x_m) \to$$
$$\to \nabla f(x^*) - 0 - \sum_{i \in j(x^*)} \lambda_i^* \nabla g_i(x^*) - \sum_{i \in j(x^*)} [\bar{\lambda}_i - \lambda_i^*] \nabla g_i(x^*) =$$
$$= - \sum_{i \in I(x^*)} [\bar{\lambda}_i - \lambda_i^*] \nabla g_i(x^*).$$

Denna summa är 0 endast om $\bar{\lambda}_i = \lambda_i^* \quad \forall i \in I(x^*)$, ty $\nabla g_i(x^*)$ är linjärt oberoende. Alltså är $\{\lambda_k\}$ konvergent, och vi kan sammanfatta läget så här: eftersom $g_i(x_k) > 0$ för alla i och k kommer $g_i(x^*) \geq 0$ att gälla. Dessutom för (x^*, λ^*) gäller att

$$\begin{aligned} \nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) &= 0^n \\ \lambda_i^* g_i(x^*) &= 0, i = 1, \ldots, m \\ \lambda_i^* \quad &\geq 0, i = 1, \ldots, m \end{aligned}$$

$\because$ I limes fås vektorer (x^*, λ^*) som tillsammans uppfyller KKT-villkoren för ursprungs problemet!

6 a) KKT: $L(x, \lambda, \mu) = \frac{1}{2} x^T Q x - q^T x + \lambda^T (Ax - b) - \mu^T x$ ger

- $\left[\begin{array}{r} Qx + \lambda^T \lambda - I\mu = q \\ \lambda, \mu \geq 0 \end{array} \right]$ (dual till.)
- $\left[\begin{array}{r} \lambda^T (Ax - b) = 0 \\ \mu^T x = 0 \end{array} \right]$ (kompl.)
- $\left[\begin{array}{r} Ax \leq b \\ x \geq 0 \end{array} \right]$ (primal till.)

KKT beskrives mängden av globalt optimala lösningar om Q är positivt semidefinit.

b) Inför en slackvariabel i $Ax \leq b$. Då fås ur KKT:

$$\left\{ \begin{array}{lr} Qx + A^T \lambda - I_n & = q \\ Ax & + Is = b \\ \quad x, \quad A, \mu, & s \geq 0 \end{array} \right] \text{ samt } \left\{ \begin{array}{l} \lambda^T s = 0 \\ \mu^T x = 0 \end{array} \right]$$

Vi indentifierar $v = \begin{bmatrix} x \\ s \end{bmatrix}; w = \begin{bmatrix} \mu \\ \lambda \end{bmatrix}; T = \begin{bmatrix} Q & 0 \\ A & I \end{bmatrix}; U = \begin{bmatrix} A^T & -I \\ 0 & 0 \end{bmatrix}; p = \begin{bmatrix} q \\ b \end{bmatrix}$.

Fas-1-metodP: Inför artificiella variabler $z^1 \in \mathbb{R}^n$ och z^2 i $\mathbb{R}^m$ och betrakta problemet (multipliera först rader så att $q, b \geq \mathbf{0}$!)

$$\text{minimera} \quad \sum_{j=1}^{n} z_j^1 + \sum_{i=1}^{m} z_i^2$$

$$\text{då} \left\{ \begin{array}{l} Qx + A^T\lambda - I_m u \qquad + z^1 = q \\ Ax \qquad\qquad\qquad + Is + z^2 = b \\ \quad , \quad \lambda, \quad m, \quad s, z^1, z^2 \geq 0 \\ \lambda^T s = 0, \mu^T x = 0 \end{array} \right]$$

Enda skillnaden mot ett Fas-1-problem i LP är kraven att $\lambda^T s = 0$ och $\mu^T x = 0$. Det betyder i själva verket att $\lambda_i \cdot s_i = 0 \quad \forall i$ och $\mu_j \cdot x_j = 0 \quad \forall j$. Att säkerställa att detta är uppfyllt är inte svårt: vi inför ett tillägg till inkommande kriteriet:

- Om $\lambda_i(s_i)$ redan finns i basen, får inte $s_i(\lambda_i)$ vara inkommande, såvida det inte inträffar att $s_i(\lambda_i)$ blir den utgående variabeln.
- Motsvarande för paret (μ_j, x_j).

7 a) $f(x) = 2x_1^3 + 2x_2^2 + 12x_1x_2 - 24x_1 - 20x_2 \Rightarrow$

$\nabla f(x) = (6x_1^2 + 12x_2 - 24; 6x_2^2 + 12x_1 - 20)^T$

$\nabla^2 f(x) = \begin{bmatrix} 12x_1 & 12 \\ 12 & 12x_2 \end{bmatrix} = 12 \begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix}$

Egenvärden hos $\begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix}$: $\det \begin{bmatrix} x_1 - \lambda & 1 \\ 1 & x_2 - \lambda \end{bmatrix} = 0 \Rightarrow \ldots \Rightarrow$

Egenvärden: $\begin{bmatrix} \lambda_1 = (x_1 + x_2)/2 + \sqrt{(\frac{1}{2}x_1 - \frac{1}{2}x_2)^2 + 1} \\ \lambda_2 = (x_1 + x_2)/2 - \sqrt{(\frac{1}{2}x_1 - \frac{1}{2}x_2)^2 + 1} \end{bmatrix} \lambda_1 \geq \lambda_1!$

$\lambda_2 \geq 0$ om och endast om $x_1, x_2 \geq 0$ och $x_1x_2 \geq 1$.
$\because f$ är konvex då $x_1, x_2 \geq 0, x_1x_2 \geq 1$

$\lambda_1 \leq 0$ om och endast om $x_1, x_2 \leq 0$ och $x_1x_2 \geq 1$. $\because f$ är konkav då $x_1, x_2 \geq 0, x_1x_2 \geq 1$

$\because f$ är följaktligen varken konvex eller konkav då $x_1x_2 < 1$.

b) $\bar{x} = (1,1)^T$. Newtons metod utnyttjar sökriktningen p från $\nabla^2 f(\bar{x})p = -\nabla f(\bar{x})$. $I\bar{x} = (1,1)^T$ är $\nabla f(\bar{x}) = (-6, -2); \nabla^2 f(\bar{x}) = \begin{bmatrix} 12 & 12 \\ 12 & 12 \end{bmatrix}$. Det existerar inte något p som uppfyller $\nabla^2 f(\bar{x})p = -\nabla f(\bar{x})$! Modifiering à lá Levenberg-Marquardt: addera en lämpligt skalad enhetsmatris till $\nabla^2 f(\bar{x})$. Lös t.ex.

$$[\nabla^2 f(\bar{x}) + \sigma I]p = -\nabla f(\bar{x}) \text{ för } \sigma = 10.$$

$$\begin{bmatrix} 22 & 12 \\ 12 & 22 \end{bmatrix} p = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \Leftrightarrow p \approx (0.32, -0.08)$$

Vi använder Armijos steglägdsregel:

$$f(\bar{x} + \ell p) \leq f(\bar{x}) + \alpha \ell \nabla f(\bar{x})^T p, \alpha \in (0, 1).$$

Med $\alpha = 0.1$ och $\ell = 1$ som startsteglägd fås:

$\nabla f(\bar{x})^T p \approx -1.74 < 0$
$f(\bar{x}) = -28; \bar{x} + \ell p \approx (1.32, 0.92); f(\bar{x} + \ell p) \approx -29.35$
$f(\bar{x}) + \alpha \ell \nabla f(\bar{x})^T p \approx -28.17 \geq -29.35$, d.v.s. steglägd 1 accepteras av Armijos steglägdsregel.

Nästa iterationspunkt är $x = (1.32, 0.92)$.