Lösningar till 5B1762 Optimeringslära för T, 20/5-06

Uppgift 1.(a)

Inför följande variabler:

x_1 = antal enheter av produkt A som tillverkas per dag.
x_2 = antal enheter av produkt B som tillverkas per dag.
x_3 = antal enheter av produkt C som tillverkas per dag.

Täckningsbidraget per dag ges då av $12x_1 + 9x_2 + 8x_3$.

Kapacitetsbegränsningen i stansavdelningen kan skrivas $\dfrac{x_1}{2000} + \dfrac{x_2}{1600} + \dfrac{x_3}{1100} \leq 8$.

Kapacitetsbegränsningen i pressavdelningen kan skrivas $\dfrac{x_1}{1000} + \dfrac{x_2}{1500} + \dfrac{x_3}{2400} \leq 8$.

Det ger oss följande problemformulering:

$$\begin{array}{rl} \text{maximera} & 12x_1 + 9x_2 + 8x_3 \\ \text{då} & \dfrac{x_1}{2000} + \dfrac{x_2}{1600} + \dfrac{x_3}{1100} \leq 8, \\ & \dfrac{x_1}{1000} + \dfrac{x_2}{1500} + \dfrac{x_3}{2400} \leq 8, \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0. \end{array}$$

Uppgift 1.(b)

Svarande mot den givna tillåtna baslösningen får vi följande simplexmultiplikatorer u_i och v_j, beräknade mha relationen $c_{ij} = u_i - v_j$ för basvariabler samt $v_4 = 0$.

c_{ij}	kund1	kund2	kund3	kund4	u_i
lev 1	16	25	36		47
lev 2			25	36	36
lev 3				25	25
lev 4				16	16
v_j	31	22	11	0	

Därefter får vi följande reducerade kostnader r_{ij} för ickebasvariablerna, uträknade mha relationen $r_{ij} = c_{ij} - u_i + v_j$.

r_{ij}	kund1	kund2	kund3	kund4	u_i
lev 1				2	47
lev 2	4	2			36
lev 3	10	6	2		25
lev 4	16	10	4		16
v_j	31	22	11	0	

Eftersom alla $r_{ij} \geq 0$ så är den föreslagna baslösningen optimal.

Uppgift 2.(a)

Vi har här ett LP-problem på standardformen

$$\begin{aligned} \text{minimera} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

där $\mathbf{A} = \begin{bmatrix} 4 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$ och $\mathbf{c}^\mathsf{T} = (4,\ 4,\ 2,\ 4,\ 4)$.

Den givna startlösningen ska ha basvariablerna x_1 och x_5, vilket innebär att $\beta = (1, 5)$ och $\delta = (2, 3, 4)$.

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$, medan $\mathbf{A}_\delta = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{bmatrix}$.

Basvariablernas värden i baslösningen ges av $\mathbf{x}_\beta = \bar{\mathbf{b}}$, där vektorn $\bar{\mathbf{b}}$ beräknas ur ekvationssystemet $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

dvs $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1.25 \\ 0.75 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna värden erhålls ur systemet $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

dvs $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (4,\ 2,\ 4) - (1,\ 1) \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{bmatrix} = (0,\ -2,\ 0)$.

Eftersom $r_{\delta_2} = r_3 = -2$ är minst, och < 0, ska vi låta x_3 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_3$ ur systemet $\mathbf{A}_\beta \bar{\mathbf{a}}_3 = \mathbf{a}_3$,

dvs $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_3 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{1.25}{0.5}, \frac{0.75}{0.5} \right\} = \frac{0.75}{0.5} = \frac{\bar{b}_2}{\bar{a}_{23}}.$$

Minimerande index är $i = 2$, varför $x_{\beta_2} = x_5$ inte längre får vara kvar som basvariabel. Dess plats tas av x_3.

Nu är alltså $\beta = (1, 3)$ och $\delta = (2, 5, 4)$.

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}$, medan $\mathbf{A}_\delta = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 4 & 3 \end{bmatrix}$.

Basvariablernas värden i baslösningen ges av $\mathbf{x}_\beta = \bar{\mathbf{b}}$, där vektorn $\bar{\mathbf{b}}$ beräknas ur ekvationssystemet $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

dvs $\begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$.

Vektorn $\mathbf{y}$ med simplexmultiplikatorerna värden erhålls ur systemet $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,
dvs $\begin{bmatrix} 4 & 0 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Reducerade kostnaderna för icke-basvariablerna ges av
$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (4,\ 4,\ 4) - (1,\ 0) \begin{bmatrix} 3 & 0 & 1 \\ 1 & 4 & 3 \end{bmatrix} = (1,\ 4,\ 3)$.

Eftersom $\mathbf{r}_\delta \geq \mathbf{0}$ så är den aktuella baslösningen optimal.

Därmed är punkten $x_1 = 0.5$, $x_2 = 0$, $x_3 = 1.5$, $x_4 = 0$, $x_5 = 0$ optimal.

Optimalvärdet är $\mathbf{c}^\mathsf{T}\mathbf{x} = 5$.

Uppgift 2.(b)

Om primala problemet är på standardformen

$$\begin{aligned} \text{minimera} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

så är det duala problemet på formen: maximera $\mathbf{b}^\mathsf{T}\mathbf{y}$ då $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$, som utskrivet blir:

$$\begin{array}{rrcrcl} \text{maximera} & 5y_1 & + & 3y_2 & & \\ \text{då} & 4y_1 & & & \leq & 4, \\ & 3y_1 & + & y_2 & \leq & 4, \\ & 2y_1 & + & 2y_2 & \leq & 2, \\ & y_1 & + & 3y_2 & \leq & 4, \\ & & & 4y_2 & \leq & 4. \end{array}$$

Det är välkänt att en optimal lösning till detta duala problem ges av vektorn $\mathbf{y}$ med "simplexmultiplikatorerna" i optimala baslösningen i (a)-uppgiften, dvs $\mathbf{y} = (1,\ 0)^\mathsf{T}$. Man kontrollerar snabbt att detta är en tillåten lösning till det duala problemet ovan. Vidare är duala målfunktionsvärdet $\mathbf{b}^\mathsf{T}\mathbf{y} = 5y_1 + 3y_2 = 5 - 0 = 5 =$ optimalvärdet för det primala problemet i (a)-uppgiften.

Uppgift 3.(a)

Vi har ett kvadratiska optimeringsproblem med linjära likhetsbivillkor på formen

$$\begin{array}{rl} \text{minimera} & \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \end{array}$$

där $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ och $\mathbf{b} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Matrisen $\mathbf{H} = \mathbf{I}$ är positivt definit, så vi har ett *konvext* QP-problem.

Optimalitetsvillkoren ges då av $\mathbf{H}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{u} = -\mathbf{c}$ och $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Eftersom $\mathbf{H} = \mathbf{I}$ så ger de första ekvationerna att $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u} - \mathbf{c}$,

som insatt i $\mathbf{A}\mathbf{x} = \mathbf{b}$ ger ekvationssystemet $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{b} + \mathbf{A}\mathbf{c}$,

dvs $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, med lösningen $\hat{\mathbf{u}} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Optimala lösningen $\hat{\mathbf{x}}$ ges sedan av $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{u}} - \mathbf{c} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$.

Uppgift 3.(b)

Vi har ett kvadratiska optimeringsproblem med linjära olikhetsbivillkor på formen

$$\begin{array}{rl} \text{minimera} & \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \end{array}$$

där $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ och $\mathbf{b} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

I den givna startpunkten $\bar{\mathbf{x}} = (1,\ 2,\ 0)^\mathsf{T}$ är bägge bivillkoret uppfyllda med likhet.
Därför startar vi med $\alpha = (1, 2)$ och γ tom.

Då är $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ och $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, och därmed $\mathbf{A}_\alpha^\mathsf{T} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Vi får svaret "JA" i Steg 1, ty $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_\alpha^\mathsf{T}\bar{\mathbf{u}}$ med $\bar{\mathbf{u}} = (1,\ -1)^\mathsf{T}$, så vi går till Steg 2.

Här konstateras att $\bar{u}_2 < 0$ (och minst), varför $\alpha_2 = 2$ flyttas över till γ–vektorn.

Sedan går vi till Steg 3 med $\alpha = (1)$, $\gamma = (2)$ och $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$.

I Steg 3 ska vi minimera $\frac{1}{2}\,\mathbf{d}^\mathsf{T}\mathbf{H}\mathbf{d} + (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})^\mathsf{T}\mathbf{d}$ under bivillkoret $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$,

Optimalitetsvillkoren för detta konvexa QP-problem med likhetsbivillkor ges av

$\mathbf{H}\mathbf{d} - \mathbf{A}_\alpha^\mathsf{T}\mathbf{u} = -(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ och $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$.

Eftersom $\mathbf{H} = \mathbf{I}$ så ger de första ekvationerna att $\mathbf{d} = \mathbf{A}_\alpha^\mathsf{T}\mathbf{u} - (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$,

som insatt i $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$ ger ekvationssystemet $\mathbf{A}_\alpha\mathbf{A}_\alpha^\mathsf{T}\mathbf{u} = \mathbf{A}_\alpha(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, dvs $2\mathbf{u} = 1$.

Lösningen till detta "ekvationssystem" är $\hat{\mathbf{u}} = 0.5$, varefter

$$\hat{\mathbf{d}} = \mathbf{A}_\alpha^\mathsf{T}\hat{\mathbf{u}} - (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} (\,0.5\,) - \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ -0.5 \\ 1 \end{pmatrix}.$$

Eftersom $\bar{\mathbf{x}} + \hat{\mathbf{d}} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.5 \\ -0.5 \\ 1.0 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.0 \end{pmatrix}$ uppfyller alla bivillkor låter vi

denna punkt bli nästa iterationspunkt $\bar{\mathbf{x}}$ och går till Steg 1.

Nu är $\alpha = (1)$, $\gamma = (2)$, $\bar{\mathbf{x}} = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.0 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \end{pmatrix}$ och $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$.

Vi får svaret "JA" i Steg 1, ty $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_\alpha^\mathsf{T}\bar{\mathbf{u}}$ med $\bar{\mathbf{u}} = 0.5$, så vi går till Steg 2.

Här konstateras att $\bar{\mathbf{u}} \geq \mathbf{0}$, varför den aktuella iterationspunkten

$\bar{\mathbf{x}} = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.0 \end{pmatrix}$, tillsammans med vektorn $\hat{\mathbf{y}} = \begin{pmatrix} 0.5 \\ 0 \end{pmatrix}$, uppfyller optimalitetsvillkoren,

dvs $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}^\mathsf{T}\hat{\mathbf{y}}$, $\mathbf{A}\bar{\mathbf{x}} \geq \mathbf{b}$, $\hat{\mathbf{y}} \geq \mathbf{0}$ och $\hat{\mathbf{y}}^\mathsf{T}(\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}) = 0$. Vi stannar därför här.

Uppgift 4.(a)

Vi ska starta i $\mathbf{x}^{(1)} = (0,\, 0)^{\mathsf{T}}$. Då är $\mathbf{h}(\mathbf{x}^{(1)}) = (-2,\, 1, -1,\, 0)^{\mathsf{T}}$ och $f(\mathbf{x}^{(1)}) = 3$.

Vidare är exempelvis $\dfrac{\partial h_3}{\partial x_2}(\mathbf{x}) = 2\,(x_2 - 1)$, så att $\dfrac{\partial h_3}{\partial x_2}(\mathbf{x}^{(1)}) = -2$.

Motsvarande kalkyler för de övriga derivatorna ger att $\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 2 & -2 \\ -2 & 2 \end{bmatrix}$.

I Gauss-Newtons metod ska man lösa ekvationssystemet

$\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\mathbf{h}(\mathbf{x}^{(1)})$

I vårt fall är $\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 16 & 0 \\ 0 & 16 \end{bmatrix}$ och $\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -8 \\ -4 \end{pmatrix}$,

så ekvationssystemet blir $\begin{bmatrix} 16 & 0 \\ 0 & 16 \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}$.

Vi prövar $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}$.

Då blir

$h_1(\mathbf{x}^{(2)}) = 9/4 + 25/16 - 4 = (36 + 25 - 64)/16 = -3/16,$
$h_2(\mathbf{x}^{(2)}) = 1/4 + \;9/16 - 1 = (4 + 9 - 16)/16 = -3/16,$
$h_3(\mathbf{x}^{(2)}) = 9/4 + \;9/16 - 3 = (36 + 9 - 48)/16 = -3/16,$
$h_4(\mathbf{x}^{(2)}) = 1/4 + 25/16 - 2 = (4 + 25 - 32)/16 = -3/16,$

$f(\mathbf{x}^{(2)}) = 9/128 < f(\mathbf{x}^{(1)})$ (med bred marginal) så steget $t_1 = 1$ gick bra.

Därmed har vi utfört en fullständig iteration med Gauss-Newtons metod.

Uppgift 4.(b)

Vid Newtons metod, utgående från samma startpunkt $\mathbf{x}^{(1)} = (0,\, 0)^{\mathsf{T}}$ som ovan,

ska man i stället lösa ekvationssystemet $\nabla^2 f(\mathbf{x}^{(1)})\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^{\mathsf{T}}$.

För ickelinjära minsta-kvadratproblem kan detta system skrivas

$$\left(\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}^{(1)}) + \sum_{i=1}^{m} h_i(\mathbf{x}^{(1)})\mathbf{H}_i(\mathbf{x}^{(1)})\right)\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\mathbf{h}(\mathbf{x}^{(1)}).$$

I vårt fall är $\mathbf{H}_i(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ för $i = 1, 2, 3, 4$, oberoende av $\mathbf{x}$.

Därmed är $\displaystyle\sum_{i=1}^{m} h_i(\mathbf{x}^{(1)})\mathbf{H}_i(\mathbf{x}^{(1)}) = \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix}$, så att

$$\nabla^2 f(\mathbf{x}^{(1)}) = \nabla\mathbf{h}(\mathbf{x}^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(\mathbf{x}^{(1)}) + \sum_{i=1}^{m} h_i(\mathbf{x}^{(1)})\mathbf{H}_i(\mathbf{x}^{(1)}) = \begin{bmatrix} 12 & 0 \\ 0 & 12 \end{bmatrix},$$

som är positivt definit.

Ekvationssystemet som ska lösas kan därmed skrivas

$\begin{bmatrix} 12 & 0 \\ 0 & 12 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$.

Vi prövar $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1 \mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$.
Då blir

$h_1(\mathbf{x}^{(2)}) = 25/9 + 16/9 - 4 = (25 + 16 - 36)/9 = 5/9,$
$h_2(\mathbf{x}^{(2)}) = 1/9 + 4/9 - 1 = (1 + 4 - 9)/9 = -4/9,$
$h_3(\mathbf{x}^{(2)}) = 25/9 + 4/9 - 3 = (25 + 4 - 27)/9 = 2/9,$
$h_4(\mathbf{x}^{(2)}) = 1/9 + 16/9 - 2 = (1 + 16 - 18)/9 = -1/9,$

$f(\mathbf{x}^{(2)}) = 23/81 < f(\mathbf{x}^{(1)})$ (med bred marginal) så steget $t_1 = 1$ gick bra.

Därmed har vi utfört en fullständig iteration med Newtons metod.

Man kan konstatera att på detta problem är resultatet (mätt i målfunktionsvärde) efter den första iterationen bättre med Gauss-Newtons metod än med Newtons metod, trots att G-N är enklare än N. Det är dock inte alltid så.

Uppgift 5.

Låt $f(\mathbf{x}) = |\,\mathbf{x}-\mathbf{p}\,|^2 = (\mathbf{x}-\mathbf{p})^\mathsf{T}(\mathbf{x}-\mathbf{p}) = \mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{p}^\mathsf{T}\mathbf{x} + 1$,

och $g(\mathbf{x}) = |\,\mathbf{x}-\mathbf{q}\,|^2 - 1 = (\mathbf{x}-\mathbf{q})^\mathsf{T}(\mathbf{x}-\mathbf{q}) - 1 = \mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{q}^\mathsf{T}\mathbf{x}$.

Då kan vårt givna problem skrivas: minimera $f(\mathbf{x})$ då $g(\mathbf{x}) \leq 0$.

Gradienterna ges av $\nabla f(\mathbf{x}) = 2\,\mathbf{x}^\mathsf{T} - 2\,\mathbf{p}^\mathsf{T}$ och $\nabla g(\mathbf{x}) = 2\,\mathbf{x}^\mathsf{T} - 2\,\mathbf{q}^\mathsf{T}$.

Vidare ges andraderivatsmatriserna av $\nabla^2 f(\mathbf{x}) = 2\,\mathbf{I}$ och $\nabla^2 g(\mathbf{x}) = 2\,\mathbf{I}$.

Eftersom båda dessa är positivt definita så är både f och g konvexa funktioner. Vårt optimeringsproblem är alltså ett konvext problem. Bra!

KKT-villkoren för detta konvexa problem med enbart ett bivillkor är:

KKT-1: $\nabla f(\mathbf{x}) + y\nabla g(\mathbf{x}) = \mathbf{0}^\mathsf{T}$,

KKT-2: $g(\mathbf{x}) \leq 0$,

KKT-3: $y \geq 0$,

KKT-4: $y\,g(\mathbf{x}) = 0$.

som i vårt speciella fall kan skrivas

KKT-1: $\mathbf{x} - \mathbf{p} + y\cdot(\mathbf{x}-\mathbf{q}) = \mathbf{0}$,

KKT-2: $\mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{q}^\mathsf{T}\mathbf{x} \leq 0$,

KKT-3: $y \geq 0$,

KKT-4: $y\cdot(\mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{q}^\mathsf{T}\mathbf{x}) = 0$.

Vi ska undersöka två fall: $y = 0$ och $y > 0$. (Fallet $y < 0$ utesluts av KKT-3).

Antag först att $y = 0$.

Då ger KKT-1 att $\mathbf{x} = \mathbf{p}$, som insatt i vänsterledet i KKT-2 ger att $\mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{q}^\mathsf{T}\mathbf{x} = \mathbf{p}^\mathsf{T}\mathbf{p} - 2\,\mathbf{q}^\mathsf{T}\mathbf{p} = 1 - 0 = 1$, vilket bryter mot olikheten i KKT-2. Alltså finns det ingen lösning till KKT-villkoren med $y = 0$.

Antag fortsättningsvis att $y > 0$.

Då ger KKT-4 att $\mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{q}^\mathsf{T}\mathbf{x} = 0$, medan KKT-1 ger att $\mathbf{x} = \dfrac{\mathbf{p}+\mathbf{q}\,y}{1+y}$.

Tillsammans ger dessa relationer att

$$0 = \mathbf{x}^\mathsf{T}\mathbf{x} - 2\,\mathbf{q}^\mathsf{T}\mathbf{x} = \frac{(\mathbf{p}+\mathbf{q}\,y)^\mathsf{T}(\mathbf{p}+\mathbf{q}\,y)}{(1+y)^2} - \frac{2\,\mathbf{q}^\mathsf{T}(\mathbf{p}+\mathbf{q}\,y)}{1+y} = \frac{1+y^2}{(1+y)^2} - \frac{2\,y}{1+y} = \frac{1-2y-y^2}{(1+y)^2}.$$

Vi får alltså ekvationen $y^2 + 2y - 1 = 0$, med lösningarna $y = -1+\sqrt{2}$ och $y = -1-\sqrt{2}$.

Men den andra av dessa är negativ, så den enda återstående lösningen är $\hat{y} = \sqrt{2} - 1$.

Motsvarande $\mathbf{x}$ ges av $\hat{\mathbf{x}} = \dfrac{\mathbf{p}+\mathbf{q}\,\hat{y}}{1+\hat{y}} = \dfrac{1}{\sqrt{2}}\,(\mathbf{p} + \mathbf{q}(\sqrt{2}-1)) = \mathbf{q} + \dfrac{1}{\sqrt{2}}\cdot(\mathbf{p}-\mathbf{q})$.

Då uppfyller $\hat{\mathbf{x}}$ och $\hat{y}$ KKT-villkoren ovan. Eftersom såväl f som g är konvexa funktioner så betyder det att $\hat{\mathbf{x}}$ är en globalt optimal lösning till det betraktade problemet.