Lösningar till SF1861/SF1851 Optimeringslära, 17/8-12

Uppgift 1.(a) Flödesbalansvillkoren i de fem noderna kan skrivas $\mathbf{Ax} = \mathbf{b}$, där

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 35 \\ 40 \\ -20 \\ -30 \\ -25 \end{pmatrix}.$$

Variabelvektorn ges av $\mathbf{x} = \begin{pmatrix} x_{13} \\ x_{14} \\ x_{15} \\ x_{23} \\ x_{24} \\ x_{25} \end{pmatrix}$, där variabeln x_{ij} betecknar flödet i bågen (i, j).

Att bågarna är (enkel-)riktade ger kravet att $\mathbf{x} \geq \mathbf{0}$.

Flödets kostnad (som ska minimeras) ges av $\mathbf{c}^\mathsf{T}\mathbf{x}$, där $\mathbf{c} = \begin{pmatrix} c_{13} \\ c_{14} \\ c_{15} \\ c_{23} \\ c_{24} \\ c_{25} \end{pmatrix} = \begin{pmatrix} 19 \\ 14 \\ 18 \\ 23 \\ 16 \\ 21 \end{pmatrix}$.

Om det primala problemet är på formen: minimera $\mathbf{c}^\mathsf{T}\mathbf{x}$ då $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$,
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 här blir:

$$\begin{aligned} \text{maximera} \quad & 35y_1 + 40y_2 - 20y_3 - 30y_4 - 25y_5 \\ \text{då} \quad & y_1 - y_3 \leq 19, \\ & y_1 - y_4 \leq 14, \\ & y_1 - y_5 \leq 18, \\ & y_2 - y_3 \leq 23, \\ & y_2 - y_4 \leq 16, \\ & y_2 - y_5 \leq 21. \end{aligned}$$

Uppgift 1.(b) Endast svaren ges här:

minimera $3x_1 - 4x_2$ då $-1 \leq x_1 \leq 1$, $-1 \leq x_2 \leq 1$, har optimal lösning $x_1 = -1$, $x_2 = 1$.

minimera $3x_1 - 4x_2$ då $x_1 + x_2 \leq 1$, $x_1 \geq -1$, $x_2 \geq -1$, har optlösning $x_1 = -1$, $x_2 = 2$.

minimera $3x_1 - 4x_2$ då $x_1^2 + x_2^2 \leq 1$, har optimal lösning $x_1 = -0.6$, $x_2 = 0.8$.

minimera $3x_1 - 4x_2$ då $|x_1| + |x_2| \leq 1$, har optimal lösning $x_1 = 0$, $x_2 = 1$.

Uppgift 2.(a) Problemet är på standardformen

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

där $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ och $\mathbf{c} = (2,\ 3,\ 6,\ 4)^\mathsf{T}$.

I startlösningen ska enligt uppgiftslydelsen x_1, x_2 och x_3 vara basvariabler, dvs $\beta = (1, 2, 3)$ och $\nu = (4)$.

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

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 $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

(Eftersom $\mathbf{A}_\beta$ är en triangulär matris såg vi lösningen utan några beräkningar.)

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

dvs $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$.

(Eftersom $\mathbf{A}_\beta^\mathsf{T}$ är en triangulär matris såg vi lösningen utan några beräkningar.)

Reducerade kostnaderna för den enda icke-basvariabeln ges av

$$r_4^\mathsf{T} = c_4 - \mathbf{y}^\mathsf{T}\mathbf{a}_4 = 4 - (2,\ 1,\ 5) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -1.$$

Eftersom $r_4 = -1$ är < 0 ska vi låta x_4 bli ny basvariabel.

Då behöver vi beräkna vektorn $\bar{\mathbf{a}}_4$ ur systemet $\mathbf{A}_\beta \bar{\mathbf{a}}_4 = \mathbf{a}_4$,

dvs $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \\ \bar{a}_{34} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, med lösningen $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \\ \bar{a}_{34} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

Det största värde som den nya basvariabeln x_4 kan ökas till ges av

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i4}} \mid \bar{a}_{i4} > 0 \right\} = \min \left\{ \frac{1}{1},\ -\ ,\ \frac{1}{1} \right\}.$$

Minimerande index är $i = 1$ eller $i = 3$, varför antingen $x_{\beta_1} = x_1$ eller $x_{\beta_3} = x_3$ ska lämna basen. Vi väljer (efter slantsingling) att låta x_3 lämna basen.

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

Motsvarande basmatris ges av $\mathbf{A}_\beta = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

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 $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, med lösningen $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

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

dvs $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$, med lösningen $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$.

Reducerade kostnaderna för den enda icke-basvariabeln ges av

$r_3^{\mathsf{T}} = c_3 - \mathbf{y}^{\mathsf{T}} \mathbf{a}_3 = 6 - (2,\ 1,\ 4) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 1.$

Eftersom $r_3 \geq 0$ är den aktuella baslösningen optimal. (Alla reducerade kostnader är ≥ 0.)
Alltså är punkten $\hat{\mathbf{x}} = (0, 1, 0, 1)^{\mathsf{T}}$ en optimal lösning till problemet.
Optimalvärdet är $\mathbf{c}^{\mathsf{T}} \hat{\mathbf{x}} = 7$.

Uppgift 2.(b) Eftersom reducerade kostnaden för den enda ickebasvaribeln x_3 är strikt större än noll så är varje annan tillåten lösning till problemet strikt sämre (med strikt högre målfunktionsvärde) än den erhållna optimallösningen i (a)-uppgiften ovan. $\hat{\mathbf{x}}$ är alltså den enda optimala lösningen.

Uppgift 2.(c) 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

$$\text{maximera} \ \ \mathbf{b}^{\mathsf{T}} \mathbf{y} \ \ \text{då} \ \ \mathbf{A}^{\mathsf{T}} \mathbf{y} \leq \mathbf{c},$$

som utskrivet blir:

$$\begin{aligned} \text{maximera} \quad & y_1 + y_2 + y_3 \\ \text{då} \quad & y_1 \leq 2, \\ & y_1 + y_2 \leq 3, \\ & y_2 + y_3 \leq 6, \\ & y_3 \leq 4. \end{aligned}$$

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 $\mathbf{y} = (2,\ 1,\ 4)^{\mathsf{T}}$.
Man kontrollerar snabbt att detta är en tillåten lösning till det duala problemet.
Vidare är optimalvärdet $\mathbf{b}^{\mathsf{T}} \mathbf{y} = 7 =$ optimalvärdet för det primala problemet ovan.

Uppgift 2.(d)
Vektorn $\mathbf{y} = (y_1, y_2, y_3)^\mathsf{T}$ är en optimal lösning till det duala problemet ovan om och endast om $\mathbf{y}$ är en tillåten lösning till duala problemet med målfunktionsvärdet 7, dvs om och endast om $\mathbf{y}$ uppfyller följande fem villkor:

(1) $y_1 + y_2 + y_3 = 7,$
(2) $y_1 \leq 2,$
(3) $y_1 + y_2 \leq 3,$
(4) $y_2 + y_3 \leq 6$
(5) $y_3 \leq 4.$

Villkoren (1) och (3) medför att $y_3 = 7 - (y_1 + y_2) \geq 7 - 3 = 4,$
vilket tillsammans med (5) medför att

(6) $y_3 = 4.$

Villkoren (1) och (6) medför att

(7) $y_1 + y_2 = 3,$

medan villkoren (4) och (6) medför att

(8) $y_2 \leq 2.$

Å andra sidan, om $\mathbf{y}$ uppfyller villkoren (2), (6), (7) och (8), så uppfyller $\mathbf{y}$ samtliga de ursprungliga villkoren (1)–(5).
Det betyder att $\mathbf{y}$ är en optimal lösning till duala problemet om och endast om $\mathbf{y}$ uppfyller följande fyra villkor:

(2) $y_1 \leq 2,$
(8) $y_2 \leq 2,$
(7) $y_1 + y_2 = 3,$
(6) $y_3 = 4.$

Vektorn $\mathbf{y}$ är en lösning till dessa villkor om och endast om
$\mathbf{y} = (1{+}t,\ 2{-}t,\ 4)^\mathsf{T}$, där $t \in [0, 1]$.

Exempel: $\mathbf{y} = (2, 1, 4)^\mathsf{T}$, $\mathbf{y} = (1, 2, 4)^\mathsf{T}$, $\mathbf{y} = (1.3,\ 1.7,\ 4)^\mathsf{T} \ldots$

Det finns alltså oändligt många optimala lösningar till det duala problemet.

Uppgift 3.(a)

Problemet kan skrivas på formen: minimera $\frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ då $\mathbf{A}\mathbf{x} = \mathbf{b}$,

där $\mathbf{H} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Gauss-Jordans metod tillämpad på systemet $\mathbf{A}\mathbf{x} = \mathbf{b}$ ger med några enkla radoperationer

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

ur vilket fram går att allmänna lösningen till $\mathbf{A}\mathbf{x} = \mathbf{b}$ erhålls genom att sätta den enda icke-trappstegsvariablen x_4 till ett godtyckligt tal v varefter trappstegsvariablernas värden ges av $x_1 = 1 - v$, $x_2 = v$ och $x_3 = 1 - v$.

Detta kan sammanfattas som att allmänna lösningen till $\mathbf{A}\mathbf{x} = \mathbf{b}$ ges av

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} v = \bar{\mathbf{x}} + \mathbf{z}\, v,$$

där $\bar{\mathbf{x}}$ är en tillåten lösning till $\mathbf{A}\mathbf{x} = \mathbf{b}$ medan vektorn $\mathbf{z}$ är en bas för nollrummet till $\mathbf{A}$.

Variabelbytet från $\mathbf{x}$ till den enda variabeln v leder till en kvadratisk envariabelfunktion vars unika minpunkt erhålls genom att lösa den linjära ekvationen

$$(\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z})\, v = -\mathbf{z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}),$$

förutsatt att $\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z}$ är positivt definit (dvs > 0 i detta envariabelfall).

Vi får att $\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z} = 16 > 0$ och $-\mathbf{z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = 4$,

så den unika lösningen till ekvationen ovan är $\hat{v} = 4/16 = 0.25$,

och den unika optimala lösningen till problemet ges av $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = \begin{pmatrix} 0.75 \\ 0.25 \\ 0.75 \\ 0.25 \end{pmatrix}$.

Uppgift 3.(b)

Lagrangevillkoren till problemet att minimera $\frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ då $\mathbf{b} - \mathbf{A}\mathbf{x} = \mathbf{0}$

ges av ekvationerna $\begin{array}{lcr} \mathbf{H}\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{u} & = & -\mathbf{c} \\ \mathbf{A}\mathbf{x} & = & \mathbf{b} \end{array}$.

Vi ska här ge två alternativa tillvägagångssätt för att hitta en lösning $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ till detta ekvationssystem. Det första alternativet förutsätter inte att man löst (a)-uppgiften, medan det andra förslaget förutsätter just detta.

Lösningsalternativ 1:

Eftersom $\mathbf{H}$ är diagonal med strikt positiva diagonalelement kan man enkelt beräkna $\mathbf{H}^{-1}$, och då är $\mathbf{H}\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{u} = -\mathbf{c}$ ekvivalent med att $\mathbf{x} = \mathbf{H}^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{u} - \mathbf{H}^{-1}\mathbf{c} = \mathbf{H}^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{u}$, ty $\mathbf{c} = \mathbf{0}$.

Om man sätter in detta uttryck för $\mathbf{x}$ i de återstående ekvationerna $\mathbf{A}\mathbf{x} = \mathbf{b}$ erhålls ekvationssystemet $\mathbf{A}\mathbf{H}^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{b}$, som utskrivet blir

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/6 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Multiplikation av bägge leden med 6, samt vanliga matrismultiplikationer, ger att detta

system kan skrivas $\begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 3 \\ 0 & 3 & 4 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix}$, med den unika lösningen $\hat{\mathbf{u}} = \begin{pmatrix} 1.5 \\ 0 \\ 1.5 \end{pmatrix}$.

Motsvarande $\mathbf{x}$–del av lösningen ges sedan av $\hat{\mathbf{x}} = \mathbf{H}^{-1}\mathbf{A}^{\mathsf{T}}\hat{\mathbf{u}} = \begin{pmatrix} 0.75 \\ 0.25 \\ 0.75 \\ 0.25 \end{pmatrix}$,

dvs samma $\hat{\mathbf{x}}$ som i (a)-uppgiften ovan.

Nu är $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ en lösning till Lagrangevillkoren. (Som framgår av räkningarna ovan är det i själva verket den enda lösningen till dessa optimalitetsvillkor.)

Lösningsalternativ 2:

Vårt aktuella QP-problem har konvex målfunktion (ty $\mathbf{H}$ är positivt definit) och linjära likhetsbivillkor. Då är $\hat{\mathbf{x}} \in I\!R^4$ en optimal lösning till problemet om och endast om $\hat{\mathbf{x}}$ utgör $\mathbf{x}$–delen av en lösning $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ till Lagrangevillkoren.

Men enligt (a)-uppgiften är $\hat{\mathbf{x}} = (0.75,\, 0.25,\, 0.75,\, 0.25)^{\mathsf{T}}$ den enda optimala lösningen till problemet. Därmed måste $\mathbf{x}$–delen av varje lösning till Lagrangevillkoren vara just $\hat{\mathbf{x}}$. Det återstår att bestämma motsvarande $\hat{\mathbf{u}}$ så att $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ uppfyller Lagrangevillkoren, dvs så att

$\mathbf{A}^{\mathsf{T}}\hat{\mathbf{u}} = \mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{H}\hat{\mathbf{x}} = (1.5,\, 1.5,\, 1.5,\, 1.5)^{\mathsf{T}}$, vilket utskrivet blir

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.5 \\ 1.5 \end{pmatrix}, \text{ med den unika lösningen } \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 0 \\ 1.5 \end{pmatrix}.$$

Uppgift 4.(a)

Byt beteckning och kalla den sökta konstanten c för x.

Då vill vi alltså minimera $f(x) = \frac{1}{2}\,\mathbf{h}(x)^{\mathsf{T}}\mathbf{h}(x) = \frac{1}{2}(h_1(x)^2 + h_2(x)^2)$,

där $\mathbf{h}(x) = \begin{pmatrix} h_1(x) \\ h_2(x) \end{pmatrix}$, med

$$h_1(x) = \frac{1}{1 + x\,t_1} - w_1 = \frac{1}{1+x} - 0.46 \ \text{ och}$$

$$h_2(x) = \frac{1}{1 + x\,t_2} - w_2 = \frac{1}{1+3x} - 0.22.$$

Detta är ett ickelinjärt MK-problem med $n = 1$ (en enda variablel x)
och $m = 2$ (två termer i kvadratsumman).

Derivering ger att $\nabla\mathbf{h}(x) = \begin{bmatrix} h_1'(x) \\ h_2'(x) \end{bmatrix}$, där $h_1'(x) = \dfrac{-1}{(1+x)^2}$ och $h_2'(x) = \dfrac{-3}{(1+3x)^2}$.

Vi ska enligt uppgiftslydelsen starta i $x^{(1)} = 1$. Då är

$\mathbf{h}(x^{(1)}) = \begin{pmatrix} 0.04 \\ 0.03 \end{pmatrix}$ och $\nabla\mathbf{h}(x^{(1)}) = \begin{bmatrix} -1/4 \\ -3/16 \end{bmatrix}$.

I Gauss-Newtons metod ska man lösa $\nabla\mathbf{h}(x^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(x^{(1)})\,d = -\nabla\mathbf{h}(x^{(1)})^{\mathsf{T}}\mathbf{h}(x^{(1)})$.

Vi har att $\nabla\mathbf{h}(x^{(1)})^{\mathsf{T}}\nabla\mathbf{h}(x^{(1)}) = (-1/4)^2 + (-3/16)^2 = 25/256$

och $-\nabla\mathbf{h}(x^{(1)})^{\mathsf{T}}\mathbf{h}(x^{(1)}) = (4/16)(4/100) + (3/16)(3/100) = 25/1600$,

så vi får ekvationen $(25/256)\,d = 25/1600$, med lösningen $d^{(1)} = 256/1600 = 0.16$.

Vi prövar som vanligt med $t_1 = 1$, så att $x^{(2)} = x^{(1)} + t_1 d^{(1)} = 1 + 0.16 = 1.16$. Då blir

$$h_1(x^{(2)}) = \frac{1}{2.16} - 0.46 = \frac{1 - 0.46 \cdot 2.16}{2.16} = \frac{0.0064}{2.16} < 0.04 = h_1(x^{(1)}) \ \text{ och}$$

$$h_2(x^{(2)}) = \frac{1}{4.48} - 0.22 = \frac{1 - 0.22 \cdot 4.48}{4.48} = \frac{0.0144}{4.48} < 0.03 = h_2(x^{(1)}).$$

Eftersom $|h_1(x^{(2)})| < |h_1(x^{(1)})|$ och $|h_2(x^{(2)})| < |h_2(x^{(1)})|$ så är $f(x^{(2)}) < f(x^{(1)})$,
vilket betyder att $t_1 = 1$ accepteras.

Därmed har vi utfört en iteration med Gauss-Newtons metod och erhållit förslaget $c = 1.16$,
vilket är ett bättre värde än startgissningen $c = 1$.

Uppgift 4.(b)

Vid Newtons metod, utgående från samma startpunkt $x^{(1)} = 1$ som ovan,
ska man i stället bestämma d ur ekvationen $f''(x^{(1)})d = -f'(x^{(1)})$.

Vi har att $-f'(x^{(1)}) = -h_1'(x^{(1)})h_1(x^{(1)}) - h_2'(x^{(1)})h_2(x^{(1)}) = 25/1600$, och

$f''(x^{(1)}) = h_1'(x^{(1)})^2 + h_2'(x^{(1)})^2 + h_1(x^{(1)})h_1''(x^{(1)}) + h_2(x^{(1)})h_2''(x^{(1)})$.

Men $h_1''(x) = \dfrac{2}{(1+x)^3}$ och $h_2''(x) = \dfrac{18}{(1+3x)^3}$

så $h_1''(x^{(1)}) = 2/8$ och $h_2''(x^{(1)}) = 18/64$.

Därmed är $f''(x^{(1)}) = h_1'(x^{(1)})^2 + h_2'(x^{(1)})^2 + +h_1(x^{(1)})h_1''(x^{(1)}) + h_2(x^{(1)})h_2''(x^{(1)}) =$

$= (-1/4)^2 + (-3/16)^2 + (4/100)(2/8) + (3/100)(18/64) > (-1/4)^2 + (-3/16)^2 = 25/256$.

Lösningen $d^{(1)}$ till Newtonekvationen $f''(x^{(1)})d = -f'(x^{(1)})$ uppfyller därför

$0 < d^{(1)} < 256/1600 = 0.16$,

och eftersom stegparametern $t_1 \leq 1$ så kommer den nya iterationspunkten att uppfylla

$x^{(2)} = x^{(1)} + t_1 d^{(1)} < 1 + 0.16 = 1.16$.

Det värde på konstanten c som en iteration med Newtons metod föreslår är alltså < 1.16,
dvs mindre än det värde som en iteration med Gauss-Newtons metod föreslog.

(Anmärkning: I detta exempel visar det sig att det värde på c som en iteration med Gauss-Newtons metod föreslår är bättre än motsvarande värde för Newtons metod, men detta är inget generellt resultat. Med en annan startgissning än $c = 1$ kan det hända att en iteration med Newtons metod lyckas bättre än en iteration med Gauss-Newtons metod. Det är heller inte alltid fallet att Newtons metod föreslår ett kortare steg än Gauss-Newtons metod.)

Uppgift 5.

Inför funktionen

$$f(\mathbf{x}) = \frac{1}{|x_2 - x_1|} + \frac{1}{|x_3 - x_1|} + \frac{1}{|x_4 - x_1|} + \frac{1}{|x_3 - x_2|} + \frac{1}{|x_4 - x_2|} + \frac{1}{|x_4 - x_3|},$$

och kalla den föreslagna punkten för $\bar{\mathbf{x}}$, dvs $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = (0,\ 1,\ 2,\ 3)$.

För att $\bar{\mathbf{x}}$ ska kunna vara en globalt optimal lösning så måste $\bar{\mathbf{x}}$ åtminstone vara en lokalt optimal lösning, så vi börjar med att undersöka om $\bar{\mathbf{x}}$ är en lokalt optimal lösning till problemet att minimera $f(\mathbf{x})$ under bivillkoren $0 \leq x_1 < x_2 < x_3 < x_4 \leq 3$.

Vi betraktar därför en liten omgivning till punkten $\bar{\mathbf{x}}$,
nämligen mängden av punkter $\mathbf{x} \in I\!R^4$ som uppfyller $\|\mathbf{x} - \bar{\mathbf{x}}\| < 0.1$.

För alla dessa $\mathbf{x}$ är $x_1 < x_2 < x_3 < x_4$, och därmed

$$f(\mathbf{x}) = \frac{1}{x_2 - x_1} + \frac{1}{x_3 - x_1} + \frac{1}{x_4 - x_1} + \frac{1}{x_3 - x_2} + \frac{1}{x_4 - x_2} + \frac{1}{x_4 - x_3}, \quad \text{utan beloppstecken!}$$

Funktionen f kan då deriveras, och de partiella derivatorna av f ges av

$$\frac{\partial f}{\partial x_1}(\mathbf{x}) = \frac{1}{(x_2 - x_1)^2} + \frac{1}{(x_3 - x_1)^2} + \frac{1}{(x_4 - x_1)^2}, \quad \text{speciellt} \quad \frac{\partial f}{\partial x_1}(\bar{\mathbf{x}}) = \frac{1}{1} + \frac{1}{4} + \frac{1}{9} > 0,$$

$$\frac{\partial f}{\partial x_2}(\mathbf{x}) = \frac{-1}{(x_2 - x_1)^2} + \frac{1}{(x_3 - x_2)^2} + \frac{1}{(x_4 - x_2)^2}, \quad \text{speciellt} \quad \frac{\partial f}{\partial x_2}(\bar{\mathbf{x}}) = -\frac{1}{1} + \frac{1}{1} + \frac{1}{4} > 0,$$

$$\frac{\partial f}{\partial x_3}(\mathbf{x}) = \frac{-1}{(x_3 - x_1)^2} + \frac{-1}{(x_3 - x_2)^2} + \frac{1}{(x_4 - x_3)^2}, \quad \text{speciellt} \quad \frac{\partial f}{\partial x_3}(\bar{\mathbf{x}}) = -\frac{1}{4} - \frac{1}{1} + \frac{1}{1} < 0,$$

$$\frac{\partial f}{\partial x_4}(\mathbf{x}) = \frac{-1}{(x_4 - x_1)^2} + \frac{-1}{(x_4 - x_2)^2} + \frac{-1}{(x_4 - x_3)^2}, \quad \text{speciellt} \quad \frac{\partial f}{\partial x_4}(\bar{\mathbf{x}}) = -\frac{1}{9} - \frac{1}{4} - \frac{1}{1} < 0.$$

Först undersöker vi om det lönar sig att ändra x_1 lite från dess aktuella värde $\bar{x}_1 = 0$.
Att $\dfrac{\partial f}{\partial x_1}(\bar{\mathbf{x}}) > 0$ medför att om vi ökar x_1 lite så ökar $f(\mathbf{x})$, medan om vi minskar x_1 lite så minskar $f(\mathbf{x})$. Men x_1 får inte bli < 0 så vi kan inte minska x_1. Vi kan endast öka x_1, vilket dock leder till en ökning av $f(\mathbf{x})$. Det lönar sig alltså inte med någon liten ändring av x_1.

Sedan undersöker vi om det lönar sig att ändra x_2 lite från dess aktuella värde $\bar{x}_2 = 1$.
Att $\dfrac{\partial f}{\partial x_2}(\bar{\mathbf{x}}) > 0$ medför att om vi minskar x_2 lite så minskar $f(\mathbf{x})$, och det är tillåtet att minska x_2 lite, ty bivillkoren $0 \leq x_1 < x_2 < x_3 < x_4 \leq 3$ kommer fortfarande att vara uppfyllda om minskningen av x_2 är tillräckligt liten.

Därav följer att $\bar{\mathbf{x}}$ inte är en lokalt optimal lösning till problemet, och därmed inte heller en globalt optimal lösning!

Kontroll: Antag att $\mathbf{x} = (\bar{x}_1,\ \bar{x}_2 - \delta,\ \bar{x}_3,\ \bar{x}_4) = (0,\ 1 - \delta,\ 2,\ 3)$, där $\delta > 0$ men "litet".

Då är $f(\mathbf{x}) = \dfrac{1}{1 - \delta} + \dfrac{1}{2} + \dfrac{1}{3} + \dfrac{1}{1 + \delta} + \dfrac{1}{2 + \delta} + \dfrac{1}{1} \approx f(\hat{\mathbf{x}}) - 0.25\,\delta < f(\hat{\mathbf{x}})$,

där vi gjort en första ordningens Taylorutveckling kring $\delta = 0$.