Tentamen TMA945 Tillämpad optimeringslära 990528

1 a) Efter införande av slackvariabler är problemet ekvivalent med problemet att

$$
\begin{aligned}
&\text{minimera } f(x) := -x_1 + 2x_2 + x_3 \\
&\text{då } \left[\begin{array}{l}
2x_1 + \ \ x_2 - \ \ x_3 + s_1 \qquad = 7 \\
-x_1 + 2x_2 + +3x_3 \qquad - s_2 = 3 \\
x_1 \quad x_2, \quad x_3, \quad s_1, \quad s_2 \geq 0
\end{array}\right.
\end{aligned}
$$

Ett Fas-I problem skapas genom tillägg av en artificiell variabel i det andra bivillkoret:

$$
\begin{aligned}
&\text{minimera } h(a) := a \\
&\text{då } \quad 2x_1 + \ \ x_2 - \ \ x_3 + s_1 \qquad\qquad = 7 \\
&\qquad -x_1 + 2x_2 + 3x_3 \qquad - s_2 + a = 3 \\
&\qquad\ x_1, \quad x_2, \quad x_2, \quad s_1, \quad s_2, \quad a \geq 0
\end{aligned}
$$

Startbas: s_1, a; målfunktionen uttryckt i övriga

variabler: $h(a) = a = 3 + x_1 - 2x_2 - 3x_3 + s_2$.

Tablå:

bas	$-h$	x_1	x_2	x_3	s_1	s_2	a	$\bar{b}$	
$-h$	1	1	-2	-3	0	1	0	-3	x_3 ink.
s_1	0	2	1	-1	1	0	0	7	a utg.
a	0	-1	2	$\boxed{3}$	0	-1	1	3	
$-h$	1	0	0	0	0	0	1	0	red. kostn. ≥ 0 optimum i Fas-I tillåten bas i s_1, x_3
s_1	0	5/3	5/3	0	1	$-1/3$	1/3	8	
x_3	0	$-1/3$	2/3	1	0	$-1/3$	1/3	1	inför f, uttryckt i icke-basvar.

$$
\begin{aligned}
f(x) &= -x_1 + 2x_2 + x_3 \\
&= -x_1 + 2x_2 + (1 + \frac{1}{3}x_1 - \frac{2}{3}x_2 + \frac{1}{3}s_2) \\
&= 1 - \frac{2}{3}x_1 + \frac{4}{3}x_2 + \frac{1}{3}s_2
\end{aligned}
$$

Fas-II: bas	$-f$	x_1	x_2	x_3	s_1	s_2	$\bar{b}$	
$-f$	1	$-2/3$	$4/3$	0	0	0	-1	x_1 ink.
s_1	0	$\boxed{5/3}$	$5/3$	0	1	$-1/3$	8	s_1 utg.
x_3	0	$-1/3$	$2/3$	1	0	$-1/3$	1	
$-f$	1	0	2	$2/5$	$2/5$	$1/5$	$1/5$	red. kostn. ≥ 0 Optimum i Fas-II
x_1	0	1	1	0	$3/5$	$-1/5$	$24/5$	
x_3	0	0	1	1	$1/5$	$-2/5$	$13/5$	

$$x^* = (24/5, 0, 13/5)^T; z^* = -11/5$$

b) Skuggpriserna ges av de optimala värdena hos variablerna i motsvarande duala problem. Tecknen hos dem är ≤ 0 respektive ≥ 0, och deras numeriska värden (sånär som på tecken) återfinns som reducerade kostnader hos slackvariablerna. Följaktligen är $y^* = (-2/5, 1/5)$. Om nu 3 ersätts med $3+\epsilon$ i högerledet för bivillkor två kommer det optimala värdet förändras med $\epsilon/5$.

Index: $i =$ individ $(i = 1, \dots, 6)$

$j =$ veckodag $(j = 1, \dots, 6)$

Kostnader: $p_{ij} =$ poäng för dag j satt av individ i

$$c_i = \begin{cases} 1 & \text{om individ } i \text{ behärskar espressomaskinen} \\ 0 & \text{annars} \end{cases}$$

Modell:

2 a)

- Variabler:

$$x_{ij} = \begin{cases} 1 & \text{om individ } i \text{ arbetar dag } j \\ 0 & \text{annars} \end{cases}$$

- Bivillkor:

$$\begin{aligned} &\textstyle\sum_i x_{ij} = 4, j = 1, \dots, 6 (\text{ personalbehov}) \\ &\textstyle\sum_i c_i x_{ij} \geq 1, j = 1, \dots, 6 (\text{ maskinkompetens}) \\ &\textstyle\sum_j x_{ij} \leq 4, i = 1, \dots, 6 (\text{ arbetsbegränsning}) \\ &x_{ij} \in \{0, 1\}, \forall i, j \end{aligned}$$

- Målfunktion: maximera lyckan $= \sum_i \sum_j p_{ij} x_{ij}$

b) Ny variabel: $z =$ lyckan hos den sämst lottade.

Ny modell: maximera z.

$$\begin{aligned} \text{då} \quad &\textstyle\sum_j p_{ij} x_{ij} \geq z, i = 1, \dots, 6 \\ &\textstyle\sum_i c_{ij} = 4, j = 1, \dots, 6 \\ &\textstyle\sum_i c_{ij} x_{ij} \geq 1, j = 1, \dots, 6 \\ &x_{ij} \in \{0, 1\}, \forall i, j, z \text{ fri} \end{aligned}$$

Kommentar: den egentliga målfunktionen är att maximera den lägsta lyckan d.v.s. att maximera $f(x) := \min_i(\sum_j p_{ij} x_{ij})$. Eftersom denna funktion inte är linjär konverteras den enligt ovanstående med hjälp av en extra variabel.

3 a)

$$s_{ij} = \text{ flöde på båge } (i,j)$$
$$f = \text{flödesstyrka genom grafen}$$

Maxflödesproblemet:
maximera f

$$\text{då} \begin{cases} \displaystyle\sum_{j:(i,j)\in E} x_{ij} - \sum_{j:(j,i)\in E} x_{ij} = \begin{cases} f & i = s \\ 0 & \neq s,t \\ -f & i = t \end{cases} \\ \\ u_{ij} \geq x_{ij} \geq 0 \quad \forall (i,j) \in E \\ f \quad \text{fri} \end{cases}$$

b) Inför dualvariabler π_i för nodjämviktsvillkoren och α_{ij} för kapacitetsvillkoren.

Dualt problem:
minimera $\displaystyle\sum_{(i,j)\in E} u_{ij}\alpha_{ij}$

$$\text{då} \left[\begin{aligned} \pi_i - \pi_j + \alpha_{ij} &\geq 0, \quad \forall (i,j) \in E \\ \pi_t - \pi_s &= 1 \\ \alpha_{ij} &\geq 0. \quad \forall (i,j) \in E \\ \pi_i \text{ fri}, &\quad i \in N \end{aligned} \right.$$

Låt $S \in N$ vara sådant att $s \in S, t \notin S$. Då beskriver mängderna S och $N \backslash S$ ett snitt i grafen. Maxflödes-minusnittsteoremet säger att f^* (den maximala flödesstyrkan) är lika med den minimala kapaciteten hos något snitt i grafen, vilket kan uttryckas som $\sum_{\substack{(i,j)\in E \\ i\in S \\ j\notin s}} u_{ij}$, vilken minimeras över $S \subset N$.

Motivering mha LP-dualitet:

Sätt

$$\pi_i = \begin{cases} 0, & i \in s \quad (0 \text{ på } s\text{-sidan}, \\ 1, & i \notin s \quad 1 \text{ på } t\text{-sidan}) \end{cases}$$

$$\alpha_{ij} = \begin{cases} 1, & (i,j) \in E, i \in S, j \notin s \\ 0 & \end{cases} \qquad \text{(bågar över snittet)}$$

Detta är en dualt tillåten lösning, och målfunktionsvärdet är lika med $\sum_{\substack{(i,j)\in E \\ i\in s, j\notin s}} u_{ij}$, d.v.s. styrkan hos snittet definerat av nodmändgen s. Svag dualitet visar att f^* inte kan överstiga styrkan hos något snitt. Den minimala styrkan motsvarar ett flöde x där lflödet av maximalt (u_{ij}) på varje båge som passerar över snittet. Eftersom $f = \sum_{\substack{(i,j)\in s \\ i\in s, j\notin s}} u_{ij}$ gäller där måste x vara optimalt och f den maximala flödesstyrkan.

---

4. $f(x) = (x_1 + 2x_\alpha - 3)^2 + (x_1 - \alpha)^2$;

$$\nabla f(x) = \begin{bmatrix} 4x_1 + 4x_2 - 10 \\ 4x_1 + 8x_\alpha - 12 \end{bmatrix}; \nabla^2 f(x) = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix}; x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

---

a) $p_0 = -\nabla f(x_0) = \begin{bmatrix} 10 \\ 12 \end{bmatrix}; x_1 = x_0 + \alpha_0 p_0 = 2\alpha_0 \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$$\alpha_0 = -\frac{\nabla f(x_0)^T p_0}{p_0^T \nabla^2 f(x_0) p_0} = \ldots = \frac{61}{628} \Rightarrow x_1 = \frac{122}{628} \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

$\nabla f(x_1) = \frac{2}{157} \begin{bmatrix} -114 \\ 95 \end{bmatrix}$; $p \in R^\alpha$ är en avtagande riktning i x_1 om $p^T \nabla f(x_1) <$

0, d.v.s., om

$$p^T \begin{bmatrix} -114 \\ 95 \end{bmatrix} < 0 \Leftrightarrow -114p_1 + 95p_\alpha < 0$$

$\Rightarrow x_1$ är <u>inte</u> optimallösning, ty $f(x_1 + \tilde{p}) < f(x_1)$ om $-114\tilde{p}_1 + 95\tilde{p}_2 < 0$ och $\|\tilde{p}\|$ är tillräckligt liten.

$$p_0 = -(\nabla^2 f(x_0))^{-1} \nabla f(x_0) = -\frac{1}{16} \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} -10 \\ -12 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$

$$x_1 = x_0 + \alpha_0 p_0 = \frac{\alpha_0}{\alpha} \begin{bmatrix} 4 \\ 1 \end{bmatrix}; \alpha_0 = \frac{-\nabla f(x_0)^T p_0}{p_0^T \nabla^2 f(x_0) p_0} = \ldots = 1 \text{ (ty kvadratisk).}$$

$$\Rightarrow x_1 = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}; \nabla f(x_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow p^T \nabla f(x_1) = 0 \not< 0 \text{ för alla } p \in R^\alpha \Rightarrow$$

inga avtagande riktningar existerar. f är konvex, ty $\nabla^2 f(x) = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix}$ är positivt definite för alla $x \in R^2 \Rightarrow x$ är globalt optimum. (f är strikt konvex $\Rightarrow x_1$ är <u>unikt</u> optimum).

5 a) Låt $y_k, k \in K$, vara mängden av samatliga extrempunkter till mängden av tillåtna lösningar i LP-dualen. Då kan vi skriva

$$v(b) = \begin{array}{c} \max c^T x \\ \text{då } Ax \le b \\ x \ge 0 \end{array} = \begin{array}{c} \min b^T y \\ \text{då } A^T y \le c \\ y \ge 0 \end{array} = \min_{k \in K} b^T y_k.$$

Enligt antagandet är (säg) y_k den unika optimala extrempunkten. Då följer att funktionen $v(\cdot)$ är differentierbar i en omgivning av b, och gradienten $\nabla v(b) = y_{k^*}$. Speciellt gäller att den partiella derivatan $\dfrac{\partial v(b)}{\partial b_i}$ ges av $(y_{k^*})_i$, d.v.s. element i i den optimala duallösningen. I känslighetsanalysen utnyttjas att $v(b) = c_B^T B^{-1} b$ för en optimal bas B, och att $(y^*)^T = c_B^T B^{-1}$ är en optimalduallösning. Formeln för skygpriset, $y_i^* = (c_B^T B^{-1})_i$ är alltså korrekt.

b) Om y_{k^*} inte är den enda optimala extrem punkten är inte $v(\dot{})$ differentierbar i punkten b. Om vi är intresserade av att finna det korrekta skuggpriset för villkor i måste formeln i a) förändras. Intressanta blir nu alla de ectrempunkter y_k som ger värdet $v(b)$ (de <u>aktiva</u> extrempunkterna). Det korrekta skuggpriset är då en riktningsderivata av $v(\cdot)$ i b med avseende på en önskad förändringsriktning δ:

$$v'(b;\delta) := \lim_{t\downarrow 0} \frac{1}{t}(v(b+t\delta) - v(b)).$$

(I de fall då $v(\cdot)$ är differentierbar i b gäller att $v'(b;\delta) = \delta^T \nabla v(b)$, vilket betyder att användandet av riktningsderivatan vid härledningen av skuggpriset är det mest generella och är <u>alltid</u> korrekt.) Om vi speciellt söker skuggpriset för villkor i låter vi $\delta = \delta_{+i} := (0,0,\dots,0,1,0,\dots,o)^T$ där bara element i är nollskilt, om de är en <u>ökning</u> av höger ledet vi undersöker, medan vi låter $\delta = \delta_{-i} := (0,0,\dots,0,-1,0,\dots,0)^T$ om vi vill undersöka en minskning. Detta är en viktig observation: om inte $v(\cdot)$ är differntierbar är inte säkert $v'(b;\delta_{+i}) = -v'(b_j\delta_{-i})$ eftersom höger- och vänsterderivatorna inte sammanfaller, och vi måste alltså skilja mellan dessa fall.

För konvexa funktioner (som $v(\cdot)$ är ett exempel på) gäller formeln $v'(b_j\delta) = \max\{\delta^T\xi | \xi \in \partial v(b)\}$, där $\partial v(b)$ är mängden av alla möjliga lutningar hos $v(\cdot)$ i b (de så kallade subgradienterna). Speciellt fås att

$$v'(b_j\delta_{+i}) = \max\{(y_{y^*}); |y_{k^*} \text{ är en optimal duallösning}\}.$$

alltså det <u>största</u> värdet som dualvariabeln y_i har i någon optimallösning.

Notering: Vi kan också, mer informellt, i det duala problemet se att om y_{k^*} är unik gäller det också vid små förändringar av b (lutningen hos målfunktionen) medan om fler än en extrempunkt y_k är optimal för något b förändrar kanske z^* mer oförutsägbart. Detta ger en annan bild av differentierbarheten av z^*.

6. minimera $f(x) := 61.8 + 5.72x_1 + 0.0175(x_3)^{0.85} + 0.0094(x_4)^{0.75} + 0.006t, xx_3$ då

$$\begin{aligned}
t_2 x_1 &\geq d_1 t_1 + d_2(t_2 - t_1) \\
x_2 &\geq p_0 \\
x_2 &= 36.25\frac{(d_2 - x_1)(t_2 - t_1)}{t_1} \ln\left(\frac{x_2}{p_0}\right) \\
x_1 &= 348,300\frac{(d_2 - x_1)(t_2 - t_1)}{x_2} \\
x_j &\geq 0, \forall j
\end{aligned}$$

$\Rightarrow$

minimera $f(x) := 61.8 + 5.72x_1 + 0.0175x_3^{0.85} + 0.0094x_4^{0.75} + 0.0036x_3$

då

$$
\begin{aligned}
g_1(x) :=& x_1 - 17.5 \geq 0 \\
g_2(x) :=& x_2 - 200 \geq 0 \\
h_1(x) :=& x_3 - 24\frac{1}{6}(40 - x_1)\ln\left(\frac{x_2}{200}\right) = 0 \\
h_2(x) :=& x_4 - 139,320(40 - x_1)/x_2 = 0 \\
g_3, \ldots, g_6(x) :=& x_1, \ldots, x_4 \geq 0
\end{aligned}
$$

Vi kontrollerar om KKT-villkoren är uppfyllda i
$\bar{x} = (17.5, 473.7, 468.8, 6618)^T$:

$$g_1(\bar{x}) = 0, g_2(\bar{x}) > 0, h_1(\bar{x}) \approx 0, h_2(\bar{x}) \approx 0, g_3, \ldots, g_6(\bar{x}) > 0.$$

Aktiva olikhetsvillkor alltså enbart bivillkor 1.

$$
\begin{aligned}
\nabla g_1(\bar{x}) &= (1, 0, 0, 0)^T \\
\nabla H_1(\bar{x}) &= (24\frac{1}{6}\ln(\bar{x}_2/200), 24\frac{1}{6}(40 - \bar{x}_1)/(200\bar{x}), 1, 0)^T \approx (20.84, 0.005739, 1, 0)^T \\
\nabla H_2(\bar{x}) &= (139,320/\bar{x}_2, 139,320(40 - \bar{x}_1)/\bar{x}_2^2, 0, 1)^T \approx (294.1, 13.97, 0, 1)^T \\
\nabla f(\bar{x}) &= (5.72, 0, 0.85 \cdot 0.0175\bar{x}_3^{-0.15} + 0.0036, 0.75 \cdot 0.0094\bar{x}_4^{-0.25}) \approx \\
&\quad (5.72, 0, 0.00513, 0.0007896).
\end{aligned}
$$

Finn en lösning till

$$\nabla f(\bar{x}) - \lambda_1 \nabla g_1(\bar{x}) - \mu_1 \nabla h_1(\bar{x}) - \mu_2 \nabla h_2(\bar{x}) = \mathbf{0}$$

$$
\begin{bmatrix}
1 & 20.84 & 294.1 \\
0 & 0.005739 & 13.97 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\lambda_1 \\ \mu_1 \\ \mu_2
\end{bmatrix}
=
\begin{bmatrix}
5.72 \\ 0 \\ 0.009513 \\ 0.0007816
\end{bmatrix}
\Rightarrow
$$

$\bar{\mu}_1 = 0.009513(\geq 0!)$ och $\bar{\mu}_2 = 0.0007816(\geq 0!)$ ur rad 3 - 4, vilket i rad 1 ger att $\bar{\lambda}_1 = 5.292(\geq 0!)$. Ekvation 2 uppfylls inte exakt med dessa val, men felet är litet:

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = \nabla f(\bar{x}) - \bar{\lambda}_1 \nabla g_1(\bar{x}) - \bar{\mu}_1 \nabla h_1(\bar{x}) - \bar{\mu}_2 \nabla h_3(\bar{x}) = (0, 0.011, 0, 0)^T$$

vilket innebär att $\|\nabla_x L(\bar{x}\bar{\lambda}, \bar{\mu})\| = 0.011$. En ytterligare bild av att $(\bar{x}, \bar{\lambda}, \bar{\mu})$ är att betrakta som en acceptabelt nära stationär punkt till L är att det vanligt förekommande avbrottskriteriet i obegränsad optimering, applicerat på $\min_x L(x, \bar{\lambda} < \bar{\mu})$,

$$\|\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu})\| \leq \epsilon(1 + |L(\bar{x}, \bar{\lambda}, \bar{\mu})|), \epsilon > 0$$

är uppfyllt med $\epsilon \approx 6 \cdot 10^{-5}$. Eftersom $\bar{x}$ är primalt tillåten (approximativt), $(\bar{\lambda}, \bar{\mu}) \geq \mathbf{0}$, $\bar{x}$ och $(\bar{\lambda}, \bar{\mu})$ är komplementära och $(\bar{x}, \bar{\lambda}, \bar{\mu})$ är en stationär punkt till $L(\cdot, \bar{\lambda}, \bar{\mu})$ (approximativt), betraktar vi $(\bar{x}, \bar{\lambda}, \bar{\mu})$ som en approximativ KKT-punkt.

Betraktar vi sedan $\nabla^2_{xx} L(\bar{x}, \bar{\lambda}, \bar{\mu})$ finner vi att den är mycket nära en nollmatris. Vi uppfyller alltså också den andra ordningens nödvändiga villkor för ett lokalt minimum (åtminstone approximativt).

Studerar vi nu problemet närmare ser man att x_1 dominerar kostnadsbidraget kraftigt. Vi kan också se att $\bar{\lambda}_1$ är stor, vilket indikerar att en minskning av dess värde skulle ha blivit effekten av en lägre undre gräns. Om vi betraktar x_1 som fixt på sin lägsta nivå kan vi betrakta problemet som ett optimeringsproblem enbart i x_1, eftersom x_2 och x_4 ges av värdena hos x_1 och x_2. Målfunktionen, satt i x_2, är av formen $k_1(\text{lu } x_2)^{0.85} + k_2 x_2^{-0.75}$ för några $k_1, k_2 > -$. För de stora värden av x_2 som krävs är funktionen mycket flack, d.v.s. linjär. Detta antyder också av att $\nabla_w L(\bar{x}, \bar{\lambda}, \bar{\mu})$ är nära noll. I det intressanta området är alltså optimeringsproblemet i det närmaste att betrakta som ett LP-problem, snarare om det (icke-konvex) ILP-problem som vi ser i urspringsformuleringen. Vi kan troligen lita på att $\bar{x}$ är en bra lösning, efatersom KKT-villkoren är både nödvändiga och *tillräckliga* för global optimalitet hos LP-problem. (Det faktum att utskriften gavs med få signifikanta siffror gör att vi inte kan räkna med att kunna uppfylla optimalitetsvillkoren exakt, och eftersom ursprungsdata optimalitetsvillkoren exakt, och eftersom ursprungsdata troligen är bekräftade med fel är $\bar{x}$ att betrakta som ett tillräckligt bra närmevärde.)

Notering och en alternativ analys:

Vi kan se att modellen är olyckligt formulerad. Till exempel ser vi att x_1 dominerar kostnaden totalt, till priset av att övriga variablers värden blir nästan försumbara. Ett bättre angreppssätt hade varit att skala variablerna så att deras inflytande blivit mer jämbördiga.

Utdata från programmet kan också användas på ett annorlunda sätt än vi gjorde. Eftersom x_3 och x_4 är funktioner av x_1 och x_2 borde en rimligare approximation av deras optimala värden ha utnyttjat sambanden (2c) och (2d) givet värdena hos x_1 och x_2. Med denna modifiering hade vi vunnit flera saker. (1) Vi hade uppfyllt alla bivillkor, speciellt de fysikaliska sambanden (2c) och (2d). (2) Vi hade därmed sluppit flelet i KKT-villkoren ($\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = \mathbf{0}$) vars värde i analysen föranleder en viss betänksamhet. Notera att felet (residualen) som är 0.011, är flera storleksordningar större än högerleden. En numerisk analytiker skulle hävda att lösningen $(\bar{\lambda}, \bar{\mu})$ (speciellt $\bar{\mu}$) inte har särskilt många korrekta siffror...; vi hade också fått en annorlunda lösning om vi hade angripit systemet annorlunda. Vi noterar slutligen att med utnyttjar de av en omskrivning av f i (x_1, x_2) kan visa att KKT-villkoren gäller, inklusive andra ordningens villkore.

7 a) Betrakta problemet

$$
\text{(P)} \qquad \begin{array}{ll} \text{minimera }_x \in X f(x) & \\ \text{då} \quad \begin{array}{ll} g_i(x) \leq g_i(\bar{x}), & i \in I := \{i = 1, \ldots, m | \bar{y}_i > 0\} \\ g_i(x) \leq b_i, & i \notin I, \text{ där } b_i \geq g_i(\bar{x}). \end{array} \end{array}
$$

Vi visar att $(\bar{x}, \bar{y})$ är en sadelpunkt till (P), och detta medför direkt att $\bar{x}$ är optimal i (P). (Beviset för det senare kan utnyttja resultatet i b).) Låt

$$
L(x, y) = f(x) + \sum_{i \in I} y_i(g_i(x) - g_i(\bar{x})) + \sum_{i \notin I} y_i(g_i(x) - b_i), x \in X, y \in \mathbb{R}^m_+.
$$

$(\bar{x}, \bar{y})$ är en sadelpunkt till L på $x \times \mathbb{R}^m_+$ om

$$
L(\bar{x}, y) \overset{(1)}{\leq} L(\bar{x}, \bar{y}) \overset{(2)}{\leq} L(x, \bar{y}), \quad \forall (x, y) \in X \times \mathbb{R}^m_+.
$$

Olikhet 1: Fixera $x = \bar{x}$ och studera problemet att $max_{y \in \mathbb{R}^m_+} L(\bar{x}, y) = f(\bar{x}) + \sum_{i \notin I} y_i(g_i(\bar{x}) - b_i)$. Eftersom $y_i, i \in I$, inte förekommer i $L(\bar{x}, y)$ kan vi sätta $y_i = \bar{y}_i, i \in I$. Enligt förutsättningarna är $b_i \geq g_i(\bar{x}), i \notin I$. Det är därför optimalt att välja $y_i = 0, i \notin I$. Detta bevisar att olikhet 1 gäller.

Olikhet 2: Fixera $y = \bar{y}$ och studera problemet att $\text{minimera}_{x \in X} L(x, \bar{y}) = f(x) + \sum_{i \in I} \bar{y}_i(g_i(x) - g_i(\bar{x}))$. Eftersom $\bar{y}_i \cdot g_i(\bar{x})$ är konstant är detta problemet ekvivalent med att $\text{minimera}_{x \in X} f(x) + \sum_{i=1}^m \bar{y}_i g_i(x)$, vilket enligt förutsättningarna har optimallösning $\bar{x}$. Detta bevisar att olikhet 2 gäller.

b) Resultatet följer direkt av svaga dualsatsen:

$$
\begin{aligned} L(\bar{y}) \leq f^*, \forall \bar{y} \in \mathbb{R}^m_+ &\Rightarrow f(\bar{x}) + \sum_{i=1}^m \bar{y}_i g_i(\bar{x}) \leq f^* \\ &\Leftrightarrow f(\bar{x}) - f^* \leq \sum_{i=1}^m \bar{y}_i g_i(\bar{x}). \end{aligned}
$$

Om $\bar{x}$ är tillåten (d.v.s. $g_i(\bar{x}) \leq 0, i = 1, \ldots, m$) är högerledet icke-negativt, vilket betyder att $\bar{x}$ är en optimal lösning $(f(\bar{x}) = f^*)$ om $\sum_{i=1}^m \bar{y}_i g_i(\bar{x}) = 0$, d.v.s. om $\bar{x}$ och $\bar{y}$ är komplementära. Sammanfattningsvis är $\bar{x}$ optimal i problemet om:

i) $\bar{x}$ är primalt tillåten: $g_i(\bar{x}) \leq 0, i = 1, \ldots, m$

ii) $\bar{y}$ är dualt tillåten: $\bar{y}_i \geq 0, i = 1, \ldots, m$

iii) $\bar{x}$ löser problemet att minimera$_{x \in X} L(x, \bar{y})$

iv) $\bar{x}$ och $\bar{y}$ är komplementära: $\bar{y}_i \cdot g_i(\bar{x}) = 0, i = 1, \ldots, m.$

Relation till KKT-villoren: KKT-villkoren formuleras för en specialiserad fform av problemet där $X = \mathbb{R}^n$ och där f och $g_i, i = 1, \ldots, m$ är differentierbara funktioner. Under dessa förutsättningar motsvarar KKt-villkoren optimalitetsvillkoren ovan precis, sånär som på en viktig skillnad: i KKT-villkoren ersätts ii) med kravet att $\bar{x}$ är en stationär punkt till Lagrangefunktionen. Det kravet är *svagare* än iii), vilket innebär att $\bar{x}$ inte bara uppfyller första ordningens krav på att $\bar{x}$ är ett lokalt minimum till $L(\cdot, \bar{y})$ utan att faktiskt $\bar{x}$ är ett globalt minimum till $L(\cdot, \bar{y})$. Med det starkare kravet följer också det starkare resultaet att $\bar{x}$ är ett globalt minimum till f i urspringsproblemet, medan KKT-villkoren, som vi vet, bara medför detta under extra förutsättningar om problemets konvexitet.