Statisk Optimering

Jesper Michael Møller

Matematisk Institut, Universitetsparken 5, DK–2100 København
E-mail address: moller@math.ku.dk
URL: http://www.math.ku.dk/~moller

Indhold

KAPITEL 1

Ikke-lineær optimering

1. Hvad er ikke-lineær optimering?

Ikke-lineær optimering handler om løsning af ikke-lineære optimeringsprogrammer. Et ikke-lineært optimeringsprogram er et optimeringsproblem hvor det gælder om at maksimere eller minimere en objektfunktion under bibetingelser giver ved identiteter eller uligheder. Problemet kaldes for ikke-lineært hvis vi ikke forudsætter at de involverede funktioner er lineære. Vi antager i stedet at funktionerne er differentiable (lokalt approksimativt lineære) eller konvekse.

2. Konvekse og konkave funktioner $\mathbf{R}^n \to \mathbf{R}$

Vi rekapitulerer nogle fakta om konvekse funktioner.

Definition 2.1 (Konveks mængde). *En delmængde A af $\mathbf{R}^n$ er konveks hvis $\lambda_0 x_0 + \lambda_1 x_1$ for all punkter $x_0, x_1 \in A$ og alle reelle tal $\lambda_0, \lambda_1 \geq 1$ med $\lambda_0 + \lambda_1 = 1$.*

Lad $A \subseteq \mathbf{R}^n$ være en konveks delmængde af $\mathbf{R}^n$.

Definition 2.2 (Konveks og konkav funktion). *En reel funktion $f\colon A \to \mathbf{R}$ på A er* konveks *hvis epigrafen*

$$\{(x,t) \in A \times \mathbf{R} \mid t \geq f(x)\}$$

er en konveks delmængde af $\mathbf{R}^{n+1}$, og konkav *hvis subgrafen*

$$\{(x,t) \in A \times \mathbf{R} \mid t \leq f(x)\}$$

er en konveks delmængde af $\mathbf{R}^{n+1}$, dvs hvis $-f$ er konveks.

En alternativ definition er at f er konveks hvis og kun hvis $f(\lambda_0 x_0 + \lambda_1 x_1) \leq \lambda_0 f(x_0) + \lambda_1 f(x_1)$ og konkav hvis og kun hvis $f(\lambda_0 x_0 + \lambda_1 x_1) \geq \lambda_0 f(x_0) + \lambda_1 f(x_1)$ når $\lambda_0 \geq 0, \lambda_1 \geq 0$ og $\lambda_0 + \lambda_1 = 1$.

Grafen for en konveks (konkav) funktion ligger altid over (under) tangentplanen:

Sætning 2.3. *Lad $A \subseteq \mathbf{R}^n$ være en åben konveks delmængde og $f\colon A \to \mathbf{R}$ en reel C^1-funktion på A. Følgende betingelser er ækvivalente:*

(1) f er konveks
(2) $f(\lambda_0 x_0 + \lambda_1 x_1) \leq \lambda_0 f(x_0) + \lambda_1 f(x_1)$ for all punkter $x_0, x_1 \in A$ og alle reelle tal $\lambda_0 \geq 0, \lambda_1 \geq 0$ med $\lambda_0 + \lambda_1 = 1$.
(3) $\forall x, y \in A\colon f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$
(4) $\forall x \in A\colon \nabla^2 f(x)$ er positiv semidefinit (hvis f er C^2)

Tilsvarende,

(5) f er konkav
(6) $f(\lambda_0 x_0 + \lambda_1 x_1) \geq \lambda_0 f(x_0) + \lambda_1 f(x_1)$.
(7) $\forall x, y \in A\colon f(y) \leq f(x) + \nabla f(x) \cdot (y - x)$
(8) $\forall x \in A\colon \nabla^2 f(x)$ er negativ semidefinit (hvis f er C^2)

Bevis. (1) Antag at f er konveks. Lad x og y være punkter i A. Definition 2.2 siger at $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ når $0 \leq t \leq 1$. Anvender vi $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=1}$ får vi $\nabla f(x) \cdot (x - y) \leq f(x) - f(y)$. Se [8, Sætning 4.6.1] for beviset for at den anden implikation $\Longleftarrow$ også gælder.
(2) [8, Sætning 4.3.1, 4.5.1] □

Mængden af konvekse (konkave) funktioner lukket under dannelse af sum og multiplikation med positiv konstant.

Sætning 2.4. [8, Sætning 4.3.1] *Lad S være en symmetrisk $n \times n$ matrix.*

(1) S er diagonaliserbar og alle egenværdier for S er reelle
(2) S er positiv definit (positiv semidefinit) $\iff$ Alle egenværdier for S er positive (ikke-negative)

(3) (Sylvesters kriterium) S er positiv definit $\iff$ Alle ledende principale underdeterminanter for S er positive
(4) S er positiv semidefinit $\iff$ Alle principale underdeterminanter for S er ikke-negative

En principal delmatrix fås ved at slette nogle rækker og de samme søjler fra S. (Der er ret mange af dem!)

3. Implicit given funktion sætning

Implicit given funktion sætning beskriver strukturen af løsningmængden for et system af m ligninger med $n+m$ ubekendte.

SÆTNING 3.1 (Implicit given funktion sætning). *Lad $F\colon \mathbf{R}^{n+m} \to \mathbf{R}^m$ være en C^1-funktion og lad x_0 være et punkt i $\mathbf{R}^{n+m}$ med $F(x_0) = 0$. Antag at $\mathrm{rk}(DF(x_0)) = m$ (dvs de m gradienter $\nabla F_1(x_0), \dots, \nabla F_m(x_0)$ for koordinatfunktionerne i F er lineært uafhængige eller, ækvivalent, at der er m lineært uafhængige søjler i Jacobimatricen $DF(x_0)$). De m ligninger $F(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) = 0$ kan lokalt bruges til at udtrykke m koordinater som funktioner af de øvrige n koordinater. Hvis vi siger at de sidste m søjler i $DF(x_0)$ er lineært uafhængige, så kan de sidste m koordinater udtrykkes ved de første n koordinater: Der findes en C^1-funktion $L\colon \mathbf{R}^n \to \mathbf{R}$ sådan at*

$$F(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) = 0 \iff (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) \in F^{-1}(0)$$
$$\iff (x_{n+1}, \dots, x_{n+m}) = L(x_1, \dots, x_n)$$

lokalt nær x_0.

Vi siger at løsningsfunktionen L er implicit givet ved ligningssystemet $F(x) = 0$. Her er to tilføjelser til sætningen.

(1) Da $F(x_1, \dots, x_n, L(x_1, \dots, x_n)) = 0$ giver Kædereglen at

$$\frac{\partial F(x_1, \dots, x_n, L(x_1, \dots, x_n))}{\partial(x_1, \dots, x_n)} = \frac{\partial F}{\partial(x_1, \dots, x_n)} + \frac{\partial F}{\partial(x_{n+1}, \dots, x_{n+m})} \frac{\partial L}{\partial(x_1, \dots, x_n)} = 0$$

eller at

$$\frac{\partial L}{\partial(x_1, \dots, x_n)} = -\Big(\frac{\partial F}{\partial(x_{n+1}, \dots, x_{n+m})}\Big)^{-1} \frac{\partial F}{\partial(x_1, \dots, x_n)}$$

(2) Tangentrummet $T_{x_0}F^{-1}(0)$ består af alle tangenter $u'(t)$ til kurver $u\colon \mathbf{R} \to F^{-1}(0)$. Det er klart at $DF(x_0)(T_{x_0}F^{-1}(0)) = 0$ for $F(u(t)) = 0$ for alle kurver $u(t)$ på løsningsmængden $F^{-1}(0)$. Men faktisk er

$$T_{x_0}F^{-1}(0) = DF(x_0)^{-1}(0) = \mathrm{span}\{\nabla F_1(x_0), \dots, \nabla F_m(x_0)\}^\perp$$

for begge vektorrum har dimension n. Vi benytter at $DF(x_0)x = (\nabla F_1(x_0) \cdot x, \cdots, \nabla F_m(x_0) \cdot x)$ for det sidste lighedstegn. Normalrummet til løsningsmængden er derfor

$$T_{x_0}F^{-1}(0)^\perp = \mathrm{span}\{\nabla F_1(x_0), \dots, \nabla F_m(x_0)\}$$

Dvs at de m lineært uafhængige gradienter $\nabla F_1(x_0), \dots, \nabla F_m(x_0)$ står vinkelret på løsningsmængden $F^{-1}(0)$.

(3) Tangentrummet afsat i x_0

$$x_0 + T_{x_0}F^{-1}(0) = \{x \in \mathbf{R}^n \mid DF(x_0)(x - x_0) = 0\}$$

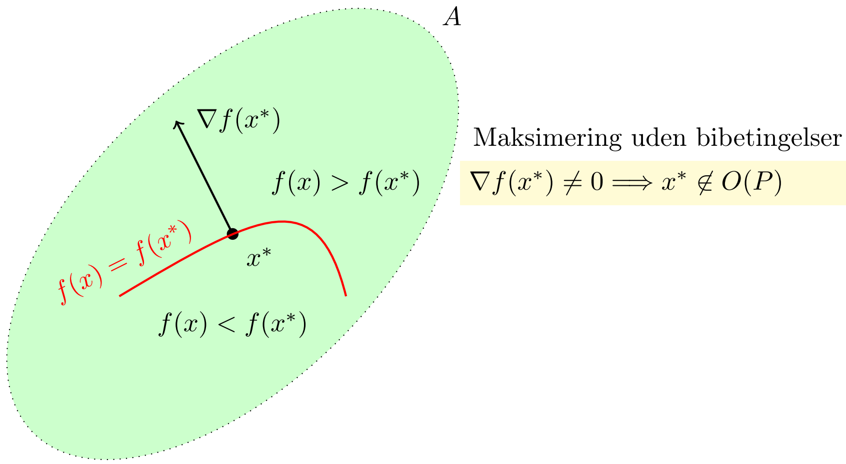
består af de vektorer der afsat fra x_0 er ortogonale på koordinatfunktionernes gradienter.

EKSEMPEL 3.2. Ligningen $x_1^2 - x_2 = 0$ bestemmer nær $(0,0)$, x_2 som en funktion af x_1, nemlig $x_2 = x_1^2$, men ikke x_1 som en funktion af x_2. Ligningen $x_1^2 - x_2^2 = 0$ bestemmer nær $(0,0)$ hverken x_1 som en funktion af x_2 eller x_2 som en funktion af x_1.

EKSEMPEL 3.3 (Tangent til graf). Lad $f\colon \mathbf{R}^n \to \mathbf{R}$ være en C^1-funktion. Grafen for f

$$\{(x,y) \in \mathbf{R}^n \times \mathbf{R} \mid f(x) = y\} = F^{-1}(0)$$

er løsningsmængden for funktionen $F\colon \mathbf{R}^n \times \mathbf{R}^1 \to \mathbf{R}$, $F(x,y) = f(x) - y$. Grafens tangentplan i $(x_0, f(x_0))$ har ligning $\begin{pmatrix}\nabla f(x_0) & -1\end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$ eller $y - y_0 = \nabla f(x_0) \cdot (x - x_0)$.



FIGUR 1. Optimering uden bibetingelser

4. Optimering uden bibetingelser

Vi betragter i dette afsnit optimeringsproblemet

$$\text{(P)} \qquad \text{Maksimér } f$$

hvor

- A er åben en delmængde af $\mathbf{R}^n$
- f er en reel C^1-funktion defineret på A

Problemets mulige og optimale løsninger er

$$M(P) = A$$
$$O(P) = \{x^* \in M(P) \mid f(x^*) \geq f(x) \text{ for alle } x \in M(P)\}$$

De optimale løsninger er de mulige løsninger som optimerer objektfunktionen.

SÆTNING 4.1 (Nødvendig[1] betingelse **NB**). *Hvis $x^* \in O(P)$ er en optimal løsning til optimeringsproblemet (P), så er $\nabla f(x^*) = 0$:*

$$O(P) \subseteq \{x^* \in M(P) \mid \nabla f(x^*) = 0\}$$

BEVIS. Lad $x^* \in M(P)$ være en mulig løsning. Antag at gradienten $\nabla f(x^*)$ *ikke* er 0. Så findes en kurve $u(t)$ gennem x^* til tiden $t = 0$ sådan at $\frac{\mathrm{d}fu}{\mathrm{d}t} = \nabla f(x^*) \cdot \frac{\mathrm{d}u}{\mathrm{d}t}(0) \neq 0$. Altså har fu *ikke* lokalt ekstremum i 0. $\square$

Dette sker meget ofte at $\nabla f(x^*) = 0$ uden at x^* er en optimal løsning. Tænk bare på en funktion $f\colon \mathbf{R} \to \mathbf{R}$ med mange lokale minima og maxima.

Vi kan ikke forvente at de nødvendige betingelser, som er lokale, også er tilstrækkelige. For at opnå en tilstrækkelig betingelse bliver vi nødt til at stille globale krav til objektfunktionen. Feks opnår vi en tilstrækkelig betingelse ved at forlange konkavitet.

SÆTNING 4.2 (Tilstrækkelig betingelse **TB**). *Antag at A er en åben* konveks *delmængde af $\mathbf{R}^n$ og at f er en* konkav *C^1-funktion på A. Hvis $\nabla f(x^*) = 0$, så er x^* en optimal løsning til optimeringsproblemet (P):*

$$O(P) = \{x \in M(P) \mid \nabla f(x^*) = 0\}$$

BEVIS. Antag at f er konkav og at $\nabla f(x^*) = 0$. Ifølge Sætning 2.3.(7) er

$$f(x) \leq f(x^*) + \nabla f(x^*) \cdot (x - x^*) = f(x^*)$$

for alle $x \in A$. Dvs at x^* er et globalt maksimumspunkt for f på A. Den modsatte inklusion blev vist i Sætning 4.1. $\square$

[1] 'A er tilstrækkelig for B' og 'B er nødvendig for A' betyder $A \implies B$

5. Optimering under lighedsbibetingelser

Vi betragter i dette afsnit optimeringsproblemet

$$\text{(P)} \qquad \text{Maksimér } f \text{ under bibetingelserne } g_1 = b_1, \dots, g_m = b_m$$

hvor

- A er en åben delmængde af $\mathbf{R}^n$
- $f\colon A \to \mathbf{R}$ og $g = (g_1, \dots, g_m)\colon A \to \mathbf{R}^m$ er C^1-funktioner defineret på A
- $b = (b_1, \dots, b_m) \in \mathbf{R}^m$

Problemets mulige og optimale løsninger er

$$M(P) = \{x \in A \mid g_1(x) = b_1, \dots, g_m(x) = b_m\} = g^{-1}(0)$$
$$O(P) = \{x^* \in M(P) \mid f(x^*) \geq f(x) \text{ for alle } x \in M(P)\}$$

De mulige løsninger er de punkter i A som opfylder alle bibetingelser og de optimale løsninger er mulige løsninger som optimerer objektfunktionen.

Eksempel 5.1 (Optimering af objektfunktion på graf). Lad $h\colon \mathbf{R}^{n-1} \to \mathbf{R}$ være en C^1-funktion og $G(h) = \{(x, y) \in \mathbf{R}^{n-1} \times \mathbf{R} \mid y = h(x)\}$ grafen for h i $\mathbf{R}^{n-1} \times \mathbf{R} = \mathbf{R}^n$. Lad også $f\colon \mathbf{R}^n \to \mathbf{R}$ være en C^1-funktion på $\mathbf{R}^n$. Problemet

$$(P) \qquad \text{Maksimér } f \text{ under bibetingelsen } g(x, y) = h(x) - y = 0$$

går ud på at maksimere objektfunktionen f på $M(P)$ som er grafen $G(h)$ for h. Gradienten $\nabla g = (\nabla h, -1)$ står vinkelret på niveaufladen $M(P) = g^{-1}(0)$. Hvis $\nabla f \notin \operatorname{span}\{\nabla g\}$, så er ∇f *ikke* vinkelret på grafen og derfor findes en kurve $u(t)$ i $\mathbf{R}^n$ så

$$\frac{\mathrm{d}}{\mathrm{d}t} f(u(t), h(t)) = \nabla f(u(t), h(u(t))) \cdot (u'(t), \nabla h \cdot u'(t))$$

ikke er 0. Altså har f hverken maksimum eller minimum. Vi har altså vist at

$$O(P) \subseteq \{(x, y) \in \mathbf{R}^n \mid \nabla f(x, y) \in \operatorname{span}\{\nabla g(x, y)\}\}$$

Dette er et specialtilfælde af Lagranges sætning.

Definition 5.2 (**LICQ**). *En mulig løsning $x \in M(P)$ opfylder* **LICQ** *(Linear Independence Constraint Qualification) hvis gradienterne for bibetingelserne er lineært uafhængige i x, dvs hvis $\{\nabla g_i(x) \mid i = 1, \dots, m\}$ er en lineært uafhængig mængde af vektorer i $\mathbf{R}^n$.*

Hemmeligheden ved Lagranges sætning er at $M(P)$ er beskrevet i Implicit Funktion Sætning, specielt at normalrummet i x til $M(P)$ er udspændt af gradienterne $\nabla g_1(x), \dots, \nabla g_m(x)$ til bibetingelserne når x opfylder **LICQ**, se Implicit Funktion Sætning 3.1.

Proposition 5.3. *Antag at $x \in M(P)$ er en mulig løsning som opfylder* **LICQ**. *For en vektor v i $\mathbf{R}^n$ gælder:*

$$v \text{ er vinkelret på } M(P) \text{ i } x \iff v \in \operatorname{span}\{\nabla g_1(x), \dots, \nabla g_m(x)\}$$

Sætning 5.4 (Lagrange nødvendig betingelse **LNB**). *Der gælder*

$$O(P) \subseteq \{x \in M(P) \mid \ x \text{ opfylder ikke } \mathbf{LICQ}\} \cup \{x \in M(P) \mid \nabla f(x) \in \operatorname{span}\{\nabla g_i(x)\}\}$$

Bevis. Vi skal vise at

$$\{x \in M(P) \mid \ x \text{ opfylder } \mathbf{LICQ} \text{ og } \nabla f(x) \notin \operatorname{span}\{\nabla g_i(x)\}\} \subseteq M(P) - O(P)$$

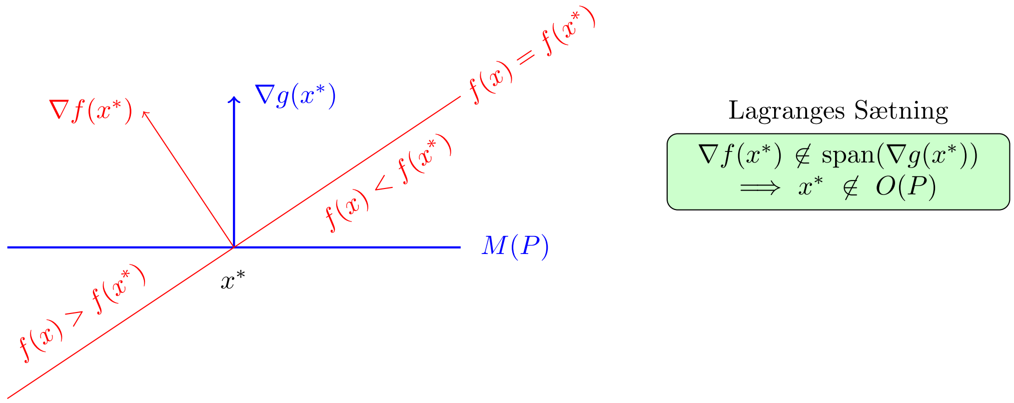
Lad $x \in M(P)$ være en mulig løsning sådan at gradienterne i x for bibetingelserne er lineært uafhængige og objektfunktionens gradient $\nabla f(x)$ *ikke* ligger i underrummet $\operatorname{span}\{\nabla g_1(x), \dots, \nabla g_m(x)\}$, dvs at $\nabla f(x)$ *ikke* er vinkelret på $M(P)$ i x. Der findes da en kurve $u(t)$ på $M(P)$ gennem x til tiden $t = 0$ så $\frac{\mathrm{d}(fu)}{\mathrm{d}t}(0) = \nabla f(x) \cdot \frac{\mathrm{d}u}{\mathrm{d}t}(0) \neq 0$. Så har fu ikke lokalt ekstremum i 0 og x er ikke en optimal løsning til (P). $\square$

Hvis vi lægger et konkavitetskrav på så får vi en tilstækkelig betingelse for at x løser (P).

Sætning 5.5 (Lagrange tilstrækkelig betingelse **LTB**). *Antag at A er en åben konveks delmængde af $\mathbf{R}^n$ og at f og $g_1, \dots, g_m$ er C^1-funktioner på A. Så gælder*

$$\{x \in M(P) \mid \exists u \in \mathbf{R}^m\colon f - u^t g\colon A \to \mathbf{R} \text{ er konkav og } D(f - u^t g)(x) = 0\} \subseteq O(P)$$

Bevis. Antagelsen er at der findes $u \in \mathbf{R}^m$ så $A \to \mathbf{R}\colon x \to f(x) - u^t g(x)$ er en konkav funktion og $D(f - u^t g)(x) = 0$ for en mulig løsning $x \in M(P)$. Ifølge Sætning 4.2 er da $f(x') - u^t g(x') \leq f(x) - u^t g(x)$ for alle $x' \in A$. Men da er $f(x') - u^t b \leq f(x) - u^t b$ for alle $x \in M(P)$. Altså er x en optimal løsning til maksimeringsproblemet (P). $\square$



FIGUR 2. **LNB**

Vi har at $x \in M(P)$ og $\nabla f(x) \in \text{span}\{\nabla g_i(x)\}\}$ hvis og kun hvis der findes $u = (u_1, \ldots, u_m) \in \mathbf{R}^m$ så (x, u) opfylder *Lagrange betingelserne*:

Lagrange betingelser

(a) $g(x) = b$ eller $g_1(x) = b_1, \ldots, g_m(x) = b_m$ (mulig løsning)
(b) $Df(x) = u^t Dg(x)$ eller $Df(x) = u_1 Dg_1(x) + \cdots + u_m Dg_m(x)$ (optimalitet)

Vi reformulerer *Lagrange nødvendig betingelse*:

SÆTNING 5.6 (Lagrange nødvendig betingelse **LNB** II).

$$O(P) \subseteq \{x^* \in M(P) \mid \; x^* \textit{ opfylder ikke } \mathbf{LICQ}\} \cup \{x^* \in A \mid \exists u^* \in \mathbf{R}^m \colon (x^*, u^*) \textit{ opfylder } \mathbf{L} \textbf{ betingelser}\}$$

SÆTNING 5.7 (Lagrange tilstrækkelig betingelse **LTB** II). *Antag at A er en åben konveks delmængde af $\mathbf{R}^n$ og at f og $g_1, \ldots, g_m$ er C^1-funktioner på A. Så gælder*

$$\{x^* \in A \mid \exists u^* \in \mathbf{R}^m \colon (x^*, u^*) \textit{ opfylder } \mathbf{L} \textbf{ betingelser} \textit{ og } f - u^* \cdot g \textit{ er konkav}\} \subseteq O(P)$$

EKSEMPEL 5.8. Vi ser på et optimeringsproblem i dimension 2,

$$\text{(P)} \qquad \text{Maksimér } f(x_1, x_2) = -x_1 - x_2^2 \text{ under bibetingelsen } g(x_1, x_2) = x_1 = 0$$

De mulige løsninger, $M(P) = \{(x_1, x_2) \mid x_1 = 0\}$, er x_2-aksen. Da $\nabla g = (1, 0)$ er **LICQ** opfyldt for alle mulige løsninger. Derfor er

$$O(P) \subseteq \{(x_1, x_2) \in M(P) \mid (-1, -2x_2) \in \text{span}\{(1, 0)\}\} = \{(0, 0)\}$$

Da objektfunktionen er konkav har vi endda $O(P) = \{(0, 0)\}$ ifølge Sætning 4.2.

EKSEMPEL 5.9 (**LICQ**). Vi ser på optimeringsproblemet i dimension 3,

$$\text{(P)} \qquad \text{Maksimér } f(x_1, x_2, x_3) = -x_1 - x_3^2 \text{ under bibetingelserne } g_1(x_1, x_2, x_3) = 0, g_2(x_1, x_2, x_3) = 0$$

og vi vil tage to forskellige par af bibetingelser.

Under de to bibetingelser $g_1(x_1, x_2, x_3) = x_1$, $g_2(x_1, x_2, x_3) = x_2$ i $\mathbf{R}^3$ er de mulige løsninger $M(P) = \{(0, 0, x) \mid x \in \mathbf{R}\}$ tredjeaksen i $\mathbf{R}^3$. Gradienterne $\nabla g_1 = (1, 0, 0)$ og $\nabla g_2 = (0, 1, 0)$ er lineært uafhængige, så **LICQ** er opfyldt i hele $M(P)$. De to gradienter udspænder $\text{span}\{(1, 0, 0), (0, 1, 0)\} = \mathbf{R}^2 \times 0$. Da $\nabla f = (-1, 0, -2x_3)$ siger **LNB** at $O(P) \subseteq \{(0, 0, 0)\}$, dvs at $(0, 0, 0)$ er den eneste mulighed for en optimal løsning. Det er faktisk en optimal løsning da $x_3 = 0$ er et maksimimumspunkt for $f(0, 0, x_3) = -x_3^2$, $x_3 \in \mathbf{R}$.

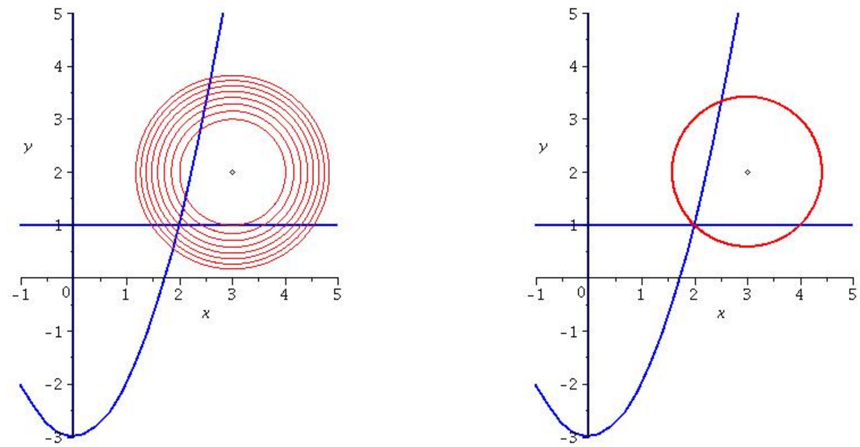
Under de to bibetingelser $g_1(x_1, x_2, x_3) = x_1$, $g_2(x_1, x_2, x_3) = x_1 + x_2^2$ i $\mathbf{R}^3$ er de mulige løsninger $M(P) = \{(0, 0, x) \mid x \in \mathbf{R}\}$ igen tredjeaksen i $\mathbf{R}^3$. Gradienterne $\nabla g_1 = (1, 0, 0)$ og $\nabla g_2 = (1, 2x_2, 0)$ er lineært afhængige, faktisk identiske, i hele $M(P)$ så **LICQ** er ikke opfyldt i noget punkt af $M(P)$. Derfor siger **LNB** bare at $O(P) \subseteq M(P)$ så **LNB** er uden indhold. Vi ved fra før at $(0, 0, 0)$ er en optimal løsning; den opfylder ikke **LICQ**.

6. Optimering under ulighedsbibetingelser

Vi betragter i dette afsnit maksimeringsproblemet

$$\text{(P)} \qquad \text{Maksimér } f \text{ under bibetingelserne } g_1 \leq b_1, \ldots, g_m \leq b_m$$

hvor



FIGUR 3. Eksempel 6.1

- A er en åben delmængde af $\mathbf{R}^n$
- $f\colon A \to \mathbf{R}$ og $g = (g_1, \dots, g_m)\colon A \to \mathbf{R}^m$ er C^1-funktioner defineret på A
- $b = (b_1, \dots, b_m) \in \mathbf{R}^m$

Problemets mulige og optimale løsninger er

$$M(P) = \{x \in A \mid g_1(x) \le b_1, \dots, g_m(x) \le b_m\}$$
$$O(P) = \{x^* \in M(P) \mid f(x^*) \ge f(x) \text{ for alle } x \in M(P)\}$$

Hvis $x \in M(P)$ er en mulig løsning, siger vi at bibetingelsen g_i er *aktiv* i x hvis $g_i(x) = b_i$ og *slæk* i x hvis $g_i(x) < b_i$.

EKSEMPEL 6.1. Minimeringsproblemet, illustreret i Figur 3,

$$\text{(P)}\quad \text{Minimer } (x_1-3)^2 + (x_2-2)^2 \text{ under bibetingelser}$$

$$\begin{aligned} x_1^2 - x_2 - 3 &\le 0 \\ x_2 - 1 &\le 0 \\ -x_1 &\le 0 \end{aligned}$$

går ud på at finde den mindste cirkel med centrum i $(3,2)$ som skærer mængden

$$M(P) = \{(x_1,x_2) \in \mathbf{R}^2 \mid x_1^2 - x_2 - 3 \le 0, x_2 - 1 \le 0, -x_1 \le 0\}$$

af mulige løsninger. En optimal løsning er et punkt i $M(P)$ tættest muligt på $(3,2)$.

Maksimeringsproblemet

$$\text{(P)}\quad \text{Maksimér } (x_1-3)^2 + (x_2-2)^2 \text{ under bibetingelser}$$

$$\begin{aligned} x_1^2 - x_2 - 3 &\le 0 \\ x_2 - 1 &\le 0 \\ -x_1 &\le 0 \end{aligned}$$

går ud på at bestemme det punkt i $M(P)$ der ligger længst væk fra $(3,2)$. Objektfunktionens gradientvektorfelt $\nabla f(x)$ stråler radialt ud fra $(3,2)$. **KKTNB** siger at hvis x^* er optimal så ligger objektfunctionens gradient $\nabla f(x^*)$ i den positive kegle udpsændt af gradienterne for de aktive bibetingelser. Man kun nu tegne sig frem til en løsning til (P). **[Gør det!]**

1.1. Kegler og Farkas's lemma.

Lad $v_1, \dots, v_m$ være m vektorer i $\mathbf{R}^n$

DEFINITION 6.2 (Kegle). *Keglen udspændt af $v_1, \dots, v_m$ er mængden*

$$\operatorname{cone}(\{v_1, \dots, v_m\}) = \{\lambda_1 v_1 + \cdots + \lambda_m v_m \mid \lambda_1, \dots, \lambda_m \ge 0\}$$

af alle postive linearkombinationer af vektorerne. (Vi sætter keglen af den tomme mængde til at være $\operatorname{cone}(\emptyset) = \{0\}$*.)*

LEMMA 6.3 (Farkas' lemma). *Lad $v_1, \dots, v_m$ og b være vektorer i $\mathbf{R}^n$. Så gælder*

$$b \notin \operatorname{cone}(\{v_1, \dots, v_m\}) \iff \exists y \in \mathbf{R}^n\colon y \cdot v_1 \ge 0, \dots, y \cdot v_m \ge 0, y \cdot b < 0$$

Farkas' lemma siger at en vektor b der ikke ligger i keglen kan skilles fra keglen med en hyperplan (med normalvektor y).

1.2. Karush–Kuhn–Tucker betingelserne. Karush–Kuhn–Tucker sætningerne giver nødvendige og tilstrækkelige betingelser for optimalitet i (P).

Hvem viste Karush–Kuhn–Tuckers sætning?

Den originale artikel af Kuhn og Tucker

DEFINITION 6.4 (Mulige kurver). *En mulig kurve gennem den mulige løsning $x \in M(P)$ er en C^1-kurve $s\colon (-\varepsilon, \varepsilon) \to A$ defineret nær 0 sådan at $s(0) = x$ og, for alle bibetingelser g_i som er aktive i x, er $g_i(s(t)) \leq b_i$ for alle $t \leq 0$, $g_i(s(t)) \geq b_i$ for alle $t \geq 0$.*

Lad $s(t)$ være en mulig kurve gennem den mulige løsning $x \in M(P)$. Hvis g_i er aktiv i x så er $\nabla g_i(x) \cdot s'(0) = \frac{\mathrm{d}}{\mathrm{d}t} g_i(s(t))\big|_{t=0} \geq 0$ da $g_i(s(t))$ er voksende i 0. Vi har altså at

$$\{s'(0) \mid s \text{ mulig kurve gennem } x\} \subseteq \{y \in \mathbf{R}^n \mid \nabla g_i(x) \cdot y \geq 0 \text{ for alle } i \text{ med } g_i(x) = 0\} \tag{6.5}$$

men det ville være vidunderligt med et lighedstegn her!

DEFINITION 6.6 (**LICQ**). *En mulig løsning $x \in M(P)$ opfylder* **LICQ** *(Linear Independence Constraint Qualification) hvis gradienterne for de* aktive *bibetingelser er lineært uafhængige i x, dvs hvis*

$$\{\nabla g_i(x) \mid 1 \leq i \leq m, g_i(x) = 0\}$$

er en lineært uafhængig mængde af vektorer.

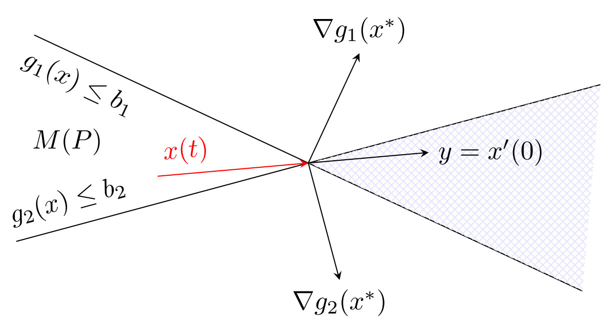
Man kan vise at hvis **LICQ** gælder i $x \in M(P)$ så er der faktisk lighedstegn i (6.5): Tangenterne til mulige kurver er præcis de vektorer som ligger på den positive side af hyperplanerne bestemt ved gradienterne for de aktive bibetingelser. (Det samme gælder klart hvis alle bibetingelser er affine funktioner.)

SÆTNING 6.7 (**KKTNB** I). [5, Theorem 1.27][4, Theorem 3.2]. *Der gælder*

$$O(P) \subseteq \{x \in M(P) \mid x \textit{ opfylder ikke } \mathbf{LICQ}\} \cup \{x \in M(P) \mid \nabla f(x) \in \operatorname{cone}(\{\nabla g_i(x) \mid g_i(x) = 0\})\}$$

BEVIS. Lad $x \in M(P)$ være en mulig løsning som opfylder **LICQ** og $\nabla f(x) \notin \operatorname{cone}(\{\nabla g_i(x) \mid g_i(x) = 0\})$. Påstanden er at x ikke er optimal.

Fra Farkas' Lemma (Lemma 6.3) ved vi at der findes en vektor y som ligger på den positive side af alle aktive bibetingelser og har $y \cdot \nabla f(x^*) < 0$. **LICQ** garanterer at $y = s'(0)$ er tangent for en mulig kurve gennem x. Da $\frac{\mathrm{d}}{\mathrm{d}t} f(s(t))\big|_{t=1} = \nabla f(x) \cdot y < 0$, er $f(s(t))$ en aftagende funktion af t. For $t < 0$ er $f(s(t)) > f(s(0)) = f(x)$ så x er ikke en optimal løsning.



□

Karush–Kuhn–Tuckers *tilstrækkelige* betingelser hviler igen på konkavitet [5, Theorem 1.28].

SÆTNING 6.8 (**KKTTB**). *Antag at A er en åben* konveks *delmængde af $\mathbf{R}^n$. Så er*

$$\{x \in M(P) \mid \exists u \in \mathbf{R}^m \colon u \geq 0,\ u^t(g(x) - b) = 0,\ Df(x) = u^t Dg(x) \textit{ og } f - u^t g \textit{ er konkav}\} \subseteq O(P)$$

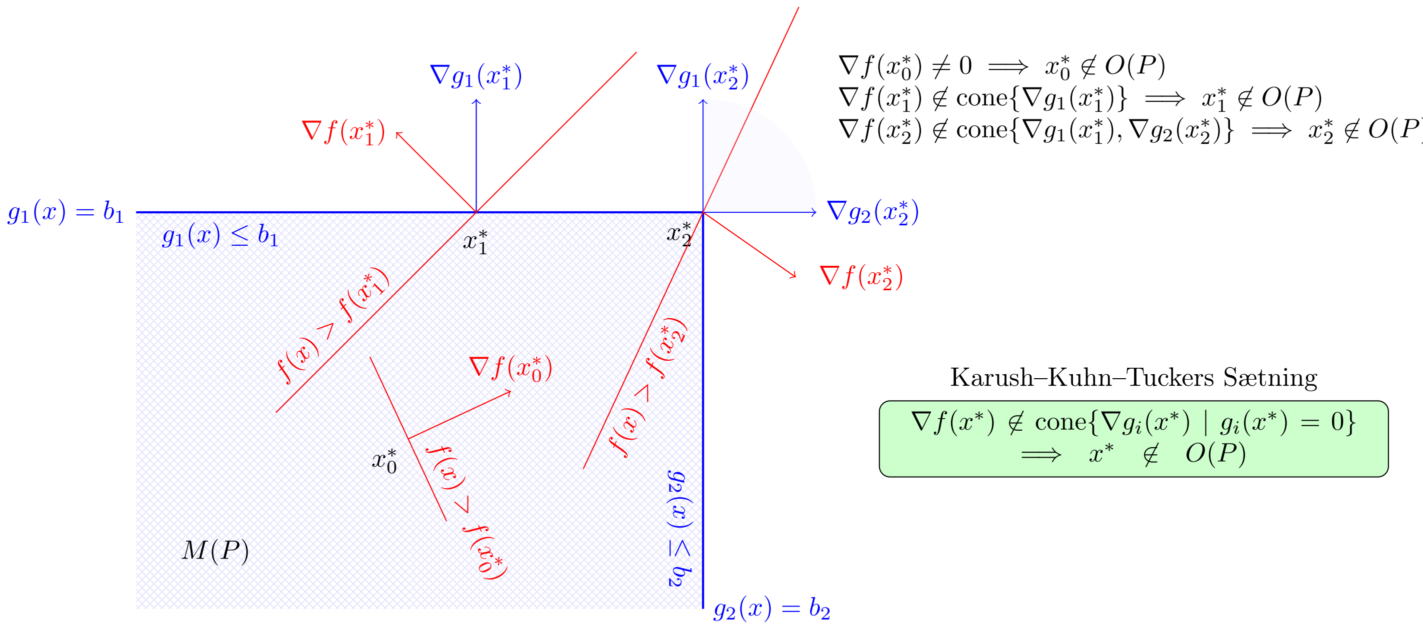
BEVIS. Da $f - u^t g$ er konkav på den konvekse mængde A og gradienten i $x \in M(P)$ er 0, så er x et globalt maksimumspunkt ifølge **TB**:

$$\forall x' \in A \colon (f - u^t g)(x) \geq (f - u^t g)(x')$$

På venstresiden er $u^t g(x) = u^t b$. På højresiden er $u^t g(x') \leq u^t b$ for all $x' \in M(P)$. Altså er

$$\forall x' \in M(P) \colon f(x) - u^t b = (f - u^t g)(x) \geq (f - u^t g)(x') = f(x') - u^t g(x) \geq f(x') - u^t b$$

men dvs at $f(x) \geq f(x')$ for alle $x' \in M(P)$, eller $x \in O(P)$. □



FIGUR 4. **KKTNB**

Vi har at $x \in M(P)$ og $\nabla f(x) \in \mathrm{cone}(\{\nabla g_i(x) \mid g_i(x) = 0\})$ hvis og kun hvis der findes $u = (u_1, \dots, u_m) \in \mathbf{R}^m$ så (x, u) opfylder *Karush–Kuhn–Tucker betingelserne*:

KKT betingelser

(a) $u \geq 0$ eller $u_1 \geq 0, \dots, u_m \geq 0$ (positivitet)
(b) $g(x) \leq b$ eller $g_1(x) \leq b_1, \dots, g_m(x) \leq b_m$ (mulig løsning)
(c) $u^t(g(x) - b) = 0$ eller $u_1(g_1(x) - b_1) = 0, \dots, u_m(g_m(x) - b_m) = 0$ (komplementær slæk)
(d) $Df(x) = u^t Dg(x)$ eller $Df(x) = u_1 Dg_1(x) + \cdots + u_m Dg_m(x)$ (optimalitet)

Vi reformulerer Karush–Kuhn–Tuckers sætninger:

SÆTNING 6.9 (**KKTNB** II).

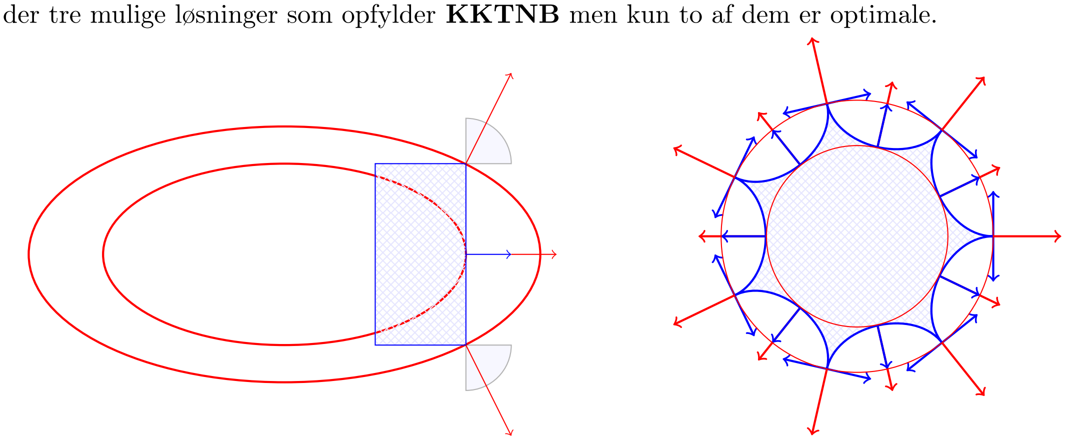
$$O(P) \subseteq \{x \in M(P) \mid x \textit{ opfylder ikke } \mathbf{LICQ}\} \cup \{x \in A \mid \exists u \in \mathbf{R}^m \colon (x, u) \textit{ opfyder } \mathbf{KKT\ betingelser}\}$$

SÆTNING 6.10 (**KKTTB II**). *Antag at A er en åben* konveks *delmængde af $\mathbf{R}^n$. Så er*

$$\{x \in M(P) \mid \exists u \in \mathbf{R}^m \colon (u, x) \textit{ opfyder } \mathbf{KKT\ betingelser} \textit{ og } f - u^t g \textit{ er konkav}\} \subseteq O(P)$$

I praksis er det så godt som umuligt at bruge **KKTNB** til at finde $O(P)$ da det kræver en analyse af de $2^m = \sum_{0 \leq k \leq m} \binom{m}{k}$ tilfælde med k aktive bibetingelser for $k = 0, \dots, m$.

EKSEMPEL 6.11. Der kan være mulige løsninger som opfylder **KKTNB** uden at være optimale. I eksemplet til venstre er der tre mulige løsninger som opfylder **KKTNB** men kun to af dem er optimale.



Eksemplet til højre viser at der også kan være optimale løsninger som ikke opfylder Karush–Kuhn–Tucker betingelserne, dvs ikke ligger i $\{x^* \in M(P) \mid \nabla f(x^*) \in \text{cone}(\{\nabla g_i(x^*) \mid g_i(x^*) = 0\})\}$, og derfor ikke opfylder **LICQ**.

EKSEMPEL 6.12. Lad $f\colon \mathbf{R} \to \mathbf{R}$ være en C^1-funktion af en variabel. Problemet

$$(P) \qquad \text{Maksimér } f \text{ under bibetingelser } a - x \leq 0,\ x - b \leq 0$$

går ud på at finde maksimum for f på intervallet $[a, b] = M(P)$ hvor $a < b$. Gradienterne for de aktive bibetingelser er lineært uafhængige da højst en bibetingelse er aktiv i et givet punkt. Hvis f har maksimum i $x^* \in [a, b]$ så findes $u_1, u_2 \in \mathbf{R}$ så

- $u_1 \geq 0$, $u_2 \geq 0$
- $a \leq x \leq b$
- $u_1(a - x) = 0$, $u_2(x - b) = 0$
- $f'(x^*) = -u_1 + u_2$

Det betyder at enten er x^* et indre punkt (og $u_1 = 0$, $u_2 = 0$) med $f'(x^*) = 0$ eller $x^* = a$ er det venstre endepunkt (og $u_2 = 0$) og $f'(a) \leq 0$ eller $x^* = b$ er det højre endepunkt (og $u_1 = 0$) og $f'(b) \geq 0$.

BEMÆRKNING 6.13 (Minimeringsproblemet).

$$(\text{P}) \qquad \text{Minimér } f \text{ under bibetingelserne } g_1 \leq 0, \ldots, g_m \leq 0$$

løses ved at maksimere $-f$ under bibetingelserne $g_1 \leq 0, \ldots, g_m \leq 0$. Hvis x^* opfylder **LICQ** og er en minimator, så findes ifølge Sætning 6.7 en vektor $u \in \mathbf{R}^m$ så (x^*, u) opfylder Karush–Kuhn–Tucker betingelserne

(a) $u \geq 0$ (positivitet)
(b) $g(x^*) \leq 0$ (mulig løsning)
(c) $u^t g(x^*) = 0$ (komplementær slæk)
(d) $\nabla(f)(x^*) = -u^t \nabla g(x^*)$ (optimalitet)

Og hvis (x^*, u) opfylder ovenstående betingelser og $A \to \mathbf{R}\colon x \to f(x) + u^t g(x)$ er konveks så er x^* en løsning til minimeringsproblemet.

Karush–Kuhn–Tuckers Sætning indeholder både optimering uden bibetingelser (Sætning 4.2) og Lagranges Sætning (Sætning 5.4) som specialtilfælde ($g_1 \leq b_1, \ldots, g_m \leq b_m, -g_1 \leq -b_1, \ldots, -g_m \leq -b_m$).
[Vi ser lige bort fra at LICQ ikke er opfyldt - se [4].]

7. Det generelle optimeringsproblem

Lad

- $I = \{1, \ldots, m\}$, $I^* \subseteq I$
- $J = \{1, \ldots, n\}$, $J^* \subseteq J$
- A er en åben delmængde af $\mathbf{R}^J$
- $f\colon A \to \mathbf{R}$, $g = (g_i)_{i \in I}\colon A \to \mathbf{R}^I$

Vi sætter

$$g = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} : \mathbf{R}^n \to \mathbf{R}^m \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbf{R}^m$$

og vi skriver $[I^*]g(x)$ for $(g_i(x))_{i \in I^*}$, $[I^*]b = (b_i)_{i \in I^*}$ og $[J^*]x = (x_j)_{j \in J^*}$.

Vi ser først på optimeringsproblemet

$$(\text{P}) \qquad \text{Maksimér } f(x) \text{ under bibetingelserne } [I^*]g(x) \leq [I^*]b,\ [I - I^*]g(x) = [I - I^*]b$$

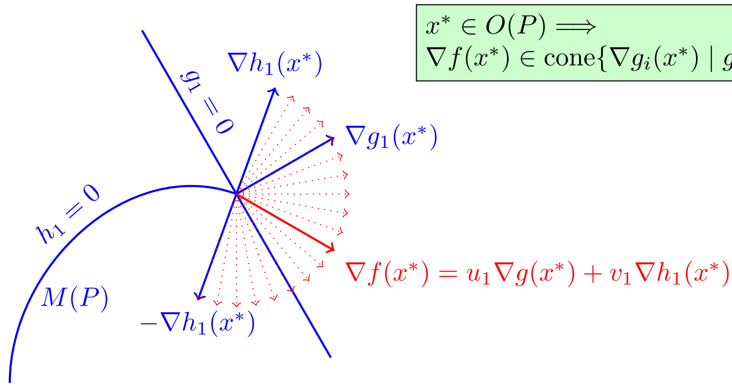
med både uligheds- og lighedsbibetingelser.

PROPOSITION 7.1. *Lad $x \in M(P)$ være en optimal løsning til (P) som opfylder* **LICQ**. *Så findes en vektor $y \in \mathbf{R}^m$ så*

(a) $y^t[I^*] \geq 0$
(b) $[I^*]g(x) \leq [I^*]b$, $[I - I^*]g(x) = [I - I^*]b$
(c) $y^t(g(x) - b) = 0$
(d) $Df(x) = y^t Dg(x)$

Generel Karush–Kuhn–Tucker

$x^* \in O(P) \Longrightarrow$
$\nabla f(x^*) \in \text{cone}\{\nabla g_i(x^*) \mid g_i(x^*) = 0\} + \text{span}\{\nabla h_k(x^*)\}$



Figur 5. **KKTNB**

Det generelle maksimeringsproblem er

(P) Maksimér $f(x)$ under bibetingelserne $[I^*]g(x) \leq [I^*]b$, $[I - I^*]g(x) = [I - I^*]b$, $[J^*]x \geq 0$

De *mulige* løsninger

$$M(P) = \{x \in A \mid [I^*]g(x) \leq [I^*]b, \quad [I = I^*]g(x) = [I - I^*]b, \quad [J^*]x \geq 0\}$$

er de punkter i A som opfylder alle bibetingelser samt fortegnskrav, og de *optimale* løsninger

$$O(P) = \{x^* \in M(P) \mid f(x^*) \geq f(x) \text{ for alle } x \in M(P)\}$$

er de mulige løsninger med størst mulig f-værdi.

Vi siger at en mulig løsning $x \in M(P)$ opfylder **LICQ** (Linear Independence Constraint Qualification) hvis gradienterne for de aktive bibetingelser er lineært uafhængige i x, dvs hvis

$$\{\nabla g_i(x) \mid i \in I, g_i(x) = b_i\} \cup \{-\nabla x_j(x) \mid j \in J^*, x_j^* = 0\}$$

er en lineært uafhængig mængde af n-vektorer.

Sætning 7.2 (Karush–Kuhn–Tucker nødvendig betingelse **KKTNB**). [**1**, Propositions 3.3.1, 3.3.16] [**6**, Chp 2.4] [**5**] [**8**, Thm 8.10.1] *Antag at A er en åben delmængde af $\mathbf{R}^n$ og at f, $g_1, \ldots, g_m$ er C^1-funktioner defineret på A. Hvis $x \in M(P)$ er en optimal løsning til programmet (P) som opfylder* **LICQ** *så findes en vektor $y \in \mathbf{R}^m$ sådan at (x,y) opfylder KKT-betingelserne:*

(a) $[I^*]g(x) \leq [I^*]b$, $[I - I^*]g(x) = [I - I^*]b$, $[J^*]x \geq 0$ *(mulig løsning)*
(b) $y^t Dg(x)[J^*] \geq Df(x)[J^*]$, $y^t Dg(x)[J - J^*] = Df(x)[J - J^*]$, $y^t[I^*] \geq 0$ *(optimalitet og positivitet)*
(c) $y^t(g(x) - b) = 0$, $(Df(x) - y^t Dg(x))x = 0$ *(komplementær slæk)*

Bevis. Fra (versionen i Proposition 7.1) af **KKTNB** med uligheder og ligheder i bibetingelserne og med de ekstra bibetingelser $-x_j \leq 0$ for $j \in J^*$, får vi at der findes $y \in \mathbf{R}^m$ og $v \in \mathbf{R}^n$ så

(a) $y^t[I^*] \geq 0$ og $v^t[J^*] \geq 0$ (positivitet)
(b) $[I^*]g(x) \leq [I^*]b$, $[I - I^*]g(x) = [I - I^*]b$, $[J^*]x \geq 0$ (mulig løsning)
(c) $y^t(g(x) - b) = 0$ og $v^t[J^*][J^*]x = 0$ (komplementær slæk)
(d) $Df(x) = y^t Dg(x) - v^t[J^*][J^*]I_n$ (optimalitet)

Vi reorganiserer nu denne information. Fra (d), $y^t Dg(x)[J^*] = Df(x)[J^*] + v^t[J^*][J^*]I_n[J^*] \geq Df(x)[J^*]$ og $y^t Dg(x)[J - J^*] = Df(x)[J - J^*] + v^t[J^*][J^*]I_n[J - J^*] = Df(x)[J - J^*]$. Fra (a), $y^t[I^*] \geq 0$. Det er punkt (b) i sætningen. Fra (c), $y^t(g(x) - b) = 0$. Fra (c) og (d), $0 = -v^t[J^*][J^*]x = -v^t[J^*][J^*]I_n x = (Df(x) - y^t Dg(x))x$. Det er punkt (c) i sætningen. Punkt (a) i sætningen er (b). □

Sætning 7.3 (Karush–Kuhn–Tucker tilstrækkelig betingelse **KKTTB**). *Hvis $x \in M(P)$, $y \in \mathbf{R}^m$ og*

- *A er konveks*
- *KKT-betingelserne (a)–(c) fra Sætning 7.2 er opfyldt for (x, y)*
- *$f - y^t g$ er en konkav funktion på den konvekse mængde A*

så er x optimal.

Disse generelle sætninger dækker Sætning 4.1, 5.4, 6.7 og Sætning 4.2, 5.5, 6.10 som specialtilfælde.

EKSEMPEL 7.4 (Neoklassisk model). I et marked med n goder til priser $p^t = (p_1, \dots, p_n)$ står en forbruger med indkomst I overfor problemet

$$(P(I)) \qquad \text{Maksimér } U(x) \text{ under bibetingelserne } p^t x \le I,\ x \ge 0$$

hvor $U(x)$ er nytten ved forbrug $x = (x_1, \dots, x_n)$. Ekstremværdisætningen garanterer at der er et optimalt forbrug. Det er rimeligt at forlange $\frac{\partial U}{\partial x_i} > 0$ og $\frac{\partial^2 U}{\partial x_i^2} < 0$. Hvis vi antager at der findes et optimalt forbrug med $x > 0$, giver (c) at $DU(x) = y^t p^t$ og $p^t x = I$ da $y > 0$. Det følger at $I = p^t x = \frac{DU(x)x}{y}$ eller at Lagrangemultiplikatoren $y = \frac{DU(x)x}{I}$.

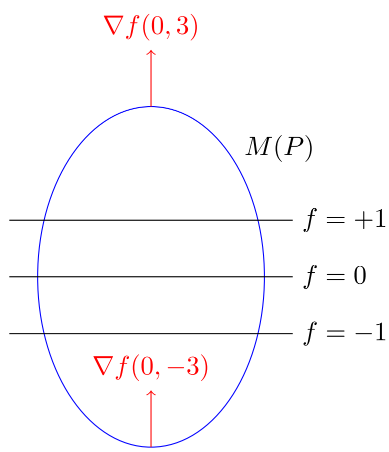
8. Eksempler

På Skagen Nordstrand, det nordligste punkt i Danmark, står kompasnålen vinkelret på kysten.

EKSEMPEL 8.1 (Lagrange sætning 5.4).

$$\text{(P)} \qquad \text{Maksimér } f(x_1, x_2) = x_2 \text{ under bibetingelsen } g(x_1, x_2) = \frac{1}{4}x_1^2 + \frac{1}{9}x_2^2 - 1 = 0$$

Problemet er at finde det højeste punkt på ellipsen $M(P) = g_1^{-1}(1)$. Den eneste optimale løsning er $x^* = (0, 3)$. Lad os nu se om **LNB** og **LTB** giver det rigtige svar.



Gradienterne for objektfunktionen og bibetingelsen er

$$\nabla f = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \nabla g = \begin{pmatrix} \frac{1}{2}x_1 \\ \frac{2}{9}x_2 \end{pmatrix}$$

$M(P)$, mængden af mulige løsninger, er en ellipse. Bibetingelsens gradient er lineært uafhængig, dvs $\neq 0$, på hele $M(P)$. **LNB** fra Sætning 5.4 giver derfor at

$$O(P) \subseteq \{(x_1^*, x_2^*) \in M(P) \mid \exists u^* \colon \begin{pmatrix} 0 \\ 1 \end{pmatrix} = u^* \begin{pmatrix} \frac{1}{2}x_1^* \\ \frac{2}{9}x_2^* \end{pmatrix}\}$$

Vektorerne i mængden på højresiden er kandidater til optimale løsninger. Efter en lille overvejelse ser vi at den består af de to punkter $(0, \pm 3)$. I disse to punkter er

$$\nabla f(0, -3) = -\frac{3}{2}\nabla g(0, -3), \qquad \nabla f(0, 3) = \frac{3}{2}\nabla g(0, 3),$$

Vi ved altså nu at $O(P) \subseteq \{(0, -3), (0, +3)\}$. Men $(0, -3)$ er ikke optimalt fordi $f(0, -3) = -3 < 3 = f(0, 3)$. Altså er $O(P) \subseteq \{(0, +3)\}$. Det betyder at der er enten ingen eller præcis ét optimal løsning. Hvordan kan vi nu argumentere for at $(0, 3)$ faktisk er optimal? Her er to muligheder:

- Da ellipsen $M(P)$ er kompakt og f er kontinuert, så har f en største værdi på $M(P)$, dvs at $O(P) \neq \emptyset$.
- Da $f - \frac{3}{2}g$ er konkav og $\nabla(f - \frac{3}{2}g)(0, 3) = 0$, så er $(0, 3)$ optimal ifølge **LTB** fra Sætning 5.5.

EKSEMPEL 8.2 (Lagrange sætning 5.4).

$$\text{(P)} \qquad \text{Maksimér } x_1^2 + x_2^2 + x_3^2 \text{ under bibetingelserne } x_1^2 + x_2^2 + 4x_3^2 = 1,\ x_1 + 3x_2 + 2x_3 = 1$$

De mulige løsninger $M(P) = \{x_1^2+x_2^2+4x_3^2 = 1, x_1+3x_2+2x_3 = 1\}$ er kompakt, så den kontinuerte objektfunktion $f(x_1,x_2,x_3) = x_1^2+x_2^2+x_3^2$ har både et maksimum og et minimum på $M(P)$. Gradienterne

$$\nabla g_1 = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 8x_3 \end{pmatrix}, \quad \nabla g_2 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

er lineært uafhængige i planen $\{(x_1,x_2,x_3) \mid x_1+3x_2+2x_3 = 1\}$ og specielt også i hele $M(P)$. **LNB** siger derfor at $O(P)$ er en delmængde af mængden af de $(x_1,x_2,x_3) \in M(P)$ for hvilke der findes $u_1,u_2,u_3 \in \mathbf{R}$ sådan at Lagrange betingelserne

(a) $x_1^2+x_2^2+4x_3^2 = 1$, $x_1+2x_2+3x_3 = 1$

(b) $\begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix} = u_1 \begin{pmatrix} 2x_1 \\ 2x_2 \\ 8x_3 \end{pmatrix} + u_2 \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$

er opfyldt. Der er to muligheder,

$$(x_1,x_2,x_3,u_1,u_2) = (\frac{3}{\sqrt{10}}, -\frac{1}{\sqrt{10}}, 0, 1, 0), \qquad (x_1,x_2,x_3,u_1,u_2) = (\frac{1}{5}\zeta, \frac{3}{5}\zeta, \zeta, \frac{7}{22}, -\frac{3}{11}\zeta), \quad \zeta = \sqrt{\frac{5}{22}}$$

som må være maksimum ($f = 1$) og minimum ($f = \frac{35}{110}$) af f på $M(P)$.

Eksempel 8.3. [8, Eksempel 2, p 261]

(P) Maksimér $f(x,y) = x^2+2y$ under bibetingelserne $g_1(x,y) = x^2+y^2 \le 5$, $g_2(x,y) = -y \le 0$

Den kontinuerte objektfunktion $f(x,y) = x^2+2y$ har både et maksimum og et minimum på den kompakte mængde $M(P) = \{x^2+y^2 \le 5, \quad y \ge 0\}$ (se Figur 6). Gradienterne er

$$\nabla f = \begin{pmatrix} 2x \\ 2 \end{pmatrix}, \quad \nabla g_1 = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \quad \nabla g_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Det er ikke svært at se at **LICQ** er opfyldt for all mulige løsninger. **KKTNB** siger derfor at for enhver optimal løsning, $(x,y) \in O(P)$, findes $u,v \in \mathbf{R}$ så

(a) $u \ge 0$, $v \ge 0$

(b) $x^2+y^2 \le 5$, $y \ge 0$

(c) $u(x^2+y^2-5) = 0$, $vy = 0$

(d) $\begin{pmatrix} 2x \\ 2 \end{pmatrix} = u \begin{pmatrix} 2x \\ 2y \end{pmatrix} + v \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

Da der er 2 bibetingelser kan analysen deles ind i $2^2 = 4$ tilfælde: Ingen bibetingelser er aktive, netop en bibetingelse er aktiv, begge bibetingelser er aktive. I dette tilfælde finder vi at (a)–(d) har de tre løsninger $(x,y,u,v) = (\pm 2, 1, 1, 0), (0, \sqrt{5}, 1/\sqrt{5}, 0)$. Vi ved nu at de optimale løsninger ligger blandt de tre punkter

$$\{(-2,1),(2,1),(0,\sqrt{5})\}$$

Da $f(\pm 2, 1) = 6$ og $f(0,\sqrt{5}) = 2\sqrt{5} < 6$ er de optimale løsninger de to punkter $(\pm 2, 1)$.

Eksempel 8.4 (De nødvendige Karush–Kuhn–Tucker betingelser er ikke tilstrækkelige).

(P) Maksimer $x_1^2+x_2^2$ under bibetingelsen $x_1 \le 1$

De mulige løsninger $M(P) = \{(x_1,x_2) \in \mathbf{R}^2 \mid x_1 \le 1\}$ er alle punkter med første koordinat højst 1. Da gradienten $\nabla g(x_1,x_2) = (1,0)$ ikke er nulvektoren, er **LICQ** opfyldt i hele $M(P)$. Objektfunktionens gradient er $\nabla f(x_1,x_2) = (2x_1, 2x_2)$. **KKTNB** siger at $O(P) \subseteq \{(0,0),(1,0)\}$. (Antag nemligt at (x_1,x_2) er en optimal løsning. **KKTNB** siger at hvis $x_1 < 1$ så er $(x_1,x_2) = (0,0)$, og hvis $x_1 = 1$ så er $(2, 2x_2) = u(1,0)$ for et $u \ge 0$.) Men det er oplagt at der ikke findes optimale løsninger, da $f(x_1,x_2)$ er ikke opad begrænset på $M(P)$.

Eksempel 8.5 (Eksamen 07/08 Opg 3). Lad Q være en kvadratisk positiv semidefinit $(n \times n)$-matrix, A en $(m \times n)$-matrix og $b \in \mathbf{R}^m$, $c \in \mathbf{R}^n$ vektorer. Betragt det konvekse kvadratiske optimeringsproblem

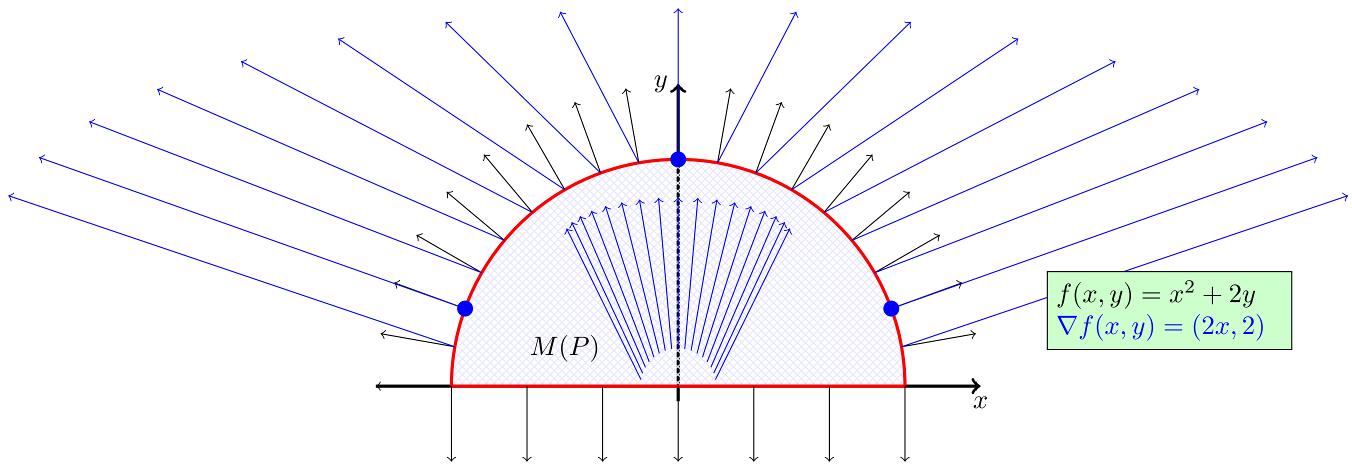
(P) Minimer $c \cdot x + \frac{1}{2}x^t Q x$ under bibetingelserne $Ax \le b$

Karush–Kuhn–Tucker betingelserne er

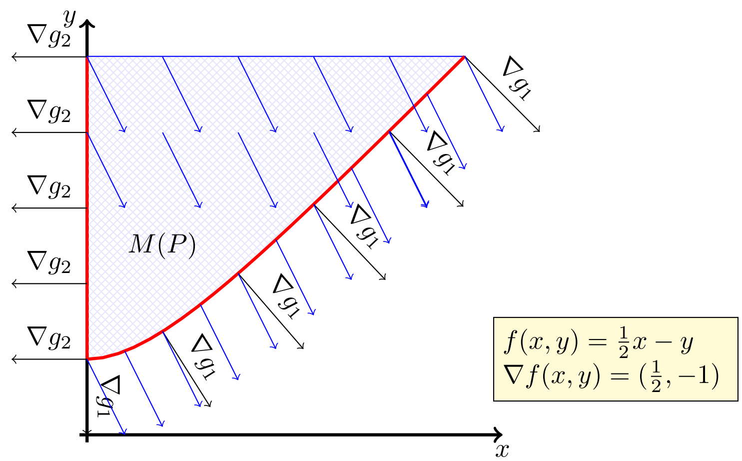
(1) $u \ge 0$

(2) $Ax \le b$

(3) $u \cdot (Ax - b) = 0$



FIGUR 6. Eksempel 8.3



FIGUR 7. Eksempel 8.6

(4) $c + Qx + A^t u = 0$

Ifølge Sætning 6.7 og Sætning 6.10 er x^* en løsning til (P) hvis og kun hvis der findes u^* så (x^*, u^*) opfylder Karush–Kuhn–Tucker betingelserne.

EKSEMPEL 8.6 (Konveks optimering). [8, Eksempel 1, p 260]

$$\text{(P)} \qquad \text{Maksimér } \frac{1}{2}x - y \text{ under bibetingelserne } x + e^{-x} - y \le 0,\ -x \le 0$$

Bibetingelsernes gradienter er altid lineært uafhængige. Karush–Kuhn–Tucker betingelserne er

(1) $u \ge 0$, $v \ge 0$
(2) $x + e^{-x} \le y$, $x \ge 0$
(3) $u(x + e^{-x} - y) = 0$, $vx = 0$
(4) $\begin{pmatrix} 1/2 \\ -1 \end{pmatrix} - u \begin{pmatrix} 1 - e^{-x} \\ -1 \end{pmatrix} - v \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Figur 7 viser at det eneste punkt som opfylder betingelserne er $(x, y, u, v) = (\ln(2), \ln(2) + \frac{1}{2}, 1, 0)$. (Det følger også af betingelserne: Fra (4) får vi $u = 1$ så (x, y) ligger på grafen for $x + e^{-x}$. Hvis $x > 0$ er $v = 0$ så $\frac{2}{2} - (1 - e^{-x}) = 0$ og $x = \ln(2)$.) Da $f(x, y) - g_1(x, y) = -\frac{1}{2}x - \exp(-x)$ er konkav er $(x, y) = (\ln(2), \ln(2) + \frac{1}{2})$ faktisk en optimal løsning til (P).

EKSEMPEL 8.7 (Lagrange multiplikatorer er skyggepriser for bibetingelser). [8, p 245, §8.6] I maksimeringsproblemet

$$\text{(P}(b)) \qquad \text{Maximer } f \text{ under bibetingelser } g_1 = b_1, \dots, g_m = b_m$$

afhænger bibetingelserne af de (exogene) variable $b = (b_1, \ldots, b_m)$. Hvad sker der med de optimale løsninger når den exogene variable b ændres?

Lagrangebetingelserne for optimeringsproblemerne P(b) er $F(x, u, b) = 0$ hvor $F\colon \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}^m \times \mathbf{R}^n$ er funktionen

$$F(x,u,b) = \begin{pmatrix} g(x) - b \\ \nabla\big(f - \sum u_j g_j\big)(x) \end{pmatrix} = \begin{pmatrix} g_1(x) - b_1 \\ \vdots \\ g_m(x) - b_m \\ \frac{\partial}{\partial x_1}\big(f - \sum u_j g_j\big)(x) \\ \vdots \\ \frac{\partial}{\partial x_n}\big(f - \sum u_j g_j\big)(x) \end{pmatrix} = \begin{pmatrix} g(x) - b \\ \nabla\big(f - \sum u^t g\big)(x) \end{pmatrix}$$

Vi kan derfor sige

$$(x,u) \text{ opfylder } \mathbf{LNB} \text{ for } \mathrm{P}(b) \iff F(x,u,b) = 0$$

Kan vi finde $(x, u) = (x(b), u(b))$ i disse ligninger som en funktion af de exogene variable b? Ja, det kan vi ifølge Implicit Funktion Sætning forudsat at Jacobi matricen for F efter de endogene variable (x, u) er invertibel. Vi vil da have

$$\frac{\partial F}{\partial(x,u)} \frac{\partial \begin{pmatrix} x \\ u \end{pmatrix}}{\partial b} + \frac{\partial F}{\partial b} = 0 \quad \text{eller} \quad \frac{\partial \begin{pmatrix} x \\ u \end{pmatrix}}{\partial b}(b) = -\frac{\partial F}{\partial(x,u)}(x^*, u^*, b)^{-1} \frac{\partial F}{\partial b}(x^*, u^*, b) \tag{8.8}$$

hvor

$$\frac{\partial \begin{pmatrix} x \\ u \end{pmatrix}}{\partial b} = \begin{pmatrix} \frac{\partial x}{\partial b} \\ \frac{\partial u}{\partial b} \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial b_1} & \cdots & \frac{\partial x_1}{\partial b_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial b_1} & \cdots & \frac{\partial x_n}{\partial b_n} \\ \frac{\partial u_1}{\partial b_1} & \cdots & \frac{\partial u_1}{\partial b_n} \\ \vdots & & \vdots \\ \frac{\partial u_m}{\partial b_1} & \cdots & \frac{\partial u_m}{\partial b_n} \end{pmatrix}$$

De partielle afledte efter de endogene variable (x, u)

$$\frac{\partial F}{\partial(x,u)} = \frac{\partial \begin{pmatrix} g - b \\ \nabla\big(f - \sum u^t g\big) \end{pmatrix}}{\partial(x,u)} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} & 0 & \cdots & 0 \\ \frac{\partial^2}{\partial x_1 \partial x_1}\big(f - \sum u_j g_j\big) & \cdots & \frac{\partial^2}{\partial x_n \partial x_1}\big(f - \sum u_j g_j\big) & -\frac{\partial g_1}{\partial x_1} & \cdots & -\frac{\partial g_n}{\partial x_1} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2}{\partial x_1 \partial x_n}\big(f - \sum u_j g_j\big) & \cdots & \frac{\partial^2}{\partial x_n \partial x_n}\big(f - \sum u_j g_j\big) & -\frac{\partial g_1}{\partial x_n} & \cdots & -\frac{\partial g_n}{\partial x_n} \end{pmatrix} \tag{8.9}$$

$$= \begin{pmatrix} \frac{\partial g}{\partial x} & 0 \\ H\big(f - \sum u^t g\big) & -(\frac{\partial g}{\partial x})^t \end{pmatrix}$$

er en kvadratisk $(n + m) \times (n + m)$ matrix. De partielle afledte efter de exogene variable b

$$\frac{\partial F}{\partial b} = -E_{m+n}[1, m] = \begin{pmatrix} -E_m \\ 0_{n,m} \end{pmatrix} \tag{8.10}$$

er -1 gange de m første søjler i $(m + n) \times (m + n)$ enhedsmatricen E_{m+n}.

I dette tilfælde giver ligningerne (8.8), hvis vi indsætter udtrykkene fra (8.9) og (8.10), at

$$\begin{pmatrix} \frac{\partial g}{\partial x} & 0 \\ H\big(f - u^t g\big) & -(\frac{\partial g}{\partial x})^t \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial b} \\ \frac{\partial u}{\partial b} \end{pmatrix} = \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix}$$

som giver at

$$\frac{\partial g}{\partial x}\frac{\partial x}{\partial b} = E_m, \qquad H\big(f - u^t g\big)\frac{\partial x}{\partial b} = \Big(\frac{\partial g}{\partial x}\Big)^t \frac{\partial u}{\partial b}$$

og at $(m+n)\times m$ matricen

$$\frac{\partial\begin{pmatrix}x\\u\end{pmatrix}}{\partial b}=-\left(\frac{\partial F}{\partial(x,u)}\right)^{-1}(-E_{m+n}[1,m])=\left(\frac{\partial F}{\partial(x,u)}\right)^{-1}[1,m]=\begin{pmatrix}\frac{\partial g}{\partial x} & 0\\ H\big(f-u^tg\big) & -(\frac{\partial g}{\partial x})^t\end{pmatrix}^{-1}[1,m]$$

er de m første søjler i den inverse af $(m+n)\times(m+n)$-matricen $\left(\frac{\partial F}{\partial(x,u)}\right)^{-1}$.

Hvordan afhænger den optimale løsning $(x^*(b),u^*(b))$ **af de exogene variable** b**?:**

$$\begin{pmatrix}x^*(b+\Delta b)\\u^*(b+\Delta b)\end{pmatrix}\approx\begin{pmatrix}x^*(b)\\u^*(b)\end{pmatrix}+\frac{\partial\begin{pmatrix}x\\u\end{pmatrix}}{\partial b}\Delta b=\begin{pmatrix}x^*(b)\\u^*(b)\end{pmatrix}+\begin{pmatrix}\frac{\partial g}{\partial x} & 0\\ H\big(f-u^tg\big) & -(\frac{\partial g}{\partial x})^t\end{pmatrix}^{-1}[1,m]\Delta b$$

er en approximativ løsning til Lagrangeligningerne.

Hvordan ændres $f(x^*(b))$ **når** b **ændres?:**

$$f(x^*(b+\Delta b))\approx f(x^*(b))+\frac{\partial f(x^*(b))}{\partial b}\Delta b=f(x^*(b))+u^*(b)^t\Delta b$$

fordi

$$\frac{\partial f(x(b))}{\partial b}=\frac{\partial f}{\partial x}\frac{\partial x}{\partial b}=\sum_{j=1}^m u_j\frac{\partial g_j}{\partial x}\frac{\partial x}{\partial b}=\sum_{j=1}^m u_j[j]\frac{\partial g}{\partial x}\frac{\partial x}{\partial b}=\sum_{j=1}^m u_j\,[j]E_m=\begin{pmatrix}u_1 & u_2 & \cdots & u_m\end{pmatrix}=u^t$$

da Lagrangebetingelsen $\frac{\partial f}{\partial x}=\sum u_j\frac{\partial g_j}{\partial x}$ er opfyldt. (I denne formel står $[j]A$ for række j i matrix A).

Økonomer tolker $x^*(b)$ som den optimale strategi, der giver den optimale nytteværdi $f(x^*(b))$, under givne exogene variable b. Formlen for $x^*(b+\Delta b)$ informerer om ændringen i den optimale strategi og formlen for $f(x^*(b+\Delta b))$ om ændringen i den optimale nytte under en lille ændring Δb i de exogene variable.

Lagrange multiplikatorerne $u^*(b)$ kaldes for bibetingelsernes skyggepriser. Lad $p=(p_1,\dots,p_m)$ være de faktiske priser for de m exogene variable $b=(b_1,\dots,b_m)$. Det kræver en investering på $p^t\Delta b$ at ændre b til $b+\Delta b$ og det giver en ændring i nytteværdi som er $u^*(b)^t\Delta b$. Investeringsgevinsten er altså

$$(u^*(b)-p)^t\Delta b$$

Investeringen er profitabel hvis dette tal er positivt.

Læg mærke til at $u^*(b)$ for komparativ statistik er mere interesant end $x^*(b)$ fordi systemet af sig selv finder frem til $x^*(b)$ mens $u^*(b)$ fortæller om nytteværdiens følsomhed for ændringer i de exogene variable. Det kan feks være et politisk ønske at ændre på $f(x^*(b))$. Den størst mulige effekt opnås ved at tage ændringen Δb i samme retning som $u^*(b)$.

Værdifunktionen er

$$V(b_1,\dots,b_m)=\sup\{f(x_1,\dots,x_n)\mid g(x_1,\dots,x_n)=(b_1,\dots,b_m)\}$$

og de marginale værdifunktioner

$$\frac{\partial V}{\partial b_j}=u_j$$

er lig med Lagrangemultiplikatorerne som derfor udpeger den bedste strategi for at øge værdifunktionen.

9. Matematisk økonomiske modeller

I økonomisk teori antager man at vi har opnået et maksimum og man ønsker at sige noget om hvad der der sker når de exogene variable ændres. Vi antager et maksimum er blevet antaget og vi ønsker at sige noget om hvilke konsekvenser vi kan drage af det.

Eksempel 9.1. [8, Eksempel 2 p 271] Et elektricitetsværk inddeler året i n perioder. Lad x_i være produktionen i periode i og p_i prisen for elektricitet i den periode, $1\le i\le n$. Værkets kapacitet k er konstant over hele året. Værkets udgifter består af driftsomkostninger $C(x)=C(x_1,\dots,x_n)$, som kun afhænger af produktionsmønsteret x, og kapacitetsomkostninger $D(k)$, som kun afhænger af kapaciteten k. Værkets årlige profit er

$$\pi(x,k)=p\cdot x-C(x)-D(k)$$

og profitmaksimeringsproblemet er

$$\text{(P)}\qquad \text{Maksimér } \pi(x,k) \text{ under bibetingelserne } 0\le x_1\le k,\dots,0\le x_n\le k, k\ge 0$$

Der er her $2n+1$ bibetingelser, $x_i - k \leq 0$ og $-x_i \leq 0$ for $1 \leq i \leq n$, og $-k \leq 0$. Hvis (x,k) er et årligt produktionsmønster med maksimal profit så er Karush–Kuhn–Tucker betingelserne opfyldt:

(a) $u_1, \dots, u_n \geq 0$, $v_1, \dots, v_n \geq 0$, $w \geq 0$
(b) $0 \leq x_i \leq k$ for $1 \leq i \leq n$, og $k \geq 0$
(c) $u_i(x_i - k) = 0$ og $v_i x_i = 0$ for $1 \leq i \leq n$, og $wk = 0$
(d) $\nabla\pi(x,k) = \sum_i u_i \nabla(x_i - k) + \sum_i v_i \nabla(-x_i) + w\nabla(-k)$

Ligning (d) siger at

$$p_i - \frac{\partial C}{\partial x_i} = \frac{\partial \pi}{\partial x_i} = u_i - v_i \quad (1 \leq i \leq n) \qquad -\frac{\partial D}{\partial k} = \frac{\partial \pi}{\partial k} = -\sum_i u_i - w$$

Lad os antage at produktionen $x_i > 0$ i alle perioder og at kapaciteten $k > 0$. Så er $v_1 = \cdots = v_n = 0$, $w = 0$ ifølge (c) og derfor

$$p_i - \frac{\partial C}{\partial x_i} = u_i \quad (1 \leq i \leq n), \qquad \frac{\partial D}{\partial k} = \sum_i u_i$$

og også $u_i = 0$ i de perioder hvor $x_i < k$. Vi konkluder at $\frac{\partial C}{\partial x_i} \leq p_i$ i alle perioder, $\frac{\partial C}{\partial x_i} = p_i$ i perioder med fuld kapacitetsudnyttelse, og

$$\frac{\partial D}{\partial k} = \sum_{1 \leq i \leq n} (p_i - \frac{\partial C}{\partial x_i}) = \sum_{\substack{1 \leq i \leq n \\ x_i = k}} (p_i - \frac{\partial C}{\partial x_i})$$

Dvs at i perioder hvor produktionen ligger under den maksimale kapacitet så dækker salgsprisen lige akkurat de marginale driftsomkostninger. I en periode i med spidsbelastning overstiger salgsprisen de marginale driftsomkostninger med u_i. Summen af disse overskridelser over alle spidsbelastningsperioder er lig med den marginale kapacitetsomkostning.

Eksempel 9.2 (Fortolkning af Lagrange multiplikatoren i en økonomi med to goder). Lise lever af hamburgere og cola. Prisen for en hamburger er p_1 og prisen for en cola er p_2. Lises indkomst er I.

Sæt $A = \{(x_1, x_2) \mid x_1 > 0, x_2 > 0\}$ og lad funktionerne $f, g \colon A \to \mathbf{R}$ være funktionerne $f(x_1, x_2) = a_1 \ln(x_1) + a_2 \ln(x_2)$ og $g(x_1, x_2) = p_1 x_1 + p_2 x_2$. Funktionen $f(x_1, x_2)$ er (Cobb–Douglas) nytteværdien af, og funktionen $g(x_1, x_2)$ er prisen, for x_1 hamburgere og x_2 colaer. Cobb–Douglas funktionen er konkav. De positive konstanter a_1 og a_2 afspejler Lises preferencer for hamburgere eller colaer. Ved at re-skalere f med $\frac{1}{a_1 + a_2}$ om nødvendigt kan vi godt antage at $a_1 + a_2 = 1$. Lise er en prisbevidst forbruger og hun, eller den 'usynlige hånd', har derfor løst optimeringsproblemet

$$(\mathrm{P}(I)) \qquad \text{Maksimér } f(x_1, x_2) \text{ under bibetingelsen } p_1 x_1 + p_2 x_2 = I$$

De mulige løsninger $M(P(I)) = \{(x_1, x_2) \in A \mid p_1 x_1 + p_2 x_2 = I\}$ er et åbent linjestykke. Gradienten for bibetingelsen $\nabla g = (p_1, p_2)$ er $\neq 0$ på hele $M(P(I))$ så **LICQ** er opfyldt for alle mulige løsninger. **LNB** (Sætning 5.4) siger at hvis (x_1, x_2) er optimal så findes en Lagrange multiplikator $u \in \mathbf{R}$ så

(1) $p_1 x_1 + p_2 x_2 - I = 0$
(2) $\begin{pmatrix} \frac{a_1}{x_1} \\ \frac{a_2}{x_2} \end{pmatrix} = u \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$

Af (2) og (1) får vi

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{u} \begin{pmatrix} \frac{a_1}{p_1} \\ \frac{a_2}{p_2} \end{pmatrix} \quad \text{og} \quad I = p_1 x_1 + p_2 x_2 = \frac{1}{u}(a_1 + a_2) = \frac{1}{u}$$

Den eneste mulige optimale løsning som funktion af indkomsten I og Lagrange multiplikator er derfor

$$\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix}(I) = I \begin{pmatrix} \frac{a_1}{p_1} \\ \frac{a_2}{p_2} \end{pmatrix}, \quad u = \frac{1}{I}$$

LTB siger at dette faktisk er en optimal løsning. Altså er $O(P) = \{(Ia_1/p_1, Ia_2/p_2)\}$.

Hvordan ændres det optimale forbrug $(x_1^*(I), x_2^*(I))$ når I ændres?:

$$\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix}(I + \Delta I) \approx \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix}(I) + \begin{pmatrix} \frac{a_1}{p_1} \\ \frac{a_2}{p_2} \end{pmatrix} \Delta I$$

Hvordan ændres den optimale nytteværdi $f(x^*(I))$ når I ændres?: Kædereglen giver

$$\frac{\mathrm{d}}{\mathrm{d}I}f(x_1^*(I),x_2^*(I)) = \frac{\partial f}{\partial(x_1,x_2)}(x_1^*(I),x_2^*(I))\begin{pmatrix}\frac{\mathrm{d}x_1^*}{\mathrm{d}I}\\ \frac{\mathrm{d}x_s^*}{\mathrm{d}I}\end{pmatrix} = \frac{1}{I}\begin{pmatrix}p_1 & p_2\end{pmatrix}\begin{pmatrix}a_1/p_1\\ a_2/p_2\end{pmatrix} = \frac{1}{I}$$

Den maksimale nytteværdi

$$f(x_1^*(I+\Delta I), x_2^*(I+\Delta I)) \approx f(x_1^*(I),x_2^*(I)) + \frac{1}{I}\Delta I = f(x_1^*(I),x_2^*(I)) + \frac{\Delta I}{I}$$

vokser omvendt proportionalt med I: Den marginale nytteværdi af en lønforhøjelse er omvendt proportional med lønniveauet.

Lagrange multiplikatoren $u^*(I) = \frac{1}{I}$ er i dette eksempel lig med den marginale nytteværdi – det er ikke tilfældigt! (Se Eksempel 8.7.) Nyttefunktionen er

$$V(I) = \sup\{f(x_1,x_2) \mid (x_1,x_2) \in P(I)\}$$

og vi fandt at den marginale nytteærdi til $\frac{\partial V}{\partial I}$ til at være lig med Lagrange multiplikatoren u for optimeringsproblemet $P(I)$.

EKSEMPEL 9.3. Lad a_1 og a_2 være positive konstanter så $a_1 + a_2 = 1$ (Cobb–Douglas). Sæt $f(x_1,x_2) = a_1\ln(x_1) + a_2\ln(x_2)$, $g_1(x_1,x_2) = 2x_1 + x_2$, $g_2(x_1,x_2) = x_1 + 2x_2$ for $x_1 > 0$ og $x_2 > 0$. Vi forestiller os at $f(x_1,x_2)$ er nytteværdien og at $g_1(x_1,x_2)$ og $g_2(x_1,x_2)$ er forbruget af to resurcer (energi og areal) ved produktion af x_1 enheder af produkt nr 1 og x_2 enheder af produkt nr 2. Der er 300 energienheder og 450 arealenheder til rådighed for produktionen.

(P) Maksimér f under bibetingelserne $g_1 \le 300$, $g_2 \le 450$

Hvis a_2 er stor, er det en fordel at producе produkt nr 2 som kræver meget jord. Derfor vil økonomiens begrænsning ligge i mængden af tilgængelig jord (g_2 er aktiv) mens skyggeprisen på energi vil være $u_1 = 0$ da g_1 ikke er aktiv.

Hvis a_2 er lille, dvs a_1 er stor, er det en fordel at produce produkt nr 1 som kræver meget energi. Derfor vil økonomiens begrænsning ligge i mængden af tilgængelig energi (g_1 er aktiv) mens skyggeprisen på jord vil være $u_2 = 0$ da g_2 ikke er aktiv.

Hvis a_2 har en mellemværdi så er det en fordel at producere en blanding af de to produkter.

Karush–Kuhn–Tucker–betingelser

(1) $u_1 \ge 0$, $u_2 \ge 0$
(2) $2x_1 + x_2 \le 300$, $x_1 + 2x_2 \le 450$
(3) $u_1(2x_1 + x_2 - 300) = 0$, $u_2(x_1 + 2x_2 - 450) = 0$
(4) $\frac{a_1}{x_1} - 2u_1 - u_2 = 0$, $\frac{a_2}{x_2} - u_1 - 2u_2 = 0$

Overvej hvad betingelse (4) betyder geometrisk på linjerne $2x_1 + x_2 = 300$ og $x_1 + 2x_2 = 450$ og i deres skæringspunkt, $(50, 200)$. Se Figur 8.

Den økonomiske tolkning af betingelse (3) om komplementær slæk er at hvis energibetingelse g_1 er slæk, $g_1(x_1,x_2) < 0$, så er der stadig energi til rådighed og derfor er skyggeprisen på energi lig 0, dvs $u_1 = 0$.

Ved en systematisk behandling af Karush–Kuhn–Tucker–betingelserne må vi gå igennem de fire muligheder:

$u_1 = 0$, $u_2 = 0$: Begge bibetingelser inaktive. Hverken energi eller jord er fuld udnyttet. Det lyder ikke som et sandsynligt maksimum. Her er $a_1 = 0$ og $a_2 = 0$. Umuligt.

$u_1 = 0$, $u_2 > 0$: Den anden bibetingelse er aktiv. Jorden er er fuldt udnyttet, $x_1 + 2x_2 = 450$. Her er $a_2 \ge 8/9$ stor.

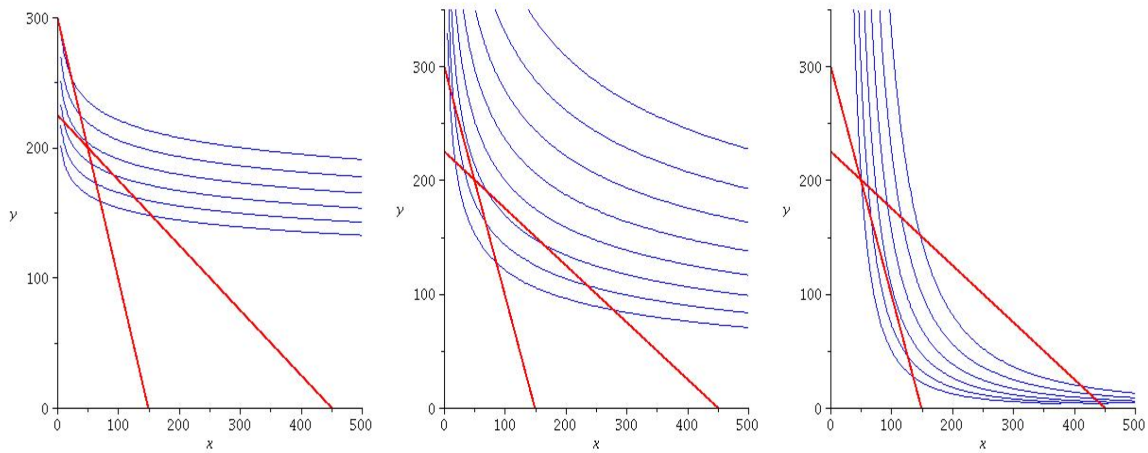
$u_1 > 0$, $u_2 = 0$: Den første bibetingelse er aktiv. Energi er fuldt udnyttet, $2x_1 + x_2 = 300$. Her er $a_2 \le 2/3$ lille.

$u_1 > 0$, $u_2 > 0$: Begge bibetingelser er aktive. Både energi og jord er fuldt udnyttet. Her er $2/3 < a_2 < 8/9$ og $(x_1,x_2) = (50, 200)$.

Hvis feks $a_2 = 11/12$ så er $u_1 = 0$ og $u_2 > 0$. Modellen siger at da $u_2 > 0$ er al jord er udnyttet, men der er stadig energi som ikke er udnyttet. Skyggeprisen på jord er positiv, mens skyggeprisen på energi er $u_1 = 0$ da tilførsel af mere energi ikke vil øge objektfunktionen.

EKSEMPEL 9.4. [8] Ved produktionsstrategi $v = (v_1,\dots,v_n) \ge 0$, afsætningspris $p > 0$ og omkostninger $q = (q_1,\dots,q_n) > 0$ er profitten

$$\pi(p,q,v) = pf(v) - q\cdot v, \qquad p > 0, q > 0, v \ge 0$$



FIGUR 8. $a_2 = 11/12$ ($u_1 = 0$, $u_2 > 0$), $a_2 = 3/4$ ($u_1 > 0$, $u_2 > 0$) og $a_2 = 1/3$ ($u_1 > 0$, $u_2 = 0$)

hvor $f(v)$ er produktionen ved strategi v. Vi antager at f er en C^∞-funktion. Desuden antager vi at for ethvert reelt tal $a > 0$ findes der et tal K_a sådan at $\nabla f(v) \cdot \frac{v}{|v|} < a$ når $|v| > K_a$.

For givne $k > 0$, $c > 0$ og $a > 0$ sæt

$$X_{(k,c)} = \{(p,q) \in \mathbf{R}_{>0} \times \mathbf{R}^n_{>0} \mid 0 < p < k,\ (c,\dots,c) < q\}, \qquad Y_{(k,c)} = \mathbf{R}^n_{\geq 0} \cap \overline{B(0; K_{\frac{c}{k\sqrt{n}}})}$$

Så har funktionen

$$\pi\colon X_{(k,c)} \times Y_{(k,c)} \to \mathbf{R}$$

en kontinuert værdifunktion [8]

$$V(p,q) = \sup_{v \in Y_{(k,c)}} \pi(p,q,v)\colon X_{(k,c)} \to \mathbf{R}$$

Men når $(p,q) \in X_{(k,c)}$ er

$$\max_{v \in Y_{(k,c)}} \pi(p,q,v) = \sup_{v \geq 0} \pi(p,q,v)$$

For at se det, lad $u \geq 0$ være en enhedsvektor. Den afledte i t, hvor $t > K_{\frac{c}{k\sqrt{n}}}$, af funktionen $t \to \pi(p,q,tv) = pf(tv) - tq \cdot v$, er

$$p\nabla f(tu) \cdot u - q \cdot u \leq p\nabla f(tu) \cdot u - \frac{c}{\sqrt{n}} \leq p\frac{c}{k\sqrt{n}} - \frac{c}{\sqrt{n}} \leq \frac{c}{\sqrt{n}} - \frac{c}{\sqrt{n}} = 0$$

Dette viser at $t \to \pi(p,q,tv)$ er aftagende når $t > K_{\frac{c}{k\sqrt{n}}}$.

Da ethvert punkt $(p,q) \in \mathbf{R}_{>0} \times \mathbf{R}^n_{>0}$ ligger i en af de åbne mængder $X_{(k,c)}$, ser vi at værdifunktionen

$$V\colon \mathbf{R}_{>0} \times \mathbf{R}^n_{>0} \to \mathbf{R}, \qquad V(p,q) = \sup_{v \geq 0} \pi(p,q,v),$$

er veldefineret og kontinuert: Der findes en optimal produktionsstrategi og den maksimale profit afhænger kontinuert af salgspris og omkostningspriser. Det er ikke givet at den optimale produktionsstrategi er entydig eller kontinuert afhængig af priserne.

EKSEMPEL 9.5. Cobb–Douglas production function

Google search på nonlinear optimization

KAPITEL 2

Lineær optimering

1. Hvad er lineær optimering?

Hvad er et lineært maksimeringsprogram? Lad

- A være en $m \times n$ matrix
- b en m-vektor
- c en n-vektor
- I^* er en delmængde af $I = \{1, \dots, m\}$
- J^* er en delmængde af $J = \{1, \dots, n\}$

Mængden I vil blive brugt til at indicere problemets bibetingelser, og mængden J til at indicere problemets variable $x = (x_j)_{j \in J}$. Et generelt lineært program har formen

$$\text{(P)} \qquad \text{Max } c^t x \text{ under bibetingelser } \begin{array}{ll} [i]Ax \leq [i]b & i \in I^* \\ [i]Ax = [i]b & i \notin I^* \\ [j]x \geq 0 & j \in J^* \end{array}$$

Bibetingelser indiceret ved I^* er uligheder og variable indiceret ved J^* er ≥ 0. Vi vil også skrive[1]

$$\text{(P)} \qquad \text{Max } c^t x \text{ under bibetingelser } [I^*]Ax \leq [I^*]b,\ [I - I^*]Ax = [I - I^*]b,\ [J^*]x \geq 0$$

for at gøre det lidt mere kompakt.

En *tilladt løsning* eller *mulig løsning* til (P) er en vektor $x \in \mathbf{R}^J$ som opfylder alle bibetingelser. Vi skriver

$$M(P) = \{x \in \mathbf{R}^J \mid [I^*]Ax \leq [I^*]b, \quad [I - I^*]Ax = [I - I^*]b, \quad [J^*]x \geq 0\}$$

for det konvekse polyeder af alle mulige løsninger til (P). Den *optimale værdi* for (P) er

$$\sup(P) = \sup\{c^t x \mid x \in M(P)\}$$

og en *optimal løsning* for (P) er en tilladt vektor $x^* \in M(P)$ sådan at $f(x^*) = \sup(P)$, dvs $f(x^*) \geq f(x)$ for alle $x \in M(P)$. Vi skriver

$$O(P) = \{x^* \in M(P) \mid c^t x^* = \sup(P)\}$$

for mængden af optimale løsninger.

Eksempel 1.1. Her er et eksempel på et generelt lineært program

$$\begin{aligned} &\text{(P)} \quad \text{Maksimér } x_1 + 5x_2 - 2x_3 \text{ under bibetingelser} \\ &-2x_1 + 3x_2 \leq 3 \\ &x_1 + 2x_2 - x_3 = 4 \\ &x_1 \geq 0, \quad x_2 \geq 0 \end{aligned}$$

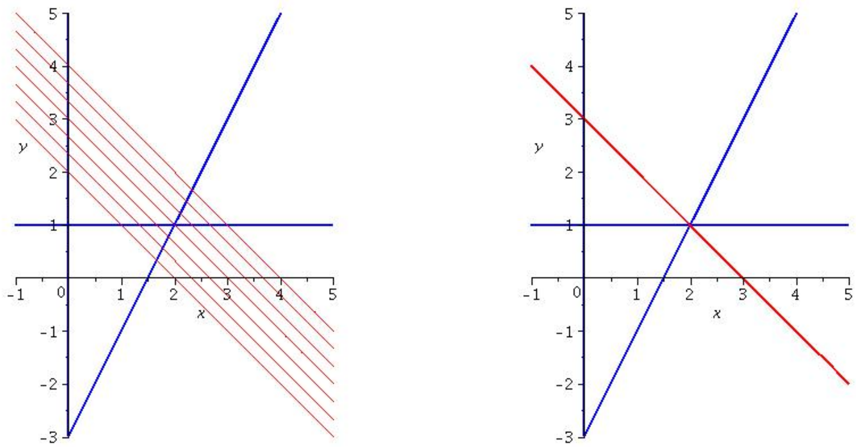
hvor $I = \{1, 2\}$, $I^* = \{1\}$, $J = \{1, 2, 3\}$, $J^* = \{1, 2\}$ og

$$c = \begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix}, \qquad A = \begin{pmatrix} -2 & 3 & 0 \\ 1 & 2 & -1 \end{pmatrix}, \qquad b = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Her er nogle nyttige regneregler:

- $[i](AB) = ([i]A)B$ og $[I^*](AB) = ([I^*]A)B$
- $(AB)[j] = A(B[j])$ og $(AB)[J^*] = A(B[J^*])$

[1] Vi skriver $[I^*]A$ for I^*-rækkerne og $A[J^*]$ for J^*-søjlerne i A



FIGUR 1. Geometrisk løsning af LP fra Eksempel 2.1

- $Ax = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} A[1] & \cdots & A[n] \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j\in J}(A[j])x_j = x_1A[1] + \cdots + x_nA[n]$ er en linearkombination i $\mathbf{R}^m$ af de n *søjler* i A.
- $y^tA = \begin{pmatrix} y_1 & \cdots & y_m \end{pmatrix} A = \begin{pmatrix} y_1 & \cdots & y_m \end{pmatrix}\begin{pmatrix} [1]A \\ \vdots \\ [m]A \end{pmatrix} = \sum_{i\in I} y_i[i]A = y_1[1]A + \cdots + y_m[m]A$ er en linearkombination i $\mathbf{R}^n$ af de m *rækker* i A
- $y^tAx = \sum_{i\in I, j\in J} y_i([i]A[j])x_j$

Vi skal se at ethvert lineært maksimeringsprogram (P) har et dualt minimeringsprogram (P') som i Definition 3.1. Teorien, som handler om dualiteten mellem det primale program (P) og det duale program (P'), kulminerer med Dualitetssætningen 8.1.

Kanoniske programmer og standard programmer er specielle lineære programmer. Vi vil først omtale kanoniske programmer. Dernæst vil vi se på standard programmer. Vi formulerer dualitet først for standard programmer og dernæst for generelle programmer. Svag dualitet er en banalitet; stærk dualitet en subtilitet.

2. Geometrisk løsning

Løsning af et lineært program kan fortolkes geometrisk. Vi skal her se et eksempel i dimension 2 hvor det er lettest at se ideen.

EKSEMPEL 2.1. Maksimeringsproblemet

$$\text{(P)} \quad \text{Maksimér } x_1 + x_2 \text{ under bibetingelser}$$
$$2x_1 - x_2 \le 3$$
$$x_2 \le 1$$
$$x_1 \ge 0$$

går ud på at finde et punkt i mængden

$$M(P) = \{(x_1, x_2) \in \mathbf{R}^2 \mid 2x_1 - x_2 - 3 \le 0, x_2 \le 1, x_1 \ge 0\}$$

med størst mulig koordinatsum – se Figur 1. Geometrisk kommer det ud på at parallelforskyde linjen $x_1 + x_2 = 0$ i retning $(1, 1)$ lige indtil den forlader $M(P)$. En optimal løsning er et punkt i $M(P)$ med maximal koordinatsum.

3. Generelle lineære programmer og deres duale

I et *generelt* lineært maksimeringsprogram kan der både være bibetingelser givet ved uligheder og ved ligheder og der kan både være fortegnskrav og ikke være fortegnskrav på de variable.

DEFINITION 3.1. *Et generelt lineært program har formen*

$$(P) \qquad \text{Maksimér } c^t x \text{ under bibetingelser } [I^*]Ax \leq [I^*]b,\ [I-I^*]Ax = [I-I^*]b,\ [J^*]x \geq 0$$

og det duale program er

$$(P') \qquad \text{Minimér } y^t b \text{ under bibetingelser } y^t A[J^*] \geq c^t[J^*],\ y^t A[J-J^*] = c^t[J-J^*],\ y^t[I^*] \geq 0$$

I tableauet for et generelt lineært program (hvor vi for nemheds skyld antager at $I^* = \{1,\dots,i\}$ og $J^* = \{1,\dots,j\}$) markerer vi I^*-rækker og J^*-søjler med stjerner

$$\begin{array}{cc|cccccc|c}
 & & * & \dots & * & & & & \\
 & & x_1 \geq 0 & \dots & x_j \geq 0 & x_{j+1} & \dots & x_n & b \\ \hline
* & y_1 \geq 0 & a_{11} & \dots & a_{1j} & a_{1j+1} & \dots & a_{1n} & \leq b_1 \\
\vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & \\
* & y_i \geq 0 & a_{i1} & \dots & a_{ij} & a_{ij+1} & \dots & a_{in} & \leq b_i \\
 & y_{i+1} & a_{i+11} & \dots & a_{i+1j} & a_{i+1j+1} & \dots & a_{i+1n} & = b_{i+1} \\
 & \vdots & \vdots & & \vdots & & \vdots & \vdots & \\
 & y_m & a_{m1} & \dots & a_{mj} & a_{mj+1} & \dots & a_{mn} & = b_m \\ \hline
 & c^t & \geq c_1 & \dots & \geq c_j & = c_{j+1} & \dots & = c_n &
\end{array}$$

Rækkerne, med $x_j$erne tilføjet, viser de primale bibetingelser og søjlerne, med $y_i$erne tilføjet, de duale bibetingelser.

Det duale program er motiveret af den generelle version Sætning 7.2 af Karush–Kuhn–Tucker.

EKSEMPEL 3.2. Det generelle lineære program fra Eksempel 1.1 har

$$\text{(P')} \quad \text{Minimér } 3y_1{+}4y_2 \text{ under bibetingelser}$$
$$\begin{aligned} -2y_1 + y_2 &\geq 1 \\ 3y_1 + 2y_2 &\geq 5 \\ -y_2 &= -2 \\ y_1 &\geq 0 \end{aligned}$$

som sit duale program. De mulige løsninger er $M(P') = \{(y_1, 2) \mid \frac{1}{3} \leq y_1 \leq \frac{1}{2}\}$ og den optimale løsning er $O(P') = \{(\frac{1}{3}, 2)\}$ da y_1 skal være så lille som muligt. Den optimale værdi er $\inf(P') = 9$.

4. Dualitetslemmaet

Vi betragter et *generelt* lineært program og dets duale program som i Definition 3.1.

LEMMA 4.1 (Dualitetslemma). *Lad $x \in M(P)$ og $y \in M(P')$ være mulige løsninger. Så gælder:*

(1) $c^t x \leq y^t Ax \leq y^t b$

(2) Hvis $c^t x \geq y^t b$, så er x og y optimale løsninger, ulighedstegnene i (1) er lighedstegn, og $\sup(P) = \inf(P')$.

BEVIS. Vi observerer at

$$(c^t - y^tA)x = (c^t - y^tA)[J^*][J^*]x + (c^t - y^tA)[J-J^*][J-J^*]x = (c^t - y^tA)[J^*][J^*]x \tag{4.2}$$

$$y^t(Ax-b) = y^t[I^*][I^*](Ax-b) + y^t[I-I^*][I-I^*](Ax-b) = y^t[I^*][I^*](Ax-b) \tag{4.3}$$

Begge tal er ≤ 0 da de er er produktet af en rækkevektor ≤ 0 med en søjlevektor ≥ 0. Vi har defor at $c^x \leq y^t Ax$ og $y^t Ax \leq y^t b$. □

KOROLLAR 4.4. *$O(P) \neq \emptyset \iff O(P') \neq \emptyset$, og hvis det sker, så er $\sup(P) = \inf(P')$.*

BEVIS. Lad $x \in O(P)$ være en optimal løsning. Ifølge **KKTNB** findes $y \in M(P')$ som opfylder KKT-betingelserne i Sætning 7.2. Fra punkt (c) ser vi at $c^t x = y^t Ax = y^t b$. Men så er y optimal og de to optimale værdier er ens. □

KOROLLAR 4.5. *Lad $x \in M(P)$ og $y \in M(P')$ være tilladte løsninger. Følgende betingelser er ækvivalente:*

(1) $x \in O(P)$ og $y \in O(P')$

(2) $c^t x \geq y^t b$

(3) $c^t x = y^t b$

(4) $c^t x = y^t Ax = y^t b$

(5) $(c^t - y^tA)x = 0$ og $y^t(Ax-b) = 0$

(6) $(c^t - y^tA)[J^*][J^*]x = 0$ og $y^t[I^*][I^*](Ax-b) = 0$

Maksimér $c^t x$	$Ax \leq b,\ x \geq 0$	$Ax \leq b$	$Ax = b,\ x \geq 0$	$Ax = b$
Minimér $y^t b$	$y^t A \geq c^t,\ y^t \geq 0$	$y^t A = c^t,\ y^t \geq 0$	$y^t A \geq c^t$	$y^t A = c^t$
Type	$I^* = I,\ J^* = J$	$I^* = I,\ J^* = \emptyset$	$I^* = \emptyset,\ J^* = J$	$I^* = \emptyset,\ J^* = \emptyset$
Navn	standard		kanonisk	

TABEL 1. Liste over primale og duale programmer

(7) $\forall j \in J^* \colon y^t A[j] = c_j$ *eller* $x_j = 0$ *og* $\forall i \in I^* \colon y_i = 0$ *eller* $[i]Ax = b_i$
(8) $\forall j \in J \colon y^t A[j] = c_j$ *eller* $x_j = 0$ *og* $\forall i \in I \colon y_i = 0$ *eller* $[i]Ax = b_i$

BEVIS. Vi har lige set at (2) $\implies$ (1) og faktisk også at (2)–(8) er ækvivalente: Vi bruger (4.2) og (4.3) til (5) $\iff$ (6). Vi har (7) $\iff$ (8) fordi denne betingelse er automatisk opfyldt for $j \in J - J^*$ og $i \in I - I^*$.
(1) $\implies$ (3): Da x er optimal findes ifølge de generelle KKT-betingelser i Sætning 7.2 en mulig $z \in M(P')$ med $c^t x = z^t b$. Så er z optimal (Lemma 4.1.(2)). Da også y er optimal, er $y^t b = z^t b$. Altså, $c^t x = y^t b$. $\square$

EKSEMPEL 4.6. Tableauet

		$*$	$*$		
		x_1	x_2	x_3	
$*$	y_1	-2	3	0	≤ 3
	y_2	1	2	-1	$= 4$
		≥ 1	≥ 5	$= -2$	

beskriver det programmet fra Eksempel 1.1 og dets duale program

$$
\begin{array}{ll}
\text{(P)} & \text{Max } x_1 + 5x_2 - 2x_3 \text{ ub} \\
& -2x_1 + 3x_2 \leq 3 \\
& x_1 + 2x_2 - x_3 = 4 \\
& x_1 \geq 0, \quad x_2 \geq 0
\end{array}
\qquad
\begin{array}{ll}
\text{(P')} & \text{Min } 3y_1 + 4y_2 \text{ ub} \\
& -2y_1 + y_2 \geq 1 \\
& 3y_1 + 2y_2 \geq 5 \\
& -y_2 = -2 \\
& y_1 \geq 0
\end{array}
$$

I minimeringsprogrammet (P') er $y_2 = 2$. (P') går derfor ud på at minimere $3y_1 + 8$ under bibetingelserne $-2y_1 \geq -1$, $3y_1 \geq 1$, $y_1 \geq 0$. Mulighedsområdet er $M(P') = [1/3, 1/2]$, så en optimal løsning er $(y_1, y_2) = (1/3, 2)$ og den optimale værdi er $\inf(P') = 9$. Vi ved nu (Korollar 4.4) at der også findes en optimal løsning (x_1, x_2, x_3) til (P) med optimal værdi $x_1 + 5x_2 - 2x_3 = 9$. Komplementær slæk (Korollar 1.(8)) giver at da $-2y_1 + y_2 = -2/3 + 2 = 4/3 > 1$ er $x_1 = 0$, og da $y_1 > 0$ er $-2x_1 + 3x_2 = 3$. Altså er $x_2 = 1$ og så er $x_3 = -2$.

5. Kanoniske programmer

I et kanonisk maksimeringsprogram er alle bibetingelser givet ved ligheder og der er et fortegnskrav på alle variable. Det betyder at $I^* = \emptyset$ og $J^* = J$.

DEFINITION 5.1. *Et* kanonisk *lineært program har formen*

$$(P) \qquad \textit{Maksimér } c^t x \textit{ under bibetingelser } Ax = b,\ x \geq 0$$

og det duale program har formen

$$(P) \qquad \textit{Minimér } y^t b \textit{ under bibetingelser } y^t A \geq c^t$$

I et kanonisk programs tableau

	$*$	$\dots$	$*$	$\dots$	$*$	
	$x_1 \geq 0$	$\dots$	$x_j \geq 0$	$\dots$	$x_n \geq 0$	b
y_1	a_{11}	$\dots$	a_{1j}	$\dots$	a_{1n}	$= b_1$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
y_i	a_{i1}	$\dots$	a_{ij}	$\dots$	a_{in}	$= b_i$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
y_m	a_{m1}	$\dots$	a_{mj}	$\dots$	a_{mn}	$= b_m$
c^t	$\geq c_1$	$\dots$	$\geq c_j$	$\dots$	$\geq c_n$	

er der stjerner på alle primale variable (x_j), da de er underlagt positivitskravet, men ingen stjerner på de duale variable (y_i), da ingen bibetingelser er uligheder. Mulighedsområdet $M(P)$ er fællesmængden af den positive kegle udspændt a basisvektorerne ($x \geq 0$) og et lineært underrum ($Ax = b$).

DEFINITION 5.2. *En mulig løsning* $x \in \mathbf{R}^n$*,*

$$x \geq 0, \qquad b = Ax = \sum_{j \in J} x_j A[j] = \sum_{\substack{j \in J \\ x_j > 0}} x_j A[j]$$

til det kanoniske program (P) er en basisløsning *hvis de* benyttede *søjler*

$$\{A[j] \mid j \in J, \, x_j > 0\}$$

er lineært uafhængige.

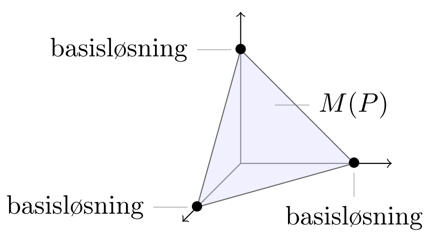
Der er kun endeligt mange basisløsninger:

(1) Der er kun endeligt mange delmængder J^* af J så søjlerne i matricen $A[J^*]$ er lineært uafhængige;
(2) For hver $J^* \subseteq J$ som i (1) findes højst en vektor $x \in \mathbf{R}^n$ med $A[J^*][J^*]x = b$, $[J^*]x > 0$, $[J - J^*]x = 0$.

Her er den geometriske tolkning (som er vigtig for simplex algoritmen).

SÆTNING 5.3. *En mulig løsning $x \in M(P)$ er en basisløsning hvis og kun hvis x er et hjørne i det konvekse polyeder $M(P)$.*

En mulig løsning $x \in M(P)$ er et *hjørne* i det konvekse polyeder $M(P)$ hvis x ikke ligger mellem to andre punkter fra $M(P)$.



(P) Max $c^t x$ ub $Ax = 1$, $x \geq 0$

$$A = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \, b = (1)$$

$$M(P) = \{(x_1, x_2, x_3) \geq 0 \mid x_1 + x_2 + x_2 = 1\}$$

Et af de tre hjørner er optimalt

Fordelen ved kanoniske systemer er

SÆTNING 5.4. *Hvis (P) har en optimal løsning, så har (P) også en optimal basisløsning.*

BEVIS. Lad x være en optimal løsning til det kanoniske maksimerings program (P) (Definition 5.1). Lad $J^* = \{j \in J \mid x_j > 0\}$ være mængden af de index $j \in J$ hvor $x_j > 0$.

Hvis de benyttede søjler $\{A[j] \mid j \in J^*\}$ er lineært uafhængige, så er x en basisløsning og vi er færdige.

Antag derfor at de benyttede søjler $\{A[j] \mid j \in J^*\}$ er lineært afhængige. Vælg $y \in \mathbf{R}^n$ så $Ay = 0$, y benytter de samme søjler som x og $y_j > 0$ for mindst et j. Vi påstår at $c^t y = 0$: Når λ er tæt ved 0, lad os sige $|\lambda| < \varepsilon$, så $x - \lambda y$ en mulig løsning fordi $x - \lambda y \geq 0$ og $A(x - \lambda y) = Ax - \lambda Ay = Ax = b$. Men så er

$$c^t(x - \lambda y) = c^t x - \lambda c^t y \leq c^t x \text{ dvs } \lambda c^t y \geq 0$$

da x er en optimal løsning. Men her kan λ have begge fortegn, så det kan kun være rigtigt hvis faktisk $c^t y = 0$.

Da $c^t y = 0$ ser vi at alle punkter af formen $x - \lambda y$ er optimale løsninger når $x - \lambda y \geq 0$, dvs $x_j - \lambda y_j \geq 0$ for alle j. Hvis $y_j \leq 0$ er dette OK. Hvis $y_j > 0$ (og det sker for mindst et j) forlanger vi at $\lambda \leq \frac{x_j}{y_j}$. Hvis vi sætter

$$\lambda = \min\{\frac{x_j}{y_j} \mid j \in J^*, \, y_j > 0\}$$

så er $\lambda > 0$ og $x - \lambda y$ er en mulig, faktisk optimal, løsning. Nu er $x_j - \lambda y_j = 0$ for mindst et $j \in J^*$. Vi har altså fundet en optimal løsning som benytter færre søjler.

Hvis de benyttede søjler er linerært uafhængige, er vi færdige. Hvis ikke, kører vi den samme procedure igen. Denne proces stopper; i hvert fald ved den tomme mængde, som er lineært uafhængig. □

Da der kun er endeligt mange basisløsninger er løsningen af kanoniske programmer fuldstændig algoritmisk og dermed afsluttet set fra et teoretisk matematisk synspunkt. Den eneste generelle løsningsmetode vi har går på at omskrive til at kanonisk program, finde alle basisløsninger, og vælge den basisløsning der maksimerer objektfunktionen. Fra Sætning 5.4 ved vi at det er en optimal løsning.

Ethvert lineært program er ækvivalent med et kanonisk program. Men det kanoniske program har fået flere variable og vil ofte i praksis være mere uhåndterligt end det oprindelige program. Det er en af ulemperne ved kanoniske programmer. En anden (rent æstetisk?) ulempe er at der ikke er symmetri mellem I, I^* og J, J^* (se Tabel 1).

Ethvert generelt lineært program kan omformuleres til et kanonisk program ved at erstatte

- alle ubegrænsede variable $[J-J^*]x$ med $[J-J^*]x - [J-J^*]u$ og forlange $[J-J^*]x \geq 0$, $[J-J^*]u \geq 0$
- alle ulighedsbibetingelser $[I^*]Ax \leq [I^*]b$ med lighedsbibetingelser $[I^*]Ax + [I^*]v = [I^*]b$ hvor $v \geq 0$

Ved omskrivningen introduceres $|J-J^*|+|I^*|$ nye variable: u_j, $j \in J-J^*$, og v_i, $i \in I^*$, men antallet af bibetingelser er uændret.

Definition 5.5. *Det kanoniske program associeret til det generelle program (P) fra Definition 3.1 er*

$$(P_{\text{kan}}) \qquad \mathit{Max} \begin{pmatrix} c \\ -[J-J^*]c \\ 0 \end{pmatrix}^t \begin{pmatrix} x \\ u \\ v \end{pmatrix} \text{ under bibetingelser } \begin{pmatrix} A & -A[J-J^*] & E_I[I^*] \end{pmatrix} \begin{pmatrix} x \\ u \\ v \end{pmatrix} = b, \ \begin{pmatrix} x \\ u \\ v \end{pmatrix} \geq 0$$

Her er $u = (u_j)_{j \in J-J^*}$, $v = (v_i)_{i \in I^*}$, $[J-J^*]c$ er $J-J^*$ rækkerne i c, $E_I[I^*]$ er I^*-søjlerne i $I \times I$ enhedsmatricen E_I, og $A[J-J^*]$ er $J-J^*$ søjlerne i A.

Eksempel 5.6. Det kanoniske program (P_{kan}) associeret til (P) fra Eksempel 1.1 er

$$\text{Max } \begin{pmatrix} 1 & 5 & -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u_3 \\ v_1 \end{pmatrix} \text{ under bibetingelser } \begin{pmatrix} -2 & 3 & 0 & 0 & 1 \\ 1 & 2 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u_3 \\ v_1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u_3 \\ v_1 \end{pmatrix} \geq 0$$

Ingen basisløsninger til det kanoniske program benytter 0 eller 1 søjle i matricen. (Der findes ikke $x_j > 0$ så $x_j A[j] = b$.) Der er 4 basisløsninger som benytter 2 søjler: De 4 løsninger til $A[J^*][J^*]x = b$, $[J^*]x > 0$, $[J-J^*]x = 0$, hvor $|J^*| = 2$ og $A[J^*]$ har lineært uafhængige søjler, er

- $x^t = (\frac{6}{7}, \frac{11}{7}, 0, 0, 0)$ med $c^t x = \frac{61}{7}$
- $x^t = (4, 0, 0, 0, 11)$ med $c^t x = 4$
- $x^t = (0, 1, 0, 2, 0)$ med $c^t x = 9$
- $x^t = (0, 0, 0, 4, 3)$ med $c^t x = 8$

Forudsat at vi ved at der er optimale løsninger til det kanoniske program, så er $x^t = (0, 1, 0, 2, 0)$ med $c^t x = 9$ en optimal basisløsning. En optimal løsning til det oprindelige, generelle lineære program er så $x^t = (0, 1, -2)$ med samme optimale værdi, 9. Det er også den optimale værdi i det duale program (P') (Eksempel 3.2).

6. Standard programmer

I et standard minimeringsprogram er alle bibetingelser givet ved uligheder og der er et fortegnskrav på alle variable. Det betyder at $I^* = I$ og $J^* = J$.

Definition 6.1. *Et lineært* standard maksimeringsprogram *har formen*

$$(P) \qquad \mathit{Maksimér}\ c^t x \text{ under bibetingelser } Ax \leq b,\ x \geq 0$$

og det duale minimeringsprogram er

$$(P') \qquad \mathit{Minimér}\ y^t b \text{ under bibetingelser } y^t A \geq c^t,\ y^t \geq 0$$

I standard programmets tableau

$$\begin{array}{cc|ccccc|c}
 & & * & \dots & * & \dots & * & \\
 & & x_1 & \dots & x_j & \dots & x_n & b \\
\hline
* & y_1 & a_{11} & \dots & a_{1j} & \dots & a_{1n} & \leq b_1 \\
 & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
* & y_i & a_{i1} & \dots & a_{ij} & \dots & a_{in} & \leq b_i \\
 & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
* & y_m & a_{m1} & \dots & a_{mj} & \dots & a_{mn} & \leq b_m \\
\hline
 & c & \geq c_1 & \dots & \geq c_j & \dots & \geq c_n &
\end{array}$$

er der stjerner på alle variable (x_j) og (y_i) da de alle skal være ≥ 0. Rækkerne i tableauet viser de primale bibetingelser og søjlerne viser de duale bibetingelser.

EKSEMPEL 6.2. Her er et eksempel på et standardprogram

$$\begin{aligned}(\mathrm{P})\ \text{Maksimér } x_1 - 2x_2 - 3x_3 - x_4 &\text{ under bibetingelser}\\ x_1 - x_2 - 2x_3 - x_4 &\leq 4\\ 2x_1 + x_3 - 4x_4 &\leq 2\\ -2x_1 + x_2 + x_4 &\leq 1\\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 &\geq 0,\end{aligned}$$

Ethvert generelt lineært program kan omformuleres til et standard program ved at erstatte

- de ubegrænsede variable $[J-J^*]x$ med $[J-J^*]x - [J-J^*]u$ og forlange $[J-J^*]x \geq 0$, $[J-J^*]u \geq 0$
- lighederne $[I-I^*]Ax = [I-I^*]b$ med ulighederne $[I-I^*]Ax \leq [I-I^*]b$, $[I-I^*]Ax \geq [I-I^*]b$, eller $[I-I^*]Ax \leq [I-I^*]b$, $-[I-I^*]Ax \leq -[I-I^*]b$

Ved omskrivningen introduceres $|J-J^*|$ nye variable u_j, $j \in J-J^*$, og $|I-I^*|$ nye bibetingelser.

DEFINITION 6.3. *Standard programmet associeret til det generelle program (P) fra Definition 3.1 er*

$$\mathit{Max} \begin{pmatrix} c \\ -[J-J^*]c \end{pmatrix}^t \begin{pmatrix} x \\ u \end{pmatrix} \ \textit{under bibetingelser} \begin{pmatrix} A & -A[J-J^*] \\ -[I-I^*]A & [I-I^*]A[J-J^*] \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \leq \begin{pmatrix} b \\ -[I-I^*]b \end{pmatrix}, \begin{pmatrix} x \\ u \end{pmatrix} \geq 0$$

hvor $u = (u_j)_{j \in J-J^*}$.

EKSEMPEL 6.4. Standard programmet associeret til (P) fra Eksempel 1.1 er

$$\text{Max } \begin{pmatrix} 1 & 5 & -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u_3 \end{pmatrix} \text{ under bibetingelser } \begin{pmatrix} -2 & 3 & 0 & 0 \\ 1 & 2 & -1 & 1 \\ -1 & -2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u_3 \end{pmatrix} \leq \begin{pmatrix} 3 \\ 4 \\ -4 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u_3 \end{pmatrix} \geq 0$$

7. Farkas' lemma

Farkas' lemma siger at hvis en vektor $b \in \mathbf{R}^m$ ikke ligger i den positive kegle udspændt af endeligt mange vektorer i $\mathbf{R}^m$, så findes en hyperplan som separerer den positive kegle og vektor b. Hyperplanen, $y^\perp$, er givet ved sin normalvektor $y \in \mathbf{R}^m$. I formuleringen af af Farkas's lemma sættes de endeligt mange, n, vektorer i $\mathbf{R}^m$, som udspænder keglen, op som søjler i en matrix $(m \times n)$ matrix A.

Lad A være en $(m \times n)$-matrix og $b \in \mathbf{R}^m$ en m-vektor.

SÆTNING 7.1 (Farkas' lemma). *Netop et af følgende to tilfælde indtræffer:*

(I) Der findes $x \in \mathbf{R}^n$ *så* $Ax = b$, $x \geq 0$.
(II) Der findes $y \in \mathbf{R}^m$ *så* $y^tA \geq 0$, $y^tb < 0$

I (I) har vi m-søjlen

$$Ax = \begin{pmatrix} A[1] & \cdots & A[n] \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A[1]x_1 + \cdots + A[n]x_n \in \mathbf{R}^m$$

som er en linearkombination af søjlerne i A og i (II) har vi n-rækken

$$y^tA = \begin{pmatrix} (y^tA)[1] & \cdots & (y^tA)[n] \end{pmatrix} = \begin{pmatrix} y^t(A[1]) & \cdots & y^t(A[n]) \in \mathbf{R}^n \end{pmatrix}$$

som er ys prikprodukter med søjlerne.

(I) og (II) kan ikke begge indtræffe for det ville give $y^tAx = (y^tA)x \geq 0$ og $y^tAx = y^t(Ax) = y^tb < 0$.

Vi vil ikke bevise Farkas' lemma her men bare nøjes med at se på et eksempel.

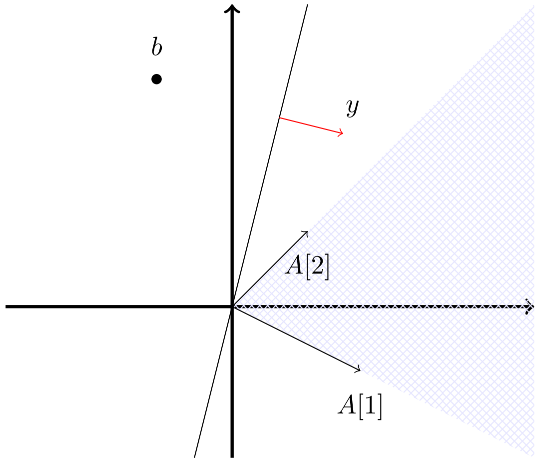
EKSEMPEL 7.2. Lad

$$C(A) = \{x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid x_1 \geq 0, x_2 \geq 0\}, \qquad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

være den konvekse kegle udspændt af søjlerne i 2×2-matricen A. Hvis (I) ikke gælder så betyder det at $b \notin C(A)$. Vælg en hyperplan $H(y) = y^\perp$ sådan at $C(A)$ ligger på den positive side og b på den negative side. (Se Figur 2.) Det betyder at $y^tA \geq 0$ og $y^tb < 0$.

Vi skal bruge en variation af Farkas' lemma med uligheder alle steder:

KOROLLAR 7.3. *Netop et af følgende to tilfælde indtræffer:*



$$Ax = (A[1], A[2]) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A[1]x_1 + A[2]x_2$$
$$y^t A = y^t(A[1], A[2]) = (y^t A[1], y^t A[2])$$

FIGUR 2. Farkas lemma

(I) Der findes $x \in \mathbf{R}^n$ så $Ax \leq b$, $x \geq 0$.
(II) Der findes $y \in \mathbf{R}^m$ så $y^t A \geq 0$, $y^t b < 0$, $y \geq 0$.

BEVIS. Det første tilfælde betyder at der findes $\begin{pmatrix} x \\ z \end{pmatrix} \in \mathbf{R}^n \times \mathbf{R}^m$ sådan at $Ax + z = b$, $\begin{pmatrix} x \\ z \end{pmatrix} \geq 0$. Ligningen $Ax + z = b$ kan også skrives

$$\begin{pmatrix} A & E \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = b$$

Farkas lemma siger at hvis dette ikke sker, så findes $y \in \mathbf{R}^m$ så $y^t \begin{pmatrix} A & E \end{pmatrix} \geq 0$ og $y^t b < 0$. Da $y^t \begin{pmatrix} A & E \end{pmatrix} = \begin{pmatrix} y^t A & y^t \end{pmatrix}$ betyder det at $y^t A \geq 0$, $y^t b < 0$, $y \geq 0$. □

8. Dualitetssætningen

Tabel 2 viser de ni forhold som der kunne tænkes at være mellem duale programmer. Dualitetssætningen siger at kun fire af de ni tilfælde faktisk forekommer.

SÆTNING 8.1 (Dualitetssætning). *Lad (P) og (P') være duale lineære programmer som i Definition 3.1. Netop én af følgende fire situationer vil gælde:*

(I) (P) og (P') har optimale løsninger og $\sup(P) = \inf(P')$.
(II) $M(P) \neq \emptyset$, $M(P') = \emptyset$ og $\sup(P) = \infty$.
(III) $M(P) = \emptyset$, $M(P') \neq \emptyset$ og $\inf(P') = -\infty$
(IV) $M(P) = \emptyset$, $M(P') = \emptyset$

KOROLLAR 8.2 (Den stærke Dualitetssætning). *Følgende fem betingelser er ækvivalente:*

(1) (P) og (P') har optimale løsninger og $\sup(P) = \inf(P')$
(2) $M(P) \neq \emptyset$ og $M(P') \neq \emptyset$
(3) (P) har en optimal løsning
(4) (P') har en optimal løsning
(5) $M(P) \neq \emptyset$ og $\sup(P) < \infty$
(6) $M(P') \neq \emptyset$ og $\inf(P') > -\infty$

BEVIS. Dualitetssætningen 8.1 giver at (1) $\iff$ (2) $\iff$ (3) $\iff$ (4) $\iff$ (5) $\iff$ (6) fordi de alle siger at vi er i 'normaltilfældet' (I). □

Dualitetssætningen 8.1 siger at når $M(P) \neq \emptyset$ og $M(P') \neq \emptyset$ så er der optimale løsninger til begge programmer og de optimale værdier er ens.

BEVIS FOR SÆTNING 8.1. Da ethvert program kan omformuleres til et standardprogram kan vi godt antage at (P) og (P') er duale standardprogrammer som i Definition 6.1.

	$M(P) \neq \emptyset$, $\sup(P) < \infty$	$M(P) \neq \emptyset$, $\sup(P) = \infty$	$M(P) = \emptyset$
$M(P') \neq \emptyset$, $\inf(P') > -\infty$	$\times$		
$M(P') \neq \emptyset$, $\inf(P') = -\infty$			$\times$
$M(P') = \emptyset$		$\times$	$\times$

TABEL 2. Dualitetssætningen

Korollar 4.5 siger at (I) gælder hvis og kun hvis der findes $x \in \mathbf{R}^n$ og $y \in \mathbf{R}^m$ så

$$\begin{pmatrix} A & 0 \\ 0 & -A^t \\ -c^t & b^t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} b \\ -c \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \geq 0$$

Her står nemlig at $x \in M(P)$, $y \in M(P')$ og $c^t x \geq b^t y$.

Antag nu at situation (I) *ikke* er tilfældet. Vi skal så vise at et af tilfældene (II)–(IV) indtræffer.

Da (I) *ikke* gælder, så siger Farkas' lemma (Korollar 7.3) at der findes $v \in \mathbf{R}^m$, $u \in \mathbf{R}^n$ og $\alpha \in \mathbf{R}$ så

$$\begin{pmatrix} v^t & u^t & \alpha \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -A^t \\ -c^t & b^t \end{pmatrix} \geq 0, \qquad \begin{pmatrix} v^t & u^t & \alpha \end{pmatrix} \begin{pmatrix} b \\ -c \\ 0 \end{pmatrix} < 0, \qquad \begin{pmatrix} v \\ u \\ \alpha \end{pmatrix} \geq 0$$

Her står at

$$Au \leq \alpha b, \quad v^t A \geq \alpha c^t, \quad v^t b < c^t u, \quad u \geq 0, \quad v \geq 0, \quad \alpha \geq 0$$

Vi kan ikke have $\alpha > 0$ for det ville sige at $\alpha^{-1} u \in M(P)$, $\alpha^{-1} v \in M(P')$ og $(\alpha^{-1} v)^t b < c^t(\alpha^{-1} u)$ i modstrid med Dualitetslemmaet 4.1.(1). Vi har altså at $\alpha = 0$, dvs

$$Au \leq 0, \quad v^t A \geq 0, \quad b^t v < c^t u, \quad u \geq 0, \quad v \geq 0$$

Vi deler nu ind i tilfældene $c^t u > 0$ og $c^t \leq 0$, hver med to undertilfælde.

$c^t u > 0$: I dette tilfælde er $M(P) = \emptyset$: Hvis y er en mulig løsning til (P'), da er $y^t A \geq c^t$ og $y \geq 0$; men

$$y^t A \geq c^t, \quad y \geq 0 \Longrightarrow c^t u \leq (y^t A) u = y^t (Au) \leq 0$$

i modstrid med antagelsen om at $c^t u > 0$. Der er nu to muligheder:

$M(P) = \emptyset$: Tilfælde (IV).

$M(P) \neq \emptyset$: Vælg en mulig løsning $x \in M(P)$, dvs $Ax \leq b$, $x \geq 0$. Så er $x + \lambda u \in M(P)$ for alle $\lambda > 0$ for $A(x + \lambda u) = Ax + \lambda Au \leq Ax \leq b$ og $x + \lambda u \geq 0$. Desuden har vi at objektfunktionen

$$c^t(x + \lambda u) = c^t x + \lambda c^t u \to \infty \text{ for } t \to \infty$$

fordi $c^t u > 0$. Altså er vi i tilfælde (II).

$c^t u \leq 0$: Nu er $v^t b < u^t c \leq 0$. I dette tilfælde er $M(P) = \emptyset$ fordi

$$Ax \leq b, \quad x > 0 \Longrightarrow v^t b \geq v^t (Ax) = (v^t A) x \geq 0$$

i modstrid med antagelsen om at $v^t b < 0$. Der er nu to muligheder:

$M(P') = \emptyset$: Tilfælde (IV).

$M(P') \neq \emptyset$: Vælg $y \in M(P')$, dvs $y^t A \geq c^t$, $y \geq 0$. Så er $y + \lambda v \in M(P')$ for alle $\lambda > 0$ fordi $(y + \lambda v)^t A = y^t A + \lambda v^t A \geq y^t A \geq c^t$ og $y + \lambda v \geq 0$. Desuden har vi at objektfunktionen

$$(y + \lambda v)^t b = y^t b + \lambda v^t b \to -\infty \text{ for } t \to \infty$$

fordi $v^t b < 0$. Altså er vi i tilfælde (III).

$\square$

EKSEMPEL 8.3 (MASO eksamen vinter 2007/2008, Opgave 5). Tableauet

		$*$			
		x_1	x_2	x_3	
$*$	y_1	3	-5	1	≤ 1
$*$	y_2	-1	4	2	≤ 2
$*$	y_3	2	1	-3	≤ -2
		≥ 7	$= -1$	$= -4$	

beskriver de generelle lineære programmer

(P) Max $7x_1 - x_2 - 4x_3$ under bibetingelser

$$\begin{aligned} 3x_1 - 5x_2 + x_3 &\leq 1 \\ -x_1 + 4x_2 + 2x_3 &\leq 2 \\ 2x_1 + x_2 - 3x_3 &\leq -2 \\ x_1 &\geq 0 \end{aligned}$$

(P') Min $y_1 + 2y_2 - 2y_3$ under bibetingelser

$$\begin{aligned} 3y_1 - y_2 + 2y_3 &\geq 7 \\ -5y_1 + 4y_2 + y_3 &= -1 \\ y_1 + 2y_2 - 3y_3 &= -4 \\ y_1 \geq 0, y_2 \geq 0, y_3 &\geq 0 \end{aligned}$$

De to ligninger blandt bibetingelserne i (P') giver

$$\begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 2 & -3 \end{pmatrix}^{-1} \begin{pmatrix} -1 + 5y_1 \\ -4 - y_1 \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 3 & 1 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} -1 + 5y_1 \\ -4 - y_1 \end{pmatrix} = \begin{pmatrix} y_1 - \frac{1}{2} \\ y_1 + 1 \end{pmatrix}$$

Minimeringsprogrammet (P') er derfor ækvivalent med

(P') Min $y_1 - 3$ under bibetingelser

$$\begin{aligned} 4y_1 + \frac{5}{2} &\geq 7 \\ y_2 &= y_1 - \frac{1}{2} \\ y_3 &= y_1 + 1 \\ y_1 \geq 0, y_2 \geq 0, y_3 &\geq 0 \end{aligned}$$

$$\begin{aligned} &M(P') = \{(y_1, y_1 - \frac{1}{2}, y_1 + 1) \mid y_1 \geq \frac{9}{8}\} \\ &O(P') = \{(\frac{9}{8}, \frac{5}{8}, \frac{17}{8})\} \\ &\inf(P') = -\frac{15}{8} \end{aligned}$$

Da (P') har optimale løsninger gælder det samme for (P) (Korollar 4.4). Da alle tre koordinater i den optimale løsning for (P') er positive giver komplementær slæk (Korollar 1.(8)) at alle tre bibetingelser i (P) aktive. Det giver at $x = (\frac{1}{4}, \frac{1}{8}, \frac{7}{8})$ er den eneste optimale løsning til (P). Vi checker lige at objektfunktionens værdi i $(\frac{1}{4}, \frac{1}{8}, \frac{7}{8})$ er $-\frac{15}{8}$ som den skulle være.

Flere eksempler (med Maple).

Der findes mange (gratis) computer algebra programmer til behandling af lineære optimering. Google search på Linear optimization

Litteratur

1. Dimitri P. Bertsekas, *Nonlinear programming*, second ed., Athena Scientific Optimization and Computation Series, Athena Scientific, Belmont, MA, 1999. MR 3444832
2. Neil Cameron, *Introduction to linear and convex programming*, Australian Mathematical Society Lecture Series, vol. 1, Cambridge University Press, Cambridge, 1985. MR 811634 (87k:90156)
3. Vašek Chvátal, *Linear programming*, A Series of Books in the Mathematical Sciences, W. H. Freeman and Company, New York, 1983. MR 717219 (86g:90062)
4. Eustaquio, Karas, and Ribeiro, *Constraint qualifications for nonlinear programming*, Unpublished note.
5. Reiner Horst, Panos M. Pardalos, and Nguyen V. Thoai, *Introduction to global optimization*, second ed., Nonconvex Optimization and its Applications, vol. 48, Kluwer Academic Publishers, Dordrecht, 2000. MR 1799654 (2001h:90001)
6. Mikuláš Luptáčik, *Mathematical optimization and economic analysis*, Springer Optimization and Its Applications, vol. 36, Springer, New York, 2010. MR 2555210
7. Jiří Matoušek and Bernd Gärtner, *Understanding and using linear programming*, Universitext, Springer-Verlag, Berlin, 2007.
8. Knut Sydsaeter, *Matematisk analyse. Bind II*, second ed., Universitetsforlaget, Oslo, 1978, Scandinavian University Books. MR 502888 (80c:00004)