Mélykúti Bence

Young-programozás

2005. április 5.

Bevezetés

- Ismerkedés a Young-programozás alapjaival
- Áttekintés: az LP komplementaritási problémája
- A Young-programozási feladat (az LP-feladat felépítését utánozva)
- A Young-programozás dualitáselmélete
  (Kas Péter, Klafszky Emil, Mályusz Levente, Gökhan Izbirak, 2000.)
- Kitekintés: a Young-programozás és más problémacsaládok viszonya
  (Zehra Boratas, Illés Tibor, Kas Péter, 2002.)

A Young-egyenlőtlenség

Legyen $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ folytonos, szigorúan monoton csökkenő függvény.

$$\psi := \varphi^{-1}$$

$(u, v) \in \mathbb{R}^2$ esetén

$$\begin{aligned} S_\varphi(u, v) : &= (u - \psi(v))v - \int_{\psi(v)}^{u} \varphi(t)\,\mathrm{d}t \\ &= (v - \varphi(u))u - \int_{\varphi(u)}^{v} \psi(t)\,\mathrm{d}t \end{aligned}$$

Def.: $S_\varphi(u, v)$ az $(u, v) \in \mathbb{R}^2$ pont φ grafikonjára vonatkoztatott *illeszkedési pontatlansága.*

$(x, z) \in \mathbb{R}^{2n}$ esetén $S_\varphi(x, z) := \sum_{j=1}^{n} S_\varphi(x_j, z_j)$.

Young-egyenlőtlenség: $S_\varphi \geq 0$.

A Young-programozás S_φ minimalizálását fogja jelenteni.

A lineáris programozási feladat

$A \in \mathbb{R}^{m\times n}$ mátrix. $x = (x_1, \dots, x_n)^T$, $z = (z_1, \dots, z_n)^T$, $y = (y_1, \dots, y_m)^T$ ismeretlen, $\widehat{x}, \widehat{z} \in \mathbb{R}^n$ rögzített vektorok.

$$S(x,z) = \sum_{j=1}^{n} S(x_j, z_j) := \sum_{j=1}^{n} x_j z_j$$

Megj.: Ez olyan φ-nek felel meg, amelynek görbéje

$$\{(x,z) \in \mathbb{R}^2 | x \geq 0, z \geq 0, x^T z = 0\}.$$

$$P := \{x | Ax = A\widehat{x}, x \geq 0\}, \quad D := \{z | z = \widehat{z} + A^T y, z \geq 0\},$$

$$P_+ := \{x | Ax = A\widehat{x}, x > 0\}, \quad D_+ := \{z | z = \widehat{z} + A^T y, z > 0\}.$$

$\mathcal{L} := A$ mátrix sortere (oszlopvektorokként)$= A^T$ oszloptere. Ekkor $\mathcal{L}^\perp = A$ nulltere.

$$x \in \widehat{x} + \mathcal{L}^\perp \iff Ax = A\widehat{x}$$

$$z \in \widehat{z} + \mathcal{L} \iff z = \widehat{z} + A^T y$$

Az LP-feladat:

(I) *(egyensúlyi forma)* Keresendő x, y, z, amelyekkel

$$\begin{array}{ll} \text{(L1)} & Ax = A\widehat{x}, \quad z = \widehat{z} + A^T y \\ \text{(L2)} & x \geq 0, \quad z \geq 0 \\ \text{(L3)} & x^T z = 0 \end{array}$$

(II) *(optimalizálási forma)* Keresendő x, y, z megoldás:

$$\min\{x^T z | x \in P, z \in D\}$$

(III) *(primál-duál forma)* Keresendő x, y, z megoldás:

$$\min\{x^T \widehat{z} | x \in P\}, \quad \min\{\widehat{x}^T z | z \in D\}$$

Lemma:
(1) Ha x és z kielégíti (L1)-et, akkor $x^T\widehat{z} + \widehat{x}^T z = x^T z + \widehat{x}^T\widehat{z}$.
(2) Ha x és z kielégíti (L1)-et, (L2)-t, akkor $x^T\widehat{z} + \widehat{x}^T z \geq \widehat{x}^T\widehat{z}$.
(3) Ha x és z kielégíti (L1)-et, (L2)-t és (L3)-at, akkor $x^T\widehat{z} + \widehat{x}^T z = \widehat{x}^T\widehat{z}$.
Biz.: x és z kielégíti (L1)-et $\Longrightarrow$
$\Longrightarrow x = \widehat{x} + w$, ahol $w \in \mathcal{L}^\perp$; $z = \widehat{z} + y$, ahol $y \in \mathcal{L}$.
(1) mindkét oldalát felírjuk ezekkel, és kijön, hogy egyenlők.
(2) és (3) triviálisan következik (1)-ből.

Köv.: *(gyenge dualitás)* Ha $x \in P$ és $z \in D$, és teljesül a Lemma (3) pontja ($x^T \widehat{z} + \widehat{x}^T z = \widehat{x}^T \widehat{z}$), akkor x primál, z duál optimális megoldás.

Tétel:
(1) Ha (L1) és (L2) megoldható, akkor (L1), (L2) és (L3) is.
(2) Ha (L1) és (L2) megoldható, akkor létezik olyan (x^*, z^*) megoldása (III)-nak, amelyre $x^{*T} z^* = 0$.
(3) Ha $\widehat{x}, \widehat{z} \geq 0$, speciálisan mind a primál, mind a duál feladatnak létezik megengedett megoldása, akkor létezik mindkettőnek optimális megoldása ($x^* \in P, z^* \in D$), amelyekre $x^{*T} z^* = 0$.
(Nem bizonyítjuk.)

Köv.: *(erős dualitás)* Ha $x \in P$ és $z \in D$ optimális megoldásai a primál és a duál feladatnak, akkor a Lemmában (3) teljesül: $x^T \widehat{z} + \widehat{x}^T z = \widehat{x}^T \widehat{z}$.

A Young-programozás, a dualitás

Legyen $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ folytonos, szigorúan monoton csökkenő függvény;

$$\lim_{t\to 0+} \varphi(t) = \infty, \quad \lim_{t\to\infty} \varphi(t) = 0.$$

$$\psi := \varphi^{-1}$$

Állítások: Legyen $x, z \in \mathbb{R}_+$.
(1) $S_\varphi(x,z) = S_\psi(z,x)$
(2) $S_\varphi(x,z)$ szigorúan konvex x-ben és z-ben
(3) $S_\varphi \geq 0$
(4) $S_\varphi(x,z) = 0 \iff x = \psi(z)$
(5) $\lim_{x\to\infty} S_\varphi(x,z) = \infty$ minden fix z-re
(6) $\lim_{z\to\infty} S_\varphi(x,z) = \infty$ minden fix x-re
(7) $\partial_x S_\varphi(x,z) = z - \varphi(x)$
(8) $\partial_z S_\varphi(x,z) = x - \psi(z)$
(9) $\lim_{x\to 0+} \partial_x S_\varphi(x,z) = -\infty$ minden fix z-re
(10) $\lim_{z\to 0+} \partial_z S_\varphi(x,z) = -\infty$ minden fix x-re
(11) $\lim_{x\to\infty} \partial_x S_\varphi(x,z) = z$ minden fix z-re
(12) $\lim_{z\to\infty} \partial_z S_\varphi(x,z) = x$ minden fix x-re

Tegyük fel, hogy $\widehat{x} > 0$ és $\widehat{z} > 0$. Ekkor $P_+ \neq \emptyset$, $D_+ \neq \emptyset$.

A Young-feladat:

(I) *(egyensúlyi forma)* Keresendő x, y, z, amelyekkel

$$\begin{array}{ll}(\mathrm{Y1}) & Ax = A\widehat{x}, \quad z = \widehat{z} + A^T y \\ (\mathrm{Y2}) & x > 0, \quad z > 0 \\ (\mathrm{Y3}) & \forall j \in \{1, 2, \ldots, n\} : \quad x_j = \psi(z_j)\end{array}$$

(II) *(optimalizálási forma)* Keresendő x, y, z megoldás:

$$\min\{\sum_{j=1}^{n} S_\varphi(x_j, z_j) | x \in P_+, z \in D_+\}$$

(III) *(primál-duál forma)* Keresendő x, y, z megoldás:

$$\min\{\sum_{j=1}^{n} S_\varphi(x_j, \widehat{z}_j) | x \in P_+\}, \quad \min\{\sum_{j=1}^{n} S_\varphi(\widehat{x}_j, z_j) | z \in D_+\}$$

Lemma: Tegyük fel, hogy $\widehat{x} > 0$ és $\widehat{z} > 0$.
(1) Ha x és z kielégíti (Y1)-et, (Y2)-t, akkor

$$S_\varphi(x, \widehat{z}) + S_\varphi(\widehat{x}, z) = S_\varphi(x, z) + S_\varphi(\widehat{x}, \widehat{z}).$$

(2) Ha x és z kielégíti (Y1)-et, (Y2)-t, akkor

$$S_\varphi(x, \widehat{z}) + S_\varphi(\widehat{x}, z) \geq S_\varphi(\widehat{x}, \widehat{z}).$$

(3) Ha x és z kielégíti (Y1)-et, (Y2)-t és (Y3)-at, akkor

$$S_\varphi(x, \widehat{z}) + S_\varphi(\widehat{x}, z) = S_\varphi(\widehat{x}, \widehat{z}).$$

Biz.: x és z kielégíti (Y1)-et $\Longrightarrow$
$\Longrightarrow x = \widehat{x} + w$, ahol $w \in \mathcal{L}^\perp$; $z = \widehat{z} + y$, ahol $y \in \mathcal{L}$.
S_φ definíciója alapján (1) mindkét oldalát felírjuk ezekkel, és kijön, hogy egyenlők.
(2) és (3) triviálisan következik (1)-ből.

Köv.: *(gyenge dualitás)* Ha $x \in P_+$ és $z \in D_+$, és teljesül a Lemma (3) pontja ($S_\varphi(x, \widehat{z}) + S_\varphi(\widehat{x}, z) = S_\varphi(\widehat{x}, \widehat{z})$), akkor x primál, z duál optimális megoldás.

Tétel: Feltettük, hogy $\widehat{x} > 0$ és $\widehat{z} > 0$.
(1) (Y1), (Y2), (Y3) megoldható.
(2) Létezik pontosan egy olyan (x^*, z^*) megoldása (II)-nek, amelyre (Y1) és (Y2) teljesül (azaz $x^* \in P_+$ és $z^* \in D_+$), és $\sum_{j=1}^n S_\varphi(x_j^*, z_j^*) = 0$.
(3) Létezik pontosan egy olyan $x^* \in P_+$ és $z^* \in D_+$, amelyekre $\forall j \in \{1, 2, \ldots, n\} : \quad x_j^* = \psi(z_j^*)$.

Megj.: $S_\varphi \geq 0$, $S_\varphi(x_j, z_j) = 0 \iff x_j = \psi(z_j)$ miatt ezek ekvivalensek.

A konvex programozás általános elméletének felhasználása helyett direkt bizonyítást adunk (3)-ra. Adható bizonyítás a primál és a duál oldalon is, és ezek nagyon hasonlók.

Biz.: (A duál oldali bizonyítás.)
áll.: A duál célfüggvénynek egyértelmű minimumhelye van D_+-ban.
biz.: $S_\varphi(x,z)$ szigorúan konvex x-ben és z-ben $\Longrightarrow$
$\Rightarrow$ a duál célfv.-nek egyértelmű min.helye van D-ben, jelöljük z^*-gal.
Belátjuk, hogy $z^* > 0$.
$\mathcal{J} := \{j | z_j^* = 0\}$. Indirekt tegyük fel, hogy $\mathcal{J} \neq \emptyset$.

$$z(t) := z^* + t(\widehat{z} - z^*) \quad (0 \leq t \leq 1)$$

$$\lim_{t\to 0+} \partial_t S_\varphi(\widehat{x}, z(t)) = \lim_{t\to 0+} \sum_{j=1}^{n} \partial_{z_j} S_\varphi(\widehat{x}, z(t))(\widehat{z}_j - z_j^*) \overset{(8)}{=}$$

$$\overset{(8)}{=} \lim_{t\to 0+} \sum_{j=1}^{n} (\widehat{x}_j - \psi(z_j(t)))(\widehat{z}_j - z_j^*) = -\infty,$$

ugyanis ha $j \in \mathcal{J}$, akkor $z_j(t) \xrightarrow{t\to 0} z_j^* = 0$, $\lim_{s\to 0+} \psi(s) = \infty$, és $\widehat{z}_j - z_j^* > 0$. A $j \notin \mathcal{J}$ indexű tagok pedig végesek.
Az, hogy a $j \in \mathcal{J}$ koordinátákban a parciális derivált $-\infty$, ellentmond z^* optimalitásának. Tehát $z^* \in D_+$.

Legyen $\forall j \in \{1, 2, \ldots, n\}: \quad x_j^* = \psi(z_j^*)$.

áll.: x^* primál optimális megoldás.
biz.: $x_j^* = \psi(z_j^*) > 0$ nyilvánvaló.
Indirekt feltéve, hogy $Ax^* \neq A\widehat{x}$, valamely i-re $a^{(i)}x^* \neq a^{(i)}\widehat{x}$.

$$z(\vartheta) := z^* + \vartheta a^{(i)T} \quad (\vartheta \in \mathbb{R})$$

Elég kicsi ϑ esetén $z(\vartheta) \in D_+$, és

$$\partial_\vartheta S_\varphi(\widehat{x}, z(\vartheta))\Big|_{\vartheta=0} = \sum_{j=1}^{n} (\widehat{x}_j - x_j^*) a^{(i)}{}_j = a^{(i)}\widehat{x} - a^{(i)}x^* \neq 0,$$

ami ellentmond annak, hogy z^* duál optimális megoldás. ■

(A primál oldali bizonyítás.)

áll.: A primál célfüggvénynek egyértelmű minimumhelye van P_+-ban.

biz.: $S_\varphi(x,z)$ szigorúan konvex x-ben és z-ben $\Longrightarrow$

$\Rightarrow$ a primál célfv.-nek egyértelmű min.helye van P-ben, jelöljük x^*-gal. Belátjuk, hogy $x^* > 0$.

$\mathcal{J} := \{j | x_j^* = 0\}$. Indirekt tegyük fel, hogy $\mathcal{J} \neq \emptyset$.

$$x(t) := x^* + t(\widehat{x} - x^*) \quad (0 \leq t \leq 1)$$

$$\lim_{t\to 0+} \partial_t S_\varphi(x(t), \widehat{z}) = \lim_{t \to 0+} \sum_{j=1}^{n} \partial_{x_j} S_\varphi(x(t), \widehat{z})(\widehat{x}_j - x_j^*) \overset{(7)}{=}$$

$$\overset{(7)}{=} \lim_{t\to 0+} \sum_{j=1}^{n} (\widehat{z}_j - \varphi(x_j(t)))(\widehat{x}_j - x_j^*) = -\infty,$$

ugyanis ha $j \in \mathcal{J}$, akkor $x_j(t) \xrightarrow{t\to 0} x_j^* = 0$, $\lim_{s\to 0+} \varphi(s) = \infty$, és $\widehat{x}_j - x_j^* > 0$. A $j \notin \mathcal{J}$ indexű tagok pedig végesek.

Az, hogy a $j \in \mathcal{J}$ koordinátákban a parciális derivált $-\infty$, ellentmond x^* optimalitásának. Tehát $x^* \in P_+$.

Legyen $\forall j \in \{1, 2, \ldots, n\}: \quad z_j^* = \varphi(x_j^*)$.

áll.: z^* duál optimális megoldás.
biz.: $z_j^* = \varphi(x_j^*) > 0$ nyilvánvaló.
Indirekt feltéve, hogy $z^* - \widehat{z} \neq A^T y$ semmilyen y-ra, azaz $z^* - \widehat{z} \notin \mathcal{L}$, létezik $\widetilde{x} \in \mathcal{L}^\perp$, amelyre $\widetilde{x}^T (z^* - \widehat{z}) \neq 0$.

$$x(\vartheta) := x^* + \vartheta \widetilde{x} \quad (\vartheta \in \mathbb{R})$$

Elég kicsi ϑ esetén $x(\vartheta) \in P_+$, és

$$\partial_\vartheta S_\varphi(x(\vartheta), \widehat{z}) \Big|_{\vartheta=0} = \widetilde{x}^T (\widehat{z} - z^*) \neq 0,$$

ami ellentmond annak, hogy x^* primál optimális megoldás. ■

Köv.: *(erős dualitás)* Ha $x \in P_+$ és $z \in D_+$ optimális megoldásai a primál és a duál feladatnak, akkor a Lemmában (3) teljesül, azaz $S_\varphi(x, \widehat{z}) + S_\varphi(\widehat{x}, z) = S_\varphi(\widehat{x}, \widehat{z})$.

Köv.: $\widehat{x} > 0$ és $\widehat{z} > 0$ továbbra is álljon fent. Ekkor a primál és a duál optimum bizonyos korlátozás mellett nem függ $\widehat{x}$ és $\widehat{z}$ pontos megválasztásától.

Pontosan fogalmazva:
Egy és csak egy olyan $x \in \widehat{x} + \mathcal{L}^\perp$, $z \in \widehat{z} + \mathcal{L}$ létezik, hogy
$\forall j \in \{1, 2, \ldots, n\}: \quad z_j = \varphi(x_j)$,
és $\forall \check{x} \in \widehat{x} + \mathcal{L}^\perp$, $\check{z} \in \widehat{z} + \mathcal{L}$: $S_\varphi(x, \check{z}) + S_\varphi(\check{x}, z) = S_\varphi(\check{x}, \check{z})$.
Biz.: Ez a Tétel azonnali következménye.

Kicsit messzebbről nézve

Lineáris feltételrendszerű, szeparábilis, konvex, nemlineáris optimalizálási probléma, *separable convex nonlinear optimization problem* (SCNP):
$A \in \mathbb{R}^{m\times n}, \quad b \in \mathbb{R}^m, \quad \forall j \in \{1,2,\dots,n\}: \quad f_j$ konvex

$$\min \sum_{j=1}^{n} f_j(x_j)$$

$$Ax = b$$

$$x \geq 0$$

Egyik legfontosabb speciális esete:
Entrópia-programozás, *entropy optimization problem*, (EOP)
$\forall j \in \{1,2,\dots,n\}: \quad f_j \in C^2(\mathbb{R}_+)$ és

$$f_j''(x_j) = \mu_j(x_j^{d_j} + q_j),$$

ahol $\mu_j \geq 0$, $q_j \geq 0$, $d_j \in \mathbb{R}$.
Példák: $x\log(x) - x$, $-\log(x) + x$, $-\frac{x^\alpha}{\alpha} + x\ (\alpha < 1, \alpha \neq 0)$.

Általánosított entrópia-programozás, *extended entropy programming problem*, (EEP)
$\forall j \in \{1, 2, \ldots, n\} : \quad g_j \in C^3(\mathbb{R})$, konvex és

$$|g_j'''(x_j)| \leq \frac{\kappa_j g_j''(x_j)}{x_j},$$

ahol $\kappa_j > 0$. Maga az f_j függvény pedig $f_j(x_j) = c_j x_j + g_j(x_j)$ alakú. Másként írva a cél:

$$\min c^T x + \sum_{j=1}^{n} g_j(x_j)$$

Young-programozás
$\widehat{z} \in \mathbb{R}^n$, $\widehat{z} > 0$

$$f_j(x_j) := S_\varphi(x_j, \widehat{z}_j) = (x_j - \psi_j(\widehat{z}_j))\widehat{z}_j - \int_{\psi_j(\widehat{z}_j)}^{x_j} \varphi(t)\,\mathrm{d}t$$

Ezek kapcsolata:
Áll.: Az (EOP) osztály része az (EEP) osztálynak.
Áll.: A Young jó nagy része az SCNP-nek. Pontosan fogalmazva:
Ha $\forall j \in \{1, 2, \ldots, n\}$: $f_j : \mathbb{R}_+ \to \mathbb{R}_+$ folytonosan differenciálható, szigorúan konvex, (egyetlen) minimumhelye $x_j^* > 0$, és $f_j(x_j^*) = 0$, valamint $\lim_{t\to 0+} f_j'(t) = -\infty$ és $\lim_{t\to\infty} f_j'(t) =: z_j > 0$,
akkor $\forall j \in \{1, 2, \ldots, n\}$-re létezik $\varphi_j : \mathbb{R}_+ \to \mathbb{R}_+$ folytonos, szigorúan monoton csökkenő függvény, amelyre

$$\lim_{t\to 0+} \varphi_j(t) = \infty, \quad \lim_{t\to\infty} \varphi_j(t) = 0 \quad \text{és} \quad f_j(x_j) = S_{\varphi_j}(x_j, \widehat{z}_j).$$

Megj.: A legtöbb népszerű entrópia-függvény (EOP) osztály és a Young-osztály metszetében van.
Megj.: Az id^{2n} $(n \geq 1)$ függvények (EOP)-ben vannak, de nincsenek a Young-osztályban.

Tétel.: (Zehra Boratas, Illés Tibor, Kas Péter, 2002.) A $\varphi(t) := \frac{(2+\sin t)^{\frac{1}{t}}}{t}$-ből kapott Young-program nem (EEP)-beli. Sőt, a logaritmikus barrier-függvénye nem self-concordant.