Μεμ-293 Θεωρία Βελτιστοποίησης

Εξ αποστάσεως item 13

Πέμπτη 21/5/2020

(zoom meeting)

(5μμ-7μμ)

B 13.1

Το δυϊκό πρόβλημα γραμμικού προγραμματισμού

ΟΡΙΣΜΟΙ: Έστω $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $f: \mathbb{R}^n \to \mathbb{R}$ με $f(x) = (c,x)$ $\forall x \in \mathbb{R}^n$.
$g: \mathbb{R}^m \to \mathbb{R}$ με: $g(y) = (b,y)$ $\forall y \in \mathbb{R}^m$, και $A \in \mathbb{R}^{m \times n}$.

Το πρόβλημα γραμμικού προγραμματισμού: $\max\limits_{x \in S'} f(x)$ όπου:

$$S' := \{x \in \mathbb{R}^n : x \geq 0 \text{ και } Ax \boxed{\leq} b\}$$

λέμε ότι είναι ένα πρόβλημα γραμμικού προγραμματισμού σε ημικανονική μορφή. Το δυϊκό ενός προβλήματος σε ημικανονική μορφή έχει επιδιώκωση: $\min\limits_{y \in \hat{S}} g(y)$ όπου: $\tilde{S} := \{y \in \mathbb{R}^m : y \geq 0, A^T y \boxed{\geq} c\}$.

ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Στην ημικανονική μορφή δεν είναι απαραίτητο να έχουμε $b \geq 0$

2. Κάθε πρόβλημα γραμμικού προγραμματισμού μπορεί να διατυπωθεί ισοδύναμα σε ημικανονική μορφή. Ένας περιορισμός της μορφής $\sum\limits_{j=1}^{n} A_{ij} x_j \geq b_i$ παίρνει, προφανώς την μορφή $\sum\limits_{j=1}^{n} (-A_{ij}) x_j \leq -b_i$.

   Ένας ισοτικός περιορισμός $\sum\limits_{j=1}^{n} A_{ij} x_j = b_i$, αντικαθίσταται από τις ανισότητες:

$$\sum_{j=1}^{n} A_{ij} x_j \leq b_i$$

$$\sum_{j=1}^{n} A_{ij} x_j \geq b_i \Leftrightarrow \sum_{j=1}^{n} (-A_{ij}) x_j \leq -b_i$$

(B 13.2)

Θεώρημα:

Για κάθε $x \in S$ και $y \in \tilde{S}$ ισχύει ότι: $f(x) \boxed{\leq} g(y)$.

Απόδειξη: Έστω $x \in S$ και $y \in \tilde{S}$. Τότε $x \geq 0, y \geq 0$, $b \boxed{\geq} Ax$ και $A^T y \boxed{\geq} c$.

Άρα: $g(y) = (b,y) = \sum_{j=1}^{m} b_j y_j \underset{x \geq 0}{\boxed{\geq}} \sum_{j=1}^{m} (Ax)_j y_j = (Ax, y) = (x, A^T y) \underset{x \geq 0}{\boxed{\geq}} (c, x) = f(x)$. ■

Θεώρημα: Έστω $x^* \in S$ και $y^* \in \tilde{S}$, τέτοια ώστε: $f(x^*) = g(y^*)$.

Τότε το x^* είναι σημείο ολικού μεγίστου της f στο S και το y^* σημείο ολικού ελαχίστου της g στο $\tilde{S}$.

Απόδ.

Για κάθε $x \in S$ έχουμε: $f(x) \boxed{\leq} g(y^*) = f(x^*)$.

Για κάθε $y \in \tilde{S}$ έχουμε: $g(y^*) = f(x^*) \boxed{\leq} g(y)$. ■

ΘΕΩΡΗΜΑ: Αν το πρόβλημα $\max\limits_{x\in S} f(x)$ έχει βέλτιστη βασική μη εκφυλισμένη λύση, τότε το πρόβλημα $\min\limits_{y\in \tilde{S}} g(y)$ έχει βέλτιστη λύση.

Απόδειξη:

Έχουμε ως δεδομένο ότι το πρόβλημα $\max\limits_{x\in S} f(x)$ έχει βέλτιστη βασική μη εκφυλισμένη λύση $x^*\in S$.

Προσθέτουμε στις γραμμές του A το μοναδιαίο πίνακα, δηλ. προσθέτουμε επιπλέον μεταβλητές $(x_{n+i})_{i=1}^{m}$, και διαμορφώνουμε το πρόβλημα $\max\limits_{z\in D} \tilde{f}(z)$, όπου: $D=\{z\in\mathbb{R}^{n+m} : [A|I_m]z=b,\ z\geq 0\}$ και $\tilde{f}(z)=f(z_1,\dots,z_n)$ $\forall z\in\mathbb{R}^{n+m}$

Επειδή, $Ax^*\leq b$, μπορούμε να βρούμε $y^*\in\mathbb{R}^m$ τ.ω. $y^*\geq 0$ και $Ax^*+y^*=b$. Ορίζουμε $z^*\in\mathbb{R}^{n+m}$ με:

$$z_j^*=\begin{cases} x_j^*, & j=1,\dots,n \\ y^*_{j-n}, & j=n+1,\dots,n+m \end{cases},\quad j=1,\dots,n+m.$$

Τότε: $z^*\geq 0$ και $[A|I_m]z^*=Ax^*+y^*=b$. Άρα: $z^*\in D$ με: $\tilde{f}(z^*)=f(x^*)$.

Έστω: $z\in D$. Τότε: ορίζουμε $x\in\mathbb{R}^n$ με: $x_j=z_j$, $j=1,\dots,n$, και $y\in\mathbb{R}^m$ με: $y_j=z_{n+j}$, $j=1,\dots,m$. Επειδή, $z\geq 0$, έπεται ότι: $x\geq 0$ και $y\geq 0$. Επιπλέον, $[A|I_m]z=b \Rightarrow Ax+y=b \Rightarrow Ax\leq b$. Έτσι: $x\in S$.

Άρα: $\tilde{f}(z)=f(x)\leq f(x^*)\leq \tilde{f}(z^*)$. Αυτό σημαίνει ότι το παραπάνω πρόβλημα σε σχεδόν (διότι δεν έχουμε $b\geq 0$) κανονική μορφή έχει βέλτιστη λύση.

B

(13.4

Επειδή υπάρχει βέλτιστη λύση $z^* \in D$ υπάρχει βέλτιστη βασική εφικτή [μη εκφυλισμένη] λύση z^B (το γεγονός ότι το b μπορεί να μην είναι θετικό δεν επηρεάζει το σχετικό θεώρημα). Ας είναι $I=\{i_\ell\}_{\ell=1}^{m}$ οι βασικές συντεταγμένες του z^B με $1\le i_1<i_2<\dots<i_m\le n+m$. Άρα οι αντίστοιχες βασικές στήλες του $[A|I_m]$ θα είναι: $(a_{i_\ell})_{\ell=1}^{m}$. Επομένως ο πίνακας $B=[a_{i_1}|\;\;|a_{i_m}]$ είναι ένας $m\times m$ αντιστρέψιμος πίνακας. Έτσι: $z^B=\sum_{\ell=1}^{m} z^B_{i_\ell} e_{i_\ell}$. Επειδή $[A|I_m]z^B=b$, έπεται ότι $B\begin{bmatrix} z^B_{i_1}\\ \vdots \\ z^B_{i_m}\end{bmatrix}=b \Leftrightarrow$

$\Leftrightarrow \begin{bmatrix} z^B_{i_1}\\ \vdots \\ z^B_{i_m}\end{bmatrix}=B^{-1}b$. Επιπλέον, ορίζουμε: $\begin{bmatrix} y^k_{i_1}\\ \vdots \\ y^k_{i_m}\end{bmatrix}=B^{-1}a_k$ και $y^k_j=0$ όταν $j\notin I$, $1\le k\le n+m$. Επειδή το z^B είναι βέλτιστη εφικτή μη εκφυλισμένη λύση σε πρόβλημα εύρεσης μεγίστου έπεται: $\tilde{f}(y^k)-c_k\ge 0$, $k=1,\dots,n+m$. Έστω $c^B=\begin{bmatrix} c_{i_1}\\ \vdots \\ c_{i_m}\end{bmatrix}$ όπου: $c_j=0$, $j=n+1,\dots,n+m$. Τότε:

$$\sum_{\ell=1}^{m} c^B_{i_\ell} y^k_{i_\ell}\ge c_k \Leftrightarrow \sum_{\ell=1}^{m} c^B_{i_\ell}(B^{-1}a_k)_{i_\ell}\ge c_k$$

$k=1,\dots,n+m$ $\qquad$ $k=1,\dots,n+m$

Όταν $k=n+1,\dots,n+m$, τότε: $c_k=0$. Έτσι: $\sum_{\ell=1}^{m} c^B_{i_\ell}(B^{-1}a_k)_{i_\ell}\ge 0 \Leftrightarrow \sum_{\ell=1}^{m} c^B_{i_\ell}(B^{-1}\underbrace{e_k}_{\in\mathbb{R}^m \text{ στοιχείο της κανονικής βάσης}})_{i_\ell}\ge 0$. Έτσι: $\boxed{(c^B)^T B^{-1}\ge 0}$

Όταν $k=1,\dots,n$, τότε $\sum_{\ell=1}^{m} c^B_{i_\ell}(B^{-1}a_k)_{i_\ell}\ge c_k$. Άρα: $\boxed{(c^B)^T B^{-1}A\ge c}$

Ορίζουμε: $\underbrace{w^T}_{} = \underbrace{(c_B)^T}_{1\times m} \underbrace{B^{-1}}_{m\times m}$. Από τα προηγούμενα έχουμε: $w \geq 0$ και $w^T A \geq c$, άρα: $w \in \tilde{S}$. (13,5

Επιπλέον,
$g(w) = b^T w = w^T b = (c_B)^T B^{-1} b = (c_B)^T \begin{pmatrix} z^B_{i_1} \\ z^B_{i_m} \end{pmatrix} = f(z^B) = f(x^*)$.

Άρα: η w είναι βέλτιστη λύση του δυϊκού προβλήματος

$\boxed{\Xi}$

Θεώρημα: Αν η f δεν είναι φραγμένη στο S, τότε $\tilde{S} = \phi$

Απόδ. Έστω ότι: $\tilde{S} \neq \phi$. Τότε υπάρχει $w \in \tilde{S}$ και: $g(w) \geq f(x)$ $\forall x \in S$ Άτοπο.

(4) Θεώρημα: Αν $x^* \in S$ είναι η βέλτιστη λύση του $\max_S f$ και $w^* \in \hat{S}$ η βέλτιστη λύση του $\min_{\hat{S}} g$

τότε: $x^*_i \left(\sum_{j=1}^{m} A_{ji} w^*_j - c_i \right) = 0, \quad i = 1, \dots, n$

Απόδ. Η προς απόδειξη σχέση σημαίνει: $f(x^*) = (A^T w^*, x^*) = .$ Σημειώνουμε ότι: $(A^T w^*, x^*) = (w^*, Ax^*) = (w^*, b) = g(w^*)$

Στην προηγούμενη απόδειξη: $(w^*)^T = c_B^T B^{-1} \Leftrightarrow (w^*)^T B = c_B^T \Leftrightarrow (w^*_1, \dots, w^*_m) [\alpha_{i1} | \cdots | \alpha_{im}] = (c_{i_1}, \dots, c_{i_m})$, που είναι η ζητούμενη όταν $i \in I$. Όταν $i \notin I$ τότε $x^*_i = 0$. ∎