Μ293 ΘΕΩΡΙΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

7η εξ αποστάσεως διάλεξη

Πέμπτη 30/4/2020

5μμ-7μμ

(zoom)

version 1

B7.1

$$\min_{x \in S} f(x) \qquad f: \mathbb{R}^n \to \mathbb{R}, \text{ με: } f(x) = (c, x) \ \forall x \in \mathbb{R}^n \quad \text{όπου: } c \in \mathbb{R}^n \text{ και } c \neq 0$$

$$S = \{ x \in \mathbb{R}^n : Ax = b,\ x \geq 0 \} \quad \text{όπου: } A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^n,\ b \geq 0$$

Αναδρομή:

Θεώρημα:

α) Όταν $S \neq \emptyset$, τότε υπάρχει βασική εφικτή λύση $x_B \in S$

β) Όταν υπάρχει $x^* \in S$ με: $f(x^*) = \min_S f$, τότε υπάρχει βασική εφικτή λύση $x^*_B \in S$ με: $f(x^*_B) = f(x^*)$.

Θεώρημα

Έστω $x \in S$. Τότε: (x είναι ένα ακραίο σημείο του S) ανν (x είναι μια βασική εφικτή λύση) του.

Πόρισμα 1.

Όταν $S \neq \emptyset$, τότε το S έχει τουλάχιστον ένα ακραίο σημείο.

Πόρισμα 2.

Όταν υπάρχει $x^* \in S$ τ.ω $f(x^*) = \min_{x \in S} f(x)$, τότε υπάρχει ακραίο σημείο $z^* \in S$ τ.ω $f(z^*) = f(x^*)$.

Πόρισμα 3.

Το S έχει πεπερασμένο πλήθος ακραίων σημείων.

Τέλος Αναδρομής

Συμπέρασμα: Όταν το S είναι μη κενό και φραγμένο το πρόβλημα ελαχιστοποίησης $\min_{x \in S} f(x)$ λύνεται βρίσκοντας τα ακραία σημεία του S και εντοπίζοντας σε ποιό από αυτά λαμβάνει η f την ελάχιστη τιμή.

Μας ενδιαφέρει να καταλήξουμε σε μία αλγοριθμική διαδικασία που χρειαζόμαστε επιπλέον αποτελέσματα.

Έστω ότι $\operatorname{rank}(A)=m$. (Όταν $\operatorname{rank}(A) < m$, τότε ο A έχει μία ή περισσότερες γραμμές γραμμικώς εξαρτημένες δηλ. κάποιοι ισοτικοί περιορισμοί πλεονάζουν). Έστω ακόμα $x^B \in S$ μία μη εκφυλισμένη βασική εφικτή λύση. Έτσι υπάρχουν m ακριβώς μη μηδενικές συντεταγμένες $(x^B_{i_\ell})_{\ell=1}^{m}$ και έστω $I:=\{i_\ell : \ell=1,\dots,m\}$. Έστω $(a_k)_{k=1}^{n} \subset \mathbb{R}^m$ οι στήλες του A. Επιπλέον, οι στήλες $(a_{i_\ell})_{\ell=1}^{m}$ είναι γραμμικώς ανεξάρτητες και αποτελούν μία βάση του $\mathbb{R}^m$. Έτσι έχουμε

$$a_k = \sum_{\ell=1}^{m} y^k_{i_\ell} a_{i_\ell}, \quad k=1,\dots,n,$$

δηλ. κάθε στήλη του A γράφεται κατά μοναδικό τρόπο ως γραμμικός συνδυασμός των στοιχείων της βάσης $(a_{i_\ell})_{\ell=1}^{m}$. Στη συνέχεια ορίζουμε $(y^k)_{k=1}^{n} \subset \mathbb{R}^n$ ως εξής:

$$y^k_j = \begin{cases} y^k_j & \text{όταν } j \in I \\ 0 & \text{όταν } j \notin I \end{cases}, \quad j=1,\dots,n.$$

που σημαίνει ότι $Ay^k = a_k$, $k=1,\dots,n$. Επιπλέον ορίζουμε τις βοηθητικές ποσότητες:

$$z^k := f(y^k) = \sum_{\ell=1}^{m} c_{i_\ell} y^k_{i_\ell}, \quad k=1,\dots,n.$$

ΘΕΩΡΗΜΑ 1 Έστω x^B βασική εφικτή λύση.

Όταν: $\boxed{z^k \le c_k \quad \text{για } k=1,\dots,n}$, τότε $\boxed{\text{η βασική εφικτή λύση } x^B \text{ είναι σημείο ολικού ελαχίστου της } f \text{ στο } S.}$

(B 7.3

Απόδειξη: Ο σκοπός μας είναι να δείξουμε ότι $f(x^B) \le f(v) \quad \forall v \in S$.

Έστω $v \in S$. Τότε $v \geq 0$ και $Av = b$. Επομένως έχουμε:

$$b = Av = \sum_{k=1}^{n} v_k a_k = \sum_{k=1}^{n} v_k \left(\sum_{l=1}^{m} y_{il}^k a_{il}\right)$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{m} v_k y_{il}^k a_{il}$$

$$= \sum_{l=1}^{m} \left(\sum_{k=1}^{n} v_k y_{il}^k\right) a_{il}.$$

και

$$b = Ax^B = \sum_{l=1}^{m} x_{il}^B a_{il}$$

Άρα: $x_{il}^B = \sum_{k=1}^{n} v_k y_{il}^k$, $l = 1, \dots, m$.

(B 7.4

Έτσι:

$$
\begin{aligned}
f(x^B) = (c, x^B) &= \sum_{j=1}^{n} c_j x_j^B \\
&= \sum_{\ell=1}^{m} c_{i_\ell} x_{i_\ell}^B \\
&= \sum_{\ell=1}^{m} c_{i_\ell} \left(\sum_{k=1}^{n} v_k y_{i_\ell}^k \right) \\
&= \sum_{k=1}^{n} v_k \left(\sum_{\ell=1}^{m} c_{i_\ell} y_{i_\ell}^k \right) \\
&= \sum_{k=1}^{n} v_k \left(\sum_{j=1}^{n} c_j y_j^k \right) = \sum_{k=1}^{n} v_k \, (c, y^k) \\
&= \sum_{k=1}^{n} v_k \, f(y^k) \\
&= \sum_{k=1}^{n} \underbrace{v_k}_{\geq 0} z^k \leq \sum_{k=1}^{m} v_k \cdot c_k = (c, v) = f(v).
\end{aligned}
$$

$\boxtimes$

Σημείωση. Όταν το x^B είναι μία εκφυλισμένη βασική λύση τότε $x^B=0$ ή το x^B έχει $1\leq p<m$ θετικές συντεταγμένες $(x_{i_\ell})_{\ell=1}^{p}$ και επομένως οι $(a_{i_\ell})_{\ell=1}^{p}$ είναι γραμμικώς ανεξάρτητες. Επειδή $\mathrm{rank}(A)=m$, μπορούμε στην περίπτωση $x^B=0$ να βρούμε m γραμμικώς ανεξάρτητες στήλες $(a_{i_\ell})_{\ell=1}^{m}$ και στην περίπτωση $x^B\neq 0$ να βρούμε στήλες $(a_{i_\ell})_{\ell=p+1}^{m}$ έτσι ώστε οι στήλες $(a_{i_\ell})_{\ell=1}^{m}$ να είναι γραμμικώς ανεξάρτητες. Στη συνέχεια ορίζουμε: $I=\{i_\ell\}_{\ell=1}^{m}$. Μέσα από αυτό το πλαίσιο ισχύει το προηγούμενο θεώρημα. $\square$

B.7.5

Θεώρημα 2.1 Έστω x^B μία μη εκφυλισμένη βασική λύση.

Όταν. $\boxed{\text{υπάρχει } k_0\in\{1,\dots,n\} \text{ τ.ω. } z^{k_0}>c_{k_0}}$ τότε

α) $k_0\notin I$

β) η μη εκφυλισμένη βασική λύση x^B δεν είναι σημείο ολικού ελαχίστου της f στο S.

Απόδειξη:

β) Αρκεί να βρούμε $y\in S$ τ.ω. $f(y)<f(x^B)$.

Από την υπόθεση του θεωρήματος έχουμε: $z^{k_0}>c_{k_0}$ για κάποιο $k_0\{1,\dots,n\}$. Επειδή $Ax^B=b$ έχουμε:

$$b=\sum_{\ell=1}^{m}x^B_{i_\ell}a_{i_\ell}$$

Επίσης έχουμε: $a_{k_0}=\sum_{\ell=1}^{m}y^{k_0}_{i_\ell}a_{i_\ell}$. Άρα

$$b-\theta a_{k_0}=\sum_{\ell=1}^{m}(x^B_{i_\ell}-\theta y^{k_0}_{i_\ell})a_{i_\ell}\quad \forall\theta\in\mathbb{R}$$

$$\Leftrightarrow\quad b=\sum_{\ell=1}^{m}(x^B_{i_\ell}-\theta y^{k_0}_{i_\ell})a_{i_\ell}+\theta a_{k_0}\quad\forall\theta\in\mathbb{R}$$

Επίσης:

$$f(x^B)-\theta(z^{k_0}-c_{k_0}) = f(x^B)-\theta f(y^{k_0})+\theta c_{k_0}$$
$$= \sum_{j=1}^{n} c_j x_j^B - \theta \sum_{j=1}^{n} c_j y_j^{k_0} + \theta c_{k_0}$$
$$= \sum_{l=1}^{m} c_{i_l} x_{i_l}^B - \theta \sum_{l=1}^{m} c_{i_l} y_{i_l}^{k_0} + \theta c_{k_0} = \sum_{l=1}^{m} c_{i_l}(x_{i_l}^B - \theta y_{i_l}^{k_0}) + \theta c_{k_0}$$
$$= f(x^B-\theta y^{k_0})+\theta c_{k_0} \quad \forall \theta\in\mathbb{R}.$$

Σημ. $e_{k_0}\in\mathbb{R}^n$ με: $(e_{k_0})_j = \begin{cases}1 & j=k_0\\ 0 & j\neq k_0\end{cases}$ για $j=1,\dots,n$. Είναι το k_0-στο στοιχείο της κανονικής βάσης του $\mathbb{R}^n$.

Για κάθε $\theta\in\mathbb{R}$ ορίζουμε: $\boxed{\hat{x}(\theta) := x^B-\theta y^{k_0}+\theta\cdot e_{k_0}}$. Τότε η παραπάνω σχέση παίρνει τη μορφή:

$$f(x^B)-\theta(z^{k_0}-c_{k_0}) = f(\hat{x}(\theta)), \quad \forall \theta\in\mathbb{R}.$$

Επιπλέον έχουμε:

$$A\hat{x}(\theta) = \theta a_{k_0} + \sum_{l=1}^{m}(x_{i_l}^B-\theta y_{i_l}^{k_0})a_{i_l} = b, \quad \forall\theta\in\mathbb{R}.$$

Περίπτωση 1 $\quad y_{i_l}^{k_0}\le 0$ για $l=1,\dots,m$ (δηλ. $y^{k_0}\le 0$)

Τότε για οποιοδήποτε $\theta>0$ έχουμε: $\hat{x}(\theta)\ge 0$ καθώς

$$(\hat{x}(\theta))_j = \begin{cases} 0, & \text{όταν } j\notin (I\cup\{k_0\}) \\ \underbrace{x_j^B - \theta \underbrace{y_j^{k_0}}_{\le 0}}_{\ge 0} + \underbrace{\theta (e_{k_0})_j}_{\ge 0} \ge x_j^B \ge 0, & \text{όταν } j\in I\cup\{k_0\}\end{cases}, \quad j=1,\dots,n.$$

Επομένως $\hat{x}(\theta)\in S$ $\forall\theta>0$, καθώς $\hat{x}(\theta)\ge 0$ και $A\hat{x}(\theta)=b$ για κάθε $\theta>0$.

Επιπλέον, έχουμε: $f(\hat{x}(\theta)) = f(x^B) - \underbrace{\theta}_{>0}\underbrace{(z^{k_0}-c_{k_0})}_{>0} < f(x^B)$, $\forall\theta>0$, και:

$$\boxed{f(\hat{x}(\theta)) = f(x^B)-\theta|z^{k_0}-c_{k_0}| \xrightarrow[\theta\to+\infty]{} -\infty}$$

που σημαίνει ότι η f δεν έχει σημείο ολικού ελαχίστου στο S.

(B.7.7

Περίπτωση 2: υπάρχει $l_0 \in \{1,\dots,m\}$ τ.ω $y^{k_0}_{i_{l_0}} > 0$.

Τότε ορίζουμε: $$\theta^0 := \min \underbrace{\left\{ \frac{x^B_j}{y^{k_0}_j} : j \in I \text{ και } y^{k_0}_j > 0 \right\}}_{\neq \emptyset} > 0$$ (Έχουμε $\theta^0 > 0$ επειδή η x^B δεν είναι εκφυλισμένη βασική λύση)

Επομένως: $f(\hat{x}(\theta^0)) = f(x^B) - \theta^0 |z^{k_0} - c_{k_0}| < f(x_B)$ Έχουμε ήδη ότι $A\hat{x}(\theta^0) = b$, επομένως χρειαζόμαστε να εξασφαλίσουμε ότι: $\hat{x}(\theta^0) \geq 0$ για να έχουμε $\hat{x}(\theta^0) \in S$. Παρατηρούμε ότι:

$$(\hat{x}(\theta^0))_j = x^B_j - \theta^0 y^{k_0}_j + \theta^0 (e_{k_0})_j \geq x^B_j - \theta^0 y^{k_0}_j = \begin{cases} 0 & \text{όταν } j \notin I \\ x^B_j + \theta^0 |y^{k_0}_j| > 0 & \text{όταν } j \in I \text{ και } y^{k_0}_j \leq 0 \\ y^{k_0}_j \left(\frac{x^B_j}{y^{k_0}_j} - \theta^0 \right) \geq 0 & \text{όταν } j \in I \text{ και } y^{k_0}_j > 0 \end{cases}, \; j = 1, \dots, n.$$

Πράγματι $\hat{x}(\theta^0) \geq 0$.

α). Ας υποθέσουμε ότι $k_0 \in I$. Επειδή τα $\{a_j\}_{j \in I}$ είναι γραμμικώς ανεξάρτητα

έπεται ότι: $$y^{k_0}_j = \begin{cases} 0 & j \notin I \\ 0 & j \in I \text{ και } j \neq k_0 \\ 1 & j \in I \text{ και } j = k_0 \end{cases}, \; j = 1, \dots, n,$$ διότι πρέπει να ισχύει: $a^{k_0} = \sum_{j \in I} y^{k_0}_j a_j$. Άρα: $z^{k_0} = f(y^{k_0}) = c_{k_0}$, άτοπο. ∎

Επομένως: $k_0 \notin I$.

(B.7.8

ΠΟΡΙΣΜΑ 1: Έστω $S \neq \emptyset$ και x^B μη εκφυλισμένη βασική εφικτή λύση.

$\boxed{z^k \leq c_k,\ k=1,\dots,n}$ ανν $\boxed{x^B \text{ είναι σημείο ολικού ελαχίστου της } f \text{ στο } S}$

ΠΟΡΙΣΜΑ 2: Έστω: $S \neq \emptyset$, S φραγμένο και x^B μια μη εκφυλισμένη βασική εφικτή λύση.

Όταν υπάρχει $k_0 \in \{1,\dots,n\}$ τ.ω. $z^{k_0} > c_{k_0}$, τότε υπάρχει βασική εφικτή λύση y^B καλύτερη από την x^B δηλ. $f(y^B) < f(x^B)$.

ΠΡΟΕΠΙΣΚΟΠΗΣΗ ΤΗΣ ΑΠΟΔΕΙΞΗΣ: Με βάση την απόδειξη του Θεωρήματος 2 πιο πάνω, επειδή η f δεν μπορεί να είναι φραγμένη (καθώς το S είναι φραγμένο) αναγκαστικά το y^{k_0} έχει τουλάχιστον μία αυστηρά θετική συντεταγμένη και μπορεί να κατασκευαστεί διάνυσμα $\hat{x}(\theta^0) \in S$ τ.ω $f(\hat{x}(\theta^0)) < f(x^B)$. Επιπλέον, το $\hat{x}(\theta^0)$ είναι βασική εφικτή λύση. ◇

Σημ. Η νέα βασική εφικτή λύση $y^B = \hat{x}(\theta^0)$ μπορεί να είναι εκφυλισμένη. Αν δεν είναι σημείο ολικού ελαχίστου της f στο S (βλ. Θεώρημα 1) τότε: δεν είναι σίγουρο ότι η διαδικασία του Θεωρήματος 2 μπορεί να τη βελτιώσει διότι μπορεί να προκύψει $\theta^0 = 0$.