ΜΕΜ 293 ΘΕΩΡΙΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

9η εξ αποστάσεως διάλεξη

Πέμπτη 7/5/2025

(5μμ-7μμ)

(zoom)

Version 2

Αναδρομή

Διατύπωση προβλήματος:

$$f:\mathbb{R}^n \to \mathbb{R} \quad \text{με} \quad f(x)=\langle c,x\rangle \quad \forall x\in\mathbb{R}^n$$

$$c\in\mathbb{R}^n \quad \text{με} \quad c\neq 0$$

$$\min_{x\in S} f(x)$$

$$S := \left\{ x\in\mathbb{R}^n : \begin{array}{l} Ax=b \\ x\geq 0 \end{array} \right\} \qquad A\in\mathbb{R}^{m\times n}, \quad b\geq 0$$

"κανονική" μορφή του προβλήματος γραμμικού προγραμματισμού.

$$A=[\underbrace{a^1,\dots,a^n}_{\text{στήλες του } A}]$$

οι οποίες είναι στοιχεία του $\mathbb{R}^m$

Όταν $S=\phi$ τότε το πρόβλημα δεν είναι καλά ορισμένο. Όταν $S\neq\phi$ και μη φραγμένο τότε η f μπορεί να μην έχει σημείο ολικού ελαχίστου. Όταν $S\neq\phi$ και S φραγμένο, επειδή το S είναι κλειστό, έπεται ότι η f έχει σημείο ολικού ελαχίστου στο S.

Έστω x^B μία βασική εφικτή λύση. Αυτό σημαίνει ότι υπάρχει σύνολο δεικτών $I\subset\{1,\dots,n\}$ με m ακριβώς στοιχεία τ.ω. $x^B_i=0$ $\forall i\notin I$, και τα $\{a^k\}_{k\in I}$ να είναι γραμμικώς ανεξάρτητα. Όταν $x^B_i>0$ $\forall i\in I$, τότε λέμε ότι το x^B είναι μία μη εκφυλισμένη βασική λύση, διαφορετικά λέμε ότι μία εκφυλισμένη βασική λύση.

Έχουμε ήδη δείξει ότι αν το πρόβλημα ελαχιστοποίησης έχει λύση τότε υπάρχει σημείο ολικού ελαχίστου της f στο S η οποία είναι βασική εφικτή λύση. Επομένως αναζητώντας για ένα αλγόριθμο εύρεσης σημείου ολικού ελαχίστου μπορεί να είναι η αναζήτηση των "βέλτιστης" βασικής εφικτής λύσης.

Επίσης έχουμε δείξει ότι οι βασικές εφικτές λύσεις είναι ακραία σημεία του S.

Αναγκαία συνθήκη για την ύπαρξη μιάς μη εκφυλισμένης βασικής εφικτής λύσης είναι $\mathrm{rank}(A)=m$. Επειδή $\mathrm{rank}(A)\leq\min\{n,m\}$ αναγκαστικά $n\geq m$. Αποφεύγουμε την περίπτωση $m=n$ διότι τότε $S=\phi$ ή $S=\{x^0\}$ που είναι ειδικές περιπτώσεις.

Υπάρχουν εκφυλισμένες βασικές λύσεις όταν: $b\neq 0$ και όταν το b είναι μη μηδενικό και μπορεί να γραφεί ως γραμμικός συνδυασμός λιγότερο από m το πλήθος γραμμικά ανεξάρτητων στηλών του A με θετικούς συντελεστές.

Αναδρομή

Έστω x^B μία βασική εφικτή λύση. Τότε υπάρχει $I \subset \{1,\dots,n\}$ με: $card(I) = m$ τ.ω.

$x^B_i = 0 \ \forall i \notin I$ και $(a^k)_{k\in I}$ να είναι γραμμικώς ανεξάρτητα δηλ. να αποτελούν μία βάση B του $\mathbb{R}^m$.

Έτσι: $\boxed{a^k = \sum_{i\in I} y^k_i a^i, \ k=1,\dots,n. \quad \text{με: } \sum_{i\in I} |y^k_i| > 0}$ Ορίζουμε: $y^k \in \mathbb{R}^n$ ως εξής: $\boxed{y^k_j = \begin{cases} y^k_j & \text{όταν } j \in I \\ 0 & \text{όταν } j \notin I \end{cases}, \ j=1,\dots,n.}$

Επιπλέον ορίζουμε: $z^k = f(y^k)$ για $k = 1,\dots,n$, και παρατηρούμε ότι: $\boxed{Ay^k = a^k, \ k=1,\dots,n}$

Θεώρημα: $z^k \leq c_k$ για $k=1,\dots,n \implies$ η x^B είναι σημείο ολικού ελαχίστου της f στο S.

Θεώρημα: Αν x^B μη εκφυλισμένη βασική λύση και υπάρχει $k_0 \in \{1,\dots,n\}$ τ.ω. $z^{k_0} > c_{k_0}$ $\implies$

α) η x^B δεν είναι σημείο ολικού ελαχίστου της f στο S.

β) αν $y^{k_0} \leq 0$ τότε η f δεν έχει ολικό ελάχιστο

γ) $k_0 \notin I$

Σημ. Όταν $k_0 \in I$ τότε $y^{k_0}_j = \begin{cases} 0 & j \neq k_0 \\ 1 & j = 1 \end{cases}$. Άρα $z^{k_0} = f(y^{k_0}) = c_{k_0}$, δηλ. δεν μπορεί να ισχύει: $z^{k_0} > c_{k_0}$.

Αναδρομή.

ΠΟΡΙΣΜΑ 1 Έστω x^B μία μη εκφυλισμένη βασική εφικτή λύση.

το x^B είναι σημείο ολικού ελαχίστου της f στο S	ανν	$z^k \leq c^k$ για $k=1,\dots,n$

ΠΟΡΙΣΜΑ 2. Υποθέτουμε ότι $S \neq \emptyset$ και ότι η f έχει σημείο ολικού ελαχίστου στο S.
Έστω x^B μία μη εκφυλισμένη βασική εφικτή λύση ότι υπάρχει: $k_0 \notin I$ τ.ω. $z^{k_0} > c_{k_0}$.
Τότε υπάρχει βασική εφικτή λύση w^B με $f(w^B) < f(x^B)$.

Τέλος Αναδρομής

Απόδειξη. Ορίζουμε: $\hat{x}(\theta) = x^B - \theta y^{k_0} + \theta e_{k_0}$ για κάθε $\theta > 0$. Τότε: $f(\hat{x}(\theta)) = f(x^B) - \theta |z^{k_0} - c_{k_0}| < f(x^B)$ $\forall \theta > 0$.

Επιπλέον: $A\hat{x}(\theta) = \underset{\sim}{Ax^B} - \theta \underset{\sim}{Ay^{k_0}} + \theta \underset{\sim}{Ae_{k_0}} = \underset{\sim}{b} - \theta \underset{\sim}{a^{k_0}} + \theta \underset{\sim}{a^{k_0}} = b$ $\forall \theta > 0$.

Όταν $y^{k_0} \leq 0$ τότε: $\hat{x}_j(\theta) = x_j^B + \theta |y_j^{k_0}| + \theta \delta_{k_0 j} \geq x_j^B \geq 0$, $j=1,\dots,n$, $\forall \theta > 0$. Άρα $\hat{x}(\theta) \geq 0$ $\forall \theta > 0$. Όμως $\lim_{\theta \to +\infty} f(\hat{x}(\theta)) = -\infty$, που σημαίνει ότι η f δεν έχει σημείο ολικού ελαχίστου στο S. Άτοπο.

Άρα: $y_{l_0}^{k_0} > 0$ για κάποιο $l_0 \in I$. Ορίζουμε $\theta^0 = \min\left\{ \frac{x_j^B}{y_j^{k_0}} : j \in I \text{ και } y_j^{k_0} > 0 \right\} > 0$. Τότε:

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

ΔΝ. $\hat{x}(\theta^0) \geq 0$. Άρα $\hat{x}(\theta^0)$ είναι εφικτή λύση και $f(\hat{x}(\theta^0)) < f(x^B)$.

Έστω: $j_0 \in I$ με $y^{k_0}_{j_0} > 0$ τ.ω. $\theta^0 = \frac{x^B_{j_0}}{y^{k_0}_{j_0}}$. Τότε: $(\hat{x}(\theta^0))_{j_0} = x^B_{j_0} - \theta^0 y^{k_0}_{j_0} = 0$.

Έτσι ενώ το x^B έχει m θετικές συντεταγμένες $(x^B_j)_{j\in I}$, το $\hat{x}(\theta^0)$ έχει το πολύ m θετικές συντεταγμένες με δείκτες πού περιέχονται στο σύνολο $J := \{k_0\} \cup (I \setminus \{j_0\})$.

Συνεπώς το $\hat{x}(\theta^0)$ "χάνει" τη μη μηδενική συντεταγμένη στη θέση j_0 και "κερδίζει" τη μη μηδενική συντεταγμένη στη θέση k_0. Διατύπωνεται και ως εξής: ((από τις βασικές μεταβλητές φεύγει η j_0 και εισέρχεται η k_0)).
Όταν το θ^0 ως minimum λαμβάνεται για διαφορετικά $j\in I$ με $y^{k_0}_j > 0$, δηλ. το j_0 δεν είναι μοναδικό, τότε η $\hat{x}(\theta^0)$ είναι μια εκφυλισμένη βασική εφικτή λύση.

Μένει να εξασφαλίσουμε ότι η $\hat{x}(\theta^0)$ είναι βασική εφικτή λύση, δηλ. τα $(a^\kappa)_{\kappa\in J}$ είναι γραμμικώς ανεξάρτητα. Έστω ότι είναι γραμμικώς εξαρτημένα. Τότε υπάρχουν $(\gamma_j)_{j\in J}$ τ.ω.

$$\gamma_{k_0} a^{k_0} + \sum_{\substack{j\in J \\ j\neq k_0}} \gamma_j a^j = 0 \quad \text{και} \quad \sum_{j\in J} |\gamma_j| > 0$$

Αν $\gamma_{k_0} = 0$ τότε: $\sum_{\substack{j\in J \\ j\neq k_0}} \gamma_j a^j = 0$ που σημαίνει ότι $\gamma_j = 0 \ \forall j \in J\setminus\{k_0\}$, επειδή $J\setminus\{k_0\} \subseteq I$ και τα $(a^j)_{j\in I}$ είναι γραμμ. ανεξάρτητα. Άρα άτοπο. Επομένως $\boxed{\gamma_{k_0} \neq 0}$

Όμως:

$$a^{k_0} = \sum_{j\in I} y_j^{k_0} a^j \Rightarrow \gamma_{k_0} a^{k_0} = \sum_{j \in I} \gamma_{k_0} y_j^{k_0} a^j$$

$$\Rightarrow -\sum_{\substack{j\in J \\ j\neq k_0}} \gamma_j a^j = \sum_{j\in I} \gamma_{k_0} y_j^{k_0} a^j$$

$$\Rightarrow -\sum_{j \in I \setminus \{j_0\}} \gamma_j a^j = \sum_{j\in I} \gamma_{k_0} y_j^{k_0} a^j$$

$$\Rightarrow \sum_{j\in I\setminus\{j_0\}} \left(\gamma_j + \gamma_{k_0} y_j^{k_0}\right) a^j + \gamma_{k_0} y_{j_0}^{k_0} a^{j_0} = 0$$

Λόγω της γραμμ. ανεξαρτησίας των $(a^j)_{j\in I}$ έπεται ότι: $\gamma_{k_0} y_{j_0}^{k_0} = 0$. Όμως: $y_{j_0}^{k_0} > 0$ και $\gamma_{k_0} \neq 0$, άρα καταλήξαμε σε άτοπο. Επομένως το $w^B = \hat{x}(\theta^0)$ είναι βασική εφικτή καλύτερη της x^B καθώς $f(w^B) < f(x^B)$.

Ουσιαστικά η απόδειξη μας οδηγεί στη μέθοδο Simplex

Μέθοδος Simplex

Η μέθοδος Simplex ξεκινά από μία μη εκφυλισμένη βασική εφικτή λύση, και ελέγχει αν είναι σημείο ολικού ελαχίστου της f. Αν δεν είναι υπολογίζει μία καλύτερη βασική εφικτή λύση. (βλ. Πόρισμα 1 και την απόδειξη του Πορίσματος 2).

Παράδειγμα.

$\min \ (-5x_1+4x_2)$

$$\begin{aligned} x_1 - x_2 + x_3 \quad\quad\quad &= 6 \\ 3x_1 - 2x_2 \quad + x_4 \quad &= 24 \\ -2x_1 + 3x_2 \quad\quad + x_5 &= 9 \end{aligned}$$

Εδώ $n=5$ και $m=3$. Επιπλέον έχουμε: $C=(-5,4,0,0,0)\in\mathbb{R}^5$, $b=\begin{bmatrix}6\\24\\9\end{bmatrix}>0 \ \in\mathbb{R}^3$

$$A=\begin{bmatrix}1&-1&1&0&0\\3&-2&0&1&0\\-2&3&0&0&1\end{bmatrix}\in\mathbb{R}^{3\times5}$$

με $\mathrm{rank}(A)=3$ καθώς οι στήλες 3,4,5 είναι προφανώς γραμμικά ανεξάρτητες.

		c	c_1	c_2	c_3	c_4	c_5	
B	c^B	x^B	a_1	a_2	a_3	a_4	a_5	θ
a_3	c_3	x^B_3	y^1_3	y^2_3	y^3_3	y^4_3	y^5_3	
a_4	c_4	x^B_4	y^1_4	y^2_4	y^3_4	y^4_4	y^5_4	
a_5	c_5	x^B_5	y^1_5	y^2_5	y^3_5	y^4_5	y^5_5	
		$f(x^B)$	z^1	z^2	z^3	z^4	z^5	z^k
			z^1-c_1	z^2-c_2	z^3-c_3	z^4-c_4	z^5-c_5	z^k-c_k

Επειδή ο A περιέχει τον 3x3 μοναδιαίο πίνακα (βλ. στήλες 3,4,5) μια προφανής βασική λύση είναι: $x^B=b_1a_3+b_2a_4+b_3a_5=(0,0,6,24,9)$, η οποία αντιστοιχεί στις γραμμικώς ανεξάρτητες στήλες $B=\{a_3,a_4,a_5\}$ και οι βασικές συντεταγμένες έχουν δείκτες $I=\{3,4,5\}$. Η μέθοδος Simplex αναπτύσσεται σε tableaux και κάθε tableau αντιστοιχεί σε ένα στάδιο της μεθόδου. Κάτω από το B βάζουμε σε κάθε γραμμή το σύμβολο a_i με $i\in I$. Κάτω από το c^B βάζουμε σε κάθε γραμμή, δίπλα στο a_i, την τιμή c_i, για $i\in I$. Κάτω από το x^B στην ίδια γραμμή με το a_i βάζουμε το x^B_i για $i\in I$. Κάτω από τη στήλη a_i και στη γραμμή της στήλης a_j γράφουμε τα y^i_j για $i=1,\dots,5$ και $j\in I$.

Αν για κάποιο $k_0\notin I$ έχουμε $z^{k_0}-c_{k_0}>0$, τότε κάτω από το θ στη γραμμή της a_j γράφουμε το πηλίκο $\frac{x^B_j}{y^{k_0}_j}$ όταν $y^{k_0}_j>0$, διαφορετικά το αφήνουμε κενό.

B.9.7

		c	-5	4	0	0	0	
B	c^B	x^B	a_1	a_2	a_3	a_4	a_5	θ
a_3	0	6	$\boxed{1}$	-1	1	0	0	6/1
a_4	0	24	3	-2	0	1	0	24/3
a_5	0	9	-2	3	0	0	1	—
		0	0	0	0	0	0	z^k
			5	-4	0	0	0	z^k-c_k

$k_0=1$

$j_0=3$ $\quad \theta^0=6$

$y^{k_0}_{j_0}=y^1_3=1$ οδηγός

tableaux 1

$z^1-c_1=5>0$ άρα $k_0=1$

$$\theta^0=\min\left\{\frac{x^B_j}{y^{k_0}_j} : j\in I \text{ και } y^{k_0}_j>0\right\}=\min\Big\{\underbrace{\tfrac{6}{1}}_{j=3},\underbrace{\tfrac{24}{3}}_{j=4}\Big\}=\min\{6,8\}=6 . \text{ Άρα } j_0=3$$

Έτσι φεύγει η στήλη a_3 από τη βάση B και αντικαθίσταται από την στήλη a_1. Η νέα βασική εφικτή λύση είναι:

$$w^B=x^B-\theta^0 y^{k_0}+\theta^0 e_{k_0} = x^B-6y^1+6e_1$$

$$=6e_3+24e_4+9e_5+6e_1-6[e_3+3e_4-2e_5]$$

$$=6e_1+\cancel{6e_3}+24\hat{e}_4+9\check{e}_5-\cancel{6e_3}-18\hat{e}_4+12\check{e}_5$$

$$=6e_1+6e_4+21e_5=(6,0,0,6,21),$$ η οποία είναι μη εκφυλισμένη.

Επίσης: $f(w^B)=-30<0=f(x^B)$

Στο επόμενο tableau ξεκινάμε με την μη εκφυλισμένη βασική λύση w^B στη θέση της x^B (δεν αλλάζουμε το συμβολισμό στο tableau) και πρέπει να βρούμε τους συντελεστές των μη βασικών στηλών ως προς τις νέες βασικές στήλες. Σ' αυτή τη διαδικασία βασική παράμετρος είναι ο θετικός αριθμός $y^{k_0}_{j_0}=y^1_3$ που καλείται οδηγός (pivot).