Θεωρία Γραμμικού Προγραμματισμού

Διάλεξη 3: 15.10.2014 Διδάσκων: Σταύρος Κολλιόπουλος

Γραφείς: Κατερίνα Δρόσου, Γιώργος Παπαδημητρίου

3.1 Ακραία Σημεία - Αλγεβρικός Ορισμός

Έστω ένα σύνολο $P \subset \mathbb{R}^n$ και $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$ τ.ω. $a_i^T x \geq b_i, i \in M_1$, $a_i^T x \leq b_i, i \in M_2$, $a_i^T x = b_i, i \in M_3$, όπου τα M_1, M_2, M_3 είναι πεπερασμένα σύνολα δεικτών.

Ορισμός 3.1. *Έστω $x^* \in \mathbb{R}^n$ μία λύση που ικανοποιεί τον περιορισμό $a_i^T x^* = b_i$ για κάποιο $i \in M_1 \uplus M_2 \uplus M_3$, ο αντίστοιχος περιορισμός καλείται* ενεργός (active) *στο x^*.*

Θεώρημα 3.1. *Έστω μία λύση $x^* \in \mathbb{R}^n$ και το σύνολο $I = \{i \mid a_i^T x^* = b_i\}$. Οι παρακάτω προτάσεις είναι ισοδύναμες:*

1. *Υπάρχουν n διανύσματα στο $\{a_i \mid i \in I\}$ τα οποία είναι γραμμικά ανεξάρτητα.*
2. *Ο γραμμικός υπόχωρος $\operatorname{span} \bigcup_{i \in I} \{a_i\}$ είναι ο $\mathbb{R}^n$.*
3. *Το σύστημα $a_i^T x = b_i$, $i \in I$, έχει μοναδική λύση.*

Απόδειξη: Το ότι η πρόταση (1) ισοδυναμεί με την πρόταση (2) είναι προφανές.

Θα δείξουμε ότι το (2) ισοδυναμεί με το (3).

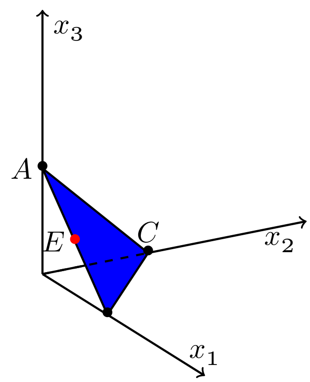
Αν δεν ισχύει το (3) τότε δεν ισχύει το (2), διότι: Αν το σύστημα έχει λύσεις τις x^1, x^2 με $x^1 \neq x^2$, τότε: ορίζουμε $d := x^1 - x^2$ και $a_i^T (x^1 - x^2) = 0, \forall i \in I$. Άρα $d \perp a_i, \forall i \in I$. Άρα $d \notin \operatorname{span} \bigcup_{i \in I} \{a_i\}$.

Αν δεν ισχύει το (2) τότε δεν ισχύει το (3), διότι: Αν $\operatorname{span} \bigcup_{i \in I} \{a_i\} \neq \mathbb{R}^n$, δηλαδή $\operatorname{span} \bigcup_{i \in I} \{a_i\} \subset \mathbb{R}^n$, διαλέγουμε $d \neq 0$ κάθετο στον υπόχωρο. Αν $a_i^T \hat{x} = b_i$, $\forall i \in I$, τότε $a_i^T (\hat{x} + b) = b_i$, $\forall i \in I$. Δηλαδή το συτημα έχει πολλαπλές λύσεις, άρα δεν ισχύει το (3).

Επομένως δείξαμε ότι και το (3) είναι ισοδύναμο με το (2). ∎

Παρατήρηση 3.1. *Θα λέμε ότι* οι περιορισμοί είναι γραμμικά ανεξάρτητοι, *αν τα αντίστοιχα διανύσματα συντελεστών a_i είναι γραμμικά ανεξάρτητα.*

Παράδειγμα για την έννοια του ενεργού περιορισμού:



Έστω το πολύεδρο του σχήματος, $P = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 1 \text{ με } x_1, x_2, x_3 \geq 0\}$. Στο σημείο C έχουμε 3 ενεργούς περιορισμούς: $x_1 + x_2 + x_3 = 1$, $x_1 = 0$ και $x_3 = 0$. Στο σημείο E έχουμε 2 ενεργούς περιορισμούς: $x_1 + x_2 + x_3 = 1$ και $x_2 = 0$.

Ορισμός 3.2. *Έστω πολύεδρο $P \subseteq \mathbb{R}^n$ που ορίζεται από ισότητες και ανισότητες. Έστω $x^* \in \mathbb{R}^n$. Τότε:*

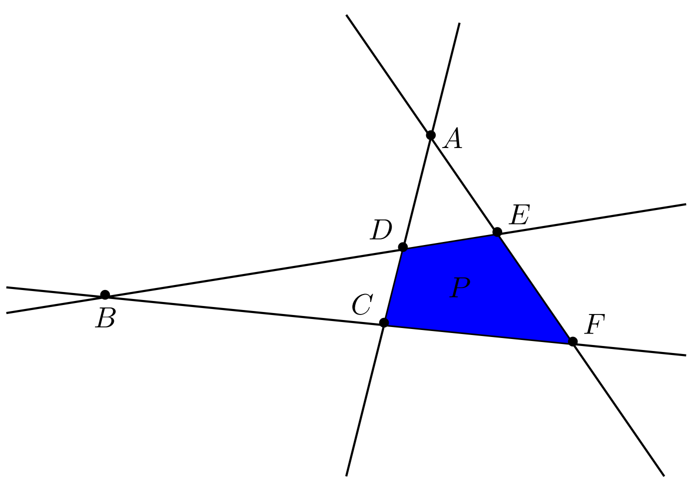
1. *Η λύση x^* καλείται* βασική λύση (β.λ.) *αν*

   (α′) όλες οι ισότητες είναι ενεργές (στο x^*).

   (β′) από όλους τους ενεργούς περιορισμούς, n εξ αυτών είναι γραμμικά ανεξάρτητοι.

2. *Η λύση x^* καλείται* βασική εφικτή λύση (β.ε.λ.) *αν είναι βασική και επιπλέον $x^* \in P$.*

Παράδειγμα 3.1. *Στο σχήμα τα A, B, C, D, E, F είναι β.λ. ενώ (μόνο) τα C, D, E, F β.ε.λ.*



Σχήμα 3.1: Παράδειγμα β.λ. και β.ε.λ.

Θεώρημα 3.2. *Έστω ένα πολύεδρο $P \subseteq \mathbb{R}^n$, με $P \neq \emptyset$ και $x^* \in P$. Οι παρακάτω προτάσεις είναι ισοδύναμες:*

1. *x^* είναι κορυφή του P*
2. *x^* είναι ακραίο σημείο του P.*
3. *x^* είναι β.ε.λ.*

Απόδειξη: Χωρίς βλάβη της γενικότητας υποθέτουμε ότι το P ορίζεται από περιορισμούς της μορφής $a_i^T x \geq b_i$ και $a_i^T x = b_i$.

α): $(1) \Rightarrow (2)$

Έστω x^* κορυφή του P. Υπάρχει $c \in \mathbb{R}^n$ τ.ω. $c^T x^* < c^T y$, $\forall y \in P$ και $y \neq x$. Αν $y \in P$, $z \in P$, $y \neq x^*$, $z \neq x^*$ με $0 \leq \lambda \leq 1$. Θα δείξουμε ότι το x^* δεν είναι κυρτός συνδυασμός των y, z.

$$\left.\begin{array}{l} c^T x^* < c^T y \\ c^T x^* < c^T z. \end{array}\right\} \Rightarrow c^T x^* < c^T(\lambda y + (1-\lambda)z) \Rightarrow x^* \neq \lambda y + (1-\lambda)z.$$

Άρα το x^* ακραίο σημείο.

β): $(2) \Rightarrow (3)$

Έστω x^* δεν είναι β.ε.λ. Θα δείξουμε ότι δεν είναι ακραίο σημείο. Έστω το σύνολο $I = \{i \mid a_i^T x^* = b_i\}$. Αφού x^* δεν είναι β.ε.λ. αυτό σημαίνει ότι $\mathrm{span}\{a_i \mid i \in I\} \subset \mathbb{R}^n$. Τότε θα υπάρχει $d \in \mathbb{R}^n$ με $d \neq 0$ τ.ω. $a_i^T d = 0, \forall i \in I$. Έστω $\epsilon > 0$ ένας «μικρός» αριθμός. Ορίζουμε $y := x^* + \epsilon d$ και $z := x^* - \epsilon d$. Έχουμε επίσης ότι το x^* είναι μεταξύ των y και z, αφού $x^* = \frac{y+z}{2}$. Δηλαδή το x^* είναι κυρτός συνδυασμός των y και z. Αρκεί να δείξουμε ότι τα δύο αυτά σημεία ανήκουν κι αυτά στο P (άρα το x^* όχι ακραίο σημείο).

Διακρίνουμε δύο περιπτώσεις. Για τα $i \in I$: $a_i^T y = a_i^T(x^* + \epsilon d) = a_i^T x^* = b_i = a_i^T z$.

Αν $i \notin I$, έχουμε: $a_i^T x^* > b_i$, οπότε $a_i^T y = a_i^T(x^* + \epsilon d) = a_i^T x^* + a_i^T \epsilon d$. Άρα για "μικρό" ϵ, $a_i^T y > b_i$.

Άρα το $y \in P$. Ομοίως και $z \in P$.

γ): $(3) \Rightarrow (1)$

Έστω x^* β.ε.λ. Ορίζουμε και πάλι το σύνολο $I = \{i \mid a_i^T x^* = b_i\}$.

Επίσης έστω $c := \sum_{i \in I} a_i$.

Έχουμε:

$$c^T x^* = \sum_{i \in I} a_i^T x^* = \sum_{i \in I} b_i \tag{3.1}$$

Επίσης, $\forall x \in P$, ισχύει $a_i^T x \geq b_i$ οπότε:

$$c^T x = \sum_{i \in I} a_i^T x \geq \sum_{i \in I} b_i \tag{3.2}$$

Άρα από τις (3.1) και (3.2) προκύπτει ότι η x^* είναι η βέλτιστη λύση στο πρόβλημα:

$$\min\{c^T x \mid x \in P\}. \tag{3.3}$$

Μένει να δείξουμε ότι το x^* είναι και μοναδικό.

Η σχέση (3.2) ισχύει με ισότητα ανν $a_i^T x = b_i, \forall i \in I$ και αφού x^* β.ε.λ., υπάρχουν n γραμμικά ανεξάρτητοι ενεργοί περιορισμοί στο x^*. Άρα από Θεώρημα 3.1 το x^* μοναδική λύση στο σύστημα $a_i^T x = b_i$. ∎

Πόρισμα 3.1. *Το πολύεδρο $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$, $A \in \mathbb{R}^{m \times n}$ έχει πεπερασμένο αριθμό β.ε.λ.*

Απόδειξη: Σε κάθε β.ε.λ. αντιστοιχούν n γραμμικά ανεξάρτητοι περιορισμοί που ικανοποιούνται με ισότητα. Οποιοιδήποτε n τέτοιοι ορίζουν ένα μοναδικό σημείο. Επομένως διαφορετικές βασικές λύσεις αντιστοιχούν σε διαφορετικές n-άδες γραμμικά ανεξάρτητων περιορισμών. Άρα έχουμε το πολύ $\binom{m}{n}$ βασικές εφικτές λύσεις. ∎

Δύο διακεκριμένες β.λ. στο $\mathbb{R}^n$ καλούνται *γειτονικές* αν $n-1$ γραμμικά ανεξάρτητοι περιορισμοί είναι ενεργοί και στις δύο.

Παρατήρηση 3.2. *Ο Αλγόριθμος Simplex, αυτό που κάνει είναι να επισκέπτεται γειτονικές βασικές εφικτές λύσεις για να βρει την βέλτιστη.*

Στο Σχήμα 3.1 οι κορυφές D, E είναι γειτονικές στην B και οι A, C γειτονικές στην D.