Θεωρία Γραμμικού Προγραμματισμού

Διάλεξη 2: 14.10.2014 Διδάσκων: Σταύρος Κολλιόπουλος

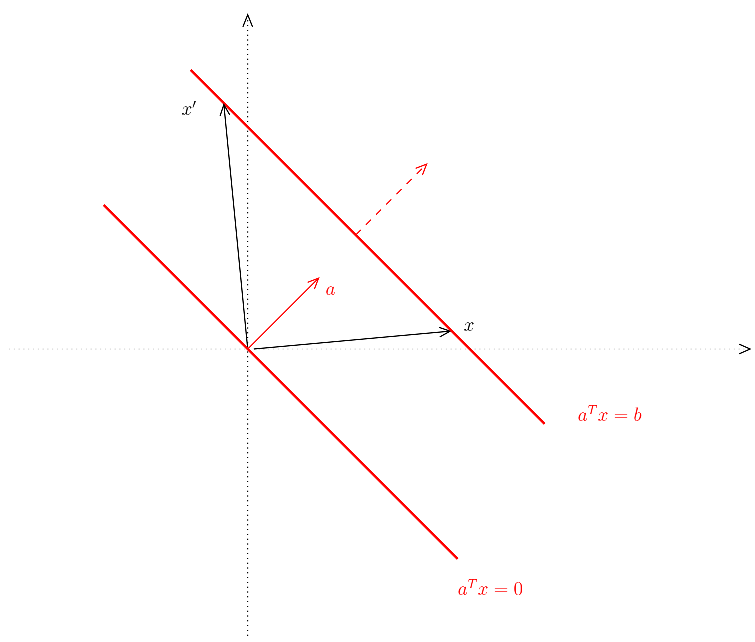
Γραφείς: Κατερίνα Δρόσου, Γιώργος Παπαδημητρίου & Σ. Κ.

2.1 Γραμμικά Προγράμματα

Ένα γραμμικό πρόγραμμα είναι ένα πρόβλημα βελτιστοποίησης μιας γραμμικής συνάρτησης $c : \mathbb{R}^n \to \mathbb{R}$ υπό γραμμικούς περιορισμούς. Η μεταβλητή $x \in \mathbb{R}^n$ παίρνει τιμές στην τομή ενός πεπερασμένου συνόλου ημιχώρων.

Ορισμός 2.1 *Δοθέντος διανύσματος $a \in \mathbb{R}^n$ με $a \neq 0$ και $b \in \mathbb{R}$ καλούμε:*

1. Υπερεπίπεδο (hyperplane) *ένα σύνολο της μορφής* $\{x \in \mathbb{R}^n \mid a^T x = b\}$.
2. Ημίχωρο (halfspace) *ένα σύνολο της μορφής* $\{x \in \mathbb{R}^n \mid a^T x \leq b\}$.



Σχήμα 2.1: Το διάνυσμα a είναι κάθετο στο υπερεπίπεδο $\{x \mid a^T x = 0\}$.

Στο Σχήμα 2.1 τα x, x' είναι διανύσματα που ανήκουν στο υπερεπίπεδο $\{x \mid a^T x = b\}$. Η Η διαφορά τους είναι κάθετη στο διάνυσμα a μιας και $a^T x = a^T x' \Rightarrow a^T (x - x') = 0$. Άμεση συνέπεια του Ορισμού 2.1 είναι ότι κάθε ημίχωρος είναι πολύεδρο και κάθε υπερεπίπεδο είναι κι αυτό πολύεδρο, ως τομή δύο ημιχώρων.

Ορισμός 2.2 Γραμμικό πρόγραμμα (linear program) *καλείται ένα πρόβλημα βελτιστοποίησης της μορφής*

$$\min\{c^T x \mid Ax \leq b\} \tag{2.1}$$

όπου $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ *και* $c \in \mathbb{R}^n$. *Το πολύεδρο* $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ *καλείται* σύνολο των εφικτών λύσεων (feasible set) *ή* εφικτή περιοχή *του γραμμικού προγράμματος. Το γραμμικό πρόγραμμα καλείται* εφικτό *αν η εφικτή του περιοχή είναι μη κενή και* ανέφικτο *ειδάλλως.*

Συχνά χρησιμοποιούμε τη συντομογραφία LP αντί του όρου γραμμικό πρόγραμμα. *Γραμμικό πρόγραμμα σε πρότυπη μορφή (standard form)* καλείται ένα γραμμικό πρόγραμμα της μορφής

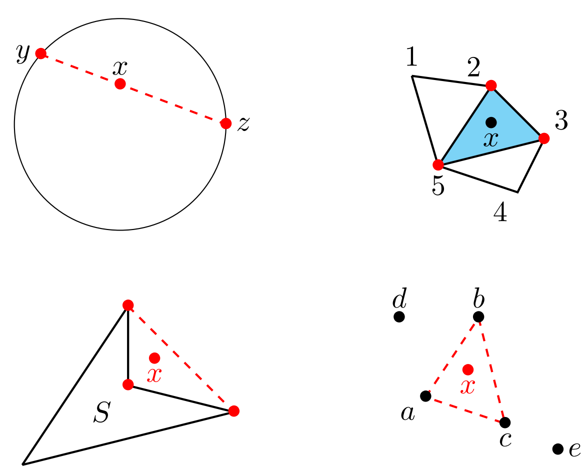
$$\min\{c^T x \mid Ax = b,\ x \geq 0\}.$$

Οποιοδήποτε LP της μορφής (2.1) μπορεί να μετατραπεί σε ένα ισοδύναμο σε πρότυπη μορφή. Προσθέτουμε μία μεταβλητή περιθωρίου (slack variable) σε κάθε ανίσωση και περιορισμό μη αρνητικότητας για κάθε μεταβλητή περιθωρίου. Οι ανισότητες $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$, $i \in [m]$, μετατρέπονται σε $\sum_{j=1}^{n} a_{ij}x_j + s_i = b_i$ όπου για κάθε μεταβλητή περιθωρίου πρέπει $s_i \geq 0$. Αφήνεται ως άσκηση η μετατροπή που απαιτείται για τις x μεταβλητές.

Λέμε ότι ένα πολύεδρο $P \subseteq \mathbb{R}^n$ είναι σε *πρότυπη μορφή* αν ορίζεται ως $P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$.

2.2 Θεώρημα Καραθεοδωρή

Παρατήρηση 2.1 *Στο* $\mathbb{R}^n$, $n+2$ *ή παραπάνω σημεία είναι πάντα αφινικά εξαρτημένα. Διότι, αν* $x_1, x_2, \ldots, x_n, x_{n+1}, x_{n+2}$ *αφινικά ανεξάρτητα, τότε* $x_2 - x_1, \ldots, x_{n+1} - x_1, x_{n+2} - x_1$ *γραμμικά ανεξάρτητα.*



Πίνακας 2.1: Παραδείγματα συνόλων στο $\mathbb{R}^2$. Το σημείο x είναι πάντοτε κυρτός συνδυασμός το πολύ 3 σημείων του συνόλου.

Παρατηρήστε ότι στο θεώρημα που ακολουθεί το σύνολο S δεν χρειάζεται να είναι ούτε πεπερασμένο ούτε κυρτό.

Θεώρημα 2.1 (Θεώρημα Καραθεοδωρή) *Έστω σύνολο $S \subseteq \mathbb{R}^n$. Κάθε $x \in \mathrm{conv}(S)$ μπορεί να γραφεί ως κυρτός συνδυασμός το πολύ $n+1$ αφινικά ανεξάρτητων σημείων του S.*

Απόδειξη. Έστω $x \in \mathrm{conv}(S)$. Από τον Ορισμό 1.4, υπάρχουν $\lambda_1, \ldots, \lambda_t > 0$ με $\sum_{j=1}^t \lambda_j = 1$ τ. ώ. $x = \lambda_1 x_1 + \cdots + \lambda_t x_t$ και $x_1, \ldots, x_t \in S$.

Αν $x_1, \ldots, x_t$ αφινικά ανεξάρτητα τελειώσαμε. Αν $x_1, \ldots, x_t$ αφινικά εξαρτημένα, τότε υπάρχουν $\mu_1, \ldots, \mu_t$ με $\sum_j |\mu_j| \neq 0$ και $\sum_{j=1}^t \mu_j = 0$ τ. ώ. $\sum_{j=1}^t \mu_j x_j = 0$. Χωρίς βλάβη της γενικότητας (χβτγ) υποθέτουμε ότι $\mu_1 > 0$. Υπάρχει $a \geq 0$ τ. ώ. το σημείο x γράφεται ως εξής:

$$x = \sum_{j=1}^{t} (\lambda_j - a\mu_j) x_j.$$

Αν $a = 0$, έχουμε τον αρχικό κυρτό συνδυασμό. Θεωρούμε τη μέγιστη τιμή $a_0 \geq 0$ του a για την οποία $\lambda_j - a_0\mu_j \geq 0$ για κάθε j. Αυτή η τιμή είναι καλά ορισμένη αφού $a_0 \leq \lambda_1/\mu_1$. Επειδή $\sum_{j=1}^t \mu_j x_j = 0$ παίρνουμε

$$x = \sum_{j=1}^{t} (\lambda_j - a_0\mu_j) x_j \tag{2.2}$$

Όλοι οι συντελεστές του γραμμικού συνδυασμού στο δεξί μέλος της (2.2) είναι μη αρνητικοί και επίσης

$$\sum_{j=1}^{t} (\lambda_j - a_0\mu_j) = \sum_{j=1}^{t} \lambda_j - a_0 \sum_{j=1}^{t} \mu_j = 1$$

Πήραμε το x ως κυρτό συνδυασμό το πολύ $t-1$ σημείων. Επαναλαμβάνουμε την παραπάνω διαδικασία μέχρι το x να προκύψει ως κυρτός συνδυασμός αφινικά ανεξάρτητων σημείων. ∎

Η απόδειξη της ακόλουθης πρότασης αφήνεται ως άσκηση.

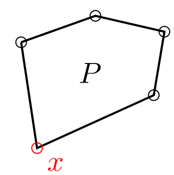
Θεώρημα 2.2 (Θεώρημα Καραθεοδωρή για κώνους) *Έστω σύνολο $S \subseteq \mathbb{R}^n$. Κάθε $x \in \mathrm{cone}(S)$ μπορεί να γραφεί ως κωνικός συνδυασμός το πολύ n γραμμικά ανεξάρτητων σημείων του S.*

2.3 Ακραία Σημεία, Κορυφές

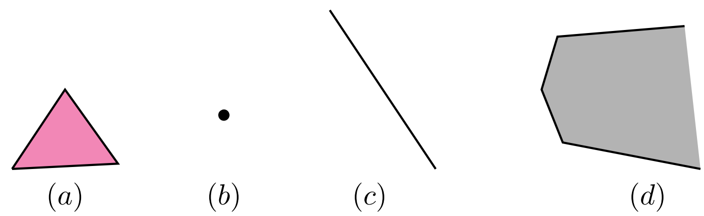
Ορισμός 2.3 *Έστω πολύεδρο P. Ένα σημείο $x \in P$ καλείται* ακραίο σημείο (extreme point) *του P αν δεν υπάρχουν $y, z \in P$ με $y, z \neq x$ και $\lambda \in (0,1)$, τ. ώ. $x = \lambda y + (1-\lambda) z$.*

Στο πολύεδρο του Σχήματος 2.2, τα κυκλωμένα σημεία είναι ακραία. Παρατηρούμε ότι δεν υπάρχουν σημεία z, y που να ανήκουν στο πολύεδρο και το x να είναι κυρτός συνδυασμός αυτών. Ένα σύνολο $S \subseteq \mathbb{R}^n$ καλείται *φραγμένο* αν υπάρχει σταθερά k τ. ώ. $\forall x \in S$, $\|x\|_\infty \leq k$.

Ορισμός 2.4 (Πολύτοπα) *Ένα σύνολο σημείων στο $\mathbb{R}^n$ καλείται* πολύτοπο *αν είναι το κυρτό κάλυμα πεπερασμένου πλήθους διανυσμάτων.*



Σχήμα 2.2: Τα ακραία σημεία του πολυέδρου P.



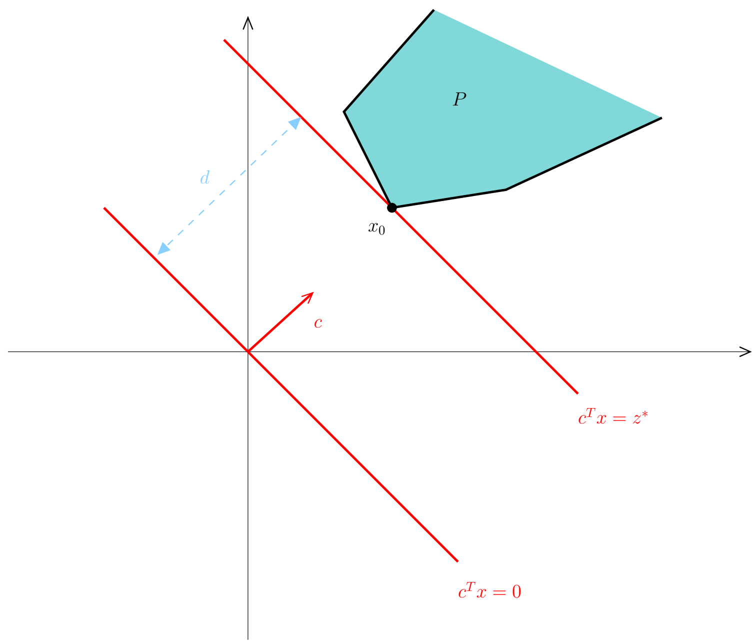
Σχήμα 2.3: Παραδείγματα πολυέδρων. Τα (a) και (b) είναι και πολύτοπα.

Από τον Ορισμό 2.4 προκύπτει ότι ένα πολύτοπο είναι φραγμένο σύνολο. Στη Διάλεξη 10 θα αποδείξουμε τυπικά ότι (i) ένα σύνολο P είναι πολύτοπο αν και μόνο αν είναι φραγμένο πολύεδρο και (ii) κάθε πολύτοπο είναι το κυρτό κάλυμμα των ακραίων σημείων του. Βλ. Σχήμα 2.3 για παραδείγματα.

Ένα σημείο x ενός πολυέδρου P καλείται κορυφή αν υπάρχει κάποιο γραμμικό πρόγραμμα με εφικτή περιοχή P για το οποίο το x είναι η μοναδική βέλτιστη λύση. Ισοδύναμα έχουμε τον ακόλουθο ορισμό.

Ορισμός 2.5 *Έστω $P \subseteq \mathbb{R}^n$ πολύεδρο. Ένα σημείο $x \in P$ καλείται* κορυφή (vertex) *του P αν υπάρχει $c \in \mathbb{R}^n$ τ. ώ. $\forall y \in P, y \neq x : c^T x < c^T y$.*

Δείτε το παράδειγμα στο Σχήμα 2.4.



Σχήμα 2.4: Στην κορυφή x_0 του πολυέδρου P ελαχιστοποιείται η αντικειμενική συνάρτηση $c^T x$ με βέλτιστη τιμή z^*. Η ποσότητα z^* είναι ανάλογη της απόστασης d.