Σημειώσεις
Κυρτής Ανάλυσης

Α. ΓΙΑΝΝΟΠΟΥΛΟΣ

Σ. ΠΑΠΑΔΟΠΟΥΛΟΥ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΦΘΙΝΟΠΩΡΟ 2000

Περιεχόμενα

1 Πραγματικές κυρτές συναρτήσεις **5**
- 1.1 Πραγματικές κυρτές συναρτήσεις . . . 5
- 1.2 Χαρακτηρισμοί των κυρτών συναρτήσεων . . . 11
- 1.3 Παραγωγίσιμες κυρτές συναρτήσεις . . . 14
- 1.4 Θεωρήματα για ολοκληρώματα . . . 15
- 1.5 Η συζυγής συνάρτηση . . . 18
- 1.6 Ασκήσεις . . . 24

2 Κυρτά σύνολα στον Ευκλείδειο χώρο **29**
- 2.1 Τοπολογία του Ευκλείδειου χώρου . . . 29
- 2.2 Κυρτά σύνολα στον Ευκλείδειο χώρο . . . 33
- 2.3 Υπερεπίπεδα και διαχωριστικά θεωρήματα . . . 39
- 2.4 Πολυεδρικοί κώνοι - το λήμμα του Farkas . . . 46
- 2.5 Ασκήσεις . . . 51

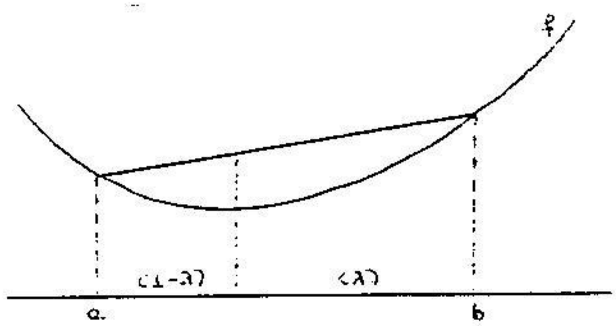
3 Κυρτές συναρτήσεις στον $\mathbb{R}^n$ **55**
- 3.1 Ορισμοί . . . 55
- 3.2 Συνέχεια και παραγωγισιμότητα . . . 61
- 3.3 Συναρτήσεις στήριξης . . . 67
- 3.4 Υποδιαφορικά . . . 73
- 3.5 Συζυγής συνάρτηση . . . 78
- 3.6 Το θεώρημα δυϊσμού του Fenchel . . . 81
- 3.7 Ασκήσεις . . . 83

4 Εφαρμογές **87**
- 4.1 Θεωρία παιγνίων . . . 87
- 4.2 Κυρτός προγραμματισμός . . . 92
- 4.3 Ανάλυση πινάκων . . . 98
- 4.4 Ασκήσεις . . . 105

Κεφάλαιο 1

Πραγματικές κυρτές συναρτήσεις

Σε αυτό το κεφάλαιο, με I συμβολίζουμε ένα (κλειστό, ανοικτό ή ημιανοικτό, πεπερασμένο ή άπειρο) διάστημα στο $\mathbb{R}$.



Σχήμα 1

1.1 Πραγματικές κυρτές συναρτήσεις

§1. **Ορισμός** Έστω $f : I \to \mathbb{R}$ μία συνάρτηση.

(α) Η f λέγεται κυρτή αν

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b) \tag{1}$$

για κάθε $a, b \in I$ και κάθε $\lambda \in \mathbb{R}$ με $0 < \lambda < 1$. Στο Σχήμα 1 μπορείτε να δείτε την γεωμετρική σημασία του ορισμού: η χορδή που έχει σαν άκρα τα σημεία $(a, f(a))$ και $(b, f(b))$ δεν είναι πουθενά κάτω από το γράφημα της f.

(β) Η f λέγεται *γνήσια κυρτή* αν είναι κυρτή και έχουμε γνήσια ανισότητα στην (1) για κάθε $a, b \in I$ με $a \neq b$.

§2. Ισοδύναμοι τρόποι με τους οποίους μπορεί να περιγραφεί η κυρτότητα της $f : I \to \mathbb{R}$ είναι οι εξής:

(α) Για κάθε $a, b, x \in I$ με $a < x < b$,

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b).$$

Παρατηρήστε ότι το δεξιό μέλος αυτής της ανισότητας ισούται με

$$f(a) + \frac{f(b)-f(a)}{b-a}(x-a).$$

(β) Για κάθε $a, b \in I$ και κάθε $\lambda, \mu \in \mathbb{R}$ με $\lambda > 0$, $\mu > 0$ και $\lambda + \mu = 1$,

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b).$$

§3. Οι παρακάτω απλές ιδιότητες των κυρτών συναρτήσεων είναι άμεσες συνέπειες του ορισμού:

(α) Αν $f, g : I \to \mathbb{R}$ είναι δύο κυρτές συναρτήσεις και $a \geq 0$, $b \geq 0$, τότε η $af + bg$ είναι κυρτή συνάρτηση.

(β) Το άθροισμα πεπερασμένων το πλήθος κυρτών συναρτήσεων είναι κυρτή συνάρτηση.

(γ) Το κατά σημείο όριο μιάς συγκλίνουσας ακολουθίας κυρτών συναρτήσεων είναι κυρτή συνάρτηση.

(δ) Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, αν $x_1, \ldots, x_m \in I$ και $\lambda_1, \ldots, \lambda_m \geq 0$ με $\lambda_1 + \ldots + \lambda_m = 1$, έχουμε $\sum_{i=1}^m \lambda_i x_i \in I$ και

$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i).$$

(ε) Έστω $\{f_j : j \in J\}$ τυχούσα οικογένεια κυρτών συναρτήσεων $f_j : I \to \mathbb{R}$. Ορίζουμε $f : I \to \mathbb{R}$ με $f(x) = \sup\{f_j(x) : j \in J\}$. Αν η f είναι πεπερασμένη παντού στο I, τότε η f είναι κυρτή.

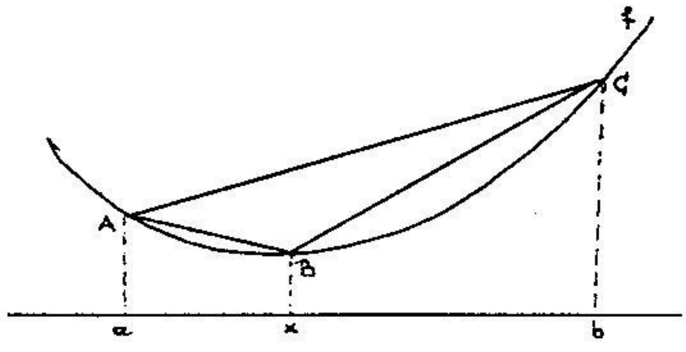
Οι αποδείξεις αυτών των ισχυρισμών αφήνονται σαν ασκήσεις για τον αναγνώστη.

Θεώρημα 1.1 *Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Τότε,*

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x} \tag{2}$$

για κάθε $a, b, x \in I$ με $a < x < b$. Αντίστροφα, αν κάποια από τις δύο ανισότητες στην (2) ισχύει για κάθε $a < x < b$ στο I, τότε η f είναι κυρτή.

Αν η f είναι γνήσια κυρτή, τότε έχουμε γνήσιες ανισότητες στην (2).



Σχήμα 2

Στο Σχήμα 2 μπορείτε να δείτε την γεωμετρική σημασία αυτού του θεωρήματος: η κλίση του AB είναι μικρότερη από την κλίση του AC, η οποία με τη σειρά της είναι μικρότερη από την κλίση του BC.

Απόδειξη: Αφού η f είναι κυρτή, έχουμε

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b). \tag{3}$$

Από αυτήν την ανισότητα βλέπουμε ότι

$$f(x) - f(a) \leq \frac{a-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) = \frac{x-a}{b-a}[f(b) - f(a)],$$

το οποίο αποδεικνύει την αριστερή ανισότητα στην (2). Για την δεξιά ανισότητα εργαζόμαστε με ανάλογο τρόπο. Αν η f είναι γνήσια κυρτή, τότε έχουμε γνήσια ανισότητα στην (3), άρα και στην (2) □

§4. Συμβολίζουμε με $\mathrm{int}(I)$ το εσωτερικό του I. Θεωρούμε μία κυρτή συνάρτηση $f : I \to \mathbb{R}$ και ένα σημείο $c \in \mathrm{int}(I)$. Αν $[a,b]$ είναι ένα υποδιάστημα του I τέτοιο ώστε $a < c < b$ από το Θεώρημα 1.1 έχουμε

$$\frac{f(c) - f(a)}{c-a} \leq \frac{f(x) - f(c)}{x-c}$$

για κάθε $x \in (c,b]$. Πάλι από το Θεώρημα 1.1, η συνάρτηση

$$x \to \frac{f(x) - f(c)}{x-c}$$

είναι αύξουσα στο $(c,b]$. Επομένως, η δεξιά παράγωγος

$$f'_{-}(c) := \lim_{x \downarrow c} \frac{f(x) - f(c)}{x-c}$$

υπάρχει. Με ανάλογο τρόπο μπορούμε να δείξουμε ότι η αριστερή παράγωγος $f'_-(c)$ υπάρχει.

Αν $a < c < d < b$, τότε για αρκετά μικρό θετικό h έχουμε

$$\frac{f(c) - f(c-h)}{h} \le \frac{f(c+h) - f(c)}{h} \le \frac{f(d) - f(d-h)}{h}.$$

Παίρνοντας όρια καθώς $h \downarrow 0$, συμπεραίνουμε ότι

$$f'_-(c) \le f'_+(c) \le f'_-(d).$$

Επίσης, αν $x > c$ έχουμε

$$\frac{f(x) - f(c)}{x - c} \ge \frac{f(c+h) - f(c)}{h}$$

για κάθε $0 < h < x - c$, οπότε αφήνοντας το $h \downarrow 0$ βλέπουμε ότι

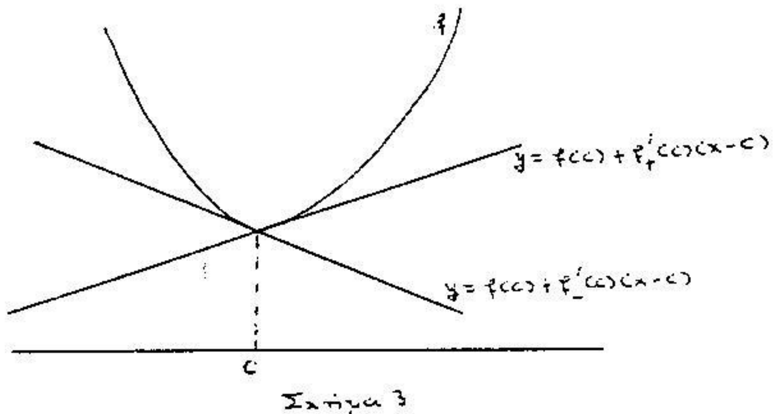
$$f(x) \ge f(c) + f'_+(c)(x - c) \ge f(c) + f'_-(c)(x - c).$$

ενώ αν $x < c$ έχουμε

$$\frac{f(x) - f(c)}{x - c} \le \frac{f(c-h) - f(c)}{-h}$$

για κάθε $0 < h < c - x$, οπότε αφήνοντας το $h \downarrow 0$ βλέπουμε ότι

$$f(x) \ge f(c) + f'_-(c)(x - c) \ge f(c) + f'_-(c)(x - c).$$



Σχήμα 3

Έχουμε λοιπόν αποδείξει το εξής θεώρημα.

Θεώρημα 1.2 *Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, η f έχει δεξιά και αριστερή παράγωγο σε κάθε εσωτερικό σημείο του I και οι f'_- και f'_+ είναι αύξουσες συναρτήσεις στο $\mathrm{int}(I)$. Αν $c \in \mathrm{int}(I)$, τότε*

$$f'_-(c) \le f'_+(c)$$

και

$$f(x) \geq f(c) + f'_-(c)(x-c), \qquad f(x) \geq f(c) + f'_+(c)(x-c)$$

για κάθε $x \in I$. □

Ειδικότερα, η ύπαρξη των πλευρικών παραγώγων εξασφαλίζει ότι η f είναι συνεχής στο εσωτερικό του I:

Θεώρημα 1.3 *Κάθε κυρτή συνάρτηση* $f : I \to \mathbb{R}$ *είναι συνεχής στο* $\mathrm{int}(I)$.

Απόδειξη: Έστω $c \in \mathrm{int}(I)$. Τότε, για μικρά $h > 0$ έχουμε $c+h, c-h \in I$ και

$$f(c+h) = f(c) + \frac{f(c+h) - f(c)}{h} \cdot h \to f(c) + f'_+(c) \cdot 0 = f(c)$$

όταν $h \downarrow 0$, ενώ, τελείως ανάλογα,

$$f(c-h) = f(c) + \frac{f(c-h) - f(c)}{-h} \cdot (-h) \to f(c) + f'_-(c) \cdot 0 = f(c)$$

όταν $h \downarrow 0$. Άρα, η f είναι συνεχής στο c. □

Παρατήρηση Έστω $f : [a,b] \to \mathbb{R}$ κυρτή συνάρτηση. Η απόδειξη του Θεωρήματος 1.2 δείχνει ότι, σε αυτήν την περίπτωση, οι $f'_+(a)$ και $f'_-(b)$ υπάρχουν αν επιτρέπουμε και τις τιμές $+\infty$ και $-\infty$ σαν όρια.

§5. Έστω $f : I \to \mathbb{R}$ και I_0 ένα υποδιάστημα του I. Λέμε ότι η f είναι *συνάρτηση Lipschitz* στο I_0 αν υπάρχει σταθερά $K > 0$ με την ιδιότητα $|f(x) - f(y)| \leq K|x-y|$ για κάθε $x, y \in I_0$. Από την συνθήκη αυτή έπεται ότι η f είναι ομοιόμορφα συνεχής στο I_0.

Θεώρημα 1.4 *Έστω* $f : I \to \mathbb{R}$ *κυρτή συνάρτηση και* $[a,b] \subset \mathrm{int}(I)$. *Τότε, η* f *είναι συνάρτηση Lipschitz στο* $[a,b]$.

Απόδειξη: Υπάρχουν $c, d \in I$ τέτοια ώστε $c < a < b < d$. Τότε, αν $a \leq x < y \leq b$, από το Θεώρημα 1.2 έχουμε

$$f'_+(a) \leq f'_+(x) \leq \frac{f(x) - f(y)}{x - y} \leq f'_-(y) \leq f'_-(b).$$

Έπεται ότι $|f(x) - f(y)| \leq K|x-y|$, όπου $K := \max\{|f'_+(a)|, |f'_-(b)|\}$. □

Παρατήρηση Σημειώνουμε ότι η f δεν είναι αναγκαστικά συνάρτηση Lipschitz σε ολόκληρο το I, ακόμα και αν υποθέσουμε ότι η f είναι φραγμένη. Επίσης, η f δεν είναι αναγκαστικά συνεχής στο I, ακόμα και αν υποθέσουμε ότι το I είναι κλειστό και φραγμένο διάστημα (Άσκηση 7).

§6. Περνάμε τώρα στη μελέτη της παραγωγισιμότητας των κυρτών συναρτήσεων.

Θεώρημα 1.5 *Έστω* f *κυρτή συνάρτηση. Τότε,*

(α) *η* f'_- *είναι συνεχής από αριστερά και η* f'_+ *είναι συνεχής από δεξιά στο* $\mathrm{int}(I)$.

(β) *υπάρχουν το πολύ αριθμήσιμα το πλήθος σημεία στα οποία η* f *δεν έχει παράγωγο.*

Απόδειξη: (α) Λόγω της συνέχειας της f στο $\text{int}(I)$ (Θεώρημα 1.3), για κάθε $x < y \in \text{int}(I)$ έχουμε

$$\frac{f(y)-f(x)}{y-x} = \lim_{z\downarrow x}\frac{f(y)-f(z)}{y-z} \geq \lim_{z\downarrow x} f'_+(z).$$

(το τελευταίο όριο υπάρχει γιατί η f'_+ είναι αύξουσα: δείτε την Άσκηση 16). Παίρνοντας όριο καθώς $y \downarrow x$, βλέπουμε ότι

$$f'_+(x) \geq \lim_{z\downarrow x} f'_+(z).$$

Όμως η f'_+ είναι αύξουσα (Θεώρημα 1.2), επομένως

$$f'_+(x) \leq \lim_{z\downarrow x} f'_+(z).$$

Άρα $f'_+(x) = \lim_{z\downarrow x} f'_-(z)$, το οποίο αποδεικνύει τη συνέχεια της f'_- από δεξιά. Η συνέχεια της f'_- από αριστερά αποδεικνύεται εντελώς ανάλογα.

(β) Από το Θεώρημα 1.2, έχουμε

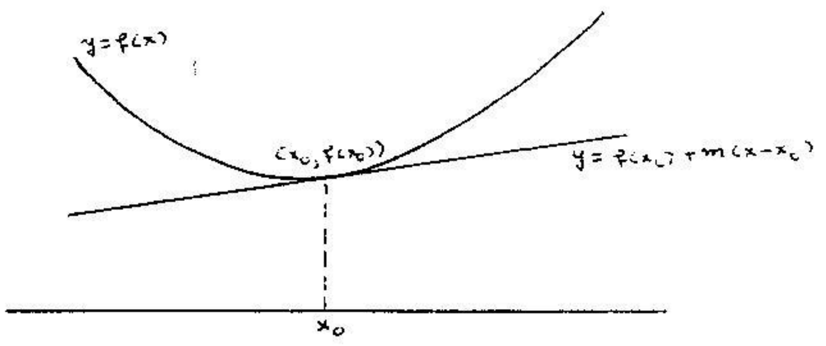
$$f'_+(x) \leq f'_-(y) \leq f'_-(z)$$

για κάθε $x, y, z \in \text{int}(I)$ με $x < y < z$. Αν η f'_- είναι συνεχής στο y, τότε

$$f'_-(y) = \lim_{x\uparrow y} f'_+(x) = \lim_{z\downarrow y} f'_-(z) = f'_-(y),$$

το οποίο σημαίνει ότι η f έχει παράγωγο στο y. Δηλαδή, αν η f δεν έχει παράγωγο σε κάποιο σημείο του $\text{int}(I)$, τότε η αύξουσα συνάρτηση f'_+ έχει ασυνέχεια σε αυτό. Αυτό αποδεικνύει το (β), γιατί μία αύξουσα συνάρτηση μπορεί να παρουσιάζει το πολύ αριθμήσιμα σημεία ασυνέχειας (Άσκηση 16). □

§7. **Ορισμός** Έστω I ανοικτό διάστημα και $f : I \to \mathbb{R}$ μία συνάρτηση. Αν $x_0 \in I$ και $h : \mathbb{R} \to \mathbb{R}$ γραμμική συνάρτηση, λέμε ότι η h *στηρίζει* την f στο x_0 αν $h(x_0) = f(x_0)$ και $h(x) \leq f(x)$ για κάθε $x \in I$. Σε αυτήν την περίπτωση, λέμε ότι η f έχει *στήριγμα* στο x_0.



Σχήμα 4

Θεώρημα 1.6 *Έστω I ανοικτό διάστημα και $f : I \to \mathbb{R}$. Τότε, η f είναι κυρτή αν* *και μόνο αν έχει στήριγμα σε κάθε σημείο του I.*

Απόδειξη: Υποθέτουμε πρώτα ότι η f έχει στήριγμα σε κάθε σημείο του I. Έστω $x, y \in I$ και $0 < \lambda < 1$. Αν h είναι ένα στήριγμα της f στο $\lambda x + (1-\lambda)y$, τότε

$$f(\lambda x + (1-\lambda)y) = h(\lambda x + (1-\lambda)y) = \lambda h(x) + (1-\lambda)h(y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Επομένως, η f είναι κυρτή.

Ας υποθέσουμε τώρα ότι η f είναι κυρτή. Έστω $c \in I$ και m τυχών αριθμός με την ιδιότητα $f'_-(c) \leq m \leq f'_+(c)$. Ορίζουμε

$$h(x) = f(c) + m(x-c).$$

Έστω $y, z \in I$ με $y < c < z$. Τότε,

$$\frac{f(y) - f(c)}{y - c} \leq f'_-(c) \leq m$$

άρα $f(y) \geq f(c) + m(y-c) = h(y)$, και

$$\frac{f(z) - f(c)}{z - c} \geq f'_+(c) \geq m$$

άρα $f(z) \geq f(c) + m(z-c) = h(z)$. Δηλαδή, η h στηρίζει την f στο c. $\square$

Θεώρημα 1.7 *Έστω I ανοικτό διάστημα και $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, η f είναι παραγωγίσιμη στο $c \in I$ αν και μόνο αν έχει μοναδικό στήριγμα στο c.*

Απόδειξη: Υποθέτουμε πρώτα ότι η f έχει παράγωγο στο c. Έστω ότι η $h(x) = f(c) + m(x-c)$ στηρίζει την f στο c. Τότε, αν $y, z \in I$ και $y < c < z$,

$$\frac{f(y) - f(c)}{y - c} \leq m \leq \frac{f(z) - f(c)}{z - c}.$$

Παίρνοντας $y \uparrow c$ και $z \downarrow c$, συμπεραίνουμε ότι $m = f'(c)$. Άρα, η f έχει μοναδικό στήριγμα στο c.

Για την αντίστροφη κατεύθυνση, παρατηρούμε ότι η $h(x) = f(c) + m(x-c)$ στηρίζει την f στο c για κάθε $m \in [f'_-(c),\ f'_+(c)]$ (δείτε την απόδειξη του προηγούμενου θεωρήματος). Αφού η f έχει μοναδικό στήριγμα στο c, πρέπει να ισχύει η $f'_-(c) = f'_+(c)$. Δηλαδή, η f έχει παράγωγο στο c. $\square$

1.2 Χαρακτηρισμοί των κυρτών συναρτήσεων

Σε αυτήν την παράγραφο δίνουμε δύο χαρακτηρισμούς της κυρτότητας για συνεχείς συναρτήσεις. Με τη βοήθεια τέτοιων χαρακτηρισμών είναι συχνά ευκολότερο να ελέγξουμε ότι κάποια (συγκεκριμένη) συνεχής συνάρτηση είναι κυρτή.

Θεώρημα 1.8 *Έστω $f : I \to \mathbb{R}$ συνεχής συνάρτηση που ικανοποιεί την*

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{2}[f(a)+f(b)] \tag{4}$$

για κάθε $a, b \in I$. Τότε, η f είναι κυρτή.

Απόδειξη: Δείχνουμε πρώτα με επαγωγή ως προς n ότι αν $N = 2^n$ και $a_1, \ldots, a_N \in I$ τότε

$$f\left(\frac{a_1+a_2+\ldots+a_N}{N}\right) \leq \frac{1}{N}\sum_{i=1}^{N} f(a_i). \tag{5}$$

Από την υπόθεσή μας, η (5) ισχύει για $n = 1$. Αν υποθέσουμε ότι ισχύει για $n \leq k$ και $a_1, \ldots, a_{2^k}, a_{2^k+1}, \ldots, a_{2^{k+1}} \in I$, τότε

$$\begin{aligned}
f\left(\frac{1}{2^{k+1}}\sum_{i=1}^{2^{k+1}} a_i\right) &= f\left(\frac{1}{2}\cdot\frac{1}{2^k}\sum_{i=1}^{2^k} a_i + \frac{1}{2}\cdot\frac{1}{2^k}\sum_{j=1}^{2^k} a_{2^k+j}\right)\\
&\leq \frac{1}{2}f\left(\frac{1}{2^k}\sum_{i=1}^{2^k} a_i\right) + \frac{1}{2}f\left(\frac{1}{2^k}\sum_{j=1}^{2^k} a_{2^k+j}\right)\\
&\leq \frac{1}{2}\cdot\frac{1}{2^k}\sum_{i=1}^{2^k} f(a_i) + \frac{1}{2}\cdot\frac{1}{2^k}\sum_{j=1}^{2^k} f(a_{2^k+j})\\
&= \frac{1}{2^{k+1}}\sum_{i=1}^{2^{k+1}} f(a_i)
\end{aligned}$$

όπου χρησιμοποιήσαμε διαδοχικά την επαγωγική υπόθεση για $n = 1$ και $n = k$.

Στη συνέχεια δείχνουμε ότι αν η (5) ισχύει για $N = m \geq 2$ τότε ισχύει και για $N = m-1$. Πράγματι, αν $a_1, \ldots, a_{m-1} \in I$, ορίζουμε

$$a_m = \frac{1}{m-1}(a_1+a_2+\ldots+a_{m-1})$$

και παρατηρούμε ότι

$$a_m = \frac{1}{m}(a_1+a_2+\ldots+a_m).$$

Τότε, εφαρμόζοντας την (5) για τους $a_1, \ldots, a_m$ έχουμε

$$f(a_m) = f\left(\frac{a_1+\ldots+a_m}{m}\right) \leq \frac{1}{m}\sum_{i=1}^{m-1} f(a_i) + \frac{1}{m}f(a_m),$$

απ' όπου παίρνουμε

$$f(a_m) \leq \frac{1}{m-1}\sum_{i=1}^{m-1} f(a_i).$$

Δηλαδή, η (5) ισχύει για $N = m-1$. Συνδυάζοντας τους δύο ισχυρισμούς βλέπουμε ότι η (5) ισχύει για κάθε $N \in \mathbb{N}$.

Έστω τώρα $a, b \in I$ και $k, n \in \mathbb{N}$ με $k < n$. Από την (5) βλέπουμε ότι

$$f\left(\frac{k}{n}a + \frac{n-k}{n}b\right) \leq \frac{1}{n}(kf(a) + (n-k)f(b)) = \frac{k}{n}f(a) + \frac{n-k}{n}f(b),$$

δηλαδή

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b) \tag{6}$$

αν $\lambda \in \mathbb{Q}$, $0 < \lambda < 1$. Χρησιμοποιώντας τη συνέχεια της f συμπεραίνουμε ότι η (6) εξακολουθεί να ισχύει αν $\lambda \in \mathbb{R}$, $0 < \lambda < 1$. Άρα, η f είναι κυρτή. $\square$

Θεώρημα 1.9 *Μία συνεχής συνάρτηση $f : I \to \mathbb{R}$ είναι κυρτή αν και μόνο αν για κάθε υποδιάστημα $[w, z]$ του I υπάρχει τουλάχιστον ένα $\theta \in (0,1)$ για το οποίο*

$$f(\theta w + (1-\theta)z) \leq \theta f(w) + (1-\theta)f(z).$$

Απόδειξη: Αν η f είναι κυρτή στο I τότε για κάθε υποδιάστημα $[w, z]$ του I και για κάθε $\theta \in (0,1)$, έχουμε

$$f(\theta w + (1-\theta)z) \leq \theta f(w) + (1-\theta)f(z).$$

Αντίστροφα τώρα υποθέτουμε ότι η f δεν είναι κυρτή. Τότε, υπάρχουν $\theta_0 \in (0,1)$ και $a_1, b_1 \in I, a_1 < b_1$, τέτοια ώστε

$$f(\theta_0 a_1 + (1-\theta_0)b_1) > \theta_0 f(a_1) + (1-\theta_0)f(b_1).$$

Θέτουμε $x_0 = \theta_0 a_1 + (1-\theta_0)b_1$. Η f είναι συνεχής στο I, άρα υπάρχει $\delta > 0$ τέτοιο ώστε για κάθε $\theta \in (\theta_0 - \delta, \theta_0 + \delta)$ να ισχύει

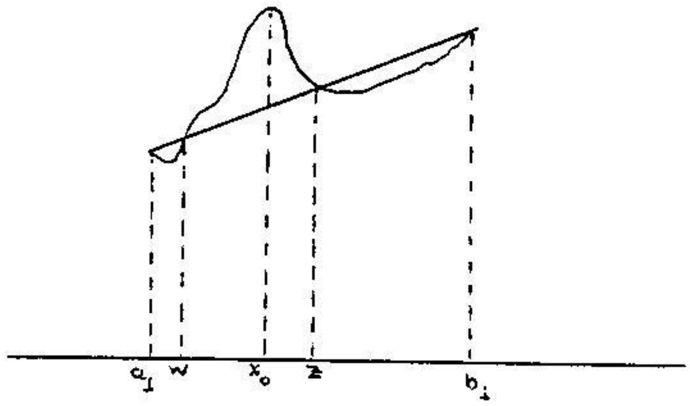
$$f(\theta a_1 + (1-\theta)b_1)) > \theta f(a_1) + (1-\theta)f(b_1). \tag{$*$}$$

Ορίζουμε

$$w = \inf\{x = \theta a_1 + (1-\theta)b_1 : f(x) > \theta f(a_1) + (1-\theta)f(b_1)\},$$

και

$$z = \sup\{x = \theta a_1 + (1-\theta)b_1 : f(x) > \theta f(a_1) + (1-\theta)f(b_1)\}.$$



Σχήμα 5

Τότε, $a_1 \leq w < z \leq b_1$ από την (*), αν $w = \theta' a_1 + (1-\theta')b_1$ και $z = \theta'' a_1 + (1-\theta'')b_1$ έχουμε

$$f(w) = \theta' f(a_1) + (1-\theta')f(b_1), \quad f(z) = \theta'' f(a_1) + (1-\theta'')f(b_1).$$

και για κάθε $y = \theta a_1 + (1-\theta)b_1 \in [w, z]$ ισχύει

$$f(y) > \theta f(a_1) + (1-\theta)f(b_1).$$

Από την υπόθεση όμως, υπάρχει $y = \lambda w + (1-\lambda)z = (\lambda\theta' + (1-\lambda)\theta'')a_1 + (\lambda(1-\theta') + (1-\lambda)(1-\theta''))b_1 \in [w, z]$ τέτοιο ώστε

$$f(y) \leq \lambda f(w) + (1-\lambda)f(z) = (\lambda\theta' + (1-\lambda)\theta'')f(a_1) + (\lambda(1-\theta') + (1-\lambda)(1-\theta''))f(b_1),$$

το οποίο είναι άτοπο. □

1.3 Παραγωγίσιμες κυρτές συναρτήσεις

§1. Στην περίπτωση που η f είναι δύο φορές παραγωγίσιμη στο ανοικτό διάστημα I, έχουμε έναν απλό χαρακτηρισμό της κυρτότητας μέσω της δεύτερης παραγώγου.

Θεώρημα 1.10 *Έστω I ανοικτό διάστημα και $f : I \to \mathbb{R}$ δύο φορές παραγωγίσιμη συνάρτηση. Τότε, η f είναι κυρτή αν και μόνο αν $f''(x) \geq 0$ για κάθε $x \in I$.*

Απόδειξη: Υποθέτουμε πρώτα ότι η f είναι κυρτή. Από το Θεώρημα 1.2 η f' είναι αύξουσα, επομένως $f''(x) \geq 0$ για κάθε $x \in I$.

Αντίστροφα, έστω ότι $f'' \geq 0$ στο I. Αν $x < y \in I$ και $0 < \lambda < 1$, χρησιμοποιώντας το θεώρημα μέσης τιμής του απειροστικού λογισμού βρίσκουμε ξ_1, ξ_2 με

$x < \xi_1 < \lambda x + (1-\lambda)y < \xi_2 < y$ και ξ_3 με $\xi_1 < \xi_2 < \xi_3$, τέτοια ώστε

$$\begin{aligned}
&f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda)f(y)\\
&= \lambda[f(\lambda x + (1-\lambda)y) - f(x)] + (1-\lambda)[f(\lambda x + (1-\lambda)y) - f(y)]\\
&= \lambda(1-\lambda)(y-x)f'(\xi_1) + (1-\lambda)\lambda(x-y)f'(\xi_2)\\
&= \lambda(1-\lambda)(y-x)(\xi_1-\xi_2)f''(\xi_3) \leq 0.
\end{aligned}$$

Αυτό αποδεικνύει ότι η f είναι κυρτή. □

Παρατήρηση Από την απόδειξη προκύπτει ότι η f είναι γνήσια κυρτή αν $f''(x) > 0$ για κάθε $x \in I$. Το αντίστροφο δεν ισχύει: η συνάρτηση $x \mapsto x^4$ είναι γνήσια κυρτή στο $\mathbb{R}$, αλλά $f''(0) = 0$.

§2. **Ανισότητες**

Το Θεώρημα 1.10 μάς βοηθάει να δώσουμε πολλά απλά παραδείγματα κυρτών συναρτήσεων, και μέσω αυτών των συναρτήσεων μπορούμε να αποδείξουμε ανισότητες που πολλές φορές δεν φαίνονται και τόσο απλές με την πρώτη ματιά. Για παράδειγμα, η ανισότητα

$$x^{\lambda}y^{\mu} \leq \lambda x + \mu y \tag{7}$$

ισχύει για κάθε $x, y > 0$ και $\lambda, \mu > 0$ με $\lambda + \mu = 1$. Για να αποδείξουμε αυτήν την ανισότητα, μπορούμε να χρησιμοποιήσουμε το γεγονός ότι η συνάρτηση $x \mapsto e^x$ είναι (γνήσια) κυρτή, στη μορφή

$$\exp(\lambda \log x + \mu \log y) \leq \lambda \exp(\log x) + \mu \exp(\log y).$$

Η (7) εμφανίζεται πολύ συχνά στη μορφή

$$x^{1/p}y^{1/q} \leq \frac{1}{p}x + \frac{1}{q}y \tag{8}$$

ή στη μορφή

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \tag{9}$$

όπου $x, y > 0$, $p, q > 1$ και $1/p + 1/q = 1$. Στην ειδική περίπτωση $p = q = 2$, η (8) είναι η πολύ γνωστή ανισότητα $2\sqrt{xy} \leq x + y$.

1.4 Θεωρήματα για ολοκληρώματα

§1. Κάθε κυρτή συνάρτηση $f : (a, b) \to \mathbb{R}$ έχει δεξιά παράγωγο σε όλα τα σημεία του (a, b). Το θεώρημα που ακολουθεί χαρακτηρίζει την f σαν αόριστο ολοκλήρωμα της δεξιάς παραγώγου της.

Θεώρημα 1.11 *Μία συνάρτηση $f:(a,b)\to\mathbb{R}$ είναι κυρτή αν και μόνο αν αναπαρίσταται στη μορφή*

$$f(x)=f(c)+\int_c^x g(t)dt, \qquad c,x\in(a,b) \tag{10}$$

όπου $g:(a,b)\to\mathbb{R}$ είναι μία αύξουσα συνεχής από δεξιά συνάρτηση.

Απόδειξη: Υποθέτουμε πρώτα ότι η f είναι κυρτή και θεωρούμε $c,x\in(a,b)$. Εξετάζουμε την περίπτωση $c<x$ (όμοια εργαζόμαστε για την περίπτωση $x<c$). Από τα Θεωρήματα 1.2 και 1.5, οι f'_+ και f'_- υπάρχουν και είναι αύξουσες και η f'_+ είναι συνεχής από δεξιά ενώ η f'_- συνεχής από αριστερά. Λόγω μονοτονίας είναι και οι δύο ολοκληρώσιμες κατά Riemann.

Θεωρούμε τυχούσα διαμέριση $P=\{c=y_0<y_1<\ldots<y_m=x\}$ του $[c,x]$. Παρατηρούμε ότι

$$f'_-(y_{k-1})\leq f'_+(y_{k-1})\leq\frac{f(y_k)-f(y_{k-1})}{y_k-y_{k-1}}\leq f'_-(y_k)\leq f'_+(y_k)$$

για κάθε $k=1,\ldots,m$.

Γράφουμε Ξ_1 για την επιλογή σημείων $\{y_0,\ldots,y_{m-1}\}$ και Ξ_2 για την επιλογή σημείων $\{y_1,\ldots,y_m\}$. Τότε, τα αντίστοιχα αθροίσματα Riemann ικανοποιούν τις ανισότητες

$$\begin{aligned}\Sigma(f'_-,P,\Xi_1) &\leq \Sigma(f'_+,P,\Xi_1)\\ &\leq \sum_{k=1}^m[f(y_k)-f(y_{k-1})]=f(x)-f(c)\\ &\leq \Sigma(f'_-,P,\Xi_2)\leq\Sigma(f'_+,P,\Xi_2).\end{aligned}$$

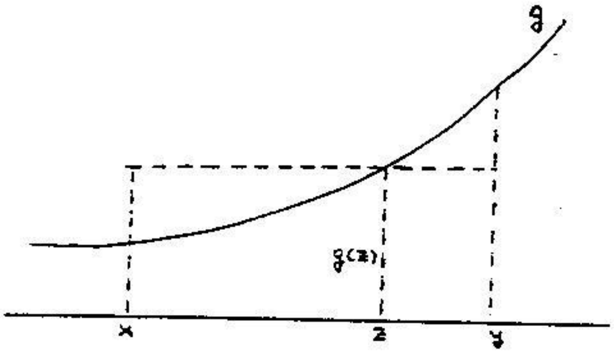
Παίρνοντας το πλάτος της P να τείνει στο 0, βλέπουμε ότι

$$\int_c^x f'_-(t)dt\leq\int_c^x f'_+(t)dt\leq f(x)-f(c)\leq\int_c^x f'_-(t)dt\leq\int_c^x f'_+(t)dt.$$

Επομένως,

$$f(x)=f(c)+\int_c^x f'_+(t)dt=f(c)+\int_c^x f'_-(t)dt.$$

Για την αντίστροφη κατεύθυνση, ας υποθέσουμε ότι η (10) ισχύει για κάποια αύξουσα (συνεχή από δεξιά) συνάρτηση g.



Σχήμα 6

Έστω $x, y \in (a, b)$ με $x < y$ και $\lambda \in (0,1)$. Θέτουμε $z = \lambda x + (1-\lambda)y$. Τότε (δείτε το Σχήμα 6),

$$\begin{aligned} f(z) - \lambda f(x) - (1-\lambda)f(y) &= \lambda[f(z) - f(x)] - (1-\lambda)[f(y) - f(z)] \\ &= \lambda \int_x^z g(t)dt - (1-\lambda)\int_z^y g(t)dt \\ &\leq \lambda(z-x)g(z) - (1-\lambda)(y-z)g(z). \end{aligned}$$

Η τελευταία ποσότητα ισούται με 0, γιατί $z - x = (\lambda - 1)(x-y)$ και $y - z = \lambda(y - x)$. □

§2. Έστω $f : [a, b] \to \mathbb{R}$ κυρτή συνάρτηση και $a_1, \ldots, a_n \in [a, b]$. Όπως έχουμε ήδη δεί,

$$f\left(\frac{a_1 + \ldots + a_n}{n}\right) \leq \frac{1}{n}\sum_{i=1}^{n} f(a_i). \tag{12}$$

Η ανισότητα αυτή μάς λέει ότι η τιμή της f στον μέσο όρο n αριθμών είναι μικρότερη ή ίση από τον μέσο όρο των αντίστοιχων τιμών. Θα δείξουμε τώρα το ολοκληρωτικό ανάλογό της.

Θεώρημα 1.12 (ανισότητα του Jensen) *Έστω $f : (a, b) \to \mathbb{R}$ κυρτή συνάρτηση και $g : [c, d] \to (a, b)$ συνεχής συνάρτηση. Τότε,*

$$f\left(\frac{1}{d-c}\int_c^d g(x)dx\right) \leq \frac{1}{d-c}\int_c^d f(g(x))dx.$$

Απόδειξη: Θέτουμε

$$p := \frac{1}{d-c}\int_c^d g(x)dx.$$

Τότε, $p \in (a, b)$. Από το Θεώρημα 1.6 έχουμε

$$f(y) \geq f(p) + f'_+(p)(y - p)$$

για κάθε $y \in (a, b)$, επομένως

$$f(g(x)) \geq f(p) + f'_+(p)(g(x) - p)$$

για κάθε $x \in [c, d]$. Ολοκληρώνοντας την τελευταία ανισότητα στο $[c, d]$ παίρνουμε το ζητούμενο. □

1.5 Η συζυγής συνάρτηση

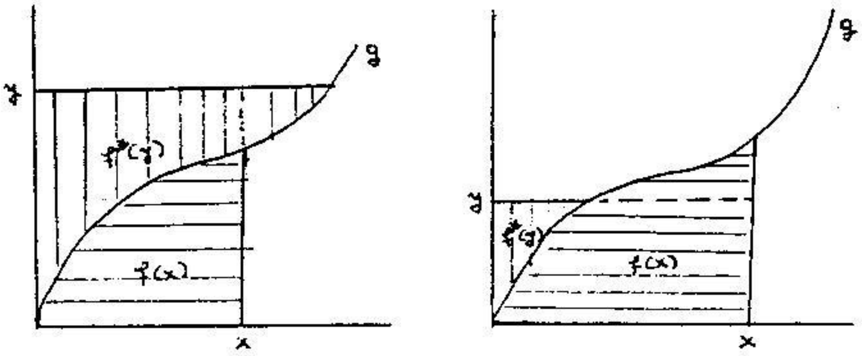
Η έννοια του δυϊσμού θα εμφανιστεί πολλές φορές σε αυτό το μάθημα. Πολύ συχνά, όταν αντιμετωπίζουμε κάποιο μαθηματικό πρόβλημα καταφεύγουμε σε κάποιο άλλο που, αν και φαίνεται τελείως διαφορετικό, αντικατοπτρίζει όλες τις πτυχές του αρχικού και αναλύεται ευκολότερα. Για παράδειγμα, μερικές φορές είναι πιό εύκολο να δείξουμε ότι ένα σύνολο είναι κλειστό αποδεικνύοντας ότι το συμπλήρωμά του είναι ανοικτό: οι δύο αυτές έννοιες είναι δυϊκές. Στη στοιχειώδη θεωρία συνόλων, σε κάθε θεώρημα που διατυπώνεται μέσω ενώσεων, τομών και συμπληρωμάτων αντιστοιχεί ένα δυϊκό θεώρημα που παίρνουμε αν εναλλάξουμε τις ενώσεις με τομές και το κενό σύνολο με το σύνολο ως προς το οποίο παίρνουμε συμπληρώματα. Δεδομένου ότι οι αποδείξεις δύο δυϊκών θεωρημάτων δεν παρουσιάζουν πάντα την ίδια δυσκολία, υπάρχουν δύο λόγοι για να χρησιμοποιήσουμε τον δυϊσμό: μπορεί να αποδείξουμε ευκολότερα αυτό που θέλουμε ή να κερδίσουμε χωρίς κόπο ένα λιγότερο προφανές θεώρημα.

Σε αυτήν την παράγραφο θα αντιστοιχίσουμε σε κάθε κυρτή συνάρτηση μία άλλη, την *συζυγή* της κυρτή συνάρτηση.

Έστω $g : [0, +\infty) \to [0, +\infty)$ μία γνήσια αύξουσα και συνεχής συνάρτηση με $g(0) = 0$ και $\lim_{x\to\infty} g(x) = +\infty$. Τότε, η g^{-1} υπάρχει και έχει τις ίδιες ιδιότητες με την f. Αν ορίσουμε

$$f(x) = \int_0^x g(s)ds, \qquad f^*(y) = \int_0^y g^{-1}(s)ds,$$

τότε οι f και f^* είναι κυρτές συναρτήσεις στο $[0, +\infty)$.



Σχήμα 7

Από τον ορισμό των f και f^* ή και από το ίδιο το Σχήμα 7, μπορούμε να πάρουμε αρκετά συμπεράσματα. Το πρώτο από αυτά είναι η *ανισότητα του Young*:

(α) $xy \leq f(x) + f^*(y)$ για κάθε $x, y \geq 0$.
(β) $xy = f(x) + f^*(y)$ αν και μόνο αν $y = g(x) = f'(x)$.
(γ) $(f^*)' = (f')^{-1}$.
(δ) $f^{**} = f$.
(ε) $f^*(y) = \sup_{x \geq 0}[xy - f(x)]$.

Λέμε ότι η f^* είναι η *συζυγής συνάρτηση* της f. Έχουμε λοιπόν ορίσει την έννοια της συζυγούς συνάρτησης για μία κλάση κυρτών συναρτήσεων, εκείνων που αναπαρίστανται σαν ολοκληρώματα κάποιων ειδικών αυξουσών συναρτήσεων. Σκοπός μας είναι να ορίσουμε την συζυγή συνάρτηση για τυχούσα κυρτή συνάρτηση, με τέτοιον τρόπο ώστε να διατηρούνται κατά κάποιον τρόπο όλες οι ιδιότητες (α)-(ε). Ορίζουμε λοιπόν την f^* μέσω της (ε).

Ορισμός Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Η *συζυγής συνάρτηση* f^* της f ορίζεται από την

$$f^*(y) = \sup_{x \in I}[xy - f(x)]$$

και έχει πεδίο ορισμού το $I^* = \{y \in \mathbb{R} : f^*(y) < \infty\}$.

Όπως θα δούμε, η f^* είναι κυρτή συνάρτηση. Έχει όμως άλλη μία ιδιότητα που θα παίξει σημαντικό ρόλο στη συνέχεια: είναι *κλειστή*. Λέμε ότι μία συνάρτηση $g : J \to \mathbb{R}$ είναι *κλειστή* αν το *σύνολο στάθμης* $L_a = \{x \in J : g(x) \leq a\}$ είναι κλειστό υποσύνολο του $\mathbb{R}$ για κάθε $a \in \mathbb{R}$ [ένα $B \subseteq \mathbb{R}$ λέγεται κλειστό αν ικανοποιεί το εξής: αν $b_m \in B$ και $b_m \to b \in \mathbb{R}$, τότε $b \in B$.]

Θεώρημα 1.13 *Αν η $f : I \to \mathbb{R}$ είναι κυρτή, τότε η συζυγής της συνάρτηση $f^* : I^* \to \mathbb{R}$ είναι κυρτή και κλειστή.*

Απόδειξη: Παρατηρούμε πρώτα ότι $I^* \neq \emptyset$: παίρνουμε $x_0 \in \mathrm{int}(I)$ και $m \in [f'_-(x_0), f'_+(x_0)]$. Τότε, $f(x) \geq f(x_0) + m(x - x_0)$ για κάθε $x \in I$. Άρα,

$$xm - f(x) \leq x_0 m - f(x_0)$$

για κάθε $x \in I$, το οποίο δείχνει ότι

$$f^*(m) = \sup_{x \in I}[xm - f(x)] \leq x_0 m - f(x_0) < +\infty,$$

δηλαδή, $m \in I^*$. Δεν είναι δύσκολο να ελέγξουμε ότι το I^* είναι διάστημα: αν $y < z \in I^*$ και $w \in (y, z)$, τότε

$$xw - f(x) \leq \max\{xy - f(x), xz - f(x)\}$$

για κάθε $x \in I$, άρα

$$\begin{aligned} f^*(w) &= \sup_{x \in I}[xw - f(x)] \leq \max\{\sup_{x \in I}[xy - f(x)], \sup_{x \in I}[xz - f(x)]\} \\ &\leq \max\{f^*(y), f^*(w)\} < +\infty. \end{aligned}$$

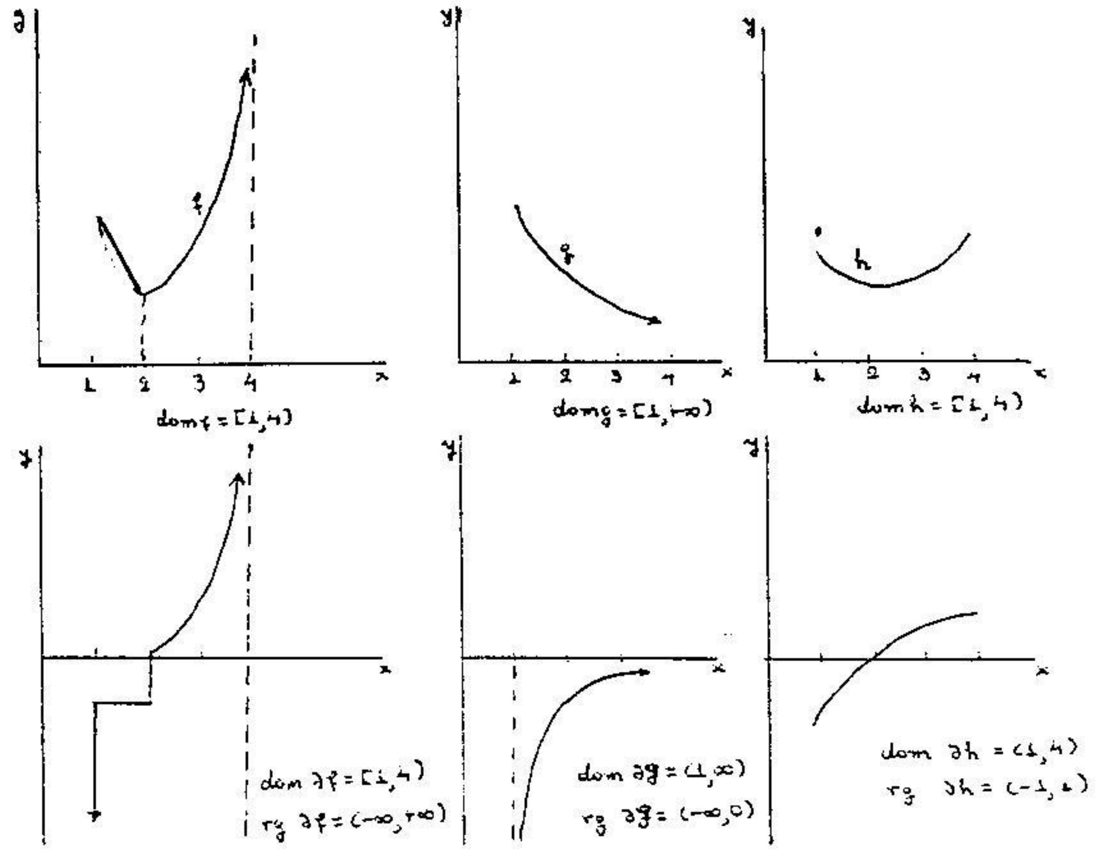
Το γεγονός ότι η $f^* : I^* \to \mathbb{R}$ είναι supremum γραμμικών συναρτήσεων αποδεικνύει ότι η f^* είναι κυρτή. Τέλος, δείχνουμε ότι η f^* είναι κλειστή: έστω $a \in \mathbb{R}$ και $y_m \in L_a$ με $y_m \to y \in \mathbb{R}$. Για κάθε $x \in I$ έχουμε

$$xy_m - f(x) \leq f^*(y_m) \leq a \Longrightarrow xy - f(x) = \lim_{m \to \infty}[xy_m - f(x)] \leq a.$$

Άρα, $f^*(y) = \sup_{x \in I}[xy - f(x)] \leq a$ δηλαδή $y \in L_a$. □

Αφού η f^{**} είναι πάντα κλειστή συνάρτηση, είναι τώρα φανερό ότι ο αρχικός μας στόχος να ορίσουμε την f^* έτσι ώστε να ικανοποιούνται οι (α)-(ε) δεν είναι γενικά εφικτός. Αν η f δεν είναι κλειστή, τότε η $f^{**} = f$ αποκλείεται να ισχύει. Θα πρέπει λοιπόν να περιοριστούμε στις κλειστές κυρτές συναρτήσεις $f : I \to \mathbb{R}$.

Ένα πρώτο ερώτημα είναι πόσο περιορισμένη είναι αυτή η κλάση των κλειστών κυρτών συναρτήσεων. Στο Σχήμα που ακολουθεί βλέπετε τρείς κυρτές συναρτήσεις.



Σχήμα 8

Οι πρώτες δύο είναι κλειστές, ενώ η τρίτη όχι. Με λίγη σκέψη βλέπουμε ότι αυτό που μπορεί να «χαλάσει» την ιδιότητα της κλειστότητας είναι η συμπεριφορά της f στα άκρα του I. Η f είναι κλειστή αν και μόνο αν είναι συνεχής στα άκρα του I που ανήκουν στο I και έχει όριο το $-\infty$ καθώς πλησιάζουμε κάποιο πεπερασμένο άκρο του I που δεν ανήκει στο I.

Ορισμός Έστω I ένα διάστημα και $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Για κάθε $x \in I$, ορίζουμε το *υποδιαφορικό* $\partial f(x)$ στο x να είναι το σύνολο όλων των $m \in \mathbb{R}$ για τα οποία η γραμμική συνάρτηση $y \to f(x) + m(y - x)$ στηρίζει την f στο x. Αν $x \in \text{int}(I)$, τότε το $\partial f(x)$ είναι μη κενό: $\partial f(x) = [f'_-(x), f'_+(x)]$. Το $\partial f(x)$ είναι μονοσύνολο αν και μόνο αν η f έχει παράγωγο στο x.

Άρα, η ∂f είναι μία πλειότιμη συνάρτηση με πεδίο ορισμού $\text{dom}(\partial f)$ το σύνολο των $x \in I$ στα οποία η f έχει στήριγμα. Παρατηρήστε ότι $\text{int}(I) \subseteq \text{dom}(\partial f) \subseteq I$. Το πεδίο τιμών $\text{rg}(\partial f)$ είναι το σύνολο των κλίσεων των συναρτήσεων στήριξης. Αν η f είναι κλειστή, δεν είναι δύσκολο να ελέγξετε ότι αν το I έχει πεπερασμένο δεξιό άκρο τότε το $\mathrm{rg}(\partial f)$ εκτείνεται ως το $+\infty$, ενώ αν το I έχει πεπερασμένο αριστερό άκρο τότε το $\mathrm{rg}(\partial f)$ εκτείνεται ως το $-\infty$.

Στο δεύτερο μισό του Σχήματος 8 σχεδιάσαμε το υποδιαφορικό καθεμιάς από τις τρείς κυρτές συναρτήσεις που βλέπετε στο πρώτο μισό. Σε κάθε περίπτωση, το γράφημα παριστάνει μία συνεχή αύξουσα «συνάρτηση» και ενδέχεται να περιέχει κατακόρυφα ή οριζόντια ευθύγραμμα τμήματα.

Ορισμός Ένα υποσύνολο Γ του $\mathbb{R}\times\mathbb{R}$ λέγεται *αύξον* αν $(x_2-x_1)(y_2-y_1)\geq 0$ οποτεδήποτε $(x_1,y_1),(x_2,y_2)\in\Gamma$. Το Γ είναι *μεγιστικό αύξον σύνολο* αν είναι αύξον και δεν περιέχεται γνήσια σε άλλο αύξον σύνολο.

Παρατηρήστε ότι αν το Γ είναι αύξον, τότε και το σύνολο $\Gamma^{-1}:=\{(y,x):(x,y)\in\Gamma\}$ είναι αύξον. Επίσης, αν το Γ είναι μεγιστικό, το ίδιο ισχύει και για το Γ^{-1}.

Θεώρημα 1.14 *Έστω I ένα διάστημα και $f:I\to\mathbb{R}$ κλειστή κυρτή συνάρτηση. Τότε, το γράφημα $\Gamma_f:=\{(x,y):x\in I, y\in\partial f(x)\}|$ του ∂f είναι μεγιστικό αύξον σύνολο.*

Απόδειξη: Δείχνουμε πρώτα ότι το Γ_f είναι αύξον. Αν $x_1<x_2$ στο I και $(x_1,y_1),(x_2,y_2)\in\Gamma_f$, τότε

$$y_2\geq f'_-(x_2)\geq f'_+(x_1)\geq y_1$$

άρα $(x_2-x_1)(y_2-y_1)\geq 0$. Τελείως ανάλογα εξετάζουμε την περίπτωση $x_1<x_2$, ενώ η περίπτωση $x_1=x_2$ είναι προφανής.

Για να δείξουμε ότι το Γ_f είναι μεγιστικό, πρέπει να δείξουμε ότι αν $(x_1,y_1)\notin\Gamma_f$ τότε υπάρχει $(x,y)\in\Gamma_f$ τέτοιο ώστε $(x-x_1)(y-y_1)<0$. Αντικαθιστώντας την f με την κλειστή κυρτή συνάρτηση $h(x)=f(x+x_1)-xy_1$ βλέπουμε ότι μπορούμε να κάνουμε την επιπλέον υπόθεση $x_1=y_1=0$. Ζητάμε λοιπόν να δείξουμε ότι: αν $(0,0)\notin\Gamma_f$ τότε υπάρχει $(x,y)\in\Gamma_f$ τέτοιο ώστε $xy<0$.

Διακρίνουμε δύο περιπτώσεις:

(α) αν $0\notin I$ τότε είτε $x<0$ για κάθε $x\in I$ ή $x>0$ για κάθε $x\in I$. Αν $x<0$ για κάθε $x\in I$ και $b\leq 0$ το δεξιό άκρο του I, τότε το σύνολο τιμών του ∂f εκτείνεται ως το $+\infty$ γιατί η f είναι κλειστή, άρα υπάρχει $(x,y)\in\Gamma_f$ με $y>0$. Έπεται ότι $xy<0$. Όμοια δουλεύουμε αν $x>0$ για κάθε $x\in I$.

(β) Αν $0\in I$, τότε από την υπόθεση ότι $(0,0)\notin\Gamma_f$ έπεται ότι είτε $f'_-(0)>0$ ή $f'_+(0)<0$. Σε κάθε περίπτωση, η f δεν έχει ελάχιστο στο 0, άρα υπάρχει $x_2\in I$ τέτοιο ώστε $f(x_2)<f(0)$.

Αν $x_2<0$, από την

$$0<f(0)-f(x_2)=\int_{x_2}^{0}f'_+(t)dt$$

βλέπουμε ότι υπάρχει $x_3\in(x_2,0)$ τέτοιο ώστε $f'_+(x_3)>0$. Τότε, $(x_3,f'_+(x_3))\in\Gamma_f$ και $x_3\cdot f'_+(x_3)<0$.

Αν $x_2>0$, δουλεύουμε ανάλογα χρησιμοποιώντας την

$$0>f(x_2)-f(0)=\int_{0}^{x_2}f'_+(t)dt.$$

□

Μπορούμε τώρα να αποδείξουμε το κεντρικό θεώρημα αυτής της παραγράφου.

Θεώρημα 1.15 *Έστω I ένα διάστημα και $f : I \to \mathbb{R}$ κλειστή κυρτή συνάρτηση. Τότε, η $f^* : I^* \to \mathbb{R}$ είναι επίσης κυρτή και κλειστή, και*

(α) $xy \leq f(x) + f^*(y)$ για κάθε $x \in I$ και $y \in I^*$.

(β) $xy = f(x) + f^*(y)$ αν και μόνο αν $y \in \partial f(x)$.

(γ) $\Gamma_{f^*} = (\Gamma_f)^{-1}$.

(δ) $f^{**} = f$.

Απόδειξη: Από το Θεώρημα 1.13 ξέρουμε ότι η f^* είναι κυρτή και κλειστή. Το (α) είναι άμεση συνέπεια του ορισμού της f^*.

Για να αποδείξουμε το (β), παρατηρούμε ότι μία κυρτή συνάρτηση $g : I \to \mathbb{R}$ έχει ελάχιστο στο $x \in I$ αν και μόνο αν $0 \in \partial g(x)$. Τώρα,

$$-f^*(y) = -\sup_{x\in I}[xy - f(x)] = \inf_{x\in I}[f(x) - xy].$$

Όμως η $g(x) = f(x) - xy$ είναι κυρτή, άρα παίρνει ελάχιστη τιμή στο x αν και μόνο αν $0 \in \partial g(x)$, δηλαδή αν και μόνο αν $y \in \partial f(x)$ (γιατί;). Άρα, $-f^*(y) = f(x) - xy$ αν και μόνο αν $y \in \partial f(x)$, το οποίο είναι το (β).

Για το (γ) δείχνουμε πρώτα ότι

$$(\Gamma_f)^{-1} \subseteq \Gamma_{f^*}.$$

Αν $(y, x) \in (\Gamma_f)^{-1}$, τότε $(x, y) \in \Gamma_f$. Άρα, $y \in \partial f(x)$, δηλαδή $f(z) \geq f(x) + y(z - x)$ για κάθε $z \in I$. Τότε,

$$\forall z \in I \quad zy - f(z) \leq xy - f(x) \Longrightarrow f^*(y) \leq xy - f(x)$$

άρα

$$f^*(y) - xy \leq -f(x) \leq f^*(w) - xw \ \ \forall w \in I^* \Longrightarrow x \in \partial f^*(y)$$

από το (β). Έπεται ότι $(y, x) \in \Gamma_{f^*}$. Όμως το Γ_f είναι μεγιστικό αύξον σύνολο από το Θεώρημα 1.14, άρα και το $(\Gamma_f)^{-1}$ είναι μεγιστικό αύξον σύνολο. Αφού το Γ_{f^*} είναι αύξον, συμπεραίνουμε ότι

$$(\Gamma_f)^{-1} = \Gamma_{f^*}.$$

Για το (δ), παρατηρούμε πρώτα ότι

$$\Gamma_{f^{**}} = (\Gamma_{f^*})^{-1} = ((\Gamma_f)^{-1})^{-1} = \Gamma_f.$$

Αυτό αποδεικνύει ότι οι f και f^{**} έχουν το ίδιο πεδίο ορισμού I. Αφού οι f, f^{**} είναι κλειστές, για να δείξουμε ότι $f = f^{**}$, αρκεί να δείξουμε ότι $f(c) = f^{**}(c)$ για κάθε $c \in \mathrm{int}(I)$. Επιλέγουμε $c \in \mathrm{int}(I)$ και $y \in I^*$ με $(c, y) \in \Gamma_f = (\Gamma_{f^*})^{-1}$. Τότε, $(c, y) \in \Gamma_f$ άρα

$$cy = f(c) + f^*(y),$$

και $(y, c) \in \Gamma_{f^*}$ άρα

$$cy = f^*(y) + f^{**}(c),$$

το οποίο δείχνει την $f(c) = f^{**}(c)$. □

1.6 Ασκήσεις

Ομάδα Α

1. Δείξτε ότι τα (α) και (β) της 1.1(§2) είναι ισοδύναμα με τον ορισμό της κυρτότητας της $f : I \to \mathbb{R}$.

2. Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, αν $x_1, \ldots, x_m \in I$ και $\lambda_1, \ldots, \lambda_m \geq 0$ με $\lambda_1 + \ldots + \lambda_m = 1$, έχουμε $\sum_{i=1}^m \lambda_i x_i \in I$ και

$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i).$$

3. Δείξτε ότι το κατά σημείο όριο μιάς συγκλίνουσας ακολουθίας κυρτών συναρτήσεων είναι κυρτή συνάρτηση.

4. Έστω $\{f_j : j \in J\}$ τυχούσα οικογένεια κυρτών συναρτήσεων $f_j : I \to \mathbb{R}$. Ορίζουμε $f : I \to \mathbb{R}$ με $f(x) = \sup\{f_j(x) : j \in J\}$. Αν η f είναι πεπερασμένη παντού στο I, τότε η f είναι κυρτή.

5. Έστω $f, g : \mathbb{R} \to \mathbb{R}$ κυρτές συναρτήσεις. Υποθέτουμε ακόμα ότι η g είναι αύξουσα. Δείξτε ότι η $g \circ f$ είναι κυρτή.

6. Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Δείξτε ότι

$$f(x_1 + \delta) - f(x_1) \leq f(x_2 + \delta) - f(x_2)$$

για κάθε $x_1 < x_2 \in I$ και $\delta > 0$ για το οποίο $x_1 + \delta, x_2 + \delta \in I$.

7. Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Δείξτε με ένα παράδειγμα ότι η f δεν είναι αναγκαστικά συνάρτηση Lipschitz σε ολόκληρο το I, ακόμα και αν υποθέσουμε ότι η f είναι φραγμένη. Επίσης, δείξτε ότι η f δεν είναι αναγκαστικά συνεχής στο I, ακόμα και αν υποθέσουμε ότι το I είναι κλειστό και φραγμένο διάστημα.

8. Έστω $f : I \to \mathbb{R}$ κυρτή συνάρτηση. Δείξτε ότι:

(α) αν η f έχει ολικό μέγιστο σε κάποιο εσωτερικό σημείο του I, τότε η f είναι σταθερή.

(β) αν η f έχει ολικό ελάχιστο στο $c \in I$, τότε η c είναι φθίνουσα αριστερά από το c και αύξουσα δεξιά από το c.

(γ) αν η f έχει τοπικό ελάχιστο, τότε αυτό το ελάχιστο είναι ολικό ελάχιστο.

(δ) αν η f είναι γνήσια κυρτή, τότε έχει το πολύ ένα σημείο ολικού ελαχίστου.

9. Έστω $f : \mathbb{R} \to \mathbb{R}$ κυρτή συνάρτηση. Αν η f είναι άνω φραγμένη, τότε είναι σταθερή.

10. Δείξτε ότι κάθε κυρτή συνάρτηση ορισμένη σε φραγμένο διάστημα είναι κάτω φραγμένη.

11. Έστω $f : (a, b) \to \mathbb{R}$ κυρτή συνάρτηση και $c \in (a, b)$. Δείξτε ότι η f είναι παραγωγίσιμη στο c αν και μόνο αν

$$\lim_{h \downarrow 0} \frac{f(c+h) + f(c-h) - 2f(c)}{h} = 0.$$

12. (α) Έστω $x_1, \dots, x_n$ και $r_1, \dots, r_n$ θετικοί πραγματικοί αριθμοί με $r_1 + \dots + r_n = 1$. Δείξτε ότι

$$\prod_{i=1}^{n} x_i^{r_i} \leq \sum_{i=1}^{n} r_i x_i.$$

(β) Δείξτε ότι ο γεωμετρικός μέσος n θετικών πραγματικών αριθμών είναι μικρότερος ή ίσος από τον αριθμητικό τους μέσο.
(γ) Έστω $p, q > 1$ με $1/p + 1/q = 1$ και $a_1, \dots, a_n$, $b_1, \dots, b_n$ θετικοί πραγματικοί αριθμοί. Δείξτε την *ανισότητα του Hölder*:

$$\sum_{i=1}^{n} a_i b_i \leq \left(\sum_{i=1}^{n} a_i^p \right)^{1/p} \left(\sum_{i=1}^{n} b_i^q \right)^{1/q}.$$

Όταν $p = q = 2$, παίρνουμε την ανισότητα Cauchy-Schwarz.

13. Έστω $h : \mathbb{R} \to \mathbb{R}$ συνάρτηση με θετικές τιμές, τέτοια ώστε η $\log h$ να είναι κυρτή. Δείξτε ότι η h είναι κυρτή. Δώστε παράδειγμα θετικής κυρτής συνάρτησης $g : \mathbb{R} \to \mathbb{R}$ της οποίας ο λογάριθμος δεν είναι κυρτή συνάρτηση.

14. Μία θετική συνάρτηση f λέγεται *λογαριθμικά κυρτή* αν η $\log f$ είναι κυρτή.

(α) Έστω f δύο φορές παραγωγίσιμη θετική συνάρτηση. Δείξτε ότι η f είναι λογαριθμικά κυρτή αν και μόνο αν $f'' f \geq (f')^2$.

(β) Αν οι f και g είναι λογαριθμικά κυρτές και $a, b > 0$, τότε η $af + bg$ είναι λογαριθμικά κυρτή.

(γ) Έστω $a_1, \dots, a_n > 0$. Δείξτε ότι η συνάρτηση

$$x \mapsto \log \left(a_1^x + \dots + a_n^x \right)$$

είναι κυρτή στο $\mathbb{R}$.

15. Αποδείξτε τις ιδιότητες (α)-(ε) των f και f^* που ορίστηκαν στην αρχή της Παραγράφου 1.5.

16. Έστω $h : I \to \mathbb{R}$ αύξουσα συνάρτηση.
(α) Δείξτε ότι για κάθε $c \in \mathrm{int}(I)$ υπάρχουν τα $\lim_{z \uparrow c} h(z)$ και $\lim_{z \downarrow c} h(z)$, και

$$\lim_{z \uparrow c} h(z) \leq \lim_{z \downarrow c} h(z).$$

(β) Δείξτε ότι υπάρχουν το πολύ άπειρα αριθμήσιμα σημεία του I στα οποία η h είναι ασυνεχής.

17. Βρείτε τα $\partial f(x)$, $\partial f^*(y)$, $f^*(y)$, I^* αν
(α) $f(x) = x^2$, $I = \mathbb{R}$.
(β) $f(x) = |x|$, $I = \mathbb{R}$.
(γ) $f(x) = e^x$, $I = \mathbb{R}$.
(δ) $f(x) = |x|^p / p$, $I = \mathbb{R}$, $p > 1$.

Ομάδα Β

18. Έστω $a_1, \dots, a_m > 0$ με $\sum_{i=1}^m a_i = 1$ και $x_1, \dots, x_m > 0$. Για κάθε $t \neq 0$ ορίζουμε

$$M_t(\vec{x}, \vec{a}) = \left(a_1 x_1^t + \dots + a_m x_m^t\right)^{1/t},$$

ενώ για $t = 0$ θέτουμε

$$M_0(\vec{x}, \vec{a}) = x_1^{a_1} \dots x_m^{a_m}.$$

Δείξτε ότι

$$\lim_{t \to 0} M_t(\vec{x}, \vec{a}) = M_0(\vec{x}, \vec{a})$$

και ότι η $M_t(\vec{x}, \vec{a})$ είναι αύξουσα συνάρτηση του t.

19. Με το συμβολισμό της Άσκησης 17, δείξτε ότι αν $x_k = \max\{x_1, \dots, x_m\}$, τότε

$$a_k^{1/t} x_k \leq M_t(\vec{x}, \vec{a}) \leq x_k$$

για κάθε $t > 0$. Ποιό είναι το $\lim_{t \to \infty} M_t(\vec{x}, \vec{a})$; Σχεδιάστε τη γραφική παράσταση της $M_t(\vec{x}, \vec{a})$. Ποιά είναι η συμπεριφορά της στο $-\infty$;

20. Δείξτε ότι αν η $f : (0, +\infty) \to \mathbb{R}$ είναι κυρτή και $x_1, \dots, x_m, y_1, \dots, y_m > 0$, τότε

$$(x_1 + \dots + x_m) f\left(\frac{y_1 + \dots + y_m}{x_1 + \dots + x_m}\right) \leq \sum_{i=1}^m x_i f\left(\frac{y_i}{x_i}\right).$$

Δείξτε ότι η $f(x) = (1 + x^p)^{1/p}$ είναι κυρτή στο $(0, +\infty)$ όταν $p \geq 1$, και συμπεράνετε ότι

$$((x_1 + \dots + x_m)^p + (y_1 + \dots + y_m)^p)^{1/p} \leq \sum_{i=1}^m (x_i^p + y_i^p)^{1/p}.$$

21. Δείξτε ότι η συνάρτηση $-\sin x$ είναι κυρτή στο $[0, \pi]$. Χρησιμοποιώντας το δείξτε ότι η μέγιστη περίμετρος n-γώνου που εγγράφεται στο μοναδιαίο κύκλο είναι $2n \sin(\pi/n)$.

22. Έστω $f : [0, 2\pi] \to \mathbb{R}$ κυρτή συνάρτηση. Δείξτε ότι για κάθε $k \geq 1$,

$$\frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx \geq 0.$$

23. Δείξτε ότι υπάρχει κυρτή συνάρτηση $f : [0, 1] \to \mathbb{R}$ που δεν έχει παράγωγο σε κανένα σημείο ενός πυκνού υποσυνόλου του $[0, 1]$.
[*Υπόδειξη:* Θεωρήστε μία αρίθμηση $\{r_k : k \in \mathbb{N}\}$ των ρητών του $[0, 1]$ και ορίστε

$$g(t) = \sum_{\{k : r_k \leq t\}} \frac{1}{2^k}.$$

Η g είναι αύξουσα: με τη βοήθειά της ορίστε κατάλληλη f.]

24. Έστω $f : (a, b) \to \mathbb{R}$ συνεχής συνάρτηση. Δείξτε ότι η f είναι κυρτή αν και μόνο αν

$$f(x) \leq \frac{1}{2h} \int_{-h}^{h} f(x+t)dt$$

για κάθε διάστημα $[x - h, x + h] \subset (a, b)$.

25. Έστω $f : I \to \mathbb{R}$ θετική κοίλη συνάρτηση. Δείξτε ότι η $1/f$ είναι κυρτή.

26. Έστω $f : [0, +\infty)$ κυρτή, μη αρνητική συνάρτηση με $f(0) = 0$. Ορίζουμε $F : [0, +\infty) \to \mathbb{R}$ με $F(0) = 0$ και

$$F(x) = \frac{1}{x} \int_{0}^{x} f(t)dt.$$

Δείξτε ότι η F είναι κυρτή.

Κεφάλαιο 2

Κυρτά σύνολα στον Ευκλείδειο χώρο

2.1 Τοπολογία του Ευκλείδειου χώρου

§1. Σε αυτήν την παράγραφο θα θυμηθούμε βασικά στοιχεία της τοπολογίας του Ευκλείδειου χώρου $\mathbb{R}^n$. Σταθεροποιούμε ένα σύστημα συντεταγμένων και θεωρούμε το *εσωτερικό γινόμενο*

$$\langle x, y\rangle = x_1y_1 + \ldots + x_ny_n$$

όπου $x = (x_1, \ldots, x_n)$ και $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Η *νόρμα* του x συμβολίζεται με $\|x\|$ και ορίζεται από την

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Γεωμετρικά, $\|x\|$ είναι η απόσταση του x από την αρχή των αξόνων. Με βάση τους ορισμούς του εσωτερικού γινομένου και της νόρμας ελέγχουμε εύκολα τις εξής απλές, αλλά σημαντικές, ιδιότητές τους.

(α) $\langle x, x\rangle \geq 0$ και $\langle x, x\rangle = 0 \Longleftrightarrow x = 0$
(β) $\langle x, y\rangle = \langle y, x\rangle$
(γ) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
(δ) $\langle ax, y\rangle = a\langle x, y\rangle$
(ε) $\|x\| > 0$ αν $x \neq 0$
(ζ) $\|ax\| = |a| \cdot \|x\|$

για κάθε $x, y, z \in \mathbb{R}^n$ και κάθε $a \in \mathbb{R}$.

Ανισότητα Cauchy-Schwarz *Αν $x, y \in X$, τότε*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}.$$

Ισότητα ισχύει αν και μόνο αν τα x και y είναι γραμμικά εξαρτημένα.

Απόδειξη: Αν $y = 0$, τότε η ανισότητα ισχύει σαν ισότητα, και τα $0, x$ είναι γραμμικά εξαρτημένα.

Έστω $y \neq 0$. Ορίζουμε $P : \mathbb{R} \to \mathbb{R}$ με $P(\lambda) = \langle x - \lambda y, x - \lambda y\rangle$. Τότε, $P(\lambda) \geq 0$ για κάθε $\lambda \in \mathbb{R}$, δηλαδή

$$\langle x, x\rangle - 2\lambda\langle x, y\rangle + \lambda^2\langle y, y\rangle \geq 0, \quad \lambda \in \mathbb{R}.$$

Αυτό σημαίνει ότι η διακρίνουσα του τριωνύμου $P(\lambda)$ είναι μικρότερη ή ίση του 0. Δηλαδή, $4\langle x, y\rangle^2 - 4\langle x, x\rangle\langle y, y\rangle \leq 0$, απ' όπου παίρνουμε την

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}.$$

Ισότητα έχουμε αν και μόνο αν η διακρίνουσα του P είναι 0, δηλαδή, αν και μόνο αν το P έχει διπλή ρίζα λ_0. Όμως,

$$P(\lambda_0) = 0 \Longleftrightarrow x = \lambda_0 y,$$

δηλαδή, αν τα x, y είναι γραμμικά εξαρτημένα. □

Άμεση συνέπεια της ανισότητας Cauchy-Schwarz είναι η τριγωνική ανισότητα για τη νόρμα: αν $x, y \in \mathbb{R}^n$, τότε

$$\|x + y\| \leq \|x\| + \|y\|.$$

Για την απόδειξη αυτής της ανισότητας παρατηρούμε ότι

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \end{aligned}$$

από τις ιδιότητες του εσωτερικού γινομένου και την ανισότητα Cauchy-Schwarz.

§2. Όπως αναφέραμε, η νόρμα μετράει αποστάσεις από την αρχή των αξόνων. Μπορεί όμως να χρησιμοποιηθεί και για την μέτρηση της απόστασης οποιωνδήποτε δύο σημείων. Ορίζουμε την *Ευκλείδεια απόσταση* $d(x, y)$ των $x, y \in \mathbb{R}^n$ μέσω της

$$d(x, y) = \|x - y\|.$$

Δηλαδή,

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}.$$

Χρησιμοποιώντας τις ιδιότητες της νόρμας, ελέγχουμε εύκολα ότι:

(α) $d(x, y) > 0$ αν $x \neq y$.
(β) $d(y, x) = d(x, y)$.
(γ) $d(x, y) \leq d(x, z) + d(z, y)$.
(δ) $d(ax, ay) = |a| d(x, y)$.

(ε) $d(x+z, y+z) = d(x,y)$.

§3. Η Ευκλείδεια απόσταση επάγει μία τοπολογία στον $\mathbb{R}^n$. Για κάθε $x \in \mathbb{R}^n$ και κάθε $r > 0$, η *ανοικτή μπάλα* $D(x,r)$ με κέντρο το x και ακτίνα r είναι το σύνολο

$$D(x,r) = \{y \in \mathbb{R}^n : d(x,y) < r\}.$$

Έστω S υποσύνολο του $\mathbb{R}^n$. Το $x \in S$ λέγεται *εσωτερικό σημείο* του S αν υπάρχει $r > 0$ τέτοιος ώστε $D(x,r) \subset S$. Ένα σύνολο S λέγεται *ανοικτό* αν κάθε σημείο του είναι εσωτερικό σημείο. Παρατηρήστε ότι κάθε ανοικτή μπάλα είναι ανοικτό σύνολο. Η οικογένεια όλων των ανοικτών υποσυνόλων του $\mathbb{R}^n$ είναι η *συνήθης τοπολογία* του $\mathbb{R}^n$.

Αν S είναι ένα υποσύνολο του $\mathbb{R}^n$, τότε η οικογένεια όλων των συνόλων U της μορφής $U = V \cap S$ όπου V ανοικτό σύνολο, λέγεται *σχετική τοπολογία του* S. Για παράδειγμα, αν θεωρήσουμε το υποσύνολο $S = \{(x_1, x_2) : x_1 \geq 0\}$ του επιπέδου, τότε το

$$U = \{(x_1, x_2) : 0 \leq x_1 < 1,\ 0 < x_2 < 1\}$$

είναι ανοικτό στην σχετική τοπολογία του S, αλλά δεν είναι ανοικτό σύνολο στον $\mathbb{R}^2$.

Παρατηρήστε ότι:

(α) Το κενό σύνολο $\emptyset$ και ο $\mathbb{R}^n$ είναι ανοικτά σύνολα.

(β) Η ένωση ανοικτών συνόλων είναι ανοικτό σύνολο.

(γ) Η τομή πεπερασμένων το πλήθος ανοικτών συνόλων είναι ανοικτό σύνολο. Δεν είναι δύσκολο να δώσουμε παράδειγμα άπειρης αριθμήσιμης τομής ανοικτών συνόλων που δεν είναι ανοικτό σύνολο: για κάθε $x \in \mathbb{R}^n$, έχουμε $\bigcap_{n=1}^{\infty} D(x, 1/n) = \{x\}$ το οποίο δεν είναι ανοικτό.

Ένα σύνολο S λέγεται *κλειστό* αν το συμπλήρωμά του $\tilde{S} = \mathbb{R}^n \setminus S$ είναι ανοικτό. Παρατηρήστε ότι:

(α) Το $\emptyset$ και ο $\mathbb{R}^n$ είναι κλειστά σύνολα.

(β) Η τομή κλειστών συνόλων είναι κλειστό σύνολο.

(γ) Η ένωση πεπερασμένων το πλήθος κλειστών συνόλων είναι κλειστό σύνολο.

Είδαμε ότι το κενό σύνολο και ο χώρος $\mathbb{R}^n$ είναι ταυτόχρονα ανοικτά και κλειστά σύνολα. Στην πραγματικότητα, είναι τα μόνα υποσύνολα του $\mathbb{R}^n$ που έχουν αυτήν την ιδιότητα. Από την άλλη πλευρά, δεν είναι καθόλου δύσκολο να δώσουμε παραδείγματα υποσυνόλων του $\mathbb{R}^n$ που δεν είναι ούτε ανοικτά ούτε κλειστά.

Ένα σύνολο S λέγεται *φραγμένο* αν υπάρχουν $x \in \mathbb{R}^n$ και $R > 0$ τέτοια ώστε $S \subset D(x,R)$. Το *εσωτερικό* ενός συνόλου S είναι το σύνολο όλων των εσωτερικών σημείων του S και συμβολίζεται με $\mathrm{int}(S)$. Το $\mathrm{int}(S)$ είναι το μεγαλύτερο ανοικτό υποσύνολο του S. Η *κλειστή θήκη* ενός συνόλου S είναι το σύνολο όλων των $x \in \mathbb{R}^n$ για τα οποία υπάρχει ακολουθία $\{x_m\} \subset S$ με $x_m \to x$, και συμβολίζεται με $\overline{S}$. Το $\overline{S}$ είναι το μικρότερο κλειστό σύνολο που περιέχει το S. Για παράδειγμα, αν S είναι το σύνολο όλων των σημείων του $\mathbb{R}^n$ που έχουν ρητές συντεταγμένες, τότε $\mathrm{int}(S) = \emptyset$ και $\overline{S} = \mathbb{R}^n$.

Το *σύνορο* του S είναι το σύνολο $\mathrm{bd}(S) = \overline{S} \cap \overline{\mathbb{R}^n \setminus S}$. Ισοδύναμα, $x \in \mathrm{bd}(S)$ αν για κάθε $r > 0$ η $D(x,r)$ έχει μη κενή τομή τόσο με το S όσο και με το $\mathbb{R}^n \setminus S$.

Σημείωση Στη συνέχεια όλα τα σύνολα υποτίθενται μη κενά χωρίς αυτό να αναφέρεται, εκτός αν τονίζεται το αντίθετο ή χρειάζεται να τονίσουμε ότι κάποιο σύνολο είναι μη κενό.

§4. Μία συνάρτηση $f : A \to \mathbb{R}^m$ λέγεται *συνεχής* στο $x \in A$ αν για κάθε $\varepsilon > 0$ υπάρχει $\delta > 0$ τέτοιο ώστε

$$y \in D(x, \delta) \cap A \Longrightarrow f(y) \in D(f(x), \varepsilon).$$

Η f λέγεται *συνεχής* αν είναι συνεχής σε κάθε $x \in A$. Η f είναι συνεχής αν και μόνο αν για κάθε ανοικτό $U \subset \mathbb{R}^m$ το $f^{-1}(U)$ είναι ανοικτό υποσύνολο του A με τη σχετική τοπολογία.

Υπάρχει άλλος ένας ισοδύναμος ορισμός της συνέχειας της f στο x, τον οποίο θα χρησιμοποιούμε συχνά. Η f είναι συνεχής στο $x \in A$ αν για κάθε ακολουθία $\{x_m\}$ στο A με $x_m \to x$ έπεται ότι $f(x_m) \to f(x)$.

Εύκολα ελέγχουμε ότι οι εξής συναρτήσεις είναι συνεχείς:

(α) Η συνάρτηση $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ που ορίζεται από την $f(x, y) = x + y$.

(β) Η συνάρτηση $f_z : \mathbb{R}^n \to \mathbb{R}^n$ που ορίζεται από την $f_z(x) = x + z$, όπου $z \in \mathbb{R}^n$ σταθερό.

(γ) Η συνάρτηση $f_a : \mathbb{R}^n \to \mathbb{R}^n$ που ορίζεται από την $f_a(x) = ax$, όπου $a \in \mathbb{R}$ σταθερό.

(δ) Η συνάρτηση $f : \mathbb{R} \to \mathbb{R}^n$ που ορίζεται από την $f(t) = tx + (1 - t)y$, όπου $x, y \in \mathbb{R}^n$ σταθερά.

§5. Αν $A, B \subset \mathbb{R}^n$ και $t \in \mathbb{R}$, ορίζουμε τα σύνολα

$$A + B := \{x + y : \ x \in A, \ y \in B\}$$

και

$$tA := \{tx : \ x \in A\}.$$

Σε πλήρη γενικότητα, δεν περιμένουμε ισότητες της μορφής $A + A = 2A$. Για παράδειγμα, αν $A = \{x, y\}$, τότε $A + A = \{2x, x + y, 2y\}$ και $2A = \{2x, 2y\}$: αυτά τα δύο σύνολα δεν έχουν λόγο να είναι ίσα.

Αν $A = \{x\}$ τότε γράφουμε $x + B$ για το σύνολο $\{x\} + B$. Το $x + B$ προκύπτει με μεταφορά κάθε σημείου του B κατά x. Παρατηρήστε ότι

$$A + B = \bigcup_{x \in A} (x + B) = \bigcup_{y \in B} (A + y).$$

Σύνολα της μορφής $x + tA$ με $t \neq 0$ λέγονται *ομοιοθετικά* προς το A ή *ομοιόθετα* του A.

§6. Ένα υποσύνολο K του $\mathbb{R}^n$ λέγεται *συμπαγές* αν είναι κλειστό και φραγμένο. Ένας ισοδύναμος ορισμός μπορεί να δοθεί με τη βοήθεια των *ανοικτών καλύψεων*: μία συλλογή ανοικτών συνόλων $\{U_i : i \in I\}$ λέγεται ανοικτή κάλυψη του συνόλου S αν $S \subset \bigcup_{i \in I} U_i$. Μία υποοικογένεια $\{U_j : j \in J\}$ της $\{U_i : i \in I\}$ λέγεται *υποκάλυψη* της $\{U_i : i \in I\}$ για το S αν $S \subset \bigcup_{j \in J} U_j$. Αποδεικνύεται ότι:

Θεώρημα Heine-Borel *Ένα υποσύνολο K του $\mathbb{R}^n$ είναι συμπαγές αν και μόνο αν κάθε ανοικτή του κάλυψη έχει πεπερασμένη υποκάλυψη.*

Ένας άλλος ισοδύναμος ορισμός της συμπάγειας είναι ο εξής: το K είναι συμπαγές αν και μόνο αν για κάθε ακολουθία $\{x_m\} \subset K$ μπορούμε να βρούμε υπακολουθία $\{x_{k_m}\}$ της $\{x_m\}$ και $x \in K$ έτσι ώστε $x_{k_m} \to x$.

Η απόδειξη των ισοδυναμιών αφήνεται για τις ασκήσεις. Στη συνέχεια θα χρησιμοποιούμε ελεύθερα και τους τρείς ορισμούς. Κλείνουμε αυτήν την παράγραφο με δύο βασικές προτάσεις για συνεχείς συναρτήσεις σε συμπαγή σύνολα.

Πρόταση 2.1 *Έστω $f : K \to \mathbb{R}^m$ συνεχής συνάρτηση, όπου K συμπαγές υποσύνολο του $\mathbb{R}^n$. Τότε, το $f(K)$ είναι συμπαγές υποσύνολο του $\mathbb{R}^m$.*

Απόδειξη: Θεωρούμε τυχούσα ανοικτή κάλυψη $\{U_i : i \in I\}$ του $f(K)$. Αφού η f είναι συνεχής, κάθε $V_i = f^{-1}(U_i)$ είναι ανοικτό σύνολο (ως προς K) και αφού $\bigcup_{i\in I} f^{-1}(U_i) \supset f^{-1}(f(K)) \supseteq K$, η οικογένεια $\{V_i : i \in I\}$ είναι ανοικτή κάλυψη του K. Το K είναι συμπαγές, άρα υπάρχουν $i_1, \ldots, i_m \in I$ τέτοιοι ώστε $K \subset V_{i_1} \cup \ldots \cup V_{i_m}$. Τότε, $f(K) \subset U_{i_1} \cup \ldots \cup U_{i_m}$. Βρήκαμε πεπερασμένη υποκάλυψη της τυχούσας ανοικτής κάλυψης $\{U_i : i \in I\}$ του $f(K)$, άρα το $f(K)$ είναι συμπαγές. □

Πόρισμα *Έστω $f : K \to \mathbb{R}$ συνεχής συνάρτηση, όπου K συμπαγές υποσύνολο του $\mathbb{R}^n$. Τότε, η f παίρνει ελάχιστη και μέγιστη τιμή στο K. Δηλαδή, υπάρχουν $x_1, x_2 \in K$ τέτοια ώστε*

$$f(x_1) \leq f(x) \leq f(x_2)$$

για κάθε $x \in K$. Ειδικότερα, η f είναι φραγμένη στο K: υπάρχει $M > 0$ με την ιδιότητα $|f(x)| \leq M$ για κάθε $x \in K$. □

2.2 Κυρτά σύνολα στον Ευκλείδειο χώρο

§1. **Ορισμός** (α) Έστω $A \neq \emptyset$ υποσύνολο του $\mathbb{R}^n$. Λέμε ότι το A είναι *κυρτό* αν για κάθε $x, y \in A$ και κάθε $\lambda \in [0,1]$ έχουμε

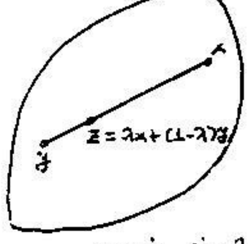
$$\lambda x + (1-\lambda)y \in A.$$

Το σύνολο $\{\lambda x + (1-\lambda)y : \lambda \in [0,1]\}$ είναι το ευθύγραμμο τμήμα με άκρα τα x και y, επομένως ένα σύνολο είναι κυρτό αν περιέχει το ευθύγραμμο τμήμα που συνδέει οποιαδήποτε δύο σημεία του.

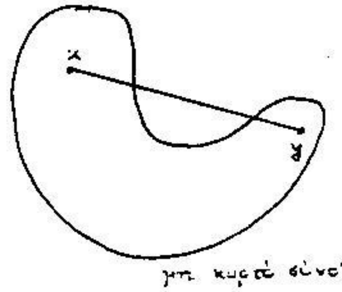
(β) Λέμε ότι το A είναι *αφφινικό* αν για κάθε $x, y \in A$ και κάθε $\lambda \in \mathbb{R}$ έχουμε

$$\lambda x + (1-\lambda)y \in A.$$

Το σύνολο $\{\lambda x + (1-\lambda)y : \lambda \in \mathbb{R}\}$ είναι η ευθεία που περνάει από τα x και y, επομένως ένα σύνολο είναι αφφινικό αν περιέχει την ευθεία που περνάει από οποιαδήποτε δύο σημεία του. Κάθε αφφινικό σύνολο είναι προφανώς κυρτό. Το αντίστροφο όμως δεν ισχύει.



κυρτό σύνολο



μη κυρτό σύνολο

Θεώρημα 2.1 *Ένα σύνολο $A \subseteq \mathbb{R}^n$ είναι αφφινικό αν και μόνο αν είναι μεταφορά κάποιου γραμμικού υποχώρου του $\mathbb{R}^n$.*

Απόδειξη: Έστω ότι $A = x + W = \{x + w : w \in W\}$ όπου W γραμμικός υπόχωρος του $\mathbb{R}^n$. Αν $y_1, y_2 \in A$, υπάρχουν $w_1, w_2 \in W$ τέτοια ώστε $y_1 = x + w_1$ και $y_2 = x + w_2$. Για κάθε $\lambda \in \mathbb{R}$ έχουμε

$$\lambda y_1 + (1-\lambda) y_2 = \lambda(x + w_1) + (1-\lambda)(x + w_2) = x + (\lambda w_1 + (1-\lambda) w_2) \in A$$

γιατί $\lambda w_1 + (1-\lambda) w_2 \in W$ (ο W είναι υπόχωρος). Άρα, το A είναι αφφινικό.

Αντίστροφα, ας υποθέσουμε ότι το A είναι αφφινικό. Θεωρούμε τυχόν $x_0 \in A$ και ορίζουμε $W = -x_0 + A$. Έστω $w_1, w_2 \in W$ και $\lambda \in \mathbb{R}$. Θα δείξουμε ότι $w_1 + \lambda w_2 \in W$. Γράφουμε $w_1 = -x_0 + y_1$, $w_2 = -x_0 + y_2$ όπου $y_1, y_2 \in A$. Τότε,

$$w_1 + \lambda w_2 = -x_0 + \lambda\left(2\left(\frac{1}{2}y_1 + \frac{1}{2}y_2\right) - x_0\right) + (1-\lambda) y_1.$$

Αφού το A είναι αφφινικό, έχουμε $z := \frac{1}{2}y_1 + \frac{1}{2}y_2 \in A$, και αφού $x_0 \in A$ έπεται ότι $2z + (-1)x_0 \in A$. Πάλι από την αφφινικότητα του A, βλέπουμε ότι

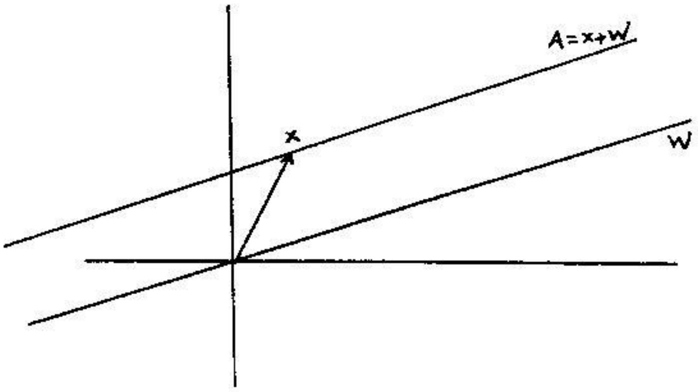
$$u := \lambda\left(2\left(\frac{1}{2}y_1 + \frac{1}{2}y_2\right) - x_0\right) + (1-\lambda) y_1 \in A.$$

άρα το $w_1 + \lambda w_2 = -x_0 + u \in W$.

Αφού $0 \in W$, παίρνοντας $w_1 = 0$ έχουμε: $w \in W \Longrightarrow \lambda w \in W$ για κάθε $\lambda \in \mathbb{R}$. Παίρνοντας $\lambda = 1$ έχουμε: $w_1, w_2 \in W \Longrightarrow w_1 + w_2 \in W$. Άρα, το W είναι γραμμικός υπόχωρος του $\mathbb{R}^n$. □

Παρατήρηση Ο υπόχωρος W του Θεωρήματος 2.1 ορίζεται μονοσήμαντα από το A

(Άσκηση 1).



Ορισμός Αν $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ και $\sum_{i=1}^m \lambda_i = 1$, το $x = \sum_{i=1}^m \lambda_i x_i$ λέγεται *αφφινικός συνδυασμός* των $x_1, \ldots, x_m$. Αν επιπλέον $\lambda_i \geq 0$, το x λέγεται *κυρτός συνδυασμός* των $x_1, \ldots, x_m$.

Θεώρημα 2.2 *Ένα σύνολο $A \subseteq \mathbb{R}^n$ είναι κυρτό (αφφινικό) αν και μόνο αν κάθε κυρτός (αφφινικός) συνδυασμός σημείων του A ανήκει στο A.*

Απόδειξη: Η μία κατεύθυνση είναι απλή: αν κάθε κυρτός συνδυασμός σημείων του A ανήκει στο A, τότε για κάθε $x, y \in A$ και $\lambda \in [0,1]$, παίρνοντας $\lambda_1 = \lambda$ και $\lambda_2 = 1 - \lambda$ βλέπουμε ότι ο κυρτός συνδυασμός $\lambda x + (1-\lambda)y \in A$, δηλαδή το A είναι κυρτό.

Για την αντίστροφη κατεύθυνση, θεωρούμε ένα κυρτό σύνολο A, $x_1, \ldots, x_m \in A$ και $\lambda_i \geq 0$ με $\sum_{i=1}^m \lambda_i = 1$ και δείχνουμε ότι $\sum_{i=1}^m \lambda_i x_i \in A$. Η απόδειξη θα γίνει με επαγωγή ως προς το m. Ο ορισμός της κυρτότητας του A εξασφαλίζει την περίπτωση $m = 2$. Υποθέτουμε ότι ο ισχυρισμός αληθεύει για οποιονδήποτε κυρτό συνδυασμό από m ή λιγότερα σημεία, και θεωρούμε τον κυρτό συνδυασμό $x = \sum_{i=1}^{m+1} \lambda_i x_i$, $x_i \in A$. Αφού $\sum_{i=1}^{m+1} \lambda_i = 1$, κάποιος από τους λ_i είναι μικρότερος από 1. Αλλάζοντας την διάταξη των λ_i, μπορούμε να υποθέσουμε ότι $\lambda_{m+1} < 1$. Τότε,

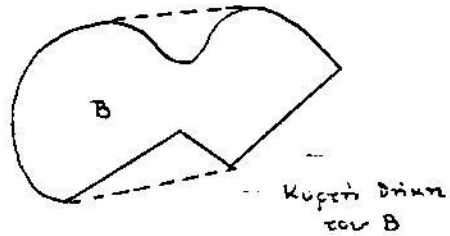
$$\begin{aligned} x &= (1 - \lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1} \\ &=: (1 - \lambda_{m+1}) y + \lambda_{m+1} x_{m+1}. \end{aligned}$$

Αφού $\sum_{i=1}^m \lambda_i = 1 - \lambda_{m+1}$, το y είναι κυρτός συνδυασμός m σημείων του A. Από την επαγωγική υπόθεση παίρνουμε $y \in A$, και χρησιμοποιώντας ξανά την επαγωγική υπόθεση για τα y και x_{m+1} ($m = 2$) συμπεραίνουμε ότι $x \in A$. Η απόδειξη του χαρακτηρισμού για αφφινικά σύνολα είναι όμοια. □

Θεώρημα 2.3 *Οποιαδήποτε τομή (αφφινικών) κυρτών υποσυνόλων του $\mathbb{R}^n$ είναι* *(αφφινικό) κυρτό σύνολο. Αν $\{A_m\}$ είναι μία αύξουσα ακολουθία (αφφινικών) κυρτών υποσυνόλων του $\mathbb{R}^n$, τότε η ένωση $\bigcup_{m=1}^{\infty} A_m$ είναι (αφφινικό) κυρτό σύνολο.*

Απόδειξη: Άσκηση 9. □

§2. **Ορισμός** Έστω $B \subset \mathbb{R}^n$. Η τομή όλων των κυρτών υποσυνόλων του $\mathbb{R}^n$ που περιέχουν το B λέγεται *κυρτή θήκη* του B και συμβολίζεται με $\mathrm{co}(B)$. Όμοια, η τομή όλων των αφφινικών υποσυνόλων του $\mathbb{R}^n$ που περιέχουν το B λέγεται *αφφινική θήκη* του B και συμβολίζεται με $\mathrm{aff}(B)$. Σύμφωνα με το Θεώρημα 2.3, η $\mathrm{co}(B)$ είναι κυρτό σύνολο και η $\mathrm{aff}(B)$ είναι αφφινικό σύνολο.



Θεώρημα 2.4 *Για κάθε $B \subset \mathbb{R}^n$, η κυρτή (αφφινική) θήκη του B αποτελείται ακριβώς από όλους τους κυρτούς (αφφινικούς) συνδυασμούς στοιχείων του B.*

Απόδειξη: Θα δείξουμε μόνο τον ισχυρισμό για την κυρτή θήκη. Γράφουμε C για το σύνολο όλων των κυρτών συνδυασμών σημείων του B. Αφού η $\mathrm{co}(B)$ είναι κυρτό σύνολο και περιέχει το B, από το Θεώρημα 2.2 βλέπουμε ότι $C \subseteq \mathrm{co}(B)$. Αντίστροφα, αν $x = \sum_{i=1}^{k} a_i x_i$ και $y = \sum_{j=1}^{m} b_j y_j$ είναι δύο σημεία του C, τότε για κάθε $\lambda \in [0,1]$ έχουμε

$$\lambda x + (1-\lambda) y = \sum_{i=1}^{k} (\lambda a_i) x_i + \sum_{j=1}^{m} ((1-\lambda) b_j) y_j \in C$$

γιατί $x_i, y_j \in B$ και $\sum_{i=1}^{k} \lambda a_i + \sum_{j=1}^{m} (1-\lambda) b_j = 1$. Δηλαδή, το C είναι ένα κυρτό σύνολο που περιέχει το B. Από τον ορισμό της $\mathrm{co}(B)$ παίρνουμε $\mathrm{co}(B) \subseteq C$. □

Το Θεώρημα 2.4 βελτιώνεται ως εξής: κάθε σημείο της κυρτής θήκης $\mathrm{co}(B)$ του B γράφεται σαν κυρτός συνδυασμός το πολύ $n+1$ σημείων του B. Θα δείξουμε μάλιστα ένα κάπως ισχυρότερο αποτέλεσμα. Για το σκοπό αυτό, ορίζουμε την *διάσταση* ενός κυρτού συνόλου B στον $\mathbb{R}^n$. Αν το B είναι αφφινικό, τότε υπάρχει μεταφορά $x + B$ του B που είναι γραμμικός υπόχωρος του $\mathbb{R}^n$ (Θεώρημα 2.1). Τότε, η διάσταση του B ορίζεται να είναι η διάσταση του υποχώρου $x + B$. Αν το B είναι τυχόν κυρτό σύνολο, ορίζουμε σαν διάσταση του B την διάσταση της αφφινικής του θήκης.

Θεώρημα του Καραθεοδωρή. *Αν $B \subseteq \mathbb{R}^n$ και η $\mathrm{co}(B)$ έχει διάσταση m, τότε για κάθε $z \in \mathrm{co}(B)$ υπάρχουν $x_0, \ldots, x_s \in B$ ($s \leq m$) τέτοια ώστε το z να είναι κυρτός συνδυασμός τους.*

Απόδειξη: Έστω $z \in \mathrm{co}(B)$. Τότε, υπάρχουν $x_0, \ldots, x_k \in B$ και $a_i > 0$ με $\sum_{i=0}^k a_i = 1$ τέτοια ώστε $z = \sum_{i=0}^k a_i x_i$. Υποθέτουμε ότι $k > m$ και ορίζουμε $C = \{x_0, \ldots, x_k\}$. Παρατηρούμε ότι η διάσταση της αφφινικής θήκης του C είναι μικρότερη ή ίση από $m < k$, άρα τα $x_1 - x_0, \ldots, x_k - x_0$ είναι γραμμικά εξαρτημένα. Επομένως, υπάρχουν $b_1, \ldots, b_k$ όχι όλα μηδέν, τέτοια ώστε $\sum_{i=1}^k b_i(x_i - x_0) = 0$. Αν ορίσουμε $b_0 = -\sum_{i=1}^k b_i$, τότε

$$\sum_{i=0}^k b_i x_i = 0, \qquad \sum_{i=0}^k b_i = 0.$$

Αφού όλα τα a_i είναι γνήσια θετικά, μπορούμε να βρούμε $t > 0$ με την ιδιότητα $c_i = a_i - tb_i \geq 0$ για κάθε i και $c_l = 0$ για κάποιο $l \leq k$. Τότε,

$$z = \sum_{i=0}^k a_i x_i = \sum_{i=0}^k (c_i + tb_i) x_i = \sum_{i \neq l} c_i x_i$$

και

$$\sum_{i \neq l} c_i = \sum_{i=0}^k c_i = \sum_{i=0}^k (a_i - tb_i) = \sum_{i=0}^k a_i = 1.$$

Άρα, το z γράφεται σαν κυρτός συνδυασμός λιγότερων από $k+1$ σημείων του B. Τώρα, είτε $k = m$ οπότε έχουμε τελειώσει ή συνεχίζουμε με τον ίδιο τρόπο. Τελικά, θα πετύχουμε να γράψουμε το z σαν κυρτό συνδυασμό $m+1$ ή λιγότερων σημείων του B. $\square$

Ορισμός Τα $x_0, x_1, \ldots, x_m$ λέγονται αφφινικά ανεξάρτητα αν τα $x_1 - x_0, \ldots, x_m - x_0$ είναι γραμμικά ανεξάρτητα, δηλαδή αν η διάσταση της αφφινικής θήκης του $\{x_0, x_1, \ldots, x_m\}$ είναι ίση με m. Ισοδύναμα, αν οποτεδήποτε $\lambda_0 x_0 + \lambda_1 x_1 + \ldots + \lambda_m x_m = 0$ και $\lambda_0 + \ldots + \lambda_m = 0$, έπεται ότι $\lambda_0 = \ldots = \lambda_m = 0$ (Άσκηση 2).

Τα απλούστερα κυρτά σύνολα είναι οι κυρτές θήκες πεπερασμένων το πλήθος σημείων, δηλαδή τα κυρτά σύνολα της μορφής $\mathrm{co}\{x_0, \ldots, x_m\}$, $x_i \in \mathbb{R}^n$. Κάθε τέτοιο σύνολο λέγεται *πολύτοπο*. Αν τα $x_1 - x_0, \ldots, x_m - x_0$ είναι γραμμικά ανεξάρτητα στον $\mathbb{R}^n$ (οπότε υποχρεωτικά $m \leq n$), τότε το $\mathrm{co}\{x_0, x_1, \ldots, x_m\}$ λέγεται m-simplex με κορυφές τα $x_0, x_1, \ldots, x_m$. Ένα 0-simplex είναι απλώς ένα σημείο, ένα 1-simplex είναι ένα ευθύγραμμο τμήμα, ένα 2-simplex είναι ένα τρίγωνο και ούτω καθεξής. Κάθε σημείο x ενός m-simplex γράφεται με μοναδικό τρόπο σαν κυρτός συνδυασμός της μορφής $x = \sum_{i=1}^m \lambda_i x_i$. Οι συντελεστές λ_i λέγονται *βαρυκεντρικές συντεταγμένες* του x.

Εξετάζουμε τώρα τις βασικές τοπολογικές ιδιότητες των κυρτών συνόλων.

Θεώρημα 2.5 *Αν B είναι ένα κυρτό σύνολο, τότε το $\mathrm{int}(B)$ και το $\overline{B}$ είναι κυρτά σύνολα.*

Απόδειξη: Έστω $x, y \in \mathrm{int}(B)$ και $\lambda \in (0,1)$. Αν $z = \lambda x + (1-\lambda)y$, για κάθε $w \in \mathbb{R}^n$ μπορούμε να γράφουμε

$$z + w = \lambda(x+w) + (1-\lambda)(y+w).$$

Αφού το $\lambda > 0$ είναι δεδομένο και $x, y \in \mathrm{int}(B)$, μπορούμε να βρούμε $\delta > 0$ τέτοιο ώστε $\|w\| < \delta \Longrightarrow x+w, y+w \in B$. Όμως τότε, από την κυρτότητα του B βλέπουμε ότι $D(z,\delta) \subseteq B$. Άρα, $z \in \mathrm{int}(B)$ και το $\mathrm{int}(B)$ είναι κυρτό.

Έστω τώρα $x, y \in \overline{B}$ και $\lambda \in (0,1)$. Υπάρχουν $x_m, y_m \in B$ με $x_m \to x$ και $y_m \to y$. Το B είναι κυρτό, άρα $\lambda x_m + (1-\lambda)y_m \in B$ για κάθε m. Έπεται ότι

$$\lambda x + (1-\lambda)y = \lim_{m\to\infty}(\lambda x_m + (1-\lambda)y_m) \in \overline{B}.$$

Άρα, το $\overline{B}$ είναι κυρτό. □

Θεώρημα 2.6 *Έστω B υποσύνολο του $\mathbb{R}^n$. Αν το B είναι ανοικτό, τότε η κυρτή του θήκη $\mathrm{co}(B)$ είναι ανοικτό σύνολο. Αν το B είναι συμπαγές, τότε η κυρτή του θήκη είναι συμπαγής.*

Απόδειξη: Υποθέτουμε πρώτα ότι το B είναι ανοικτό και θεωρούμε τυχόν στοιχείο $x = \sum_{i=1}^m \lambda_i x_i$ της $\mathrm{co}(B)$. Για κάθε $w \in \mathbb{R}^n$ έχουμε $x + w = \sum_{i=1}^m \lambda_i(x_i + w)$ και αφού κάθε $x_i \in B$ και το B είναι ανοικτό, μπορούμε να βρούμε $\delta > 0$ με την ιδιότητα

$$\|w\| < \delta \Longrightarrow x_i + w \in B, \quad i = 1, \ldots, m.$$

Από την κυρτότητα της $\mathrm{co}(B)$ βλέπουμε ότι

$$\|w\| < \delta \Longrightarrow x + w = \sum_{i=1}^m \lambda_i(x_i + w) \in \mathrm{co}(B),$$

δηλαδή $D(x,\delta) \subseteq \mathrm{co}(B)$. Άρα, η $\mathrm{co}(B)$ είναι ανοικτό σύνολο.

Ένας άλλος τρόπος απόδειξης είναι ο εξής: αφού $\mathrm{co}(B) \supseteq B$, έχουμε

$$\mathrm{int}(\mathrm{co}(B)) \supseteq \mathrm{int}(B) = B$$

γιατί το B είναι ανοικτό. Όμως το $\mathrm{int}(\mathrm{co}(B))$ είναι κυρτό σύνολο και περιέχει το B, άρα $\mathrm{int}(\mathrm{co}(B)) \supseteq \mathrm{co}(B)$. Αυτό σημαίνει ότι η $\mathrm{co}(B)$ είναι ανοικτό σύνολο.

Έστω τώρα ότι το B είναι συμπαγές. Αυτό σημαίνει ότι το B είναι κλειστό και φραγμένο. Ειδικότερα, υπάρχει $M > 0$ τέτοιος ώστε $\|x\| \leq M$ για κάθε $x \in B$. Αν $y \in \mathrm{co}(B)$, τότε $y = \sum_{i=1}^m \lambda_i x_i$ όπου $x_i \in B$, $\lambda_i > 0$ και $\sum_{i=1}^m \lambda_i = 1$, άρα

$$\|y\| \leq \sum_{i=1}^m \lambda_i \|x_i\| \leq M \sum_{i=1}^m \lambda_i = M,$$

άρα η $\mathrm{co}(B)$ είναι φραγμένο σύνολο.

Για να δείξουμε ότι η $\mathrm{co}(B)$ είναι κλειστό σύνολο, θεωρούμε ακολουθία $\{x_s\}$ σημείων της $\mathrm{co}(B)$ με $x_s \to x \in \mathbb{R}^n$. Από το Θεώρημα του Καραθεοδωρή μπορούμε να υποθέσουμε ότι κάθε x_s γράφεται στη μορφή

$$x_s = \sum_{i=1}^{n+1} \lambda_{s,i} x_{s,i}$$

όπου $x_{s,i} \in B$, $\lambda_{s,i} \geq 0$ και $\sum_{i=1}^{n+1} \lambda_{s,i} = 1$. Οι ακολουθίες $\lambda_{s,i}$, $x_{s,i}$ είναι φραγμένες ακολουθίες στο $\mathbb{R}$ και στον $\mathbb{R}^n$ αντίστοιχα, επομένως μπορούμε να βρούμε αύξουσα ακολουθία δεικτών $s_1 < s_2 < \ldots < s_r < \ldots$ και $\lambda_i \in \mathbb{R}$, $x_i \in \mathbb{R}^n$ τέτοια ώστε

$$\lim_{r\to\infty} \lambda_{s_r,i} \to \lambda_i, \quad \lim_{r\to\infty} x_{s_r,i} \to x_i, \quad i = 1, \ldots, n+1.$$

Τότε,

$$x = \lim_{r\to\infty} x_{s_r} = \lim_{r\to\infty} \sum_{i=1}^{n+1} \lambda_{s_r,i} x_{s_r,i} = \sum_{i=1}^{n+1} \lambda_i x_i$$

και

$$\sum_{i=1}^{n+1} \lambda_i = 1 = \lim_{r\to\infty} \sum_{i=1}^{n+1} \lambda_{s_r,i} = 1.$$

Αφού το B είναι κλειστό, κάθε $x_i \in B$ και αφού το x είναι κυρτός συνδυασμός των x_i, έχουμε $x \in \mathrm{co}(B)$. Άρα, η $\mathrm{co}(B)$ είναι κλειστό σύνολο. □

Παρατήρηση Η κυρτή θήκη κλειστού συνόλου δεν είναι πάντα κλειστό σύνολο. Πάρτε σαν παράδειγμα την $\mathrm{co}(A)$ όπου $A = \{(x,0) : x \in \mathbb{R}\} \bigcup \{(0,1)\}$ στο $\mathbb{R}^2$ (δείτε και την Άσκηση 17).

Θα κλείσουμε αυτήν την παράγραφο με κάποιες παρατηρήσεις για τη σχετική τοπολογία. Το *σχετικό εσωτερικό* ενός κυρτού συνόλου B είναι το εσωτερικό του συνόλου αν το δούμε σαν υποσύνολο της αφφινικής του θήκης. Αποτελείται δηλαδή από όλα τα σημεία $x \in \mathrm{aff}(B)$ για τα οποία υπάρχει $\varepsilon > 0$ τέτοιο ώστε: αν $y \in \mathrm{aff}(B)$ και $\|x - y\| < \varepsilon$, τότε $y \in B$. Για παράδειγμα, το σχετικό εσωτερικό του δίσκου $\{(r,s,1) \in \mathbb{R}^3 : r^2 + s^2 \leq 1\}$ είναι το σύνολο $\{(r,s,1) \in \mathbb{R}^3 : r^2 + s^2 < 1\}$, ενώ το εσωτερικό του είναι το κενό σύνολο. Ανάλογα μπορεί κανείς να ορίσει το σχετικό σύνορο ενός κυρτού συνόλου. Ένα κυρτό σύνολο λέγεται *σχετικά ανοικτό* αν συμπίπτει με το σχετικό εσωτερικό του. Πολλά από τα θεωρήματα που θα αποδείξουμε στη συνέχεια αφορούν ανοικτά σύνολα. Συνήθως, εξακολουθούν να ισχύουν αν αντικαταστήσουμε τη λέξη «ανοικτό» με τη λέξη «σχετικά ανοικτό». Απλώς κοιτάμε το σχετικά ανοικτό σύνολό μας σαν ανοικτό στον υπόχωρο που προκύπτει με κατάλληλη μεταφορά της αφφινικής του θήκης.

2.3 Υπερεπίπεδα και διαχωριστικά θεωρήματα

§1. **Ορισμοί** (α) Μία συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}^m$ λέγεται *γραμμική* αν

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

για κάθε $x, y \in \mathbb{R}^n$ και $\lambda, \mu \in \mathbb{R}$.

Η f λέγεται *αφφινική* αν

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

για κάθε $x, y \in \mathbb{R}^n$ και $\lambda, \mu \in \mathbb{R}$ με $\lambda + \mu = 1$.

(β) Μία γραμμική συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$ λέγεται *γραμμικό συναρτησοειδές*. Αν $f : \mathbb{R}^n \to \mathbb{R}$ είναι ένα γραμμικό συναρτησοειδές και $a_i = f(e_i)$, τότε για κάθε $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ έχουμε

$$f(x) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} x_i a_i = \langle x, u \rangle$$

όπου $u = (a_1, \ldots, a_n)$.

(γ) Ένα αφφινικό υποσύνολο H του $\mathbb{R}^n$ λέγεται υπερεπίπεδο αν $\dim(H) = n-1$. Η σχέση ανάμεσα σε υπερεπίπεδα και γραμμικά συναρτησοειδή περιγράφεται από το εξής θεώρημα.

Θεώρημα 2.7 *Έστω $H \subset \mathbb{R}^n$. Το H είναι υπερεπίπεδο αν και μόνο αν υπάρχουν μη μηδενικό γραμμικό συναρτησοειδές $f : \mathbb{R}^n \to \mathbb{R}$ και $t \in \mathbb{R}$ τέτοια ώστε $H = \{x : f(x) = t\}$.*

Απόδειξη: Υποθέτουμε πρώτα ότι το H είναι υπερεπίπεδο. Αν $x_0 \in H$, τότε $H = x_0 + W$ όπου W (μονοσήμαντα ορισμένος) γραμμικός υπόχωρος του $\mathbb{R}^n$ με $\dim(W) = n-1$. Θεωρούμε ένα μοναδιαίο διάνυσμα v_0 κάθετο στον W και ορίζουμε $f : \mathbb{R}^n \to \mathbb{R}$ με $f(x) = \langle x, v_0 \rangle$. Τότε,

$$H = \{x : f(x) = \langle x_0, v_0 \rangle\}.$$

Πράγματι, αν $x \in H$ τότε $x = x_0 + w$ για κάποιο $w \in W$, άρα

$$f(x) = \langle x_0 + w, v_0 \rangle = \langle x_0, v_0 \rangle + \langle w, v_0 \rangle = \langle x_0, v_0 \rangle$$

αφού $\langle w, v_0 \rangle = 0$. Αντίστροφα, αν $f(x) = \langle x, v_0 \rangle = \langle x_0, v_0 \rangle$, τότε $\langle x - x_0, v_0 \rangle = 0$ άρα $x - x_0 \in W$, οπότε $x \in x_0 + W = H$.

Αντίστροφα, έστω $f(x) = \langle x, u \rangle$ μη μηδενικό γραμμικό συναρτησοειδές (δηλαδή, $u \neq 0$) και $t \in \mathbb{R}$. Έστω $x_0 \in \mathbb{R}^n$ με $f(x_0) = t$ (τέτοιο x_0 υπάρχει, πάρτε για παράδειγμα το $tu/\|u\|^2$). Τότε, $f(x) = t$ αν και μόνο αν $\langle x - x_0, u \rangle = 0$ δηλαδή αν $x - x_0 \in W$ όπου $W = u^{\perp} = \{w : \langle u, w \rangle = 0\}$. Ο W είναι γραμμικός υπόχωρος του $\mathbb{R}^n$ διάστασης $(n-1)$, άρα το $\{x : f(x) = t\} = x_0 + W$ είναι υπερεπίπεδο. $\square$

§2. **Ορισμοί** (α) Αν $H = \{x : \langle x, u \rangle = t\}$ είναι ένα υπερεπίπεδο, λέμε ότι τα σύνολα $H^- = \{x : \langle x, u \rangle \leq t\}$ και $H^+ = \{x : \langle x, u \rangle \geq t\}$ είναι οι *ημίχωροι* που ορίζονται από το H.

(β) Αν A και B είναι υποσύνολα του $\mathbb{R}^n$, λέμε ότι το υπερεπίπεδο H *διαχωρίζει* τα A και B αν αυτά περιέχονται σε διαφορετικούς ημιχώρους που ορίζονται από το H. Λέμε ότι τα A και B διαχωρίζονται γνήσια αν υπάρχουν $t < s$ και $u \in \mathbb{R}^n$ τέτοια ώστε $A \subseteq \{x : \langle u, x \rangle \leq t\}$ και $B \subseteq \{x : \langle u, x \rangle \geq s\}$. Δηλαδή, αν μπορούμε να βρούμε δύο παράλληλα υπερεπίπεδα τα οποία να διαχωρίζουν τα A και B. Παρατηρήστε ότι αναγκαία συνθήκη για να διαχωρίζονται γνήσια δύο σύνολα είναι να είναι ξένα, ενώ δύο δύνολα που απλώς διαχωρίζονται δεν είναι αναγκαστικά ξένα. Για παράδειγμα οποιαδήποτε εφαπτομένη σε έναν κύκλο του $\mathbb{R}^2$ διαχωρίζει τον κύκλο από το σημείο επαφής. Από την άλλη πλευρά, αν δύο σύνολα (ακόμα κι αν είναι κλειστά και κυρτά) είναι ξένα, αυτό δεν σημαίνει ότι διαχωρίζονται γνήσια. Πάρτε για παράδειγμα τα

$$A = \{(x,y) \in \mathbb{R}^2 : x > 0, y \geq 1/x\}, \qquad B = \{(x,y) \in \mathbb{R}^2 : x > 0, y \leq -1/x\}.$$

Αν όμως το ένα από τα δύο σύνολα είναι συμπαγές, τότε έχουμε γνήσιο διαχωρισμό.

Θεώρημα 2.8 *Έστω A και B ξένα, μη-κενά και κυρτά υποσύνολα του $\mathbb{R}^n$. Αν το A είναι συμπαγές και το B κλειστό, τότε τα A και B διαχωρίζονται γνήσια.*

Απόδειξη: Θα χρειαστούμε το εξής λήμμα.

Λήμμα *Έστω B κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$ και $x \notin B$. Υπάρχει μοναδικό $y_0 \in B$ τέτοιο ώστε*

$$\|x - y_0\| = d(x,B) := \inf\{\|x-y\| : y \in B\}.$$

Απόδειξη: Για κάθε $m \in \mathbb{N}$ μπορούμε να βρούμε $y_m \in B$ τέτοιο ώστε

$$(*) \qquad d(x,B) \leq \|x - y_m\| < d(x,B) + \frac{1}{m}.$$

Παρατηρούμε ότι

$$2\|x-y_m\|^2 + 2\|x-y_s\|^2 = \|2x-(y_m+y_s)\|^2 + \|y_m - y_s\|^2.$$

Αφού $(y_m+y_s)/2 \in B$, παίρνουμε $\|2x-(y_m+y_s)\|^2 \geq 4d^2(x,B)$, άρα

$$\|y_m - y_s\|^2 \leq 2\left(d + \frac{1}{m}\right)^2 + 2\left(d + \frac{1}{s}\right)^2 - 4d^2 \to 0$$

όταν $m, s \to \infty$, όπου $d = d(x,B)$. Δηλαδή, η $\{y_m\}$ είναι ακολουθία Cauchy.

Κάθε ακολουθία Cauchy στον $\mathbb{R}^n$ συγκλίνει, άρα υπάρχει $y_0 \in \mathbb{R}^n$ με $y_m \to y_0$. Αφού το B είναι κλειστό, έχουμε $y_0 \in B$ και από την $(*)$, $\|x-y_0\| = \lim_m \|x-y_m\| = d(x,B)$.

Για τη μοναδικότητα, ας υποθέσουμε ότι $\|x-y\| = \|x-y_0\| = d$ για κάποιο $y \in B$. Τότε, αφού $(y+y_0)/2 \in B$,

$$4d^2 = 2\|x-y\|^2 + 2\|x-y_0\|^2 = \|2x-(y+y_0)\|^2 + \|y-y_0\|^2 \geq 4d^2 + \|y-y_0\|^2,$$

δηλαδή $\|y-y_0\| \leq 0$, άρα $y = y_0$. $\square$

Μπορούμε τώρα να αποδείξουμε το Θεώρημα 2.8: Θεωρούμε τη συνάρτηση $x \mapsto d(x,B)$ ορισμένη στο A. Η συνάρτηση απόστασης από σύνολο είναι συνεχής (Άσκηση

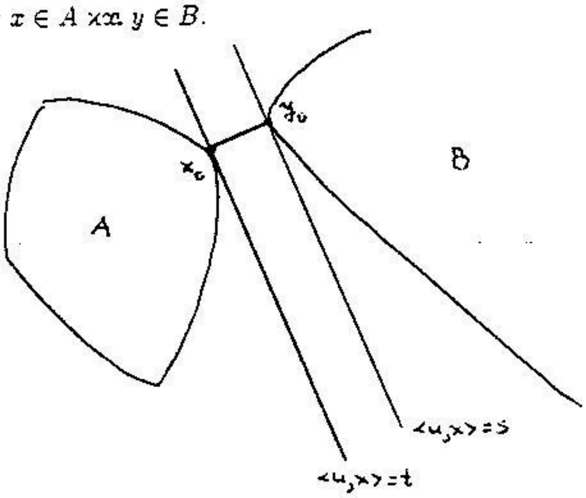
19) και παίρνει γνήσια θετικές τιμές στο A, γιατί τα A και B είναι ξένα και το B κλειστό. Άρα, υπάρχει $x_0 \in A$ στο οποίο παίρνουμε την ελάχιστη τιμή της:

$$d(x_0, B) \leq d(x, B), \quad x \in A.$$

Από το Λήμμα, υπάρχει μοναδικό $y_0 \in B$ τέτοιο ώστε $\|x_0 - y_0\| = d(x_0, B)$. Δηλαδή, βρήκαμε $x_0 \in A$ και $y_0 \in B$ με

$$\|x_0 - y_0\| = d(x_0, B) \leq d(x, B) \leq \|x - y\|$$

για κάθε $x \in A$ και $y \in B$.



Θεωρούμε το μη μηδενικό διάνυσμα $u = y_0 - x_0$ και θέτουμε $t = \langle u, x_0 \rangle$ και $s = \langle u, y_0 \rangle$. Είναι $s - t = \|u\|^2 > 0$ και τα

$$\{x : \langle u, x \rangle = t\}, \quad \{x : \langle u, x \rangle = s\}$$

είναι παράλληλα υπερεπίπεδα, κάθετα προς το u, τα οποία περνούν από τα x_0 και y_0 αντίστοιχα (βλέπε σχήμα). Θα δείξουμε ότι $\langle u, x \rangle \leq t$ για κάθε $x \in A$ και $\langle u, y \rangle \geq s$ για κάθε $y \in B$, οπότε τα A και B διαχωρίζονται γνήσια.

Έστω $x \in A$. Για κάθε $\lambda \in (0, 1)$ έχουμε $x_0 + \lambda(x - x_0) = (1 - \lambda)x_0 + \lambda x \in A$, άρα $\|x_0 + \lambda(x - x_0) - y_0\| \geq \|x_0 - y_0\|$. Υψώνοντας στο τετράγωνο έχουμε

$$\|x_0 - y_0\|^2 + 2\lambda\langle x - x_0, x_0 - y_0 \rangle + \lambda^2\|x - x_0\|^2 \geq \|x_0 - y_0\|^2,$$

δηλαδή,

$$2\langle x - x_0, x_0 - y_0 \rangle + \lambda\|x - x_0\|^2 \geq 0,$$

και αφήνοντας το $\lambda \to 0^+$ παίρνουμε $\langle x - x_0, x_0 - y_0 \rangle \geq 0$, δηλαδή

$$\langle x, y_0 - x_0 \rangle \leq \langle x_0, y_0 - x_0 \rangle = t.$$

Όμοια αποδεικνύουμε ότι $\langle u, y \rangle \geq s$ για κάθε $y \in B$. $\square$

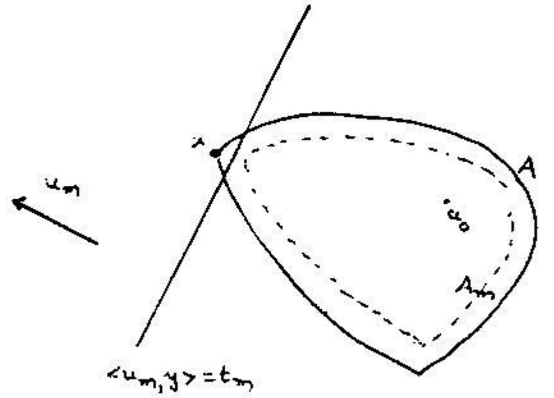
Χρησιμοποιώντας το Θεώρημα 2.8 μπορούμε τώρα να αποδείξουμε το βασικό διαχωριστικό θεώρημα για κυρτά σύνολα στον $\mathbb{R}^n$.

Λήμμα *Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ με $\text{int}(A) \neq \emptyset$ και $x \notin \text{int}(A)$. Τότε υπάρχει μη μηδενικό γραμμικό συναρτησοειδές f τέτοιο ώστε $f(x) \geq f(a)$ για κάθε $a \in A$.*

Απόδειξη: Έστω $a_0 \in \text{int}(A)$. Για κάθε $m \geq 2$ ορίζουμε

$$A_m = \{a_0 + (1 - 1/m)(a - a_0) : a \in A\}.$$

Τότε, $A_m \subseteq \text{int}(A)$ για κάθε m, η $\{A_m\}$ είναι αύξουσα ακολουθία κυρτών υποσυνόλων του A, και $\bigcup_{m=2}^{\infty} \overline{A_m} \supseteq A$ (άσκηση).



Για κάθε $m \geq 2$ έχουμε $x \notin \overline{A_m}$ (γιατί;), οπότε από το Θεώρημα 2.8 τα $\{x\}$ και $\overline{A_m}$ διαχωρίζονται γνήσια. Αυτό σημαίνει ότι υπάρχουν μοναδιαίο διάνυσμα u_m και $t_m \in \mathbb{R}$ τέτοια ώστε

$$(*) \qquad \langle u_m, a_0\rangle \leq \sup_{a \in \overline{A_m}} \langle u_m, a\rangle \leq t_m < \langle u_m, x\rangle.$$

Η $\{u_m\}$ και η $\{t_m\}$ είναι φραγμένες ακολουθίες: $\|u_m\| = 1$ και $-\|a_0\| \leq t_m \leq \|x\|$ από την $(*)$ και την ανισότητα Cauchy-Schwarz. Άρα, υπάρχουν $m_1 < \dots < m_k < \dots$ με $u_{m_k} \to u \in \mathbb{R}^n$ και $t_{m_k} \to t \in \mathbb{R}$.

Για κάθε $a \in A$ υπάρχει k_0 τέτοιος ώστε $a \in \overline{A_{m_k}}$ για κάθε $k \geq k_0$. Από την $(*)$,

$$\begin{aligned} \langle u, a\rangle &= \lim_{k\to\infty} \langle u_{m_k}, a\rangle \\ &\leq \lim_{k\to\infty} t_{m_k} = t \\ &\leq \lim_{k\to\infty} \langle u_{m_k}, x\rangle = \langle u, x\rangle. \end{aligned}$$

Το συναρτησοειδές $f(x) = \langle u, x\rangle$ ικανοποιεί το ζητούμενο. $\Box$

Θεώρημα 2.9 *Έστω A και B κυρτά υποσύνολα του $\mathbb{R}^n$ με $\text{int}(A) \neq \emptyset$ και $\text{int}(A) \cap B = \emptyset$. Τότε, υπάρχει υπερεπίπεδο που διαχωρίζει τα A και B.*

Απόδειξη: Η συνθήκη $\mathrm{int}(A) \cap B = \emptyset$ είναι ισοδύναμη με την

$$0 \notin \mathrm{int}(A) - B.$$

Το $C := \mathrm{int}(A) - B$ είναι ανοικτό και κυρτό σύνολο (δείτε τις Ασκήσεις 8 και 10). Από το Λήμμα, υπάρχει $u \in \mathbb{R}^n$ τέτοιο ώστε $\langle u, c\rangle \leq \langle u, 0\rangle = 0$ για κάθε $c \in C$. Δηλαδή,

$$\langle u, a - b\rangle \leq 0 \Longrightarrow \langle u, a\rangle \leq \langle u, b\rangle$$

για κάθε $a \in \mathrm{int}(A)$ και $b \in B$. Από τη συνέχεια του εσωτερικού γινομένου,

$$\langle u, a\rangle \leq \langle u, b\rangle$$

για κάθε $a \in A$ και $b \in B$. Έπεται ότι

$$t_0 := \sup_{a\in A} \langle u, a\rangle \leq t_1 := \inf_{b\in B} \langle u, b\rangle.$$

Άρα, το υπερεπίπεδο $\{x : \langle u, x\rangle = t\}$ διαχωρίζει τα A και B για κάθε $t \in [t_0, t_1]$. $\square$

§3. **Ορισμός** Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$. Λέμε ότι το υπερεπίπεδο H *στηρίζει* το A στο x_0, αν $x_0 \in A \cap H$ και το A περιέχεται σε κάποιον από τους δύο ημιχώρους που ορίζει το H. Εφαρμόζοντας το Θεώρημα 2.9 με $B = \{x_0\}$ όπου x_0 είναι ένα σημείο του συνόρου του κυρτού συνόλου A, παίρνουμε το εξής:

Θεώρημα 2.10 *Έστω $x_0 \in \mathrm{bd}(A)$, όπου A κυρτό υποσύνολο του $\mathbb{R}^n$ με μη κενό εσωτερικό. Τότε, υπάρχει υπερεπίπεδο H που στηρίζει το A στο x_0.* $\square$

Παρατηρήσεις (α) Ένα πόρισμα του Θεωρήματος 2.10 είναι το εξής: κάθε κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$ με μη κενό εσωτερικό είναι η τομή των ημιχώρων που το περιέχουν (Άσκηση 27).
(β) Το Θεώρημα 2.10 ισχύει και χωρίς την υπόθεση ότι $\mathrm{int}(A) \neq \emptyset$. (Άσκηση 23).

§4. **Ορισμός** Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ και $x_0 \in A$. Το x_0 λέγεται *ακραίο σημείο* του A αν δεν περιέχεται στο εσωτερικό κανενός ευθύγραμμου τμήματος του οποίου τα άκρα ανήκουν στο A. Δηλαδή, αν ισχύει το εξής: αν $x_1, x_2 \in A$ και $\lambda \in [0, 1]$ με $x_0 = \lambda x_1 + (1 - \lambda)x_2$, τότε είτε $x_0 = x_1$ ή $x_0 = x_2$.

Για παράδειγμα, τα ακραία σημεία ενός δίσκου είναι όλα τα σημεία του συνόρου του και τα ακραία σημεία ενός κύβου είναι οι κορυφές του. Ένα κλειστό κυρτό σύνολο μπορεί να μην έχει ακραία σημεία: πάρτε για παράδειγμα οποιονδήποτε ημίχωρο.

Θεώρημα 2.11 (Minkowski) *Έστω A κυρτό και συμπαγές υποσύνολο του $\mathbb{R}^n$. Τότε, το A είναι η κυρτή θήκη των ακραίων σημείων του.*

Απόδειξη: Η απόδειξη γίνεται με επαγωγή ως προς την διάσταση d του A (δηλαδή, την διάσταση της αφφινικής του θήκης).

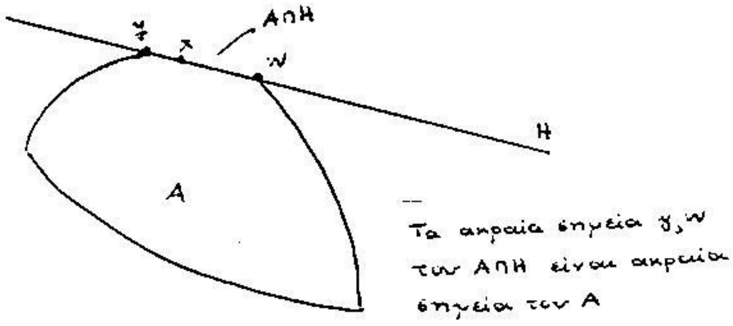
Αν $d = 0$, τότε το A είναι μονοσύνολο και ο ισχυρισμός του θεωρήματος ισχύει προφανώς. Υποθέτουμε λοιπόν ότι η διάσταση του A είναι $k > 0$ και ότι το θεώρημα ισχύει για όλα τα κυρτά συμπαγή υποσύνολα του $\mathbb{R}^n$ που έχουν διάσταση μικρότερη από k.

Το A είναι φραγμένο, άρα το σχετικό σύνορο του A αν το δούμε σαν υποσύνολο της $\mathrm{aff}(A)$ είναι μη κενό. Παίρνουμε $x \in A$ και διακρίνουμε δύο περιπτώσεις:

(α) Το x ανήκει στο σχετικό σύνορο του A στην $\mathrm{aff}(A)$. Από το Θεώρημα 2.10, υπάρχουν $u \in \mathbb{R}^n$ και $t \in \mathbb{R}$ τέτοια ώστε

$$\langle u, a\rangle \leq t = \langle u, x\rangle$$

για κάθε $a \in A$. Θεωρούμε το υπερεπίπεδο $H = \{y : \langle u, y\rangle = t\}$ και το $A \cap H$. Το $A \cap H$ είναι μη κενό, κυρτό και συμπαγές σύνολο με διάσταση μικρότερη από k. Άρα, είναι η κυρτή θήκη των ακραίων σημείων του. Αφού $x \in A \cap H$, το x γράφεται σαν κυρτός συνδυασμός ακραίων σημείων του $A \cap H$. Αν λοιπόν δείξουμε ότι κάθε ακραίο σημείο του $A \cap H$ είναι ακραίο σημείο του A, θα έχουμε δείξει ότι το x είναι στην κυρτή θήκη των ακραίων σημείων του A.



Έστω w ακραίο σημείο του $A \cap H$ και $y, z \in A$ και $\lambda \in [0, 1]$ με $w = \lambda y + (1-\lambda)z$. Τότε,

$$t = \langle u, w\rangle = \lambda\langle u, y\rangle + (1-\lambda)\langle u, z\rangle \leq \lambda t + (1-\lambda)t = t,$$

και αυτό σημαίνει ότι $\langle u, y\rangle = \langle u, z\rangle = t$, δηλαδή, $y, z \in A \cap H$. Όμως το w είναι ακραίο σημείο του $A \cap H$, άρα είτε $w = y$ ή $w = z$. Αυτό αποδεικνύει ότι το w είναι ακραίο σημείο του A, και συμπληρώνει την απόδειξη στην περίπτωση που το x είναι στο σχετικό σύνορο του A.

(β) Το x ανήκει στο σχετικό εσωτερικό του A στην $\mathrm{aff}(A)$. Παίρνουμε τυχόν $y \in A$ με $y \neq x$ και θεωρούμε την ευθεία που ορίζουν τα x και y. Αφού το A είναι συμπαγές, η τομή αυτής της ευθείας με το A είναι ένα ευθύγραμμο τμήμα του οποίου τα άκρα x_1, x_2 ανήκουν στο σχετικό σύνορο του A. Από το (α), καθένα από τα x_1, x_2 είναι κυρτός συνδυασμός ακραίων σημείων του A. Το x είναι κυρτός συνδυασμός των x_1 και x_2, άρα είναι κι αυτό κυρτός συνδυασμός ακραίων σημείων του A. $\square$

2.4 Πολυεδρικοί κώνοι - το λήμμα του Farkas

§1. **Ορισμός** Ένα υποσύνολο K του $\mathbb{R}^n$ λέγεται *κώνος* αν για κάθε $x \in K$ και κάθε $\lambda > 0$ έχουμε $\lambda x \in K$. Το σημείο 0, το οποίο λέγεται *κορυφή* του κώνου, μπορεί να ανήκει ή να μην ανήκει σε αυτόν (αν στον ορισμό βάζαμε $\lambda \geq 0$ θα απαιτούσαμε η κορυφή να συμπεριλαμβάνεται στον κώνο).

Δεν είναι δύσκολο να διαπιστώσετε ότι το $K \subseteq \mathbb{R}^n$ είναι κυρτός κώνος αν και μόνο αν για κάθε $x, y \in K$ και κάθε $\lambda > 0$ ισχύει ότι $x + y \in K$ και $\lambda x \in K$.

Ορισμός Έστω K ένας κώνος στον $\mathbb{R}^n$. Το *πολικό* του K είναι το σύνολο

$$K^\circ := \{y \in \mathbb{R}^n : \forall x \in K \ \ \langle x, y \rangle \leq 0\}.$$

Το *διπολικό* του K είναι το σύνολο

$$K^{\circ\circ} = (K^\circ)^\circ = \{x \in \mathbb{R}^n : \forall y \in K^\circ \ \ \langle x, y \rangle \leq 0\}.$$

Παρατηρήσεις Εύκολα ελέγχουμε τα εξής:

(α) Τα K° και $K^{\circ\circ}$ είναι κυρτοί κώνοι και περιέχουν το 0.

(β) $K \subseteq K^{\circ\circ}$.

(γ) Το πολικό K° του K είναι κλειστό σύνολο: πράγματι, αν $y_m \in K^\circ$ και $y_m \to y$, τότε για κάθε $x \in K$ έχουμε

$$\langle x, y \rangle = \lim_{m \to \infty} \langle x, y_m \rangle \leq 0$$

από τη συνέχεια του εσωτερικού γινομένου και το γεγονός ότι $\langle x, y_m \rangle \leq 0$ για κάθε m. Άρα, $y \in K^\circ$.

(δ) Αν K_1, K_2 είναι δύο κώνοι και $K_1 \subseteq K_2$, τότε $K_2^\circ \subseteq K_1^\circ$. Αφού $K \subseteq K^{\circ\circ}$, παίρνουμε $K^{\circ\circ\circ} \subseteq K^\circ$. Από την άλλη πλευρά, $K^\circ \subseteq (K^\circ)^{\circ\circ} = K^{\circ\circ\circ}$ λόγω του (β). Άρα, $K^\circ = K^{\circ\circ\circ}$ για κάθε κώνο K.

(ε) Αφού ο $K^{\circ\circ}$ είναι κλειστό σύνολο και περιέχει το K, βλέπουμε ότι $\overline{K} \subseteq K^{\circ\circ}$. Στην περίπτωση που ο αρχικός κώνος K είναι κυρτός, έχουμε ισότητα:

Θεώρημα 2.12 *Αν K είναι ένας μη κενός κυρτός κώνος, τότε $K^{\circ\circ} = \overline{K}$.*

Απόδειξη: Είδαμε ότι $\overline{K} \subseteq K^{\circ\circ}$. Έστω ότι υπάρχει $y \in K^{\circ\circ} \setminus \overline{K}$. Το $\overline{K}$ είναι κλειστό κυρτό σύνολο, άρα διαχωρίζεται γνήσια από το $\{y\}$. Δηλαδή, υπάρχουν $u \in \mathbb{R}^n$ και $t \in \mathbb{R}$ με την ιδιότητα

$$\sup_{x \in \overline{K}} \langle u, x \rangle = t < \langle u, y \rangle.$$

Παρατηρήστε ότι δεν μπορεί να υπάρχει $x_0 \in \overline{K}$ με $\langle u, x_0 \rangle > 0$, γιατί τότε θα είχαμε $\sup_{x \in \overline{K}} \langle u, x \rangle = +\infty$. (Το $\overline{K}$ είναι κώνος οπότε θα είχαμε $\sup_{x \in \overline{K}} \langle u, x \rangle \geq \lambda \langle u, x_0 \rangle$ για κάθε $\lambda > 0$.)

Άρα $t \leq 0$ (για την ακρίβεια, $t = 0$). Έπεται ότι $u \in (\overline{K})^\circ \subseteq K^\circ$, αλλά $\langle u, y \rangle > 0$ ενώ $y \in K^{\circ\circ}$, το οποίο είναι άτοπο. $\square$

§2. Αν $x = (\xi_1, \ldots, \xi_s) \in \mathbb{R}^s$ και $\xi_i \geq 0$ για κάθε $i = 1, \ldots, s$, θα γράφουμε $x \geq 0$. Με P^s συμβολίζουμε το σύνολο $\{x \in \mathbb{R}^s : x \geq 0\}$ όλων των $x \in \mathbb{R}^s$ που έχουν μη αρνητικές συντεταγμένες.

Θυμηθείτε ότι οι *κλειστοί ημίχωροι* του $\mathbb{R}^n$ είναι τα σύνολα της μορφής

$$\{x \in \mathbb{R}^n : \langle x, u \rangle \leq t\}$$

όπου $u \in \mathbb{R}^n$ και $t \in \mathbb{R}$. Αν $t = 0$, τότε λέμε ότι ο κλειστός ημίχωρος έχει *συνοριακό σημείο το 0.*

Ορισμός *Πολυεδρικός κώνος* στον $\mathbb{R}^n$ είναι μία τομή πεπερασμένων το πλήθος κλειστών ημιχώρων του $\mathbb{R}^n$ που έχουν συνοριακό σημείο το 0.

Κάθε πολυεδρικός κώνος στον $\mathbb{R}^n$ εκφράζεται σαν το σύνολο των λύσεων ενός πεπερασμένου συστήματος ανισοτήτων

$$\langle x, u_i \rangle \leq 0, \quad 1 \leq i \leq k$$

όπου $u_1, \ldots, u_k \in \mathbb{R}^n$. Αν ορίσουμε την απεικόνιση $T : \mathbb{R}^n \to \mathbb{R}^k$ με

$$T(x) = (\langle x, u_1 \rangle, \ldots, \langle x, u_k \rangle),$$

τότε

$$K = \{x \in \mathbb{R}^n : T(x) \leq 0\} = -T^{-1}(P^k).$$

Ορισμός Ένας κυρτός κώνος K στον $\mathbb{R}^n$ λέγεται *πεπερασμένα παραγόμενος* αν υπάρχει ένα πεπερασμένο σύνολο $\{u_1, \ldots, u_k\} \subset \mathbb{R}^n$ με την ιδιότητα

$$K = \{t_1 u_1 + \ldots + t_k u_k : t_1, \ldots, t_k \geq 0\}.$$

Τα διανύσματα $u_1, \ldots, u_k$ λέγονται *γεννήτορες* του K. Παρατηρήστε ότι το σύνολο όλων των μη αρνητικών γραμμικών συνδυασμών πεπερασμένων το πλήθος διανυσμάτων είναι πάντα κυρτός κώνος, άρα ο ορισμός μας είναι συνεπής.

Ισοδύναμα, οι πεπερασμένα παραγόμενοι κυρτοί κώνοι είναι ακριβώς τα σύνολα της μορφής

$$K = \{T(x) : x \geq 0\} = T(P^k)$$

όπου $k \in \mathbb{N}$ και $T : \mathbb{R}^k \to \mathbb{R}^n$ μία γραμμική απεικόνιση.

Θεώρημα 2.13 *Το πολικό ενός πεπερασμένα παραγόμενου κυρτού κώνου στον $\mathbb{R}^n$ είναι πολυεδρικός κώνος.*

Απόδειξη: Έστω $K = S(P^k)$ ένας πεπερασμένα παραγόμενος κυρτός κώνος, όπου $S(t_1, \ldots, t_k) = t_1 u_1 + \ldots + t_k u_k$ για κάποια $u_1, \ldots, u_k \in \mathbb{R}^n$.

Τότε,

$$y \in K^\circ \Longleftrightarrow \sum_{i=1}^{k} t_i \langle u_i, y \rangle \leq 0$$

για κάθε $t_1, \ldots, t_k \geq 0$.

Επιλέγοντας $t_i = 1$ και $t_j = 0$ αν $j \neq i$, βλέπουμε ότι $\langle u_i, y\rangle \leq 0$ για κάθε $i = 1, \ldots, k$. Αντίστροφα, αν $\langle u_i, y\rangle \leq 0$ για κάθε i, τότε $\sum t_i \langle u_i, y\rangle \leq 0$ για κάθε $t_1, \ldots, t_k \geq 0$. Δηλαδή,

$$y \in K^\circ \iff \langle u_i, y\rangle \leq 0, \quad i = 1, \ldots, k.$$

Αυτό αποδεικνύει ότι το K° είναι πολυεδρικός κώνος. □

Το βασικό αποτέλεσμα αυτής της παραγράφου μάς λέει ότι η κλάση των πολυεδρικών κώνων συμπίπτει με την κλάση των πεπερασμένα παραγόμενων κυρτών κώνων. Για την απόδειξη θα χρειαστούμε δύο προτάσεις.

Πρόταση 2.2 *Έστω $T : \mathbb{R}^n \to \mathbb{R}^k$ μία γραμμική απεικόνιση. Θεωρούμε $b \in \mathbb{R}^k$ και ορίζουμε $A = \{x \in \mathbb{R}^n : T(x) \leq b\}$. Τότε,*

(α) το A έχει το πολύ πεπερασμένα το πλήθος ακραία σημεία.

(β) αν το A είναι φραγμένο και μη κενό, τότε το A είναι η κυρτή θήκη πεπερασμένων το πλήθος σημείων. (είναι «πολύτοπο».)

Απόδειξη: (α) Έστω $b = (b_1, \ldots, b_k)$. Τότε, $x \in A$ αν και μόνο αν

$$\langle x, u_i\rangle \leq b_i, \quad i = 1, \ldots, k$$

όπου u_i τα διανύσματα γραμμές του T (δηλαδή, $u_i = T^t(e_i)$). Έστω E το σύνολο των ακραίων σημείων του A. Για κάθε $x \in E$ ορίζουμε

$$I(x) = \{i \leq k : \langle x, u_i\rangle < b_i\}.$$

Θα δείξουμε ότι αν $x \neq y$ στο E, τότε $I(x) \neq I(y)$. Αυτό αποδεικνύει το ζητούμενο, γιατί τα υποσύνολα του $\{1, \ldots, k\}$ είναι πεπερασμένα το πλήθος.

Έστω ότι $I(x) = I(y)$. Τότε, υπάρχει $\epsilon > 0$ τέτοιο ώστε

$$(*) \qquad \langle x + \epsilon(x - y), u_i\rangle \leq b_i$$

για κάθε $i \in I(x)$.

Αν πάλι $i \notin I(x)$, τότε $\langle x, u_i\rangle = \langle y, u_i\rangle = b_i$, επομένως $\langle x + \epsilon(x - y), u_i\rangle = b_i$. Άρα, η $(*)$ ισχύει για κάθε $i \leq k$. Δηλαδή, $x + \epsilon(x - y) \in A$. Όμως,

$$x = \frac{\epsilon}{1+\epsilon} y + \frac{1}{1+\epsilon}[x + \epsilon(x - y)],$$

και αφού το x είναι ακραίο σημείο του A, υποχρεωτικά έχουμε $x = y$.

(β) Το A είναι κλειστό και κυρτό. Αν είναι φραγμένο, τότε είναι συμπαγές κυρτό σύνολο. Άρα, το A είναι η κυρτή θήκη των ακραίων σημείων του, τα οποία είναι πεπερασμένα το πλήθος από το (α). □

Πρόταση 2.3 *Έστω K ένας πεπερασμένα παραγόμενος κυρτός κώνος στον $\mathbb{R}^n$. Τότε,*

(α) το K είναι η ένωση πεπερασμένων το πλήθος πεπερασμένα παραγόμενων κυρτών κώνων, καθένας από τους οποίους έχει ένα γραμμικά ανεξάρτητο σύνολο γεννητόρων.

(β) το K είναι κλειστό.

Απόδειξη: (α) Η απόδειξη θυμίζει την απόδειξη του Θεωρήματος του Καραθεοδωρή. Υπάρχουν $u_1, \ldots, u_k \in \mathbb{R}^n$ τέτοια ώστε

$$K = \{t_1u_1 + \ldots + t_ku_k : t_1, \ldots, t_k \geq 0\}.$$

Έστω $0 \neq x \in K$. Τότε, $x = s_1w_1 + \ldots + s_mw_m$ με $s_i > 0$, όπου $w_1, \ldots, w_m$ κάποια από τα u_i. Αν τα $w_1, \ldots, w_m$ είναι γραμμικά εξαρτημένα, υπάρχουν $r_1, \ldots, r_m$ όχι όλα μηδέν, τέτοια ώστε

$$r_1w_1 + \ldots + r_mw_m = 0.$$

Τότε, για κάθε $\rho \in \mathbb{R}$,

$$x = (s_1 - \rho r_1)w_1 + \ldots + (s_m - \rho r_m)w_m.$$

Μπορούμε να βρούμε ρ_0 για το οποίο $s_i - \rho_0 r_i \geq 0$ για κάθε $i \leq m$ και τουλάχιστον ένα $s_i - \rho_0 r_i = 0$. Δηλαδή, μπορούμε να γράψουμε το x σαν θετικό γραμμικό συνδυασμό λιγότερων από m εκ των w_i. Η διαδικασία αυτή μπορεί να συνεχιστεί έως ότου γράψουμε το x σαν θετικό γραμμικό συνδυασμό κάποιων γραμμικά ανεξάρτητων w_i. Το $\{u_1, \ldots, u_k\}$ έχει πεπερασμένα το πλήθος διαφορετικά υποσύνολα, οπότε το K περιέχεται στην ένωση των πεπερασμένων το πλήθος κυρτών κώνων που παράγουν τα γραμμικά ανεξάρτητα υποσύνολα του $\{u_1, \ldots, u_k\}$. Από την άλλη πλευρά, κάθε τέτοιος κώνος προφανώς περιέχεται στο K, άρα έχουμε ισότητα.

(β) Κάθε κυρτός κώνος που παράγεται από γραμμικά ανεξάρτητα διανύσματα είναι κλειστός (Άσκηση 32), άρα το ζητούμενο είναι συνέπεια του (α): πεπερασμένη ένωση κλειστών συνόλων είναι κλειστό σύνολο. □

Θεώρημα 2.14 *Έστω $K \subseteq \mathbb{R}^n$. Το K είναι πολυεδρικός κώνος αν και μόνο αν είναι πεπερασμένα παραγόμενος κυρτός κώνος.*

Απόδειξη: Έστω πρώτα $K = \{x : T(x) \leq 0\}$ ένας πολυεδρικός κώνος, όπου $T : \mathbb{R}^n \to \mathbb{R}^k$ γραμμική απεικόνιση. Τότε,

$$K = T^{-1}(T(\mathbb{R}^n) \cap (-P^k)).$$

Ορίζουμε $A = T(\mathbb{R}^n) \cap (-P^k)$ και $B = \{y = (\eta_1, \ldots, \eta_k) \in A : \sum_{i=1}^k \eta_i \geq -1\}$. Το B είναι μη κενό γιατί $0 \in B$, και

$$A = \{sy : s \geq 0, y \in B\}.$$

Το B μπορεί να περιγραφεί σαν το σύνολο των λύσεων του συστήματος ανισοτήτων

$$\begin{cases} \langle y, u_i \rangle \leq 0 & , i \leq s \\ \langle y, -u_i \rangle \leq 0 & , i \leq s \\ \langle y, e_i \rangle \leq 0 & , i \leq k \\ \langle y, -e \rangle \leq 1 & \end{cases}$$

όπου $u_1, \dots, u_s$ βάση του κάθετου υποχώρου του $T(\mathbb{R}^n)$ στον $\mathbb{R}^k$, $e_1, \dots, e_k$ η συνήθης ορθοκανονική βάση του $\mathbb{R}^k$, και $e = (1, 1, \dots, 1)$. Εύκολα ελέγχουμε ότι το B είναι φραγμένο, άρα από την Πρόταση 2.2 υπάρχουν $y_1, \dots, y_m \in \mathbb{R}^k$ τέτοια ώστε

$$B = \mathrm{co}\{y_1, \dots, y_m\}.$$

Έπεται ότι το A είναι πεπερασμένα παραγόμενος κώνος με γεννήτορες τα $y_1, \dots, y_m$.

Βρίσκουμε $x_1, \dots, x_m \in \mathbb{R}^n$ με $T(x_i) = y_i$, $i = 1, \dots, m$. Τότε, για κάθε $x \in K$ υπάρχουν $t_1, \dots, t_m \geq 0$ τέτοια ώστε

$$T(x) = \sum_{i=1}^{m} t_i y_i = \sum_{i=1}^{m} t_i T(x_i),$$

άρα

$$x - \sum_{i=1}^{m} t_i x_i \in T^{-1}(\{0\}).$$

Έπεται ότι ο K είναι πεπερασμένα παραγόμενος κώνος με γεννήτορες τα $x_1, \dots, x_m$ και $\pm z_1, \dots, \pm z_r$ όπου $\{z_1, \dots, z_r\}$ βάση του $T^{-1}(\{0\})$.

Αντίστροφα, έστω K ένας πεπερασμένα παραγόμενος κυρτός κώνος. Από την Πρόταση 2.3 ο K είναι κλειστός, άρα $K = (K^\circ)^\circ$. Από το Θεώρημα 2.13, το K° είναι πολυεδρικός κώνος, επομένως το πρώτο μέρος της απόδειξης μάς λέει ότι το K° είναι πεπερασμένα παραγόμενος κώνος. Χρησιμοποιώντας για μία ακόμα φορά το Θεώρημα 2.13, βλέπουμε ότι το $K = (K^\circ)^\circ$ είναι πολυεδρικός κώνος. □

Παρατήρηση Η απόδειξη δείχνει ότι οι κώνοι $\{T(x) : x \geq 0\}$ και $\{y : T^t(y) \leq 0\}$ είναι ο ένας πολικό του άλλου.

Μία εφαρμογή του Θεωρήματος 2.14 είναι το *λήμμα του Farkas*: έστω $A : \mathbb{R}^n \to \mathbb{R}^k$ μία γραμμική απεικόνιση. Ορίζουμε

$$K = \{x \in \mathbb{R}^k : A^t(x) \leq 0\}$$

όπου A^t ο ανάστροφος πίνακας του A. Τότε, $b \in K^\circ$ αν και μόνο αν για κάθε $x \in \mathbb{R}^k$ ισχύει η συνεπαγωγή

$$A^t(x) \leq 0 \Longrightarrow \langle b, x \rangle \leq 0.$$

Από την παρατήρηση μετά το Θεώρημα 2.14 βλέπουμε ότι $K^\circ = \{A(y) : y \geq 0\}$, άρα έχουμε αποδείξει το εξής:

Λήμμα του Farkas *Έστω $A : \mathbb{R}^n \to \mathbb{R}^k$ μία γραμμική απεικόνιση. Για κάθε $b \in \mathbb{R}^k$, υπάρχει $y \in \mathbb{R}^n$, $y \geq 0$ τέτοιο ώστε $A(y) = b$ αν και μόνο αν ισχύει η συνεπαγωγή*

$$A^t(x) \leq 0 \Longrightarrow \langle b, x \rangle \leq 0.$$

□

Το λήμμα του Farkas διατυπώνεται συχνά στην ακόλουθη μορφή:

Θεώρημα 2.15 *Για κάθε γραμμική απεικόνιση $A : \mathbb{R}^n \to \mathbb{R}^k$ και κάθε $b \in \mathbb{R}^k$, είτε* *(α) το σύστημα $A^t(x) \leq 0$ και $\langle b, x\rangle > 0$ έχει λύση, ή*

(β) η εξίσωση $b = A(y)$ έχει λύση $y \geq 0$,

αλλά τα (α) και (β) δεν μπορούν να συμβούν ταυτόχρονα. $\square$

2.5 Ασκήσεις

1. Δείξτε ότι ο υπόχωρος W του Θεωρήματος 2.1 ορίζεται μονοσήμαντα από το A.

2. Δείξτε ότι τα $x_0, x_1, \ldots, x_m$ είναι αφφινικά ανεξάρτητα αν και μόνο αν οποτεδήποτε $\lambda_0 x_0 + \lambda_1 x_1 + \ldots + \lambda_m x_m = 0$ και $\lambda_0 + \ldots + \lambda_m = 0$, έπεται ότι $\lambda_0 = \ldots = \lambda_m = 0$.

3. Για κάθε $x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ συμβολίζουμε με x^* το σημείο $(\xi_1, \ldots, \xi_n, 1)$ του $\mathbb{R}^{n+1}$. Δείξτε ότι τα $x_1, \ldots, x_m$ είναι αφφινικά ανεξάρτητα στον $\mathbb{R}^n$ αν και μόνο αν τα $x_1^*, \ldots, x_m^*$ είναι γραμμικά ανεξάρτητα στον $\mathbb{R}^{n+1}$.

4. Έστω $x, y \in \mathbb{R}^n$ και $\lambda, \mu > 0$ με $\lambda + \mu = 1$. Δείξτε ότι

$$\lambda\|x\|^2 + \mu\|y\|^2 = \|\lambda x + \mu y\|^2 + \lambda\mu\|x - y\|^2.$$

Συμπεράνατε ότι

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Δείξτε επίσης ότι αν $\|x\| = \|y\| = 1$ και $x \neq y$, τότε $\|\lambda x + \mu y\| < 1$.

5. Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$. Δείξτε ότι το σύνολο

$$A_\epsilon = \{x \in \mathbb{R}^n : d(x, A) \leq \epsilon\}$$

είναι κυρτό για κάθε $\epsilon > 0$.

6. Έστω A κλειστό υποσύνολο του $\mathbb{R}^n$. Αν το A δεν είναι κυρτό, δείξτε ότι υπάρχουν $x, y \in A$ τέτοια ώστε $A \cap (x, y) = \emptyset$, όπου (x, y) το ανοικτό ευθύγραμμο τμήμα με άκρα τα x και y.

7. Έστω A κυρτό σύνολο στον $\mathbb{R}^n$. Δείξτε ότι το A είναι κλειστό αν και μόνο αν το $A \cap \ell$ είναι κλειστό για κάθε ευθεία ℓ στον $\mathbb{R}^n$.

8. Έστω A και B κυρτά υποσύνολα του $\mathbb{R}^n$. Δείξτε ότι τα $A + B$ και tA, $t \in \mathbb{R}$ είναι κυρτά σύνολα.

Δείξτε ότι $(t + s)A = tA + sA$ για κάθε $t, s \geq 0$.

9. Έστω $\{A_m\}$ ακολουθία κυρτών υποσυνόλων του $\mathbb{R}^n$.

(α) Αν $A_m \subseteq A_{m+1}$, $m = 1, 2, \ldots$, τότε το $\bigcup_{m=1}^{\infty} A_m$ είναι κυρτό.

(β) Το $\liminf A_m = \bigcup_{k=1}^{\infty} \bigcap_{m=k}^{\infty} A_m$ είναι πάντα κυρτό.

10. (α) Δείξτε με ένα παράδειγμα ότι το άθροισμα δύο κλειστών συνόλων δεν είναι πάντα κλειστό σύνολο.

(β) Έστω A και B υποσύνολα του $\mathbb{R}^n$. Αν το A είναι κλειστό και το B συμπαγές, τότε το $A + B$ είναι κλειστό.

(γ) Έστω A και B υποσύνολα του $\mathbb{R}^n$. Αν το A είναι ανοικτό, τότε το $A+B$ είναι ανοικτό.

11. (Radon) Έστω $m \geq n+2$ και $x_1, \ldots, x_m \in \mathbb{R}^n$. Δείξτε ότι υπάρχει διαμέριση του $\{1, \ldots, m\}$ σε δύο ξένα υποσύνολα I και J, τέτοια ώστε

$$\mathrm{co}\{x_i : i \in I\} \cap \mathrm{co}\{x_j : j \in J\} \neq \emptyset.$$

12. Έστω A μη κενό φραγμένο υποσύνολο του $\mathbb{R}^n$. Δείξτε ότι τα A και $\mathrm{co}(A)$ έχουν την ίδια διάμετρο.

13. (Gauss-Lucas) Έστω $P(z) = a_n z^n + \ldots + a_1 z + a_0$, $a_n \neq 0$, μη σταθερό μιγαδικό πολυώνυμο. Δείξτε ότι οι ρίζες της παραγώγου P' του P ανήκουν στην κυρτή θήκη του συνόλου των ριζών του P.

14. Έστω A, B υποσύνολα του $\mathbb{R}^n$. Δείξτε ότι

$$\mathrm{co}(A+B) = \mathrm{co}(A) + \mathrm{co}(B).$$

15. Έστω $A \subset \mathbb{R}^n$. Δείξτε ότι $\mathrm{co}(\mathrm{int}(A)) \subseteq \mathrm{int}(\mathrm{co}(A))$. Ισχύει πάντα ισότητα;

16. Δείξτε ότι κάθε συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$ γράφεται σαν τομή από κλειστές μπάλες του $\mathbb{R}^n$.

κλειστό

17. Η κυρτή θήκη ενός κλειστού συνόλου δεν είναι πάντα [redacted] σύνολο, ακόμα και στον $\mathbb{R}^2$. Πάρτε για παράδειγμα την κυρτή θήκη του $U = \{(t,s) : t^2 s^2 = 1, s > 0\}$. Τι ισχύει στο $\mathbb{R}$;

18. Για μία διαφορετική απόδειξη του δεύτερου μέρους του Θεωρήματος 2.6, θεωρήστε το

$$U = \{(\lambda_1, \ldots, \lambda_{n+1}) \in \mathbb{R}^{n+1} : \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1\}$$

και την απεικόνιση $f : A \times \ldots \times A \times U \to \mathbb{R}^n$ ($(n+1)$-αντίτυπα του A) με

$$f(x_1, \ldots, x_{n+1}, \lambda_1, \ldots, \lambda_{n+1}) = \sum_{i=1}^{n+1} \lambda_i x_i.$$

Δείξτε πρώτα ότι το πεδίο ορισμού είναι συμπαγές και η f συνεχής. Κατόπιν, χρησιμοποιήστε το Θεώρημα του Καραθεοδωρή.

19. Έστω $A \neq \emptyset$. Ορίζουμε $d(\cdot, A) : \mathbb{R}^n \to \mathbb{R}^+$ με

$$x \mapsto d(x, A) := \inf\{\|x - a\| : a \in A\}.$$

Δείξτε ότι $|d(x_1, A) - d(x_2, A)| \leq \|x_1 - x_2\|$ για κάθε $x_1, x_2 \in \mathbb{R}^n$, άρα η $d(\cdot, A)$ είναι ομοιόμορφα συνεχής.

20 Αν $A \subseteq \mathbb{R}^n$ κλειστό και $x \notin A$, δείξτε ότι $d(x, A) > 0$.

21. Δώστε διαφορετική απόδειξη του Θεωρήματος 2.10 ακολουθώντας τα εξής βήματα:

(α) Έστω ότι $x_0 = 0$ και $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$. Υπάρχει $y_0 \in S$ το οποίο είναι το πιό απομακρυσμένο από το A.

(β) Το x_0 είναι το πλησιέστερο σημείο του A προς το y_0.

(γ) Το υπερεπίπεδο $\{z : \langle y_0 - x_0, z - x_0 \rangle = 0\}$ στηρίζει το A στο x_0.

22. Δείξτε ότι τα $A = \{(0, y, 1) : y \in \mathbb{R}\}$ και $B = \{(x, y, z) : x, y, z \geq 0, xy \geq z^2\}$ είναι ξένα κλειστά υποσύνολα του $\mathbb{R}^3$. Δείξτε ότι υπάρχει μοναδικό υπερεπίπεδο που διαχωρίζει τα A και B, και ότι αυτό το υπερεπίπεδο έχει μη κενή τομή τόσο με το A όσο και με το B.

23. Δείξτε ότι το Θεώρημα 2.10 ισχύει και χωρίς την υπόθεση $\mathrm{int}(A) \neq \emptyset$. Το Θεώρημα 2.9 χρειάζεται την υπόθεση (για παράδειγμα, πάρτε B ανοικτό και $A = \{x_0\} \subseteq B$).

24. Αν το υπερεπίπεδο H στηρίζει το κυρτό $A \subset \mathbb{R}^n$ στο $x_0 \in \mathrm{bd}(A)$, τότε το H δεν περιέχει εσωτερικά σημεία του A.

25. Αν τα A, B είναι ξένα κυρτά υποσύνολα του $\mathbb{R}^n$, το A είναι κλειστό και $\mathbb{R}^n = A \cup B$, τότε το A είναι ημίχωρος.

26. Αν το κλειστό σύνολο $A \subset \mathbb{R}^n$ έχει μη κενό εσωτερικό και από κάθε σημείο του συνόρου του περνάει κάποιο υπερεπίπεδο στήριξης, τότε το A είναι κυρτό.

27. Κάθε κλειστό κυρτό σύνολο A στον $\mathbb{R}^n$ είναι η τομή των ημιχώρων που το περιέχουν.

28. Έστω A υποσύνολο του $\mathbb{R}^n$. Δείξτε ότι τα ακραία σημεία του $\mathrm{co}(A)$ είναι ακριβώς τα σημεία $a \in A$ για τα οποία $a \notin \mathrm{co}(A \setminus \{a\})$.

29. (α) Δείξτε ότι το σύνολο των ακραίων σημείων ενός κλειστού κυρτού υποσυνόλου του $\mathbb{R}^2$ είναι κλειστό.

(β) Δώστε παράδειγμα συμπαγούς κυρτού συνόλου στον $\mathbb{R}^3$ του οποίου το σύνολο των ακραίων σημείων δεν είναι κλειστό.

30. Έστω A κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$ και H υπερεπίπεδο που στηρίζει το A στο x_0. Αν το x είναι ακραίο σημείο του $A \cap H$, τότε το x είναι ακραίο σημείο του A.

31. Έστω A κυρτό σύνολο στον $\mathbb{R}^n$ και $x_0 \in \mathrm{bd}(A)$ με $x_0 \in A$. Λέμε ότι το x_0 είναι *εκτεθειμένο* σημείο του A αν υπάρχει υπερεπίπεδο στήριξης H του A στο x_0 με $H \cap A = \{x_0\}$. Δείξτε ότι κάθε εκτεθειμένο σημείο του A είναι ακραίο σημείο του A, αλλά το αντίστροφο δεν ισχύει ακόμα και στον $\mathbb{R}^2$.

32. Δείξτε ότι κάθε κυρτός κώνος που παράγεται από γραμμικά ανεξάρτητα διανύσματα είναι κλειστός .

33. Αποδείξτε το Λήμμα του Farkas εφαρμόζοντας διαχωριστικό θεώρημα στα $\{b\}$ και $T(P^n)$.

34. (Gordan) Για κάθε γραμμική απεικόνιση $T : \mathbb{R}^n \to \mathbb{R}^k$ συμβαίνει ακριβώς ένα από τα εξής:

(α) η $T^t(x) > 0$ έχει λύση $x \in \mathbb{R}^k$,

(β) η $T(y) = 0$, $y \geq 0$, $y \neq 0$ έχει λύση $y \in \mathbb{R}^n$.

Κεφάλαιο 3

Κυρτές συναρτήσεις στον $\mathbb{R}^n$

3.1 Ορισμοί

§1. Έστω $A \subseteq \mathbb{R}^n$ και $f : A \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$. Το σύνολο

$$\{(x,t) : x \in A, t \in \mathbb{R}, t \geq f(x)\}$$

λέγεται *επιγράφημα* της f και συμβολίζεται με $\mathrm{epi}(f)$.

Μία συνάρτηση $f : A \to \overline{\mathbb{R}}$ λέγεται *κυρτή* αν το επιγράφημα $\mathrm{epi}(f)$ της f είναι κυρτό υποσύνολο του $\mathbb{R}^{n+1}$. Η $g : A \to \overline{\mathbb{R}}$ λέγεται *κοίλη* αν η $-g$ είναι κυρτή. Μία *αφφινική* συνάρτηση στο A είναι μία συνάρτηση που είναι ταυτόχρονα κυρτή και κοίλη και παίρνει πεπερασμένες τιμές στο A.

Παρατήρηση Το $E = \mathrm{epi}(f)$ έχει τις εξής δύο ιδιότητες:

(α) Αν $(x,t) \in E$ και $s > t$, τότε $(x,s) \in E$ (φανερό, αφού $f(x) \leq t < s$).

(β) Αν $\lambda = \inf\{t \in \mathbb{R} : (x,t) \in E\} \in \mathbb{R}$, τότε $(x,\lambda) \in E$ (αφού $\lambda \in \mathbb{R}$, υπάρχουν $t \in \mathbb{R}$ με $f(x) \leq t$ και το ελάχιστο τέτοιο t είναι το $f(x)$).

Αντίστροφα, έστω $E \subseteq \mathbb{R}^{n+1}$ με τις ιδιότητες (α) και (β). Ορίζουμε

$$(*) \qquad f(x) = \inf\{t \in \mathbb{R} : (x,t) \in E\}, \quad x \in \mathbb{R}^n$$

όπου $\inf(\emptyset) = +\infty$. Τότε, η f παίρνει τιμές στο $\overline{\mathbb{R}}$ και χρησιμοποιώντας μόνο την ιδιότητα (α) του E βλέπουμε ότι

$$\{(x,t) \in \mathbb{R}^{n+1} : t > f(x)\} \subseteq E \subseteq \mathrm{epi}(f).$$

Αν τώρα υποθέσουμε επιπλέον ότι το E είναι κυρτό, τότε η συνάρτηση f που ορίζεται από την $(*)$ είναι κυρτή. Πράγματι, έστω $(x,t) \in \mathrm{epi}(f)$, $(y,s) \in \mathrm{epi}(f)$ και $\lambda \in (0,1)$. Πρέπει να δείξουμε ότι $(\lambda x + (1-\lambda)y, \lambda t + (1-\lambda)s) \in \mathrm{epi}(f)$.

Αν $t > f(x)$ και $s > f(y)$, τότε $(x,t),(y,s) \in E$ και αφού το E είναι κυρτό παίρνουμε

$$(\lambda x + (1-\lambda)y, \lambda t + (1-\lambda)s) \in E \subseteq \mathrm{epi}(f).$$

Στη γενική περίπτωση ($t \geq f(x)$, $s \geq f(y)$), θεωρούμε $t_n > t$ και $s_n > s$ με $t_n \to t$ και $s_n \to s$. Τότε, $(\lambda x+(1-\lambda)y, \lambda t_n+(1-\lambda)s_n) \in \mathrm{epi}(f)$, δηλαδή $f(\lambda x+(1-\lambda)y) \leq \lambda t_n+(1-\lambda)s_n$, και αφήνοντας το $n \to \infty$ παίρνουμε $f(\lambda x+(1-\lambda)y) \leq \lambda t+(1-\lambda)s$.

Το *ουσιώδες πεδίο ορισμού* μιάς κυρτής συνάρτησης στο A είναι η προβολή του επιγραφήματος της f στον $\mathbb{R}^n$

$$\mathrm{dom}(f) = \{x \in \mathbb{R}^n : \exists t : (x,t) \in \mathrm{epi}(f)\} = \{x : f(x) < +\infty\}.$$

Το $\mathrm{dom}(f)$ είναι κυρτό υποσύνολο του $\mathbb{R}^n$, γιατί είναι γραμμική εικόνα του κυρτού συνόλου $\mathrm{epi}(f)$. Η *διάσταση* της f είναι η διάσταση του $\mathrm{dom}(f)$. Είναι φανερό ότι η κυρτότητα της f είναι ισοδύναμη με την κυρτότητα του περιορισμού της στο $\mathrm{dom}(f)$. Όλη η πληροφορία βρίσκεται στον περιορισμό της f στο $\mathrm{dom}(f)$ και το αρχικό πεδίο ορισμού A δεν παίζει κανένα ρόλο από μόνο του.

Υπάρχουν σοβαροί λόγοι για τους οποίους δεν είναι αρκετό να μελετήσουμε την κλάση όλων των κυρτών συναρτήσεων που έχουν κάποιο σταθερό κυρτό σύνολο σαν το κοινό ουσιώδες πεδίο ορισμού τους. Ονομάζουμε λοιπόν κυρτή συνάρτηση μία κυρτή συνάρτηση που ορίζεται σε ολόκληρο τον $\mathbb{R}^n$ αλλά επιτρέπεται να απειρίζεται. Όταν μία κυρτή συνάρτηση κατασκευάζεται σύμφωνα με κάποιους τύπους, οι τύποι αυτοί θα προσδιορίζουν το ουσιώδες πεδίο ορισμού της, αφού θα προσδιορίζουν τα σημεία στα οποία η συνάρτηση παίρνει την τιμή $+\infty$.

Για τους υπολογισμούς στους οποίους υπεισέρχονται το $+\infty$ και το $-\infty$, θα υιοθετήσουμε τους προφανείς κανόνες: $x + (+\infty) = +\infty$ αν $x \in \mathbb{R}$, $x \cdot (-\infty) = -\infty$ αν $x > 0$ κλπ. Συμφωνούμε επίσης ότι

$$0 \cdot (+\infty) = (+\infty) \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0 = 0.$$

Η παράσταση $+\infty + (-\infty)$ είναι αόριστη.

§2. Μία κυρτή συνάρτηση f λέγεται *κανονική* αν το επιγράφημά της είναι μη κενό και δεν περιέχει κατακόρυφες ευθείες. Δηλαδή, αν υπάρχει τουλάχιστον ένα x για το οποίο $f(x) < +\infty$ και αν $f(y) > -\infty$ για κάθε y. Ισοδύναμα, η f είναι κανονική αν το $\mathrm{dom}(f)$ είναι μη κενό και ο περιορισμός της f στο $\mathrm{dom}(f)$ παίρνει πεπερασμένες τιμές.

Ένας άλλος τρόπος να σκεφτόμαστε μία κανονική κυρτή συνάρτηση είναι σαν μία συνάρτηση που προκύπτει αν πάρουμε μία πεπερασμένη κυρτή συνάρτηση που ορίζεται σε ένα μη κενό κυρτό σύνολο A και την επεκτείνουμε σε ολόκληρο τον $\mathbb{R}^n$ θέτοντας $f(x) = +\infty$ αν $x \notin A$.

Μία κυρτή συνάρτηση που δεν είναι κανονική λέγεται *μη κανονική*. Οι κανονικές κυρτές συναρτήσεις είναι το αντικείμενο της μελέτης μας, όμως μη κανονικές συναρτήσεις προκύπτουν από τις κανονικές σε πολλές φυσιολογικές καταστάσεις, οπότε είναι προτιμότερο να τις δεχτούμε στη συνέχεια. Ένα παράδειγμα μη κανονικής κυρτής συνάρτησης που δεν παίρνει μόνο τις τιμές $\pm\infty$, είναι η $f : \mathbb{R} \to \overline{\mathbb{R}}$ που ορίζεται από την

$$f(x) = \begin{cases} -\infty & , |x| < 1 \\ 0 & , |x| = 1 \\ +\infty & , |x| > 1. \end{cases}$$

§3. Έστω $f : A \to \overline{\mathbb{R}}$. Από τον ορισμό, η f είναι κυρτή στο A αν και μόνο αν το

$$(1-\lambda)(x,t) + \lambda(y,s) = ((1-\lambda)x + \lambda y, (1-\lambda)t + \lambda s)$$

ανήκει στο $\mathrm{epi}(f)$ οποτεδήποτε τα (x,t) και (y,s) ανήκουν στο $\mathrm{epi}(f)$ και $\lambda \in [0,1]$. Με άλλα λόγια, αν $x, y \in A$, $\lambda \in [0,1]$ και $f(x) \leq t$, $f(y) \leq s$, τότε

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)t + \lambda s.$$

Η συνθήκη αυτή μπορεί να διατυπωθεί με διάφορους τρόπους:

Θεώρημα 3.1 *Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ και $f : A \to (-\infty, +\infty]$. Τότε η f είναι κυρτή στο A αν και μόνο αν*

$$f((1-\lambda)x + \lambda y) < (1-\lambda)a + \lambda b$$

οποτεδήποτε $x, y \in A$, $f(x) < a$, $f(y) < b$ και $\lambda \in [0,1]$. □

Αν η f παίρνει πραγματικές τιμές, παίρνουμε έναν απλούστερο χαρακτηρισμό.

Θεώρημα 3.2 *Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ και $f : A \to \mathbb{R}$. Τότε η f είναι κυρτή στο A αν και μόνο αν*

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y)$$

για κάθε $x, y \in A$. □

Μία χρήσιμη παραλλαγή προκύπτει αν εφαρμόσουμε το Θεώρημα 2.2 στο επιγράφημα της f:

Θεώρημα 3.3 (Ανισότητα του Jensen) *Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ και $f : A \to (-\infty, +\infty]$. Τότε η f είναι κυρτή στο A αν και μόνο αν*

$$f(\lambda_1 x_1 + \ldots + \lambda_m x_m) \leq \lambda_1 f(x_1) + \ldots + \lambda_m f(x_m)$$

για κάθε $x_1, \ldots, x_m \in A$ και $\lambda_i \geq 0$ με $\sum_{i=1}^m \lambda_i = 1$. □

Οι κοίλες συναρτήσεις ικανοποιούν τις αντίστροφες ανισότητες με τις ίδιες υποθέσεις. Οι αφφινικές συναρτήσεις ικανοποιούν τις ανισότητες σαν ισότητες.

§4. Υπάρχουν πολλές χρήσιμες αντιστοιχίες μεταξύ κυρτών συνόλων και κυρτών συναρτήσεων. Η απλούστερη από αυτές αντιστοιχίζει σε κάθε υποσύνολο A του $\mathbb{R}^n$ την *δείκτρια συνάρτηση* δ_A που ορίζεται από την

$$\delta_A(x) = \begin{cases} 0 & , x \in A \\ +\infty & , x \notin A. \end{cases}$$

Το επιγράφημα της δ_A είναι ένας «ημικύλινδρος» με διατομή το A. Εύκολα ελέγχουμε ότι το A είναι κυρτό σύνολο αν και μόνο αν η δ_A είναι κυρτή στον $\mathbb{R}^n$. Οι δείκτριες συναρτήσεις παίζουν σημαντικό ρόλο στην κυρτή ανάλυση, ανάλογο με αυτόν που παίζουν οι χαρακτηριστικές συναρτήσεις συνόλων σε άλλους κλάδους της Ανάλυσης.

Ένα άλλο παράδειγμα της σχέσης μεταξύ κυρτών συνόλων και κυρτών συναρτήσεων μάς δίνει η συνάρτηση απόστασης

$$d(x,A)=\inf\{\|x-a\|\ : a\in A\},\qquad x\in\mathbb{R}^n.$$

Θεώρημα 3.4 *Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$. Η συνάρτηση απόστασης $d(x,A)$ είναι κυρτή.*

Απόδειξη: Για κάθε A ισχύει $d(x,A)\geq d(x,\overline{A})$, αφού $A\subseteq\overline{A}$. Αντίστροφα, αν $b\in\overline{A}$, τότε υπάρχουν $a_n\in A$ με $a_n\to b$. Τότε,

$$\|x-b\|=\lim_n\|x-a_n\|\geq d(x,A).$$

Επομένως, $d(x,\overline{A})\geq d(x,A)$, δηλαδή $d(x,A)=d(x,\overline{A})$. Άρα, αφού το $\overline{A}$ είναι κυρτό όταν το A είναι κυρτό, για την απόδειξη του Θεωρήματος αρκεί να περιοριστούμε στην περίπτωση κλειστού κυρτού συνόλου A.

Έστω $x,y\in\mathbb{R}^n$, $t\in[0,1]$. Από το Λήμμα που χρησιμοποιήθηκε για το Θεώρημα 2.8, υπάρχουν $a,b\in A$ με $\|x-a\|=d(x,A)$ και $\|y-b\|=d(y,A)$. Τότε $ta+(1-t)b\in A$, άρα

$$\begin{aligned} d(tx+(1-t)y,A) &\leq \|tx+(1-t)y-ta-(1-t)b\| \\ &\leq t\|x-a\|+(1-t)\|y-b\|=td(x,A)+(1-t)d(y,A). \end{aligned}$$

□

§5. Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$. Χρησιμοποιώντας τον ορισμό της κυρτής συνάρτησης μπορούμε εύκολα να δείξουμε τα εξής:

(α) Αν $f_i: A\to\mathbb{R}$ $(i\in I)$ είναι κυρτές συναρτήσεις, τότε η $f:A\to\mathbb{R}$ που ορίζεται από την

$$f(x)=\sup\{f_i(x): i\in I\}$$

είναι κυρτή συνάρτηση.

(β) Αν $f,g:A\to\mathbb{R}$ είναι κυρτές συναρτήσεις και $a,b\geq 0$, τότε η $af+bg$ είναι κυρτή συνάρτηση.

(γ) Αν $f:A\to\mathbb{R}$ είναι μία κυρτή συνάρτηση και $g:I\to\mathbb{R}$ είναι μία αύξουσα κυρτή συνάρτηση που ορίζεται σε ένα διάστημα $I\subseteq\mathbb{R}$ που περιέχει την εικόνα $f(A)$ του A μέσω της f, τότε η σύνθεση $g\circ f:A\to\mathbb{R}$ είναι κυρτή.

(δ) Έστω $f,g:\mathbb{R}^n\to\overline{\mathbb{R}}$ κυρτές συναρτήσεις, όχι ταυτοτικά ίσες με $+\infty$. Το σύνολο $\mathrm{epi}(f)+\mathrm{epi}(g)$ είναι κυρτό. Ορίζουμε την *ελαχιστική συνέλιξη* $f\Box g$ των f και g από την

$$(f\Box g)(x)=\inf\{\lambda:(x,\lambda)\in\mathrm{epi}(f)+\mathrm{epi}(g)\}.$$

Η $f\Box g : \mathbb{R}^n \to \overline{\mathbb{R}}$ είναι κυρτή συνάρτηση. Στην περίπτωση που οι f, g δεν παίρνουν την τιμή $-\infty$, η $f\Box g$ δίνεται και από την

$$(f\Box g)(x) = \inf\{f(x_1) + g(x_2) : x_1 + x_2 = x\} = \inf\{f(y) + g(x-y) : y \in \mathbb{R}^n\}.$$

Παράδειγμα: Έστω A μη κενό κυρτό υποσύνολο του $\mathbb{R}^n$. Τότε,

$$d(x, A) = \inf_{y\in A} \|x - y\| = \inf\{\|x-y\| + \delta_A(y)\} = (f\Box\delta_A)(x)$$

όπου $f(x) = \|x\|$ και δ_A η δείκτρια συνάρτηση του A. Οι f και δ_A είναι κυρτές συναρτήσεις, άρα η $d(\cdot, A)$ είναι κυρτή.

§6. **Ορισμός** Έστω A κυρτό σύνολο και $f : A \to \mathbb{R}$. Αν $x_0 \in A$ και $T : \mathbb{R}^n \to \mathbb{R}$ είναι ένας αφφινικός μετασχηματισμός, λέμε ότι ο T *στηρίζει* την f στο x_0 αν $T(x_0) = f(x_0)$ και $T(x) \leq f(x)$ για κάθε $x \in A$. Η γεωμετρική ερμηνεία του ορισμού είναι η εξής: το σύνολο

$$\{(x, T(x)) : x \in \mathbb{R}^n\}$$

είναι ένα υπερεπίπεδο στον $\mathbb{R}^{n+1}$ που περνάει από το σημείο $(x_0, f(x_0))$ και βρίσκεται κάτω από το γράφημα της f.

Θεώρημα 3.5 *Έστω A ανοικτό, κυρτό υποσύνολο του $\mathbb{R}^n$ και $f : A \to \mathbb{R}$. Τότε, η f είναι κυρτή αν και μόνο αν έχει στήριγμα σε κάθε σημείο του A.*

Απόδειξη: Έστω ότι η f είναι κυρτή. Θεωρούμε ένα $x_0 \in A$. Το $(x_0, f(x_0))$ ανήκει στο σύνορο του $\mathrm{epi}(f)$, αφού προσεγγίζεται από τα σημεία $(x_0, f(x_0) - \frac{1}{n})$.

Αφού το $\mathrm{epi}(f)$ είναι κυρτό, από το Θεώρημα 2.10 υπάρχει υπερεπίπεδο H στον $\mathbb{R}^{n+1}$, που στηρίζει το $\mathrm{epi}(f)$ στο $(x_0, f(x_0))$ (σύμφωνα με την παρατήρηση (β) που ακολουθεί το Θεώρημα 2.10, δεν χρειάζεται να ελέγξουμε αν το $\mathrm{epi}(f)$ έχει μη κενό εσωτερικό). Επομένως, υπάρχουν γραμμικό συναρτησοειδές $\phi : \mathbb{R}^n \to \mathbb{R}$, $a \in \mathbb{R}$, $\gamma \in \mathbb{R}$ έτσι ώστε $\phi(x_0) + af(x_0) = \gamma$ και $\phi(x) + at \geq \gamma$ για κάθε $(x, t) \in \mathrm{epi}(f)$. Επιπλέον, το ϕ είναι μη μηδενικό ή $a \neq 0$.

Από τα παραπάνω προκύπτει ότι $a > 0$. Κατ΄ αρχήν, αν $(x, t) \in \mathrm{epi}(f)$, $s \geq t$, ισχύει και $(x, s) \in \mathrm{epi}(f)$. Επομένως, η ανωτέρω ανισότητα συνεπάγεται ότι $a \geq 0$. Έστω ότι ισχύει $a = 0$. Τότε το ϕ έχει ελάχιστη τιμή στο A (την γ στο σημείο x_0). Στην περίπτωση αυτή όμως το ϕ δεν είναι μηδενικό. Άρα, υπάρχει $y \in \mathbb{R}^n$ με $\phi(y) > 0$. Αφού το A είναι ανοικτό, υπάρχει $\lambda > 0$ με $x_0 + \lambda y \in A$, οπότε θα ίσχυε $\phi(x_0 + \lambda y) > \phi(x_0)$.

Αφού $a > 0$, από την ανισότητα $\phi(x) + af(x) \geq \gamma$, $x \in A$, παίρνουμε

$$f(x) \geq \frac{\gamma}{a} - \frac{1}{a}\phi(x), \quad x \in A.$$

Άρα, η αφφινική συνάρτηση $T(x) = \frac{\gamma}{a} - \frac{1}{a}\phi(x)$, $x \in \mathbb{R}^n$, στηρίζει την f στο x_0.

Έστω τώρα ότι για κάθε $a \in A$ υπάρχει αφφινική συνάρτηση $T_a : \mathbb{R}^n \to \mathbb{R}$ που στηρίζει την f στο a. Τότε ισχύει $f = \sup_a T_a$. Άρα, η f είναι κυρτή. $\square$

Στην επόμενη παράγραφο θα δούμε ότι η *μοναδικότητα* του στηρίγματος μιάς κυρτής συνάρτησης σε ένα σημείο είναι ισοδύναμη με την διαφορισιμότητά της σε αυτό.

Παρατήρηση Ένα υπερεπίπεδο στον $\mathbb{R}^{n+1}$ λέγεται κατακόρυφο αν έχει τη μορφή $H = \{(x,t) : \phi(x) = \gamma\}$ για κάποιο μη μηδενικό γραμμικό συναρτησοειδές $\phi : \mathbb{R}^n \to \mathbb{R}$ και κάποιο $\gamma \in \mathbb{R}$ ή, ισοδύναμα, αν για κάθε $(x,t) \in H$ ισχύει $\{x\} \times \mathbb{R} \subseteq H$.

Η απόδειξη του Θεωρήματος 3.5 δείχνει ότι αν $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$ κυρτή και $x \in A$, τότε η f έχει στήριγμα στο x αν και μόνο αν υπάρχει μη κατακόρυφο υπερεπίπεδο στον $\mathbb{R}^{n+1}$ που στηρίζει το $\mathrm{epi}(f)$ στο $(x, f(x))$. Ειδικότερα, τέτοιο υπερεπίπεδο υπάρχει αν $x \in \mathrm{int}(A)$.

§7. Πολλές από τις συναρτήσεις που εμφανίζονται συχνά στην κυρτότητα είναι συναρτήσεις f με πραγματικές τιμές, που ορίζονται σε έναν κυρτό κώνο K (συχνά τον ίδιο τον $\mathbb{R}^n$) και ικανοποιούν την ισότητα $f(\lambda x) = \lambda f(x)$ για κάθε $x \in A$ και κάθε $\lambda > 0$. Αυτές οι συναρτήσεις λέγονται *θετικά ομογενείς*. Η νόρμα $\|\cdot\|$ μάς δίνει ένα σημαντικό παράδειγμα θετικά ομογενούς συνάρτησης. Θα δούμε περισσότερα παραδείγματα στις επόμενες παραγράφους.

Θεώρημα 3.6 *Έστω f θετικά ομογενής συνάρτηση που ορίζεται σε έναν κυρτό κώνο K του $\mathbb{R}^n$. Τότε, η f είναι κυρτή αν και μόνο αν $f(x+y) \leq f(x) + f(y)$ για κάθε $x, y \in K$.*

Απόδειξη: Άσκηση. □

Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ και $f : A \to \mathbb{R}$. Για κάθε $a \in \mathbb{R}$, το *σύνολο στάθμης L_a ύψους a της f* είναι το

$$L_a = \{x \in A : f(x) \leq a\}.$$

Αν η $f : A \to \mathbb{R}$ είναι κυρτή, τότε κάθε L_a είναι κυρτό σύνολο. Πράγματι, αν $x, y \in L_a$ και $\lambda, \mu \geq 0$ με $\lambda + \mu = 1$, τότε

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) \leq \lambda a + \mu a = a,$$

δηλαδή $\lambda x + \mu y \in L_a$. Υπάρχουν μη κυρτές συναρτήσεις (όπως η $f : \mathbb{R} \to \mathbb{R}$ με $f(x) = x^3$) των οποίων όλα τα σύνολα στάθμης είναι κυρτά. Στην περίπτωση όμως μιάς μη αρνητικής θετικά ομογενούς συνάρτησης, αν έστω και ένα σύνολο στάθμης (π.χ. το L_1) είναι κυρτό, αυτό εξασφαλίζει την κυρτότητα της f.

Θεώρημα 3.7 *Έστω μία μη αρνητική θετικά ομογενής συνάρτηση f που ορίζεται σε έναν κυρτό κώνο K του $\mathbb{R}^n$. Αν το $\{x \in K : f(x) \leq 1\}$ είναι κυρτό, τότε η f είναι κυρτή.*

Απόδειξη: Θα αποδείξουμε ότι $f(x+y) \leq f(x) + f(y)$ για κάθε $x, y \in K$.

Έστω $L = \{x \in K : f(x) \leq 1\}$. Αν $f(x) > 0$, $f(y) > 0$, τότε $x/f(x) \in L$ και $y/f(y) \in L$, άρα

$$\frac{x+y}{f(x)+f(y)} = \frac{f(x)}{f(x)+f(y)}\frac{x}{f(x)} + \frac{f(y)}{f(x)+f(y)}\frac{y}{f(y)} \in L,$$

άρα

$$f\left(\frac{x+y}{f(x)+f(y)}\right) \leq 1.$$

Αν $f(x)=0$, $f(y)>0$, αρκεί να αποδειχθεί ότι $f(x+y) \leq f(y)+\varepsilon$ για κάθε $\varepsilon > 0$. Ισχύει $x/\varepsilon \in L$ και $y/f(y) \in L$, άρα

$$\frac{x+y}{\varepsilon+f(y)} = \frac{\varepsilon}{\varepsilon+f(y)}\frac{x}{\varepsilon} + \frac{f(y)}{\varepsilon+f(y)}\frac{y}{f(y)} \in L,$$

άρα

$$f\left(\frac{x+y}{\varepsilon+f(y)}\right) \leq 1.$$

Αν $f(x)=0$, $f(y)=0$, αποδεικνύεται όμοια ότι $f(x+y) \leq \varepsilon$ για κάθε $\varepsilon > 0$. □

Ολοκληρώνουμε αυτήν την παράγραφο με μία εφαρμογή αυτού του αποτελέσματος. Έστω $p \geq 1$. Ορίζουμε

$$f(x) = f(\xi_1, \ldots, \xi_n) = (\xi_1^p + \ldots + \xi_n^p)^{1/p}$$

για κάθε $x=(\xi_1,\ldots,\xi_n)$ με $\xi_i \geq 0$. Η f είναι μη αρνητική και θετικά ομογενής. Επίσης, η f^p είναι κυρτή συνάρτηση σαν άθροισμα των κυρτών συναρτήσεων $\xi_1^p, \ldots, \xi_n^p$. Άρα, το σύνολο στάθμης

$$\{x : f(x) \leq 1\} = \{x : f^p(x) \leq 1\}$$

είναι κυρτό. Από το Θεώρημα 3.7, η f είναι κυρτή. Άμεση συνέπεια είναι η ανισότητα του Minkowski

$$\left(|\xi_1+\eta_1|^p + \ldots + |\xi_n+\eta_n|^p\right)^{1/p} \leq \left(|\xi_1|^p + \ldots + |\xi_n|^p\right)^{1/p} + \left(|\eta_1|^p + \ldots + |\eta_n|^p\right)^{1/p},$$

για κάθε $\xi_i, \eta_i \in \mathbb{R}$.

3.2 Συνέχεια και παραγωγισιμότητα

Σε αυτήν την παράγραφο θα μελετήσουμε τη συνέχεια και την διαφορισιμότητα των κυρτών συναρτήσεων.

Θεώρημα 3.8 *Έστω f κυρτή συνάρτηση ορισμένη σε ένα μη κενό ανοικτό σύνολο $A \subseteq \mathbb{R}^n$. Τότε, η f είναι συνεχής στο A.*

Απόδειξη: Έστω $x_0 \in A$. Για την απόδειξη της συνέχειας της f στο x_0, αρκεί να περιοριστούμε στην περίπτωση στην οποία η f έχει ολικό ελάχιστο στο x_0. Γιατί, αν T είναι ένα στήριγμα της f στο x_0, τότε η $f-T$ είναι κυρτή και έχει ολικό ελάχιστο στο x_0. Άρα, αφού η T είναι συνεχής, η συνέχεια της f στο x_0 έπεται από τη συνέχεια της $f-T$ στο x_0.

Έστω λοιπόν ότι ισχύει $f(x) \geq f(x_0)$, $x \in A$. Σύμφωνα με το Θεώρημα 1.3, η f είναι συνεχής χωριστά ως προς κάθε μεταβλητή, άρα $\lim_{\lambda\to 0} f(x_0 + \lambda e_i) = f(x_0)$ για κάθε $i = 1, \ldots, n$.

Έστω (x_k) ακολουθία σημείων του A με $x_k \to x_0$. Τότε ισχύει $x_{ki} \to x_{0i}$ για κάθε $i = 1, \ldots, n$ (όπου $x_k = (x_{k1}, \ldots, x_{kn})$). Αφού

$$x_k = x_0 + (x_k - x_0) = x_0 + \sum_{i=1}^{n}(x_{ki} - x_{0i})e_i = \frac{1}{n}\sum_{i=1}^{n}(x_0 + n(x_{ki} - x_{0i})e_i),$$

από την κυρτότητα της f έπεται ότι

$$f(x_0) \leq f(x_k) \leq \frac{1}{n}\sum_{i=1}^{n} f(x_0 + n(x_{ki} - x_{0i})e_i).$$

Από τη χωριστή συνέχεια της f έπεται ότι το δεξιό μέλος της ανισότητας όντως συγκλίνει στο $f(x_0)$, άρα $f(x_k) \to f(x_0)$. □

Ορισμός Έστω $\{e_1, \ldots, e_n\}$ η συνήθης ορθοκανονική βάση του $\mathbb{R}^n$. Αν A είναι ένα ανοικτό υποσύνολο του $\mathbb{R}^n$ και $f : A \to \mathbb{R}$, η i-στή μερική παράγωγος της f στο $x \in A$ ορίζεται από την

$$\frac{\partial f}{\partial x_i}(x) = \lim_{\lambda\to 0}\frac{f(x + \lambda e_i) - f(x)}{\lambda}$$

αν αυτό το όριο υπάρχει. Αν $y \in \mathbb{R}^n$, η παράγωγος της f στην κατεύθυνση του y είναι το όριο

$$\lim_{\lambda\to 0}\frac{f(x + \lambda y) - f(x)}{\lambda}$$

αν βέβαια υπάρχει.

Αν το A είναι ανοικτό και κυρτό και η $f : A \to \mathbb{R}$ κυρτή συνάρτηση, τότε οι κατά κατεύθυνση παράγωγοι στο $x \in A$ μπορεί να μην υπάρχουν, οι πλευρικές όμως κατά κατεύθυνση παράγωγοι

$$f'(x; y) := \lim_{\lambda\to 0^+}\frac{f(x + \lambda y) - f(x)}{\lambda}$$

υπάρχουν για κάθε $x \in A$ και $y \neq 0$. Αυτό προκύπτει ως εξής: το

$$I = \{\lambda \in \mathbb{R} : x + \lambda y \in A\}$$

είναι ένα ανοικτό διάστημα του $\mathbb{R}$ που περιέχει το μηδέν. Η συνάρτηση $g : I \to \mathbb{R}$ με $g(a) = f(x + ay)$ είναι κυρτή: αν $a, b \in I$ και $\lambda, \mu \geq 0$ με $\lambda + \mu = 1$, τότε

$$\begin{aligned} g(\lambda a + \mu b) &= f(x + (\lambda a + \mu b)y) = f(\lambda(x + ay) + \mu(x + by)) \\ &\leq \lambda f(x + ay) + \mu f(x + by) = \lambda g(a) + \mu g(b). \end{aligned}$$

Αφού η g είναι κυρτή, υπάρχει η

$$g'_+(0) = \lim_{\lambda\to 0^+}\frac{g(\lambda) - g(0)}{\lambda} = \lim_{\lambda\to 0^+}\frac{f(x + \lambda y) - f(x)}{\lambda}.$$

Επίσης,

$$\lim_{\lambda\to 0^-}\frac{f(x+\lambda y)-f(x)}{\lambda} = -\lim_{\lambda\to 0^+}\frac{f(x+\lambda(-y))-f(x)}{\lambda} = -f'(x;-y)$$

δηλαδή, οι δύο πλευρικές παράγωγοι στο x υπάρχουν και η παράγωγος της f στην κατεύθυνση του y υπάρχει αν και μόνο αν

$$f'(x;y) = -f'(x;-y).$$

Παράδειγμα: Θεωρούμε την κυρτή συνάρτηση $f(x)=\|x\|^2=\xi_1^2+\ldots+\xi_n^2$. Για κάθε $y=(\eta_1,\ldots,\eta_n)\in\mathbb{R}^n$ έχουμε

$$\begin{aligned} f'(x;y) &= \lim_{\lambda\to 0^+}\frac{\|x+\lambda y\|^2-\|x\|^2}{\lambda}\\ &= \lim_{\lambda\to 0^+}\frac{2\lambda\langle x,y\rangle+\lambda^2\|y\|^2}{\lambda}\\ &= 2\langle x,y\rangle. \end{aligned}$$

Αφού $-f'(x;-y)=-2\langle x,-y\rangle=f'(x;y)$, η f έχει παράγωγο στο x σε κάθε κατεύθυνση y. Επίσης, η συνάρτηση $f'(x;\cdot):\mathbb{R}^n\to\mathbb{R}$ με $y\mapsto f'(x;y)$ είναι γραμμική για κάθε $x\in\mathbb{R}^n$.

Θεώρημα 3.9 *Έστω f κυρτή συνάρτηση ορισμένη σε ένα ανοικτό κυρτό σύνολο A και έστω $x\in A$. Τότε, η $f'(x;\cdot):\mathbb{R}^n\to\mathbb{R}$ είναι θετικά ομογενής κυρτή συνάρτηση και*

$$f'(x;y)\geq -f'(x;-y)$$

για κάθε $y\in\mathbb{R}^n$. Αν η f έχει παράγωγο στο x στην κατεύθυνση του y, τότε $f'(x;\lambda y)=\lambda f'(x;y)$ για κάθε $\lambda\in\mathbb{R}$.

Απόδειξη: Για τη θετική ομογένεια παρατηρούμε ότι, αν $t>0$,

$$\begin{aligned} f'(x;ty) &= \lim_{\lambda\to 0^+}\frac{f(x+\lambda ty)-f(x)}{\lambda}\\ &= t\lim_{\lambda\to 0^+}\frac{f(x+\lambda ty)-f(x)}{t\lambda} = tf'(x;y). \end{aligned}$$

Η κυρτότητα προκύπτει από τις παρακάτω σχέσεις: αν $y_1,y_2\in\mathbb{R}^n$, $t\in(0,1)$,

$$\begin{aligned} f'(x;ty_1+(1-t)y_2) &= \lim_{\lambda\to 0^+}\frac{f(t(x+\lambda y_1)+(1-t)(x+\lambda y_2))-f(x)}{\lambda}\\ &\leq \lim_{\lambda\to 0^+}\left(t\frac{f(x+\lambda y_1)-f(x)}{\lambda}+(1-t)\frac{f(x+\lambda y_2)-f(x)}{\lambda}\right). \end{aligned}$$

Για $x\in A$, $y\in\mathbb{R}^n$, θεωρούμε τη συνάρτηση $g(\lambda)=f(x+\lambda y)$, $\lambda\in(-\varepsilon,\varepsilon)$ (όπου $\varepsilon>0$ τέτοιο ώστε $D(x,\varepsilon)\subseteq A$). Τότε,

$$f'(x;y)=g'_+(0),\qquad -f'(x;-y)=g'_-(0).$$

Αφού η g είναι κυρτή, ισχύει $g'_-(0) \leq g'_+(0)$. Τέλος, αν η f έχει παράγωγο στο x στην κατεύθυνση του y και $\lambda < 0$, έχουμε

$$f'(x;\lambda y) = -f'(x;-\lambda y) = \lambda f'(x;y).$$

$\square$

Έστω A ανοικτό υποσύνολο του $\mathbb{R}^n$. Αν $f : A \to \mathbb{R}$ και $x \in A$, λέμε ότι η f είναι *διαφορίσιμη* στο x αν υπάρχει $z \in \mathbb{R}^n$ (το οποίο είναι αναγκαστικά μοναδικό) με την ιδιότητα

$$\lim_{u\to 0} \frac{f(x+u) - f(x) - \langle z, u\rangle}{\|u\|} = 0.$$

Αν ένα τέτοιο z υπάρχει, λέγεται *ανάδελτα* της f στο x και συμβολίζεται με $\nabla f(x)$. Αν το $\nabla f(x)$ υπάρχει, τότε για κάθε $y \neq 0$ έχουμε

$$\begin{aligned} 0 &= \lim_{\lambda\to 0} \frac{|f(x+\lambda y) - f(x) - \langle \nabla f(x), \lambda y\rangle|}{\|\lambda y\|} \\ &= \lim_{\lambda\to 0} \frac{1}{\|y\|}\left|\frac{f(x+\lambda y) - f(x)}{\lambda} - \langle \nabla f(x), y\rangle\right|, \end{aligned}$$

δηλαδή η f έχει παράγωγο στο x σε οποιαδήποτε κατεύθυνση y. Επιπλέον, $f'(x;y) = \langle \nabla f(x), y\rangle$, δηλαδή η $f'(x;\cdot)$ είναι γραμμική.

Η ύπαρξη των κατά κατεύθυνση παραγώγων της f στο x σε *όλες* τις διευθύνσεις y δεν εξασφαλίζει ότι η f είναι διαφορίσιμη στο x ούτε ότι η $f'(x;\cdot)$ είναι γραμμική. Όταν όμως η f είναι κυρτή, έχουμε το εξής αποτέλεσμα.

Θεώρημα 3.10 *Έστω ότι η f είναι κυρτή συνάρτηση, ορισμένη στο ανοικτό και κυρτό σύνολο A. Αν οι μερικές παράγωγοι $\partial f/\partial x_1, \ldots, \partial f/\partial x_n$ υπάρχουν στο x, τότε η f είναι διαφορίσιμη στο x.*

Απόδειξη: Αρκεί να αποδειχθεί ότι

$$\lim_{u\to 0} \frac{f(x+u) - f(x) - \langle \nabla f(x), u\rangle}{\|u\|} = 0.$$

Θεωρούμε $r > 0$ έτσι ώστε $D(x,r) \subseteq A$. Ορίζουμε την συνάρτηση

$$\phi(u) = f(x+u) - f(x) - \langle \nabla f(x), u\rangle, \qquad u \in \mathbb{R}^n,\ \|u\| < \varepsilon.$$

Τότε, αν $|\lambda| < \varepsilon$,

$$\frac{\phi(\lambda e_i)}{\lambda} = \frac{1}{\lambda}\left(f(x+\lambda e_i) - f(x) - \lambda\frac{\partial f}{\partial x_i}(x)\right) = \frac{f(x+\lambda e_i) - f(x)}{\lambda} - \frac{\partial f}{\partial x_i}(x).$$

Άρα,

$$\lim_{\lambda\to 0} \frac{\phi(\lambda e_i)}{\lambda} = 0, \qquad i = 1, \ldots, n. \tag{1}$$

Επειδή η ϕ είναι κυρτή, αν $0 < \|u\| < \varepsilon$ έχουμε

$$\begin{aligned}
\frac{\phi(u)}{\|u\|} &= \frac{1}{\|u\|}\phi\left(\sum_{i=1}^{n} u_i e_i\right) = \frac{1}{\|u\|}\phi\left(\frac{1}{n}\sum_{i=1}^{n} n u_i e_i\right) \\
&\leq \sum_{i=1}^{n} \frac{|u_i|}{\|u\|}\left|\frac{\phi(n u_i e_i)}{n u_i}\right| \leq \sum_{i=1}^{n}\left|\frac{\phi(n u_i e_i)}{n u_i}\right|.
\end{aligned}$$

Αφού $0 = \phi(0) \leq (\phi(u) + \phi(-u))/2$, έχουμε $\phi(u) \geq -\phi(-u)$. Άρα, από την προηγούμενη ανισότητα έπεται ότι

$$-\sum_{i=1}^{n}\left|\frac{\phi(-n u_i e_i)}{n u_i}\right| \leq \frac{-\phi(-u)}{\|u\|} \leq \frac{\phi(u)}{\|u\|} \leq \sum_{i=1}^{n}\left|\frac{\phi(n u_i e_i)}{n u_i}\right|.$$

Από την (1) τώρα έπεται ότι $\lim_{u\to 0} \frac{\phi(u)}{\|u\|} = 0$ (Αν $u_i = 0$, θεωρούμε τους αντίστοιχους όρους των παραπάνω αθροισμάτων ίσους με 0). □

Το επόμενο θεώρημα χαρακτηρίζει την διαφορισιμότητα μιάς κυρτής συνάρτησης σε ένα σημείο μέσω των συναρτήσεων στήριξης σε αυτό. Θα χρειαστούμε ένα λήμμα:

Λήμμα *Έστω A κυρτό υποσύνολο του $\mathbb{R}^n$ με $\mathrm{int}(A) \neq \emptyset$, $x_0 \in \mathrm{bd}(A)$, και L ευθεία που περιέχει το x_0 και δεν τέμνει το $\mathrm{int}(A)$. Τότε υπάρχει υπερεπίπεδο H που στηρίζει το A στο x_0 και περιέχει την L.*

Απόδειξη: Τα A, L είναι κυρτά σύνολα και $L \cap \mathrm{int}(A) = \emptyset$. Από το Θεώρημα 2.9 έπεται ότι υπάρχει υπερεπίπεδο H που διαχωρίζει τα A και L. Δηλαδή, υπάρχει μη μηδενικό γραμμικό συναρτησοειδές $\phi : \mathbb{R}^n \to \mathbb{R}$ και $\gamma \in \mathbb{R}$ έτσι ώστε: $\phi(a) \geq \gamma$ αν $a \in A$, $\phi(x) \leq \gamma$ αν $x \in L$, και $\phi(z) = \gamma$ αν $z \in H$.

Το L έχει τη μορφή $L = \{x_0 + ty : t \in \mathbb{R}\}$. Τότε, $\phi(x_0) + t\phi(y) \leq \gamma$ για κάθε $t \in \mathbb{R}$, άρα $\phi(y) = 0$. Έπεται ότι $\phi(x_0 + ty) = \phi(x_0) \leq \gamma$ για κάθε $t \in \mathbb{R}$.

Αφού όμως $x_0 \in \overline{A}$, έχουμε $\phi(x_0) \geq \gamma$. Άρα, $\phi(x_0) = \gamma = \phi(x_0 + ty)$ για κάθε $t \in \mathbb{R}$. Δηλαδή, $L \subseteq H$. □

Θεώρημα 3.11 *Έστω f κυρτή συνάρτηση ορισμένη στο ανοικτό και κυρτό σύνολο A. Τότε, η f είναι διαφορίσιμη στο $x_0 \in A$ αν και μόνο αν έχει μοναδικό στήριγμα στο x_0.*

Απόδειξη: Έστω T στήριγμα της f στο x_0. Αφού η T είναι αφφινική και $T(x_0) = f(x_0)$, η T έχει τη μορφή

$$T(x) = \langle a, x - x_0\rangle + f(x_0)$$

για κάποιο $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Τότε, για $i = 1, \ldots, n$ ισχύει

$$\begin{aligned}
a_i &= \frac{\partial T}{\partial x_i}(x_0) = \lim_{\lambda\to 0^+} \frac{T(x_0 + \lambda e_i) - T(x_0)}{\lambda} \\
&\leq \lim_{\lambda\to 0^+} \frac{f(x_0 + \lambda e_i) - f(x_0)}{\lambda} = \frac{\partial f}{\partial x_i}(x_0).
\end{aligned}$$

Επιπλέον,

$$\begin{aligned} a_i &= \lim_{\lambda\to 0^-} \frac{T(x_0+\lambda e_i)-T(x_0)}{\lambda} \\ &\geq \lim_{\lambda\to 0^-} \frac{f(x_0+\lambda e_i)-f(x_0)}{\lambda} = \frac{\partial f}{\partial x_i}(x_0). \end{aligned}$$

Άρα $a = \nabla f(x_0)$. Αυτό σημαίνει ότι για οποιοδήποτε στήριγμα T της f στο x_0 θα ισχύει

$$T(x) = \langle \nabla f(x_0), x - x_0\rangle + f(x_0).$$

Για την αντίστροφη κατεύθυνση, σύμφωνα με το Θεώρημα 3.10 αρκεί να αποδειχθεί ότι υπάρχουν οι μερικές παράγωγοι

$$\frac{\partial f}{\partial x_i}(x_0), \qquad i = 1,\dots,n.$$

Έστω $\varepsilon > 0$ τέτοιο ώστε $D(x_0,\varepsilon) \subseteq A$. Θεωρούμε τη συνάρτηση $g(\lambda) = f(x_0+\lambda e_i)$, $\lambda \in (-\varepsilon,\varepsilon)$. Αρκεί να αποδειχθεί ότι η g είναι παραγωγίσιμη στο 0, δηλαδή (Θεώρημα 1.7) ότι η g έχει μοναδικό στήριγμα στο 0. Άρα, αρκεί να αποδειχθεί ότι αν η g έχει δύο διαφορετικά στηρίγματα στο 0, τότε η f έχει δύο διαφορετικά στηρίγματα στο x_0.

Θα αποδειχθεί ότι, αν h είναι στήριγμα της g στο 0, υπάρχει στήριγμα T της f στο x_0 έτσι ώστε $T(x_0+\lambda e_i) = h(\lambda)$, $\lambda \in \mathbb{R}$. Αφού διαφορετικές h_1, h_2 δίνουν διαφορετικές T_1, T_2 αντίστοιχα, προκύπτει το ζητούμενο.

Έστω λοιπόν h στήριγμα της g στο 0. Η h είναι αφφινική, άρα έχει τη μορφή $h(\lambda) = a\lambda + b$. Θεωρούμε την ευθεία

$$L = \{(x_0+\lambda e_i, h(\lambda)) : \lambda \in \mathbb{R}\} = \{(x_0, b) + \lambda(e_i, a) : \lambda \in \mathbb{R}\}.$$

Αφού $b = h(0) = g(0) = f(x_0)$, το $(x_0, f(x_0))$ ανήκει στην L. Ανήκει επίσης στο $\mathrm{bd}(\mathrm{epi}(f))$. Επιπλέον ισχύει $\mathrm{int}(\mathrm{epi}(f)) \neq \emptyset$, αφού το σύνολο $\{(x,t) : x \in A, t > f(x)\}$ είναι ανοικτό, λόγω της συνέχειας της f. Επίσης,

$$L \cap \mathrm{int}(\mathrm{epi}(f)) = \emptyset.$$

Πράγματι, αν $(x,t) \in \mathrm{int}(\mathrm{epi}(f))$ και $x \in A$, τότε $t > f(x)$, ενώ τα σημεία της L έχουν τη μορφή $(x_0+\lambda e_i, h(\lambda))$. Αν $x_0+\lambda e_i \in A$, αφού η h στηρίζει την g παίρνουμε

$$h(\lambda) \leq g(\lambda) = f(x_0+\lambda e_i).$$

Από το Λήμμα προκύπτει ότι υπάρχει υπερεπίπεδο H στον $\mathbb{R}^{n+1}$ το οποίο στηρίζει το $\mathrm{epi}(f)$ στο $(x_0, f(x_0))$ και περιέχει την L. Τότε (βλέπε απόδειξη του Θεωρήματος 3.5), υπάρχει αφφινική συνάρτηση $T : \mathbb{R}^n \to \mathbb{R}$ τέτοια ώστε $H = \{(x, T(x)) : x \in \mathbb{R}^n\}$. Η T στηρίζει την f. Επιπλέον, για κάθε $\lambda \in \mathbb{R}$ ισχύει $(x_0+\lambda e_i, h(\lambda)) \in L$, άρα $T(x_0+\lambda e_i) = h(\lambda)$. □

Στο Κεφάλαιο 1 είδαμε ότι μία συνάρτηση f που έχει δεύτερη παράγωγο είναι κυρτή αν και μόνο αν η f'' είναι μη αρνητική. Δίνουμε τώρα ένα ανάλογο κριτήριο για κυρτές συναρτήσεις n μεταβλητών που έχουν μερικές παραγώγους δεύτερης τάξης.

Θεώρημα 3.12 *Έστω f συνάρτηση που είναι ορισμένη σε ένα ανοικτό κυρτό σύνολο A και έχει συνεχείς μερικές παραγώγους δεύτερης τάξης σε κάθε $x \in A$. Τότε, η f είναι κυρτή αν και μόνο αν για κάθε $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right] (x) z_i z_j \geq 0.$$

Απόδειξη: Έστω ότι η f είναι κυρτή. Θεωρούμε $x \in A$, $\varepsilon > 0$ με $D(x, \varepsilon) \subseteq A$ και $z = (z_1, \ldots, z_n)$ με $\|z\| = 1$. Τότε, η συνάρτηση $g(\lambda) = f(x + \lambda z)$, $\lambda \in (-\varepsilon, \varepsilon)$ είναι κυρτή και δύο φορές παραγωγίσιμη. Από τον κανόνα της αλυσίδας παίρνουμε

$$g''(\lambda) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} (x + \lambda z) z_i z_j.$$

Αφού η g είναι κυρτή, ισχύει

$$0 \leq g''(0) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} (x) z_i z_j.$$

Δηλαδή, η ζητούμενη ανισότητα ισχύει για κάθε z με $\|z\| = 1$. Αν $z \neq 0$, η ανισότητα ισχύει για το $z / \|z\|$, άρα και για το z.

Αντίστροφα, έστω ότι ισχύουν οι θεωρούμενες ανισότητες. Για να αποδειχθεί ότι η f είναι κυρτή, αρκεί να αποδειχθεί ότι για κάθε $x, y \in A$ η $g(\lambda) = f(x + \lambda(y - x))$, $\lambda \in [0, 1]$ είναι κυρτή, άρα ότι $g''(\lambda) \geq 0$, $\lambda \in [0, 1]$. Όπως προηγουμένως,

$$g''(\lambda) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} (x + \lambda(y - x))(y_i - x_i)(y_j - x_j) \geq 0.$$

□

Ας υποθέσουμε ότι η f είναι όπως στο προηγούμενο θεώρημα. Ο πίνακας $(\partial^2 f / \partial x_i \partial x_j)(x)$ λέγεται *Εσσιανή* της f στο x. Από τις υποθέσεις μας για την f, ο πίνακας αυτός είναι συμμετρικός. Έχουμε λοιπόν αποδείξει ότι η f είναι κυρτή αν και μόνο αν η Εσσιανή της είναι συμμετρικός θετικά ημιορισμένος πίνακας για κάθε $x \in A$.

3.3 Συναρτήσεις στήριξης

§1. Έστω K κυρτό και συμπαγές υποσύνολο του $\mathbb{R}^n$, $u \in \mathbb{R}^n$, $u \neq 0$. Ας θεωρήσουμε το συναρτησοειδές ϕ με $\phi(x) = \langle u, x \rangle$, $x \in \mathbb{R}^n$. Σε αυτό αντιστοιχούν τα υπερεπίπεδα $H_\gamma = \{x \in \mathbb{R}^n : \phi(x) = \gamma\}$, $\gamma \in \mathbb{R}$. Δύο από αυτά στηρίζουν το K, τα

$$H_1 = \{x \in \mathbb{R}^n : \phi(x) = \min\{\phi(z) : z \in K\}\}$$

και
$$H_2 = \{x \in \mathbb{R}^n : \phi(x) = \max\{\phi(z) : z \in K\}\}.$$
Ορίζουμε μία συνάρτηση στον $\mathbb{R}^n$ δίνοντας σε κάθε u την τιμή του γ που αντιστοιχεί στο H_2. Ορίζουμε δηλαδή
$$h_K(u) = \max\{\langle u, x\rangle : x \in K\}, \qquad u \in \mathbb{R}^n.$$
Η h_K λέγεται *συνάρτηση στήριξης* του K.

Παραδείγματα (1) Έστω $K = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : \max\{|x_i| : i = 1, \ldots, n\} \leq 1\}$ (δηλαδή το K είναι ένας «κύβος» με κέντρο την αρχή των αξόνων και μήκος ακμής 2). Για κάθε $u \in \mathbb{R}^n$ ισχύει
$$\begin{aligned} h_K(u) &= \max\{\langle u, x\rangle : x \in K\} = \max\{\langle u, x\rangle : |x_i| \leq 1, i = 1, \ldots, n\} \\ &\leq \sum_{i=1}^n |u_i|. \end{aligned}$$
Επιπλέον, αν $x_i = \mathrm{sgn}(u_i)$, τότε $x = (x_1, \ldots, x_n) \in K$ και
$$\langle u, x\rangle = \sum_{i=1}^n \mathrm{sgn}(u_i) u_i = \sum_{i=1}^n |u_i|.$$
Συνεπώς,
$$h_K(u) = \sum_{i=1}^n |u_i|.$$
(2) Έστω $K = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i=1}^n |x_i| \leq 1\}$. (Το K είναι πολύεδρο με κορυφές τα $\pm e_i$, $i = 1, \ldots, n$. Για $n = 1$ είναι τετράγωνο, για $n = 2$ οκτάεδρο). Για κάθε $u \in \mathbb{R}^n$ ισχύει
$$\begin{aligned} h_K(u) &= \max\{\langle u, x\rangle : x \in K\} = \max\{\langle u, x\rangle : \sum_{i=1}^n |x_i| \leq 1\} \\ &\leq \max\{|u_i| : i = 1, \ldots, n\}. \end{aligned}$$
Επιπλέον, αν $\max\{|u_i| : i = 1, \ldots, n\} = |u_{i_0}|$ και $x = (\mathrm{sgn} u_{i_0}) e_{i_0}$, τότε $x \in K$ και $\langle u, x\rangle = |u_{i_0}|$. Άρα,
$$h_K(u) = \max\{|u_i| : i = 1, \ldots, n\}.$$

Ιδιότητες της συνάρτησης στήριξης

(α) Η h_K είναι κυρτή (ως supremum αφφινικών συναρτήσεων) και θετικά ομογενής.
(β) $h_{\lambda K} = \lambda h_K$, αν $\lambda > 0$.
(γ) $h_{K_1 + K_2} = h_{K_1} + h_{K_2}$. Πράγματι,
$$\begin{aligned} h_{K_1+K_2}(u) &= \max\{\langle u, x_1 + x_2\rangle : x_1 \in K_1, x_2 \in K_2\} \\ &= \max\{\langle u, x_1\rangle + \langle u, x_2\rangle : x_1 \in K_1, x_2 \in K_2\} \\ &= \max\{\langle u, x_1\rangle : x_1 \in K_1\} + \max\{\langle u, x_2\rangle : x_2 \in K_2\}. \end{aligned}$$

Πρόταση 3.1 *Αν K_1, K_2 κυρτά συμπαγή υποσύνολα του $\mathbb{R}^n$, τότε $K_1 \subseteq K_2$ αν και μόνο αν $h_{K_1} \leq h_{K_2}$.*

Απόδειξη: Η μία συνεπαγωγή είναι προφανής από τον ορισμό της h_K.

Έστω τώρα ότι δεν ισχύει $K_1 \subseteq K_2$, δηλαδή ότι υπάρχει $x \in K_1$ με $x \notin K_2$. Τότε τα x, K_2 χωρίζονται γνήσια από ένα υπερεπίπεδο, άρα υπάρχει $u \in \mathbb{R}^n$ έτσι ώστε $\langle u, y\rangle < \langle u, x\rangle$ για κάθε $y \in K_2$. Τότε,

$$h_{K_2}(u) = \max\{\langle u, y\rangle : y \in K_2\} < \langle u, x\rangle \leq h_{K_1}(u).$$

Άρα δεν ισχύει $h_{K_1} \leq h_{K_2}$. □

Πόρισμα 1 *Αν K κυρτό συμπαγές, τότε $0 \in K$ αν και μόνο αν $h_K \geq 0$.*

Απόδειξη: $h_{\{0\}} = 0$. □

Πόρισμα 2 *Η συνάρτηση στήριξης καθορίζει το σύνολο K. Δηλαδή, αν K_1, K_2 κυρτά συμπαγή με $h_{K_1} = h_{K_2}$, τότε $K_1 = K_2$.* □

Θεώρημα 3.13 *Έστω $h : \mathbb{R}^n \to \mathbb{R}$ κυρτή και θετικά ομογενής. Τότε υπάρχει (μοναδικό) κυρτό συμπαγές υποσύνολο K του $\mathbb{R}^n$ με $h_K = h$.*

Απόδειξη: Θεωρούμε το σύνολο

$$K = \{x \in \mathbb{R}^n : \langle u, x\rangle \leq h(u), \forall u \in \mathbb{R}^n\}.$$

Για κάθε $u \in \mathbb{R}^n$ το σύνολο $\{x \in \mathbb{R}^n : \langle u, x\rangle \leq h(u)\}$ είναι ημίχωρος (ή ολόκληρος ο $\mathbb{R}^n$), άρα κυρτό. Συνεπώς και το K είναι κυρτό ως τομή κυρτών συνόλων.

Επειδή η h είναι συνεχής, το K είναι κλειστό. Επιπλέον, είναι συμπαγές, αφού για κάθε $x \in K$ και $i = 1, \ldots, n$ ισχύει

$$-h(e_i) \leq -\langle -e_i, x\rangle = x_i = \langle e_i, x\rangle \leq h(e_i).$$

Είναι φανερό ότι $h_K(u) \leq h(u)$. Για την αντίστροφη ανισότητα, θεωρούμε ένα $u \in \mathbb{R}^n$ και θα αποδείξουμε ότι υπάρχει $x \in K$ τέτοιο ώστε $\langle y, x\rangle \leq h(y)$ για κάθε $y \in \mathbb{R}^n$ και $h(u) \leq \langle u, x\rangle$, δηλαδή $h(u) = \langle u, x\rangle$.

Ισοδύναμα, θα αποδείξουμε ότι υπάρχει γραμμικό συναρτησοειδές $\phi : \mathbb{R}^n \to \mathbb{R}$ που στηρίζει την h στο u. Αφού η h είναι κυρτή, υπάρχει αφφινική ϕ με την ιδιότητα αυτή. Αρκεί λοιπόν να αποδειχθεί ότι $\phi(0) = 0$.

Θεωρούμε τη συνάρτηση

$$\psi(\lambda) = h(\lambda u) - \phi(\lambda u) = \lambda h(u) - \phi(\lambda u), \quad \lambda \in (0, +\infty).$$

Η ψ είναι αφφινική, και $\psi \geq 0$, $\psi(1) = 0$. Δηλαδή η ψ είναι κοίλη συνάρτηση στο $(0, +\infty)$, που έχει ελάχιστη τιμή. Συνεπώς είναι σταθερή, άρα $\lambda h(u) = \phi(\lambda u)$, $\lambda \in (0, +\infty)$. Για $\lambda \to 0$ παίρνουμε $\phi(0) = 0$. □

§2. Για κυρτά κλειστά σύνολα που περιέχουν το 0 στο εσωτερικό τους ορίζεται και η *συνάρτηση στάθμης* (gauge function), η οποία μπορεί να θεωρηθεί δυϊκή της συνάρτησης στήριξης.

Έστω L κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$ με $0 \in \mathrm{int}(L)$. Έστω $x \in \mathbb{R}^n$. Τότε υπάρχουν θετικά λ έτσι ώστε $x \in \lambda L$ (Αν $r > 0$ με $D(0,r) \subseteq L$, αρκεί να ισχύει $\|x\| < \lambda r$). Η συνάρτηση στάθμης του L ορίζεται ως εξής:

$$g_L(x) = \inf\{\lambda > 0 : x \in \lambda L\}.$$

Παρατήρηση Επειδή το L είναι κυρτό, αν $x \in \lambda L$, $0 < \lambda < \mu$, τότε ισχύει και $x \in \mu L$, αφού $x = \lambda z$ με $z \in L$, άρα

$$x = \mu\left(\frac{\lambda}{\mu}z + \left(1 - \frac{\lambda}{\mu}\right)0\right) \Longrightarrow x \in \mu L.$$

Συνεπώς το σύνολο $\Lambda = \{\lambda > 0 : x \in \lambda L\}$ έχει τη μορφή $(a, +\infty)$ ή $[a, +\infty)$, όπου $a \geq 0$.

Αν $a > 0$, το σύνολο Λ είναι κλειστό. Γιατί, αν $\lambda_n \in \Lambda$, $n = 1, 2, \ldots$ και $\lambda_n \to \lambda$, τότε υπάρχουν $z_n \in L$ με $x = \lambda_n z_n$, $n = 1, 2, \ldots$. Αφού $\lambda \geq a > 0$, ισχύει $z_n \to x/\lambda$. Αφού το L είναι κλειστό, ισχύει $x/\lambda \in L$, άρα $\lambda \in \Lambda$.

Συνεπώς, το Λ έχει τη μορφή $(0, +\infty)$ ή $[a, +\infty)$, $a > 0$.

Ιδιότητες της συνάρτησης στάθμης

(1) $g_L(x) \geq 0$, $x \in \mathbb{R}^n$.
(2) $g_L(x) = 0$ αν και μόνο αν $\lambda x \in L$ για κάθε $\lambda > 0$. Ειδικότερα, $g_L(0) = 0$.
(3) $g_{\mu L} = \frac{1}{\mu} g_L$, $\mu > 0$. Πράγματι,

$$g_{\mu L}(x) = \inf\{\lambda > 0 : x \in \lambda\mu L\} = \inf\{\frac{\rho}{\mu} > 0 : x \in \rho L\} = \frac{1}{\mu} g_L(x).$$

(4) $L = \{x \in \mathbb{R}^n : g_L(x) \leq 1\}$.
Απόδειξη: Έστω $x \in \mathbb{R}^n$ με $g_L(x) \leq 1$. Αν $g_L(x) = 0$, τότε $\lambda x \in L$ για κάθε $\lambda > 0$, άρα $x \in L$. Αν $g_L(x) > 0$, από την παρατήρηση πιό πάνω έπεται ότι το σύνολο Λ έχει τη μορφή $[a, +\infty)$, όπου $a = g_L(x)$. Έπεται ότι $1 \in \Lambda$.

Ο αντίστροφος εγκλεισμός είναι προφανής από τον ορισμό της g_L. $\square$

(5) $\mu L = \{x \in \mathbb{R}^n : g_L(x) \leq \mu\}$, $\mu > 0$. Άμεσο από τα (3) και (4).
(6) Η g_L είναι θετικά ομογενής. Γιατί, για $\mu > 0$ και $x \in \mathbb{R}^n$ ισχύει

$$\begin{aligned} g_L(\mu x) &= \inf\{\lambda > 0 : \mu x \in \lambda L\} = \inf\{\lambda > 0 : x \in (\lambda/\mu) L\} \\ &= \inf\{\rho\mu > 0 : x \in \rho L\} = \mu g_L(x). \end{aligned}$$

(7) Η g_L είναι κυρτή.
Απόδειξη: Έστω $x, y \in \mathbb{R}^n$, $t \in (0,1)$. Θα αποδείξουμε ότι

$$g_L(tx + (1-t)y) \leq t g_L(x) + (1-t) g_L(y).$$

Αν το δεύτερο μέλος είναι 0, ισχύει $x \in \lambda L$, $y \in \lambda L$ για κάθε $\lambda > 0$. Άρα $tx + (1 - t)y \in \lambda L$ για κάθε $\lambda > 0$, και συνεπώς $g_L(tx + (1-t)y) = 0$.

Έστω ότι το δεύτερο μέλος της ανισότητας είναι διάφορο του 0. Υπάρχουν $\lambda_n > 0$, $\mu_n > 0$ έτσι ώστε

$$\lambda_n \to g_L(x), \quad \mu_n \to g_L(y), \quad x \in \lambda_n L, \quad y \in \mu_n L, \quad n = 1, 2, \ldots.$$

Τότε, $tx + (1-t)y \in (t\lambda_n + (1-t)\mu_n)L$, δηλαδή υπάρχουν $z_n \in L$, $n = 1, 2, \ldots$ με

$$tx + (1-t)y = (t\lambda_n + (1-t)\mu_n)z_n.$$

Έπεται ότι

$$z_n \to \frac{1}{tg_L(x) + (1-t)g_L(y)}[tx + (1-t)y].$$

Αφού το L είναι κλειστό, το όριο αυτό ανήκει στο L, απ΄ όπου έπεται ότι

$$g_L(tx + (1-t)y) \leq tg_L(x) + (1-t)g_L(y).$$

(8) Αν K, L κυρτά κλειστά υποσύνολα του $\mathbb{R}^n$ με $0 \in \text{int}(K) \cap \text{int}(L)$ και $g_K = g_L$, τότε $K = L$. Αυτό είναι άμεση συνέπεια της (4).

Θεώρημα 3.14 *Έστω $g : \mathbb{R}^n \to \mathbb{R}$ θετικά ομογενής κυρτή συνάρτηση, $g \geq 0$. Έστω $L = \{x \in \mathbb{R}^n : g(x) \leq 1\}$. Τότε το L είναι κυρτό, κλειστό, $0 \in \text{int}(L)$ και $g = g_L$.*

Απόδειξη: Το L είναι κυρτό ως σύνολο στάθμης κυρτής συνάρτησης. Επίσης είναι κλειστό, αφού η g είναι συνεχής.

Αφού η g είναι θετικά ομογενής, ισχύει $g(0) = 0$ και η συνέχεια της g συνεπάγεται ότι το σύνολο $\{x \in \mathbb{R}^n : g(x) < 1\}$ είναι ανοικτό. Άρα, $0 \in \text{int}(L)$.

Από την ιδιότητα (4) έπεται ότι $\{x \in \mathbb{R}^n : g(x) \leq 1\} = \{x \in \mathbb{R}^n : g_L(x) \leq 1\}$. Από την ισότητα αυτή θα εξαχθεί η ισότητα των g και g_L.

Έστω $x \in \mathbb{R}^n$. Αν $g(x) = 0$, τότε για κάθε $\lambda > 0$ έχουμε $g(\lambda x) = 0 \leq 1$, άρα $\lambda x \in L$. Έπεται ότι $g_L(x) = 0$. Αν $g_L(x) = 0$, τότε $\lambda x \in L$ για κάθε $\lambda > 0$, άρα $g_L(\lambda x) \leq 1$ για κάθε $\lambda > 0$. Έπεται ότι $\lambda g(x) = g(\lambda x) \leq 1$ για κάθε $\lambda > 0$, συνεπώς $g(x) = 0$.

Έστω $g(x) > 0$, $g_L(x) > 0$. Τότε $g(x/g(x)) = 1$, οπότε

$$g_L\left(\frac{x}{g(x)}\right) = \frac{g_L(x)}{g(x)} \leq 1.$$

Όμοια προκύπτει ότι $g(x)/g_L(x) \leq 1$, άρα $g(x) = g_L(x)$. $\square$

§3. Η έννοια του πολικού ενός κυρτού συνόλου οδηγεί σε μία σχέση μεταξύ των συναρτήσεων στήριξης και των συναρτήσεων στάθμης.

Ορισμός Έστω K κυρτό υποσύνολο του $\mathbb{R}^n$. Το *πολικό* του K είναι το σύνολο

$$K^\circ = \{y \in \mathbb{R}^n : \langle x, y \rangle \leq 1 \; \forall x \in K\}.$$

Παρατηρήσεις 1. Αν το K είναι κυρτός κώνος, ο παραπάνω ορισμός είναι ισοδύναμος με αυτόν της Παραγράφου 2.4. Γιατί, σε αυτήν την περίπτωση, αν $\langle x, y \rangle \leq 1$ για κάθε $x \in K$, ισχύει και $\lambda \langle x, y \rangle \leq 1$ για κάθε $\lambda > 0$, οπότε $\langle x, y \rangle \leq 0$.

2. Το K° είναι κυρτό και κλειστό σύνολο (αυτό προκύπτει ακριβώς όπως και στην περίπτωση κώνου).

3. $0 \in K^\circ$.

4. Αν $0 \in K$, τότε $K^{\circ\circ} = \overline{K}$. Ειδικότερα, αν το K είναι κλειστό τότε $K^{\circ\circ} = K$.

Απόδειξη: Από τον ορισμό βλέπουμε εύκολα ότι $K^{\circ\circ} \supseteq K$. Αφού το $K^{\circ\circ}$ είναι κλειστό, ισχύει $K^{\circ\circ} \supseteq \overline{K}$.

Για να αποδειχθεί η ισότητα, ας υποθέσουμε ότι υπάρχει $x \in K^{\circ\circ}$ το οποίο δεν ανήκει στο $\overline{K}$. Τότε, από το Θεώρημα 2.8, υπάρχει υπερεπίπεδο που διαχωρίζει γνήσια το x από το $\overline{K}$. Δηλαδή, υπάρχουν $0 \neq y \in \mathbb{R}^n$ και $\gamma \in \mathbb{R}$ τέτοια ώστε $\langle z, y\rangle < \gamma$ για κάθε $z \in \overline{K}$ και $\langle x, y\rangle > \gamma$.

Αφού $0 \in K$, ισχύει $\gamma > 0$. Έπεται ότι $\langle z, y/\gamma\rangle < 1$ για κάθε $z \in \overline{K}$, άρα $y/\gamma \in K^\circ$. Αλλά $\langle x, y/\gamma\rangle > 1$, το οποίο είναι άτοπο αφού $x \in K^{\circ\circ}$. □

Θεώρημα 3.15 *Έστω K κυρτό συμπαγές υποσύνολο του $\mathbb{R}^n$ με $0 \in K$. Τότε $0 \in \operatorname{int}(K^\circ)$ και $h_K = g_{K^\circ}$. Αντίστροφα, έστω L κυρτό κλειστό υποσύνολο του $\mathbb{R}^n$ με $0 \in \operatorname{int}(L)$. Τότε το L° είναι συμπαγές και ισχύει $h_{L^\circ} = g_L$.*

Απόδειξη: Έστω K όπως παραπάνω. Για να ισχύει $0 \in \operatorname{int}(K^\circ)$, αρκεί να υπάρχει $\varepsilon > 0$ τέτοιο ώστε: αν $\|y\| < \varepsilon$, τότε $\langle x, y\rangle \leq 1$ για κάθε $x \in K$. Ισχύει $\langle x, y\rangle \leq \|x\| \cdot \|y\|$, επομένως αρκεί να θέσουμε $\varepsilon = 1/M$, όπου $M = \max\{\|x\| : x \in K\}$.

Η h_K είναι κυρτή, θετικά ομογενής και $h_K \geq 0$, αφού $0 \in K$. Από το Θεώρημα 3.14, ισχύει $h_K = g_L$ άν

$$L = \{y \in \mathbb{R}^n : h_K(y) \leq 1\} = \{y \in \mathbb{R}^n : \langle x, y\rangle \leq 1 \ \forall x \in K\},$$

άρα $L = K^\circ$.

Αν τώρα L είναι σύνολο όπως στο δεύτερο μέρος του Θεωρήματος, υπάρχει $\varepsilon > 0$ με $D(0, \varepsilon) \subseteq L$. Αν $y \in L^\circ$, $y \neq 0$, ισχύει $\frac{\varepsilon}{2\|y\|} y \in D(0, \varepsilon)$, άρα $\langle y, \frac{\varepsilon}{2\|y\|} y\rangle \leq 1$, δηλαδή $\|y\| \leq 2/\varepsilon$. Άρα το L° είναι φραγμένο, επομένως συμπαγές. Από το πρώτο μέρος του Θεωρήματος έπεται ότι $h_{L^\circ} = g_{L^{\circ\circ}} = g_L$. □

Αν λοιπόν περιοριστούμε σε κυρτά συμπαγή σύνολα K με $0 \in \operatorname{int}(K)$, τα πολικά αυτών των συνόλων ανήκουν επίσης σε αυτήν την κατηγορία και $h_K = g_{K^\circ}$, $h_{K^\circ} = g_K$.

Παράδειγμα Έστω

$$K = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} |x_i| \leq 1\}.$$

Στην §1 είδαμε ότι $h_K(y) = \max\{|y_i| : i = 1, \ldots, n\}$. Άρα το πολικό του K είναι ο κύβος

$$L = \{y \in \mathbb{R}^n : h_K(y) \leq 1\} = [-1, 1]^n.$$

3.4 Υποδιαφορικά

§1. **Ορισμός** Έστω $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Θεωρούμε $x \in \mathbb{R}^n$ στο οποίο η f είναι πεπερασμένη.

(α) Ένα $z \in \mathbb{R}^n$ λέγεται *υποπαράγωγος* της f στο x αν

$$f(x+u) \geq f(x) + \langle u, z \rangle$$

για κάθε $u \in \mathbb{R}^n$.

(β) Το σύνολο όλων των υποπαραγώγων της f στο x λέγεται *υποδιαφορικό* της f στο x και συμβολίζεται με $\partial f(x)$. Το $\partial f(x)$ είναι κυρτό υποσύνολο του $\mathbb{R}^n$. Λέμε ότι η f είναι *υποδιαφορίσιμη* στο x αν $\partial f(x) \neq \emptyset$. Αν η f δεν παίρνει πεπερασμένη τιμή στο x, τότε θέτουμε $\partial f(x) = \emptyset$.

(γ) Το *υποδιαφορικό της* f είναι η συνολοσυνάρτηση $\partial f : x \mapsto \partial f(x)$.

(δ) Το *πεδίο ορισμού* $\mathrm{dom}(\partial f)$ *του* ∂f είναι το σύνολο $\{x \in \mathbb{R}^n : \partial f(x) \neq \emptyset\}$.

Παραδείγματα (α) Έστω $f : (a,b) \to \mathbb{R}$ κυρτή συνάρτηση. Απο την απόδειξη του Θεωρήματος 1.6, η f είναι υποδιαφορίσιμη σε κάθε $c \in (a,b)$ και $\partial f(c) = [f'_-(c), f'_+(c)]$.

(β) Ορίζουμε $f : \mathbb{R} \to \mathbb{R}$ θέτοντας $f(x) = -\sqrt{1-x^2}$ αν $|x| \leq 1$ και $f(x) = +\infty$ αλλιώς. Η f είναι κανονική κυρτή συνάρτηση και είναι διαφορίσιμη στο $(-1,1)$. Έχουμε $-1, 1 \in \mathrm{dom}(f)$ αλλά η f δεν είναι υποδιαφορίσιμη στα $-1, 1$.

(γ) Θεωρούμε την $f(x) = \|x\|$. Η f δεν είναι διαφορίσιμη στο 0, είναι όμως υποδιαφορίσιμη. Το $\partial f(0)$ αποτελείται από όλα τα $z \in \mathbb{R}^n$ για τα οποία

$$\|y\| \geq \langle y, z \rangle, \qquad y \in \mathbb{R}^n.$$

Έπεται ότι $\partial f(0) = \{z \in \mathbb{R}^n : \|z\| \leq 1\}$.

§2. Θα προσπαθήσουμε τώρα να δώσουμε γεωμετρική περιγραφή της έννοιας της υποδιαφορισιμότητας. Γράφουμε (x,a), όπου $x \in \mathbb{R}^n$ και $a \in \mathbb{R}$, για το τυχόν στοιχείο του $\mathbb{R}^{n+1}$. Τότε, κάθε υπερεπίπεδο H του $\mathbb{R}^{n+1}$ περιγράφεται από μία εξίσωση της μορφής

$$\langle x, x_0 \rangle + \lambda a_0 = b_0$$

όπου $(x_0, a_0) \neq (0,0)$ και $b_0 \in \mathbb{R}$.

Λέμε ότι το H είναι *κατακόρυφο* αν $a_0 = 0$ (οπότε $x_0 \neq 0$). Αν το H δεν είναι κατακόρυφο, τότε η εξίσωση του H παίρνει τη μορφή

$$\lambda = \langle x, -\frac{x_0}{a_0} \rangle + \frac{b_0}{a_0}$$

δηλαδή το H είναι το γράφημα της συνεχούς αφφινικής συνάρτησης

$$x \mapsto \langle x, -\frac{x_0}{a_0} \rangle + \frac{b_0}{a_0}.$$

Παρατήρηση Από την απόδειξη του Θεωρήματος 3.5 προκύπτει ότι: η συνάρτηση $f : A \to \mathbb{R}$ έχει στήριγμα στο σημείο $x \in A$ αν και μόνο αν υπάρχει μη κατακόρυφο επίπεδο στήριξης του $\mathrm{epi}(f)$ στο σημείο $(x, f(x))$.

Θεώρημα 3.16 *Έστω f κανονική κυρτή συνάρτηση στον $\mathbb{R}^n$ και $x_0 \in \mathrm{dom}(f)$. Η f είναι υποδιαφορίσιμη στο x_0 αν και μόνο αν υπάρχει μη-κατακόρυφο υπερεπίπεδο που στηρίζει το $\mathrm{epi}(f)$ στο $(x_0, f(x_0))$.*

Απόδειξη: Από τον ορισμό της υποπαραγώγου είναι φανερό ότι το $z \in \mathbb{R}^n$ είναι υποπαράγωγος της f στο x_0 αν και μόνο αν ισχύει

$$f(y) \geq f(x_0) + \langle y - x_0, z\rangle, \qquad y \in \mathbb{R}^n.$$

Αυτό είναι ισοδύναμο με το ότι η αφφινική συνάρτηση $T(y) = f(x_0) + \langle y - x_0, z\rangle$, $y \in \mathbb{R}^n$, είναι στήριγμα της f στο x_0. Το συμπέρασμα του θεωρήματος προκύπτει τώρα άμεσα από την προηγούμενη παρατήρηση. □

Παρατήρηση Το επιχείρημα που χρησιμοποιήθηκε στην απόδειξη δείχνει ότι υπάρχει 1-1 και επί αντιστοιχία μεταξύ του $\partial f(x_0)$ και του συνόλου των στηριγμάτων της f στο x_0.

Θεώρημα 3.17 *Έστω f κανονική κυρτή συνάρτηση στον $\mathbb{R}^n$. Τότε, η f είναι υποδιαφορίσιμη σε κάθε σημείο του $\mathrm{int}(\mathrm{dom}(f))$.*

Απόδειξη: Άμεση από τα Θεωρήματα 3.5 και 3.16. □

Πόρισμα *Κάθε κυρτή συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$ είναι υποδιαφορίσιμη.* □

§3. Περνάμε τώρα στην περιγραφή του υποδιαφορικού μιάς κυρτής συνάρτησης.

Θεώρημα 3.18 *Έστω $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ κανονική κυρτή συνάρτηση και $x_0 \in \mathrm{int}(\mathrm{dom} f)$. Τότε, $z \in \partial f(x_0)$ αν και μόνο αν*

$$f'(x_0; x) \geq \langle x, z\rangle$$

για κάθε $x \in \mathbb{R}^n$.

Απόδειξη: Έστω $z \in \partial f(x_0)$, δηλαδή $f(x_0 + u) \geq f(x_0) + \langle u, z\rangle$ για κάθε $u \in \mathbb{R}^n$. Ειδικότερα, αν $x \in \mathbb{R}^n$, ισχύει

$$f(x_0 + \lambda x) \geq f(x_0) + \langle \lambda x, z\rangle,$$

άρα

$$\frac{f(x_0 + \lambda x) - f(x_0)}{\lambda} \geq \langle x, z\rangle$$

για κάθε $\lambda > 0$. Επομένως,

$$f'(x_0; x) \geq \langle x, z\rangle.$$

Για το αντίστροφο, έστω $u \in \mathbb{R}^n$. Θεωρούμε το διάστημα

$$I = \{\lambda \in \mathbb{R} : x_0 + \lambda u \in \mathrm{dom} f\}$$

(είναι διάστημα, που περιέχει το 0 στο εσωτερικό του), και τη συνάρτηση

$$g(\lambda) = f(x_0 + \lambda u), \quad \lambda \in I.$$

Τότε,

$$g'_+(0) = f'(x_0; u) \quad , \quad g'_-(0) = -f'(x_0; -u).$$

Από την υπόθεση συμπεραίνουμε ότι

$$g'_+(0) \geq \langle z, u\rangle \geq g'_-(0),$$

δηλαδή $\langle z, u\rangle \in \partial g(0)$. Έπεται ότι $g(\lambda) \geq g(0) + \lambda\langle z, u\rangle$ για κάθε $\lambda \in I$, δηλαδή

$$f(x_0 + \lambda u) \geq f(x_0) + \langle z, \lambda u\rangle, \quad \lambda \in I.$$

Αν $1 \in I$, παίρνουμε $f(x_0 + u) \geq f(x_0) + \langle z, u\rangle$. Αν $1 \notin I$, ισχύει $f(x_0 + u) = +\infty$, οπότε πάλι ισχύει η ανισότητα. Άρα, $z \in \partial f(x_0)$. □

Παρατήρηση Οι υποθέσεις ότι η f είναι κανονική κυρτή και ότι το x_0 είναι εσωτερικό σημείο του $\mathrm{dom} f$ δεν είναι απαραίτητες στο παραπάνω θεώρημα. Χρησιμοποιήθηκαν μόνο στην απόδειξη της δεύτερης κατεύθυνσης, επειδή οι κυρτές συναρτήσεις μιάς μεταβλητής έχουν συζητηθεί μόνο στην περίπτωση που παίρνουν πραγματικές τιμές.

Αν υποτεθεί η ανισότητα $f'(x_0; x) \geq \langle x, z\rangle$ για κάθε $x \in \mathbb{R}^n$, αποδεικνύεται ότι η f δεν παίρνει την τιμή $-\infty$ (και αφού $f(x_0) \in \mathbb{R}$, είναι κανονική). Αν $f(x) = -\infty$ για κάποιο x, αποδεικνύεται ότι $f'(x_0; x - x_0) = -\infty$, το οποίο αντιφάσκει στην παραπάνω ανισότητα.

Αν το x_0 είναι συνοριακό σημείο του $\mathrm{dom}(f)$, η παραπάνω απόδειξη μεταφέρεται σε αυτήν την περίπτωση, αφού οι χρησιμοποιούμενες ιδιότητες της g μεταφέρονται εύκολα και στην περίπτωση που αυτή έχει τιμές στο $\overline{\mathbb{R}}$.

Θεώρημα 3.19 *Έστω $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ κυρτή συνάρτηση, $A = \{x \in \mathbb{R}^n : f(x) \in \mathbb{R}\}$, $x_0 \in \mathrm{int}(A)$. Τότε, η f είναι διαφορίσιμη στο x_0 αν και μόνο αν το $\partial f(x_0)$ είναι μονοσύνολο. Στην περίπτωση αυτή ισχύει*

$$\partial f(x_0) = \{\nabla f(x_0)\}.$$

Απόδειξη: Άμεση από το Θεώρημα 3.11 και την παρατήρηση που ακολουθεί το Θεώρημα 3.16. □

Λήμμα *Έστω $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ κανονική κυρτή συνάρτηση, και $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ κανονική κοίλη (δηλαδή η $-g$ είναι κανονική κυρτή) με $f \geq g$. Υποθέτουμε ακόμα ότι υπάρχει $x_0 \in \mathrm{dom}(f) \cap \mathrm{dom}(-g)$ στο οποίο η f είναι συνεχής. Τότε, υπάρχει αφφινική συνάρτηση $T : \mathbb{R}^n \to \mathbb{R}$ με $f \geq T \geq g$.*

Απόδειξη: Θεωρούμε τα σύνολα $K_1 = \mathrm{epi}(f)$ και

$$K_2 = \{(y, -t) : (y, t) \in \mathrm{epi}(-g)\} = \{(y, t) : t \leq g(y)\}.$$

Τα K_1, K_2 είναι κυρτά. Επιπλέον ισχύει $\mathrm{int}(K_1) \neq \emptyset$, γιατί η συνέχεια της f στο x_0 συνεπάγεται ότι $x_0 \in \mathrm{int}(\mathrm{dom}(f))$. Από αυτό έπεται ότι $\mathrm{int}(\mathrm{epi}(f)) \neq \emptyset$ (βλέπε απόδειξη του Θεωρήματος 3.5). Επίσης, $\mathrm{int}(K_1) \cap K_2 = \emptyset$. Γιατί, αν $(y, t) \in \mathrm{int}(K_1)$, τότε $t > f(y) \geq g(y)$, άρα $(y, t) \notin K_2$. Από το Θεώρημα 2.9, υπάρχει υπερεπίπεδο H

στον $\mathbb{R}^{n+1}$ που διαχωρίζει τα K_1, K_2. Δηλαδή υπάρχουν $(y_0, \alpha) \in \mathbb{R}^{n+1}$, $(y_0, \alpha) \neq (0,0)$ και $\gamma \in \mathbb{R}$ τέτοια ώστε

$$\langle y_0, y\rangle + \alpha t \geq \gamma \geq \langle y_0, x\rangle + \alpha s$$

για κάθε $(y,t) \in K_1$, $(x,s) \in K_2$.

Από την ανισότητα $\langle y_0, x_0\rangle + \alpha t \geq \gamma$ για κάθε $t \geq f(x_0)$, προκύπτει ότι $\alpha \geq 0$. Ισχύει όμως και $\alpha > 0$. Πράγματι, αν υποθέσουμε ότι $\alpha = 0$, τότε

$$\begin{aligned}\langle y_0, x_0\rangle &= \langle y_0, x_0\rangle + \alpha f(x_0)\\ &\geq \gamma \geq \langle y_0, x_0\rangle + \alpha f(x_0)\\ &= \langle y_0, x_0\rangle,\end{aligned}$$

άρα $\langle y_0, x_0\rangle = \gamma$. Από την άλλη πλευρά όμως, $\langle y_0, x\rangle \geq \gamma$ για κάθε $x \in \mathrm{dom}(f)$. Λόγω της συνέχειας της f στο x_0, υπάρχει $\varepsilon > 0$ με $D(x_0, \varepsilon) \subseteq \mathrm{dom}(f)$. Άρα, το γραμμικό συναρτησοειδές που ορίζει το y_0 στον $\mathbb{R}^n$ έχει ελάχιστη τιμή στο $D(x_0, \varepsilon)$ (την τιμή γ στη θέση x_0). Αυτό όμως είναι δυνατόν μόνο αν $y_0 = 0$ (βλέπε απόδειξη του Θεωρήματος 3.5). Αυτό πάλι είναι άτοπο, γιατί τότε $(y_0, \alpha) = (0,0)$.

Αφού $\alpha > 0$, ισχύει

$$\langle y_0/\alpha, y\rangle + t \geq \gamma/\alpha \geq \langle y_0/\alpha, x\rangle + s$$

για κάθε $(y,t) \in K_1$, $(x,s) \in K_2$, και ειδικότερα

$$\langle y_0/\alpha, y\rangle + f(y) \geq \gamma/\alpha$$

για κάθε $y \in \mathrm{dom}(f)$ και

$$\gamma/\alpha \geq \langle y_0/\alpha, x\rangle + g(x)$$

για κάθε $x \in \mathrm{dom}(g)$.Αλλά για $y \notin \mathrm{dom}(f)$, $x \notin \mathrm{dom}(g)$, οι ανισότητες ισχύουν προφανώς. Άρα ισχύει

$$\langle y_0/\alpha, y\rangle + f(y) \geq \gamma/\alpha \geq \langle y_0/\alpha, x\rangle + g(y)$$

για κάθε $y \in \mathbb{R}^n$. Άρα, η συνάρτηση $T(y) = (\gamma/\alpha) - \langle y_0/\alpha, y\rangle$, $y \in \mathbb{R}^n$, ικανοποιεί το συμπέρασμα του Λήμματος. □

Θεώρημα 3.20 *Έστω f_1, f_2 κανονικές κυρτές συναρτήσεις στον $\mathbb{R}^n$. Τότε,*

(α) $\partial(f_1+f_2)(x) \supseteq \partial f_1(x) + \partial f_2(x)$ για κάθε $x \in \mathbb{R}^n$.

(β) αν το $\mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$ περιέχει ένα σημείο στο οποίο η f_1 είναι συνεχής, τότε

$$\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x).$$

Απόδειξη: Αν $f_1(x) = +\infty$ ή $f_2(x) = +\infty$, τότε $\partial f_1(x) = \emptyset$ ή $\partial f_2(x) = \emptyset$ και $\partial(f_1+f_2)(x) = \emptyset$, οπότε το συμπέρασμα είναι προφανές. Θα περιοριστούμε λοιπόν στην περίπτωση $x \in \mathrm{dom}(f) \cap \mathrm{dom}(g)$.

(α) Έστω $z_1 \in \partial f_1(x)$, $z_2 \in \partial f_2(x)$. Τότε, $f_1(x+u) \geq f_1(x) + \langle z_1, u\rangle$ και $f_2(x+u) \geq f_2(x) + \langle z_2, u\rangle$ για κάθε $u \in \mathbb{R}^n$, άρα

$$(f_1+f_2)(x+u) \geq (f_1+f_2)(x) + \langle z_1+z_2, u\rangle.$$

(β) Έστω $z \in \partial(f_1+f_2)(x)$. Θα βρούμε $z_1 \in \partial f_1(x)$ και $z_2 \in \partial f_2(x)$ με $z_1+z_2=z$. Δηλαδή, θα βρούμε $z_1 \in \mathbb{R}^n$ τέτοιο ώστε

$$f_1(x+u) \geq f_1(x) + \langle u, z_1\rangle$$

και

$$f_2(x+u) \geq f_2(x) + \langle u, z-z_1\rangle$$

για κάθε $u \in \mathbb{R}^n$. Ισοδύναμα ζητάμε

$$f_1(x+u) - f_1(x) \geq \langle u, z_1\rangle \geq f_2(x) - f_2(x+u) + \langle u, z\rangle, \quad u \in \mathbb{R}^n.$$

Θεωρούμε τις συναρτήσεις

$$g_1(u) = f_1(x+u) - f_1(x), \quad u \in \mathbb{R}^n$$

και

$$g_2(u) = -f_2(x+u) + f_2(x) + \langle u, z\rangle, \quad u \in \mathbb{R}^n.$$

Η g_1 είναι κανονική κυρτή και η g_2 κανονική κοίλη (ως άθροισμα της αντίθετης μιάς κανονικής κυρτής και μιάς αφφινικής). Ισχύει $g_1 \geq g_2$, γιατί

$$g_1(u) - g_2(u) = (f_1+f_2)(x+u) - (f_1+f_2)(x) - \langle u, z\rangle$$

και $z \in \partial(f_1+f_2)(x)$. Από την υπόθεση στο (β), υπάρχει $x_0 \in \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$ στο οποίο η f_1 είναι συνεχής. Τότε, $x_0 - x \in \mathrm{dom}(g_1) \cap \mathrm{dom}(g_2)$ και η g_1 είναι συνεχής στο $x_0 - x$. Άρα ικανοποιούνται οι υποθέσεις του Λήμματος, οπότε υπάρχει αφφινική συνάρτηση $T : \mathbb{R}^n \to \mathbb{R}$ τέτοια ώστε $g_1 \geq T \geq g_2$. Για $u=0$ ισχύει

$$0 = g_1(0) \geq T(0) \geq g_2(0) = 0,$$

άρα $T(0)=0$, δηλαδή η T είναι γραμμική, άρα έχει τη μορφή $T(u) = \langle u, z_1\rangle$ για κάποιο $z_1 \in \mathbb{R}^n$.

Τότε παίρνουμε

$$\begin{aligned} f_1(x+u) - f_1(x) &= g_1(u) \geq \langle u, z_1\rangle \geq g_2(u) \\ &= f_2(x) - f_2(x+u) + \langle u, z\rangle, \end{aligned}$$

δηλαδή το z_1 ικανοποιεί τις ζητούμενες ανισότητες. □

3.5 Συζυγής συνάρτηση

§1. **Ορισμός** Έστω $f:\mathbb{R}^n \to \overline{\mathbb{R}}$. Η *συζυγής* (ή πολική) συνάρτηση της f είναι η $f^*:\mathbb{R}^n \to \overline{\mathbb{R}}$ που ορίζεται από την

$$f^*(y) = \sup_{x\in\mathbb{R}^n} \big[\langle x,y\rangle - f(x)\big].$$

Η *διπολική* συνάρτηση f^{**} της f είναι η $(f^*)^*$.

Παρατηρήσεις: (α) Αν η $f^*(y)$ είναι πεπερασμένη, τότε είναι ο μικρότερος πραγματικός αριθμός a για τον οποίο έχουμε

$$f(x) \geq \langle x,y\rangle - a$$

για κάθε $x\in\mathbb{R}^n$. (β) Η f είναι supremum αφφινικών συναρτήσεων, επομένως είναι κυρτή συνάρτηση.

§2. Βασικές ιδιότητες που προκύπτουν άμεσα από τον ορισμό είναι οι εξής:

(1) Αν $f,h:\mathbb{R}^n \to \overline{\mathbb{R}}$ και $f\leq h$, τότε $f^*\geq h^*$.

(2) $(+\infty)^* = -\infty$. Αντίστροφα, αν η f^* παίρνει την τιμή $-\infty$, τότε $f\equiv +\infty$.

(3) Αν η $f:\mathbb{R}^n \to \overline{\mathbb{R}}$ παίρνει την τιμή $-\infty$ σε κάποιο σημείο, τότε $f^*\equiv +\infty$. Ειδικότερα, $(-\infty)^* = +\infty$.

Παρατηρήστε ότι, από τα (2) και (3), η σχέση $f^{**} = f$ *δεν* είναι γενικά σωστή. Δεν είναι δύσκολο να δείτε ότι $f^{**}\leq f$ για κάθε $f:\mathbb{R}^n \to \overline{\mathbb{R}}$.

(4) Αν $\{f_i : i\in I\}$ είναι μία οικογένεια συναρτήσεων $f_i:\mathbb{R}^n \to \overline{\mathbb{R}}$, τότε

$$\left(\inf_{i\in I} f_i\right)^* = \sup_{i\in I} f_i^*$$

και

$$\left(\sup_{i\in I} f_i\right)^* \leq \inf_{i\in I} f_i^*.$$

Στην τελευταία ανισότητα, δεν έχουμε γενικά ισότητα.

(5) Αν $f:\mathbb{R}^n \to \overline{\mathbb{R}}$ και $\lambda > 0$, τότε

$$(\lambda f)^*(y) = \lambda f^*\left(\frac{y}{\lambda}\right).$$

(6) Αν $f:\mathbb{R}^n \to \overline{\mathbb{R}}$ και $a\in\mathbb{R}$, τότε

$$(f+a)^* = f^* - a.$$

(7) Αν $f:\mathbb{R}^n \to \overline{\mathbb{R}}$ και $x\in\mathbb{R}^n$, τότε

$$f_x^*(y) = f^*(y) + \langle x,y\rangle,$$

όπου f_x η συνάρτηση που ορίζεται από την $f_x(z) = f(z-x)$, $z\in\mathbb{R}^n$.

(8) Αν $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, τότε

$$\inf\{f(x) : x \in \mathbb{R}^n\} = -f^*(0).$$

Παραδείγματα: 1. Θεωρούμε την $f(x) = \frac{1}{2}\|x\|^2$. Τότε,

$$\begin{aligned} f^*(y) &= \sup_x \big[\langle x,y\rangle - \frac{1}{2}\|x\|^2\big] = \sup_{t\geq 0}\sup_{\|x\|=t}\big[\langle x,y\rangle - \frac{1}{2}\|x\|^2\big] \\ &= \sup_{t\geq 0}\left(t\sup_{\|x\|=1}\langle x,y\rangle - \frac{t^2}{2}\right) = \sup_{t\geq 0}\big[t\|y\| - \frac{t^2}{2}\big] \\ &= \frac{1}{2}\|y\|^2, \end{aligned}$$

δηλαδή $f^* = f$.

2. Έστω K υποσύνολο του $\mathbb{R}^n$ και δ_K η δείκτρια συνάρτηση του K (βλέπε 3.1, §4). Τότε, για κάθε $y \in \mathbb{R}^n$ ισχύει

$$\begin{aligned} \delta_K^*(y) &= \sup_{x\in\mathbb{R}^n}(\langle x,y\rangle - \delta_K(y)) = \sup_{x\in K}(\langle x,y\rangle - \delta_K(y)) \\ &= \sup_{x\in K}\langle x,y\rangle. \end{aligned}$$

Ειδικότερα, αν το K είναι κυρτό συμπαγές, τότε η δ_K^* συμπίπτει με τη συνάρτηση στήριξης h_K του K. Αν $K = \{a\}$, τότε $\delta_K^*(y) = \langle a,y\rangle$, δηλαδή σε αυτήν την περίπτωση η συζυγής συνάρτηση είναι γραμμική.

3. Έστω $f(x) = \langle a,x\rangle$, $x \in \mathbb{R}^n$, για κάποιο $a \in \mathbb{R}^n$. Τότε,

$$f^*(y) = \sup_{x\in\mathbb{R}^n}(\langle x,y\rangle - \langle a,x\rangle) = \sup_{x\in\mathbb{R}^n}\langle x, y-a\rangle.$$

Άρα, $f^*(y) = +\infty$ αν $y \neq a$ και $f^*(y) = 0$ αν $y = a$. Δηλαδή, $f^* = \delta_{\{a\}}$.

4. Έστω $f(x) = \|x\|$, $x \in \mathbb{R}^n$. Τότε,

$$f^*(y) = \sup_{x\in\mathbb{R}^n}(\langle x,y\rangle - \|x\|).$$

Αν $\|y\| \leq 1$, τότε $\langle x,y\rangle - \|x\| \leq 0$ για κάθε $x \in \mathbb{R}^n$. Επιπλέον, το 0 είναι τιμή της $\langle x,y\rangle - \|x\|$ (για $x = 0$). Άρα, $f^*(y) = 0$.

Αν $\|y\| > 1$, τότε για $\lambda > 0$ ισχύει

$$\langle \lambda y, y\rangle - \|\lambda y\| = \lambda(\|y\|^2 - \|y\|).$$

Άρα,

$$f^*(y) \geq \sup_{\lambda>0}\lambda(\|y\|^2 - \|y\|) = +\infty.$$

Συνεπώς, $f^* = \delta_{\overline{D(0,1)}}$.

Θεώρημα 3.21 *Έστω $f, g : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ κανονικές κυρτές συναρτήσεις. Τότε,*

$$(f\Box g)^* = f^* + g^*.$$

Απόδειξη: Αν $y \in \mathbb{R}^n$, ισχύει

$$\begin{aligned}
(f\Box g)^*(y) &= \sup_{x\in\mathbb{R}^n} (\langle x,y\rangle - (f\Box g)(x)) \\
&= \sup_{x\in\mathbb{R}^n} \left(\langle x,y\rangle - \inf_{z\in\mathbb{R}^n}(f(z)+g(x-z)\right) \\
&= \sup_{x\in\mathbb{R}^n}\sup_{z\in\mathbb{R}^n} (\langle x,y\rangle - f(z) - g(x-z)) \\
&= \sup_{x\in\mathbb{R}^n}\sup_{z\in\mathbb{R}^n} (\langle x-z,y\rangle - f(z) + \langle z,y\rangle - g(x-z)) \\
&= \sup_{x\in\mathbb{R}^n}\sup_{z\in\mathbb{R}^n : f(z)\neq+\infty} (\langle x-z,y\rangle - f(z) + \langle z,y\rangle - g(x-z)) \\
&= \sup_{z\in\mathbb{R}^n : f(z)\neq+\infty} \left(\langle z,y\rangle - f(z) + \sup_{x\in\mathbb{R}^n}[\langle x-z,y\rangle - g(x-z)]\right) \\
&= \sup_{z\in\mathbb{R}^n : f(z)\neq+\infty} \left(\langle z,y\rangle - f(z) + \sup_{u\in\mathbb{R}^n}[\langle u,y\rangle - g(u)]\right) \\
&= \sup_{z\in\mathbb{R}^n : f(z)\neq+\infty} (\langle z,y\rangle - f(z) + g^*(y)) \\
&= \sup_{z\in\mathbb{R}^n} (\langle z,y\rangle - f(z)) + g^*(y) = f^*(y)+g^*(y).
\end{aligned}$$

(Ο περιορισμός $f(z)\neq+\infty$ στον σχηματισμό μερικών supremum έχει γίνει για να αποκλεισθεί η περίπτωση να σχηματίζονται αθροίσματα στα οποία εμφανίζονται και το $+\infty$ και το $-\infty$. Ο περιορισμός δεν επηρεάζει την τιμή των supremum, αφού, αν $f(z)=+\infty$, το αντίστοιχο στοιχείο που εμφανίζεται στο σχηματισμό του supremum είναι το $-\infty$). $\Box$

Παράδειγμα Έστω K κυρτό, συμπαγές υποσύνολο του $\mathbb{R}^n$. Θεωρούμε τις συναρτήσεις $f(x)=d(x,K)$ και $g(x)=\|x\|$. Τότε, $f=g\Box\delta_K$ (βλέπε 3.1, §5). Άρα,

$$f^* = g^* + \delta_K^*,$$

δηλαδή $f^*(y)=h_K(y)$ αν $\|y\|\leq 1$, και $f^*(y)=+\infty$ αν $\|y\|>1$.

§3. Έστω $f:\mathbb{R}^n\to\overline{\mathbb{R}}$. Από τον ορισμό της συζυγούς συνάρτησης, για κάθε $x,y\in\mathbb{R}^n$ έχουμε

$$f^*(y)\geq\langle x,y\rangle - f(x)$$

δηλαδή,

$$f(x)+f^*(y)\geq\langle x,y\rangle, \tag{$*$}$$

όταν έχει νόημα το πρώτο μέλος. Η $(*)$ λέγεται *ανισότητα του Fenchel.*

Θεώρημα 3.22 *Έστω $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ και x ένα σημείο στο οποίο η f παίρνει πεπερασμένη τιμή. Τότε, $y\in\partial f(x)$ αν και μόνο αν*

$$f^*(y)=\langle x,y\rangle - f(x).$$

Απόδειξη: Η ισότητα στην ανισότητα του Fenchel ισοδυναμεί με το ότι

$$\langle x,y\rangle - f(x) \geq f^*(y) = \sup_{z\in\mathbb{R}^n}\left(\langle y,z\rangle - f(z)\right).$$

Αυτό είναι ισοδύναμο με το ότι

$$\langle x,y\rangle - f(x) \geq \langle y,z\rangle - f(z)$$

για κάθε $z\in\mathbb{R}^n$, δηλαδή

$$f(z) \geq f(x) + \langle y, z-x\rangle$$

για κάθε $z\in\mathbb{R}^n$. Αυτό όμως ισχύει ακριβώς όταν $y\in\partial f(x)$. $\square$

Το Θεώρημα 3.22 δείχνει ότι τα υποδιαφορικά της f στο x είναι ακριβώς εκείνα τα $y\in\mathbb{R}^n$ για τα οποία η ανισότητα του Fenchel γίνεται ισότητα.

3.6 Το θεώρημα δυϊσμού του Fenchel

Στον κυρτό προγραμματισμό, συχνά μετασχηματίζουμε ένα πρόβλημα ελαχιστοποίησης σε ένα πρόβλημα μεγιστοποίησης το οποίο ονομάζεται *δυϊκό πρόβλημα*. Υπάρχουν διάφοροι τρόποι με τους οποίους κατασκευάζονται δυϊκά προβλήματα. Ένας από αυτούς βασίζεται στο θεώρημα δυϊσμού του Fenchel.

Θυμηθείτε ότι μία συνάρτηση $g:\mathbb{R}^n\to\overline{\mathbb{R}}$ λέγεται κοίλη αν η $-g$ είναι κυρτή. Αν η g είναι κοίλη, ορίζουμε

$$\mathrm{dom}(g) = \{x\in\mathbb{R}^n : g(x) > -\infty\}.$$

Η g λέγεται *κανονική* αν η $-g$ είναι κανονική κυρτή συνάρτηση. Η *συζυγής κοίλη συνάρτηση της* g είναι η συνάρτηση $g_*:\mathbb{R}^n\to\overline{\mathbb{R}}$ που ορίζεται από την

$$g_*(y) = \inf_{x\in\mathbb{R}^n}\big\{\langle x,y\rangle - g(x)\big\}.$$

Η g_* είναι κοίλη συνάρτηση, και $g_*(y) = -(-g)^*(-y)$ για κάθε $y\in\mathbb{R}^n$.

Θεώρημα δυϊσμού του Fenchel *Εστω $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ κανονική κυρτή συνάρτηση και $g:\mathbb{R}^n\to\overline{\mathbb{R}}$ κανονική κοίλη συνάρτηση. Τότε, αν $\mathrm{int}(\mathrm{dom}(f))\cap\mathrm{int}(\mathrm{dom}(g))\neq\emptyset$ έχουμε*

$$\inf_{x\in\mathbb{R}^n}\big\{f(x)-g(x)\big\} = \sup_{y\in\mathbb{R}^n}\big\{g_*(y)-f^*(y)\big\}.$$

Απόδειξη: Παρατηρούμε ότι ισχύει $f(x)\neq-\infty$, $f^*(y)\neq-\infty$, $g(x)\neq+\infty$, $g_*(y)\neq+\infty$ για κάθε $x,y\in\mathbb{R}^n$. Ειδικότερα, οι διαφορές $f(x)-g(x)$ και $g_*(y)-f^*(y)$ ορίζονται για κάθε $x,y\in\mathbb{R}^n$.

Από την ανισότητα του Fenchel έχουμε

$$f(x)+f^*(y) \geq \langle x,y\rangle \geq g(x)+g_*(y)$$

άρα
$$f(x) - g(x) \geq g_*(y) - f^*(y)$$
για κάθε $x, y \in \mathbb{R}^n$. Έπεται ότι
$$\inf_{x\in\mathbb{R}^n}(f(x) - g(x)) \geq \sup_{y\in\mathbb{R}^n}(g_*(y) - f^*(y)).$$
Θέτουμε $m = \inf_{x\in\mathbb{R}^n}(f(x) - g(x))$ και θα αποδείξουμε ότι $m \leq \sup_{y\in\mathbb{R}^n}(g_*(y) - f^*(y))$.

Αφού $\mathrm{dom}(f) \cap \mathrm{dom}(g) \neq \emptyset$, ισχύει $m < +\infty$. Αν $m = -\infty$, η ανισότητα είναι προφανής. Έστω λοιπόν ότι $m \in \mathbb{R}$. Θεωρούμε τις συναρτήσεις f και $g + m$. Η f είναι κανονική κυρτή συνάρτηση, η g κανονική κοίλη και $f \geq g + m$.

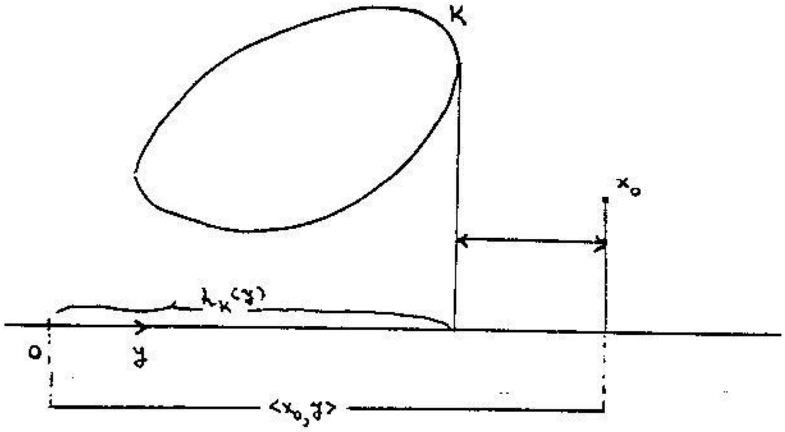
Αν $x_0 \in \mathrm{int}(\mathrm{dom}(f)) \cap \mathrm{int}(\mathrm{dom}(g))$, τότε η f είναι συνεχής στο x_0 (Θεώρημα 3.8). Άρα, ικανοποιούνται οι υποθέσεις του Λήμματος που προηγείται του Θεωρήματος 3.20. Συνεπώς, υπάρχει αφφινική συνάρτηση $T : \mathbb{R}^n \to \mathbb{R}$ με $f \geq T \geq g + m$. Δηλαδή υπάρχουν $z \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, έτσι ώστε $f(x) \geq \langle z, x\rangle + \alpha \geq g(x) + m$ για κάθε $x \in \mathbb{R}^n$. Τότε $-\alpha \geq \langle z, x\rangle - f(x)$ για κάθε $x \in \mathbb{R}^n$, δηλαδή
$$-\alpha \geq f^*(z).$$
Επίσης $\langle z, x\rangle - g(x) \geq m - \alpha$ για κάθε $x \in \mathbb{R}^n$, άρα
$$g_*(z) \geq m - \alpha.$$
Έπεται ότι $g_*(z) \geq f^*(z) + m$, άρα $\sup_{y\in\mathbb{R}^n}(g_*(y) - f^*(y)) \geq m$. □

Παράδειγμα: Έστω K κυρτό συμπαγές υποσύνολο του $\mathbb{R}^n$ με μη κενό εσωτερικό, και $x_0 \in \mathbb{R}^n$. Θεωρούμε τις συναρτήσεις $f(x) = \|x - x_0\|$, $g = -\delta_K$. Οι υποθέσεις του προηγούμενου θεωρήματος ικανοποιούνται φανερά. Ισχύει
$$\inf_{x\in\mathbb{R}^n}(f(x) - g(x)) = \inf_{x\in K}(f(x) - g(x)) = \inf_{x\in K} f(x) = d(x_0, K), \quad x \in \mathbb{R}^n.$$
Σύμφωνα με την ιδιότητα (7) της συζυγούς συνάρτησης και το παράδειγμα 4 (3.5, §2) ισχύει
$$f^*(y) = \delta_{\overline{D(0,1)}}(y) + \langle x_0, y\rangle = \begin{cases} \langle x_0, y\rangle & , \|y\| \leq 1 \\ +\infty & , \|y\| > 1 \end{cases}$$
και
$$g_*(y) = \inf_{x\in\mathbb{R}^n}(\langle x, y\rangle + \delta_K(y)) = \inf_{x\in K}\langle x, y\rangle.$$
Άρα, χρησιμοποιώντας και την συμμετρία του $\overline{D(0,1)}$ ως προς το 0, έχουμε
$$\begin{aligned}
\sup_{y\in\mathbb{R}^n}(g_*(y) - f^*(y)) &= \sup_{y\in\mathbb{R}^n}\left(\inf_{x\in K}\langle x, y\rangle - \delta_{\overline{D(0,1)}}(y) - \langle x_0, y\rangle\right) \\
&= \sup_{y\in\overline{D(0,1)}}\left(\inf_{x\in K}\langle x, y\rangle - \langle x_0, y\rangle\right) \\
&= \sup_{\|y\|\leq 1}\left(\langle x_0, y\rangle + \inf_{x\in K}\langle x, y\rangle\right) \\
&= \sup_{\|y\|\leq 1}(\langle x_0, y\rangle - h_K(y)).
\end{aligned}$$

Το θεώρημα του Fenchel δίνει λοιπόν την εξής ισότητα:

$$d(x_0, K) = \sup_{\|y\|\leq 1} (\langle x_0, y\rangle - h_K(y)).$$

Εύκολα επαληθεύεται ότι το supremum δεν αλλάζει αν θεωρήσουμε μόνον τα y με $\|y\| = 1$. Στην περίπτωση αυτή, η γεωμετρική σημασία των αριθμών πάνω από τους οποίους παίρνεται το supremum φαίνεται στο παρακάτω σχήμα.



3.7 Ασκήσεις

1. Έστω $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ κυρτή συνάρτηση. Αν το σύνολο $\mathrm{dom}(f)$ είναι ανοικτό, αποδείξτε ότι $f(x) > -\infty$ για κάθε $x \in \mathbb{R}^n$ ή $f(\mathbb{R}^n) \subseteq \{-\infty, +\infty\}$.

2. (α) Δίνεται συνάρτηση $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, για την οποία το σύνολο $\{(x, t) \in \mathbb{R}^{n+1} : t > f(x)\}$ είναι κυρτό. Αποδείξτε ότι η f είναι κυρτή.

(β) Έστω $f : \mathbb{R}^2 \to \mathbb{R}$ κυρτή συνάρτηση. Ορίζουμε τη συνάρτηση $g : \mathbb{R} \to \overline{\mathbb{R}}$ με τον τύπο

$$g(x) = \inf\{f(x, y) : y \in \mathbb{R}\}.$$

Αποδείξτε ότι η g είναι κυρτή.

3. Έστω $K \subseteq \mathbb{R}^{n+1}$ κυρτό. Για $x \in \mathbb{R}^n$ ορίζουμε

$$f(x) = \inf\{t \in \mathbb{R} : (x, t) \in K\}.$$

Αποδείξτε ότι η $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ είναι κυρτή συνάρτηση.

4. Βρείτε την $f \Box g$ στις παρακάτω περιπτώσεις:

(α) $f : \mathbb{R}^n \to \mathbb{R}$ κυρτή, $g = \delta_{\{x_0\}}$.

(β) Η f ορίζεται από την

$$f(x) = \begin{cases} -\log x & , x > 0 \\ \\ +\infty & , x \leq 0 \end{cases}$$

και η g από την

$$f(x) = \begin{cases} 0 & , x \leq 0 \\ \\ +\infty & , x > 0. \end{cases}$$

5. Αν $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ κανονικές κυρτές συναρτήσεις, αποδείξτε ότι

$$\mathrm{dom}(f \Box g) = \mathrm{dom}(f) + \mathrm{dom}(g).$$

6. Έστω K κυρτό και συμπαγές υποσύνολο του $\mathbb{R}^n$, $f : K \to \mathbb{R}$ κυρτή και συνεχής, και $\alpha = \max\{f(x) : x \in K\}$. Αποδείξτε ότι υπάρχει ακραίο σημείο x του K με $f(x) = \alpha$.

7. Αν $f : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση, αποδείξτε ότι

$$f'(x; y) \leq f(x + y) - f(x)$$

για κάθε $x, y \in \mathbb{R}^n$.

8. Έστω K κλειστό και κυρτό υποσύνολο του $\mathbb{R}^n$, και $f(x) = d(x, K)$. Έστω $x \in \mathbb{R}^n \setminus K$ και $y \in K$ με $d(x, K) = \|x - y\|$. Αποδείξτε ότι

$$\nabla f(x) = \frac{x - y}{\|x - y\|}.$$

Συμπεράνατε ότι η f είναι διαφορίσιμη στο $\mathbb{R}^n \setminus K$. [*Υπόδειξη:* Υποθέστε ότι $y = 0$. Το υπερεπίπεδο $H = \{z \in \mathbb{R}^n : \langle z, x\rangle = 0\}$ διαχωρίζει τα σύνολα K και $D(x, \|x\|)$. Για $\|h\| < \|x\|$, ισχύει

$$\left\langle x + h, \frac{x}{\|x\|}\right\rangle \leq d(x + h, K) \leq \|x + h\|.]$$

9. Έστω $f : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση, $x \in \mathbb{R}^n$. Είναι γνωστό ότι η συνάρτηση $g_x(y) = f'(x; y)$, $y \in \mathbb{R}^n$, είναι κυρτή. Άρα, για κάθε $z \in \mathbb{R}^n$ υπάρχει η πλευρική παράγωγος $g_x'(y; z)$. Ορίζουμε

$$f''(x; y; z) = g_x'(y; z).$$

Αποδείξτε ότι:
(α) $f''(x; y; y) = f'(x; y)$.
(β) $f''(x; y; -y) = -f'(x; y)$.

10. Θεωρούμε τη συνάρτηση

$$f(x) = -x_1^{a_1} \dots x_n^{a_n}, \qquad x \in \mathbb{R}^n,\ \ x_1 > 0, \dots, x_n > 0,$$

όπου $a_i > 0$, $i = 1, 2, \ldots, n$.

(α) Αποδείξτε ότι η f είναι κυρτή αν και μόνο αν $a_1 + \ldots + a_n \leq 1$. [*Υπόδειξη:* Χρησιμοποιήστε την Εσσιανή.]

(β) Αποδείξτε ότι τα σύνολα στάθμης της f είναι κυρτά.

11. Βρείτε τις συναρτήσεις στήριξης των παρακάτω συνόλων:

$$K = \{x \in \mathbb{R}^n : x_i \geq 0, i = 1, \ldots, n, x_1 + \ldots + x_n = 1\}.$$

$$L = \{(x, y) \in \mathbb{R}^2 : (x/a)^2 + (y/b)^2 \leq 1\}, \quad a, b > 0.$$

12. Βρείτε τις συναρτήσεις στάθμης των παρακάτω συνόλων:

$$K = \{(x, y) \in \mathbb{R}^2 : (x/a)^2 + (y/b)^2 \leq 1\}, \quad a, b > 0.$$

$$L = \{(x, y) \in \mathbb{R}^2 : |y| \leq 1\}.$$

$$M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq 1, |z| \leq 1\}.$$

13. Αν $K, L \subseteq \mathbb{R}^n$ κυρτά συμπαγή και $C = \mathrm{co}(K \cup L)$, αποδείξτε ότι

$$h_C(x) = \max\{h_K(x), h_L(x)\}, \quad x \in \mathbb{R}^n.$$

14. Αν $K, L \subseteq \mathbb{R}^n$ κυρτά κλειστά με $0 \in \mathrm{int}(K) \cap \mathrm{int}(L)$, αποδείξτε ότι

$$g_{K \cap L}(x) = \max\{g_K(x), g_L(x)\}, \quad x \in \mathbb{R}^n.$$

15. Έστω $h : \mathbb{R}^n \to \mathbb{R}$ κυρτή, θετικά ομογενής. Αποδείξτε ότι:

(α) $h'(u, u) = h(u)$ και $h'(u; -u) = -h(u)$.

(β) $h'(u; y) \leq h(y)$.

16. Έστω $K \subseteq \mathbb{R}^n$ κυρτό συμπαγές, και $u \in \mathbb{R}^n$, $u \neq 0$. Αν $L = \{x \in K : h_K(u) = \langle x, u \rangle\}$, αποδείξτε ότι:

(α) $h'_K(u; y) = h_L(y)$, $y \in \mathbb{R}^n$. [*Υπόδειξη:* Αν $x \in L$, χρησιμοποιώντας τη σχέση

$$\langle x, y \rangle = \frac{\langle x, u + \lambda y \rangle - \langle x, u \rangle}{\lambda}$$

συμπεράνατε ότι $h_L(y) \leq h'_K(u; y)$.

Αντίστροφα, αν $\langle x, y \rangle \leq h'_K(u; y)$, $y \in \mathbb{R}^n$, χρησιμοποιώντας την Άσκηση 15 συμπεράνατε ότι $x \in L$.]

(β) Η h_K είναι διαφορίσιμη στο u αν και μόνο αν το L είναι μονοσύνολο.

17. Αν $K \subseteq \mathbb{R}^n$ κυρτό συμπαγές, αποδείξτε ότι η h_K είναι γραμμική αν και μόνο αν το K είναι μονοσύνολο.

18. Αν $K \subseteq \mathbb{R}^n$ κυρτό συμπαγές, αποδείξτε ότι $\partial h_K(0) = K$.

19. Θεωρούμε τη συνάρτηση $f : \mathbb{R}^2 \to \mathbb{R}$ με $f(x,y) = |x| + |y|$. Αποδείξτε ότι

$$\partial f(x,y) = \begin{cases} \{(\mathrm{sgn}x, \mathrm{sgn}y)\} & , x \neq 0,\ y \neq 0 \\ [-1,1] \times [-1,1] & , x = 0,\ y = 0 \\ \{\mathrm{sgn}x\} \times [-1,1] & , x \neq 0,\ y = 0 \\ [-1,1] \times \{\mathrm{sgn}y\} & , x = 0,\ y \neq 0. \end{cases}$$

20. Έστω $f : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση. Αν $z_1 \in \partial f(x_1)$ και $z_2 \in \partial f(x_2)$, δείξτε ότι $\langle z_1 - z_2, x_1 - x_2 \rangle \geq 0$. (Δηλαδή, το υποδιαφορικό είναι «αύξουσα συνάρτηση».)

21. Έστω $f : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση. Αποδείξτε ότι:

(α) Για κάθε $x_0 \in \mathbb{R}^n$ το $\partial f(x_0)$ είναι συμπαγές. [*Υπόδειξη:* Θεωρήστε τις ανισότητες

$$f(x_0 + x) \geq f(x_0) + \langle z, x \rangle, \quad z \in \partial f(x_0),\ x \in \mathbb{R}^n,$$

και βρείτε φράγματα των $\langle z, e_i \rangle$.]

(β) $h_{\partial f(x_0)}(x) = f'(x_0; x), \quad x \in \mathbb{R}^n$. [*Υπόδειξη:* Η συνάρτηση του δευτέρου μέλους είναι κυρτή και θετικά ομογενής, άρα είναι η συνάρτηση στήριξης κάποιου κυρτού συμπαγούς συνόλου.]

22. Θεωρούμε τα σύνολα

$$A = \{(x,y) \in \mathbb{R}^2 : y > 0\} \cup \{(x,y) \in \mathbb{R}^2 : y = 0, x \geq 0\}$$

και

$$B = \{(x,y) \in \mathbb{R}^2 : y < 0\} \cup \{(x,y) \in \mathbb{R}^2 : y = 0, x \leq 0\},$$

και τις συναρτήσεις $f_1 = \delta_A$ και $f_2 = \delta_B$. Αποδείξτε ότι:

(α) $\partial f_1(0) = \{(0,t) : t \leq 0\}$.

(β) $\partial f_2(0) = \{(0,t) : t \geq 0\}$.

(γ) $f_1 + f_2 = \delta_{\{0\}}$.

(δ) $\partial(f_1 + f_2)(0) \neq \partial f_1(0) + \partial f_2(0)$.

23. Έστω $o : \mathbb{R} \to \overline{\mathbb{R}}$ άρτια συνάρτηση. Ορίζουμε τις συναρτήσεις $f(x) = o(\|x\|)$ και $g(x) = o^*(\|x\|)$, $x \in \mathbb{R}^n$. Αποδείξτε ότι $f^* = g$.

24. Έστω $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ για την οποία ισχύει $f = f^*$. Αποδείξτε ότι $f(x) = \frac{1}{2}\|x\|^2$, $x \in \mathbb{R}^n$.

25. Έστω $K \subseteq \mathbb{R}^n$ κυρτό συμπαγές.

(α) Έστω $y \notin K$, $z_0 \in K$ με $\|z_0 - y\| = d(y, K)$. Τότε, για κάθε $z \in K$ ισχύει $\langle z_0 - y, z - y \rangle > 0$. [*Υπόδειξη:* Αν $\langle z_0 - y, z - y \rangle \leq 0$, τα δύο διανύσματα σχηματίζουν αμβλεία ή ορθή γωνία, οπότε υπάρχει σημείο x στο ευθύγραμμο τμήμα με άκρα τα z, z_0 έτσι ώστε $\|x - y\| < \|z_0 - y\|$.]

(β) Αποδείξτε ότι $h_K^* = \delta_K$ (συνεπώς, $\delta_K^{**} = \delta_K$).

26. Αν K κώνος στον $\mathbb{R}^n$, αποδείξτε ότι $\delta_K^* = \delta_{K^\circ}$.

27. Έστω $f : \mathbb{R}^n \to \mathbb{R}$ και $x \in \mathbb{R}^n$ με $\partial f(x) \neq \emptyset$. Δείξτε ότι $f^{**}(x) = f(x)$.

Κεφάλαιο 4

Εφαρμογές

4.1 Θεωρία παιγνίων

Δύο παίκτες, ο A και ο B, έχουν μεγάλα αποθέματα από δραχμές, τάληρα και δεκάρικα. Αποφασίζουν να παίξουν το εξής παιχνίδι: σε κάθε γύρο, καθένας από τους δύο επιλέγει ένα κέρμα. Αν το άθροισμα των αξιών των δύο κερμάτων είναι περιττός αριθμός, τότε ο A παίρνει το κέρμα του B. Αν είναι άρτιος, τότε ο B παίρνει το κέρμα του A. Το παιχνίδι επαναλαμβάνεται με τους ίδιους όρους και κάθε παίκτης προσπαθεί να μεγιστοποιήσει τα κέρδη (ή να ελαχιστοποιήσει τις απώλειές του). Το ερώτημα είναι ποιά πρέπει να είναι η τακτική καθενός από τους παίκτες.

Για το σκοπό αυτό θεωρούμε έναν πίνακα, τον *πίνακα αποδόσεων* P, ο οποίος δίνει τα κέρδη του A για τις $3 \times 3 = 9$ πιθανές εκβάσεις του παιχνιδιού.

P:

A \ B	δ	τ	Δ
δ	−1	−1	10
τ	−5	−5	10
Δ	1	5	−10

Ο παίκτης A, βλέποντας την απόδοση 10 της πρώτης γραμμής για την οποία χρειάζεται να ρισκάρει μόνο μία δραχμή, αποφασίζει να παίζει πάντα δραχμές. Ο παίκτης B σκέφτεται αρχικά να παίζει δεκάρικα (τον συγκινεί το -10 της τρίτης στήλης), παρατηρώντας όμως την τακτική του A (χάνοντας δηλαδή ένα δεκάρικο κάθε φορά) αρχίζει κι αυτός να παίζει δραχμές. Ο παίκτης A αντιλαμβάνεται την αλλαγή και απαντά με τα δεκάρικά του. Έτσι συνεχίζεται το παιχνίδι. Μπορεί κάποιος από τους δύο παίκτες να αναπτύξει κάποια νικητήρια στρατηγική;

Πριν απαντήσουμε, ας θεωρήσουμε ένα άλλο παιχνίδι με πίνακα αποδόσεων για τον παίκτη A τον παρακάτω πίνακα Q.

Q:

A \ B	δ	τ	Δ
δ	10	0	−10
τ	0	−5	5
Δ	5	1	5

Ας υποθέσουμε ότι οι δύο παίκτες παίζουν αυτό το παιχνίδι. Ο παίκτης A, όντας συντηρητικός, αποφασίζει να παίξει τα δεκάρικά του. Η σκέψη του είναι η εξής: ότι κι αν κάνει ο B, θα κερδίζω πάντα τουλάχιστον μία δραχμή, ενώ αν έπαιζα οτιδήποτε άλλο θα μπορούσε και να χάσω. Ο παίκτης B, που είναι της αυτής νοοτροπίας, αποφασίζει να παίζει τα τάληρά του. Καταλαβαίνει ότι αυτό δεν εγγυάται πως θα νικήσει, αλλά παίζοντας έτσι ελαχιστοποιεί τις απώλειές του. Πώς σκέφτηκαν οι δύο παίκτες; Ο παίκτης A υπολόγισε το ελάχιστο κάθε γραμμής (το χειρότερο που θα μπορούσε να του συμβεί) και επέλεξε τη γραμμή για την οποία αυτό το ελάχιστο είναι μέγιστο. Όμοια, ο B υπολόγισε το μέγιστο κάθε στήλης (το χειρότερο που θα μπορούσε να του συμβεί) και επέλεξε τη στήλη για την οποία αυτό το μέγιστο είναι ελάχιστο. Τώρα η κατάσταση είναι ευσταθής. Ο A θα συνεχίσει να παίζει τα δεκάρικά του και ο B τα τάληρά του. Ο B θα συνεχίσει να χάνει, δεν θα μπορούσε όμως να κάνει κάτι διαφορετικό για να βελτιώσει τη θέση του. Η μοίρα του είναι προδιαγεγραμμένη, εκτός αν ο A χάσει τα λογικά του και μεταπηδήσει σε κάποια άλλη στρατηγική.

Η κατάσταση στο πρώτο παιχνίδι είναι τελείως διαφορετική. Έστω $Q(i,j)$ η ij-συντεταγμένη του πίνακα Q. Παρατηρήστε ότι η σταθερή θέση (3,2) του πίνακα Q έχει την ιδιότητα να είναι η μικρότερη συντεταγμένη της γραμμής της και η μεγαλύτερη συντεταγμένη της στήλης της. Είναι, όπως λέμε, ένα *σαγματικό σημείο* για τον Q. Παρατηρήστε ακόμα ότι

$$\max_i \min_j Q(i,j) = \min_j \max_i Q(i,j) = 1.$$

Από την άλλη πλευρά, ο P δεν έχει σαγματικά σημεία και, επιπλέον,

$$-1 = \max_i \min_j P(i,j) < \min_j \max_i P(i,j) = 1.$$

Οι παρατηρήσεις που κάναμε μελετώντας αυτά τα δύο απλά παιχνίδια έχουν σημασία σε ένα πολύ ευρύτερο πλαίσιο. Η ύπαρξη σαγματικού σημείου συνδέεται στενά με την ισότητα του maximin και του minimax.

Θεώρημα 4.1 *Έστω A και B τυχόντα σύνολα και $K : A \times B \to \mathbb{R}$. Τότε,*

$$\sup_{x \in A} \inf_{y \in B} K(x,y) \leq \inf_{y \in B} \sup_{x \in A} K(x,y). \tag{1}$$

Αν υπάρχει $(x_0, y_0) \in A \times B$ με την ιδιότητα

$$K(x, y_0) \leq K(x_0, y_0) \leq K(x_0, y) \tag{2}$$

για κάθε $(x,y) \in A \times B$, τότε η (1) είναι ισότητα και τα δύο μέλη της είναι ίσα με $K(x_0, y_0)$.

Απόδειξη: Για κάθε $z \in A$, $y \in B$ ισχύει

$$K(z,y) \leq \sup_x K(x,y).$$

Άρα,

$$\inf_{y \in B} K(z,y) \leq \inf_{y \in B} \sup_x K(x,y)$$

για κάθε $z \in A$. Άρα,

$$\sup_{x \in A} \inf_{y \in B} K(x,y) \leq \inf_{y \in B} \sup_{x \in A} K(x,y).$$

Αυτό αποδεικνύει την (1). Έστω τώρα (x_0, y_0) που ικανοποιεί την $K(x, y_0) \leq K(x_0, y_0) \leq K(x_0, y)$ για κάθε $(x,y) \in A \times B$. Τότε,

$$\sup_{x \in A} K(x, y_0) \leq K(x_0, y_0) \leq \inf_{y \in B} K(x_0, y).$$

Αφού

$$\inf_{y \in B} \sup_{x \in A} K(x,y) \leq \sup_{x \in A} K(x, y_0)$$

και

$$\inf_{y \in B} K(x_0, y) \leq \sup_{x \in A} \inf_{y \in B} K(x,y),$$

έπεται το ζητούμενο. □

Ένα σημείο (x_0, y_0) που ικανοποιεί την (2) λέγεται *σαγματικό σημείο* της K στο $A \times B$. Το Θεώρημα 4.1 δεν μάς δίνει κάποιον τρόπο για να διαπιστώσουμε αν υπάρχουν σαγματικά σημεία. Είναι μάλιστα πολύ δύσκολο να αποδείξουμε την ύπαρξη σαγματικού σημείου χωρίς περιορισμούς για την συνάρτηση K και τα σύνολα A και B. Ένα τέτοιο αποτέλεσμα είναι το minimax θεώρημα του von Neumann.

Θεώρημα 4.2 *Έστω $A \subset \mathbb{R}^m$ και $B \subset \mathbb{R}^n$ μη κενά, συμπαγή και κυρτά σύνολα, και έστω $K : A \times B \to \mathbb{R}$ συνεχής συνάρτηση. Αν για κάθε $y \in B$ η $K(x,y)$ είναι κοίλη συνάρτηση στο A και για κάθε $x \in A$ η $K(x,y)$ είναι κυρτή συνάρτηση στο B, τότε υπάρχει σαγματικό σημείο (x', y') για την K στο $A \times B$. Δηλαδή,*

$$\max_{x \in A} \min_{y \in B} K(x,y) = \min_{y \in B} \max_{x \in A} K(x,y) = K(x', y').$$

Απόδειξη: Κάνουμε πρώτα την επιπλέον υπόθεση ότι η $K(x,y)$ είναι γνήσια κυρτή ως προς y και γνήσια κοίλη ως προς x (για κάθε $x \in A$ και $y \in B$ αντίστοιχα). Ορίζουμε

$$g(y) = \max_{x\in A} K(x,y).$$

Σύμφωνα με τις υποθέσεις μας, υπάρχει μοναδικό σημείο $h(y) \in A$ στο οποίο πιάνεται αυτό το maximum (υπάρχει λόγω συμπάγειας, και είναι μοναδικό γιατί η $x \mapsto K(x,y)$ είναι γνήσια κοίλη).

Η συνάρτηση $h : B \to A$ είναι συνεχής στο B. Αλλιώς, θα υπήρχαν $y_0 \in B$, $\varepsilon > 0$, και μία ακολουθία $\{y_j\}$ στο B τέτοια ώστε $y_j \to y_0$ και $\|h(y_j) - h(y_0)\| > \varepsilon$ για κάθε j. Αν $x_j = h(y_j)$, $j = 0, 1, 2, \ldots$, από το γεγονός ότι η $K(x, y_0)$ είναι γνήσια κοίλη συμπεραίνουμε ότι υπάρχει $\delta > 0$ με την ιδιότητα

$$K(x_0, y_0) - K(x_j, y_j) > \delta, \quad j = 1, 2, \ldots.$$

Από την άλλη πλευρά,

$$K(x_0, y_0) - K(x_i, y_0) = [K(x_0, y_0) - K(x_0, y_i)] + [K(x_0, y_i) - K(x_i, y_0)]$$

και αφού $K(x_i, y_i)$ είναι η μέγιστη τιμή της $K(x, y_i)$ στο A, το δεξιό μέλος είναι μικρότερο από

$$[K(x_0, y_0) - K(x_0, y_i)] + [k(x_i, y_i) - K(x_i, y_0)].$$

Η K είναι ομοιόμορφα συνεχής στο συμπαγές $A \times B$, άρα για μεγάλα j οι δύο αυτοί όροι θα είναι μικρότεροι από $\delta/2$. Αυτό οδηγεί σε άτοπο.

Η συνάρτηση $g(y) = K(h(y), y)$ είναι συνεχής στο B ως σύνθεση συνεχών συναρτήσεων. Έστω y' το σημείο στο οποίο η g παίρνει την ελάχιστη τιμή της στο B. Δηλαδή,

$$g(y') = \min_{y\in B} g(y) = \min_{y\in B} \max_{x\in A} K(x,y).$$

Για κάθε $y \in B$ και κάθε $\lambda \in (0, 1)$, θέτουμε $z = (1-\lambda)y' + \lambda y$ και $x' = h(y')$, $w = h(z)$. Από τον ορισμό του y' έχουμε $g(y') \leq g(z)$, ή ισοδύναμα, $K(x', y') \leq K(w, z)$. Από την κυρτότητα της $K(x, y)$ ως προς y βλέπουμε ότι

$$\begin{aligned} K(x',y') &\leq K(w,z) \leq (1-\lambda)K(w,y') + \lambda K(w,y) \\ &\leq (1-\lambda)K(x',y') + \lambda K(w,y), \end{aligned}$$

άρα $K(x', y') \leq K(w, y)$. Αφήνοντας το $\lambda \to 0^+$, έχουμε $z \to y'$ και από την συνέχεια της h, $w \to x'$. Άρα,

$$K(x',y') \leq K(x',y),$$

και αφού $K(x, y') \leq K(x', y')$, καταλήγουμε στην

$$K(x,y') \leq K(x',y') \leq K(x',y).$$

Δηλαδή, το (x', y') είναι σαγματικό σημείο για την K.

Αν η K δεν υποτεθεί γνήσια κυρτή και κοίλη ως προς y και x αντίστοιχα, για $\varepsilon > 0$ μικρό ορίζουμε

$$K_\varepsilon(x,y) = K(x,y) - \varepsilon\|x\| + \varepsilon\|y\|,$$

οποτε κάθε K_ε ικανοποιεί τις επιπλέον υποθέσεις μας και επομένως έχει σαγματικό σημείο $(x_\varepsilon, y_\varepsilon) \in A \times B$. Χρησιμοποιώντας την συμπάγεια του $A \times B$ και παίρνοντας ακολουθία $\varepsilon_j \to 0$, βρίσκουμε υπακολουθία της $(x_{\varepsilon_j}, y_{\varepsilon_j})$ που συγκλίνει σε σαγματικό σημείο της K (άσκηση).

Τέλος, ο ισχυρισμός για την τελευταία ισότητα στην διατύπωση του θεωρήματος προκύπτει από το Θεώρημα 4.1 και την παρατήρηση ότι, στην περίπτωσή μας, τα inf και sup είναι min και max αντίστοιχα (άσκηση). □

Επιστρέφουμε στο παιχνίδι με τις δραχμές, τα τάληρα και τα δεκάρικα. Ο πίνακας αποδόσεων P του παιχνιδιού δεν είχε σαγματικά σημεία. Ονομάζουμε τις πιθανές επιλογές του κάθε παίκτη (δραχμή, τάληρο ή δεκάρικο) *αμιγείς στρατηγικές*. Κανείς από τους δύο παίκτες δεν έχει αμιγή στρατηγική που να του εξασφαλίζει τη νίκη σε κάθε γύρο. Θα μπορούσε όμως ο Α να σκεφτεί μία ακολουθία αμιγών στρατηγικών που θα μπέρδευε τον Β τόσο πολύ ώστε ο Α να κερδίζει μετά από πολλές επαναλήψεις του παιχνιδιού. Αυτή η ακολουθία δεν θα έπρεπε να ακολουθεί κάποιο κανόνα, γιατί ο Β θα απαντούσε με κατάλληλη ακολουθία δικών του στρατηγικών. Η δυνατότητα που απομένει στον Α είναι να επιλέγει *τυχαία* τις στρατηγικές του, αποφασίζοντας μόνο με ποιά πιθανότητα θα επιλέγει κάθε αμιγή στρατηγική του. Υπάρχει κάποιος βέλτιστος τρόπος για να το κάνει;

Μπορούμε να κοιτάξουμε αυτό το πρόβλημα πιό γενικά. Ας υποθέσουμε ότι οι παίκτες Α και Β έχουν m και n *αμιγείς στρατηγικές* αντίστοιχα και ότι μάς έχει δοθεί ο $m \times n$ πίνακας αποδόσεων $P = (p_{ij})$. Μία *μεικτή στρατηγική* για τον Α είναι ένα «διάνυσμα πιθανοτήτων» $x = (x_1, \ldots, x_m)$, δηλαδή ένα διάνυσμα x τέτοιο ώστε $x_i \geq 0$ και $\sum_{i=1}^m x_i = 1$. Με τον ορισμό αυτό εννοούμε ότι ο Α θα χρησιμοποιεί την i-στή στρατηγική με πιθανότητα x_i. Όμοια, το $y = (y_1, \ldots, y_n)$ είναι μία μεικτή στρατηγική του Β αν $y_j \geq 0$ και $\sum_{j=1}^n y_j = 1$. Από την στοιχειώδη θεωρία πιθανοτήτων, η *συνάρτηση αναμενόμενης απόδοσης* ορίζεται από την

$$E(x,y) = \sum_{i=1}^m \sum_{j=1}^n x_i p_{ij} y_j.$$

Ορίζουμε

$$f(x) = \min_{y \in B} E(x,y), \qquad g(y) = \max_{x \in A} E(x,y)$$

και ονομάζουμε τα x' και y' *βέλτιστες μεικτές στρατηγικές* των Α και Β αν

$$f(x') = \max_{x \in A} f(x) = \max_{x \in A} \min_{y \in B} E(x,y)$$
$$g(y') = \min_{y \in B} g(y) = \min_{y \in B} \max_{x \in A} E(x,y).$$

Το φυσιολογικό ερώτημα είναι τώρα το εξής: υπάρχουν βέλτιστες μεικτές στρατηγικές για τον Α και/ή τον Β; Το θεμελιώδες θεώρημα για τα παιχνίδια αυτού του είδους μάς δίνει καταφατική απάντηση.

Θεώρημα 4.3 *Σε κάθε παιχνίδι με δύο παίκτες όπως το παραπάνω, υπάρχουν βέλτιστες μεικτές στρατηγικές και για τους δύο παίκτες. Επιπλέον, αν x' και y' είναι βέλτιστες μεικτές στρατηγικές για τους παίκτες A και B, τότε*

$$f(x') = \max_{x\in A}\min_{y\in B} E(x,y) = \min_{y\in B}\max_{x\in A} E(x,y) = g(y')$$

και το (x', y') είναι σαγματικό σημείο για την $E(x,y)$.

Απόδειξη: Εφαρμόζουμε το Θεώρημα 4.2 με $K(x,y) = E(x,y)$ και

$$A = \{x \in \mathbb{R}^m : x_i \geq 0, \sum_{i=1}^{m} x_i = 1\},$$

$$B = \{y \in \mathbb{R}^n : y_i \geq 0, \sum_{i=1}^{n} y_i = 1\}.$$

Η E είναι γραμμική ως προς x και y, και τα A, B είναι συμπαγή και κυρτά (άσκηση). Άρα, μπορούμε να εφαρμόσουμε το Θεώρημα 4.2 και να βρούμε σαγματικό σημείο (x', y') της E στο $A \times B$. Όμως τότε, η $f(x')$ είναι μέγιστη τιμή της f και η $g(y')$ είναι ελάχιστη τιμή της g. Αυτό αποδεικνύει τον πρώτο ισχυρισμό του θεωρήματος.

Έστω x' και y' βέλτιστες μεικτές στρατηγικές των A και B. Η απόδειξη του Θεωρήματος 4.2 δείχνει ότι υπάρχει μεικτή στρατηγική x_0 τέτοια ώστε το (x_0, y') να είναι σαγματικό σημείο για την $E(x,y)$, και από αυτό έπεται ότι $f(x_0) = g(y')$. Όμως $f(x') \leq g(y')$ από την (1) στο Θεώρημα 4.1, και αν είχαμε γνήσια ανισότητα, τότε θα παίρναμε $f(x') < f(x_0)$, δηλαδή η x' δεν θα ήταν βέλτιστη. Άρα, $f(x') = g(y')$. Αφήνουμε σαν άσκηση να δείξετε ότι το (x', y') είναι σαγματικό σημείο για την $E(x,y)$ (Άσκηση 4). $\square$

Η κοινή τιμή $v' = f(x') = g(y')$ είναι η *τιμή* του παιχνιδιού. Αν $v' = 0$, το παιχνίδι είναι δίκαιο. Αν $v' > 0$, μεροληπτεί υπέρ του A, ενώ αν $v' < 0$ μεροληπτεί υπέρ του B.

Ας επιστρέψουμε στα παραδείγματά μας. Έχουν και τα δύο βέλτιστες στρατηγικές για κάθε παίκτη. Δεν θα ασχοληθούμε με την εύρεση αυτών των βέλτιστων στρατηγικών. Για το πρώτο παιχνίδι έχουμε $x' = (1/2, 0, 1/2)$ και $y' = (10/11, 0, 1/11)$, ενώ για το δεύτερο έχουμε $x' = (0, 0, 1)$ και $y' = (0, 1, 0)$. Για το πρώτο παιχνίδι έχουμε $v' = 0$ ενώ για το δεύτερο $v' = 1$. Το πρώτο παιχνίδι είναι δίκαιο, ενώ το δεύτερο μεροληπτεί υπέρ του A.

4.2 Κυρτός προγραμματισμός

Το πρόβλημα (P) του κυρτού προγραμματισμού είναι το εξής: Δίνονται ένα κυρτό σύνολο $A \subseteq \mathbb{R}^n$ και μία κοίλη συνάρτηση $f : A \to \mathbb{R}$. Να βρεθεί η μέγιστη τιμή της f στο A, κάτω από τους περιορισμούς

$$\phi_1(x) \leq 0, \ldots, \phi_m(x) \leq 0$$

όπου κάθε $\phi_i : A \to \mathbb{R}$ είναι κυρτή.

[Πολλές φορές ζητάμε την ελάχιστη τιμή μιάς κυρτής συνάρτησης κάτω από κυρτούς περιορισμούς:το πρόβλημα αυτό είναι ισοδύναμο με το προηγούμενο αφού η $-f$ είναι κοίλη αν η f είναι κυρτή.]

Πολύ συχνά το A είναι το σύνολο $P^n = \{x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n : \xi_i \geq 0\}$. Τότε, μπορούμε να διατυπώσουμε το πρόβλημά μας παραλείποντας κάθε αναφορά στο A και προσθέτοντας τη συνθήκη $x \geq 0$. Για ευκολία στο συμβολισμό, θεωρούμε την $\Phi = (\phi_1, \ldots, \phi_m) : A \to \mathbb{R}^m$ και ονομάζουμε *εφικτό σύνολο* του (P) το

$$F = \{x \in A : \Phi(x) \leq 0\}.$$

Ζητάμε λοιπόν $x_0 \in F$ τέτοιο ώστε $f(x_0) \geq f(x)$ για κάθε $x \in F$.

Θεώρημα 4.4 *Το σύνολο F των εφικτών λύσεων είναι κυρτό.*

Απόδειξη: Το F είναι τομή του A και m κυρτών συνόλων (συνόλων στάθμης των κυρτών συναρτήσεων ϕ_i, $i = 1, \ldots, m$). $\square$

Αν όλες οι συναρτήσεις του προβλήματος ήταν διαφορίσιμες, θα μπορούσαμε να χρησιμοποιήσουμε τη μέθοδο των πολλαπλασιαστών Lagrange. Θα θεωρούσαμε δηλαδή τη συνάρτηση

$$K(x, y) = f(x) - y_1\phi_1(x) - \ldots - y_m\phi_m(x). \tag{1}$$

Αν και δεν έχουμε κάνει καμμία υπόθεση διαφορισιμότητας, η συνάρτηση αυτή θα φανεί πολύ χρήσιμη. Παρατηρήστε ότι για σταθερό $y \geq 0$, η συνάρτηση $K(x, y)$ είναι κοίλη συνάρτηση του x στο A, ενώ για σταθερό $x \in A$, είναι κυρτή συνάρτηση του y. Για μία τέτοια συνάρτηση, το σημείο $(x_0, y_0) \in A \times P^m$ είναι σαγματικό σημείο αν για κάθε $(x, y) \in A \times P^m$,

$$K(x, y_0) \leq K(x_0, y_0) \leq K(x_0, y).$$

Το πρόβλημα της ύπαρξης λύσης για το πρόβλημα κυρτού προγραμματισμού (P) συνδέεται στενά με το πρόβλημα της ύπαρξης σαγματικού σημείου για την $K(x, y)$. Για να περιγράψουμε αυτή τη σχέση ακριβέστερα, χρειάζεται να ορίσουμε ακόμα μία έννοια. Ένα σημείο $x \in F$ λέγεται *γνήσια εφικτή λύση* για το (P) αν $\phi_i(x) < 0$ για κάθε $i = 1, \ldots, m$.

Θεώρημα 4.5 *Θεωρούμε το πρόβλημα κυρτού προγραμματισμού (P) και την συνάρτηση $K(x, y)$ όπως παραπάνω.*

(a) Αν η K έχει ένα σαγματικό σημείο $(x_0, y_0) \in A \times P^m$, τότε το x_0 είναι βέλτιστη λύση για το (P).

(β) Έστω ότι το (P) έχει μία γνήσια εφικτή λύση w. Τότε, αν το (P) έχει μία βέλτιστη λύση x_0, υπάρχει $y_0 \in P^m$ τέτοιο ώστε το (x_0, y_0) να είναι σαγματικό σημείο για την K.

Απόδειξη: (α) Αφού το (x_0, y_0) είναι σαγματικό σημείο, ισχύει

$$f(x) - \langle y_0, \Phi(x) \rangle \leq f(x_0) - \langle y_0, \Phi(x_0) \rangle \leq f(x_0) - \langle y, \Phi(x_0) \rangle$$

για κάθε $x \in A, y \in P^m$. Η δεξιά ανισότητα συνεπάγεται ότι $\langle y, \Phi(x_0)\rangle \leq \langle y_0, \Phi(x_0)\rangle$ για κάθε $y \in P^m$. Ειδικότερα, $\langle y_0, \Phi(x_0)\rangle \geq \lambda\langle y_0, \Phi(x_0)\rangle$ για κάθε $\lambda > 0$. Άρα,

$$\langle y_0, \Phi(x_0)\rangle = 0.$$

Τότε όμως προκύπτει ότι $\langle y, \Phi(x_0)\rangle \leq 0$ για κάθε $y \in P^m$, άρα $\Phi(x_0) \leq 0$, δηλαδή το x_0 είναι εφικτή λύση.

Αν x εφικτή λύση, η αριστερή ανισότητα δίνει

$$f(x) \leq f(x) - \langle y_0, \Phi(x)\rangle \leq f(x_0) - \langle y_0, \Phi(x_0)\rangle = f(x_0).$$

Άρα το x_0 είναι βέλτιστη λύση.

(β) Θα αποδειχθεί ότι υπάρχει $y_0 \in P^m$ έτσι ώστε

$$f(x) - \langle y_0, \Phi(x)\rangle \leq f(x_0) - \langle y_0, \Phi(x_0)\rangle \leq f(x_0)$$

για κάθε $x \in A$ (τότε, αφού $\Phi(x_0) \leq 0$, ισχύει $f(x_0) \leq f(x_0) - \langle y, \Phi(x_0)\rangle$ για κάθε $y \in P^m$, άρα το (x_0, y_0) είναι σαγματικό σημείο για την K). Αυτό είναι ισοδύναμο με το ότι

$$\langle (1, y_0), (f(x), -\Phi(x))\rangle \leq \langle (1, y_0), (f(x_0), -\Phi(x_0))\rangle \leq f(x_0)$$

για κάθε $x \in A$. Θεωρούμε τα σύνολα

$$M = \{(t, y) \in \mathbb{R}^{m+1} : \exists x \in A : t \leq f(x), y \leq -\Phi(x)\}$$

και

$$N = \{(t, y) \in \mathbb{R}^{m+1} : t \geq f(x_0), y \geq 0\}.$$

Τα σύνολα αυτά είναι κυρτά και $M \cap \mathrm{int}(N) = \emptyset$. Γιατί, αν $(t, y) \in \mathrm{int}(N) \cap M$ τότε $t > f(x_0)$. Επιπλέον, υπάρχει $x \in A$ με $t \leq f(x)$, $y \leq -\Phi(x)$, δηλαδή το x είναι εφικτή λύση. Τότε όμως, $t \leq f(x) \leq f(x_0)$, το οποίο είναι άτοπο. Επιπλέον,

$$\mathrm{int}(N) \supseteq (f(x_0), +\infty) \times \{y \in \mathbb{R}^m : y > 0\},$$

άρα $\mathrm{int}(N) \neq \emptyset$.

Από το Θεώρημα 2.9 υπάρχει υπερεπίπεδο που διαχωρίζει τα M, N. Δηλαδή υπάρχουν $(a, b) \in \mathbb{R} \times \mathbb{R}^m$, $(a, b) \neq (0, 0)$ και $\gamma \in \mathbb{R}$ τέτοια ώστε

$$at + \langle b, y\rangle \leq \gamma \leq as + \langle b, z\rangle$$

για κάθε $(t, y) \in M$, $(s, z) \in N$.

Αν $(s, z) \in N$, τότε $(s', z) \in N$ για κάθε $s' \geq s$, άρα $a \geq 0$. Όμοια προκύπτει ότι $b \geq 0$. Δεν είναι όμως δυνατόν να ισχύει $a = 0$. Γιατί, για κάθε $x \in A$ ισχύει $(f(x), -\Phi(x)) \in M$, άρα

$$af(x) + \langle b, -\Phi(x)\rangle \leq af(x_0)$$

(αφού $(f(x_0), 0) \in N$). Αν $a = 0$, έπεται ότι $\langle b, -\Phi(x)\rangle \leq 0$ για κάθε $x \in A$. Αφού υπάρχει γνήσια εφικτή λύση παίρνουμε $b = 0$, άρα $(a, b) = (0, 0)$, το οποίο είναι άτοπο.

Ισχύει λοιπόν $a > 0$ και επομένως μπορούμε να υποθέσουμε ότι $a = 1$. Τότε έχουμε

$$t + \langle b, y\rangle \leq \gamma \leq s + \langle b, z\rangle$$

για κάθε $(t, y) \in M$, $(s, z) \in N$, και ειδικότερα:

$$f(x) + \langle b, -\Phi(x)\rangle \leq \gamma \leq f(x_0) + \langle b, -\Phi(x_0)\rangle$$

για κάθε $x \in A$. Επειδή όμως $(x_0, 0) \in N$, έπεται και ότι

$$f(x_0) + \langle b, -\Phi(x_0)\rangle \leq f(x_0),$$

άρα $\langle b, \Phi(x_0)\rangle = 0$, αφού $b \geq 0$, $\Phi(x_0) \leq 0$. Άρα, αν πάρουμε $y_0 = b$, ικανοποιούνται οι ζητούμενες ανισότητες. □

Μπορούμε να διατυπώσουμε τα συμπεράσματα του Θεωρήματος 4.5 χωρίς άμεση αναφορά στη συνάρτηση $K(x, y)$. Στην απόδειξη του (α), είδαμε ότι αν το (x_0, y_0) είναι σαγματικό σημείο, τότε η

$$f(x) - y_{01}\phi_1(x) - \ldots - y_{0m}\phi_m(x)$$

παίρνει μέγιστη τιμή στο x_0, και

$$y_{01}\phi_1(x_0) + \ldots + y_{0m}\phi_m(x_0) = 0.$$

Είδαμε επίσης ότι αυτές οι συνθήκες είναι ικανές για την ύπαρξη βέλτιστης λύσης για το (P). Στο (β) είδαμε ότι αν το (P) έχει μία γνήσια εφικτή λύση, τότε οι παραπάνω συνθήκες είναι αναγκαίες για την ύπαρξη βέλτιστης λύσης στο σημείο x_0. Αφού $y_0 \geq 0$ και $\Phi(x_0) \leq 0$, η δεύτερη συνθήκη μάς λέει ότι αν $\phi_i(x_0) < 0$ τότε $y_{0i} = 0$. Μπορούμε λοιπόν να διατυπώσουμε το Θεώρημα 4.5 στην εξής μορφή:

Θεώρημα 4.6 *Υποθέτουμε ότι το πρόβλημα κυρτού προγραμματισμού (P) έχει μία γνήσια εφικτή λύση. Τότε, το (P) έχει βέλτιστη λύση στο x_0 αν και μόνο αν υπάρχει $y_0 \in P^m$ τέτοιο ώστε να ικανοποιούνται οι συνθήκες Kuhn-Tucker: η $f(x) - y_{01}\phi_1(x) - \ldots - y_{0m}\phi_m(x)$ παίρνει μέγιστη τιμή στο x_0 και αν $\phi_i(x_0) < 0$ τότε $y_{0i} = 0$.* □

Παράδειγμα Να ελαχιστοποιηθεί η συνάρτηση $x+y$ υπό τους περιορισμούς $x^2+y^2 \leq 1$, $(x-1)^2 + y^2 \leq 1$.

Το προηγούμενο θεώρημα εφαρμόζεται στη συνάρτηση $f(x, y) = -x - y$. Η συνάρτηση K είναι η

$$K(x, y, \lambda_1, \lambda_2) = -x - y - \lambda_1(x^2 + y^2 - 1) - \lambda_2((x-1)^2 + y^2 - 1),$$

$(x, y) \in \mathbb{R}^2$, $(\lambda_1, \lambda_2) \in P^2$. Η πρώτη από τις συνθήκες Kuhn-Tucker διατυπώνεται ως εξής: Η K έχει μέγιστη τιμή ως προς (x, y), ή, ισοδύναμα (επειδή η K είναι κοίλη ως προς (x, y) και το (x, y) διατρέχει ανοικτό σύνολο)

$$\frac{\partial K}{\partial x} = \frac{\partial K}{\partial y} = 0,$$

ή

$$(*) \qquad \begin{cases} -1 - 2\lambda_1 x - 2\lambda_2 x + 2\lambda_2 = 0 \\ -1 - 2\lambda_1 y - 2\lambda_2 y = 0. \end{cases}$$

Η δεύτερη συνθήκη Kuhn-Tucker είναι

$$(**) \qquad \begin{cases} \lambda_1(x^2 + y^2 - 1) = 0 \\ \lambda_2((x-1)^2 + y^2 - 1) = 0. \end{cases}$$

Οι περιορισμοί είναι

$$(***) \qquad \begin{cases} x^2 + y^2 \leq 1 \\ (x-1)^2 + y^2 \leq 1 \\ \lambda_1 \geq 0 \\ \lambda_2 \geq 0. \end{cases}$$

Πρέπει λοιπόν να λυθεί το σύστημα $(*)$, $(**)$, $(***)$.

Από τις $(*)$ δεν είναι δυνατόν να ισχύει $\lambda_1 = \lambda_2 = 0$. Αν $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, από τις $(**)$ προκύπτει ότι $x^2 + y^2 - 1 = 0 = (x-1)^2 + y^2 - 1$, άρα $x = 1/2$ και $y = \pm\sqrt{3}/2$. Τότε, από τις $(*)$ παίρνουμε

$$\lambda_1 = -\frac{\sqrt{3} \pm 1}{2\sqrt{3}},$$

το οποίο αντιβαίνει στους περιορισμούς $(***)$.

Έστω ότι $\lambda_2 = 0$. Τότε από τις $(**)$ ισχύει $x^2 + y^2 = 1$ και από τις $(*)$ ισχύει

$$2\lambda_1 x = -1 = 2\lambda_1 y$$

άρα

$$x = y = \pm\frac{1}{\sqrt{2}}.$$

Και οι δύο περιπτώσεις αποκλείονται, αφού αν $x = y = 1/\sqrt{2}$ τότε $\lambda_1 < 0$, ενώ αν $x = y = -1/\sqrt{2}$ τότε $(x-1)^2 + y^2 = 2 + \sqrt{2} > 1$.

Έστω, τέλος, $\lambda_1 = 0$. Τότε από τις $(*)$ και $(**)$ προκύπτει:

$$x = 1 \pm \frac{1}{\sqrt{2}}, \quad y = \pm\frac{1}{\sqrt{2}}, \quad \lambda_2 = \mp\frac{1}{\sqrt{2}}.$$

Η περίπτωση $\lambda_2 < 0$ αποκλείεται, οπότε η βέλτιστη λύση είναι

$$x = 1 - \frac{1}{\sqrt{2}}, \quad y = -\frac{1}{\sqrt{2}}. \qquad \square$$

Στη συνέχεια θα αποδείξουμε το θεώρημα δυϊσμού για τον κυρτό προγραμματισμό. Υπενθυμίζουμε πρώτα το πρόβλημα του κυρτού προγραμματισμού:

Αρχικό πρόβλημα (P)**:** *Να μεγιστοποιηθεί η κοίλη συνάρτηση* $f : A \to \mathbb{R}$ *υπό τους περιορισμούς* $\phi_i(x) \leq 0$*, όπου* $\phi_i : A \to \mathbb{R}$ *κυρτές συναρτήσεις,* $i = 1, \ldots, m$.

Θεωρούμε τη συνάρτηση

$$g(y) = \sup_{x \in A} K(x, y), \quad y \in P^m.$$

Για σταθερό x, η $K(x,y)$ είναι αφφινική συνάρτηση του y, άρα η g είναι το supremum μιάς οικογένειας κυρτών (αφφινικών) συναρτήσεων. Επομένως, η g είναι κυρτή. Έστω F^* το σύνολο στο οποίο η g παίρνει πεπερασμένες τιμές. Το F^* είναι το *εφικτό σύνολο* για τη συνάρτηση g.

Δυϊκό πρόβλημα (P^*)**:** *Αν* $F^* \neq \emptyset$*, να ελαχιστοποιηθεί η κυρτή συνάρτηση* $g : P^m \to \mathbb{R}$ *στο* F^*.

Η σχέση του (P) με το (P^*) περιγράφεται από το *θεώρημα δυϊσμού των Kuhn-Tucker.*

Θεώρημα 4.7 *Έστω* (P) *ένα αρχικό πρόβλημα κυρτού προγραμματισμού και* (P^*) *το δυϊκό του πρόβλημα.*

(α) Αν x *είναι μία εφικτή λύση του* (P) *και* y *μία εφικτή λύση του* (P^*)*, τότε* $f(x) \leq g(y)$.

(β) Έστω ότι το (P) *έχει μία γνήσια εφικτή λύση. Τότε, αν το πρόβλημα* (P) *έχει μία βέλτιστη λύση* x_0*, το πρόβλημα* (P^*) *έχει μία βέλτιστη λύση* y_0 *τέτοια ώστε* $f(x_0) = g(y_0)$.

Απόδειξη: (α) Αφού το x είναι εφικτή λύση του (P) και $y \geq 0$, ισχύει

$$f(x) \leq f(x) - \langle y, \Phi(x) \rangle \leq \sup_{z \in A} (f(z) - \langle y, \Phi(z) \rangle) = g(y).$$

(β) Από το Θεώρημα 4.5 έπεται ότι υπάρχει $y_0 \in P^m$ έτσι ώστε το (x_0, y_0) να είναι σαγματικό σημείο για την K, δηλαδή ισχύουν οι σχέσεις:

(i) $K(x_0, y_0) = \sup_{x \in A} K(x, y_0)$.

(ii) $K(x_0, y_0) = \inf_{y \in P^m} K(x_0, y)$.

Από το Θεώρημα 4.1 παίρνουμε τη σχέση

(iii) $K(x_0, y_0) = \inf_{y \in P^m} \sup_{x \in A} K(x, y)$

Από την (i) έπεται ότι $g(y_0) = K(x_0, y_0)$, άρα το y_0 είναι εφικτή λύση. Από τις (i) και (iii) έπεται ότι

$$g(y_0) = K(x_0, y_0) = \inf_{y \in P^m} g(y),$$

άρα το y_0 είναι βέλτιστη λύση. Τέλος, παρατηρούμε ότι

$$\inf_{y \in P^m} K(x_0, y) = \inf_{y \in P^m} (f(x_0) - \langle y, \Phi(x_0) \rangle) = f(x_0),$$

αφού $0 \in P^m$ και $\Phi(x_0) \leq 0$. Άρα η (ii) συνεπάγεται ότι $K(x_0, y_0) = f(x_0)$, άρα $g(y_0) = f(x_0)$. □

Παράδειγμα Θεωρούμε το εξής πρόβλημα γραμμικού προγραμματισμού:

(P) Να μεγιστοποιηθεί η συνάρτηση $f(x) = \langle c, x\rangle$, $x \in P^n$, υπό τους περιορισμούς $Tx \leq b$, όπου $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, T ένας $m \times n$ πίνακας.

Θεωρούμε το (P) ως πρόβλημα κυρτού προγραμματισμού. Θα διατυπώσουμε το δυϊκό πρόβλημα. Η συνάρτηση g έχει την εξής μορφή:

$$\begin{aligned} g(y) &= \sup_{x\in P^m} (f(x) - \langle y, Tx - b\rangle) \\ &= \sup_{x\geq 0}(\langle c, x\rangle - \langle y, Tx\rangle + \langle y, b\rangle) \\ &= \langle y, b\rangle + \sup_{x\geq 0}(\langle c, x\rangle - \langle T^t y, x\rangle) \\ &= \langle y, b\rangle + \sup_{x\geq 0}(\langle c - T^t y, x\rangle), \quad y \in P^m. \end{aligned}$$

Αν το διάνυσμα $c - T^t y$ έχει μία γνήσια θετική συντεταγμένη, τότε

$$\sup_{x\geq 0}(\langle c - T^t y, x\rangle) = +\infty$$

(αρκεί να πάρουμε $x = \lambda e_i$, $\lambda > 0$, αν η i-συντεταγμένη είναι γνήσια θετική). Αν $c - T^t y \leq 0$, τότε

$$\sup_{x\geq 0}(\langle c - T^t y, x\rangle) = 0.$$

Άρα το εφικτό σύνολο του δυϊκού προβλήματος είναι το $F^* = \{y \in P^m : T^t y \geq c\}$, και $g(y) = \langle y, b\rangle$, $y \in F^*$.

Άρα το δυϊκό πρόβλημα είναι το εξής:

(P^*) Να ελαχιστοποιηθεί η συνάρτηση $g(y) = \langle b, y\rangle$, $y \in P^m$, υπό τους περιορισμούς $T^t y \geq c$. □

4.3 Ανάλυση πινάκων

Σκοπός μας σε αυτήν την παράγραφο είναι να αποδείξουμε μία πολύ γενική ανισότητα για συναρτήσεις πινάκων από την οποία προκύπτουν σαν ειδικές περιπτώσεις πολλά κλασικά αποτελέσματα. Θα χρειαστούμε δύο θεωρήματα που αφορούν την κλάση Ω_n των $n \times n$ *διπλά στοχαστικών πινάκων*. Αυτοί είναι οι πίνακες (a_{ij}) που έχουν μη αρνητικές συντεταγμένες a_{ij} και ικανοποιούν τις

$$\sum_{j=1}^{n} a_{ij} = 1, \quad i = 1, \ldots, n$$

και

$$\sum_{i=1}^{n} a_{ij} = 1, \quad j = 1, \ldots, n.$$

Δηλαδή, το άθροισμα των συντεταγμένων κάθε γραμμής ή στήλης ισούται με 1. Το πρώτο θεώρημα (Birkoff, 1946) εξασφαλίζει ότι τα ακραία σημεία του Ω_n (αν το δούμε σαν υποσύνολο του $\mathbb{R}^{n^2}$) είναι οι *πίνακες μεταθέσεων*. Πίνακας μετάθεσης είναι κάθε $n \times n$ πίνακας που προκύπτει με μετάθεση των γραμμών του ταυτοτικού πίνακα: έχει ακριβώς μία συντεταγμένη ίση με 1 σε κάθε γραμμή ή στήλη, ενώ όλες οι άλλες είναι ίσες με 0.

Θεώρημα 4.8 *Το σύνολο Ω_n των διπλά στοχαστικών πινάκων είναι κυρτό υποσύνολο του γραμμικού χώρου των $n \times n$ πινάκων. Τα ακραία σημεία του Ω_n είναι οι πίνακες μεταθέσεων: κάθε διπλά στοχαστικός πίνακας είναι κυρτός συνδυασμός πινάκων μεταθέσεων.*

Απόδειξη: Εύκολα ελέγχουμε ότι το Ω_n είναι κυρτό και ότι κάθε πίνακας μετάθεσης είναι ακραίο σημείο του. Επίσης, το Ω_n είναι συμπαγές κυρτό υποσύνολο του $\mathbb{R}^{n^2}$, άρα είναι η κυρτή θήκη των ακραίων σημείων του από το Θεώρημα του Minkowski.

Αυτό που μένει να δείξουμε είναι ότι κάθε ακραίο σημείο του Ω_n είναι πίνακας μετάθεσης. Έστω $A \in \Omega_n$ ένας πίνακας που δεν είναι πίνακας μετάθεσης. Τότε υπάρχει τουλάχιστον μία γραμμή, ας πούμε η i_1-στή γραμμή, στην οποία μπορούμε να βρούμε τουλάχιστον δύο μη μηδενικές συντεταγμένες. Παίρνουμε μία από αυτές, ας πούμε την $a_{i_1 i_2}$. Τότε, $0 < a_{i_1 i_2} < 1$ γιατί το άθροισμα των στοιχείων της i_1-στής γραμμής είναι 1 και υπάρχει κι άλλο θετικό στοιχείο σε αυτήν.

Από την άλλη πλευρά, το άθροισμα των στοιχείων της i_2-στής στήλης είναι επίσης 1. Αφού $a_{i_1 i_2} < 1$, υπάρχει $a_{i_3 i_2}$ με $i_3 \neq i_1$ τέτοιο ώστε $0 < a_{i_3 i_2} < 1$. Για τον ίδιο λόγο, υπάρχει $i_4 \neq i_2$ με $0 < a_{i_3 i_4} < 1$. Συνεχίζοντας έτσι, μετά από πεπερασμένα το πλήθος βήματα θα επιστρέψουμε σε κάποια συντεταγμένη a_{ij} του A η οποία είχε ήδη επιλεγεί.

Θεωρούμε την ακολουθία των συντεταγμένων που επιλέξαμε από την πρώτη ως και την δεύτερη εμφάνιση του a_{ij} (υπολογίζουμε την πρώτη αλλά όχι την δεύτερη εμφάνισή του). Έχουμε έτσι μία πεπερασμένη διατεταγμένη ακολουθία συντεταγμένων του A με την ιδιότητα: κάθε ζευγάρι διαδοχικών όρων της περιέχει στοιχεία που (εναλλάξ) βρίσκονται στην ίδια γραμμή ή στήλη του A. Έστω a_{kl} το μικρότερο (θετικό) στοιχείο αυτής της ακολουθίας. Θεωρούμε τον πίνακα B που στις θέσεις της ακολουθίας μας έχει συντεταγμένες $+1, -1, +1, -1, \ldots$ και 0 παντού αλλού. Το άθροισμα των στοιχείων κάθε γραμμής ή στήλης του B είναι 0 (γιατί;).

Θέτουμε $A_+ = A + a_{kl}B$ και $A_- = A - a_{kl}B$. Τότε, οι A_+, A_- είναι διπλά στοχαστικοί και διαφορετικοί από τον A (γιατί;) και

$$A = \frac{A_+ + A_-}{2}.$$

Άρα, ο A δεν είναι ακραίο σημείο του Ω_n. $\square$

Έστω $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$ σταθερό διάνυσμα. Ορίζουμε

$$K_w = \{T(w) : T \in \Omega_n\}.$$

Παρατηρούμε ότι το K_w είναι κυρτό σύνολο. Επίσης, αν $T \in \Omega_n$ γράφουμε $t_1, \ldots, t_n$ για τα διανύσματα-γραμμές του T. Με αυτόν τον συμβολισμό, έχουμε το εξής.

Θεώρημα 4.9 *Αν A κυρτό με $K_w \subseteq A$ και $f : A \to \mathbb{R}$ είναι οποιαδήποτε κυρτή συνάρτηση, τότε η $g : \Omega_n \to \mathbb{R}$ που ορίζεται από την*

$$g(T) = f(T(w)) = f(\langle t_1, w\rangle, \ldots, \langle t_n, w\rangle)$$

είναι κυρτή συνάρτηση. Η ελάχιστη τιμή της g στο Ω_n πιάνεται σε κάποιον πίνακα μετάθεσης P_w.

Απόδειξη: Αν $S, T \in \Omega_n$ και $a, b > 0$ με $a + b = 1$, τότε

$$\begin{aligned} g(aS + bT) &= f[(aS + bT)(w)] = f[aS(w) + bT(w)] \\ &\leq af(S(w)) + bf(T(w)) = ag(S) + bg(T), \end{aligned}$$

δηλαδή η g είναι κυρτή. Έστω $P_1, \ldots, P_m$, $m = n!$ μία αρίθμηση των πινάκων μετάθεσης στο Ω_n. Ο τυχών $S \in \Omega_n$ γράφεται σαν κυρτός συνδυασμός

$$S = \sum_{j=1}^{m} t_j P_j$$

για κάποια $t_j \geq 0$ με $\sum_{j=1}^{m} t_j = 1$, από το Θεώρημα 4.8. Από την κυρτότητα της g,

$$g(s) \leq \sum_{j=1}^{m} t_j g(P_j) \leq \sum_{j=1}^{m} t_j g(P_w) = g(P_w),$$

όπου P_w πίνακας μετάθεσης που εξαρτάται από το w και επιλέγεται έτσι ώστε

$$g(P_w) = \max_{1 \leq j \leq m} g(P_j). \qquad \square$$

Μπορούμε τώρα να δείξουμε το κεντρικό θεώρημα της παραγράφου.

Θεώρημα 4.10 *Έστω T ένας συμμετρικός $n \times n$ πίνακας και w το διάνυσμα που προκύπτει αν διατάξουμε με τυχόντα τρόπο τις (πραγματικές) ιδιοτιμές του A. Αν A κυρτό με $K_w \subseteq A$ και $f : A \to \mathbb{R}$ είναι μία κυρτή συνάρτηση, τότε για κάθε ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ του $\mathbb{R}^n$ έχουμε*

$$f(\langle Tv_1, v_1\rangle, \ldots, \langle Tv_n, v_n\rangle) \leq f(P_w(w)) = f(w_{i_1}, \ldots, w_{i_n})$$

όπου P_w ο πίνακας μετάθεσης που μάς εξασφαλίζει το Θεώρημα 2.

Απόδειξη: Έστω $u_1, \ldots, u_n$ ιδιοδιανύσματα μήκους 1 του T που αντιστοιχούν στις ιδιοτιμές $w_1, \ldots, w_n$. Δηλαδή, $\|u_i\| = 1$ και $Tu_i = w_i u_i$. Θεωρούμε τον πίνακα S με συντεταγμένες $s_{ij} = \langle u_i, v_j\rangle^2$. Ο S είναι διπλά στοχαστικός, γιατί είναι συμμετρικός και

$$\sum_{j=1}^{n} s_{ij} = \sum_{j=1}^{n} \langle u_i, v_j\rangle^2 = \|u_i\|^2 = 1$$

για κάθε $i \leq n$. Όμως,

$$Tv_j = T\left(\sum_{i=1}^{n}\langle u_i, v_j\rangle u_i\right) = \sum_{i=1}^{n}\langle u_i, v_j\rangle w_i u_i,$$

άρα

$$\langle Tv_j, v_j\rangle = \sum_{i=1}^{n}\langle u_i, v_j\rangle^2 w_i = \langle s_j, w\rangle,$$

όπου s_j είναι η j-στή στήλη (ή, ισοδύναμα, γραμμή) του S. Συνεπώς,

$$\begin{aligned} f(\langle Tv_1, v_1\rangle, \ldots, \langle Tv_n, v_n\rangle) &= f(\langle s_1, w\rangle, \ldots, \langle s_n, w\rangle) \\ &= f(S(w)) \leq f(P_w(w)) = f(w_{i_1}, \ldots, w_{i_n}), \end{aligned}$$

όπου P_w ο πίνακας μετάθεσης που μάς εξασφαλίζει το Θεώρημα 4.9. □

Στην πραγματικότητα έχουμε αποδείξει κάτι ισχυρότερο. Αν επιλέξουμε σαν $\{v_1, \ldots, v_n\}$ τη σωστή μετάθεση $\{u_{i_1}, \ldots, u_{i_n}\}$ των ιδιοτιμών του T πετυχαίνουμε ισότητα. Δηλαδή,

$$\max f(\langle Tv_1, v_1\rangle, \ldots, \langle Tv_n, v_n\rangle) \leq f(P_w(w))$$

όπου το maximum παίρνεται πάνω από όλες τις ορθοκανονικές βάσεις του $\mathbb{R}^n$.

Επιλέγοντας κατάλληλα την συνάρτηση f στο προηγούμενο θεώρημα, παίρνουμε μία σειρά ανισοτήτων. Η πρώτη μας επιλογή είναι η

$$f(s_1, \ldots, s_n) = \sum_{i=1}^{k} s_i$$

για σταθερό $k \leq n$.

Θεώρημα 4.11 *Έστω T ένας $n \times n$ συμμετρικός πίνακας και $\{v_1, \ldots, v_n\}$ μία ορθοκανονική βάση του $\mathbb{R}^n$. Τότε,*

$$\sum_{i=1}^{k}\lambda_{n-i+1} \leq \sum_{i=1}^{k}\langle Tv_i, v_i\rangle \leq \sum_{i=1}^{k}\lambda_i$$

όπου $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ είναι οι ιδιοτιμές του T.

Απόδειξη: Η $f(s_1, \ldots, s_n) = \sum_{i=1}^{k} s_i$ είναι κυρτή στον $\mathbb{R}^n$ (η $-f$ επίσης). Από το Θεώρημα 4.10 υπάρχει μετάθεση $j_1, \ldots, j_n$ των $1, \ldots, n$ τέτοια ώστε

$$\sum_{i=1}^{k}\langle Tv_i, v_i\rangle \leq \sum_{i=1}^{k}\lambda_{j_i}.$$

Λόγω της διάταξης των λ_i, το δεξιό μέλος είναι μικρότερο ή ίσο από το $\sum_{i=1}^{k}\lambda_i$, οπότε έχουμε αποδείξει την δεξιά ανισότητα του θεωρήματος. Κάνοντας ανάλογη δουλειά με την $-f$ παίρνουμε την αριστερή ανισότητα. Παρατηρούμε ότι έχουμε ισότητα στην αριστερή ανισότητα αν το v_i, $i \leq k$ είναι ιδιοδιάνυσμα για την λ_{n-i+1}, ενώ στην δεξιά ανισότητα αν το v_i είναι ιδιοδιάνυσμα για την λ_i. $\square$

Η επόμενη επιλογή μας είναι η

$$f(s_1, \ldots, s_n) = \left(\prod_{i=1}^{k} s_i \right)^{1/k}$$

που ορίζεται για μη αρνητικά s_i και σταθερό $k \leq n$. Η f είναι κοίλη, οπότε παίρνουμε το εξής:

Θεώρημα 4.12 *Αν T είναι ένας $n \times n$ συμμετρικός πίνακας και $\{v_1, \ldots, v_n\}$ είναι μία ορθοκανονική βάση του $\mathbb{R}^n$, τότε*

$$\prod_{i=1}^{k} \lambda_{n-i+1} \leq \prod_{i=1}^{k} \langle Tv_i, v_i \rangle \leq \left(\frac{1}{k} \sum_{i=1}^{k} \lambda_i \right)^k$$

όπου $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ είναι οι ιδιοτιμές του T.

Απόδειξη: Η συνάρτηση f που χρησιμοποιούμε είναι κοίλη, άρα η $-f$ είναι κυρτή στο P^n. Από το Θεώρημα 4.10,

$$\left(\prod_{i=1}^{k} \lambda_{n-i+1} \right)^{1/k} \leq \left(\prod_{i=1}^{k} \langle Tv_i, v_i \rangle \right)^{1/k},$$

το οποίο μάς δίνει την αριστερή ανισότητα. Παρατηρούμε κι εδώ ότι έχουμε ισότητα στην αριστερή ανισότητα αν το v_i, $i \leq k$ είναι ιδιοδιάνυσμα για την λ_{n-i+1}. Για την δεξιά ανισότητα εφαρμόζουμε την ανισότητα αριθμητικού-γεωμετρικού μέσου και το Θεώρημα 4.11:

$$\left(\prod_{i=1}^{k} \langle Tv_i, v_i \rangle \right)^{1/k} \leq \frac{1}{k} \sum_{i=1}^{k} \langle Tv_i, v_i \rangle \leq \frac{1}{k} \sum_{i=1}^{k} \lambda_i. \quad \square$$

Άμεσες συνέπειες είναι δύο πολύ γνωστές ανισότητες για ορίζουσες.

Θεώρημα 4.13 (Ανισότητα του Hadamard) *Αν T είναι ένας $n \times n$ πίνακας με συντεταγμένες t_{ij}, τότε*

$$(\det T)^2 \leq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} t_{ij}^2 \right).$$

Αν ο T είναι θετικά ημιορισμένος,

$$\det T \leq \prod_{i=1}^{n} t_{ii}.$$

Απόδειξη: Δείχνουμε πρώτα την δεύτερη ανισότητα. Η ορίζουσα ενός πίνακα είναι ίση με το γινόμενο των ιδιοτιμών του, αν λοιπόν θεωρήσουμε την συνήθη ορθοκανονική βάση $\{e_i\}$ και χρησιμοποιήσουμε το Θεώρημα 4.12 με $k = n$, παίρνουμε

$$\det T = \prod_{i=1}^{n} \lambda_i \leq \prod_{i=1}^{n} \langle Te_i, e_i \rangle = \prod_{i=1} t_{ii}.$$

Αν εφαρμόσουμε αυτήν την ανισότητα για τον θετικά ημιορισμένο πίνακα $S = T^t T$ (όπου τώρα T τυχών $n \times n$ πίνακας), παίρνουμε

$$(\det T)^2 = \det S \leq \prod_{j=1}^{n} s_{jj} = \prod_{j=1}^{n} \left(\sum_{i=1}^{n} t_{ij}^2 \right). \qquad \square$$

Θεώρημα 4.14 (Ανισότητα του Minkowski) *Αν T και S είναι θετικά ημιορισμένοι $n \times n$ πίνακες, τότε*

$$[\det(T+S)]^{1/n} \geq (\det T)^{1/n} + (\det S)^{1/n}$$

και

$$\det(T+S) \geq \det T + \det S.$$

Απόδειξη: Έστω $v_1, \ldots, v_n$ ένα ορθοκανονικό σύνολο ιδιοδιανυσμάτων του $T+S$. Τότε,

$$\begin{aligned}
[\det(T+S)]^{1/n} &= \left[\prod_{i=1}^{n} \langle (T+S)v_i, v_i \rangle \right]^{1/n} \\
&= \left[\prod_{i=1}^{n} \langle Tv_i, v_i \rangle + \langle Sv_i, v_i \rangle \right]^{1/n} \\
&\geq \left[\prod_{i=1}^{n} \langle Tv_i, v_i \rangle \right]^{1/n} + \left[\prod_{i=1}^{n} \langle Sv_i, v_i \rangle \right]^{1/n} \\
&\geq (\det T)^{1/n} + (\det S)^{1/n}.
\end{aligned}$$

Η προτελευταία ανισότητα είναι συνέπεια της ανισότητας αριθμητικού-γεωμετρικού μέσου, ενώ η τελευταία προκύπτει από το Θεώρημα 4.12. Υψώνοντας στην n-στή δύναμη, παίρνουμε την $\det(T+S) \geq \det T + \det S$. $\square$

Έστω T ένας συμμετρικός $n \times n$ πίνακας με ιδιοτιμές $\lambda_1 \geq \ldots \lambda_n$. Ορίζουμε

$$s_k(T) = \sum_{i=1}^{k} \lambda_{n-i+1}, \qquad S_k(T) = \sum_{i=1}^{k} \lambda_i$$

και

$$p_k(T) = \prod_{i=1}^{k} \lambda_{n-i+1}, \qquad P_k(T) = \prod_{i=1}^{k} \lambda_i.$$

Παρατηρήστε ότι

$$\lambda_n = s_1(T) = p_1(T), \qquad \lambda_1 = S_1(T) = P_1(T)$$

και

$$\mathrm{tr}(T) = s_n(T) = S_n(T), \qquad \det T = p_n(T) = P_n(T).$$

Θεώρημα 4.15 *Οι συναρτήσεις s_k είναι κοίλες και οι συναρτήσεις S_k είναι κυρτές στην κλάση $\mathcal{S}_n$ των συμμετρικών $n \times n$ πινάκων. Δηλαδή, αν $T_1, T_2 \in \mathcal{S}_n$ και $a, b > 0$ με $a + b = 1$, τότε*

$$\begin{aligned} as_k(T_1) + bs_k(T_2) &\leq s_k(aT_1 + bT_2) \\ &\leq S_k(aT_1 + bT_2) \leq aS_k(T_1) + bS_k(T_2). \end{aligned}$$

Απόδειξη: Έστω $V = \{v_1, \ldots, v_n\}$ ορθοκανονική βάση του $\mathbb{R}^n$. Αν θέσουμε $L_V(T) = \sum_{i=1}^k \langle Tv_i, v_i \rangle$, τότε

$$-s_k(T) = \max_V[-L_V(T)], \quad S_k(T) = \max_V L_V(T).$$

Αυτό προκύπτει από την διερεύνηση των περιπτώσεων ισότητας στο Θεώρημα 4.11. Η $L_V(T)$ είναι γραμμική συνάρτηση του T, άρα οι L_V και $-L_V$ είναι κυρτές στην $\mathcal{S}_n$. Αφού η κλάση των κυρτών συναρτήσεων είναι κλειστή ως προς supremum, έπεται ότι οι s_k και S_k είναι κυρτές συναρτήσεις στην $\mathcal{S}_n$. □

Ειδικότερα, αποδείξαμε ότι

$$\lambda_n = s_1(T) = \min_{\|v\|=1} \langle Tv, v \rangle$$

και

$$\lambda_1 = S_1(T) = \max_{\|v\|=1} \langle Tv, v \rangle$$

δύο ειδικές περιπτώσεις του θεωρήματος του Fisher για τις ιδιοτιμές ενός πραγματικού συμμετρικού πίνακα.

Θεώρημα 4.16 *Η συνάρτηση p_k είναι λογαριθμικά κοίλη στην κλάση των θετικά ορισμένων πινάκων. Αν οι T_1 και T_2 είναι θετικά ορισμένοι και $a, b > 0$ με $a + b = 1$, τότε*

$$[p_k(T_1)]^a [p_k(T_2)]^b \leq p_k(aT_1 + bT_2).$$

Ειδικότερα,

$$(\det T_1)^a (\det T_2)^b \leq \det(aT_1 + bT_2).$$

Απόδειξη: Από το Θεώρημα 4.12 (και τις περιπτώσεις ισότητας στην αριστερή ανισότητα) έπεται ότι

$$p_k(T) = \min_V \prod_{i=1}^k \langle Tv_i, v_i \rangle,$$

όπου το min παίρνεται πάνω από όλες τις ορθοκανονικές βάσεις $V = \{v_1, \ldots, v_n\}$ του $\mathbb{R}^n$. Άρα,

$$\begin{aligned}
[p_k(T_1)]^a [p_k(T_2)]^b &= \left[\min_V \prod_{i=1}^{k} \langle T_1 v_i, v_i\rangle\right]^a \left[\min_V \prod_{i=1}^{k} \langle T_2 v_i, v_i\rangle\right]^b \\
&\leq \min_V \prod_{i=1}^{k} [\langle T_1 v_i, v_i\rangle^a \langle T_2 v_i, v_i\rangle^b] \\
&\leq \min_V \prod_{i=1}^{k} [a\langle T_1 v_i, v_i\rangle + b\langle T_2 v_i, v_i\rangle] \\
&= p_k(aT_1 + bT_2). \qquad \square
\end{aligned}$$

4.4 Ασκήσεις

1. Έστω ότι το (i_0, j_0) είναι σαγματικό σημείο για τον $m \times n$ πίνακα αποδόσεων $P(i,j)$. Αν x_0 και y_0 είναι τα μοναδιαία διανύσματα που έχουν 1 στην θέση i_0 και j_0 αντίστοιχα, δείξτε ότι το (x_0, y_0) είναι σαγματικό σημείο για την

$$E(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i P(i,j) y_j.$$

2. Ένας τετραγωνικός πίνακας P ικανοποιεί την $P(i,j) = -P(j,i)$ για κάθε i,j. Αν ένα παιχνίδι έχει πίνακα αποδόσεων τον P, δείξτε ότι η τιμή του παιχνιδιού είναι 0 και ότι οι δύο παίκτες έχουν τις ίδιες βέλτιστες μεικτές στρατηγικές.

3. Δείξτε ότι το σύνολο των σαγματικών σημείων της $K(x,y)$ στο Θεώρημα 4.2 είναι κυρτό στον $\mathbb{R}^m \times \mathbb{R}^n$.

4. Σε ένα παιχνίδι με δύο παίκτες όπως στο Θεώρημα 4.3, έχουμε $f(x') = g(y')$ αν και μόνο αν το (x', y') είναι σαγματικό σημείο για την $E(x,y)$. Επιπλέον, τα σύνολα των βέλτιστων στρατηγικών των Α και Β είναι κυρτά.

5. Αποδείξτε ότι η συνάρτηση $f(x) = x$, $x \in \mathbb{R}$ έχει μέγιστη τιμή υπό τον περιορισμό $x^2 \leq 0$, αλλά η $K(x,y) = x - yx^2$ δεν έχει σαγματικό σημείο.

6. Βρείτε το ελάχιστο της συνάρτησης $-6x + 2x^2 - 2xy + 2y^2$ υπό τους περιορισμούς $x + y \leq 2$, $x \geq 0$, $y \geq 0$.

7. Βρείτε το ελάχιστο της συνάρτησης $(x+1)^2 + (y-2)^2$ υπό τους περιορισμούς $0 \leq x \leq 2$, $0 \leq y \leq 1$.

8. Δίνονται διανύσματα $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ και ένας $m \times n$ πίνακας A. Θεωρούμε το πρόβλημα γραμμικού προγραμματισμού: Να ελαχιστοποιηθεί η συνάρτηση $\langle c, x\rangle$, $x \in \mathbb{R}^n$, υπό τους περιορισμούς $Ax \leq b$, $x \geq 0$. Γράψτε τις συνθήκες Kuhn-Tucker για το πρόβλημα.