ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικό και Καποδιστριακό
Πανεπιστήμιο Αθηνών

Κυρτή Ανάλυση

Ενότητα: Ακραία σημεία

Απόστολος Γιαννόπουλος

Τμήμα Μαθηματικών

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Περιεχόμενα ενότητας

6 Ακραία σημεία

6.1 Ακραία και εκτεθειμένα σημεία

Ορισμός 6.1.1 (ακραία σημεία – έδρες). *(α) Έστω C ένα κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$. Ένα σημείο $x \in C$ λέγεται ακραίο σημείο του C αν δεν περιέχεται στο εσωτερικό κάποιου ευθύγραμμου τμήματος του οποίου τα άκρα ανήκουν στο C.*

Για να ελέγξουμε ότι το $x \in C$ είναι ακραίο σημείο του C αρκεί, ισοδύναμα, να ελέγξουμε το εξής: αν $y, z \in C$ και $x = (1-t)y + tz$ για κάποιο $0 < t < 1$, τότε $y = z = x$ (παρατηρήστε ότι αν $y \neq z$ τότε το ευθύγραμμο τμήμα $[y, z]$ είναι μη τετριμμένο και για κάθε $0 < t < 1$ το $(1-t)y + tz$ είναι εσωτερικό του σημείο).

Γράφουμε $\mathrm{ext}(C)$ για το σύνολο των ακραίων σημείων του C. Παραδείγματα: τα ακραία σημεία ενός δίσκου είναι όλα τα σημεία της περιφέρειάς του και τα ακραία σημεία ενός κύβου είναι οι κορυφές του. Ένα κλειστό κυρτό σύνολο μπορεί να μην έχει ακραία σημεία: αν θεωρήσετε οποιονδήποτε κλειστό ημίχωρο $\overline{G}$ τότε $\mathrm{ext}(\overline{G}) = \emptyset$.

(β) Γενικότερα, ένα κυρτό υποσύνολο $F \subseteq C$ λέγεται έδρα του C αν ισχύει το εξής: αν $y, z \in C$ και $x = (1-t)y + tz$ για κάποιο $0 < t < 1$, τότε $y, z \in F$. Μία έδρα F του C λέγεται k-έδρα αν $\dim(\mathrm{aff}(F)) = k$. Από τους ορισμούς είναι φανερό ότι $x \in \mathrm{ext}(C)$ αν και μόνο αν το μονοσύνολο $\{x\}$ είναι 0-έδρα του C.

Λήμμα 6.1.2. *Έστω C ένα κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$. Ένα σημείο $x \in C$ είναι ακραίο σημείο του C αν και μόνο αν το σύνολο $C \setminus \{x\}$ είναι κυρτό.*

Απόδειξη. Υποθέτουμε πρώτα ότι $x \in \mathrm{ext}(C)$. Έστω $y, z \in C \setminus \{x\}$ και έστω $t \in (0,1)$. Από την κυρτότητα του C έχουμε $(1-t)y + tz \in C$. Από την άλλη πλευρά, $(1-t)y + tz \neq x$, αφού το x είναι ακραίο σημείο του C και $y, z \neq x$. Έπεται ότι $(1-t)y + tz \in C \setminus \{x\}$. Αυτό αποδεικνύει ότι το $C \setminus \{x\}$ είναι κυρτό.

Αντίστροφα, ας υποθέσουμε ότι το $C \setminus \{x\}$ είναι κυρτό. Αν $x \notin \mathrm{ext}(C)$, τότε υπάρχουν $y, z \in C$ ώστε το x να είναι εσωτερικό σημείο του $[y, z]$. Όμως τότε, $y, z \in C \setminus \{x\}$, και από την υπόθεση παίρνουμε $x \in [y, z] \subseteq C \setminus \{x\}$, το οποίο είναι άτοπο. □

Ορισμός 6.1.3 (εκτεθειμένα σημεία). *Έστω C ένα κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$. Ένα σημείο $x \in C$ λέγεται εκτεθειμένο σημείο του C αν υπάρχει υπερεπίπεδο H ώστε το H να στηρίζει το C και το x να είναι το μοναδικό κοινό σημείο των C και H: δηλαδή, $C \cap H = \{x\}$. Γράφουμε $\exp(C)$ για το σύνολο των εκτεθειμένων σημείων του C.*

Λήμμα 6.1.4. *Έστω C ένα κλειστό κυρτό υποσύνολο του $\mathbb{R}^n$. Κάθε εκτεθειμένο σημείο του C είναι ακραίο σημείο του C. Δηλαδή,*

$$\exp(C) \subseteq \mathrm{ext}(C). \tag{6.1.1}$$

Απόδειξη. Έστω $x \in \exp(C)$. Τότε, υπάρχουν $0 \neq u \in \mathbb{R}^n$ και $\alpha \in \mathbb{R}$ ώστε

$$\langle x, u \rangle = \alpha \tag{6.1.2}$$

και

$$\langle w, u \rangle < \alpha \quad \text{για κάθε} \quad w \in C \setminus \{x\}. \tag{6.1.3}$$

Υποθέτουμε ότι $x = (1-t)y + tz$ για κάποια $y, z \in C$ και $0 < t < 1$. Τότε,

$$\alpha = \langle x, u\rangle = (1-t)\langle y, u\rangle + t\langle z, u\rangle \leq (1-t)\alpha + t\alpha = \alpha. \tag{6.1.4}$$

Αφού $0 < t < 1$, αυτό σημαίνει ότι $\langle y, u\rangle = \langle z, u\rangle = \alpha$. Από την (6.1.3) έπεται ότι $y = z = x$. Δηλαδή, $x \in \mathrm{ext}(C)$. □

Σημείωση: Ο εγκλεισμός μπορεί να είναι γνήσιος. Θεωρήστε τα σύνολα

$$\begin{aligned} C_1 &= \{(x,y) \in \mathbb{R}^2 : -1 \leq x \leq 1, y = 1 + \sqrt{1-x^2}\} \\ C_2 &= \{(x,y) \in \mathbb{R}^2 : -1 \leq x \leq 1, y = -1 - \sqrt{1-x^2}\}, \end{aligned}$$

και την κυρτή τους θήκη $C = \mathrm{conv}(C_1 \cup C_2)$. Τότε, $(\pm 1, \pm 1) \in \mathrm{ext}(C) \setminus \exp(C)$. Θα δείξουμε όμως ότι στην περίπτωση ενός κυρτού συμπαγούς συνόλου, το σύνολο των εκτεθειμένων σημείων είναι πυκνό στο σύνολο των ακραίων σημείων.

Θεώρημα 6.1.5 (Straszewicz). *Έστω K ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$. Τότε,*

$$\overline{\exp(K)} \supseteq \mathrm{ext}(K). \tag{6.1.5}$$

Σημείωση: Το σύνολο των ακραίων σημείων ενός συμπαγούς κυρτού υποσυνόλου του $\mathbb{R}^n$ ($n \geq 3$) μπορεί να μην είναι κλειστό: για παράδειγμα θεωρήστε το δίσκο $D = \{(x, y, 0) \in \mathbb{R}^3 : (x-1)^2 + y^2 = 1\}$ και το συμπαγές κυρτό σύνολο $K = \mathrm{conv}(D \cup (0, 0, \pm 1))$. Δείξτε ότι το $\mathrm{ext}(K) = \exp(K)$ δεν είναι κλειστό σύνολο. Συνεπώς, ο εγκλεισμός $\overline{\exp(K)} \supseteq \mathrm{ext}(K)$ του Θεωρήματος 6.1.5 δεν μπορεί να αντικατασταθεί με ισότητα.

Η απόδειξη θα βασιστεί σε μία σειρά από λήμματα που παρουσιάζουν ανεξάρτητο ενδιαφέρον. Το πρώτο δίνει έναν χαρακτηρισμό των ακραίων σημείων ενός συμπαγούς κυρτού συνόλου.

Λήμμα 6.1.6. *Έστω K ένα κυρτό συμπαγές υποσύνολο του $\mathbb{R}^n$. Ένα σημείο $x \in K$ είναι ακραίο σημείο του K αν και μόνο αν για κάθε $r > 0$ υπάρχει ανοικτός ημίχωρος G του $\mathbb{R}^n$ ώστε $x \in G$ και $K \cap G \subseteq B(x, r)$.*

Απόδειξη. Υποθέτουμε πρώτα ότι $x \in \mathrm{ext}(K)$ και ότι $r > 0$. Τότε, το $K \setminus B(x, r)$ είναι συμπαγές, άρα το σύνολο

$$L = \mathrm{conv}(K \setminus B(x, r)) \tag{6.1.6}$$

είναι κυρτό και συμπαγές.

Παρατηρήστε ότι $x \notin L$. Αλλιώς, αφού $x \in \mathrm{ext}(K)$, από το Λήμμα 6.1.2 θα είχαμε

$$x \in \mathrm{conv}(K \setminus B(x, r)) \subseteq \mathrm{conv}(K \setminus \{x\}) = K \setminus \{x\}, \tag{6.1.7}$$

το οποίο είναι άτοπο. Αφού $x \notin L$, τα L και x διαχωρίζονται αυστηρά. Συνεπώς, υπάρχει ανοικτός ημίχωρος G ώστε $x \in G$ και $\overline{G} \cap L = \emptyset$. Τότε,

$$(K \setminus B(x, r)) \cap \overline{G} = \emptyset, \quad \text{άρα} \quad K \setminus B(x, r) \subseteq G^c. \tag{6.1.8}$$

Έπεται ότι $K \cap G \subseteq B(x, r)$.

Αντίστροφα: έστω ότι $x \notin \mathrm{ext}(K)$. Τότε, υπάρχουν $y \neq z$ στο K ώστε $x \in (y, z)$. Μπορούμε λοιπόν να βρούμε $r > 0$ αρκετά μικρό ώστε τα y και z να μην ανήκουν στην $B(x, r)$.

Από την υπόθεση, υπάρχει ανοικτός ημίχωρος G με την ιδιότητα: $x \in G$ και $K \cap G \subseteq B(x,r)$. Αφού $y,z \notin B(x,r)$, έχουμε $y,z \in G^c$. Όμως, το G^c είναι κυρτό (κλειστός ημίχωρος). Άρα, $[y,z] \subseteq G^c$. Δηλαδή, $x \in G^c$, το οποίο είναι άτοπο. □

Το επόμενο λήμμα περιγράφει μία «μέθοδο εντοπισμού εκτεθειμένων σημείων» ενός κυρτού συμπαγούς συνόλου.

Λήμμα 6.1.7. *Έστω K ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$. Έστω $w \in \mathbb{R}^n$. Θεωρούμε $z \in K$ το οποίο έχει τη μέγιστη δυνατή απόσταση από το w:*

$$\|z - w\|_2 = \max\{\|y - w\|_2 : \ y \in K\} \tag{6.1.9}$$

(παρατηρήστε ότι τέτοια σημεία υπάρχουν, αφού το K είναι συμπαγές και η $y \mapsto \|y - w\|_2$ είναι συνεχής). Τότε, $z \in \exp(K)$.

Απόδειξη. Για κάθε $y \in K$ έχουμε

$$\|z - w\|_2^2 \geq \|y - w\|_2^2 = \|(y - z) + (z - w)\|_2^2, \tag{6.1.10}$$

δηλαδή

$$\|z - w\|_2^2 \geq \|y - z\|_2^2 + 2\langle y - z, z - w\rangle + \|z - w\|_2^2. \tag{6.1.11}$$

Άρα, για κάθε $y \in K$ έχουμε

$$2(\langle y, z - w\rangle - \langle z, z - w\rangle) + \|y - z\|_2^2 \leq 0. \tag{6.1.12}$$

Έπεται ότι

$$\langle y, z - w\rangle \leq \alpha := \langle z, z - w\rangle \tag{6.1.13}$$

για κάθε $y \in K$. Δηλαδή, το

$$H = \{x \in \mathbb{R}^n : \langle x, z - w\rangle = \alpha\} \tag{6.1.14}$$

στηρίζει το K στο z. Επιπλέον, αν $\langle y, z - w\rangle = \alpha$ για κάποιο $y \in K$ τότε, επιστρέφοντας στην (6.1.12) βλέπουμε ότι $\|y - z\|_2^2 \leq 0$, δηλαδή $y = z$. Άρα, $K \cap H = \{z\}$, το οποίο δείχνει ότι το z είναι εκτεθειμένο σημείο του K. □

Χρησιμοποιώντας το Λήμμα 6.1.7 παίρνουμε το εξής.

Λήμμα 6.1.8. *Έστω K ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$. Αν G είναι ένας ανοικτός ημίχωρος του $\mathbb{R}^n$ με την ιδιότητα $K \cap G \neq \emptyset$, τότε $\exp(K) \cap G \neq \emptyset$.*

Απόδειξη. Θέτουμε $D = \mathrm{diam}(K)$ και γράφουμε τον G στη μορφή

$$G = \{x \in \mathbb{R}^n : \langle x, u\rangle < \alpha\}, \tag{6.1.15}$$

όπου $0 \neq u \in \mathbb{R}^n$ και $\alpha \in \mathbb{R}$. Έστω $x \in K \cap G$. Θα επιλέξουμε $\lambda > 0$ κατάλληλα μεγάλο και θα χρησιμοποιήσουμε το προηγούμενο Λήμμα για το $w = x + \lambda u$.

Θεωρούμε $z \in K$ το οποίο έχει τη μεγαλύτερη δυνατή απόσταση από το w. Από το Λήμμα 6.1.7 έχουμε $z \in \exp(K)$. Μένει λοιπόν να δείξουμε ότι $z \in G$. Αφού

$$\|z - w\|_2 \geq \|x - w\|_2, \tag{6.1.16}$$

αρκεί να δείξουμε ότι για κάθε $y \in K \cap G^c$ ισχύει

$$\|y - w\|_2 < \|x - w\|_2. \tag{6.1.17}$$

Παρατηρούμε ότι αν $y \in K \cap G^c$ τότε $\langle y, u\rangle \geq \alpha$, οπότε

$$\|y - w\|_2^2 = \|y - x - \lambda u\|_2^2 = \|y - x\|_2^2 - 2\lambda\langle y - x, u\rangle + \|x - w\|_2^2 \tag{6.1.18}$$

αφού $\lambda u = w - x$. Αρκεί λοιπόν να εξασφαλίσουμε ότι

$$\|y - x\|_2^2 < 2\lambda(\langle y, u\rangle - \langle x, u\rangle) \tag{6.1.19}$$

για κάθε $y \in K \cap G^c$. Αφού $\|y - x\|_2 \leq D$ και $\langle y, u\rangle - \langle x, u\rangle \geq \alpha - \langle x, u\rangle > 0$, αρκεί να επιλέξουμε το $\lambda > 0$ αρκετά μεγάλο ώστε να ισχύει η

$$D^2 < 2\lambda(\alpha - \langle x, u\rangle). \tag{6.1.20}$$

Τότε, η (6.1.17) ισχύει για κάθε $y \in K \cap G^c$, και από την (6.1.16) έχουμε $z \in \exp(K) \cap G$. □

Απόδειξη του Θεωρήματος 6.1.5: Έστω $x \in \mathrm{ext}(K)$. Θα δείξουμε ότι για κάθε $r > 0$ ισχύει $B(x, r) \cap \exp(K) \neq \emptyset$. Από το Λήμμα 6.1.6 υπάρχει ανοικτός ημίχωρος G ώστε

$$x \in G \quad \text{και} \quad K \cap G \subseteq B(x, r). \tag{6.1.21}$$

Αφού $K \cap G \neq \emptyset$, από το Λήμμα 6.1.8 υπάρχει

$$z \in \exp(K) \quad \text{ώστε} \quad z \in G. \tag{6.1.22}$$

Αφού $z \in K \cap G$, από την (6.1.21) συμπεραίνουμε ότι $z \in B^\circ(x, r)$. Δηλαδή,

$$B(x, r) \cap \exp(K) \neq \emptyset. \tag{6.1.23}$$

Αφού το $r > 0$ ήταν τυχόν, βλέπουμε ότι $x \in \overline{\exp(K)}$. □

Το βασικό αποτέλεσμα αυτής της Παραγράφου είναι το ακόλουθο θεώρημα του Minkowski.

Θεώρημα 6.1.9 (Minkowski). *Έστω K ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$. Τότε, το K είναι η κυρτή θήκη των ακραίων σημείων του:*

$$K = \mathrm{conv}(\mathrm{ext}(K)). \tag{6.1.24}$$

Απόδειξη. Η απόδειξη γίνεται με επαγωγή ως προς τη διάσταση $d = \dim(\mathrm{aff}(K))$ της αφινικής θήκης του K. Αν $d = 0$, τότε το K είναι μονοσύνολο και ο ισχυρισμός του θεωρήματος ισχύει προφανώς. Υποθέτουμε λοιπόν ότι η διάσταση του K είναι $d > 0$ και ότι το θεώρημα ισχύει για όλα τα συμπαγή κυρτά υποσύνολα του $\mathbb{R}^n$ που έχουν αφινική διάσταση μικρότερη από d.

Έστω $x \in K$. Αν το x είναι ακραίο σημείο του K τότε, προφανώς, $x \in \mathrm{conv}(\mathrm{ext}(K))$. Υποθέτουμε λοιπόν ότι $x \in K \setminus \mathrm{ext}(K)$.

Τότε, υπάρχει ευθεία ℓ ώστε το x να είναι εσωτερικό σημείο του ευθύγραμμου τμήματος $[y, z] = \ell \cap K$. Το y ανήκει στο σχετικό σύνορο του K στην $\mathrm{aff}(K)$. Τότε, υπάρχουν $0 \neq u \in \mathrm{aff}(K)$ και $\alpha \in \mathbb{R}$ ώστε

$$\langle w, u\rangle \leq \alpha = \langle y, u\rangle \tag{6.1.25}$$

για κάθε $w \in K$. Θεωρούμε το υπερεπίπεδο $H_y = \{v \in \mathrm{aff}(K) : \langle v,u\rangle = \alpha\}$ του $\mathrm{aff}(K)$ και το $K \cap H_y$. Το $K \cap H_y$ είναι μη κενό, κυρτό και συμπαγές σύνολο με διάσταση μικρότερη από d. Άρα, είναι η κυρτή θήκη των ακραίων σημείων του. Αφού $y \in K \cap H_y$, το y γράφεται σαν κυρτός συνδυασμός ακραίων σημείων του $K \cap H_y$. Αν λοιπόν δείξουμε ότι κάθε ακραίο σημείο του $K \cap H_y$ είναι ακραίο σημείο του K, θα έχουμε δείξει ότι $y \in \mathrm{conv}(\mathrm{ext}(K))$.

Έστω w ακραίο σημείο του $K \cap H_y$ και ας υποθέσουμε ότι $w = (1-t)w_1 + tw_2$ για κάποια $w_1, w_2 \in K$ και $0 < t < 1$. Τότε,

$$\alpha = \langle w,u\rangle = (1-t)\langle w_1,u\rangle + t\langle w_2,u\rangle \leq (1-t)\alpha + t\alpha = \alpha, \tag{6.1.26}$$

απ' όπου βλέπουμε ότι $\langle w_1,u\rangle = \langle w_2,u\rangle = \alpha$, δηλαδή, $w_1, w_2 \in K \cap H_y$. Όμως το w είναι ακραίο σημείο του $K \cap H_y$, άρα $w_1 = w_2 = w$. Αυτό αποδεικνύει ότι $w \in \mathrm{ext}(K)$.

Με τον ίδιο τρόπο δείχνουμε ότι $z \in \mathrm{conv}(\mathrm{ext}(K))$. Αφού $x \in [y,z]$, έπεται ότι

$$x \in \mathrm{conv}(\{y,z\}) \subseteq \mathrm{conv}(\mathrm{ext}(K)). \tag{6.1.27}$$

Το $x \in K \setminus \mathrm{ext}(K)$ ήταν τυχόν, άρα $K \subseteq \mathrm{conv}(\mathrm{ext}(K))$. □

Σημείωση: Χρησιμοποιώντας το Θεώρημα του Καραθεοδωρή βλέπουμε ότι αν K είναι ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$ τότε κάθε $x \in K$ γράφεται σαν κυρτός συνδυασμός το πολύ $n+1$ ακραίων σημείων του K.

Ένας άλλος τρόπος διατύπωσης του θεωρήματος του Minkowski είναι ο εξής.

Πόρισμα 6.1.10. *Έστω K ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$ και έστω $M \subseteq K$. Τότε,*

$$K = \mathrm{conv}(M) \quad \textit{αν και μόνο αν} \quad M \supseteq \mathrm{ext}(K). \tag{6.1.28}$$

Απόδειξη. Αφού το K είναι κυρτό και $M \subseteq K$, έχουμε $K \supseteq \mathrm{conv}(M)$. Παρατηρούμε ότι αν $M \supseteq \mathrm{ext}(K)$ τότε

$$K \supseteq \mathrm{conv}(M) \supseteq \mathrm{conv}(\mathrm{ext}(K)) = K. \tag{6.1.29}$$

Αντίστροφα, αν $K = \mathrm{conv}(M)$ και υπάρχει $x \in \mathrm{ext}(K) \setminus M$, τότε από την κυρτότητα του $K \setminus \{x\}$ (Λήμμα 6.1.2) και την $M \subseteq K \setminus \{x\}$ παίρνουμε

$$K = \mathrm{conv}(M) \subseteq \mathrm{conv}(K \setminus \{x\}) = K \setminus \{x\}, \tag{6.1.30}$$

το οποίο είναι άτοπο. □

Πόρισμα 6.1.11. *Έστω K ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$ και έστω $M \subseteq K$. Τότε,*

$$K = \overline{\mathrm{conv}}(\mathrm{exp}(K)). \tag{6.1.31}$$

Απόδειξη. Από το Θεώρημα του Straszewicz έχουμε $\mathrm{ext}(K) \subseteq \overline{\mathrm{exp}(K)}$. Παρατηρήστε ότι, γενικά, $\mathrm{conv}(\overline{A}) \subseteq \overline{\mathrm{conv}}(A)$. Άρα,

$$K = \mathrm{conv}(\mathrm{ext}(K)) \subseteq \mathrm{conv}(\overline{\mathrm{exp}(K)}) \subseteq \overline{\mathrm{conv}}(\mathrm{exp}(K)) \subseteq K. \tag{6.1.32}$$

Δηλαδή, $K = \overline{\mathrm{conv}}(\mathrm{exp}(K))$. □

Πολλές από τις εφαρμογές του Θεωρήματος του Minkowski βασίζονται στην επόμενη απλή Πρόταση.

Πρόταση 6.1.12. *Έστω K ένα μη κενό, συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$. Αν $f : K \to \mathbb{R}$ είναι μία συνεχής κυρτή συνάρτηση και αν $M = \max\{f(x) : x \in K\}$, τότε υπάρχει ακραίο σημείο z του K ώστε $f(z) = M$.*

Απόδειξη. Αφού το K είναι συμπαγές και η f είναι συνεχής, υπάρχει $x \in K$ ώστε $f(x) = M$. Αφού $x \in K = \mathrm{conv}(\mathrm{ext}(K))$, υπάρχουν $z_1, \dots, z_m \in \mathrm{ext}(K)$ και $t_i \geq 0$ με $t_1 + \cdots + t_m = 1$ ώστε $x = t_1 z_1 + \cdots + t_m z_m$. Από την κυρτότητα της f έχουμε

$$M = f(x) = f(t_1 z_1 + \cdots + t_m z_m) \leq \sum_{i=1}^{m} t_i f(z_i) \leq \max_{1 \leq i \leq m} f(z_i) \leq M. \tag{6.1.33}$$

Άρα, υπάρχει $i \leq m$ ώστε $f(z_i) = M$ (ακριβέστερα, έχουμε $f(z_i) = M$ για όλους τους δείκτες $i \leq m$ που ικανοποιούν την $t_i > 0$). □

6.2 Πολύτοπα και πολύεδρα

Στην §2.1 ορίσαμε την κλάση των πολυτόπων και την κλάση των πολυέδρων στον n-διάστατο Ευκλείδειο χώρο.

(α) **Πολύτοπο** στον $\mathbb{R}^n$ είναι η κυρτή θήκη ενός πεπερασμένου συνόλου S σημείων του $\mathbb{R}^n$.

(β) **Πολύεδρο** στον $\mathbb{R}^n$ είναι μία «πεπερασμένη τομή ημιχώρων», δηλαδή ένα σύνολο της μορφής

$$P = \{x \in \mathbb{R}^n : \langle x, u_i \rangle \leq \alpha_i \ \text{ για } \ i = 1, \dots, m\} \tag{6.2.1}$$

όπου $m \in \mathbb{N}$, $u_1, \dots, u_m$ είναι μη μηδενικά διανύσματα στον $\mathbb{R}^n$ και $\alpha_1, \dots, \alpha_m \in \mathbb{R}$.

Σκοπός μας σε αυτή την παράγραφο είναι να δείξουμε ότι η κλάση των φραγμένων πολυέδρων και η κλάση των πολυτόπων συμπίπτουν. Η απόδειξη θα βασιστεί στο εξής Λήμμα.

Λήμμα 6.2.1. *Έστω P το πολύεδρο*

$$P = \{x \in \mathbb{R}^n : \langle x, u_i \rangle \leq \alpha_i, \ i = 1, \dots, m\}. \tag{6.2.2}$$

Για κάθε $y \in P$ ορίζουμε

$$I(y) = \{i \leq m : \langle y, u_i \rangle = \alpha_i\}. \tag{6.2.3}$$

Τότε, $y \in \mathrm{ext}(P)$ αν και μόνο αν το σύνολο $\{u_i : i \in I(y)\}$ παράγει τον $\mathbb{R}^n$. Ειδικότερα, για κάθε $y \in \mathrm{ext}(P)$ έχουμε $|I(y)| \geq n$.

Απόδειξη. Υποθέτουμε πρώτα ότι για κάποιο $y \in P$ το σύνολο $\{u_i : i \in I(y)\}$ παράγει τον $\mathbb{R}^n$. Έστω ότι $y = (1-t)y_1 + t y_2$ για κάποια $y_1, y_2 \in P$ και κάποιο $0 < t < 1$. Τότε, για κάθε $i \in I(y)$ έχουμε

$$\alpha_i = \langle y, u_i \rangle = (1-t)\langle y_1, u_i \rangle + t\langle y_2, u_i \rangle \leq (1-t)\alpha_i + t\alpha_i = \alpha_i, \tag{6.2.4}$$

δηλαδή

$$\langle y_1, u_i \rangle = \langle y_2, u_i \rangle = \alpha_i \quad \text{για κάθε} \quad i \in I(y). \tag{6.2.5}$$

Έπεται ότι $(y_1 - y_2) \perp u_i$ για κάθε $i \in I(y)$. Αφού τα u_i, $i \in I(y)$ παράγουν τον $\mathbb{R}^n$, συμπεραίνουμε ότι $y_1 - y_2 = 0$. Δηλαδή, $y_1 = y_2 = y$. Αυτό δείχνει ότι $y \in \mathrm{ext}(P)$.

Αντίστροφα, ας υποθέσουμε ότι το $\{u_i : i \in I(y)\}$ δεν παράγει τον $\mathbb{R}^n$. Τότε, υπάρχει $z \neq 0$ στον $\mathbb{R}^n$ με την ιδιότητα

$$\langle z, u_i\rangle = 0 \quad \text{για κάθε} \quad i \in I(y). \tag{6.2.6}$$

Παρατηρήστε ότι, για κάθε $i \notin I(y)$ έχουμε

$$\langle y, u_i\rangle < \alpha_i. \tag{6.2.7}$$

Θεωρούμε $\varepsilon > 0$ και ορίζουμε $y_1 = y + \varepsilon z$, $y_2 = y - \varepsilon z$. Τότε, $y_i \neq y$ και

$$y = \frac{y_1 + y_2}{2}. \tag{6.2.8}$$

Επίσης, αν το $\varepsilon > 0$ είναι αρκετά μικρό, έχουμε $y_1, y_2 \in P$. Πράγματι, αν $i \in I(y)$ τότε

$$\langle y \pm \varepsilon z, u_i\rangle = \langle y, u_i\rangle \pm \varepsilon\langle z, u_i\rangle = \langle y, u_i\rangle = \alpha_i, \tag{6.2.9}$$

ενώ, αν $i \notin I(y)$ έχουμε

$$\langle y \pm \varepsilon z, u_i\rangle = \langle y, u_i\rangle \pm \varepsilon\langle z, u_i\rangle < \alpha_i \tag{6.2.10}$$

αν το $\varepsilon > 0$ είναι αρκετά μικρό ώστε να ικανοποιούνται οι

$$\varepsilon|\langle z, u_i\rangle| < \alpha_i - \langle y, u_i\rangle, \quad i \notin I(y). \tag{6.2.11}$$

Από τα παραπάνω συμπεραίνουμε ότι $y \notin \mathrm{ext}(P)$. □

Θεώρημα 6.2.2. *Κάθε φραγμένο πολύεδρο είναι πολύτοπο.*

Απόδειξη. Έστω P το φραγμένο πολύεδρο

$$P = \{x \in \mathbb{R}^n : \langle x, u_i\rangle \leq \alpha_i,\ i = 1, \ldots, m\}. \tag{6.2.12}$$

Το P είναι κλειστό και κυρτό (ως τομή κλειστών υποχώρων). Αφού είναι φραγμένο, το P είναι συμπαγές. Από το Θεώρημα του Minkowski έχουμε $P = \mathrm{conv}(\mathrm{ext}(P))$. Αν δείξουμε ότι το $\mathrm{ext}(P)$ είναι πεπερασμένο σύνολο, έπεται το Θεώρημα.

Από το Λήμμα 6.2.1, κάθε $y \in \mathrm{ext}(P)$ είναι η μοναδική λύση του συστήματος

$$\langle y, u_i\rangle = \alpha_i, \quad i \in I(y). \tag{6.2.13}$$

Η μοναδικότητα προκύπτει από το γεγονός ότι το σύνολο $\{u_i : i \in I(y)\}$ παράγει τον $\mathbb{R}^n$. Ειδικότερα, σε κάθε $y \in \mathrm{ext}(P)$ μπορούμε να αντιστοιχίσουμε μία n-άδα γραμμικά ανεξάρτητων διανυσμάτων $u_{i_1}(y), \ldots, u_{i_n}(y)$ ($i_j \in I(y)$) και η απεικόνιση $y \mapsto (u_{i_1}(y), \ldots, u_{i_n}(y))$ είναι ένα προς ένα.

Το πλήθος των δυνατών n-άδων $(u_{i_1}, \ldots, u_{i_n})$ αυτής της μορφής είναι το πολύ ίσο με $\binom{m}{n}$. Άρα,

$$|\mathrm{ext}(P)| \leq \binom{m}{n}. \tag{6.2.14}$$

Δηλαδή, το $\mathrm{ext}(P)$ είναι πεπερασμένο σύνολο. □

Για την αντίστροφη κατεύθυνση χρησιμοποιούμε τον δυϊσμό.

Θεώρημα 6.2.3. *Κάθε πολύτοπο είναι πολύεδρο.*

Απόδειξη. Έστω $P = \mathrm{conv}(\{x_1,\dots,x_m\})$ ένα πολύτοπο. Θα δουλέψουμε στην αφινική θήκη $A = \mathrm{aff}(P)$ του P. Αν δείξουμε ότι υπάρχουν $u_1,\dots,u_s \in A$ και $\alpha_i \in \mathbb{R}$ ώστε

$$P = \{x \in A : \langle x,u_i\rangle \leq \alpha_i,\ i = 1,\dots,s\}, \tag{6.2.15}$$

τότε μπορούμε να γράψουμε το P σαν τομή πεπερασμένων το πλήθος κλειστών ημιχώρων του $\mathbb{R}^n$. Πράγματι, υπάρχουν κλειστοί ημίχωροι $F_1,\dots,F_{2d}$ του $\mathbb{R}^n$ ώστε $A = F_1 \cap \cdots \cap F_{2d}$ (αν $n-d$ είναι η διάσταση της A και αν $v_1,\dots,v_d$ είναι ένα πλήρες ορθοκανονικό σύστημα διανυσμάτων κάθετων στην A, τότε η A γράφεται σαν τομή κλειστών ημιχώρων της μορφής $\{x : \langle x,v_j\rangle \leq \beta_j\}$ και $\{x : \langle x,v_j\rangle \geq \gamma_j\}$). Τότε,

$$P = F_1 \cap \cdots \cap F_{2d} \cap \{x \in \mathbb{R}^n : \langle x,u_i\rangle \leq \alpha_i,\ i = 1,\dots,s\}. \tag{6.2.16}$$

Χωρίς περιορισμό της γενικότητας υποθέτουμε λοιπόν ότι $\mathrm{int}(P) \neq \emptyset$. Μπορούμε επιπλέον να υποθέσουμε ότι $0 \in \mathrm{int}(P)$: δεν είναι δύσκολο να δείξετε ότι για κάθε $w \in \mathbb{R}^n$ το P είναι πολύεδρο αν και μόνο αν το $P + w$ είναι πολύεδρο.

Με αυτές τις υποθέσεις, το πολικό P° του P είναι ένα συμπαγές κυρτό υποσύνολο του $\mathbb{R}^n$ και $0 \in \mathrm{int}(P^\circ)$. Η βασική παρατήρηση είναι ότι

$$\begin{aligned} P^\circ &= \{y \in \mathbb{R}^n : \langle y,x\rangle \leq 1 \text{ για κάθε } x \in P\} \\ &= \{y \in \mathbb{R}^n : \langle y,x_i\rangle \leq 1 \text{ για κάθε } i = 1,\dots,m\}. \end{aligned}$$

Δηλαδή, το P° είναι ένα φραγμένο πολύεδρο στον $\mathbb{R}^n$. Από το Θεώρημα 6.2.2, το P° είναι πολύτοπο. Δηλαδή, υπάρχουν $z_1,\dots,z_r \in \mathbb{R}^n$ ώστε

$$P^\circ = \mathrm{conv}(\{z_1,\dots,z_r\}). \tag{6.2.17}$$

Τότε, το επιχείρημα που χρησιμοποιήσαμε παραπάνω δείχνει ότι

$$P^{\circ\circ} = \{x \in \mathbb{R}^n : \langle x,z_i\rangle \leq 1 \text{ για κάθε } i = 1,\dots,r\}. \tag{6.2.18}$$

Δηλαδή, το $P^{\circ\circ}$ είναι πολύεδρο. Αφού $P^{\circ\circ} = P$, έπεται το συμπέρασμα. □

6.3 Το πολύτοπο του Birkhoff

Ορισμός 6.3.1. *(α) Έστω σ μία μετάθεση του $\{1,\dots,n\}$. Ο πίνακας μετάθεσης X^σ είναι ο $n \times n$ πίνακας $X^\sigma = (x_{ij})$ με συντεταγμένες $x_{ij} = 1$ αν $\sigma(j) = i$ και $x_{ij} = 0$ αλλιώς. Δηλαδή, $x_{ij} = \delta_{i,\sigma(j)}$. Παρατηρήστε ότι κάθε πίνακας μετάθεσης έχει μία μονάδα και $n-1$ μηδενικά σε κάθε γραμμή ή στήλη του. Αντίστροφα, κάθε πίνακας αυτής της μορφής αντιστοιχεί σε κάποια μετάθεση σ του $\{1,\dots,n\}$.*

(β) Ένας $n \times n$ πίνακας $X = (x_{ij})$ λέγεται διπλά στοχαστικός αν $x_{ij} \geq 0$ για κάθε $i,j = 1,\dots,n$ και το άθροισμα των συντεταγμένων κάθε γραμμής ή στήλης του ισούται με 1. Δηλαδή,

$$\begin{aligned} \sum_{j=1}^n x_{ij} = 1 &\qquad \text{για κάθε } i = 1,\dots,n, \\ \sum_{i=1}^n x_{ij} = 1 &\qquad \text{για κάθε } j = 1,\dots,n. \end{aligned}$$

(γ) Από τον ορισμό του διπλά στοχαστικού πίνακα βλέπουμε ότι το σύνολο DS_n των διπλά στοχαστικών πινάκων είναι ένα πολύεδρο στον $\mathbb{R}^{n^2}$. Το DS_n λέγεται πολύτοπο του Birkhoff.

Εύκολα ελέγχουμε ότι κάθε πίνακας μετάθεσης είναι ακραίο σημείο του DS_n. Το θεώρημα των Birkhoff -- von Neumann ισχυρίζεται ότι το πολύτοπο DS_n δεν έχει άλλα ακραία σημεία.

Θεώρημα 6.3.2 (Birkhoff, 1946 -- von Neumann, 1953)**.** *Το σύνολο* $\mathrm{ext}(DS_n)$ *των ακραίων σημείων του πολυτόπου του Birkhoff είναι το σύνολο των* $n \times n$ *πινάκων μετάθεσης.*

Απόδειξη. Η απόδειξη που θα παρουσιάσουμε εδώ γίνεται με επαγωγή ως προς το n. Για $n = 1$ ο ισχυρισμός του Θεωρήματος ισχύει προφανώς. Έστω $n > 1$. Θεωρούμε τον αφινικό υπόχωρο

$$L = \left\{ X = (x_{ij}) \in \mathbb{R}^{n^2} : \ (\forall i \leq n) \ \sum_{j=1}^{n} x_{ij} = 1, \ (\forall j \leq n) \ \sum_{i=1}^{n} x_{ij} = 1 \right\} \tag{6.3.1}$$

του $\mathbb{R}^{n^2}$. Παρατηρήστε ότι $\dim(L) = (n-1)^2$. Πράγματι, κάθε $X = (x_{ij}) \in L$ προσδιορίζεται πλήρως από τις $(n-1)^2$ συντεταγμένες x_{ij}, $1 \leq i, j \leq n-1$, αφού οι υπόλοιπες συντεταγμένες προσδιορίζονται από τις σχέσεις

$$\begin{aligned} x_{in} &= 1 - \sum_{j=1}^{n-1} x_{ij} \\ x_{nj} &= 1 - \sum_{i=1}^{n-1} x_{ij} \\ x_{nn} &= 1 - \sum_{i=1}^{n-1} x_{in} = (2-n) + \sum_{i,j=1}^{n-1} x_{ij}. \end{aligned}$$

Αν περιοριστούμε στο χώρο L, το πολύτοπο DS_n ορίζεται από τις n^2 ανισότητες $x_{ij} \geq 0$. Ας υποθέσουμε ότι $X = (x_{ij})$ είναι ένα ακραίο σημείο του DS_n. Από το Λήμμα 6.2.1 υπάρχουν τουλάχιστον $(n-1)^2$ ζευγάρια (i, j) για τα οποία $x_{ij} = 0$. Αφού ο X είναι διπλά στοχαστικός, δεν μπορεί να έχει n μηδενικές συντεταγμένες σε κάποια γραμμή ούτε μπορεί να έχει περισσότερες από μία μη μηδενικές συντεταγμένες σε κάθε γραμμή, γιατί τότε το πλήθος των μηδενικών συντεταγμένων του θα ήταν το πολύ ίσο με $n(n-2) < (n-1)^2$. Υπάρχει λοιπόν κάποιος δείκτης $i_0 \leq n$ με την εξής ιδιότητα: υπάρχει μοναδικός δείκτης $j_0 \leq n$ ώστε $x_{i_0 j_0} \neq 0$. Αφού ο X είναι διπλά στοχαστικός, αναγκαστικά έχουμε

$$x_{i_0 j_0} = 1 \quad \text{και} \quad x_{i_0 j} = 0 \text{ αν } j \neq j_0. \tag{6.3.2}$$

Επίσης, αφού $x_{i_0 j_0} = 1$ και ο X είναι διπλά στοχαστικός, παίρνουμε

$$x_{i_0 j_0} = 1 \quad \text{και} \quad x_{i j_0} = 0 \text{ αν } i \neq i_0. \tag{6.3.3}$$

Από τις (6.3.2) και (6.3.3) είναι φανερό ότι αν διαγράψουμε την i_0-στή γραμμή και την j_0-στή στήλη του X θα προκύψει ένας διπλά στοχαστικός $(n-1) \times (n-1)$ πίνακας Y.

Παρατηρήστε ότι ο Y είναι ακραίο σημείο του DS_{n-1}. Αν υπήρχαν $Y_1 \neq Y_2$ στο DS_{n-1} και $0 < t < 1$ ώστε $Y = (1-t)Y_1 + tY_2$, τότε με «αντικατάσταση του Y από τους Y_1 και Y_2 αντίστοιχα μέσα στον X» θα παίρναμε δύο διπλά στοχαστικούς πίνακες $X_1, X_2 \in DS_n$ με την ιδιότητα: $X_1 \neq X_2$ και $X = (1-t)X_1 + tX_2$. Αυτό θα ήταν άτοπο, αφού $X \in \mathrm{ext}(DS_n)$.

Τώρα μπορούμε να εφαρμόσουμε την επαγωγική υπόθεση. Είδαμε ότι $Y \in \mathrm{ext}(DS_{n-1})$, άρα ο Y είναι πίνακας μετάθεσης. Έπεται ότι ο X είναι πίνακας μετάθεσης. □

6.3.1 Πολύτοπα μεταθέσεων

Ορισμός 6.3.3 (πολύτοπο μεταθέσεων). *Συμβολίζουμε με S_n την ομάδα των μεταθέσεων του $\{1,\dots,n\}$. Έστω $w=(w_1,\dots,w_n)\in\mathbb{R}^n$. Για κάθε μετάθεση $\sigma\in S_n$ ορίζουμε*

$$\sigma(w)=(w_{\sigma^{-1}(1)},\dots,w_{\sigma^{-1}(n)}). \tag{6.3.4}$$

Το πολύτοπο των μεταθέσεων του w είναι το πολύτοπο

$$P(w)=\mathrm{conv}(\{\sigma(w):\ \sigma\in S_n\}). \tag{6.3.5}$$

Δηλαδή, το $P(w)$ προκύπτει αν θεωρήσουμε όλα τα σημεία που παράγονται με μετάθεση των συντεταγμένων του w και πάρουμε την κυρτή τους θήκη.

Η επόμενη Πρόταση δείχνει ότι κάθε πολύτοπο μεταθέσεων είναι γραμμική εικόνα του πολυτόπου του Birkhoff.

Πρόταση 6.3.4. *Έστω $w\in\mathbb{R}^n$. Θεωρούμε τον γραμμικό μετασχηματισμό $T_w:\mathbb{R}^{n^2}\to\mathbb{R}^n$ που ορίζεται από την*

$$T_w(X)=X(w), \tag{6.3.6}$$

όπου έχουμε ταυτίσει τον $\mathbb{R}^{n^2}$ με τον γραμμικό χώρο των $n\times n$ πινάκων. Τότε,

$$T_w(DS_n)=P(w). \tag{6.3.7}$$

Απόδειξη. Από το Θεώρημα των Birkhoff -- von Neumann και από το Θεώρημα του Minkowski παίρνουμε

$$T_w(DS_n)=T_w(\mathrm{conv}(\{X^\sigma:\ \sigma\in S_n\}))=\mathrm{conv}(\{T_w(X^\sigma):\ \sigma\in S_n\}). \tag{6.3.8}$$

Παρατηρήστε ότι, για κάθε $\sigma\in S_n$,

$$T_w(X^\sigma)=X^\sigma(w)=(w_{\sigma^{-1}(1)},\dots,w_{\sigma^{-1}(n)})=\sigma(w). \tag{6.3.9}$$

Άρα,

$$T_w(DS_n)=\mathrm{conv}(\{\sigma(w):\ \sigma\in S_n\}). \tag{6.3.10}$$

Δηλαδή, $T_w(DS_n)=P(w)$. □

6.3.2 Εφαρμογές στην ανάλυση πινάκων

Αν X είναι ένας $n\times n$ πίνακας, γράφουμε $X=(X_1,\dots,X_n)$ όπου $X_1,\dots,X_n$ είναι τα διανύσματα-γραμμές του X. Στην προηγούμενη υποπαράγραφο είδαμε ότι αν $w\in\mathbb{R}^n$ και $X\in DS_n$ τότε $X(w)\in P(w)$. Από τη δομή του DS_n και την Πρόταση 6.1.12 παίρνουμε το εξής βασικό Λήμμα.

Λήμμα 6.3.5. *Έστω $w\in\mathbb{R}^n$ και C κυρτό υποσύνολο του $\mathbb{R}^n$ με $P(w)\subseteq C$. Αν $f:C\to\mathbb{R}$ είναι μία κυρτή συνάρτηση, τότε η $g:DS_n\to\mathbb{R}$ που ορίζεται από την*

$$g(X)=f(X(w))=f(\langle X_1,w\rangle,\dots,\langle X_n,w\rangle) \tag{6.3.11}$$

είναι κυρτή συνάρτηση και

$$\max(g)=\max\{f(\sigma(w)):\ \sigma\in S_n\}. \tag{6.3.12}$$

Απόδειξη. Αν $X, Y \in DS_n$ και $0 < t < 1$, τότε

$$\begin{aligned} g((1-t)X + tY) &= f[((1-t)X + tY)(w)] = f[(1-t)X(w) + tY(w)] \\ &\leq (1-t)f(X(w)) + tf(Y(w)) = (1-t)g(X) + tg(Y), \end{aligned}$$

δηλαδή η g είναι κυρτή. Από την Πρόταση 6.1.12 έπεται ότι

$$\max(g) = \max\{f(X^\sigma(w)) : \sigma \in S_n\}. \tag{6.3.13}$$

Αφού $X^\sigma(w) = \sigma(w)$, έπεται το Λήμμα. □

Χρησιμοποιώντας αυτή την παρατήρηση μπορούμε να δείξουμε πλήθος ανισοτήτων για τις ιδιοτιμές συμμετρικών πινάκων. Η απόδειξη τους βασίζεται στο εξής γενικό σχήμα.

Θεώρημα 6.3.6. *Έστω T ένας συμμετρικός $n \times n$ πίνακας και έστω w το διάνυσμα που προκύπτει αν διατάξουμε με τυχόντα τρόπο τις (πραγματικές) ιδιοτιμές του T. Αν C είναι ένα κυρτό υποσύνολο του $\mathbb{R}^n$ με $P(w) \subseteq C$ και αν $f : C \to \mathbb{R}$ είναι μία κυρτή συνάρτηση, τότε για κάθε ορθοκανονική βάση $\{v_1, \dots, v_n\}$ του $\mathbb{R}^n$ έχουμε*

$$f(\langle Tv_1, v_1\rangle, \dots, \langle Tv_n, v_n\rangle) \leq \max\{f(\sigma(w)) : \sigma \in S_n\}. \tag{6.3.14}$$

Απόδειξη. Έστω $\{u_1, \dots, u_n\}$ μία ορθοκανονική βάση από ιδιοδιανύσματα του T που αντιστοιχούν στις ιδιοτιμές $w_1, \dots, w_n$. Δηλαδή, $Tu_i = w_i u_i$. Θεωρούμε τον πίνακα X που έχει συντεταγμένες $x_{ij} = \langle u_i, v_j\rangle^2$. Ο X είναι διπλά στοχαστικός: για παράδειγμα, έχουμε

$$\sum_{j=1}^n x_{ij} = \sum_{j=1}^n \langle u_i, v_j\rangle^2 = \|u_i\|_2^2 = 1 \tag{6.3.15}$$

για κάθε $i \leq n$. Όμως,

$$Tv_j = T\left(\sum_{i=1}^n \langle u_i, v_j\rangle u_i\right) = \sum_{i=1}^n \langle u_i, v_j\rangle w_i u_i, \tag{6.3.16}$$

άρα

$$\langle Tv_j, v_j\rangle = \sum_{i=1}^n \langle u_i, v_j\rangle^2 w_i = \langle X_j, w\rangle, \tag{6.3.17}$$

όπου X_j είναι η j-στή στήλη του X. Από το Λήμμα 6.3.5,

$$\begin{aligned} f(\langle Tv_1, v_1\rangle, \dots, \langle Tv_n, v_n\rangle) &= f(\langle X_1, w\rangle, \dots, \langle X_n, w\rangle) \\ &= f(X^t(w)) \leq \max\{f(\sigma(w)) : \sigma \in S_n\}. \end{aligned}$$

Στην πραγματικότητα έχουμε αποδείξει κάτι ισχυρότερο. Αν επιλέξουμε σαν $\{v_1, \dots, v_n\}$ κατάλληλη μετάθεση $\sigma(w)$ των ιδιοτιμών του T πετυχαίνουμε ισότητα. Δηλαδή,

$$\max f(\langle Tv_1, v_1\rangle, \dots, \langle Tv_n, v_n\rangle) = \max\{f(\sigma(w)) : \sigma \in S_n\}, \tag{6.3.18}$$

όπου το αριστερό maximum παίρνεται πάνω από όλες τις ορθοκανονικές βάσεις του $\mathbb{R}^n$. □

Επιλέγοντας κατάλληλα τη συνάρτηση f στο προηγούμενο θεώρημα, παίρνουμε μία σειρά ανισοτήτων. Η πρώτη μας επιλογή είναι

$$f(s_1, \dots, s_n) = \sum_{i=1}^k s_i \tag{6.3.19}$$

για σταθερό $k \leq n$.

Θεώρημα 6.3.7. *Έστω T ένας $n \times n$ συμμετρικός πίνακας και έστω $\{v_1, \dots, v_n\}$ μία ορθοκανονική βάση του $\mathbb{R}^n$. Τότε,*

$$\sum_{i=1}^{k} \lambda_{n-i+1} \leq \sum_{i=1}^{k} \langle Tv_i, v_i \rangle \leq \sum_{i=1}^{k} \lambda_i \tag{6.3.20}$$

όπου $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ είναι οι ιδιοτιμές του T.

Απόδειξη. Η $f(s_1, \dots, s_n) = \sum_{i=1}^{k} s_i$ είναι κυρτή στον $\mathbb{R}^n$ (η $-f$ επίσης). Από το Θεώρημα 6.3.6 υπάρχει μετάθεση $\sigma \in S_n$ ώστε

$$\sum_{i=1}^{k} \langle Tv_i, v_i \rangle \leq \sum_{i=1}^{k} \lambda_{\sigma^{-1}(i)}. \tag{6.3.21}$$

Λόγω της διάταξης των λ_i, το δεξιό μέλος είναι μικρότερο ή ίσο από το $\sum_{i=1}^{k} \lambda_i$, οπότε έχουμε αποδείξει τη δεξιά ανισότητα του θεωρήματος. Δουλεύοντας ανάλογα με την $-f$ παίρνουμε την αριστερή ανισότητα. Παρατηρήστε ότι έχουμε ισότητα στην αριστερή ανισότητα αν το $v_i, i \leq k$ είναι ιδιοδιάνυσμα για την λ_{n-i+1}, ενώ στη δεξιά ανισότητα αν το v_i είναι ιδιοδιάνυσμα για την λ_i. □

Η επόμενη επιλογή μας είναι η συνάρτηση

$$f(s_1, \dots, s_n) = (s_1 s_2 \cdots s_k)^{1/k} \tag{6.3.22}$$

που ορίζεται για μη αρνητικά s_i και σταθερό $k \leq n$. Η f είναι κοίλη, οπότε παίρνουμε το εξής:

Θεώρημα 6.3.8. *Αν T είναι ένας $n \times n$ συμμετρικός πίνακας και $\{v_1, \dots, v_n\}$ είναι μία ορθοκανονική βάση του $\mathbb{R}^n$, τότε*

$$\prod_{i=1}^{k} \lambda_{n-i+1} \leq \prod_{i=1}^{k} \langle Tv_i, v_i \rangle \leq \left(\frac{1}{k} \sum_{i=1}^{k} \lambda_i \right)^k \tag{6.3.23}$$

όπου $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ είναι οι ιδιοτιμές του T.

Απόδειξη. Η συνάρτηση f που χρησιμοποιούμε είναι κοίλη, άρα η $-f$ είναι κυρτή στο P^n. Από το Θεώρημα 6.3.6,

$$\left(\prod_{i=1}^{k} \lambda_{n-i+1} \right)^{1/k} \leq \left(\prod_{i=1}^{k} \langle Tv_i, v_i \rangle \right)^{1/k}, \tag{6.3.24}$$

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 τη δεξιά ανισότητα εφαρμόζουμε την ανισότητα αριθμητικού-γεωμετρικού μέσου και το Θεώρημα 6.3.7:

$$\left(\prod_{i=1}^{k} \langle Tv_i, v_i \rangle \right)^{1/k} \leq \frac{1}{k} \sum_{i=1}^{k} \langle Tv_i, v_i \rangle \leq \frac{1}{k} \sum_{i=1}^{k} \lambda_i. \tag{6.3.25}$$ □

Άμεσες συνέπειες είναι δύο πολύ γνωστές ανισότητες για ορίζουσες.

Θεώρημα 6.3.9 (ανισότητα του Hadamard). *Αν T είναι ένας $n \times n$ πίνακας με συντεταγμένες t_{ij}, τότε*

$$(\det T)^2 \leq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} t_{ij}^2 \right). \tag{6.3.26}$$

Αν ο T είναι συμμετρικός και θετικά ημιορισμένος,

$$\det T \leq \prod_{i=1}^{n} t_{ii}. \tag{6.3.27}$$

Απόδειξη. Δείχνουμε πρώτα τη δεύτερη ανισότητα. Η ορίζουσα ενός πίνακα είναι ίση με το γινόμενο των ιδιοτιμών του, αν λοιπόν θεωρήσουμε τη συνήθη ορθοκανονική βάση $\{e_i\}$ και χρησιμοποιήσουμε το Θεώρημα 6.3.8 με $k = n$, παίρνουμε

$$\det T = \prod_{i=1}^{n} \lambda_i \leq \prod_{i=1}^{n} \langle T e_i, e_i \rangle = \prod_{i=1} t_{ii}. \tag{6.3.28}$$

Αν εφαρμόσουμε αυτή την ανισότητα για το συμμετρικό και θετικά ημιορισμένο πίνακα $S = T^t T$ (όπου τώρα T τυχών $n \times n$ πίνακας), παίρνουμε

$$(\det T)^2 = \det S \leq \prod_{j=1}^{n} s_{jj} = \prod_{j=1}^{n} \left(\sum_{i=1}^{n} t_{ij}^2 \right). \qquad \square \tag{6.3.29}$$

Θεώρημα 6.3.10 (ανισότητα του Minkowski). *Αν T και S είναι συμμετρικοί θετικά ημιορισμένοι $n \times n$ πίνακες, τότε*

$$[\det(T + S)]^{1/n} \geq (\det T)^{1/n} + (\det S)^{1/n}. \tag{6.3.30}$$

Απόδειξη. Έστω $v_1, \ldots, v_n$ μία ορθοκανονική βάση από ιδιοδιανύσματα του $T + S$. Τότε,

$$\begin{aligned}
[\det(T + S)]^{1/n} &= \left[\prod_{i=1}^{n} \langle (T + S) v_i, v_i \rangle \right]^{1/n} \\
&= \left[\prod_{i=1}^{n} \langle T v_i, v_i \rangle + \langle S v_i, v_i \rangle \right]^{1/n} \\
&\geq \left[\prod_{i=1}^{n} \langle T v_i, v_i \rangle \right]^{1/n} + \left[\prod_{i=1}^{n} \langle S v_i, v_i \rangle \right]^{1/n} \\
&\geq (\det T)^{1/n} + (\det S)^{1/n}.
\end{aligned}$$

Η προτελευταία ανισότητα είναι συνέπεια της ανισότητας αριθμητικού-γεωμετρικού μέσου, ενώ η τελευταία προκύπτει από το Θεώρημα 6.3.9. $\square$