Κεφάλαιο 3

Διαχωρίζον Υπερεπίπεδο–Λήμμα Farkas

3.1 Κλειστά Σύνολα

Ξεκινάμε με κάποιες έννοιες, γνωστές από την μαθηματική ανάλυση. Ένα σύνολο $A \subset \mathbb{R}^n$ ονομάζεται *κλειστό* αν περιέχει κάθε οριακό σημείο του. Το σύνολο ονομάζεται *συμπαγές* αν κάθε ακολουθία στοιχείων του περιέχει υπακολουθία συγκλίνουσα σε σημείο του συνόλου.

Θεώρημα 14. *Άν $A \subset \mathbb{R}^n$ ένα συμπαγές υποσύνολο του $\mathbb{R}^n$ και $f : A \to \mathbb{R}^m$ συνεχής συνάρτηση τότε το $f(A)$ είναι συμπαγές υποσύνολο του $\mathbb{R}^m$.*

Απόδειξη Έστω $\{f(a_k)\}$ μια ακολουθία σημείων του $f(A)$ και $\{a_k\}$ μια από τις ακολουθίες σημείων του A από τις οποίες θα μπορούσε να προέρχεται. (Η συνάρτηση f δεν είναι υποχρεωτικά 1–1.) Αφού το A είναι συμπαγές υπάρχει υπακολουθία $\{a_{i_k}\}$ τέτοια ώστε $a_{i_k} \to a \in A$. Συνεπώς, αφού η f είναι συνεχής, $f(a_{i_k}) \to f(a)$ και $f(a) \in f(A)$. Αυτό σημαίνει ότι το $f(A)$ είναι συμπαγές. ∎

Πόρισμα 2. *Άν $A \subset \mathbb{R}^n$ ένα συμπαγές υποσύνολο του $\mathbb{R}^n$ και $f : A \to \mathbb{R}$ είναι συνεχής συνάρτηση τότε υπάρχουν $a, b \in A$ τέτοια ώστε $f(a) = \inf\{f(x) : x \in A\}$ και $f(b) = \sup\{f(x) : x \in A\}$.*

Απόδειξη Το θεώρημα δείχνει ότι το σύνολο $f(A) = \{f(x) : x \in A\}$ είναι συμπαγές και συνεπώς κλειστό και φραγμένο. Επομένως έχει πεπερασμένο infimum και supremum. Επιπλέον τόσο το infimum όσο και το supremum ανήκουν στο $f(A)$ μια και το $f(A)$ είναι κλειστό σύνολο. Συμπεραίνουμε ότι υπάρχουν $a, b \in A$ τέτοια ώστε $f(a) = \inf f(A)$ και $f(b) = \sup f(A)$. ∎

Ορισμός 2. *Η απόσταση του σημείου x από το σύνολο A στον $\mathbb{R}^n$ ορίζεται ως $d_A(x) = \inf\{\|x-y\| : y \in A\}$.*

Πρόταση 2. *Η συνάρτηση απόστασης ικανοποιεί την συνθήκη Lipschitz*

$$| d_A(y) - d_A(x) | \leq \|x-y\|. \tag{3.1}$$

Η παραπάνω σχέση συνεπάγεται φυσικά ότι η $d_A(\cdot)$ είναι συνεχής συνάρτηση.

Απόδειξη Έστω $x, y \in \mathbb{R}^n$. Για κάθε $\epsilon > 0$ υπάρχει $a \in A$ τέτοιο ώστε $\|x-a\| < d_A(x) + \epsilon$. Επίσης

$$d_A(y) \leq \|y-a\| \leq \|y-x\| + \|x-a\| < \|y-x\| + d_A(x) + \epsilon.$$

Αφού η παραπάνω σχέση ισχύει για κάθε $\epsilon > 0$ θα ισχύει επίσης ότι $d_A(y) \leq \|y-x\| + d_A(x)$ ή

$$d_A(y) - d_A(x) \leq \|y-x\|.$$

Εναλλάσσοντας τον ρόλο των x και y θα έχουμε επίσης ότι

$$d_A(x) - d_A(y) \leq \|y-x\|.$$

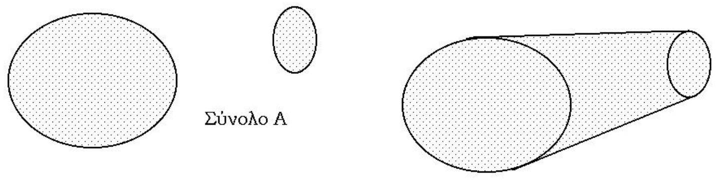
Από τις δύο σχέσεις προκύπτει η (3.1). ∎

Πρόταση 3. *Έστω A ένα μη κενό, κλειστό υποσύνολο του $\mathbb{R}^n$ και $x \in \mathbb{R}^n$. Τότε υπάρχει $a_0 \in A$ τέτοιο ώστε $d_A(x) = \|x-a_0\|$.*

Απόδειξη Υπάρχει ακολουθία $\{a_k\}$ σημείων του A τέτοια ώστε $\|x-a_k\| \to d_A(x)$. Οι συγκλίνουσες ακολουθίες στο $\mathbb{R}$ είναι φραγμένες και συνεπώς υπάρχει $r > 0$ τέτοιο ώστε $\|x-a_k\| \leq r$ για κάθε $k \in \mathbb{N}$. Συνεπώς, από την τριγωνική ιδιότητα ισχύει ότι $\|a_k\| \leq \|a_k - x\| + \|x\| \leq r + \|x\|$. Αυτό σημαίνει ότι η ακολουθία $\{a_k\}$ είναι φραγμένη και επομένως υπάρχει συγκλίνουσα υπακολουθία $\{a_{i_k}\}$ τέτοια ώστε $a_{i_k} \to a_0 \in \mathbb{R}^n$. Επειδή το A είναι κλειστό σύνολο, $a_0 \in A$. Έχουμε λοιπόν ότι $\|x-a_{i_k}\| \to \|x-a_0\|$ όταν $k \to \infty$. Επίσης, $\|x-a_{i_k}\| \to d_A(x)$ αφού $\{a_{i_k}\}$ είναι υπακολουθία της αρχικής ακολουθίας. Από την μοναδικότητα του ορίου προκύπτει επίσης ότι $d_A(x) = \|x-a_0\|$. ∎

Θεώρημα 15. *Έστω A, B, μη κενά υποσύνολα του $\mathbb{R}^n$ τέτοια ώστε το A κλειστό και το B συμπαγές. Τότε υπάρχει $a_0 \in A$, $b_0 \in B$ τέτοια ώστε $\|a_0 - b_0\| = \inf\{\|a-b\| : a \in A, b \in B\}$.*

Απόδειξη Η συνάρτηση $d_A(\cdot)$ είναι συνεχής. Αν περιορίσουμε το πεδίο ορισμού της στο συμπαγές σύνολο B τότε $d_A(b_0) = \inf\{d_A(b) : b \in B\}$ για κάποιο $b_0 \in B$. Αφού το A είναι κλειστό υπάρχει $a_0 \in A$ τέτοιο ώστε $d_A(b_0) = \|a_0 - b_0\|$. Για κάθε $a \in A$, $b \in B$, $\|a-b\| \geq d_A(b) \geq d_A(b_0) = \|a_0 - b_0\|$. Συνεπώς, $\|a_0 - b_0\| = \inf\{\|a-b\| : a \in A, b \in B\}$ αφου $a_0 \in A$, $b_0 \in B$. ∎



Σχήμα 3.1: Η κυρτή θήκη

3.2 Κυρτά Σύνολα

Ένα υποσύνολο C του $\mathbb{R}^n$ ονομάζεται *κυρτό* αν, για κάθε $x, y \in C$ και κάθε $\lambda \in [0,1]$, $\alpha x + (1-\alpha)y \in C$.

Αν a_i, $i = 1, 2, \ldots, m$ είναι διανύσματα στον $\mathbb{R}^n$, το διάνυσμα $\sum_{i=1}^m \lambda_i a_i$ ονομάζεται *κυρτός συνδυασμός* των a_i αν $\lambda_i \geq 0$ και $\sum_{i=1}^m \lambda_i = 1$.

Αν $\{A_i; i \in I\}$ είναι μια οικογένεια από κυρτά σύνολα, τότε $\bigcap_{i\in I} A_i$ είναι επίσης κυρτό σύνολο. Αυτό είναι εύκολο να το διαπιστώσει κανείς αφού για $x, y \in \bigcap_{i\in I} A_i$ θα έχουμε $x, y \in A_i$ για κάθε $i \in I$ και αφού τα A_i είναι κυρτά $\lambda x + (1-\lambda)y \in A_i$ για κάθε i άρα $\lambda x + (1-\lambda)y \in \bigcap_{i\in I} A_i$.

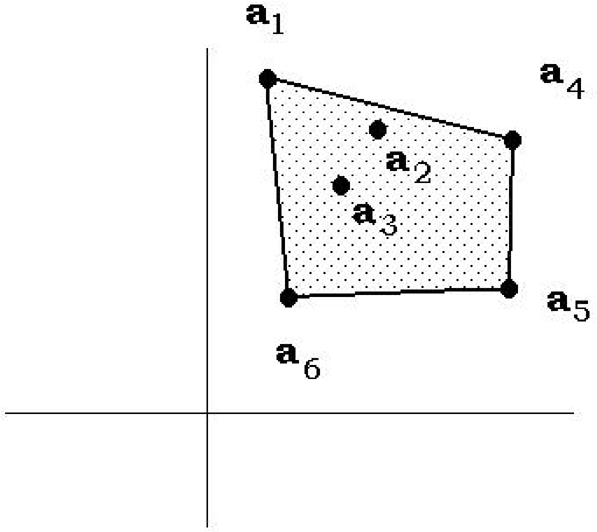
Η *κυρτή θήκη* του συνόλου $A \in \mathbb{R}^n$ την οποία συμβολίζουμε ως $\mathrm{conv}A$ είναι το μικρότερο κυρτό σύνολο που περιέχει το A δηλαδή η τομή όλων των κυρτών συνόλων που περιέχουν το A. (Η τομή αυτή είναι κυρτή με βάση την προηγούμενη παρατήρηση.)

Κάθε κυρτός συνδυασμός στοιχείων ενός κυρτού συνόλου A ανήκει στο σύνολο A. Αυτό μπορούμε να το διατυπώσουμε ως εξής:

Θεώρημα 16. *Έστω $a_1, \ldots, a_m$ σημεία ενός κυρτού συνόλου $A \in \mathbb{R}^n$. Έστω επίσης $\lambda_1, \ldots, \lambda_m \geq 0$ με $\lambda_1 + \cdots + \lambda_m = 1$. Τότε $\sum_{i=1}^m \lambda_i a_i \in A$.*

Απόδειξη Με επαγωγή ως προς m. Το θεώρημα ισχύει τετριμένα για $m = 1$ και από τον ορισμό της κυρτότητας για $m = 2$. Υποθέτουμε ότι ισχύει για $m = k \geq 2$ και θα δείξουμε ότι ισχύει για $m = k + 1$, δηλαδή ότι $x = \sum_{j=1}^{k+1} \lambda_j a_j \in A$. Αν $\mu := \lambda_1 + \cdots + \lambda_k = 0$ τότε ασφαλώς $x = a_{k+1} \in A$. Έστω λοιπόν $\mu > 0$. Ορίζουμε $y = \frac{\lambda_1}{\mu} a_1 + \cdots + \frac{\lambda_k}{\mu} a_k$. Θα ισχύει ότι $y \in A$ από την επαγωγική υπόθεση. Αλλά $x = \mu y + (1-\mu)a_{k+1} \in A$ από την κυρτότητα του A. ∎

Θεώρημα 17. *Έστω A ένα υποσύνολο του $\mathbb{R}^n$. Τότε conv A είναι το σύνολο όλων των κυρτών συνδυασμών των σημείων του A.*



Σχήμα 3.2: Η κυρτή θήκη του συνόλου $\{a_1, a_2, a_3, a_4, a_5, a_6\}$

Απόδειξη Έστω B το σύνολο όλων των κυρτών συνδυασμών των σημείων του A δηλαδή $B = \{\sum_{i=1}^{m} \lambda_i a_i; m \in \mathbb{N}, a_i \in A, \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1\}$. Είναι πολύ εύκολο να διαπιστώσει κανείς ότι το σύνολο B είναι κυρτό. Πράγματι, αν $x = \sum_{i=1}^{m} \lambda_i a_i$, $y = \sum_{i=1}^{m'} \lambda'_i a'_i$ είναι δύο κυρτοί συνδυασμοί στοιχείων του A και $\mu \in [0,1]$, $\mu' := 1 - \mu$, τότε $\mu x + \mu' y = \mu\lambda_1 a_1 + \cdots + \mu\lambda_m a_m + \mu'\lambda'_1 a'_1 + \cdots + \mu'\lambda'_{m'} a'_{m'}$ είναι επίσης κυρτός συνδυασμός στοιχείων του A και επομένως ανήκει στο B. Επίσης ισχύει προφανώς ότι $A \subset B$. Συνεπώς αφού $\mathrm{conv}A$ είναι το μικρότερο κυρτό σύνολο που περιέχει το A ισχύει ότι $\mathrm{conv}A \subset B$. Αντίστροφα, κάθε στοιχείο του B είναι κυρτός συνδυασμός στοιχείων του A και επομένως ανήκει σε κάθε κυρτό σύνολο που περιέχει το A, επομένως και στην τομή τους, δηλαδή το $\mathrm{conv}A$. ∎

Από το παραπάνω έχουμε και τα εξής προφανή πορίσματα:

Πόρισμα 3. *Αν ένα σημείο x ανήκει στην κυρτή θήκη του A τότε υπάρχει $m \in \mathbb{N}$ και σημεία $a_1, \ldots, a_m$ του A τέτοια ώστε $x = \sum_{i=1}^{m} \lambda_i a_i$, δηλαδή το x μπορεί να γραφεί ως κυρτός συνδυασμός στοιχείων του A.*

Πόρισμα 4. *Η κυρτή θήκη του συνόλου $A = \{a_1, \ldots, a_m\}$ είναι το σύνολο $\{\sum_{i=1}^{m} \lambda_i a_i, \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1\}$.*

3.3 Το θεώρημα του διαχωρίζοντος υπερεπιπέδου

Λήμμα 3. *Έστω $x, y \in \mathbb{R}^n$. Αν για κάποιο $\alpha > 0$, $\|x + \lambda y\| \geq \|x\|$ για κάθε $0 < \lambda < \alpha$, τότε $x^T y \geq 0$.*

Απόδειξη Ισχύει ότι

$$\|x\|^2 \leq \|x+\lambda y\|^2 = \|x\|^2 + 2\lambda x^T y + \lambda^2 \|y\|^2$$

και συνεπώς, αφού $\lambda > 0$, $x^T y + \frac{1}{2}\lambda\|y\|^2 \geq 0$. Αφήνοντας το $\lambda \to 0$ συμπεραίνουμε ότι $x^T y \geq 0$. ∎

Θεώρημα 18. *Έστω $A \subset \mathbb{R}^n$ μη κενό, κυρτό, κλειστό σύνολο και x σημείο του $\mathbb{R}^n$. Τότε υπάρχει ένα μοναδικό σημείο $a_0 \in A$ τέτοιο ώστε $\|x - a_0\| = \inf\{\|x - z\| : z \in A\}$. Επιπλέον, $(x-a_0)^T(a-a_0) \leq 0$ για κάθε $a \in A$.*

Απόδειξη Από προηγούμενα αποτελέσματα για κλειστά σύνολα ξέρουμε ότι υπάρχει $a_0 \in A$ τέτοιο ώστε $\|x-a_0\| = \inf\{\|x-z\| : z \in A\}$. Έστω $a \in A$ και $0 \leq \lambda \leq 1$. Η κυρτότης του A έχει ως συνέπεια ότι $(1-\lambda)a_0 + \lambda a \in A$. Ισχύει ότι

$$\|x - ((1-\lambda)a_0 + \lambda a)\| = \|(x-a_0) + \lambda(a_0 - a)\|^2 \geq \|x - a_0\|,$$

όπου, η τελευταία ανισότητα οφείλεται στον ορισμό του a_0 ως του εγγύτερου σημείου του A στο x. Από το προηγούμενο λήμμα συμπεραίνουμε ότι $(x-a_0)^T(a-a_0) \leq 0$. Για να αποδείξουμε την μοναδικότητα του a_0 έστω $a_1 \in A$ ένα άλλο σημείο τέτοιο ώστε $\|x-a_1\| = \inf\{\|x-z\| : z \in A\}$. Εφαρμόζοντας τον ίδιο συλλογισμό έχουμε $(x-a_1)^T(a-a_1) \leq 0$. Δεδομένου ότι το a είναι οποιοδήποτε σημείο του A, θέτοντας $a = a_0$ παίρνουμε

$$(x-a_1)^T(a_0-a_1) \leq 0. \tag{3.2}$$

Από την συμμετρία, εναλλάσσοντας τον ρόλο των a_0 και a_1 έχουμε

$$(x-a_0)^T(a_1-a_0) \leq 0. \tag{3.3}$$

Προσθέτοντας κατά μέλη τις ανισότητες (3.2), (3.3), παίρνουμε $(a_1-a_0)^T(a_1-a_0) = \|a_1-a_0\|^2 \leq 0$ απ' όπου προκύπτει ότι $a_1 = a_0$. ∎

Ένα υπερεπίπεδο $\alpha^T z + \beta = 0$ στον $\mathbb{R}^n$ *διαχωρίζει* δύο υποσύνολα του $\mathbb{R}^n$ αν $\alpha^T z + \beta \geq 0$ για κάθε $z \in A$ και $\alpha^T z + \beta \leq 0$ για κάθε $z \in B$. Αν οι παραπάνω ανισότητες είναι αυστηρές τότε το υπερεπίπεδο *διαχωρίζει αυστηρά* τα δύο σύνολα.

Θεώρημα 19. *Έστω A, B, ξένα, μη κενά, κυρτά σύνολα στον $\mathbb{R}^n$ όπου A κλειστό και B συμπαγές. Τότε τα A και B διαχωρίζονται αυστηρά από ένα υπερεπίπεδο στο $\mathbb{R}^n$.*

Απόδειξη Έστω $a \in A$, $b \in B$ τέτοια ώστε το a να είναι το εγγύτερο σημείο του A στο B και το b το εγγύτερο σημείο του B στο A. Αυτό είναι συνέπεια των προηγουμένων θεωρημάτων. Αφού $A \cap B = \emptyset$, $a \neq b$. Αν $x \in A$, $y \in B$, τότε από το προηγούμενο θεώρημα $(b-a)^T(x-a) \leq 0$ και $(a-b)^T(y-b) \leq 0$. Συνεπώς

$$\begin{aligned}(a-b)^T x &\geq (a-b)^T a = \frac{1}{2}\left(\|a\|^2 - \|b\|^2 + \|a-b\|^2\right) > \frac{1}{2}\left(\|a\|^2 - \|b\|^2\right) \\ &> \frac{1}{2}\left(\|a\|^2 - \|b\|^2 - \|a-b\|^2\right) = (a-b)^T b \geq (a-b)^T y.\end{aligned}$$

Θέτουμε $\theta = a - b$ και $\gamma = -\frac{1}{2}(\|a\|^2 - \|b\|^2)$. Τότε η παραπάνω σχέσεις μας δίνουν

$$\theta^T x + \gamma > 0 > \theta^T y + \gamma.$$

Συνεπώς το διαχωρίζον υπερεπίπεδο είναι το $\theta^T z + \gamma = 0$. ∎

3.4 Το Λήμμα του Farkas

Θεώρημα 20 (Λήμμα του Farkas). *Αν A είναι ένας πίνακας $m \times n$ και b ένα διάνυσμα στον $\mathbb{R}^m$ τότε ακριβώς μία από τις ακόλουθες δύο εναλλακτικές προτάσεις ισχύει*

(*i*) *Υπάρχει $x \in \mathbb{R}^n$, $x \geq 0$ τέτοιο ώστε $Ax = b$.*

(*ii*) *Υπάρχει $y \in \mathbb{R}^m$ τέτοιο ώστε $y^T A \geq 0$ και $y^T b < 0$.*

Απόδειξη Πρώτα απ' όλα διαπιστώνουμε ότι δεν μπορουν αν ισχύσουν τα (i) και (ii) ταυτόχρονα. Πράγματι, σ' αυτή την περίπτωση, για κάποιο $x \geq 0$ $Ax = b$ ενώ ταυτόχρονα υπάρχει y που να ικανοποιεί την (ii). Γι' αυτό το y έχουμε $y^T Ax = y^T b$. Το δεξί μέλος αυτής της τελευταίας σχέσης είναι από την (ii) αυστηρά αρνητικό ενώ το αριστερό είναι το εσωτερικό γινόμενο δύο μη αρνητικών διανυσμάτων, του $A^T y$ και του x.

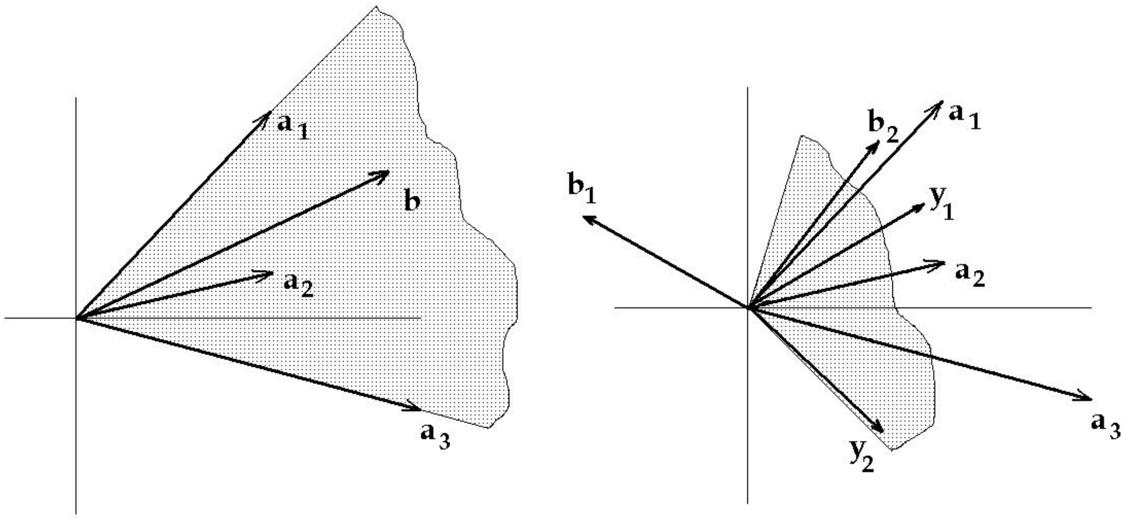
Για να αποδείξουμε το θεώρημα αρκεί να δείξουμε ότι αν δεν ισχύει η (i) τότε θα ισχύει υποχρεωτικά η (ii). Αν δεν ισχύει η (i) τότε το b βρίσκεται έξω από τον κυρτό κλειστό κώνο $C := \{Ax : x \geq 0\}$. Συνεπώς υπάρχει διαχωρίζον υπερεπίπεδο δηλαδή ένα υπερεπίπεδο $\gamma^T z + \beta = 0$ όπου $\gamma \in \mathbb{R}^m$ και $\beta \in \mathbb{R}$ τέτοιο ώστε

$$\gamma^T z + \beta > 0 \qquad \forall z \in C, \tag{3.4}$$
$$\gamma^T b + \beta < 0. \tag{3.5}$$

Αν $z \in C$ και $r > 0$ τότε και $rz \in C$, συνεπώς από την (3.4) $r\gamma^T z + \beta > 0$ για κάθε $z \in C$ και $r > 0$ ή $\gamma^T z + \beta/r > 0$. Αφήνοντας $r \to \infty$ βλέπουμε ότι $\gamma^T z \geq 0$ για όλα τα $z \in C$. Αυτό ισχύει εν προκειμένω και για τις στήλες $a_1, \ldots, a_n$ του πίνακα A. Συνεπώς $\gamma^T a_i \geq$ και μπορούμε να διαλέξουμε το διάνυσμα y που πρέπει να βρούμε για να ικανοποιήσουμε την (ii) ίσο με το γ. Ακόμη, θέτωντας $z = 0$ στην (3.4) βλέπουμε ότι $\beta > 0$. Συνεπώς από την (3.5) έχουμε $y^T b = -\beta < 0$. ∎

Οι εφαρμογές του λήμματος του Farkas είναι πάρα πολλές. Θα δώσουμε μερικές άμεσα ενώ άλλες θα δούμε στην συνέχεια.

Θεώρημα 21 (Θεώρημα εναλλακτικών του Fredholm). *Αν A είναι ένας πίνακας $m \times n$ και $b \in \mathbb{R}^m$ τότε ακριβώς μία από τις δύο εναλλακτικές προτάσεις ισχύει*



Σχήμα 3.3: Στο αριστερό σχήμα βλέπουμε την πρώτη περίπτωση του λήμματος του Farkas. Το διάνυσμα b ανήκει στον κώνο που σχηματίζουν τα a_1, a_2, a_3. Στο δεξιό σχήμα βλέπουμε την δεύτερη περίπτωση του λήμματος. Τα b_1, b_2 δεν ανήκουν στον κώνο των a_1, a_2, a_3 και επομένως υπάρχουν y_1, y_2 τέτοια ώστε $y^T a_i \geq 0$ και $y^T b < 0$.

(*i*) *Υπάρχει* $x \in \mathbb{R}^n$ *τέτοιο ώστε* $Ax = b$.

(*ii*) *Υπάρχει* $y \in \mathbb{R}^m$ *τέτοιο ώστε* $y^T A = 0$ *και* $y^T b \neq 0$.

Απόδειξη Στην πρώτη εναλλακτική πρόταση το x μπορεί να παίρνει τόσο θετικές όσο και αρνητικές τιμές. Για να εφαρμόσουμε το Λήμμα του Farkas πρέπει να μετατρέψουμε το πρόβλημα σε ένα πρόβλημα αναζήτησης θετικών λύσεων. Αυτό γίνεται εύκολα θέτοντας $x = u - v$ όπου $u, v \geq 0$. (Αυτό το τέχνασμα έχει ευρύτατη εφαρμογή όπως θα γίνει σαφές στη συνέχεια.) Η πρώτη εναλλακτική πρόταση γίνεται τότε $[A \mid -A][{u \atop v}] = b$ με $u, v \geq 0$ και είναι στη μορφή του Λήμματος του Farkas. Η εναλλακτική πρόταση του Λήμματος του Farkas είναι ότι υπάρχει $y \in \mathbb{R}^m$ τέτοιο ώστε $y^T[A \mid -A] \geq 0$ και $y^T b < 0$. Αλλά αυτό σημαίνει $y^T A \geq 0$ και $y^T A \leq 0$ ή $y^T A = 0$. Η δεύτερη συνθήκη, μπορεί απλά να γραφεί $y^T b \neq 0$. (Παρατηρείστε ότι αν για κάποιο y, $y^T A = 0$, τότε και $(-y)^T A = 0$ και συνεπώς το πρόσημο δεν παίζει ρόλο.) ∎

Το ακόλουθο θεώρημα εξασφαλίζει την ύπαρξη στάσιμης κατανομής σε μια αλυσίδα Markov με πεπερασμένο αριθμό καταστάσεων. Υπενθυμίζουμε ότι ο πίνακας $n \times n$, P, ονομάζεται *στοχαστικός* αν $P_{ij} \geq 0$ και $\sum_{j=1}^{n} P_{ij} = 1$ για κάθε $i = 1, \ldots, n$.

Θεώρημα 22. *Αν* P *είναι στοχαστικός πίνακας υπάρχει διάνυσμα γραμμής* $\pi = (\pi_1, \ldots, \pi_n)$

τέτοιο ώστε $\pi_i \geq 0$ *και* $\sum_{i=1}^n \pi_i = 1$ *το οποίο να ικανοποιεί τις εξισώσεις ισορροπίας*

$$\pi_j = \sum_{i=1}^n \pi_i P_{ij}, \qquad j = 1, \ldots, n. \tag{3.6}$$

Απόδειξη Αρκεί να δείξουμε ότι υπάρχει διάνυσμα $x \in \mathbb{R}^n$ με $x \geq 0$ τέτοιο ώστε $P^T x = x$ και $u^T x = 1$ όπου $u^T = [1, 1, \ldots, 1]$. Ισοδύναμα αρκει να δείξουμε ότι η $Ax = b$ έχει λύση $x \geq 0$ με

$$A = \begin{bmatrix} P^T - I \\ u^T \end{bmatrix}, \qquad \text{και} \qquad b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Πράγματι αν υπάρχει τέτοια λύση τότε έχουμε απλώς $x^T = \pi$. Αν δεν υπάρχει τέτοια λύση τότε θα πρέπει να ισχύει η δεύτερη εναλλακτική προταση του Farkas η οποία στην περίπτωσή μας λέει ότι θα πρέπει να υπάρχει $y \in \mathbb{R}^{n+1}$ τέτοιο ώστε $y^T A \geq 0$ και $y^T b < 0$. Αρχίζοντας από την δεύτερη σχέση γράφουμε $y^T = [z_1, \ldots, z_n, -\lambda]$ και παρατηρούμε ότι $y^T b = -\lambda < 0$, συνεπώς $\lambda > 0$. (Αυτό δικαιολογεί και την περίεργη επιλογή μας για τον συμβολισμό των συνιστωσών του διανύσματος y.) Στρέφουμε τώρα την προσοχή μας στην $y^T A \geq 0$ η οποία σημαίνει ότι

$$y^T \begin{bmatrix} P^T - I \\ u^T \end{bmatrix} \geq 0.$$

Η παραπάνω σχέση γράφεται και ως

$$\sum_{j=1}^n P_{ij} z_j - z_i - \lambda > 0, \qquad i = 1, \ldots, n. \tag{3.7}$$

Έστω m εκείνος ο δείκτης για τον οποίο $z_m = \max_{j=1,\ldots,n} z_j$. Εφόσον $z_m \geq z_j$ για κάθε j θα ισχύει και ότι

$$z_m \geq \sum_{j=1}^n P_{mj} z_j. \tag{3.8}$$

Εφαρμόζοντας την (3.7) για $i = m$ παίρνουμε την

$$\sum_{j=1}^n P_{mj} z_j > z_m + \lambda > z_m. \tag{3.9}$$

Η αντίφαση μεταξύ των (3.8) και (3.9) δείχνει ότι η δεύτερη εναλλακτική πρόταση του Farkas δεν μπορεί να ισχύει και συνεπώς αναγκαστικά ισχύει η πρώτη. ∎

Τέλος θα δώσουμε μια εναλλακτική διατύπωση του Λήμματος του Farkas.

Θεώρημα 23 (Λήμμα Farkas–Δεύτερη διατύπωση)**.** *Έστω* $A = (a_{ij})$ $m \times n$ *πίνακας. Τότε είτε το i) είτε το ii) αλλά όχι και τα δύο ισχύουν* *i) Το σημείο $0 \in \mathbb{R}^m$ περιέχεται στην κυρτή θήκη των $m+n$ σημείων*

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ \ldots, \ a_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}, \ e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, \ e_m = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

(όπου e_i το μοναδιαίο διάνυσμα στην κατεύθυνση i στον $\mathbb{R}^m$).

ii) Υπάρχουν αριθμοί $x_1, \ldots, x_m$ τέτοιοι ώστε $x_i > 0$, $\sum_{i=1}^m x_i = 1$ και $\sum_{i=1}^m x_i a_{ij} > 0$ για $j = 1, 2, \ldots, n$.

Απόδειξη Το 0 ανήκει στην κυρτή θήκη των διανυσμάτων της (i) αν υπάρχουν $s_j \geq 0$ τέτοια ώστε $\sum_{j=1}^n s_j a_j + \sum_{i=1}^m s_{m+i} e_i = 0$ και $\sum_{j=1}^{m+n} s_j = 1$. Αν $A = [a_1 \mid \cdots \mid a_n]$, η πρώτη πρόταση γράφεται επίσης ως

$$\begin{bmatrix} A \ , \ I_m \\ e^T \end{bmatrix} s = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} =: b, \qquad s \geq 0. \tag{3.10}$$

I_m είναι ο $m \times m$ μοναδιαίος πίνακας και $e^T = (1, 1, \ldots, 1)$ είναι διάνυσμα με $m+n$ στοιχεία ίσα με την μονάδα. Ακόμη, $s := (s_1, s_2, \ldots, s_{m+n})^T$ ενώ το δεξί μέλος της παραπάνω εξίσωσης είναι ένα διάνυσμα με $m+1$ στοιχεία, τα πρώτα m από τα οποία είναι 0. Παρατηρείστε ότι η συνθήκη $\sum_{j=1}^{n+m} s_j = 1$ εξασφαλίζεται απο την τελευταία γραμμή του πίνακα της (3.10). Αν δεν ισχύει η πρώτη πρόταση τότε από το λήμμα του Farkas υπάρχει $y \in \mathbb{R}^{m+1}$ τέτοιο ώστε

$$y^T \begin{bmatrix} A \ , \ I_m \\ e^T \end{bmatrix} \geq 0 \tag{3.11}$$

και $y^T b < 0$. Θέτουμε $y^T := (y_1, \ldots, y_m, -\lambda)$. Δεδομένης της μορφής του b η σχέση $y^T b < 0$ συνεπάγεται ότι $\lambda > 0$. Συνεπώς, η (3.11) συνεπάγεται ότι

$$\sum_{i=1}^m y_i a_{ij} - \lambda \geq 0, \quad j = 1, 2, \ldots, n, \tag{3.12}$$

$$y_i - \lambda \geq 0, \quad i = 1, 2, \ldots, m. \tag{3.13}$$

Αλλά αυτή η τελευταία σχέση συνεπάγεται ότι $\sum_{i=1}^m y_i a_{ij} \geq \lambda > 0$ για κάθε $j = 1, \ldots, n$, και $y_i \geq \lambda > 0$ για κάθε $i = 1, 2, \ldots, m$. Μπορούμε συνεπώς να θέσουμε

$$x_i = \frac{y_i}{\sum_{k=1}^m y_k}, \quad i = 1, 2, \ldots, m.$$

Συνεπώς, αν δεν ισχύει η πρόταση (i) τότε υποχρεωτικά ισχύει η πρόταση (ii). ∎

Κεφάλαιο 4

Θεωρία Παιγνίων

4.1 Παίγνια μηδενικού αθροίσματος σε κανονική μορφή

Στο κεφάλαιο αυτό εξετάζουμε παίγνια μηδενικού αθροίσματος σε κανονική μορφή (zero sum games in canonical form). Τα παίγνια αυτά μπορούν να περιγραφούν από ένα πίνακα που δίδει την αμοιβή που δίδει ο παίκτης II στον παίκτη I. Για παράδειγμα, στον ακόλουθο πίνακα

$$\begin{array}{cc} & \text{παίκτης II: Ελαχιστοποιεί} \\ \text{παίκτης I: Μεγιστοποιεί} & \begin{pmatrix} 5 & 1 & 3 \\ 3 & 2 & 4 \\ -3 & 0 & 1 \end{pmatrix} \end{array} \tag{4.1}$$

ο παίκτης I επιλέγει γραμμή ενώ ο παίκτης II στήλη, και από τον συνδυασμό των δύο αποφάσεων προκύπτει η αμοιβή που θα πρέπει να πληρώσει ο II στον I (η οποία βεβαίως μπορεί να είναι και αρνητική). Στο παράδειγμα της (4.1) ο παίκτης I, αν παίξει πρώτος, θα πρέπει να επιλέξει την 2η γραμμή γιατί αυτή έχει το *μεγαλύτερο ελάχιστο.* Ο παίκτης II που παίζει δεύτερος αναγκαστικά θα επιλέξει τότε την δεύτερη στήλη και θα πληρώσει 2 μονάδες στον I. Οποιαδήποτε άλλη επιλογή από την πλευρά του I οδηγεί σε μικρότερη αμοιβή για τον I. Αντίστοιχα, αν ο παίκης II παίζει πρώτος, τότε θα πρέπει να επιλέξει την στήλη με το *μικρότερο μέγιστο*, δηλαδή την στήλη 2. Τότε ο παίκτης I που παίζει δεύτερος θα πρέπει αναγκαστικά να επιλέξει την γραμμή 2 και θα πάρει αμοιβή και πάλι 2 μονάδες. Στην περίπτωση αυτή δεν έχει σημασία ποιός από τους δύο παίκτες παίζει πρώτος. Υπάρχουν όμως άλλα παίγνια στα οποία το ποιός παίζει πρώτος (δηλαδή *η πληροφορία που μπορεί να έχει ένας παίκτης για τις αποφάσεις του άλλου*) έχει σημασία. Για παράδειγμα, στην περίπτωση του παιγνίου με τον ακόλουθο πίνακα

$$\begin{array}{cc} & \text{παίκτης II: Ελαχιστοποιεί} \\ \text{παίκτης I: Μεγιστοποιεί} & \begin{pmatrix} -3 & -2 & 4 \\ 2 & 3 & 1 \\ 4 & 0 & -2 \end{pmatrix} \end{array} \tag{4.2}$$

όταν ο παίκτης I παίζει πρώτος τότε θα επιλέξει την γραμμή 2 η οποία έχει το μεγαλύτερο ελάχιστο που είναι ίσο με 1. Ο παίκτης II διαλέγει τότε την στήλη 3 και δίνει αμοιβή 1 στον I. Όταν ο παίκτης II παίζει πρώτος τότε θα επιλέξει την στήλη με το μικρότερο μέγιστο, δηλαδή την στήλη 2, οπότε ο I διαλέγει την γραμμή 2 και παίρνει 3 μονάδες από τον II.

Λήμμα 4. *Αν X, Y, δύο σύνολα και $f : X \times Y \to \mathbb{R}$ τότε*

$$\sup_{x\in X} \inf_{y\in Y} f(x,y) \leq \inf_{y\in Y} \sup_{x\in X} f(x,y).$$

Απόδειξη Ισχύει ότι $\inf_{y\in Y} f(x,y) \leq f(x,y)$ για κάθε $(x,y) \in X \times Y$. Συνεπώς, παίρνωντας supremum ως προς $x \in X$, $\sup_{x\in X} \inf_{y\in Y} f(x,y) \leq \sup_{x\in X} f(x,y)$ για όλα τα $y \in Y$. Παίρνοντας τώρα το infimum ως προς $y \in Y$ στην τελευταία σχέση συμπληρώνει την απόδειξη. ∎

Ορισμός 3. *Αν ένας πίνακας $A = (a_{ij})$ ικανοποιεί την $\max_i \min_j a_{ij} = \min_i \max_j a_{ij}$ τότε θα λέμε ότι έχει σαγματικό σημείο. Αν αντιθέτως $\max_i \min_j a_{ij} < \min_i \max_j a_{ij}$ τότε ο πίνακας δεν έχει σαγματικό σημείο.*

Από την παραπάνω συζήτηση είναι σαφές ότι, αν ο πίνακας του παιγνίου έχει σαγματικό σημείο τότε οποιαδήποτε πληροφορία έχει ένας παίκτης για την απόφαση του άλλου δεν επηρεάζει την δική του στρατηγική. Σ' αυτή την περίπτωση το παίγνιο έχει λύση με *καθαρές στρατηγικές*. Αν ο πίνακας δεν έχει σαγματικό σημείο τότε η πληροφορία που μπορεί να έχει ο ένας παίκτης για την στρατηγική του άλλου επηρεάζει και την δική του στρατηγική. Το παίγνιο δεν έχει λύση με *καθαρές στρατηγικές*.

4.2 Μεικτές στρατηγικές και το θεμελιώδες θεώρημα

Μια μεικτή στρατηγική (ή τυχαιοποιημένη στρατηγική) είναι μια κατανομή πιθανότητας πάνω στο σύνολο των στρατηγικών ενός παίκτη. Ο παίκτης δηλαδή αποφασίζει όχι ποια στρατηγική θα επιλέξει αλλά ποια κατανομή στο χώρο των στρατηγικών θα επιλέξει. Ο λόγος για τον οποίο ένας παίκτης θα μπορούσε να αποφασίσει να επιλέξει μια μεικτή στρατηγική είναι επειδή αναγκάζεται να παίξει πρώτος. Για παράδειγμα, στον πίνακα (4.2) ο οποίος δεν έχει σαγματικό σημείο όταν ο I παίζει πρώτος κερδίζει 1 μονάδα ενώ όταν παίζει δεύτερος 3 μονάδες. Αν ο παίκτης I δηλώσει ότι θα επιλέξει τις τρεις δυνατές ενέργειές του με αντίστοιχες πιθανότητες x_1, x_2, x_3, τότε ο παίκτης II αναγκάζεται να επιλέξει έτσι ώστε να ελαχιστοποιήσει την μέση αμοιβή που θα δώσει στον I δηλαδή καλείται να επιλέξει το ελάχιστο ανάμεσα στα

$$(-3x_1 + 2x_2 + 4x_3\,,\ -2x_1 + 3x_2\,,\ 4x_1 + x_2 - 2x_3)\,.$$

Για παράδειγμα, αν $x_1 = 1/10$, $x_2 = 9/10$, $x_3 = 0$ τότε ο II έχει να επιλέξει το ελάχιστο ανάμεσα στα $(1.5, 2.5, 1.3)$, συνεπώς επιλέγει την τρίτη στήλη, και δίνει αμοιβή 1.3 στον I. Άρα, όταν ο παίκτης I «παίζει πρώτος», όταν δηλαδή έχει λόγους να πιστεύει ότι ο II γνωρίζει την στρατηγική του θα προτιμήσει μια μεικτή (τυχαιοποιημένη) στρατηγική ώστε να αυξήσει τα έσοδά του. Το ίδιο βέβαια ισχύει και για τον παίκτη II που μπορεί να έχει επίσης κίνητρο να υιοθετήσει μια μεικτή στρατηγική.

Σε ένα γενικό παίγνιο που περιγράφεται από ένα πίνακα $m \times n$, $A = (a_{ij})$, αν ο παίκτης I υιοθετήσει την μεικτή στρατηγική x_i, $i = 1, \ldots, m$ και ο II την y_j, $j = 1, \ldots, n$ τότε η μέση τιμή της αμοιβής του II προς τον I είναι $x^T Ay = \sum_{j=1}^{n} \sum_{i=1}^{m} x_i a_{ij} y_j$.

Όπως είδαμε $\max_i \min_j a_{ij} \leq \min_j \max_i a_{ij}$ με ισότητα όταν ο πίνακας έχει σαγματικό σημείο. Ο χώρος όλων των μεικτών στρατηγικών του I είναι το simplex $X := \{(x_1, \ldots, x_m) : x_i \geq 0,\ \sum_{i=1}^{m} x_i = 1\}$. Παρόμοια ο χώρος των μεικτών στρατηγικών του II είναι ο $Y := \{(y_1, \ldots, y_n) : x_i \geq 0,\ \sum_{j=1}^{n} y_j = 1\}$.

Θεώρημα 24 (Von Neumann). *Κάθε παίγνιο της ανωτέρω μορφής έχει λύση με μεικτές στρατηγικές, δηλαδή*

$$\max_{x\in X} \min_{y\in Y} x^T Ay = \min_{y\in Y} \max_{x\in X} x^T Ay. \tag{4.3}$$

Απόδειξη Θέτουμε $v_I = \max_{x\in X} \min_{y\in Y} x^T Ay$ και $v_{II} = \min_{y\in Y} \max_{x\in X} x^T Ay$ με σκοπό να δείξουμε ότι $v_I = v_{II}$. Από το λήμμα 4 ισχύει ότι $v_I \leq v_{II}$, συνεπώς αρκεί να δείξουμε ότι η σχέση αυτή ικανοποιείται με ισότητα.

Έστω $A = [a_1 \mid a_2 \mid \cdots \mid a_n]$ οι στήλες του πίνακα A. Για την απόδειξη του θεωρήματος θα χρησιμοποιήσουμε την δεύτερη διατύπωση του Λήμματος του Farkas. Αν το 0 ανήκει στην κυρτή θήκη των σημείων $a_1, \ldots, a_n, e_1, \ldots, e_m$ στον $\mathbb{R}^m$ τότε υπάρχουν πραγματικοί $s_j \geq 0$, $j = 1, 2, \ldots, m+n$, τέτοιοι ώστε

$$\sum_{j=1}^{n} a_{ij} s_j + s_{n+i} = 0,\ i = 1, 2, \ldots, m,$$
$$\sum_{j=1}^{m+n} s_j = 1.$$

Αλλά δεν είναι δυνατόν να έχουμε $s_j = 0$ για κάθε $j = 1, \ldots, n$ γιατί τότε τα μοναδιαία διανύσματα e_i θα ήταν γραμμικά εξαρτημένα. Συνεπώς $\sum_{j=1}^{n} s_j > 0$ και μπορούμε να ορίσουμε τις ποσότητες

$$y_j := \frac{s_j}{\sum_{j=1}^{n} s_j} \geq 0$$

για τις οποίες ισχύει ότι $\sum_{j=1}^{n} y_j = 1$. Συνεπώς, από τα παραπάνω έχουμε ότι

$$\sum_{j=1}^{n} a_{ij} y_j = -\frac{s_{n+i}}{\sum_{j=1}^{n} s_j} \leq 0 \qquad \forall i.$$

Άρα, για την συγκεκριμμένη επιλογή των y_j από τον παίκτη II, $x^T Ay \leq 0$ για οποιαδήποτε επιλογή των x_i από τον I, συνεπώς $\max_{x\in X} x^T Ay \leq 0$ για το συγκεκριμένο $y \in Y$ και $v_{II} = \min_{y\in Y} \max_{x\in X} x^T Ay \leq 0$ κατά μείζονα λόγο. Άρα έχουμε

$$v_I \leq v_{II} \leq 0. \tag{4.4}$$

Έστω τώρα ότι η πρώτη περίπτωση του λήμματος του Farkas δεν ισχύει. Τότε θα ισχύει υποχρεωτικά η δεύτερη. Αυτό σημαίνει ότι υπάρχει $x = (x_1, \ldots, x_m)$ τέτοιο ώστε $x_i > 0$ για κάθε $i = 1, \ldots, m$, $\sum_{i=1}^{m} x_i = 1$, και $\sum_{i=1}^{m} x_i a_{ij} > 0$ για $j = 1, \ldots, n$. Συνεπώς για το συγκεκριμένο x έχουμε $x^T Ay > 0$ για κάθε y και επομένως, επαναλαμβάνοντας τον προηγούμενο συλλογισμό έχουμε $v_I > 0$ και

$$0 < v_I \leq v_{II}. \tag{4.5}$$

Εφόσον υποχρεωτικά πρέπει να ισχύει είτε η (4.4) είτε η (4.5) συμπεραίνουμε ότι είναι αδύνατον να ισχύει η $v_I \leq 0 < v_{II}$.

Έστω τώρα το παίγνιο που αντιστοιχεί στον πίνακα $B := (b_{ij})$ όπου $b_{ij} = a_{ij} + k \ \forall i, j$. Προφανώς $x^T By = x^T Ay + k$ για κάθε x, y και συνεπώς $v_I(B) = v_I(A) + k$, $v_{II}(B) = v_{II}(A) + k$. Αφού $v_I(B) \leq 0 < v_{II}(B)$ είναι αδύνατον, $v_I(B) \leq -k < v_{II}(B)$ είναι αδύνατον για κάθε $k \in \mathbb{R}$ που συνεπάγεται ότι $v_I < v_{II}$ είναι αδύνατον. Άρα, αναγκαστικά, $v_I = v_{II}$. ∎

4.3 Γραφικός υπολογισμός της τιμής

Για να κατανοήσουμε καλύτερα μερικές από τις έννοιες εξετάζουμε περιπτώσεις που ο ένας τουλάχιστον από τους δύο παίκτες έχει μόνο δύο επιλογές. Σ' αυτή την περίπτωση είναι δυνατό να επιλύσουμε γραφικά το παίγνιο. Θα εξετάσουμε το εξής παράδειγμα. Ο παίκτης II κρύβει στο χέρι του ένα νόμισμα είτε του *ενός* είτε των *δύο* ευρώ. Ο παίκτης I μαντεύει τι νόμισμα έχει κρύψει ο II και αν το πετύχει παίρνει το νόμισμα αλλιώς πληρώνει στον I 1.5 ευρώ. Το παιχνίδι παίζεται πολλές φορές. Ποιά είναι η βέλτιστη στρατηγική κάθε παίκτη και ποιά είναι η αξία του παιχνιδιού;

Ο πίνακας του παιγνίου (που δείχνει τα ποσά που πληρώνει ο II στον I) είναι

$$\begin{array}{cc} & \text{παίκτης II} \\ \text{παίκτης I} & \left[\begin{array}{rr} 1 & -1.5 \\ -1.5 & 2 \end{array}\right] \end{array}$$

Είναι σαφές ότι ο πίνακας του παιγνίου αυτού δεν έχει σαγματικό σημείο. Αν ο παίκτης I δηλώσει πρώτος τι πιστεύει ότι είναι το νόμισμα και ο παίκτης II μπορεί να κρύψει το νόμισμα μετά την δήλωση τότε ασφαλώς θα διαλέξει το αντίθετο και θα κερδίσει από τον παίκτη I 1.5 ευρώ. Αν ο παίκτης II δηλώσει πρώτος τι νόμισμα έχει κρύψει τότε ο I θα το πάρει. Ο II το γνωρίζει και γιάυτό θα κρύψει ένα ευρώ.

Ας πούμε λοιπόν ότι ο II χρησιμοποιήσει μεικτές στρατηγικές. Δηλώνει ότι θα κρύψει 1 ευρώ με πιθανότητα y_1 και 2 ευρώ με πιθανότητα $y_2 = 1 - y_1$. (Ακόμη και αν δεν το δηλώσει, αφού το παιχνίδι παίζεται συνέχεια ο I θα το καταλάβει αν συλλέξει στατιστικά στοιχεία.) Αν ο I μαντέψει 1 ευρώ τότε η μέση τιμή της αμοιβής που θα πάρει από τον II είναι $y_1 - 1.5(1 - y_1)$. Αν ο I μαντέψει 2 ευρώ τότε η μέση αμοιβή που θα πάρει από τον II είναι $-1.5y_1 + 2(1 - y_1)$. Ο παίκτης I δεν ξέρει τι νόμισμα έχει κρύψει ο II, μπορεί όμως να εκτιμήσει την πιθανότητα y_1 συνεπώς θα πεί

$$\begin{array}{ll} 1 \text{ ευρώ} & \text{αν } y_1 - 1.5(1 - y_1) > -1.5y_1 + 2(1 - y_1) \\ 2 \text{ ευρώ} & \text{αν } y_1 - 1.5(1 - y_1) \leq -1.5y_1 + 2(1 - y_1) \end{array}$$

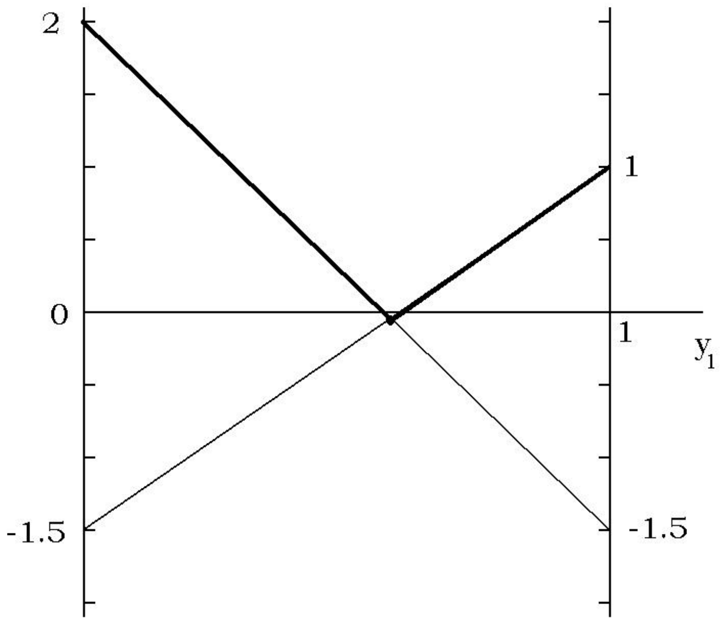
και η μέση αμοιβή που θα πάρει θα είναι

$$\max(y_1 - 1.5(1 - y_1)\ ,\ -1.5y_1 + 2(1 - y_1)) \tag{4.6}$$

Αυτό το γνωρίζει ο II (που ελαχιστοποιεί) συνεπώς θα επιλέξει το y_1 έτσι ώστε να ελαχιστοποιήσει την (4.6). Η μέση τιμή της αμοιβής που πληρώνει είναι

$$\min_{0 \leq y_1 \leq 1} \max(y_1 - 1.5(1 - y_1)\ ,\ -1.5y_1 + 2(1 - y_1)).$$

Επομένως ο II διαλέγει το y_1 έτσι ώστε $y_1 - 1.5(1 - y_1) = -1.5y_1 + 2(1 - y_1)$ ή $y_1 = \frac{7}{12}$ και η αξία του παιγνίου είναι $2 - 3.5y_1 = -1/24$. Αυτό σημαίνει ότι ο II πληρώνει κατά μέσο όρο στον I $-1/24$ ευρώ. Συνεπώς το παιχνίδι είναι επικερδές για τον II αν παίζεται βέλτιστα και του αποφέρει μέσο κέρδος 1/24 ευρώ κάθε φορά.



Σχήμα 4.1: Στο σχήμα βλέπουμε τη μέση τιμή της αμοιβής που πληρώνει ο II στον I όταν ο II χρησιμοποιεί μεικτή στρατηγική με πιθανότητες y_1 και $y_2 = 1 - y_1$. Η παχειά γραμμή είναι η $\max(y_1 - 1.5(1 - y_1)\ ,\ -1.5y_1 + 2(1 - y_1))$. Συνεπώς ο II διαλέγει y_1=το σημείο τομής και πληρώνει κατα μέσο όρο 1/24 ευρώ στον I.

4.4 Επίλυση με την βοήθεια του γραμμικού προγραμματισμού

Έστω ένας πίνακας A $m \times n$. Ο παίκτης I, ο οποίος διαλέγει γραμμές μεγιστοποιεί και έχει m επιλογές, ενώ ο παίκτης II ο οποίος διαλέγει στήλες ελαχιστοποιεί και έχει n επιλογές. Αν ο παικτης I διαλέξει την γραμμή i και ο II την στήλη j τότε ο II πληρώνει στον I το ποσό a_{ij} που δίνει ο πίνακας

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ & & & \vdots & & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ & & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}.$$

. Ας υποθέσουμε ότι ο II επιλέγει την τυχαιοποιημένη στρατηγική $(y_1, y_2, \ldots, y_j, \ldots, y_n)^T$ την οποία και ανακοινώνει στον I. Αν ο I επιλέξει την γραμμή i τότε το μέσο ποσό που θα εισπράξει από τον II είναι $a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{ij}y_j + \cdots + a_{in}y_n$. Δεδομένου ότι θέλει να μεγιστοποιήσει τα έσοδά του θα διαλέξει το i ανάλογα. Συνεπώς η αξία του παιχνιδιού είναι

$$v = \max_{i=1,2,\ldots,m} a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{ij}y_j + \cdots + a_{in}y_n$$

Ο παίκτης II θα διαλέξει τις πιθανότητες y_j έτσι ώστε να ελαχιστοποιεί το v. Συνεπώς η αξία του παιχνιδιού δίνεται από τη λύση του γραμμικού προγράμματος

$$\min v$$

$$\text{υ.π.}$$

$$\begin{aligned} y_1 + \cdots + y_j + \cdots y_n &= 1 \\ a_{11}y_1 + a_{12}y_2 \cdots + a_{1j}y_j \cdots + a_{1n}y_n &\leq v \\ \\ a_{i1}y_1 + a_{i2}y_2 \cdots + a_{ij}y_j \cdots + a_{in}y_n &\leq v \\ \\ a_{m1}y_1 + a_{m2}y_2 \cdots + a_{mj}y_j \cdots + a_{mn}y_n &\leq v \\ \\ y_j \geq 0,\ j = 1, 2, \ldots, n,\ \ v \text{ ελεύθερη}. \end{aligned}$$

Το σύστημα αυτό γράφεται ισοδύναμα ως

$$
\begin{aligned}
&\min v \\
&\text{υ.π.} \\
&\qquad y_1 + \cdots + y_j + \cdots y_n &=& \; 1 \\
&\qquad v - a_{11}y_1 - a_{12}y_2 \cdots - a_{1j}y_j \cdots - a_{1n}y_n &\geq& \; 0 \\
&\qquad v - a_{i1}y_1 - a_{i2}y_2 \cdots - a_{ij}y_j \cdots - a_{in}y_n &\geq& \; 0 \\
&\qquad v - a_{m1}y_1 - a_{m2}y_2 \cdots - a_{mj}y_j \cdots - a_{mn}y_n &\geq& \; 0 \\
&\qquad y_j \geq 0,\ j = 1, 2, \ldots, n, \ \ v \text{ ελεύθερη.}
\end{aligned}
$$

Το δυϊκό γραμμικό πρόγραμμα είναι το

$$
\begin{aligned}
&\max u \\
&\text{υ.π.} \\
&\qquad x_1 + \cdots + x_i + \cdots x_m &=& \; 1 \\
&\qquad u - a_{11}x_1 - a_{12}x_2 - \cdots + a_{1j}x_i - \cdots - a_{1n}x_m &\leq& \; 0 \\
&\qquad u - a_{i1}x_1 - a_{i2}x_2 - \cdots + a_{ij}x_i - \cdots - a_{in}x_m &\leq& \; 0 \\
&\qquad u - a_{n1}x_1 - a_{n2}x_2 - \cdots + a_{nj}x_i - \cdots + a_{nm}x_m &\leq& \; 0 \\
&\qquad x_i \geq 0,\ i = 1, 2, \ldots, m, \ \ u \text{ ελεύθερη.}
\end{aligned}
$$

Η θεωρία της δυϊκότητας εξασφαλίζει ότι η βέλτιστη τιμή του v που προκύπτει από την λύση του πρωτεύοντος προβλήματος ισούται με την βέλτιστη τιμή που προκύπτει από τη λύση του δυϊκού. Αυτή είναι η αξία του παιγνίου.

Κεφάλαιο 5

Εφαρμογές

5.1 Το Θεώρημα Arbitrage

Έχουμε διαθέσιμες n δυνατές επενδύσεις και υπάρχει το ενδεχόμενο m διαφορετικών σεναρίων (states of nature). Αν επιλέξουμε να επενδύσουμε 1 ευρώ στην επένδυση j και συμβεί το σενάριο i τότε η απόδοση είναι r_{ij}. Ένα επενδυτικό χαρτοφυλάκιο είναι ένα διάνυσμα $x \in \mathbb{R}^n$, $x = (x_1, \ldots, x_j, \ldots, x_n)^T$. Το άθροισμα $\sum_{j=1}^{n} r_{ij} x_j$ περιγράφει την απόδοση του χαρτοφυλακίου x όταν το σενάριο που συμβαίνει είναι το i. Έστω επίσης $p = (p_1, \ldots, p_n)^T$ ένα διάνυσμα τιμών των διαφόρων επενδύσεων. Οι τιμές αυτές δεν είναι όλες υποχρεωτικά θετικές. Τότε το κόστος ενός χαρτοφυλακίου είναι $p^T x$ ενώ η απόδοσή του είναι το διάνυσμα Rx στον $\mathbb{R}^m$.

Θεώρημα 25. *Ακριβώς ένα από τα ακόλουθα δύο συμβαίνουν:*

ι) Υπάρχει χαρτοφυλάκιο x για το οποίο $p^T x < 0$ και $Rx \geq 0$. Αυτή είναι η περίπτωση του arbitrage.

ιι) Υπάρχει διάνυσμα πιθανοτήτων $q \in \mathbb{R}^m$, $q \geq 0$, $\sum_{i=1}^{m} q_i = 1$, τέτοιο ώστε $q^T R = cp^T$, όπου $c > 0$ δηλαδή οι τιμές των αποδόσεων είναι ανάλογες προς τις μέσες τιμές ως προς μια κατανομή πιθανοτήτων πάνω στα διάφορα σενάρια.

Απόδειξη: Εφαρμόζουμε το Λήμμα του Farkas στον πίνακα R^T. Η μία εναλλακτική δυνατότητα είναι να υπάρχει λύση στο σύστημα $R^T v = p$ με $v \geq 0$. Ισχύει αναγκαστικά ότι $\sum_{i=1}^{m} v_i > 0$ εκτός από την τετριμμένη περίπτωση που $p = 0$. Συνεπώς, θέτωντας $c^{-1} = \sum_{i=1}^{m} v_i$, και $q = c^{-1} v$ έχουμε $q^T R = cp^T$. Αυτή είναι η περίπτωση (ii) του θεωρήματος.

Αν δεν ισχύει η πρώτη εναλλακτική δυνατότητα τότε υποχρεωτικά θα υπάρχει x τέτοιο ώστε $x^T R^T \geq 0$ και $x^T p < 0$, δηλαδή υπάρχει χαρτοφυλάκιο του οποίου το κόστος είναι αρνητικό (αφού $x^T p < 0$) αλλά η απόδοσεις είναι μη αρνητικές σε κάθε σενάριο (αφού $Rx \geq 0$).