Κεφάλαιο 3

Κυρτά σύνολα

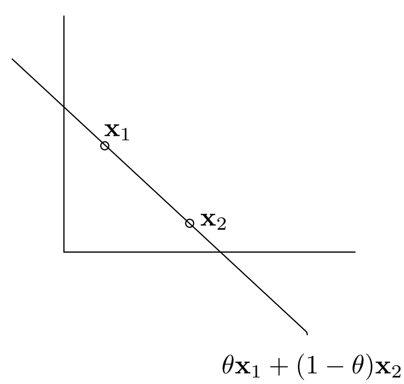
3.1 Αφφινικά σύνολα

Ορισμός 3.1.1. Ένα σύνολο $\mathbb{C} \subseteq \mathbb{R}^n$ καλείται **αφφινικό** (affine) αν τα σημεία των ευθειών που ενώνουν δύο οποιαδήποτε σημεία του $\mathbb{C}$ ανήκουν στο $\mathbb{C}$. Δηλαδή, αν $\mathbf{x}, \mathbf{y} \in \mathbb{C}$ και $\theta \in \mathbb{R}$, τότε $\theta\mathbf{x} + (1-\theta)\mathbf{y} \in \mathbb{C}$.

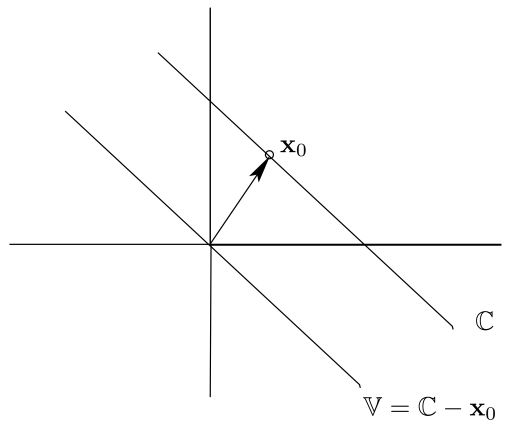
Με λόγια, το σύνολο $\mathbb{C}$ καλείται αφφινικό αν περιέχει κάθε γραμμικό συνδυασμό οποιωνδήποτε δύο σημείων του, υπό την προϋπόθεση οι συντελεστές των γραμμικού συνδυασμού να αθροίζουν στη μονάδα.

Παράδειγμα 3.1.1. Να αποδείξετε ότι το σύνολο λύσεων της εξίσωσης $\mathbf{Ax} = \mathbf{b}$ είναι αφφινικός χώρος.

Η απόδειξη έχει ως εξής. Αν το σύστημα δεν έχει λύσεις, τότε το σύνολο λύσεων είναι το $\emptyset$,



Σχήμα 3.1: Το αφφινικό σύνολο $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2$.



Σχήμα 3.2: Αφφινικό σύνολο $\mathbb{C}$ και υπόχωρος $\mathbb{V} - \mathbf{x}_0$.

το οποίο είναι αφφινικό σύνολο.

Έστω $\mathbf{x}_1$ και $\mathbf{x}_2$ λύσεις της εξίσωσης $\mathbf{Ax} = \mathbf{b}$ και $\theta \in \mathbb{R}$. Τότε

$$\begin{aligned} & \mathbf{Ax}_1 = \mathbf{b}, \quad \mathbf{Ax}_2 = \mathbf{b} \\ \Rightarrow \quad & \mathbf{A}(\theta\,\mathbf{x}_1) = \theta\,\mathbf{b}, \quad \mathbf{A}((1-\theta)\,\mathbf{x}_2) = (1-\theta)\,\mathbf{b} \\ \Rightarrow \quad & \mathbf{A}\left(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right) = \mathbf{b}. \end{aligned} \tag{3.1}$$

Συνεπώς, το σημείο $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2$ είναι λύση της εξίσωσης $\mathbf{Ax} = \mathbf{b}$. Άρα, το σύνολο λύσεων είναι αφφινικό. □

Άσκηση: Έστω $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$. Να βρείτε το σύνολο $\mathbf{aff}\{\mathbf{x}_1, \mathbf{x}_2\}$.

Ορισμός 3.1.2. Αν $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$, $\theta_1, \ldots, \theta_k \in \mathbb{R}$ με

$$\theta_1 + \cdots + \theta_k = 1,$$

τότε κάθε έκφραση της μορφής $\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k$, καλείται **αφφινικός συνδυασμός** (affine combination) των $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Χρησιμοποιώντας επαγωγή και τον ορισμό 3.1.1, μπορεί να αποδειχθεί ότι ένα αφφινικό σύνολο περιέχει κάθε αφφινικό συνδυασμό των στοιχείων του.

Δηλαδή, αν το $\mathbb{C}$ είναι αφφινικό σύνολο, $k \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{C}$, και $\theta_1 + \cdots + \theta_k = 1$, τότε $\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k \in \mathbb{C}$ (να το αποδείξετε).

Θεώρημα 3.1.1. Αν το $\mathbb{C}$ είναι αφφινικό σύνολο και $\mathbf{x}_0 \in \mathbb{C}$, τότε το σύνολο

$$\mathbb{V} = \mathbb{C} - \mathbf{x}_0 = \{\mathbf{x} - \mathbf{x}_0 \,|\, \mathbf{x} \in \mathbb{C}\} \tag{3.2}$$

είναι γραμμικός υπόχωρος.

Απόδειξη. Θα πρέπει να αποδείξουμε ότι αν $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$ και $a_1, a_2 \in \mathbb{R}$, τότε $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 \in \mathbb{V}$.

Έστω, λοιπόν, ότι $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$. Τότε, υπάρχουν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}$ τέτοια ώστε $\mathbf{v}_1 = \mathbf{x}_1 - \mathbf{x}_0$ και $\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{x}_0$.

Έστω $a_1, a_2 \in \mathbb{R}$. Τότε

$$\begin{aligned} a_1\mathbf{v}_1 + a_2\mathbf{v}_2 &= a_1(\mathbf{x}_1 - \mathbf{x}_0) + a_2(\mathbf{x}_2 - \mathbf{x}_0) \\ &= a_1\mathbf{x}_1 + a_2\mathbf{x}_2 - (a_1 + a_2)\mathbf{x}_0 \\ &= a_1\mathbf{x}_1 + a_2\mathbf{x}_2 - (a_1 + a_2)\mathbf{x}_0 + \mathbf{x}_0 - \mathbf{x}_0 \\ &= [a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + (1 - a_1 - a_2)\mathbf{x}_0] - \mathbf{x}_0 \\ &= \mathbf{x}_* - \mathbf{x}_0 \end{aligned} \tag{3.3}$$

όπου $\mathbf{x}_* := a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + (1 - a_1 - a_2)\mathbf{x}_0 \in \mathbb{C}$ (γιατί;).

Άρα, κάθε γραμμικός συνδυασμός των $\mathbf{v}_1$ και $\mathbf{v}_2$ μπορεί να εκφραστεί ως $\mathbf{x}_* - \mathbf{x}_0$, με $\mathbf{x}_* \in \mathbb{C}$, και, συνεπώς, ανήκει στο $\mathbf{V}$. □

Όμοια, μπορεί να αποδειχθεί ότι κάθε αφφινικό σύνολο $\mathbb{C}$ μπορεί να εκφραστεί σαν μετατόπιση ενός γραμμικού υποχώρου $\mathbb{V}$, δηλαδή

$$\mathbb{C} = \mathbb{V} + \mathbf{x}_0 = \{\mathbf{v} + \mathbf{x}_0 \,|\, \mathbf{v} \in \mathbb{V}\}, \tag{3.4}$$

με $\mathbf{x}_0 \in \mathbb{C}$. Το $\mathbf{x}_0$ στους παραπάνω ορισμούς δεν είναι μοναδικό. Για την ακρίβεια, μπορεί να είναι οποιοδήποτε σημείο του $\mathbb{C}$.

Η διάσταση του $\mathbb{C}$ ορίζεται ως η διάσταση του $\mathbb{V}$.

Ορισμός 3.1.3. Το σύνολο των αφφινικών συνδυασμών των στοιχείων ενός συνόλου $\mathbb{C} \subseteq \mathbb{R}^n$ καλείται **αφφινική θήκη** (affine hull) του $\mathbb{C}$ και συμβολίζεται ως $\mathbf{aff}\,\mathbb{C}$.

Δηλαδή,

$$\mathbf{aff}\,\mathbb{C} = \{\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k \,|\, \mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{C}, \theta_1 + \cdots + \theta_k = 1\}. \tag{3.5}$$

Η αφφινική θήκη ενός συνόλου $\mathbb{C}$ είναι το μικρότερο αφφινικό σύνολο το οποίο περιέχει το $\mathbb{C}$, υπό την εξής έννοια: αν $\mathbb{S}$ είναι αφφινικό σύνολο με $\mathbb{C} \subseteq \mathbb{S}$, τότε $\mathbf{aff}\mathbb{C} \subseteq \mathbb{S}$.

3.1.1 Αφφινική διάσταση και σχετικό εσωτερικό

Η **αφφινική διάσταση** (affine dimension) ενός συνόλου $\mathbb{C}$ είναι η διάσταση της αφφινικής θήκης του.

Αν η αφφινική διάσταση ενός συνόλου $\mathbb{C} \subseteq \mathbb{R}^n$ είναι μικρότερη του n, τότε το $\mathbb{C}$ ευρίσκεται μέσα στο αφφινικό σύνολο $\mathbf{aff}\,\mathbb{C} \neq \mathbb{R}^n$.

Ορίζουμε ως **σχετικό εσωτερικό** (relative interior) του συνόλου $\mathbb{C}$ το σύνολο

$$\mathbf{relint}\,\mathbb{C} = \{\mathbf{x} \in \mathbb{C} \,|\, \mathbb{B}(\mathbf{x}, r) \cap \mathbf{aff}\,\mathbb{C} \subseteq \mathbb{C}, \text{ για κάποιο } r > 0\}, \tag{3.6}$$

όπου

$$\mathbb{B}(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^n \,|\, \|\mathbf{y} - \mathbf{x}\| \leq r\}$$

και ως **σχετικό σύνορο** (relative boundary) το $\mathbf{cl}\,\mathbb{C} \setminus \mathbf{relint}\,\mathbb{C}$.

Παράδειγμα 3.1.2. Έστω το τετράγωνο

$$\mathbb{C} = \{\mathbf{x} \in \mathbb{R}^3 \,|\, -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1, x_3 = 0\}.$$

Η αφφινική θήκη του $\mathbb{C}$ είναι το σύνολο $\mathbf{aff}\,\mathbb{C} = \{\mathbf{x} \in \mathbb{R}^3 \,|\, x_3 = 0\}$. Το εσωτερικό του $\mathbb{C}$ είναι κενό (γιατί;) ενώ το σύνορό του είναι το ίδιο το $\mathbb{C}$. Όμως, το σχετικό εσωτερικό του $\mathbb{C}$ είναι το σύνολο

$$\mathbf{relint}\,\mathbb{C} = \{\mathbf{x} \in \mathbb{R}^3 \,|\, -1 < x_1 < 1, -1 < x_2 < 1, x_3 = 0\}.$$

και το σχετικό σύνορό του είναι το

$$\{\mathbf{x} \in \mathbb{R}^3 \,|\, \max\{|x_1|, |x_2|\} = 1, x_3 = 0\}.$$

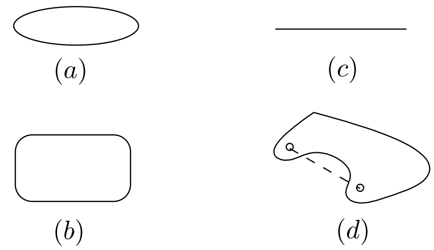
3.2 Κυρτά σύνολα

Ορισμός 3.2.1. Ένα σύνολο $\mathbb{C} \subseteq \mathbb{R}^n$ καλείται **κυρτό** (convex) αν τα σημεία των ευθυγράμμων τμημάτων τα οποία συνδέουν οποιαδήποτε δύο σημεία του $\mathbb{C}$ ανήκουν στο $\mathbb{C}$.

Δηλαδή, αν για κάθε $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}$ και $0 \leq \theta \leq 1$, έχουμε ότι

$$\theta \mathbf{x}_1 + (1 - \theta)\mathbf{x}_2 \in \mathbb{C}. \tag{3.7}$$

Μπορούμε να διαπιστώσουμε εύκολα ότι κάθε αφφινικό σύνολο είναι κυρτό.



Σχήμα 3.3: Παραδείγματα κυρτών συνόλων $(a), (b), (c)$ και μη-κυρτού συνόλου (d).

Ορισμός 3.2.2. Έστω $k \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$, και $\theta_1, \ldots, \theta_k \in \mathbb{R}$, με $\theta_i \geq 0$, για $i = 1, \ldots, k$, και $\theta_1 + \cdots + \theta_k = 1$. Το σημείο $\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k$ καλείται **κυρτός συνδυασμός** (convex combination) των $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Θεώρημα 3.2.1. Ένα σύνολο $\mathbb{C}$ είναι κυρτό αν, και μόνο αν, περιέχει όλους τους κυρτούς συνδυασμούς των στοιχείων του, δηλαδή, αν $k \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{C}$, $\theta_1, \ldots, \theta_k \in \mathbb{R}$, με $\theta_i \geq 0$, για $i = 1, \ldots, k$, και $\theta_1 + \cdots + \theta_k = 1$, τότε $\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k \in \mathbb{C}$.

Απόδειξη. (Αντίστροφο) Η απόδειξη του αντιστρόφου είναι απλή και θα αρχίσουμε με αυτή.

Αφού το $\mathbb{C}$ περιέχει τους κυρτούς συνδυασμούς των στοιχείων του για κάθε $k \in \mathbb{N}$ θα περιέχει και τους κυρτούς συνδυασμούς των στοιχείων του για $k = 2$.

Συνεπώς, είναι κυρτό σύνολο.

(Ευθύ) Η απόδειξη θα γίνει με επαγωγή.

Για $k = 2$, το θεώρημα ισχύει από τον ορισμό του κυρτού συνόλου.

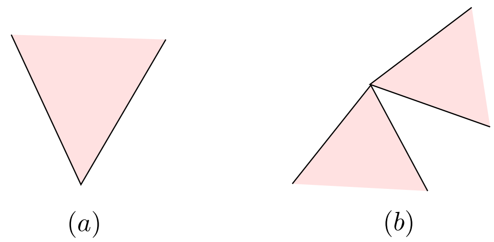
Υποθέτουμε ότι ισχύει για $k = n$.

Έστω ότι $k = n + 1$, $\mathbf{x}_i \in \mathbb{C}$, $\theta_i \in \mathbb{R}$, $\theta_i \geq 0$, για $i = 1, \ldots, n + 1$, με $\sum_{i=1}^{n+1} \theta_i = 1$ και $\mathbf{x} = \sum_{i=1}^{n+1} \theta_i \mathbf{x}_i$. Υποθέτουμε ότι $\theta_{n+1} \neq 0$ και ορίζουμε $\theta := \sum_{i=1}^{n} \theta_i$. Τότε $\theta_{n+1} = 1 - \theta$ και

$$\mathbf{x} = \sum_{i=1}^{n+1} \theta_i \mathbf{x}_i = \sum_{i=1}^{n} \theta_i \mathbf{x}_i + \theta_{n+1} \mathbf{x}_{n+1} = \theta \left(\sum_{i=1}^{n} \frac{\theta_i}{\theta} \mathbf{x}_i \right) + \theta_{n+1} \mathbf{x}_{n+1}.$$

Αν ορίσουμε $\theta'_i := \frac{\theta_i}{\theta}$, για $i = 1, \ldots, n$, παρατηρούμε ότι η έκφραση $\sum_{i=1}^{n} \theta'_i \mathbf{x}_i$ είναι κυρτός συνδυασμός n στοιχείων του $\mathbb{C}$.

Ορίζουμε $\mathbf{x}' = \sum_{i=1}^{n} \theta'_i \mathbf{x}_i$.



Σχήμα 3.4: Κυρτός κώνος (a), μη-κυρτός κώνος (b).

Από την επαγωγική υπόθεση συμπεραίνουμε ότι $\mathbf{x}' \in \mathbb{C}$.

Τέλος, διαπιστώνουμε ότι $\mathbf{x} \in \mathbb{C}$ διότι μπορεί να εκφραστεί σαν κυρτός συνδυασμός δύο σημείων του $\mathbb{C}$, ως εξής

$$\mathbf{x} = \theta \mathbf{x}' + (1-\theta)\mathbf{x}_{n+1}.$$

Συνεπώς, η πρόταση ισχύει για $k = n+1$ και η απόδειξη ολοκληρώθηκε. □

Ορισμός 3.2.3. Η **κυρτή θήκη** (convex hull) ενός συνόλου $\mathbb{C}$ είναι το σύνολο όλων των κυρτών συνδυασμών των σημείων του $\mathbb{C}$ και συμβολίζεται ως $\mathbf{conv}\,\mathbb{C}$. Δηλαδή,

$$\begin{aligned} \mathbf{conv}\,\mathbb{C} = \{\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k \,|\, \mathbf{x}_i \in \mathbb{C}, \theta_i \geq 0, i = 1, \ldots, k, \\ \theta_1 + \cdots + \theta_k = 1\}. \end{aligned} \tag{3.8}$$

Μπορεί να αποδειχθεί ότι το σύνολο $\mathbf{conv}\,\mathbb{C}$ είναι το μικρότερο κυρτό σύνολο το οποίο περιέχει το $\mathbb{C}$ (να το αποδείξετε).

Άσκηση: Έστω $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$. Να βρείτε το σύνολο $\mathbf{conv}\{\mathbf{x}_1, \mathbf{x}_2\}$.

3.3 Κώνοι

Ορισμός 3.3.1. Ένα σύνολο $\mathbb{C} \subseteq \mathbb{R}^n$ καλείται **κώνος** (cone) αν για κάθε $\mathbf{x} \in \mathbb{C}$ και $\theta \geq 0$, έχουμε ότι $\theta\mathbf{x} \in \mathbb{C}$.

Παρατηρήστε ότι αν $\mathbb{C} \neq \emptyset$, τότε $\mathbf{0} \in \mathbb{C}$ (γιατί;).

Ένα σύνολο $\mathbb{C}$ είναι κυρτός κώνος αν είναι, ταυτόχρονα, κυρτό σύνολο και κώνος (δείτε το Σχήμα 3.4).

Άσκηση: Έστω $\mathbb{C}$ κυρτός κώνος. Να αποδείξετε ότι, αν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}$ και $\theta_1, \theta_2 \geq 0$, τότε

$$\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 \in \mathbb{C}. \tag{3.9}$$

Απόδειξη: Έστω ότι $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}$ και $\theta_1, \theta_2 \geq 0$. Αν $\theta_1 = \theta_2 = 0$, τότε $\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 = \mathbf{0} \in \mathbb{C}$. Έστω $\theta_1 \neq 0$ ή $\theta_2 \neq 0$.

Τότε, εξαιτίας της κυρτότητας του $\mathbb{C}$, έχουμε ότι

$$\frac{\theta_1}{\theta_1+\theta_2}\mathbf{x}_1 + \frac{\theta_2}{\theta_1+\theta_2}\mathbf{x}_2 \in \mathbb{C}. \tag{3.10}$$

Επιπλέον, το $\mathbb{C}$ είναι κώνος, συνεπώς,

$$(\theta_1+\theta_2)\left(\frac{\theta_1}{\theta_1+\theta_2}\mathbf{x}_1 + \frac{\theta_2}{\theta_1+\theta_2}\mathbf{x}_2\right) \in \mathbb{C}. \tag{3.11}$$

Δηλαδή, $\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 \in \mathbb{C}$. □

Ορισμός 3.3.2. Έστω $k \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$, και $\theta_1, \ldots, \theta_k \in \mathbb{R}$, με $\theta_i \geq 0$, για $i = 1, \ldots, k$. Τα σημεία της μορφής

$$\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k$$

καλούνται **κωνικοί συνδυασμοί** (conic combinations) των $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Ένα σύνολο $\mathbb{C}$ είναι κυρτός κώνος αν, και μόνο αν, περιέχει όλους τους κωνικούς συνδυασμούς των στοιχείων του.

Ορισμός 3.3.3. Η **κωνική θήκη** (conic hull) ενός συνόλου $\mathbb{C}$ είναι το σύνολο των κωνικών συνδυασμών των σημείων του $\mathbb{C}$, δηλαδή, το σύνολο

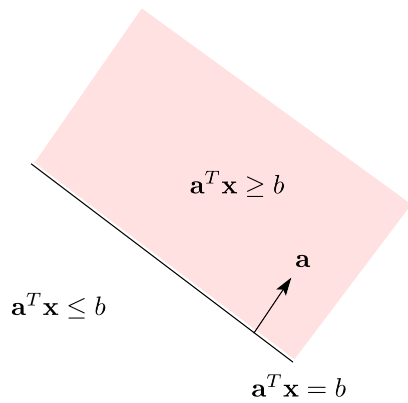
$$\{\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k \,|\, \mathbf{x}_i \in \mathbb{C}, \theta_i \geq 0, i = 1, \ldots, k\}. \tag{3.12}$$

Μπορεί να αποδειχθεί ότι η κωνική θήκη ενός συνόλου $\mathbb{C}$ είναι ο μικρότερος κυρτός κώνος ο οποίος περιέχει το $\mathbb{C}$.

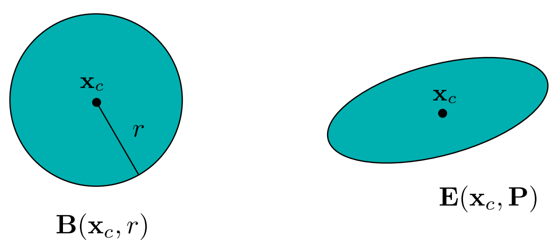
3.4 Παραδείγματα κυρτών συνόλων

Το $\emptyset$, το μονοσύνολο $\{x_0\}$ (singleton), και το $\mathbb{R}^n$ είναι αφφινικά και, συνεπώς, κυρτά σύνολα.

Ευθείες γραμμές είναι αφφινικά και, συνεπώς, κυρτά σύνολα, ενώ ευθύγραμμα τμήματα είναι κυρτά αλλά όχι αφφινικά σύνολα.



Σχήμα 3.5: Υπερεπίπεδο - Ημιχώροι.



Σχήμα 3.6: Ευκλείδεια μπάλα και ελλειψοειδές.

3.5 Υπερεπίπεδα - Ημιχώροι

Ένα **υπερεπίπεδο** (hyperplane) στο $\mathbb{R}^n$ είναι ένα σύνολο της μορφής

$$\mathbb{P} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} - b = 0\}. \tag{3.13}$$

Ένα υπερεπίπεδο διαμερίζει το $\mathbb{R}^n$ σε δύο **ημιχώρους** (half-spaces), έστω

$$\mathbb{P}^+ = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} - b \geq 0\}, \quad \mathbb{P}^- = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} - b \leq 0\}. \tag{3.14}$$

Το σύνολο $\mathbb{P}$ είναι αφφινικό και, συνεπώς, κυρτό.

Τα σύνολα $\mathbb{P}^+$ και $\mathbb{P}^-$ είναι κυρτά αλλά δεν είναι αφφινικά (να το αποδείξετε).

3.5.1 Ευκλείδειες μπάλες

Η Ευκλείδεια μπάλα στο $\mathbb{R}^n$ με κέντρο $\mathbf{x}_c$ και ακτίνα r έχει τη μορφή

$$\mathbb{B}(\mathbf{x}_c, r) = \{\mathbf{x} \in \mathbb{R}^n \,|\, \|\mathbf{x} - \mathbf{x}_c\|_2 \leq r\}. \tag{3.15}$$

Μία άλλη αναπαράσταση της Ευκλείδειας μπάλας έχει ως εξής

$$\mathbb{B}(\mathbf{x}_c, r) = \{\mathbf{x}_c + r\mathbf{x} \,|\, \|\mathbf{x}\|_2 \leq 1\}. \tag{3.16}$$

Οι Ευκλείδειες μπάλες είναι κυρτά σύνολα.

Η απόδειξη έχει ως εξής. Έστω $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{B}(\mathbf{x}_c, r)$. Αυτό σημαίνει ότι $\|\mathbf{x}_i - \mathbf{x}_c\|_2 \leq r$, για $i = 1, 2$.

Θα πρέπει να αποδείξουμε ότι για $0 \leq \theta \leq 1$, έχουμε ότι

$$\mathbf{x} := \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{B}(\mathbf{x}_c, r),$$

δηλαδή, $\|\mathbf{x} - \mathbf{x}_c\|_2 \leq r$.

Όμως,

$$\begin{aligned}
\|\mathbf{x} - \mathbf{x}_c\|_2 &= \|\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 - \mathbf{x}_c\|_2 \\
&= \|\theta(\mathbf{x}_1 - \mathbf{x}_c) + (1-\theta)(\mathbf{x}_2 - \mathbf{x}_c)\|_2 \\
&\leq \|\theta(\mathbf{x}_1 - \mathbf{x}_c)\|_2 + \|(1-\theta)(\mathbf{x}_2 - \mathbf{x}_c)\|_2 \\
&= \theta\|\mathbf{x}_1 - \mathbf{x}_c\|_2 + (1-\theta)\|\mathbf{x}_2 - \mathbf{x}_c\|_2 \\
&\leq \theta\, r + (1-\theta)\, r \\
&= r,
\end{aligned} \tag{3.17}$$

συνεπώς, $\mathbf{x} \in \mathbb{B}(\mathbf{x}_c, r)$.

3.5.2 Ελλειψοειδή

Αν $\mathbf{P} = \mathbf{P}^T \succ \mathbf{O}$ και $\mathbf{x}_c \in \mathbb{R}^n$, τότε το σύνολο

$$\mathbb{E}(\mathbf{x}_c, \mathbf{P}) = \left\{\mathbf{x} \in \mathbb{R}^n \,|\, (\mathbf{x} - \mathbf{x}_c)^T \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \leq 1\right\} \tag{3.18}$$

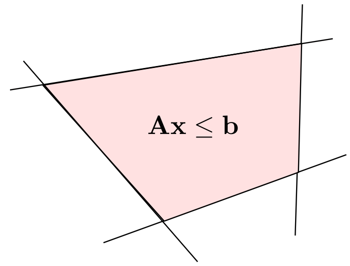
καλείται **ελλειψοειδές** (ellipsoid) με κέντρο $\mathbf{x}_c$.

Μία εναλλακτική περιγραφή του $\mathbb{E}$ έχει ως εξής

$$\mathbb{E}(\mathbf{x}_c, \mathbf{A}) = \{\mathbf{x}_c + \mathbf{A}\mathbf{u} \,|\, \|\mathbf{u}\|_2 \leq 1\}, \tag{3.19}$$

με $\mathbf{A} = \mathbf{P}^{\frac{1}{2}}$.

Τα ελλειψοειδή είναι κυρτά σύνολα. Στη συνέχεια, θα δώσουμε μία απόδειξη για την περίπτωση όπου $\mathbf{x}_c = \mathbf{0}$ (η απόδειξη γενικεύεται για κάθε $\mathbf{x}_c$).



Σχήμα 3.7: Πολύεδρο $\mathbf{Ax} \leq \mathbf{b}$.

Έστω $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{E}(\mathbf{0}, \mathbf{P})$, δηλαδή, $\mathbf{x}_i^T \mathbf{P}^{-1} \mathbf{x}_i \leq 1$, για $i = 1, 2$. Θα πρέπει να αποδείξουμε ότι, αν $0 \leq \theta \leq 1$ και $\mathbf{x} := \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2$, τότε $\mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} \leq 1$. Αρχικά, έχουμε

$$\begin{aligned} \mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} &= \left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right)^T \mathbf{P}^{-1} \left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right) \\ &= \theta^2 \mathbf{x}_1 \mathbf{P}^{-1} \mathbf{x}_1 + (1-\theta)^2 \mathbf{x}_2 \mathbf{P}^{-1} \mathbf{x}_2 + 2\theta(1-\theta) \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2 \\ &\leq \theta^2 + (1-\theta)^2 + 2\theta(1-\theta) \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2. \end{aligned} \tag{3.20}$$

Αν ορίσουμε $\mathbf{y}_i := \mathbf{P}^{-\frac{1}{2}} \mathbf{x}_i$, για $i = 1, 2$, τότε $\|\mathbf{y}_i\|_2^2 = \mathbf{x}_i^T \mathbf{P}^{-1} \mathbf{x}_i \leq 1$, για $i = 1, 2$, και

$$\begin{aligned} \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2 &= \left(\mathbf{P}^{-\frac{1}{2}} \mathbf{x}_1\right)^T \left(\mathbf{P}^{-\frac{1}{2}} \mathbf{x}_2\right) \\ &= \mathbf{y}_1^T \mathbf{y}_2 \\ &\overset{(a)}{\leq} \|\mathbf{y}_1\|_2 \|\mathbf{y}_2\|_2 \\ &\leq 1, \end{aligned} \tag{3.21}$$

όπου στο σημείο (a) χρησιμοποιήσαμε την ανισότητα Cauchy-Schwarz. Εξαιτίας του ότι $0 \leq \theta \leq 1$, έχουμε ότι $\theta(1-\theta) \geq 0$, και

$$\theta(1-\theta) \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2 \leq \theta(1-\theta). \tag{3.22}$$

Συνεπώς,

$$\mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} \leq \theta^2 + (1-\theta)^2 + 2\theta(1-\theta) = 1. \tag{3.23}$$

Άρα, το σημείο $\mathbf{x}$ ανήκει στο ελλειψοειδές $\mathbb{E}(\mathbf{0}, \mathbf{P})$ και το σύνολο $\mathbb{E}(\mathbf{0}, \mathbf{P})$ είναι κυρτό.

3.5.3 Πολύεδρα

Ως **πολύεδρο** (polyhedron) ορίζουμε τη λύση ενός πεπερασμένου πλήθους γραμμικών ανισοτήτων και γραμμικών ισοτήτων (δείτε το Σχήμα 3.7).

Για παράδειγμα, το σύνολο

$$\begin{aligned}\mathbb{D} = \{\mathbf{x} \in \mathbb{R}^n \,|\, &\mathbf{a}_j^T\mathbf{x} - b_j \leq 0, j = 1, \ldots, m, \\ &\mathbf{c}_j^T\mathbf{x} - d_j = 0, j = 1, \ldots, p\}\end{aligned} \tag{3.24}$$

είναι πολύεδρο.

Ένα πολύεδρο είναι η τομή πεπερασμένου πλήθους ημιχώρων και υπερεπιπέδων.

Αφφινικά σύνολα, ευθύγραμμα τμήματα, και ημιχώροι είναι πολύεδρα.

Να αποδείξετε ότι τα πολύεδρα είναι κυρτά σύνολα.

3.6 Πράξεις συνόλων οι οποίες διατηρούν την κυρτότητα

3.6.1 Τομή

Αν $\mathbb{S}_1, \mathbb{S}_2 \subseteq \mathbb{R}^n$ είναι κυρτά σύνολα, τότε η τομή τους $\mathbb{S}_1 \cap \mathbb{S}_2$ είναι κυρτό σύνολο. Η παραπάνω πρόταση γενικεύεται για οποιοδήποτε πλήθος (πεπερασμένο ή άπειρο) κυρτών συνόλων.

Πιο συγκεκριμένα, αν $i \in \mathcal{I}$, όπου $\mathcal{I}$ οποιοδήποτε σύνολο δεικτών, και $\mathbb{S}_i$ κυρτό σύνολο, για κάθε $i \in \mathcal{I}$, τότε το σύνολο

$$\mathbb{S} = \cap_{i\in\mathcal{I}}\, \mathbb{S}_i \tag{3.25}$$

είναι κυρτό. Η απόδειξη έχει ως εξής.

Έστω $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}$. Αυτό σημαίνει ότι $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}_i$, για $i \in \mathcal{I}$. Αν $0 \leq \theta \leq 1$, τότε, εξαιτίας της κυρτότητας των $\mathbb{S}_i$, για $i \in \mathcal{I}$, θα έχουμε ότι $\mathbf{x} := \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{S}_i$, για $i \in \mathcal{I}$.

Συνεπώς, $\mathbf{x} \in \mathbb{S}$, αποδεικνύοντας ότι το $\mathbb{S}$ είναι κυρτό σύνολο.

3.6.2 Καρτεσιανό γινόμενο

Θεώρημα 3.6.1. Αν $\mathbb{S}_1 \subseteq \mathbb{R}^n$ και $\mathbb{S}_2 \subseteq \mathbb{R}^m$ είναι κυρτά σύνολα, τότε το καρτεσιανό γινόμενο $\mathbb{S} = \mathbb{S}_1 \times \mathbb{S}_2 \subseteq \mathbb{R}^{n+m}$ είναι κυρτό.

Απόδειξη. Θα πρέπει να αποδείξουμε ότι αν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}$ και $0 \leq \theta \leq 1$, τότε $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{S}$. Έστω, λοιπόν, ότι $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}$. Τότε, υπάρχουν $\mathbf{x}_{11}, \mathbf{x}_{12} \in \mathbb{S}_1$ και $\mathbf{x}_{21}, \mathbf{x}_{22} \in \mathbb{S}_2$, τέτοια ώστε

$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{21} \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} \mathbf{x}_{12} \\ \mathbf{x}_{22} \end{bmatrix}.$$

Αφού τα σύνολα $\mathbb{S}_1$ και $\mathbb{S}_2$ είναι κυρτά θα πρέπει για, κάθε $0 \leq \theta \leq 1$, $\theta \mathbf{x}_{11} + (1-\theta)\mathbf{x}_{12} \in \mathbb{S}_1$ και $\theta \mathbf{x}_{21} + (1-\theta)\mathbf{x}_{22} \in \mathbb{S}_2$. Συνεπώς,

$$\begin{aligned} & \begin{bmatrix} \theta \mathbf{x}_{11} + (1-\theta)\mathbf{x}_{12} \\ \theta \mathbf{x}_{21} + (1-\theta)\mathbf{x}_{22} \end{bmatrix} \in \mathbb{S}_1 \times \mathbb{S}_2 \\ \Longrightarrow \quad & \theta \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{21} \end{bmatrix} + (1-\theta) \begin{bmatrix} \mathbf{x}_{12} \\ \mathbf{x}_{22} \end{bmatrix} \in \mathbb{S}_1 \times \mathbb{S}_2 \\ \Longrightarrow \quad & \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{S}, \end{aligned} \tag{3.26}$$

ολοκληρώνοντας την απόδειξη. □

3.6.3 Εικόνα αφφινικής συνάρτησης

Μία συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}^m$ καλείται αφφινική αν είναι το άθροισμα μίας γραμμικής συνάρτησης και μίας σταθεράς, δηλαδή, αν είναι της μορφής $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, με $\mathbf{A} \in \mathbb{R}^{m \times n}$ και $\mathbf{b} \in \mathbb{R}^m$.

Θεώρημα 3.6.2. Έστω ότι $\mathbb{S} \subseteq \mathbb{R}^n$ είναι κυρτό σύνολο και $f : \mathbb{R}^n \to \mathbb{R}^m$ αφφινική συνάρτηση. Τότε η εικόνα του $\mathbb{S}$ υπό την f,

$$f(\mathbb{S}) = \{f(\mathbf{x}) \,|\, \mathbf{x} \in \mathbb{S}\}, \tag{3.27}$$

είναι κυρτό σύνολο.

Απόδειξη. Θα πρέπει να αποδείξουμε ότι αν $\mathbf{y}_1, \mathbf{y}_2 \in f(\mathbb{S})$ και $0 \leq \theta \leq 1$, τότε $\mathbf{y} = \theta \mathbf{y}_1 + (1-\theta)\mathbf{y}_2 \in f(\mathbb{S})$.

Έστω λοιπόν ότι $\mathbf{y}_1, \mathbf{y}_2 \in f(\mathbb{S})$. Αυτό σημαίνει ότι υπάρχουν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}$, τέτοια ώστε $\mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \mathbf{b}$, για $i = 1, 2$.

Εξαιτίας της κυρτότητας του $\mathbb{S}$, έχουμε ότι, για $0 \leq \theta \leq 1$, $\mathbf{x} = \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{S}$. Η εικόνα του $\mathbf{x}$ δίνεται από τη σχέση

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{A}\mathbf{x} + \mathbf{b} \\ &= \mathbf{A}\left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right) + \mathbf{b} \\ &= \theta \left(\mathbf{A}\mathbf{x}_1 + \mathbf{b}\right) + (1-\theta)\left(\mathbf{A}\mathbf{x}_2 + \mathbf{b}\right) \\ &= \theta \mathbf{y}_1 + (1-\theta)\mathbf{y}_2 \\ &= \mathbf{y}. \end{aligned} \tag{3.28}$$

Συνεπώς, $\mathbf{y} = f(\mathbf{x}) \in f(\mathbb{S})$ και η απόδειξη ολοκληρώθηκε. □

Πόρισμα 3.6.1. Αν $f : \mathbb{R}^k \to \mathbb{R}^n$ είναι αφφινική συνάρτηση και $\mathbb{S} \subseteq \mathbb{R}^n$ κυρτό σύνολο, τότε η αντίστροφη εικόνα του $\mathbb{S}$ υπό την f, δηλαδή, το σύνολο

$$f^{-1}(\mathbb{S}) = \{\mathbf{x} \,|\, f(\mathbf{x}) \in \mathbb{S}\}, \tag{3.29}$$

είναι κυρτό σύνολο.

Απόδειξη. Η απόδειξη βασίζεται στο γεγονός ότι η αντίστροφη αφφινικής συνάρτησης είναι αφφινική συνάρτηση. Για παράδειγμα, αν $m > n$ και οι στήλες του $\mathbf{A}$ είναι γραμμικά ανεξάρτητες, τότε αν $\mathbf{y} = f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, η αντίστροφη συνάρτηση δίνεται από τη σχέση

$$\begin{aligned}
\mathbf{x} = f^{-1}(\mathbf{y}) &= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T(\mathbf{y} - \mathbf{b}) \\
&= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{y} - \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b} \\
&= \mathbf{A}'\mathbf{y} + \mathbf{b}',
\end{aligned} \tag{3.30}$$

με $\mathbf{A}' := \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ και $\mathbf{b}' := -\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$, η οποία είναι αφφινική συνάρτηση. □

Στη συνέχεια, θα παρουσιάσουμε μερικές σημαντικές εφαρμογές του Θεωρήματος 3.6.2.

3.6.4 Κλιμάκωση συνόλου

Πόρισμα 3.6.2. Αν $\mathbb{S} \subseteq \mathbb{R}^n$ είναι κυρτό σύνολο και $a \in \mathbb{R}$, τότε το σύνολο $a\,\mathbb{S} = \{a\mathbf{x} \,|\, \mathbf{x} \in \mathbb{S}\}$ είναι κυρτό.

Απόδειξη. Η πρόταση αυτή είναι απόρροια του Θεωρήματος 3.6.1, για αφφινική (στην πραγματικότητα, γραμμική) συνάρτηση $f : \mathbb{S} \to \mathbb{R}^n$ με $f(\mathbf{x}) = a\,\mathbf{x}$. □

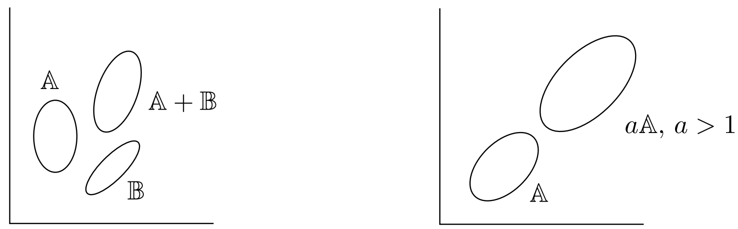
3.6.5 Μετατόπιση συνόλου

Πόρισμα 3.6.3. Αν $\mathbb{S} \subseteq \mathbb{R}^n$ είναι κυρτό σύνολο και $\mathbf{a} \in \mathbb{R}^n$, τότε το σύνολο $\mathbb{S} + \mathbf{a} = \{\mathbf{x} + \mathbf{a} \,|\, \mathbf{x} \in \mathbb{S}\}$ είναι κυρτό.

Απόδειξη. Η πρόταση αυτή είναι απόρροια του Θεωρήματος 3.6.1, για αφφινική συνάρτηση $f : \mathbb{S} \to \mathbb{R}^n$ με $f(\mathbf{x}) = \mathbf{x} + \mathbf{a}$. □

3.6.6 Προβολή συνόλου

Πόρισμα 3.6.4. Η προβολή ενός κυρτού συνόλου πάνω σε κάποιες από τις συντεταγμένες τους οδηγεί σε κυρτό σύνολο. Πιο συγκεκριμένα, αν $\mathbb{S} \subseteq \mathbb{R}^{n+m}$ είναι κυρτό σύνολο, τότε το



Σχήμα 3.8: Άθροισμα κυρτών συνόλων και κλιμάκωση κυρτού συνόλου.

σύνολο

$$\mathbb{T} = \{\mathbf{x}_1 \in \mathbb{R}^n \,|\, (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{S}\}. \tag{3.31}$$

είναι κυρτό.

Απόδειξη. Η απόδειξη αφήνεται στον αναγνώστη. □

3.6.7 Άθροισμα συνόλων

Αν $\mathbb{S}_1, \mathbb{S}_2 \subseteq \mathbb{R}^n$, τότε το άθροισμά τους ορίζεται ως

$$\mathbb{S}_1 + \mathbb{S}_2 = \{\mathbf{x}_1 + \mathbf{x}_2 \,|\, \mathbf{x}_1 \in \mathbb{S}_1, \mathbf{x}_2 \in \mathbb{S}_2\}. \tag{3.32}$$

Πόρισμα 3.6.5. Αν τα $\mathbb{S}_1$ και $\mathbb{S}_2$ είναι κυρτά, τότε το $\mathbb{S}_1 + \mathbb{S}_2$ είναι κυρτό.

Απόδειξη. Ένας τρόπος για να αποδείξουμε αυτή την πρόταση είναι να παρατηρήσουμε τα εξής:

1. το $\mathbb{S}_1 \times \mathbb{S}_2$ είναι κυρτό (καρτεσιανό γινόμενο κυρτών συνόλων),

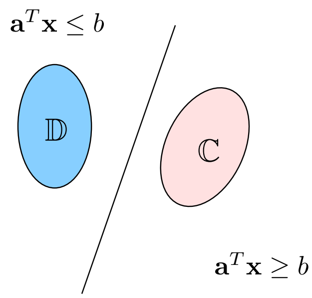
2. η εικόνα του κυρτού συνόλου $\mathbb{S}_1 \times \mathbb{S}_2$ υπό τον αφφινικό μετασχηματισμό $f : \mathbb{S}_1 \times \mathbb{S}_2 \to \mathbb{R}^n$ με τύπο

$$f(\mathbf{x_1}, \mathbf{x}_2) = [\, \mathbf{I}_n \; \mathbf{I}_n \,] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{x}_1 + \mathbf{x}_2. \tag{3.33}$$

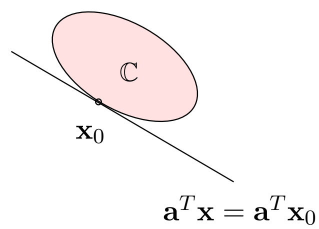
είναι το άθροισμα $\mathbb{S}_1 + \mathbb{S}_2$.

3.7 Υπερεπίπεδα διαχωρισμού

Θεώρημα 3.7.1. Έστω $\mathbb{C}, \mathbb{D} \subseteq \mathbb{R}^n$ δύο κυρτά σύνολα ξένα μεταξύ τους. Τότε, υπάρχουν μη μηδενικό διάνυσμα $\mathbf{a} \in \mathbb{R}^n$, και $b \in \mathbb{R}$, τέτοια ώστε $\mathbf{a}^T\mathbf{x} \geq b$ για κάθε $\mathbf{x} \in \mathbb{C}$ και $\mathbf{a}^T\mathbf{x} \leq b$ για κάθε $\mathbf{x} \in \mathbb{D}$ (δείτε το Σχήμα 3.9).



Σχήμα 3.9: Υπερεπίπεδο διαχωρισμού κυρτών συνόλων.



Σχήμα 3.10: Υπερεπίπεδο στήριξης κυρτού συνόλου.

Το υπερεπίπεδο $\{\mathbf{x} \,|\, \mathbf{a}^T\mathbf{x} = b\}$ καλείται **υπερεπίπεδο διαχωρισμού** (separating hyperplane) των $\mathbb{C}$ και $\mathbb{D}$.

Απόδειξη. Δείτε στις σελίδες 46–48 του βιβλίου των Boyd και Vandenberghe. Να γραφεί η απόδειξη. □

3.8 Υπερεπίπεδα στήριξης

Θεώρημα 3.8.1. Έστω μη-κενό κυρτό σύνολο $\mathbb{C} \subseteq \mathbb{R}^n$ και $\mathbf{x}_0 \in \mathbf{bd}\,\mathbb{C}$. Τότε, υπάρχει μη-μηδενικό διάνυσμα $\mathbf{a} \in \mathbb{R}^n$ τέτοιο ώστε $\mathbf{a}^T\mathbf{x} \leq \mathbf{a}^T\mathbf{x}_0$, για κάθε $\mathbf{x} \in \mathbb{C}$.

Απόδειξη. Να γραφεί η απόδειξη. □

Το Θεώρημα 3.8.1 ουσιαστικά λέει ότι το σημείο $\mathbf{x}_0$ και το σύνολο $\mathbb{C}$ διαχωρίζονται από το υπερεπίπεδο $\mathbb{P} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} = \mathbf{a}^T\mathbf{x}_0\}$.