Εθνικο Μετσοβιο Πολυτεχνειο
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Τομέας Μαθηματικών

Κυρτότητα σε χώρους Banach

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Ιωάννη Γεωργαρά

Επιβλέπων: Δημήτριος Κραββαρίτης
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012

Εθνικο Μετσοβιο Πολυτεχνειο
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Τομέας Μαθηματικών

Κυρτότητα σε χώρους Banach

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Ιωάννη Γεωργαρά

Επιβλέπων: Δημήτριος Κραββαρίτης
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012

.....

Ιωάννης Γεωργαράς
Διπλωματούχος Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών Ε.Μ.Π.

Περίληψη

Ο σκοπός αυτής της διπλωματικής εργασίας ήταν να μελετήσει τα κυρτά σύνολα, τις κυρτές συναρτήσεις και τις ιδιότητές τους. Η κυρτότητα είναι ένα πολύ σημαντικό χαρακτηριστικό, του οποίου η ύπαρξη μας εξασφαλίζει πολλές χρήσιμες ιδιότητες, γι'αυτό και βρίσκει εφαρμογή σε πολλά πεδία των μαθηματικών και άλλων επιστημών. Δόθηκαν οι αποδείξεις σε σημαντικά θεωρήματα όπως είναι το θεώρημα Καραθεοδωρή, το Mazur, το Krein-Milman και η προβολή σε κλειστό κυρτό σύνολο. Επίσης, διερευνήθηκε η συνέχεια και η παραγωγισιμότητα των κυρτών συναρτήσεων.

Abstract

The purpose of this thesis was to study convex sets, convex functions and their properties. Convexity is a very important feature, the existence of which brings many useful properties and that's why it has applications in many fields of mathematics and other sciences. Proofs of important theorems were given like Caratheodory theorem, Mazur theorem, Krein-Milman theorem and the projection on a closed convex set. Moreover, convex functions were investigated from the aspect of continuity and differentiability.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον καθηγητή μου, κ. Δ. Κραββαρίτη, για την προθυμία με την οποία δέχτηκε να αναλάβει την επίβλεψη της διπλωματικής μου και τη βοήθειά του καθ'όλη τη διάρκεια της συνεργασίας μας.

Ιωάννης Γεωργαράς

Περιεχόμενα

Κεφάλαιο 1

Κυρτά Σύνολα

Σε αυτό το κεφάλαιο θα παρουσιάσουμε κάποια βασικά χαρακτηριστικά των κυρτών συνόλων. Όπου δεν διευκρινίζεται, θεωρούμε ότι είμαστε στον $\mathbb{R}^n$ με την ευκλείδια νόρμα.

1.1 Ορισμός και Βασικές Ιδιότητες

Ξεκινάμε με τον ορισμό του κυρτού συνόλου:

Ορισμός 1.1.1. *Ένα σύνολο $C \subseteq \mathbb{R}^n$ λέγεται κυρτό αν $\forall x, x' \in C$ και $\forall \lambda \in [0,1]$ έχουμε $\lambda x + (1-\lambda)x' \in C$.*

Η γεωμετρική ερμηνεία του ορισμού είναι ότι το ευθύγραμμο τμήμα που ενώνει οποιαδήποτε δυο σημεία x, x' ενός κυρτού συνόλου (συμβολιζόμενο ως $[x, x'] := \{\lambda x + (1-\lambda)x' : 0 \leqslant \lambda \leqslant 1\}$) βρίσκεται εξ ολοκλήρου στο σύνολο.

Κάποια παραδείγματα κυρτών συνόλων είναι:

1. Τα κυρτά σύνολα σε μια διάσταση είναι ακριβώς τα διαστήματα.

2. Ένα σύνολο $A \subseteq \mathbb{R}^n$ είναι **αφινικό** αν $\lambda A + \mu A \subseteq A \quad \forall \lambda, \mu \in \mathbb{R}, \lambda + \mu = 1$. Ένας αφινικός χώρος στο $\mathbb{R}^n$ είναι ένα **υπερεπίπεδο** του $\mathbb{R}^n$ αν έχει διάσταση $n-1$. Κάθε υπερεπίπεδο και γενικότερα κάθε αφινικός χώρος είναι προφανώς κυρτά σύνολα. Έστω υπερεπίπεδο $H_{s,r} = \{x \in \mathbb{R}^n : \langle s, x\rangle = r\}$. Αυτό ορίζει τον **κλειστό ημίχωρο** $\{x \in \mathbb{R}^n : \langle s, x\rangle \leqslant r\}$ και τον **ανοιχτό ημίχωρο** $\{x \in \mathbb{R}^n : \langle s, x\rangle < r\}$ οι οποίοι επίσης είναι κυρτά σύνολα.

3. Ένας **κώνος** K είναι ένα σύνολο τέτοιο ώστε η ημιευθεία $\{\lambda x : \lambda > 0\}$ περιέχεται στο K, όταν $x \in K$. Ένας **κυρτός κώνος** είναι ένας κώνος που είναι κυρτός. Ένα απλό παράδειγμα κυρτού κώνου είναι το μη αρνητικό ορθομόριο (orthant) του $\mathbb{R}^n$

$$\Omega_+ := \{x = (\xi^1, \ldots, \xi^n) : \xi^i \geqslant 0 \quad \text{για} \quad i = 1, \ldots, n\}$$

Διαδικασίες που διατηρούν την κυρτότητα σε σύνολα

Πρόταση 1.1.2. *Έστω $\{C_j\}_{j\in J}$ μια τυχαία οικογένεια κυρτών συνόλων. Τότε η τομή $C := \cap\{C_j : j \in J\}$ είναι κυρτό σύνολο.*

Απόδειξη. Έστω $x, x' \in C$. Τότε για κάθε $j \in J$ ισχύει $x, x' \in C_j$ κι επειδή το C_j είναι κυρτό, $[x, x'] \subseteq C_j$ για κάθε $j \in J$, άρα $[x, x'] \in C$. □

Παρατήρηση: Σε αντίθεση με την τομή, η ένωση κυρτών συνόλων συνήθως δεν είναι κυρτό σύνολο.

Πρόταση 1.1.3. *Αν $C, D \subseteq \mathbb{R}^n$ είναι κυρτά σύνολα και $\mu \in \mathbb{R}$, τότε $C + D$ και μC είναι κι αυτά κυρτά.*

Απόδειξη. Έστω $\lambda, \xi \geqslant 0$ τέτοια ώστε $\lambda + \xi = 1$. Λόγω της κυρτότητας των C, D έχουμε $\lambda C + \xi C \subseteq C$ και $\lambda D + \xi D \subseteq D$. Συνεπώς $\lambda(C+D)+\xi(C+D) = (\lambda+\xi)C+(\lambda+\xi)D \subseteq C+D$. Άρα $C + D$ κυρτό.
Όμοια $\lambda(\mu C) + \xi(\mu C) = \mu(\lambda C + \xi C) \subseteq \mu C$. Άρα μC κυρτό. □

Γενικότερα, για μ_1, μ_2 πραγματικούς αριθμούς και C, D κυρτά το σύνολο $\mu_1 C + \mu_2 D$ είναι κυρτό.

Πρόταση 1.1.4. *Για $i = 1, \ldots, k$ έστω $C_i \subseteq \mathbb{R}^{n_i}$ κυρτά σύνολα. Τότε το $C_1 \times \cdots \times C_k$ είναι κυρτό σύνολο.*

Το αντίστροφο ισχύει επίσης· Το $C_1 \times \cdots \times C_k$ είναι κυρτό αν και μόνο αν κάθε C_i είναι κυρτό, όπως φαίνεται και από την επόμενη ιδιότητα. Θυμίζουμε την έννοια της αφινικής απεικόνισης: Μια απεικόνιση $A\colon \mathbb{R}^n \to \mathbb{R}^m$ λέγεται **αφινική** όταν

$$A(\lambda x + (1-\lambda)x') = \lambda A(x) + (1-\lambda)A(x')$$

για κάθε x και x' στο $\mathbb{R}^n$ και κάθε $\lambda \in \mathbb{R}$. Αυτό σημαίνει ότι η $x \mapsto A(x) - A(0)$ είναι γραμμική, οπότε μια αφινική απεικόνιση χαρακτηρίζεται από μια γραμμική απεικόνιση A_0 και ένα σημείο $y_0 := A(0) \in \mathbb{R}^m$:

$$A(x) = A_0 x + y_0 \quad \text{για κάθε} \quad x \in \mathbb{R}^m.$$

Πρόταση 1.1.5. *Έστω η αφινική απεικόνιση $A\colon \mathbb{R}^n \to \mathbb{R}^m$ και C ένα κυρτό σύνολο του $\mathbb{R}^n$. Η εικόνα $A(C)$ του C μέσω της A είναι κυρτό σύνολο στον $\mathbb{R}^m$.*

Αν D είναι κυρτό σύνολο στον $\mathbb{R}^m$, τότε η αντίστροφη εικόνα $A^{-1}(D)$ είναι κυρτό σύνολο στον $\mathbb{R}^n$.

Απόδειξη. Για $x, x' \in \mathbb{R}^n$ η εικόνα του τμήματος $[x, x']$ είναι το τμήμα $[A(x), A(x')] \subset \mathbb{R}^m$, το οποίο αποδεικνύει τον πρώτο ισχυρισμό, αλλά και το δεύτερο: Πράγματι, αν x, x' είναι τέτοια ώστε τα $A(x)$ και $A(x')$ είναι και τα δυο στο κυρτό σύνολο D, τότε κάθε σημείο του τμήματος $[x, x']$ έχει την εικόνα του στο $[A(x), A(x')] \subset D$. □

Παρακάτω ορίζουμε το εσωτερικό και την κλειστότητα ενός συνόλου για να δώσουμε στη συνέχεια κάποιες χρήσιμες τοπολογικές ιδιότητες των κυρτών συνόλων.

Ορισμός 1.1.6. *Έστω $A \subseteq \mathbb{R}^n$. Ονομάζουμε εσωτερικό του A το σύνολο*

$$intA := \{x \in \mathbb{R}^n : \exists \rho > 0 \text{ ώστε } B(x,\rho) \subseteq A\}^1$$

Τα σημεία που αποτελούν το $intA$ ονομάζονται **εσωτερικά σημεία** του A.

Ορισμός 1.1.7. *Έστω $A \subseteq \mathbb{R}^n$. Ονομάζουμε κλειστότητα του A το σύνολο*

$$\overline{A} := \{x \in \mathbb{R}^n : B(x,\varepsilon) \cap A \neq \emptyset, \forall \varepsilon > 0\}$$

ή ισοδύναμα

$$\overline{A} := \{x \in \mathbb{R}^n : \exists (x_n)_n \text{ με } x_n \in A \, \forall n \in \mathbb{N} \text{ ώστε } x_n \to x\}$$

Τα σημεία που αποτελούν το $\overline{A}$ ονομάζονται **οριακά σημεία** του A.

Πρόταση 1.1.8. *Αν C κυρτό, τότε και το $intC$ είναι κυρτό.*

Απόδειξη. Έστω $x, y \in intC$, $\lambda \in [0,1]$ και $z = \lambda x + (1-\lambda)y$. Αφού $x, y \in intC$ υπάρχουν $\varepsilon_1, \varepsilon_2 > 0$ με $B(x,\varepsilon_1) \subseteq C$ και $B(y,\varepsilon_2) \subseteq C$. Θέτουμε $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ κι έτσι $B(x,\varepsilon) \subseteq C$ και $B(y,\varepsilon) \subseteq C$.
Ισχυρισμός: $B(z,\varepsilon) \subseteq C$ και άρα $z \in intC$.
Έστω $w \in B(z,\varepsilon)$. Τότε $\|w - z\| < \varepsilon$. Θέτουμε $u = w - z$, $x_1 = x + u$, $y_1 = y + u$. Τότε $\|u\| < \varepsilon$ και $\|x_1 - x\| = \|x + u - x\| = \|u\| < \varepsilon$ άρα $x_1 \in B(x,\varepsilon)$, δηλαδή $x_1 \in C$. Ομοίως $y_1 \in C$.
Οπότε, $w = z+u = \lambda x+(1-\lambda)y+\lambda u+(1-\lambda)u = \lambda(x+u)+(1-\lambda)(y+u) = \lambda x_1+(1-\lambda)y_1$. Δηλαδή $w \in C$. □

Πρόταση 1.1.9. *Αν C κυρτό, τότε και το $\overline{C}$ είναι κυρτό.*

Απόδειξη. Έστω $x, y \in \overline{C}$, $\lambda \in [0,1]$ και $z = \lambda x+(1-\lambda)y$. Αφού $x \in \overline{C}$, υπάρχει $(x_n)_n \subseteq C$ ώστε $x_n \to x$ και ομοίως αφού $y \in \overline{C}$, υπάρχει $(y_n)_n \subseteq C$ ώστε $y_n \to y$. Επειδή το C είναι κυρτό, για το δεδομένο λ έχουμε $z_n = \lambda x_n+(1-\lambda)y_n \in C$. Επιπλέον $z_n \to \lambda x+(1-\lambda)y = z$ άρα $z \in \overline{C}$. □

1.2 Κυρτοί Συνδυασμοί και Κυρτές Θήκες

Γνωρίζουμε από τη γραμμική άλγεβρα ότι **γραμμικός συνδυασμός** των στοιχείων του $\mathbb{R}^n$ ονομάζεται το στοιχείο $\sum_{i=1}^{k} \lambda_i x_i$, όπου $\lambda_i \in \mathbb{R}$. Θα ορίσουμε τώρα μια ειδική περίπτωση γραμμικού συνδυασμού, τον **κυρτό συνδυασμό**.

[1] Συμβολίζουμε με $B(x,\rho)$ το σύνολο των σημείων που απέχουν από το x απόσταση μικρότερη του ρ.

Ορισμός 1.2.1. *Ονομάζουμε κυρτό συνδυασμό των στοιχείων $x_1, \ldots, x_k$ στον $\mathbb{R}^n$ ένα στοιχείο της μορφής*

$$\sum_{i=1}^{k} \lambda_i x_i, \quad \text{όπου} \quad \sum_{i=1}^{k} \lambda_i = 1 \text{ και } \lambda_i \geqslant 0 \text{ για } i = 1, \ldots, k.$$

Ορισμός 1.2.2. *Έστω $A \subseteq \mathbb{R}^n$. Ονομάζουμε κυρτή θήκη του A την τομή όλων των κυρτών υποσυνόλων του $\mathbb{R}^n$ που περιέχουν το A,*

$$convA := \cap\{C \subseteq \mathbb{R}^n : C \text{ κυρτό και } A \subseteq C\}.$$

Η κυρτή θήκη μπορεί να περιγραφεί και ως το σύνολο όλων των κυρτών συνδυασμών. Θα το αποδείξουμε παρακάτω με τη βοήθεια της επόμενης πρότασης:

Πρόταση 1.2.3. *Αν το $C \subseteq \mathbb{R}^n$ είναι κυρτό, τότε περιλαμβάνει κάθε κυρτό συνδυασμό στοιχείων του, δηλαδή για κάθε $k \in \mathbb{N}$, για όλα τα $x_1, \ldots, x_k \in C$ και για όλα τα $\lambda_1, \ldots, \lambda_k \in [0,1]$ με $\sum_{i=1}^{k} \lambda_i = 1$ ισχύει $\sum_{i=1}^{k} \lambda_i x_i \in C$.*

Απόδειξη. Θα το δείξουμε με επαγωγή. Για $k = 1$ έχουμε $\lambda_1 = 1$ και άρα $\lambda_1 x_1 = x_1 \in C$. Έστω ότι η πρόταση ισχύει για κάποιο $k \in \mathbb{N}$. Θα δείξουμε ότι ισχύει και για το $k+1$. Έστω $x_1, \ldots, x_{k+1} \in C$, $\lambda_1, \ldots, \lambda_{k+1} \in [0,1]$ με $\sum_{i=1}^{k+1} \lambda_i = 1$ και θέτουμε $z = \lambda_1 x_1 + \ldots + \lambda_{k+1} x_{k+1}$.
Για $\lambda_{k+1} = 1$ έχουμε $\lambda_1 = \ldots = \lambda_k = 0$. Άρα $z = \lambda_{k+1} x_{k+1} = x_{k+1} \in C$.
Για $0 \leqslant \lambda_{k+1} < 1$ έχουμε $\lambda_1 + \ldots + \lambda_k = 1 - \lambda_{k+1} \in (0,1]$. Άρα το z γράφεται $z = (1-\lambda_{k+1})\left(\frac{\lambda_1}{1-\lambda_{k+1}} x_1 + \ldots + \frac{\lambda_k}{1-\lambda_{k+1}} x_k\right) + \lambda_{k+1} x_{k+1}$. Θέτουμε $z_1 = \frac{\lambda_1}{1-\lambda_{k+1}} x_1 + \ldots + \frac{\lambda_k}{1-\lambda_{k+1}} x_k$. Το z_1 είναι κυρτός συνδυασμός k στοιχείων του C, άρα από επαγωγική υπόθεση $z_1 \in C$. Επομένως το z γράφεται $z = (1-\lambda_{k+1}) z_1 + \lambda_{k+1} x_{k+1}$, δηλαδή ανήκει στο ευθύγραμμο τμήμα $[z_1, x_{k+1}]$ κι αφού το C είναι κυρτό, έχουμε ότι $z \in C$. □

Πρόταση 1.2.4. *Έστω $A \subseteq \mathbb{R}^n$. Τότε η κυρτή θήκη του A ταυτίζεται με το σύνολο των κυρτών συνδυασμών των στοιχείων του A.*

Απόδειξη. Θέτουμε $B = \{\sum_{i=1}^{k} \lambda_i x_i : k \in \mathbb{N}, x_1, \ldots, x_k \in A, \sum_{i=1}^{k} \lambda_i = 1, \lambda_i \in [0,1]\}$. Θέλουμε να δείξουμε ότι $B = convA$.
•$B \subseteq convA$

Πράγματι, το σύνολο $convA$ είναι εξ ορισμού κυρτό που περιέχει το A. Από την πρόταση 1.2.3 έχουμε ότι το $convA$ περιέχει όλους τους κυρτούς συνδυασμούς των στοιχείων του. Επειδή $A \subseteq convA$ έχουμε ότι το $convA$ θα περιέχει και όλους τους κυρτούς συνδυασμούς των στοιχείων του A.

•$convA \subseteq B$

Παρατηρούμε ότι $A \subseteq B$. Επειδή $convA = \cap\{C \subseteq \mathbb{R}^n : C$ κυρτό και $A \subseteq C\}$, αρκεί να δειχθεί ότι το B είναι κυρτό.

Έστω $z_1, z_2 \in B$ και $\lambda \in [0,1]$. Έστω $z = \lambda z_1 + (1-\lambda)z_2$. Θα δείξουμε ότι $z \in B$. Τα z_1, z_2 ως στοιχεία του B γράφονται $z_1 = \sum_{i=1}^{k} \lambda_i x_i$ με $x_i \in A$, $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i \in [0,1]$ και $z_2 = \sum_{i=1}^{m} \mu_i y_i$ με $y_i \in A$, $\sum_{i=1}^{m} \mu_i = 1$, $\mu_i \in [0,1]$. Και το z γίνεται:

$$z = \lambda z_1 + (1-\lambda)z_2 = \lambda \sum_{i=1}^{k} \lambda_i x_i + (1-\lambda) \sum_{i=1}^{m} \mu_i y_i = \sum_{i=1}^{k} \lambda\lambda_i x_i + \sum_{i=1}^{m} (1-\lambda)\mu_i y_i.$$

Αφού $\sum_{i=1}^{k} \lambda\lambda_i + \sum_{i=1}^{m} (1-\lambda)\mu_i = \lambda \sum_{i=1}^{k} \lambda_i x_i + (1-\lambda) \sum_{i=1}^{m} \mu_i = \lambda + (1-\lambda) = 1$ και $\lambda\lambda_i, (1-\lambda)\mu_i \geqslant 0$ σημαίνει ότι $z \in B$, δηλαδή B κυρτό. □

Έστω ένα σύνολο $A \subseteq \mathbb{R}^n$. Για κάθε θετικό ακέραιο k ονομάζουμε A_k το σύνολο όλων των κυρτών συνδυασμών k στοιχείων του A. Έτσι έχουμε $A = A_1 \subset A_2 \subset \cdots \subset A_k \subset \cdots$. Τα A_k δεν είναι όλα κυρτά, όμως η ένωσή τους είναι κυρτή και συμπίπτει με την κυρτή θήκη του A. Σύμφωνα με το θέωρημα Καραθεοδωρή το k δε χρειάζεται να φτάσει το $+\infty$, αλλά αρκεί να πάρει την τιμή $n+1$, δηλαδή $A_{n+1} = convA$. Παρακάτω θα αποδείξουμε το θεώρημα, αφού πρώτα δώσουμε μια χρήσιμη για την απόδειξή του πρόταση.

Πρόταση 1.2.5. *Έστω $k \geqslant n+2$ και $x_1, \ldots, x_k \in \mathbb{R}^n$ διαφορετικά ανά δύο. Τότε υπάρχουν $\mu_1, \ldots, \mu_k \in \mathbb{R}$ όχι όλα μηδέν, τέτοια ώστε $\sum_{i=1}^{k} \mu_i = 0$ και $\sum_{i=1}^{k} \mu_i x_i = 0$.*

Απόδειξη. Για κάθε $i = 1, \ldots, k-1$ θέτουμε $y_i = x_i - x_k \neq 0$. Επειδή $y_1, \ldots, y_{k-1} \in \mathbb{R}^n$ και $k-1 \geqslant n+1$ έχουμε ότι τα $y_1, \ldots, y_{k-1}$ είναι γραμμικώς εξαρτημένα. Άρα υπάρχουν $\lambda_1, \ldots, \lambda_{k-1} \in \mathbb{R}$ όχι όλα μηδέν ώστε

$$\begin{aligned} \lambda_1 y_1 + \ldots + \lambda_{k-1} y_{k-1} = 0 \qquad & \Leftrightarrow \\ \lambda_1 (x_1 - x_k) + \ldots + \lambda_{k-1}(x_{k-1} - x_k) = 0 \qquad & \Leftrightarrow \\ \lambda_1 x_1 + \ldots + \lambda_{k-1} x_{k-1} - (\lambda_1 + \ldots + \lambda k - 1) x_k = 0 & \end{aligned}$$

Θέτουμε $\mu_1 = \lambda_1, \ldots, \mu_{k-1} = \lambda_{k-1}$, $\mu_k = -(\lambda_1 + \ldots + \lambda_{k-1})$. Τότε $\sum_{i=1}^{k} \mu_i = 0$, $\sum_{i=1}^{k} \mu_i x_i = 0$ και τουλάχιστον ένα εκ των $\mu_1, \ldots, \mu_{k-1}$ δεν είναι μηδέν. □

Θεώρημα (Κ. Καραθεοδωρή) 1.2.6. *Έστω $A \subseteq \mathbb{R}^n$. Τότε*

$$convA = \{\sum_{i=1}^{k} \lambda_i x_i : x_1, \ldots, x_k \in A,\ \sum_{i=1}^{k} \lambda_i = 1,\ \lambda_i \in [0,1],\ 1 \leqslant k \leqslant n+1\}.$$

Δηλαδή κάθε στοιχείο της κυρτής θήκης του A είναι κυρτός συνδυασμός το πολύ $n+1$ στοιχείων του A.

Απόδειξη. Έστω $z \in conv A$. Τότε υπάρχουν $k \in \mathbb{N}$, $x_1, \ldots, x_k \in A$, $\lambda_1, \ldots, \lambda_k \in [0,1]$ με $\sum_i^k \lambda_i = 1$, ώστε $z = \sum_{i=1}^k \lambda_i x_i$. Ας υποθέσουμε ότι $k > n+1$. Θα δείξουμε τότε ότι το z μπορεί να γραφεί και ως κυρτός συνδυασμός $k-1$ στοιχείων του A. Αφού $k \geqslant n+2$, από την πρόταση 1.2.5 υπάρχουν $\mu_1, \ldots, \mu_k \in \mathbb{R}$ όχι όλα μηδέν, ώστε $\sum_{i=1}^k \mu_i = 0$ και $\sum_{i=1}^k \mu_i x_i = 0$. Επομένως για κάθε $\lambda \in \mathbb{R}$ μπορούμε να γράψουμε $\sum_{i=1}^k (\lambda_i - \lambda\mu_i) = 1$ και $\sum_{i=1}^k (\lambda_i - \lambda\mu_i) x_i = z$.

Θέτουμε $\lambda = \min\{\frac{\lambda_i}{\mu_i} : \mu_i > 0\}$. Τότε για κάθε $i = 1, \ldots, k$ ισχύει $\lambda_i - \lambda\mu_i \geqslant 0$. Πράγματι, αν $\mu_i \leqslant 0 \Leftrightarrow \lambda\mu_i \leqslant 0 \Leftrightarrow -\lambda\mu_i \geqslant 0 \Leftrightarrow \lambda_i - \lambda\mu_i \geqslant 0$. Αν $\mu_i > 0$, τότε $\lambda_i - \lambda\mu_i \geqslant 0 \Leftrightarrow \lambda_i \geqslant \lambda\mu_i \Leftrightarrow \frac{\lambda_i}{\mu_i} \geqslant \lambda$ που ισχύει από την επιλογή του λ.

Άρα, το $z = \sum_{i=1}^k (\lambda_i - \lambda\mu_i) x_i$ είναι κυρτός συνδυασμός k στοιχείων του A. Επειδή $\lambda = \frac{\lambda_{i_0}}{\mu_{i_0}} \Leftrightarrow \lambda_{i_0} - \lambda\mu_{i_0} = 0$ για κάποιο $i_0 \in \{1, \ldots, k\}$, μπορούμε να γράψουμε

$$z = \sum_{\substack{i=1 \\ i \neq i_0}}^k (\lambda_i - \lambda\mu_i) x_i \quad \text{με} \quad \#\{x_i : i = 1, \ldots, k \text{ και } i \neq i_0\} = k-1.$$

□

1.3 Κλειστές Κυρτές Θήκες

Ορισμός 1.3.1. *Ένα σύνολο είναι κλειστό αν περιέχει κάθε οριακό του σημείο.*

Τα περισσότερα από τα κυρτά σύνολα που μας απασχολούν είναι κλειστά. Γι'αυτό είναι σκόπιμο να δώσουμε τον ορισμό της **κλειστής κυρτής θήκης**.

Ορισμός 1.3.2. *Έστω $A \subseteq \mathbb{R}^n$. Ονομάζουμε κλειστή κυρτή θήκη του A την τομή όλων των κλειστών κυρτών υποσυνόλων του $\mathbb{R}^n$ που περιέχουν το A,*

$$\overline{conv} A := \cap\{C \subseteq \mathbb{R}^n : C \text{ κλειστό κυρτό και } A \subseteq C\}.$$

Ένας άλλος τρόπος να περιγράψουμε την κλειστή κυρτή θήκη αποδεικνύεται παρακάτω.

Πρόταση 1.3.3. *Η κλειστή κυρτή θήκη $\overline{conv} A$ του ορισμού 1.3.2 είναι η κλειστότητα $\overline{conv A}$ της κυρτής θήκης του A.*

Απόδειξη. Επειδή το $\overline{convA}$ είναι κλειστό σύνολο που περιέχει το A, περιέχει και το $\overline{conv}A$ επίσης. Αντίστροφα, παίρνουμε ένα κλειστό κυρτό σύνολο C που περιέχει το A. Ως κυρτό, το C περιέχει το $convA$. Ως κλειστό, περιέχει επίσης το $\overline{convA}$. Αφού το C το πήραμε τυχαία, συμπεραίνουμε ότι $\cap C = \overline{conv}A \supset \overline{convA}$. □

1.4 Ακραία Σημεία και Ακραία Υποσύνολα

Ορισμός 1.4.1. *Έστω $C \subseteq \mathbb{R}^n$ κυρτό. Ένα σημείο $x \in C$ ονομάζεται ακραίο σημείο του C αν δεν υπάρχουν $x_1 \neq x_2$ στο C και $\lambda \in (0,1)$, ώστε $x = \lambda x_1 + (1-\lambda)x_2$. Το σύνολο των ακραίων σημείων του C συμβολίζεται με $extC$.*

Παραδείγματα

1. Σε ένα ευθύγραμμο τμήμα $[x, x']$ τα ακραία σημεία είναι τα x, x'.
2. Σε ένα τρίγωνο τα ακραία σημεία είναι οι 3 κορυφές του.
3. Σε έναν ημίχωρο δεν υπάρχουν ακραία σημεία.
4. Στην κλειστή μοναδιαία μπάλα $\overline{B}(0,1)$ με την ευκλείδια νόρμα τα ακραία σημεία είναι κάθε σημείο με νόρμα ίση με 1.

Θυμίζουμε κάποια χρήσιμα στοιχεία για τη συμπάγεια.

Ορισμός 1.4.2. *Ένα σύνολο A είναι συμπαγές αν για κάθε ακολουθία $(\alpha_n)_n$ με $\alpha_n \in A$ $\forall n \in \mathbb{N}$, υπάρχει $\alpha \in A$ και υπακολουθία $(\alpha_{k_n})_n$ της $(\alpha_n)_n$ με $\alpha_{k_n} \to \alpha$.*

Πρόταση 1.4.3. *Έστω A συμπαγές και $f\colon A \to \mathbb{R}$ συνεχής, τότε η f λαμβάνει μέγιστη και ελάχιστη τιμή.*

Παρακάτω αποδεικνύουμε μια χρήσιμη τοπολογική ιδιότητα του $extA$.

Πρόταση 1.4.4. *Αν $A \subseteq \mathbb{R}^n$ είναι συμπαγές και κυρτό, τότε $extA \neq \emptyset$.*

Απόδειξη. Θεωρούμε τη συνάρτηση $\phi(x) = \|x\|^2$. Επειδή το A είναι συμπαγές, από την πρόταση 1.4.3 υπάρχει $\bar{x} \in A$ που μεγιστοποιεί την ϕ. Υποθέτουμε ότι υπάρχουν $x_1, x_2 \in A$ τέτοια ώστε $\bar{x} = 1/2(x_1 + x_2)$. Αν υποθέσουμε ότι $x_1 \neq x_2$, τότε

$$\|\bar{x}\|^2 = \|\frac{1}{2}(x_1 + x_2)\|^2 < \frac{1}{2}(\|x_1\|^2 + \|x_2\|^2) \leqslant \frac{1}{2}(\|\bar{x}\|^2 + \|\bar{x}\|^2) = \|\bar{x}\|^2$$

το οποίο είναι άτοπο. Συνεπώς $x_1 = x_2 = \bar{x}$, δηλαδή $\bar{x} \in extA$. □

Θεώρημα (H. Minkowski) 1.4.5. *Έστω C συμπαγές, κυρτό στο $\mathbb{R}^n$. Τότε το C είναι η κυρτή θήκη των ακραίων σημείων του: $C = conv(extC)$.*

Ορισμός 1.4.6. *Έστω $A \subseteq \mathbb{R}^n$. Ένα μη κενό, κλειστό, κυρτό σύνολο $F \subseteq A$ λέγεται ακραίο του A αν ικανοποιεί την εξής ιδιότητα:*

$$\left.\begin{array}{c}(x_1, x_2) \in A \times A \\ \exists \lambda \in (0,1) : \lambda x_1 + (1-\lambda) x_2 \in F\end{array}\right\} \Longrightarrow [x_1, x_2] \subseteq F$$

Παραδείγματα:

1. Σε κάθε κυρτό σύνολο C τα $\emptyset$, C, καθώς και κάθε $x_0 \in extC$ είναι ακραία υποσύνολά του. Το C είναι το μοναδικό ακραίο υποσύνολο διάστασης $\dim C$, ενώ τα ακραία σημεία είναι τα μοναδικά ακραία υποσύνολα διάστασης 0.

2. Σε ένα τρίγωνο με κορυφές τα x_0, x_1, x_2 τα ακραία υποσύνολά του είναι το $\emptyset$, τα $\{x_0\}, \{x_1\}, \{x_2\}$, τα $[x_0, x_1], [x_1, x_2], [x_2, x_0]$, καθώς και το ίδιο το τρίγωνο.

3. Στην κλειστή μπάλα $\overline{B}(0,1)$ τα ακραία υποσύνολα είναι το $\emptyset$, τα ακραία σημεία $\{x\}$ με $\|x\| = 1$ και η ίδια η μπάλα $\overline{B}(0,1)$.

1.5 Προβολή σε Κλειστά Κυρτά Σύνολα

Ορισμός 1.5.1. *Έστω μη κενό $A \subseteq \mathbb{R}^n$. Έστω $x \in \mathbb{R}^n$. Ορίζουμε την απόσταση του x από το A ως*

$$d(x, A) := \inf \{d(x, \alpha) : \alpha \in A\}.$$

Δεν ισχύει γενικά ότι υπάρχει $\alpha_0 \in A$ με $d(x, \alpha_0) = d(x, A)$. Για παράδειγμα, έστω $A = (0,1)$ και $x = 2$. Τότε $d(x, A) = 1$, όμως δεν υπάρχει $\alpha \in A$ με $d(x, \alpha) = 1$. Στην επόμενη πρόταση θα αποδείξουμε ότι αρκεί το A να είναι κλειστό για να υπάρχει τέτοιο α_0.

Πρόταση 1.5.2. *Έστω μη κενό $A \subseteq \mathbb{R}^n$ κλειστό. Τότε για κάθε $x \in \mathbb{R}^n$ υπάρχει $\alpha \in A$, ώστε $d(x, A) = d(x, \alpha)$.*

Απόδειξη. Επειδή $d(x, A) := \inf \{d(x, \alpha) : \alpha \in A\}$, έπεται ότι υπάρχει $(\alpha_n)_n \subseteq A$ ώστε

$$d(x, \alpha_n) \to d(x, A) \tag{1.1}$$

Άρα η $(\alpha_n)_n$ είναι φραγμένη, διότι

$$\|\alpha_n\| \leqslant \|\alpha_n - x + x\| \leqslant \|\alpha_n - x\| + \|x\| \leqslant d(x, \alpha_n) + \|x\| \tag{1.2}$$

Η $(d(x, \alpha_n))_n$ είναι φραγμένη ως συγκλίνουσα. Άρα υπάρχει $M > 0$ τέτοιο ώστε

$$d(x, \alpha_n) \leqslant M, \ \forall n \in \mathbb{N} \tag{1.3}$$

Άρα από (1.2) και (1.3) έχουμε $\|\alpha_n\| \leqslant M + \|x\|$ για κάθε $n \in \mathbb{N}$.

Άρα από θεώρημα Bolzano - Weierstrass υπάρχει υπακολουθία $(\alpha_{k_n})_n$ της $(\alpha_n)_n$ και $\alpha \in \mathbb{R}^n$ ώστε $\alpha_{k_n} \to \alpha$ ή

$$d(\alpha_{k_n}, x) \to d(\alpha, x) \tag{1.4}$$

Επειδή A κλειστό, έχουμε $\alpha \in A$. Παρατηρούμε ότι η $(d(\alpha_{k_n}, x))_n$ είναι υπακολουθία της $(d(\alpha_n, x))_n$ και άρα από (1.1)

$$d(\alpha_{k_n}, x) \to d(x, A) \tag{1.5}$$

Από (1.4) και (1.5) έπεται ότι $d(x, A) = d(x, \alpha)$. □

Λήμμα 1.5.3. *Έστω $A \subseteq \mathbb{R}^n$ κυρτό, $x \in \mathbb{R}^n$, $\alpha_0 \in A$. Τα επόμενα είναι ισοδύναμα:*

(1) $d(x, A) = d(x, \alpha_0)$

(2) $\langle x - \alpha_0, \alpha - \alpha_0 \rangle \leqslant 0,\ \forall \alpha \in A$

Απόδειξη. (1) $\Rightarrow$ (2) Αφού $d(x, \alpha_0) = d(x, A) = \inf\{d(x, \alpha) : \alpha \in A\}$ ισχύει $d(x, \alpha_0) \leqslant d(x, \alpha)$ για κάθε $\alpha \in A$. Έστω $\alpha \in A$ και $\lambda \in (0, 1]$. Τότε $(1 - \lambda)\alpha_0 + \lambda\alpha \in A$, επειδή το A είναι κυρτό. Άρα

$$\begin{aligned}
d(x, \alpha_0) &\leqslant d(x, (1-\lambda)\alpha_0 + \lambda\alpha) \Leftrightarrow \\
\|x - \alpha_0\|^2 &\leqslant \|x - (1-\lambda)\alpha_0 + \lambda\alpha\|^2 \Leftrightarrow \\
\|x - \alpha_0\|^2 &\leqslant \|(x - \alpha_0) - \lambda(\alpha - \alpha_0)\|^2 \Leftrightarrow \\
\langle x - \alpha_0, x - \alpha_0 \rangle &\leqslant \langle (x - \alpha_0) - \lambda(\alpha - \alpha_0), (x - \alpha_0) - \lambda(\alpha - \alpha_0) \rangle \Leftrightarrow \\
\langle x - \alpha_0, x - \alpha_0 \rangle &\leqslant \langle x - \alpha_0, x - \alpha_0 \rangle + 2\langle x - \alpha_0, -\lambda(\alpha - \alpha_0) \rangle + \langle \lambda(\alpha - \alpha_0), \lambda(\alpha - \alpha_0) \rangle \Leftrightarrow \\
0 &\leqslant -2\lambda\langle x - \alpha_0, \alpha - \alpha_0 \rangle + \lambda^2\|\alpha - \alpha_0\|^2 \Leftrightarrow \\
0 &\leqslant -2\langle x - \alpha_0, \alpha - \alpha_0 \rangle + \lambda\|\alpha - \alpha_0\|^2
\end{aligned}$$

Παίρνοντας το όριο όταν $\lambda \to 0$ έχουμε ότι

$$\begin{aligned}
0 &\leqslant -2\langle x - \alpha_0, \alpha - \alpha_0 \rangle \Leftrightarrow \\
\langle x - \alpha_0, \alpha - \alpha_0 \rangle &\leqslant 0
\end{aligned}$$

(2) $\Rightarrow$ (1) Αρκεί να δείξουμε ότι $d(x, \alpha_0) \leqslant d(x, \alpha),\ \forall \alpha \in A$. Πράγματι,

$$\begin{aligned}
d(x, \alpha_0) &\leqslant d(x, \alpha) \Leftrightarrow \\
\|x - \alpha_0\|^2 &\leqslant \|x - \alpha\|^2 \\
\|x - \alpha_0\|^2 &\leqslant \|(x - \alpha_0) + (\alpha_0 - \alpha)\|^2 \Leftrightarrow \\
\|x - \alpha_0\|^2 &\leqslant \langle (x - \alpha_0) + (\alpha_0 - \alpha), (x - \alpha_0) + (\alpha_0 - \alpha) \rangle \Leftrightarrow \\
\|x - \alpha_0\|^2 &\leqslant \|x - \alpha_0\|^2 + 2\langle \alpha_0 - \alpha, x - \alpha_0 \rangle + \|\alpha_0 - \alpha\|^2 \Leftrightarrow \\
0 &\leqslant 2\langle x - \alpha_0, \alpha_0 - \alpha \rangle + \|\alpha_0 - \alpha\|^2 \Leftrightarrow \\
0 &\leqslant -2\langle x - \alpha_0, \alpha - \alpha_0 \rangle + \|\alpha_0 - \alpha\|^2 \quad \text{που ισχύει}
\end{aligned}$$

□

Πρόταση 1.5.4. *Έστω μη κενό $A \subseteq \mathbb{R}^n$ κλειστό και κυρτό. Τότε για κάθε $x \in \mathbb{R}^n$ υπάρχει μοναδικό $\alpha_0 \in A$, ώστε $d(x, A) = d(x, \alpha_0)$.*

Απόδειξη. Επειδή το A είναι κλειστό, από την πρόταση 1.5.2 έχουμε ότι υπάρχει $\alpha \in A$ με $d(x, \alpha) = d(x, A)$. Υποθέτουμε ότι υπάρχουν δύο τέτοια σημεία $\alpha_1, \alpha_2 \in A$ ώστε $d(x, \alpha_1) = d(x, \alpha_2) = d(x, A)$. Από το λήμμα 1.5.3 έχουμε $\langle x - \alpha_1, \alpha_2 - \alpha_1 \rangle \leqslant 0$ και $\langle x - \alpha_2, \alpha_1 - \alpha_2 \rangle = \langle \alpha_2 - x, \alpha_2 - \alpha_1 \rangle \leqslant 0$. Προσθέτοντας κατά μέλη τις ανισότητες έχουμε

$$\begin{aligned} \langle \alpha_2 - \alpha_1, \alpha_2 - \alpha_1 \rangle \leqslant 0 \quad &\Leftrightarrow \\ \|\alpha_2 - \alpha_1\|^2 \leqslant 0 \quad &\Leftrightarrow \\ \|\alpha_2 - \alpha_1\| = 0 \quad &\Leftrightarrow \\ \alpha_2 = \alpha_1 \quad & \end{aligned}$$

□

Παρατηρούμε ότι η ιδιότητα του κλειστού μας δίνει την ύπαρξη του α_0. Η μοναδικότητά του όμως εξασφαλίζεται από την κυρτότητα του A.

Τέλος, δίνουμε τον ορισμό της προβολής.

Ορισμός 1.5.5. *Έστω μη κενό $A \subseteq \mathbb{R}^n$ κλειστό και κυρτό. Ορίζουμε την προβολή του $\mathbb{R}^n$ στο A να είναι η απεικόνιση $p_A \colon \mathbb{R}^n \to A$ όπου για κάθε $x \in \mathbb{R}^n$, $p_A(x)$ είναι το μοναδικό σημείο του A με την ιδιότητα $\|x - p_A(x)\| = d(x, A)$ δηλαδή το p_A είναι το μοναδικό πλησιέστερο στο x σημείο του A.*

Κεφάλαιο 2

Βασικά Θεωρήματα σε Χώρους Banach

2.1 Διαχωριστικά Θεωρήματα

Έστω E ένας χώρος Banach. Υπενθυμίζουμε ότι **γραμμικό συναρτησιακό** ονομάζουμε μια γραμμική απεικόνιση ορισμένη στο E ή έναν υπόχωρό του με τιμές στο $\mathbb{R}$. Θα δώσουμε τώρα τον ορισμό για το υπερεπίπεδο.

Ορισμός 2.1.1. *Ένα υπερεπίπεδο είναι ένα σύνολο της μορφής*

$$H = \{x \in E : f(x) = \alpha\}$$

όπου f είναι ένα γραμμικό συναρτησιακό πάνω στο E, όχι ταυτοτικά μηδέν και $\alpha \in \mathbb{R}$. Λέμε ότι το H είναι ένα υπερεπίπεδο με εξίσωση $[f = \alpha]$.

Πρόταση 2.1.2. *Το υπερεπίπεδο με εξίσωση $[f = \alpha]$ είναι κλειστό αν και μόνο αν το f είναι συνεχές.*

Απόδειξη. Είναι φανερό ότι αν το f είναι συνεχές, τότε το H είναι κλειστό. Αντίστροφα ας υποθέσουμε ότι το H είναι κλειστό. Το συμπλήρωμα $\complement H$ του H είναι ανοιχτό και μη κενό (αφού $f \neq 0$). Έστω $x_0 \in \complement H$ και ας υποθέσουμε πχ ότι $f(x_0) < \alpha$. Έστω $r > 0$ τέτοιο ώστε $B(x_0, r) \subset \complement H$ όπου

$$B(x_0, r) = \{x \in E : \|x - x_0\| < r\}.$$

Τότε

$$f(x) < \alpha \quad \forall x \in B(x_0, r). \tag{2.1}$$

Πράγματι ας υποθέσουμε ότι $f(x_1) > \alpha$ για κάποιο $x_1 \in B(x_0, r)$. Το ευθύγραμμο τμήμα

$$\{x_t = (1-t)x_1 + tx_0, \quad t \in [0,1]\}$$

περιέχεται στο $B(x_0, r)$ και άρα $f(x_t) \neq \alpha \forall t \in [0,1]$. Εξάλλου, $f(x_t) = \alpha$ για $t = \dfrac{f(x_1) - \alpha}{f(x_1) - f(x_0)}$, που είναι άτοπο και άρα ισχύει η (2.1). Προκύπτει από την (2.1) ότι

$$f(x_0 + rz) < \alpha \quad \forall z \in B(0,1).$$

Επομένως το f είναι συνεχές και ισχύει $\|f\| < \dfrac{1}{r}(\alpha - f(x_0))$ □

Ορισμός 2.1.3. *Έστω $A \subset E$ και $B \subset E$. Λέμε ότι το υπερεπίπεδο H με εξίσωση $[f = \alpha]$ διαχωρίζει τα A και B με την ευρεία έννοια αν*

$$f(x) \leqslant \alpha \ \forall x \in A \quad \text{και} \quad f(x) \geqslant \alpha \ \forall x \in B.$$

Λέμε ότι το H διαχωρίζει τα A και B με τη στενή έννοια αν υπάρχει $\varepsilon > 0$ τέτοιο ώστε

$$f(x) \leqslant \alpha - \varepsilon \ \forall x \in A \quad \text{και} \quad f(x) \geqslant \alpha + \varepsilon \ \forall x \in B.$$

Παρακάτω θα διατυπώσουμε τις δυο γεωμετρικές μορφές του θεωρήματος Hahn-Banach.

Θεώρημα (Hahn-Banach, πρώτη γεωμετρική μορφή) 2.1.4. *Έστω $A \subset E$ και $B \subset E$ δυο σύνολα κυρτά, μη κενά και ξένα. Υποθέτουμε ότι το A είναι ανοιχτό. Τότε υπάρχει ένα κλειστό υπερεπίπεδο που διαχωρίζει τα A και B με την ευρεία έννοια.*

Θεώρημα (Hahn-Banach, δεύτερη γεωμετρική μορφή) 2.1.5. *Έστω $A \subset E$ και $B \subset E$ δυο σύνολα κυρτά, μη κενά και ξένα. Υποθέτουμε ότι το A είναι κλειστό και ότι το B είναι συμπαγές. Τότε υπάρχει ένα κλειστό υπερεπίπεδο που διαχωρίζει τα A και B με τη στενή έννοια.*

2.2 Θεώρημα Mazur

Θεώρημα (Mazur) 2.2.1. *Έστω E χώρος Banach και X ένα σχετικά συμπαγές υποσύνολο του E. Τότε το $convX$ είναι επίσης σχετικά συμπαγές υποσύνολο του E.*

Απόδειξη. Έστω V μια περιοχή του 0 στο E. Τότε υπάρχει μια κυρτή περιοχή του 0 ,W, στο E έτσι ώστε

$$W + W \subset V. \tag{2.2}$$

Αν το X είναι ένα σχετικά συμπαγές υποσύνολο του E, τότε υπάρχουν τελικά πολλά σημεία $x_1, \ldots, x_n$ στο X έτσι ώστε

$$X \subset \{x_1, \ldots, x_n\} + W.$$

Αν είναι $x \in convX$ τότε σύμφωνα με την Πρόταση 1.2.4 το x είναι κυρτός συνδυασμός στοιχείων $y_1, \ldots, y_n$ στο X:

$$x = \sum_{i=1}^{m} t_i y_i.$$

Υπάρχουν τότε m συντελεστές $j_1, \ldots, j_m$ έτσι ώστε

$$y_i - x_{j_i} \in W, \quad i = 1, \ldots, m.$$

Επομένως,

$$x - \sum_{i=1}^{m} t_i x_{j_i} = \sum_{i=1}^{m} t_i (y_i - x_{j_i}) \in W.$$

Παίρνουμε

$$convX \subset conv\{x_1, \ldots, x_n\} + W. \tag{2.3}$$

Έστω W' κυρτή περιοχή του 0 στο E τότε υπάρχει ένα $t > 0$ έτσι ώστε

$$\{x_1, \ldots, x_n\} \subset tW'.$$

Εξαιτίας της κυρτότητας του W' τότε και κάθε κυρτός συνδυασμός των $x_1, \ldots, x_n$ ανήκει στο tW'. Σύμφωνα πάλι με την Πρόταση 1.2.4 ισχύει

$$conv\{x_1, \ldots, x_n\} \subset tW'.$$

δηλαδή το $conv\{x_1, \ldots, x_n\}$ είναι περιορισμένο στο E και ως υποσύνολο του $sp\{x_1, \ldots, x_n\}$ ανήκει τελικά και στο διανυσματικό υπόχωρο $sp\{x_1, \ldots, x_n\}$ του E. Έτσι η $conv\{x_1, \ldots, x_n\}$ είναι σχετικά συμπαγής σε αυτό τον υπόχωρο. Υπάρχουν επομένως τελικά πολλά σημεία $z_1, \ldots, z_p$ στην $conv\{x_1, \ldots, x_n\}$ έτσι ώστε

$$conv\{x_1, \ldots, x_n\} \subset \{z_1, \ldots, z_p\} + W.$$

Από (2.2) και (2.3) προκύπτει

$$convX \subset \{z_1, \ldots, z_p\} + V,$$

δηλαδή η $convX$ είναι σχετικά συμπαγής στο E. □

Πόρισμα 2.2.2. *Έστω E χώρος Banach και $X \subset E$ σχετικά συμπαγές. Τότε ισχύει και για κάθε υποσύνολο Y του E ότι*

$$\overline{conv}(X + Y) = \overline{conv}X + \overline{conv}Y.$$

Αν το Y είναι και σχετικά συμπαγές τότε επιπλέον ισχύει

$$\overline{conv}(X \cup Y) = conv[\overline{conv}X \cup \overline{conv}Y].$$

2.3 Θεώρημα Krein-Milman

Πρόταση 2.3.1. *Έστω X χώρος Banach και K ένα μη κενό και κλειστό υποσύνολο του X.*

(1) Αν $\{A_i : i \in I\}$ είναι μια οικογένεια από ακραία υποσύνολα του K με $\cap_{i\in I} A_i = A \neq \emptyset$ τότε το A είναι ακραίο υποσύνολο του K.

(2) Αν $A \subset B \subset K$ ώστε το B είναι ακραίο υποσύνολο του K και το A ακραίο υποσύνολο του B, τότε το A είναι ακραίο υποσύνολο του K.

Απόδειξη. (1) Το A είναι κλειστό και ακραίο υποσύνολο του K ως τομή κλειστών και κυρτών. Έστω $z \in A$ και $0 < \lambda < 1$, $x, y \in K$ ώστε $z = \lambda x + (1-\lambda)y$. Τότε $z \in A_i$ για κάθε $i \in I$ και άρα $x, y \in A_i, \forall i \in I$. Συνεπώς, $x, y \in \cap_{i\in I} A_i = A$ και άρα το A είναι ακραίο υποσύνολο του K.

(2) Έστω $0 < \lambda < 1, x, y \in K$ και $z \in A$ ώστε $z = \lambda x + (1-\lambda)y$. Αφού $A \subset B$, έχουμε ότι $z \in B$ και επειδή το B είναι ακραίο υποσύνολο του K, έχουμε ότι $x, y \in B$. Επειδή το A είναι ακραίο υποσύνολο του B, έπεται ότι $x, y \in A$. Άρα το A είναι ακραίο υποσύνολο του K. □

ΠΑΡΑΤΗΡΗΣΗ. Από το (2) της προηγούμενης πρότασης έχουμε άμεσα και το εξής: Αν $A \subset K$ ακραίο υποσύνολο του K, τότε $extA \subset extK$.

Θεώρημα (Krein-Milman) 2.3.2. *Έστω X χώρος Banach και K ένα μη κενό συμπαγές κυρτό υποσύνολο του X. Τότε $K = \overline{conv}(extK)$.*

Απόδειξη. Δείχνουμε κατ'αρχήν ότι $extK \neq \emptyset$. Θέτουμε

$$\mathcal{A} = \{A \subset K : A \text{ ακραίο υποσύνολο του } K\}.$$

ΒΗΜΑ 1. Η οικογένεια $\mathcal{A}$ περιέχει ένα ελαχιστικό στοιχείο ως προς την $\subset$ δηλαδή υπάρχει $A \in \mathcal{A}$ ώστε για κάθε $B \in \mathcal{A}$ με $B \subset A$ έχουμε ότι $B = A$.

Απόδειξη. Ορίζουμε $A < B \Leftrightarrow B \subset A$ για κάθε $A, B \in \mathcal{A}$. Η $<$ είναι σχέση μερικής διάταξης στο $\mathcal{A}$. Πρέπει να δειχθεί ότι η $\mathcal{A}$ περιέχει ένα μεγιστικό (ως προς την $<$) στοιχείο. Θα εφαρμόσουμε το λήμμα Zorn. Έστω $\mathcal{C} = \{A_i : i \in I\}$ μια αλυσίδα στην $\mathcal{A}$. Αρκεί να δειχθεί ότι η $\mathcal{C}$ έχει άνω φράγμα στην $\mathcal{A}$. Θέτουμε $A = \cap_{i\in I} A_i$. Τότε το A είναι μη κενό. Πράγματι η οικογένεια $\mathcal{C}$ (επειδή είναι αλυσίδα) έχει την ιδιότητα της πεπερασμένης τομής, αποτελείται από κλειστά σύνολα και το K είναι συμπαγές. Άρα $\cap_{i\in I} A_i \neq \emptyset$. Επίσης, από πρόταση 2.3.1 έχουμε ότι $A \in \mathcal{A}$ και προφανώς $A > A_i, \forall i \in I$ αφού $A \subset A_i, \forall i \in I$ □

ΒΗΜΑ 2. Κάθε ελαχιστικό στοιχείο της $\mathcal{A}$ είναι μονοσύνολο.

Απόδειξη. Υποθέτουμε προς απαγωγή σε άτοπο ότι υπάρχει A ελαχιστικό στοιχείο της $\mathcal{A}$ με τουλάχιστον δυο διαφορετικά σημεία $x, y \in A$. Γνωρίζουμε ότι υπάρχει $x^* \in X^*$ ώστε $x^*(x) \neq x^*(y)$ και έστω $x^*(x) < x^*(y)$. Επειδή το A είναι συμπαγές (ως κλειστό υποσύνολο του συμπαγούς K) έχουμε ότι υπάρχει $\alpha \in \mathbb{R}$ ώστε $\alpha = \sup_{z\in A} x^*(z)$ και επιπλέον το σύνολο $B = \{z \in A : x^*(z) = \alpha\}$ είναι μη κενό. Θα δείξουμε ότι το B είναι ακραίο υποσύνολο του A. Έστω $y_1, y_2 \in A$ και $\lambda \in (0,1)$ με $\lambda y_1 + (1-\lambda)y_2 \in B$. Τότε $x^*(\lambda y_1 + (1-\lambda)y_2) = \alpha$.

Επειδή το x^* είναι γραμμικό, έχουμε ισοδύναμα $\lambda(x^*(y_1) - x^*(y_2)) = \alpha - x^*(y_2)$. Άρα $x^*(y_1) \geqslant x^*(y_2)$. Με τον ίδιο τρόπο, αν θέσουμε $\lambda' = 1 - \lambda$ δείχνουμε ότι $x^*(y_1) \leqslant x^*(y_2)$. Άρα $x^*(y_1) = x^*(y_2) = \alpha$. Επομένως το B είναι ακραίο υποσύνολο του A και άρα από την πρόταση 2.3.1 έχουμε ότι το B είναι ακραίο υποσύνολο του K, δηλαδή $B \in \mathcal{A}$. Επειδή $x \notin B$, έχουμε ότι $B \in \mathcal{A}$ και $B \subsetneqq A$, άτοπο αφού το A είναι ελαχιστικό ακραίο υποσύνολο του K. □

Από τα προηγούμενα βήματα έχουμε ότι $extK \neq \emptyset$, αφού αν $\{x\}$ είναι ακραίο υποσύνολο του K τότε προφανώς $x \in extK$.

Προχωρούμε τώρα να δείξουμε ότι $K = \overline{conv}(extK)$.

ΒΗΜΑ 3. $K = \overline{conv} extK$.

Απόδειξη. Θέτουμε $L = \overline{conv} extK$. Τότε το L είναι κλειστό υποσύνολο του K και άρα συμπαγές και κυρτό ως κλειστότητα του κυρτού συνόλου $convextK$. Υποθέτουμε προς απαγωγή σε άτοπο ότι $L \subsetneqq K$. Τότε υπάρχει $x \in K \setminus L$. Από το διαχωριστικό θεώρημα υπάρχει $x^* \in X^*$ ώστε

$$\sup_{z \in L} x^*(z) < x^*(x).$$

Έστω $\alpha = \sup\{x^*(y) : y \in K\}$ και $B = \{y \in K : x^*(y) = \alpha\}$. Όπως και στο Βήμα 2, το B είναι ακραίο υποσύνολο του K. Επειδή το B είναι συμπαγές κυρτό από τα δυο πρώτα βήματα έχουμε ότι $extB \neq \emptyset$. Επειδή όμως το B είναι ακραίο υποσύνολο του K, από την παρατήρηση

$$extB \subset extK \subset L.$$

Αυτό όμως είναι αδύνατο αφού αν $y \in extB$ τότε $x^*(y) = \alpha$, ενώ αν $y \in L$, $x^*(y) < x^*(x) \leqslant \alpha$. Άρα $K = \overline{conv}(extK)$. □

□

2.4 Προβολή σε Κλειστό Κυρτό Σύνολο σε Χώρο Hilbert

Στη συνέχεια με H συμβολίζουμε έναν χώρο Hilbert.

Θεώρημα 2.4.1. *Έστω $K \subset H$ ένα κλειστό, κυρτό και μη κενό. Τότε για κάθε $f \in H$ υπάρχει $u \in K$ μοναδικό τέτοιο ώστε*

$$\|f - u\| = \min_{v \in K} \|f - v\| \tag{2.4}$$

Επιπλέον το u χαρακτηρίζεται από την ιδιότητα:

$$\begin{cases} u \in K \\ \langle f - u, v - u \rangle \leqslant 0 \ \forall v \in K \end{cases} \tag{2.5}$$

Συμβολίζουμε $u = P_K f$=προβολή του f πάνω στο K.

Απόδειξη. **Ύπαρξη.**
Έστω (v_n) μια ελαχιστοποιητική ακολουθία για την (2.4) δηλαδή $v_n \in K$ και

$$d_n = \|f - v_n\| \to d = \inf_{v \in K} \|f - v\|.$$

Θα δείξουμε ότι η (v_n) είναι Cauchy. Εφαρμόζοντας την ταυτότητα του παραλληλογράμμου με $a = f - v_n, b = f - v_m$ παίρνουμε

$$\left\| f - \frac{v_n + v_m}{2} \right\|^2 + \left\| \frac{v_n - v_m}{2} \right\|^2 = \frac{1}{2}(d_n^2 + d_m^2).$$

Επειδή $\dfrac{v_n + v_m}{2} \in K$, ισχύει $\left\| f - \dfrac{v_n + v_m}{2} \right\| \geqslant d$. Επομένως,

$$\left\| \frac{v_n - v_m}{2} \right\|^2 \leqslant \frac{1}{2}(d_n^2 + d_m^2) - d^2 \quad \text{και} \quad \lim_{m,n \to \infty} \|v_n - v_m\| = 0.$$

Συνεπώς, $v_n \to u \in K$ και $d = \|f - u\|$.

Ισοδυναμία των (2.4) **και** (2.5).
Έστω $u \in K$ που ικανοποιεί την (2.5) και έστω $w \in K$. Ισχύει

$$v = (1 - t)u + tw \in K \quad \text{για} \quad t \in (0, 1]$$

και άρα

$$\|f - u\| \leqslant \|f - [(1 - t)u + tw]\| = \|(f - u) - t(w - u)\|.$$

Επομένως,

$$\|f - u\|^2 \leqslant \|f - u\|^2 - 2t\langle f - u, w - \rangle + t^2\|w - u\|^2$$

δηλαδή $2\langle f - u, w - u\rangle \leqslant t\|w - u\|^2$. Όταν $t \to 0$ παίρνουμε την (2.5). Αντίστροφα, έστω u που ικανοποιεί την (2.5). Έχουμε τότε

$$\|u - f\|^2 - \|v - f\|^2 = 2\langle f - u, v - u\rangle - \|u - v\|^2 \leqslant 0 \quad \forall v \in K$$

από όπου προκύπτει η (2.4).

Μοναδικότητα.
Έστω u_1 και u_2 που ικανοποιούν την (2.5). Έχουμε

$$\langle f - u_1, v - u_1\rangle \leqslant 0 \quad \forall v \in K \tag{2.6}$$

$$\langle f - u_2, v - u_2\rangle \leqslant 0 \quad \forall v \in K \tag{2.7}$$

Αντικαθιστώντας $v = u_2$ στην (2.6) και $v = u_1$ στην (2.7) παίρνουμε με πρόσθεση $\|u_1 - u_2\|^2 \leqslant 0$. □

Θεώρημα 2.4.2. *Έστω $M \subset H$ ένας γραμμικός κλειστός υπόχωρος και $f \in H$. Τότε το $u = P_M f$ χαρακτηρίζεται με*

$$\begin{cases} u \in M \\ \langle f - u, v \rangle = 0 \; \forall v \in M \end{cases} \tag{2.8}$$

Επιπλέον, ο P_M είναι ένας γραμμικός τελεστής.

Απόδειξη. Λόγω της (2.5) έχουμε

$$\langle f - u, v - u \rangle \leqslant 0 \quad \forall v \in M$$

και άρα

$$\langle f - u, tv - u \rangle \leqslant 0 \quad \forall v \in M, \forall t \in \mathbb{R},$$

οπότε

$$\langle f - u, v \rangle = 0 \quad \forall v \in M.$$

□

Κεφάλαιο 3

Κυρτές Συναρτήσεις στο $\mathbb{R}$

Σε αυτό το κεφάλαιο θα ασχοληθούμε με τις κυρτές συναρτήσεις ορισμένες σε ένα διάστημα I του $\mathbb{R}$.

3.1 Ορισμός

Ορισμός 3.1.1. *Μια συνάρτηση $f\colon I \to \mathbb{R}$ ονομάζεται κυρτή αν $\forall x, x' \in I$ και $\forall \lambda \in [0,1]$*

$$f(\lambda x + (1-\lambda)x') \leqslant \lambda f(x) + (1-\lambda) f(x').$$

Η f ονομάζεται αυστηρά κυρτή αν $\forall x, x' \in I$ με $x \neq x'$ και $\forall \lambda \in [0,1]$

$$f(\lambda x + (1-\lambda)x') < \lambda f(x) + (1-\lambda) f(x').$$

Η γεωμετρική ερμηνεία του ορισμού είναι ότι το ευθύγραμμο τμήμα που ενώνει τα σημεία $(x, f(x))$ και $(x', f(x'))$ βρίσκεται πάνω από το γράφημα της f. Κάποιοι ισοδύναμοι ορισμοί της κυρτότητας της $f\colon I \to \mathbb{R}$ είναι ότι ισχύει

(α′)

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

για κάθε $a, b, x \in I$ με $a < x < b$. Το δεξί μέρος της ανισότητας μπορεί να γραφεί

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

(β′)

$$f(\lambda a + \mu b) \leqslant \lambda f(a) + \mu f(b)$$

για κάθε $a, b \in I$ και για κάθε $\lambda, \mu \in \mathbb{R}$ τέτοια ώστε $\lambda, \mu > 0$, $\lambda + \mu = 1$.

Πρόταση 3.1.2. *Έστω* $f\colon I \to \mathbb{R}$. *Τότε*

$$f(\sum_{i=1}^{n} \lambda_i x_i) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)$$

για $n \in \mathbb{N}, x_i \in I, \lambda_i \geqslant 0\ (1 \leqslant i \leqslant n), \sum_{i=1}^{n} \lambda_i = 1.$

Απόδειξη. Με επαγωγή στο n. Για $n = 1$ προφανές. Έστω ότι ισχύει για $n = k$. Θα δείξουμε τότε ότι ισχύει και για $n = k+1$. Έστω $x_1, \ldots, x_{k+1} \in I$, $\lambda_1, \ldots, \lambda_{k+1} \in [0,1]$ με $\sum_{i=1}^{k+1} = 1$. Διακρίνουμε 3 περιπτώσεις:

•$\lambda_{k+1} = 1$

Τότε αφού $\lambda_i \geqslant 0$ και $\sum_{i=1}^{k+1} \lambda_i = 1$ έπεται ότι $\lambda_i = 0$ για κάθε $i = 1, \ldots, k$ και άρα

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f(x_{k+1}) = \sum_{i=1}^{k+1} \lambda_i f(x_i)$$

•$\lambda_{k+1} = 0$

Τότε $\sum_{i=1}^{k} \lambda_i = 1$ και

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leqslant \sum_{i=1}^{k} \lambda f(x_i) = \sum_{i=1}^{k+1} \lambda f(x_i)$$

•$0 < \lambda_{k+1} < 1$

Παρατηρούμε ότι $\sum_{i=1}^{k+1} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i + \lambda_{k+1} x_{k+1} = (1 - \lambda_{k+1}) \left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i\right) + \lambda_{k+1} x_{k+1}$ όπου $\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i$ είναι κυρτός συνδυασμός, διότι $\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = 1$. Επειδή f κυρτή, έπεται ότι

$$\begin{aligned} f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) &= f\left((1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i + \lambda_{k+1} x_{k+1}\right) \\ &\leqslant (1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i\right) + \lambda_{k+1} f(x_{k+1}) \end{aligned} \tag{3.1}$$

ενώ από επαγωγική υπόθεση έχουμε ότι ισχύει

$$f\left(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}x_i\right) \leqslant \sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(x_i) \tag{3.2}$$

Από (3.1) και (3.2) έχουμε ότι

$$\begin{aligned} f\left(\sum_{i=1}^{k+1}\lambda_i x_i\right) &\leqslant (1-\lambda_{k+1})f\left(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}x_i\right) + \lambda_{k+1}f(x_{k+1}) \\ &\leqslant (1-\lambda_{k+1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(x_i) + \lambda_{k+1}f(x_{k+1}) \\ &= \sum_{i=1}^{k+1}\lambda_i f(x_i) \end{aligned}$$

□

Θεώρημα 3.1.3. *Έστω* $f\colon I \to \mathbb{R}$ *κυρτή και* $a, b, x \in I,\ a < x < b$. *Τότε*

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

Αν η f *είναι αυστηρά κυρτή, τότε ισχύει η αυστηρή ανισότητα.*

Απόδειξη. Αφού η f είναι κυρτή, έχουμε

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

Από αυτή την ανισότητα συμπεραίνουμε

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

το οποίο αποδεικνύει την πρώτη ανισότητα. Ομοίως αποδεικνύεται και η δεύτερη. □

Η γεωμετρική ερμηνεία της ανισότητας είναι ότι η κλίση του γραφήματος της f μεγαλώνει, καθώς κινούμαστε προς τα δεξιά.

3.2 Συνέχεια - Παραγωγισιμότητα

Έστω $f\colon I \to \mathbb{R}$ κυρτή και $c \in intI$. Έστω $[a, b] \subset I$ τέτοιο ώστε $a < c < b$. Από το θεώρημα 3.1.3 έχουμε

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

για $x \in (c, b]$. Επίσης από το θεώρημα 3.1.3 έχουμε ότι η συνάρτηση $x \mapsto \dfrac{f(x) - f(c)}{x - c}$ είναι αύξουσα στο $(c, b]$. Δηλαδή η δεξιά παράγωγος

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

υπάρχει. Με παρόμοιο τρόπο αποδεικνύουμε και ότι η αριστερή παράγωγος $f'_-(c)$ υπάρχει.

Αν $a < c < d < b$, τότε για επαρκώς μικρό $h > 0$ έχουμε

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

Παίρνοντας το όριο $\lim\limits_{h \to 0^+}$, έχουμε

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

Έτσι έχουμε αποδείξει το ακόλουθο θεώρημα.

Θεώρημα 3.2.1. *Έστω $f \colon I \to \mathbb{R}$ κυρτή. Τότε η f έχει δεξιά και αριστερή παράγωγο σε κάθε σημείο του $intI$ και οι f'_- και f'_+ είναι αύξουσες στο $intI$. Αν $c \in intI$, έχουμε*

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

και

$$f(x) \geqslant f(c) + f'_-(c)(x - c), \quad f(x) \geqslant f(c) + f'_+(c)(x - c)$$

για κάθε $x \in I$.

Θυμίζουμε ότι η $f \colon I \to \mathbb{R}$ ονομάζεται **Lipschitz** στο $I_0 \subset I$ αν υπάρχει $K > 0$ τέτοιο ώστε $|f(x) - f(y)| \leqslant K|x - y|$ για κάθε $x, y \in I_0$. Αυτή η συνθήκη συνεπάγεται και ότι η f είναι συνεχής και ομοιόμορφα συνεχής στο I_0.

Θεώρημα 3.2.2. *Έστω $f \colon I \to \mathbb{R}$ κυρτή και $[a, b] \subset intI$. Τότε*

(1) Η f είναι Lipschitz στο $[a, b]$.

(2) Η f είναι συνεχής στο $intI$.

Απόδειξη. Υπάρχουν $c, d \in I$ τέτοια ώστε $c < a < b < d$. Από το θεώρημα 3.2.1 έχουμε

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

για $a \leqslant x \leqslant y \leqslant b$. Επομένως $|f(x) - f(y)| \leqslant K|x - y|$ όπου $K := \max(|f'_+(a)|, |f'_-(b)|$. Αυτό αποδεικνύει το (1)· το (2) είναι άμεση συνέπεια του (1). □

Σημειώνουμε ότι η f δεν είναι απαραίτητα Lipschitz στο I ακόμα και αν η f είναι φραγμένη και ότι η f δεν είναι απαραίτητα συνεχής στο I ακόμα και αν το I είναι κλειστό και πεπερασμένο.

Μια συνάρτηση που είναι Lipschitz σε ένα διάστημα $[a, b]$ είναι απολύτως συνεχής στο $[a, b]$· είναι γνωστό ότι μια τέτοια συνάρτηση είναι παραγωγίσιμη σχεδόν παντού. Έπεται από το θεώρημα 3.2.2 ότι μια κυρτή συνάρτηση είναι παραγωγίσιμη σχεδόν παντού. Παρακάτω θα αποδείξουμε μια ισχυρότερη ιδιότητα των κυρτών συναρτήσεων χωρίς τη χρήση της συνθήκης της απόλυτης συνέχειας. Θυμίζουμε ότι μια συνάρτηση f λέγεται **δεξιά συνεχής** (αντ. **αριστερά συνεχής**) σε ένα σημείο x του πεδίου ορισμού της, αν ισχύει

$$f(x) = \lim_{z \to x^+} f(z) \quad (\text{αντ. } f(x) = \lim_{z \to x^-} f(z)).$$

Θεώρημα 3.2.3. *Έστω $f \colon I \to \mathbb{R}$ κυρτή. Τότε*

(1) Στο $intI$ η f'_- είναι αριστερά συνεχής και η f'_+ είναι δεξιά συνεχής.

(2) Υπάρχει μόνο αριθμήσιμο πλήθος σημείων που η f δεν είναι παραγωγίσιμη.

Απόδειξη. (1) Λόγω της συνέχειας της f στο $intI$ (θεώρημα 3.2.2) έχουμε για κάθε $x, y, z \in intI$ με $x < z < y$

$$\frac{f(y) - f(x)}{y - x} = \lim_{z \to x^+} \frac{f(y) - f(z)}{y - z} \geqslant \lim_{z \to x^+} f'_+(z).$$

Παίρνοντας το όριο $\lim\limits_{y \to x^+}$ έχουμε

$$f'_+(x) \geqslant \lim_{z \to x^+} f'_+(z).$$

Αφού η f'_+ είναι αύξουσα (θεώρημα 3.2.1) έχουμε

$$f'_+(x) \leqslant \lim_{z \to x^+} f'_+(z).$$

Συμπεραίνουμε ότι $f'_+(x) = \lim\limits_{z \to x^+} f'_+(z)$, το οποίο δείχνει τη δεξιά συνέχεια της f'_+. Η αριστερή συνέχεια της f'_- αποδεικνύεται παρόμοια.

(2) Από θεώρημα 3.2.1 για κάθε $x, y, z \in intI$ με $x < y < z$ έχουμε

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

Αν η f'_+ είναι συνεχής στο y έχουμε

$$f'_+(y) = \lim_{x \to y^-} f'_+(x) = \lim_{z \to x^+} f'_+(z) = f'_-(y)$$

το οποίο σημαίνει ότι η f είναι παραγωγίσιμη στο y. Επομένως τα σημεία του $intI$ όπου η f δεν είναι παραγωγίσιμη είναι αυτά που η αύξουσα συνάρτηση f'_+ έχει ασυνέχεια. Αυτό αποδεικνύει το (2) διότι γνωρίζουμε ότι μια αύξουσα συνάρτηση έχει μόνο αριθμήσιμο πλήθος ασυνεχειών. □

3.3 Κεντρική Κυρτότητα

Ορισμός 3.3.1. *Μια συνάρτηση* $f\colon I \to \mathbb{R}$ *ονομάζεται κεντρικά κυρτή αν για κάθε* $a, b \in I$ *ισχύει*

$$f\left(\frac{a+b}{2}\right) \leqslant \frac{1}{2}[f(a) - f(b)]. \tag{3.3}$$

Θεώρημα 3.3.2. *Έστω* $f\colon I \to \mathbb{R}$ *κεντρικά κυρτή και συνεχής. Τότε η* f *είναι κυρτή.*

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

$$f\left(\frac{a_1 + \ldots + a_N}{N}\right) \leqslant \frac{1}{N}\sum_{i=1}^{N} f(a_i). \tag{3.4}$$

Από την (3.3) έπεται ότι η (3.4) ισχύει για $n = 1$. Υποθέτουμε ότι ισχύει για $n = k$. Θα δείξουμε ότι ισχύει για $n = k+1$. Έστω $a_1, \ldots, a_{2^k}, a_{2^k}+1, \ldots, a_{2^{k+1}} \in I$. Τότε

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

Στη συνέχεια θα δείξουμε ότι αν η (3.4) ισχύει για $N = m \geqslant 2$ τότε ισχύει και για $N = m-1$. Πράγματι, αν $a_1, \ldots, a_{m-1} \in I$, ορίζουμε

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

και παρατηρούμε ότι

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

Τότε, εφαρμόζοντας την (3.4) για τα $a_1, \ldots, a_m$ έχουμε

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

απ'όπου παίρνουμε

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

Δηλαδή η (3.4) ισχύει για $N = m-1$. Συμπεραίνουμε έτσι ότι η (3.4) ισχύει για κάθε $N \in \mathbb{N}$.

Έστω τώρα $a, b \in I$ και $k, n \in \mathbb{N}$ με $k < n$. Από την (3.4) βλέπουμε ότι

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

δηλαδή για $\lambda \in \mathbb{Q}$, $0 < \lambda < 1$

$$f(\lambda a + (1-\lambda)b) \leqslant \lambda f(a) + (1-\lambda)f(b) \tag{3.5}$$

Λόγω της συνέχειας της f συμπεραίνουμε ότι η (3.5) ισχύει και αν $\lambda \in \mathbb{R}$, $0 < \lambda < 1$. Άρα η f κυρτή. □

3.4 Παραγωγίσιμες Κυρτές Συναρτήσεις

Θεώρημα 3.4.1. *Έστω I ανοιχτό διάστημα και $f\colon I \to \mathbb{R}$ δυο φορές παραγωγίσιμη συνάρτηση. Τότε η f είναι κυρτή αν και μόνο αν $f''(x) \geqslant 0$ για κάθε $x \in I$.*

Απόδειξη. Έστω ότι η f είναι κυρτή. Από το θεώρημα 3.2.1 η f' είναι αύξουσα, επομένως $f''(x) \geqslant 0$ για κάθε $x \in I$.

Αντίστροφα, έστω ότι $f'' \geqslant 0$ στο I. Αν $x < y \in I$ και $0 < \lambda < 1$, χρησιμοποιώντας το θεώρημα μέσης τιμής βρίσκουμε ξ_1, ξ_2 με $x < \xi_1 < \lambda x + (1-\lambda)y < \xi_2 < y$ και ξ_3 με $\xi_1 < \xi_3 < \xi_2$, τέτοια ώστε

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

Δηλαδή η f είναι κυρτή. □

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

Ανισότητες

Το θεώρημα 3.4.1 μας βοηθάει να δώσουμε πολλά απλά παραδείγματα κυρτών συναρτήσεων και μέσω αυτών των συναρτήσεων μπορούμε να αποδείξουμε ανισότητες που πολλές φορές δε φαίνονται και τόσο απλές με την πρώτη ματιά. Για παράδειγμα η ανισότητα

$$x^\lambda y^\mu \leqslant \lambda x + \mu y \tag{3.6}$$

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

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

Η (3.6) εμφανίζεται πολύ συχνά στις μορφές

$$x^{1/p} y^{1/q} \leqslant \frac{1}{p} x + \frac{1}{q} y \tag{3.7}$$

και

$$xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}$$

όπου $x, y > 0,\ p, q > 1$ και $1/p + 1/q = 1$ Στην περίπτωση $p = q = 2$ η (3.7) είναι η γνωστή ανισότητα $\sqrt{xy} \leqslant \frac{1}{2}(x + y)$.

3.5 Θεωρήματα για Ολοκληρώματα

Θεώρημα 3.5.1. *Έστω $f\colon (a, b) \to \mathbb{R}$. Τότε η f είναι κυρτή αν και μόνο αν αναπαρίσταται στη μορφή*

$$f(x) = f(c) + \int_c^x g(t)\, dt, \quad c, x \in (a, b) \tag{3.8}$$

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

Απόδειξη. Υποθέτουμε πρώτα ότι η f είναι κυρτή και θεωρούμε $c, x \in (a, b)$. Εξετάζουμε την περίπτωση $c < x$ (όμοια εργαζόμαστε και για την περίπτωση $x < c$. Από τα θεωρήματα 3.2.1 και 3.2.3 οι f'_+ και f'_- υπάρχουν και είναι αύξουσες και η f'_+ είναι δεξιά συνεχής ενώ η f'_+ είναι αριστερά συνεχής. Λόγω μονοτονίας είναι και οι δυο Riemann ολοκληρώσιμες.

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

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

για κάθε $k = 1, \ldots, m$.

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

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

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

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

Επομένως,

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

Για την αντίστροφη κατεύθυνση ας υποθέσουμε ότι η (3.8) ισχύει για κάποια αύξουσα (συνεχή από δεξιά) συνάρτηση g. Έστω $x, y \in (a, b)$ με $x < y$ και $\lambda \in (0, 1)$. Θέτουμε $z = \lambda x + (1 - \lambda)y$. Τότε

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

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

Έστω $f\colon [a, b] \to \mathbb{R}$ κυρτή και $a_1, \ldots, a_n \in [a, b]$. Τότε έχουμε

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

Η ανισότητα αυτή μας λέει ότι η τιμή της f στο μέσο όρο n αριθμών είναι μικρότερη ή ίση από το μέσο όρο των αντίστοιχων τιμών. Θα δείξουμε τώρα το ανάλογο θέωρημα για τη μέση τιμή μιας συνάρτησης.

Θεώρημα (ανισότητα του Jensen) 3.5.2. *Έστω $f\colon (a, b) \to \mathbb{R}$ κυρτή συνάρτηση και $g\colon [c, d] \to (a, b)$ συνεχής. Τότε*

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

Απόδειξη. Θέτουμε

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

Τότε $p \in (a,b)$. Από το θεώρημα 3.2.1 έχουμε

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

για κάθε $y \in (a,b)$ επομένως

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

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

3.6 Η Συζυγής Συνάρτηση

Θεώρημα 3.6.1. *Μια συνάρτηση $f\colon \mathbb{R} \to \mathbb{R}$ είναι κυρτή αν και μόνο αν υπάρχει συνάρτηση $g\colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ τέτοια ώστε*

$$f(x) = \sup_{y\in\mathbb{R}}[xy - g(y)]$$

για κάθε $x \in \mathbb{R}$.

Απόδειξη. ($\Leftarrow$) Έχουμε

$$f(x) = \sup_{g(y)<+\infty}[xy - g(y)]$$

Βλέπουμε ότι η f είναι το κατά σημείο supremum μιας συλλογής κυρτών συναρτήσεων και άρα είναι κυρτή.

($\Rightarrow$) Ορίζουμε συνάρτηση $g\colon \mathbb{R} \to \mathbb{R} \cap \{+\infty\}$ με

$$g(y) = \sup_{x\in\mathbb{R}}[xy - f(x)].$$

Έστω $x_0 \in \mathbb{R}$ Για κάθε $y \in \mathbb{R}$,

$$g(y) \geqslant x_0 y - f(x_0)$$

δηλαδή

$$x_0 y - g(y) \leqslant f(x_0).$$

Επομένως

$$\sup_{y\in\mathbb{R}}[x_0 y - g(y)] \leqslant f(x_0). \tag{3.9}$$

Θέτουμε $y_0 := f'_+(x_0)$. Από το θεώρημα 3.2.1 για κάθε $x \in \mathbb{R}$

$$f(x) \geqslant f(x_0) + f'_+(x_0)(x-x_0) = f(x_0) + y_0(x-x_0)$$

δηλαδή

$$xy_0 - f(x) \leqslant x_0y_0 - f(x_0).$$

Έπεται

$$g(y_0) = x_0y_0 - f(x_0)$$

κι έτσι

$$x_0y_0 - g(y_0) = f(x_0). \tag{3.10}$$

Συνδυάζοντας τις (3.9) και (3.10) παίρνουμε το ζητούμενο. □

Η συνάρτηση g όπως ορίστηκε στο προηγούμενο θεώρημα ονομάζεται **συζυγής** της f. Η f και η g σχηματίζουν ένα ζεύγος συναρτήσεων που ικανοποιεί την ανισότητα

$$f(x) + g(y) \leqslant xy$$

για κάθε $x, y \in \mathbb{R}$.

Κεφάλαιο 4

Κυρτές Συναρτήσεις σε χώρο Banach

Σε αυτό το κεφάλαιο θα ασχοληθούμε με κυρτές συναρτήσεις ορισμένες σε έναν χώρο Banach πάνω στο $\mathbb{R}$. Παρακάτω με X συμβολίζουμε έναν χώρο Banach με νόρμα.

4.1 Κάτω ημισυνέχεια

Ορισμός 4.1.1. *Έστω A ένα σύνολο και $f\colon A \to \mathbb{R}$. Το επιγράφημα $epi(f)$ της f είναι το σύνολο*

$$\{(x, \lambda) \in A \times \mathbb{R} : f(x) \leqslant \lambda\}.$$

Ορισμός 4.1.2. *Έστω $f\colon X \to \mathbb{R}$ και $x_0 \in Q$. Η f ονομάζεται κάτω ημισυνεχής στο x_0 αν για κάθε ακολουθία $\{x_n\}$ του X ισχύει*

$$x_n \to x_0 \Rightarrow \underline{\lim} f(x_n) \geqslant f(x_0).$$

Αν η f είναι κάτω ημισυνεχής για κάθε $x_0 \in X$ θα λέμε ότι η f είναι κάτω ημισυνεχής στο X.

Μια συνεχής συνάρτηση είναι κάτω ημισυνεχής.

Θεώρημα 4.1.3. *Έστω $f\colon X \to \mathbb{R}$. Οι επόμενες προτάσεις είναι ισοδύναμες:*

(α′) Η f είναι κάτω ημισυνεχής.

(β′) Το $\{x \in X : f(x) \leqslant \lambda\}$ είναι κλειστό για κάθε $\lambda \in \mathbb{R}$.

(γ′) Το $\{x \in X : f(x) > \lambda\}$ είναι ανοιχτό για κάθε $\lambda \in \mathbb{R}$.

(δ′) Το $epi(f)$ είναι κλειστό.

Απόδειξη. (α′) ⇔ (β′)Έστω ότι η f είναι κάτω ημισυνεχής. Υποθέτουμε ότι $x_n \in f^{-1}((-\infty, \lambda])$ και ότι $x_n \to x$. Θα δείξουμε ότι $f(x) \leqslant \lambda$, οπότε το $f^{-1}((-\infty, \lambda])$ είναι κλειστό. Έστω ότι $f(x) > \lambda$ και έστω $\lambda_1 \in \mathbb{R}$ με $f(x) > \lambda_1 > \lambda$. Επειδή η f είναι κάτω ημισυνεχής έχουμε $\underline{\lim} f(x_n) \geqslant f(x)$, άρα υπάρχει υπακολουθία $\{x_{k_n}\}$ της $\{x_n\}$ με $f(x_{k_n}) > \lambda_1$ για κάθε n, άτοπο γιατί υποθέσαμε ότι $f(x_n) \leqslant \lambda$ για κάθε n. Άρα για κάθε $\lambda \in \mathbb{R}$, το $f^{-1}((\infty, \lambda]) = \{x \in X : f(x) \leqslant \lambda\}$ είναι κλειστό.

Για το αντίστροφο υποθέτουμε ότι $f^{-1}((\infty, \lambda])$ είναι κλειστό για κάθε $\lambda \in \mathbb{R}$. Αν η f δεν είναι κάτω ημισυνεχής, υπάρχει $x \in X$ και ακολουθία $\{x_n\}$ του X ώστε $x_n \to x$ και $\underline{\lim} f(x_n) < f(x)$ και έστω $\underline{\lim} f(x_n) \leqslant \lambda_1 < \lambda_2 < f(x)$. Από τις υποθέσεις αυτές έχουμε ότι υπάρχει υπακολουθία $\{x_{k_n}\}$ της x_n ώστε $f(x_{k_n}) < \lambda_2 < f(x)$ για κάθε n. Έτσι έχουμε $x_{k_n} \in f^{-1}((-\infty, \lambda_2]), x_{k_n} \to x$ και $x \notin f^{-1}((-\infty, \lambda_2])$, άτοπο γιατί υποθέσαμε ότι $f^{-1}((-\infty, \lambda])$ είναι κλειστό για κάθε λ. Άρα η f είναι άνω ημισυνεχής στο X.

(β′) ⇔ (γ′) Το συμπλήρωμα ανοικτού συνόλου είναι κλειστό και αντίστροφα.

(α′) ⇔ (δ′) Ορίζουμε $F\colon X \times \mathbb{R} \to \mathbb{R}$ με $F(x, \lambda) = f(x) - \lambda$. Η F είναι κάτω ημισυνεχής αν και μόνο αν η f είναι κάτω ημισυνεχής. Από την (β′) η κάτω ημισυνέχεια της F σημαίνει ότι το $\{(x, \lambda) : F(x, \lambda) \leqslant \lambda_1\}$ για κάθε $\lambda_1 \in \mathbb{R}$ είναι κλειστό. Αυτό αποδεικνύει το ζητούμενο, καθώς $\{(x, \lambda) : F(x, \lambda) \leqslant \lambda_1\} = \{(x, \lambda) : (x, \lambda + \lambda_1) \in epi(f)\} = epi(f) - (0, \lambda_1)$. □

Ένας εναλλακτικός ορισμός για την κάτω ημισυνέχεια μιας $f\colon X \to \mathbb{R}$ σε ένα σημείο $\alpha \in X$ είναι για κάθε $K \in \mathbb{R}, K < f(\alpha)$ να υπάρχει περιοχή U του α ώστε $f(U) > K$.

4.2 Κυρτότητα

Αντίστοιχα με τον ορισμό που δώσαμε για τις κυρτές συναρτήσεις σε ένα διάστημα I έχουμε και τον ακόλουθο ορισμό για τις συναρτήσεις που είναι ορισμένες σε έναν χώρο Banach.

Ορισμός 4.2.1. *Μια συνάρτηση $f\colon X \to \mathbb{R}$ ονομάζεται κυρτή αν $\forall x, x' \in X$ και $\forall \lambda \in (0, 1)$*

$$f(\lambda x + (1 - \lambda)x') \leqslant \lambda f(x) + (1 - \lambda) f(x').$$

Θεώρημα 4.2.2. *Η συνάρτηση $f\colon X \to \mathbb{R}$ είναι κυρτή αν και μόνο το $epi(f)$ είναι κυρτό σύνολο.*

Απόδειξη. Έστω ότι η f είναι κυρτή και ότι $(x_1, y_1), (x_2, y_2) \in epi(f)$ όπου $x_1, x_2 \in X$ και $y_1, y_2 \in \mathbb{R}$. Για κάθε $\lambda \in (0, 1)$ έχουμε

$$\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) = (\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2) \in epi(f),$$

γιατί $y_1 \geqslant f(x_1), y_2 \geqslant f(x_2)$, επομένως

$$\lambda y_1 + (1 - \lambda)y_2 \geqslant \lambda f(x_1) + (1 - \lambda) f(x_2) \geqslant f(\lambda x_1 + (1 - \lambda)x_2),$$

αφού η f είναι κυρτή. Για το αντίστροφο υποθέτουμε ότι το $epi(f)$ είναι κυρτό σύνολο. Τότε για κάθε x_1, x_2 και $\lambda \in (0,1)$ έχουμε ότι $(x_1, f(x_1)), (x_2, f(x_2)) \in epi(f)$, επομένως

$$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) =$$
$$= (\lambda x_1 + (1-\lambda)x_2, \lambda f(x_1) + (1-\lambda)f(x_2)) \in epi(f),$$

άρα

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda)f(x_2)$$

και επομένως η f είναι κυρτή. □

Ορισμός 4.2.3. *Το ουσιώδες πεδίο ορισμού μιας κυρτής συνάρτησης $f\colon X \to \mathbb{R}$ είναι το σύνολο*

$$dom(f) := \{x \in X : f(x) < +\infty\}.$$

Το $dom(f)$ είναι κυρτό υποσύνολο του X γιατί είναι γραμμική εικόνα του κυρτού συνόλου $epi(f)$. Η διάσταση της f είναι η διάσταση του $dom(f)$.

Ορισμός 4.2.4. *Έστω $f\colon X \to \mathbb{R}$ και $x_0, x \in X$. Η παράγωγος κατά κατεύθυνση της f στο x_0 κατά την κατεύθυνση x ορίζεται*

$$f'(x_0; x) := \lim_{\varepsilon \to 0^+} \frac{f(x_0 + \varepsilon x) - f(x_0)}{\varepsilon}$$

αν υπάρχει το όριο.

4.3 Συνέχεια

Ορισμός 4.3.1. *Μια συνάρτηση $f\colon X \to \mathbb{R}$ ονομάζεται τοπικά άνω (αντ. κάτω) φραγμένη σε ένα σημείο $\alpha \in X$ αν υπάρχει περιοχή του α στην οποία η f είναι άνω (αντ. κάτω) φραγμένη .*

Θεώρημα 4.3.2. *Έστω $f\colon X \to \mathbb{R}$ κυρτή συνάρτηση τοπικά άνω φραγμένη στο $\alpha \in X$. Τότε:*

(α′) Η f είναι τοπικά άνω φραγμένη σε κάθε σημείο του $int(dom(f))$.

(β′) Η f είναι συνεχής στο $int(dom(f))$.

Απόδειξη. Η μελέτη της f είναι ισοδύναμη με τη μελέτη της $x \mapsto f(x+\alpha)$, επομένως μπορούμε χωρίς βλάβη της γενικότητας να υποθέσουμε ότι $\alpha = 0$. Έστω U μια περιοχή του 0 στην οποία η f είναι άνω φραγμένη, δηλαδή $f(x) \leqslant M < +\infty$ για $x \in U$.

(α′) Έστω $x_0 \in int(dom(f))$. Υπάρχει $\lambda > 1$ τέτοιο ώστε $\lambda x_0 \in int(dom(f))$. Το σύνολο

$$W := x_0 + \left(1 - \frac{1}{\lambda}\right) U$$

είναι μια περιοχή του x_0. Αν $y \in W$, έχουμε

$$y = x_0 + \left(1 - \frac{1}{\lambda}\right) u$$

για κάποιο $u \in U$, δηλαδή

$$\begin{aligned} f(y) &= f\left(\frac{1}{\lambda}(\lambda x_0) + \left(1 - \frac{1}{\lambda}\right) u\right) \\ &\leqslant \frac{1}{\lambda} f(\lambda x_0) + \left(1 - \frac{1}{\lambda}\right) f(u) \\ &\leqslant \frac{1}{\lambda} f(\lambda x_0) + \left(1 - \frac{1}{\lambda}\right) M \end{aligned}$$

και έπεται ότι η f είναι τοπικά άνω φραγμένη στο x_0.

(β′) Έστω $0 < \varepsilon < 1$. Το σύνολο $X' := \varepsilon(U \cap (-U))$ είναι περιοχή του 0. Για κάθε $x \in X'$ έχουμε $x/\varepsilon \in U$, δηλαδή $f(x/\varepsilon) \leqslant M$ κι έτσι

$$\begin{aligned} f(x) = f\left((1-\varepsilon)\cdot 0 + \varepsilon \cdot \frac{1}{\varepsilon} x\right) &\leqslant (1-\varepsilon) f(0) + \varepsilon f\left(\frac{1}{\varepsilon} x\right) \\ &\leqslant f(0) + \varepsilon[M - f(0)]. \end{aligned}$$

Έχουμε επίσης $-x/\varepsilon \in U$, δηλαδή

$$\begin{aligned} f(0) &= f\left(\frac{1}{1+\varepsilon} x + \frac{\varepsilon}{1+\varepsilon}\left(-\frac{1}{\varepsilon} x\right)\right) \\ &\leqslant \frac{1}{1+\varepsilon} f(x) + \frac{\varepsilon}{1+\varepsilon} f\left(-\frac{1}{\varepsilon} x\right) \\ &\leqslant \frac{1}{1+\varepsilon} f(x) + \frac{\varepsilon}{1+\varepsilon} M. \end{aligned}$$

Επομένως $|f(x) - f(0)| \leqslant \varepsilon(M - f(0))$. Συμπεραίνουμε ότι η f είναι συνεχής στο 0. Συνδυάζοντας αυτό το αποτέλεσμα με το (α′) έπεται το ζητούμενο. □

Έστω $A \subset X$. Μια συνάρτηση $f\colon X \to \mathbb{R}$ ονομάζεται **Lipschitz** στο A αν για κάθε $x, y \in A$ υπάρχει $K > 0$ τέτοιο ώστε $|f(x) - f(y)| \leqslant K\|x - y\|$.

Η $f\colon X \to \mathbb{R}$ ονομάζεται **τοπικά Lipschitz** στο $A \subset X$ αν για κάθε $\alpha \in A$ υπάρχει περιοχή U του α τέτοια ώστε η f είναι Lipschitz στο $U \cap A$.

Θεώρημα 4.3.3. *Έστω $f\colon X \to \mathbb{R}$ κυρτή συνάρτηση. Αν η f είναι τοπικά άνω φραγμένη σε κάποιο σημείο του X, τότε η f είναι τοπικά Lipschitz στο $int(dom(f))$*

Απόδειξη. Έστω $\alpha \in int(dom(f))$. Από το 4.3.2 προκύπτει ότι η f είναι συνεχής στο α. Δηλαδή υπάρχουν $r_0 > 0$ και $m, M \in \mathbb{R}$ τέτοια ώστε $m \leqslant f(x) \leqslant M$ για κάθε σημείο x της κλειστής μπάλας $B(\alpha, r_0)$. Έστω $0 < r < r_0$ και $x, y \in B(\alpha, r)$. Θέτοντας $\|x - y\| = \sigma$ και

$z = y + [(r_0 - r)/\sigma](y - x)$, έχουμε $y = \lambda z + (1 - \lambda)x$ όπου $\lambda = \sigma/(\sigma + r_0 - r)$. Αφού $\|z - \alpha\| \leqslant \|y - \alpha\| + r_0 - r \leqslant r_0$, έχουμε $z \in B(\alpha, r_0)$. Επομένως $f(y) \leqslant \lambda f(z) + (1-\lambda)f(x)$ δηλαδή

$$f(y) - f(x) \leqslant \lambda(f(z) - f(x)) \leqslant \lambda(M - m) \leqslant \frac{M - m}{r_0 - r}\|x - y\|$$

το οποίο μπορεί να γραφεί

$$|f(x) - f(y)| \leqslant \frac{M - m}{r_0 - r}\|x - y\|$$

για κάθε $x, y \in B(\alpha, r)$, το οποίο ολοκληρώνει την απόδειξη. □

4.4 Παραγωγίσιμες Κυρτές Συναρτήσεις

Έστω X' ο δυϊκός του X. Έστω $f\colon X \to \mathbb{R}$ και x_0 σημείο του X όπου η f είναι πεπερασμένη. Τότε έχουμε τους παρακάτω δυο ορισμούς.

Ορισμός 4.4.1. *Η f ονομάζεται Frechet-παραγωγίσιμη (ή απλά παραγωγίσιμη) στο x_0 αν υπάρχει $x' \in X'$ τέτοιο ώστε*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - \langle x - x_0, x' \rangle}{\|x - x_0\|} = 0$$

όπου το x' είναι μοναδικό. Ονομάζεται Frechet παράγωγος (ή απλά παράγωγος) της f στο x_0 και συμβολίζεται με $f'(x_0)$ ή $df(x_0)$.

Ορισμός 4.4.2. *Η f ονομάζεται Gateaux-παραγωγίσιμη στο x_0 αν υπάρχει $x' \in X'$ τέτοιο ώστε για κάθε $x \in X$*

$$\lim_{x \to 0} \frac{f(x_0 + \varepsilon x) - f(x_0)}{\varepsilon} = \langle x, x' \rangle$$

όπου το x' είναι μοναδικό. Ονομάζεται Gateaux-διαφορικό της f στο x_0. Το συμβολίζουμε με $\nabla f(x_0)$.

Η Frechet-διαφορισιμότητα συνεπάγεται Gateaux-διαφορισιμότητα, αλλά το αντίστροφο δεν ισχύει. Ισχύει όμως για κυρτές συναρτήσεις στον $\mathbb{R}^n$.

Θεώρημα 4.4.3. *Έστω $f\colon \mathbb{R}^n \leftarrow \mathbb{R}$ δυο φορές συνεχώς παραγωγίσιμη. Ορίζουμε τον Hessian πίνακα $H(x)$ της f στο x ως εξής:*

$$H(x) := \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2(x)} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(x) \end{bmatrix}$$

όπου $x \in \mathbb{R}^n$. Τότε

$$f \text{ κυρτή} \Leftrightarrow H(x) \text{ είναι θετικά ημιορισμένος για κάθε } x \in \mathbb{R}^n.$$

Απόδειξη. Από τον ορισμό της κυρτότητας έχουμε, f κυρτή $\Leftrightarrow$ για κάθε $x, y \in \mathbb{R}^n$ η συνάρτηση $g\colon t \mapsto f(x+ty)$ από το $\mathbb{R}$ στο $\mathbb{R}$ είναι κυρτή. Από το θεώρημα 3.4.1 η g είναι κυρτή αν και μόνο αν $g''(t) \geqslant 0$ για κάθε $t \in \mathbb{R}$. Έχουμε

$$\begin{aligned} g'(t) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+ty)y_i \\ g''(t) &= \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x+ty)y_i y_j = (H(x+ty)y|y). \end{aligned}$$

Επομένως

$$f \text{ κυρτή } \Leftrightarrow (H(x+ty)y|y) \geqslant 0 \text{ όταν } x, y \in \mathbb{R}^n, t \in \mathbb{R}$$

το οποίο αποδεικνύει το ζητούμενο. $\square$

Βιβλιογραφία

[1] Jean-Baptiste Hiriart-Urruty, Claude Lemaréchal, *Fundamentals of Convex Analysis*

[2] Jürg T. Marti, *Konvexe Analysis*

[3] Haïm Brezis, *Συναρτησιακή Ανάλυση*

[4] Ιωάννης Πολυράκης, *Θέματα Ανάλυσης και Θεωρία Γενικής Ισορροπίας στην Οικονομία*

[5] Σπύρος Αργυρός, *Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης (Δεύτερη Έκδοση)*

[6] Jan van Tiel, *Convex Analysis, An Introductory Text*