Τα θεωρήματα των Radon, Καραθεοδωρή και Helly

Λεώνη Ευαγγελάτου-Δάλλα
Τμήμα Μαθηματικών, ΕΚΠΑ
ldalla@math.uoa.gr

Δήμητρα-Διονυσία Στεργιοπούλου[1]
Τμήμα Μαθηματικών, ΕΚΠΑ

Περίληψη

Τα θεωρήματα των Radon, Καραθεοδωρή, Helly περικλείουν την αρχική ιδέα για τον κλάδο της «Συνδυαστικής Γεωμετρίας» και οι αποδείξεις τους είναι μεν απλούστατες, αλλά ευρηματικές. Στο σύντομο αυτό άρθρο θα παρουσιαστούν τα θεωρήματα και οι αποδείξεις τους με τη χρήση βασικών εργαλείων.

Abstract

The theorems of Radon, Carathéodory, Helly include the initial idea about the branch of Mathematics called "Combinatorial Geometry". Their proofs are very simple, but ingenious. In this short paper these theorems and their proofs will be quoted using simple facts and instruments.

[1] Παρουσίασε το άρθρο στα πλαίσια του προπτυχιακού μαθήματος «Κυρτή Ανάλυση», Τμήμα Μαθηματικών, ΕΚΠΑ (2010-11) (http://eclass.uoa.gr/courses/MATH140/).

Εισαγωγή

Είναι συνήθως δύσκολο να εντοπιστεί η αρχή ενός κλάδου των Μαθηματικών, να καταγραφούν οι εμπνευστές και οι πρωτοπόροι των ιδεών, καθώς και το μαθηματικό περιβάλλον που τους οδήγησε στη μελέτη συγκεκριμένων θεμάτων.

Το πρακτικό εκ πρώτης όψεως αποτέλεσμα του Kirchberger το 1903 θεωρείται ότι είναι η απαρχή της Συνδυαστικής Γεωμετρίας, ενός κλάδου που απασχόλησε τους ερευνητές τόσο από θεωρητικής άποψης, όσο και από πλευράς εφαρμογών.

Στο μακρινό 1903 ο Kirchberger θεώρησε ένα κοπάδι άσπρων και μαύρων προβάτων. Απέδειξε ότι αν ανά τέσσερα μπορούν να χωριστούν σε άσπρα-μαύρα με έναν ευθύ φράκτη, τότε όλο το κοπάδι χωρίζεται με φράκτη που θα έχει από τη μια μεριά τα άσπρα και από την άλλη τα μαύρα. Το αποτέλεσμα αυτό, που μοιάζει με παιγνίδι, εντάσσεται σε αυτό που σήμερα ονομάζεται «Συνδυαστική Γεωμετρία».

Η συνέχεια ήταν εντυπωσιακή, μιας και τρεις πρωτοπόροι των αρχών του 20ου αιώνα έδωσαν τα επώνυμά τους σε θεωρήματα: Καραθεοδωρή, Radon και Helly.

Αν και μια πρώτη προσέγγιση δεν επιτρέπει να φανερωθούν οι δεσμοί που τα ενώνουν, εν τούτοις είναι αλληλένδετα και αποτελούν τα θεμέλια της «Συνδυαστικής Γεωμετρίας». Αποτελέσματα τύπου Helly έγιναν αντικείμενο έρευνας στη διάρκεια του 20ού αιώνα και κυρίως στις δεκαετίες ’60 και ’70. Θα γίνει προσπάθεια να παρουσιαστούν οι αποδείξεις όπως έχουν καταγγραφεί στην πλούσια βιβλιογραφία με κύριο σκοπό τον εντοπισμό της κεντρικής ιδέας από όπου προκύπτουν.

1 Τα θεωρήματα Radon-Καραθεοδωρή-Helly

Θα μπορούσαμε να χωρίσουμε σε γενικές γραμμές τις αποδείξεις των περίφημων αυτών θεωρημάτων σε δύο κατηγορίες: η πρώτη βασίζεται στη δομή του $\mathbb{R}^d$ ως γραμμικού χώρου εφοδιασμένου με την Ευκλείδεια απόσταση και η δεύτερη χρησιμοποιεί τη γεωμετρία του $\mathbb{R}^d$, ως γραμμικού χώρου διάστασης d.

Η αρχική απόδειξη του Καραθεοδωρή (1914) για το θεώρημά του ανήκει στην πρώτη κατηγορία, βασιζόμενη στην ύπαρξη υπερεπιπέδου στήριξης στα συνοριακά σημεία κυρτού συνόλου. Το θεώρημα του Helly διατυπώθηκε από τον ίδιο το 1913, αλλά δημοσιεύθηκε το 1923, μετά την επιστροφή του από τη Σιβηρία, όπου εκρατείτο ως αιχμάλωτος πολέμου. Η απόδειξη ανήκει και αυτή στην πρώτη κατηγορία, χρησιμοποιώντας τον διαχωρισμό με υπερεπίπεδο δύο κυρτών ξένων συνόλων.

Κατ' ουσία οι αρχικές αποδείξεις των Καραθεοδωρή [2] και Helly [3] βασίζονται στο ίδιο εργαλείο της Κυρτής Ανάλυσης: εάν K είναι κυρτό, κλειστό σύνολο του $\mathbb{R}^d$, τότε για κάθε σημείο $x \in \mathbb{R}^d$ υπάρχει μοναδικό σημείο του K που αποτελεί τη βέλτιστη προσέγγιση του x, από όλα τα σημεία του K.

Στο διάστημα της εξορίας του Helly, ο Radon (1921) [4] είχε δημοσιεύσει το θεώρημα του Helly, έχοντας πληροφορίες από τον ίδιο τον Helly το 1913. Όμως η απόδειξη του Radon ανήκει στη δεύτερη κατηγορία που αναφέραμε.

Επειδή οι μαθηματικοί αρέσκονται στην λιτότητα και την αφαίρεση, θα παρουσιάσουμε τις αποδείξεις των θεωρημάτων σύμφωνα με την βασική ιδέα του Radon.

2 Το βασικό εργαλείο

Όλοι γνωρίζουμε τον χώρο $\mathbb{R}^d$ ($d \geq 1$) ως γραμμικό χώρο διάστασης d. Για $d = 1$ ο $\mathbb{R}$ είναι το σύνολο των πραγματικών αριθμών, θεμελιωμένο με τα αξιώματα που αναφέρονται στην πρόσθεση, στον πολλαπλασιασμό, στη διάταξη και την πληρότητα. Για $d \geq 2$, ο

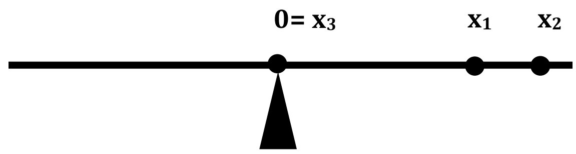
$$\mathbb{R}^d = \{(x_1, \ldots, x_d) : \ x_i \in \mathbb{R}, i = 1, \ldots, d\} ,$$

εφοδιασμένος με την πρόσθεση και το βαθμωτό πολλαπλασιασμό γίνεται γραμμικός χώρος. Ο $\mathbb{R}^d$ μπορεί να παραχθεί από σύνολο d γραμμικώς ανεξαρτήτων στοιχείων και σε οιονδήποτε σύνολο $\nu \geq d + 1$ στοιχείων, τα στοιχεία είναι γραμμικώς εξαρτημένα. Αναλυτικά, υπάρχει σύνολο $\{e_1, e_2, \ldots, e_d\}$ με την ιδιότητα ότι η σχέση $\mu_1 e_1 + \cdots + \mu_d e_d = 0$ για κάποια $\mu_1, \ldots, \mu_d \in \mathbb{R}$ ισχύει μόνο για $\mu_1 = \cdots = \mu_d = 0$ και αν $\{x_1, x_2, \ldots, x_n\} \subseteq \mathbb{R}^d$ με $n \geq d+1$, τότε υπάρχουν $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, όχι όλα μηδέν, ώστε $\lambda_1 x_1 + \cdots \lambda_n x_n = 0$. Λόγω της ιδιότητας αυτής του $\mathbb{R}^d$ ορίζουμε τη διάστασή του να είναι ίση με d.

Ο Radon τροποποιώντας εμπνευσμένα αυτήν και μόνο την ιδιότητα του $\mathbb{R}^d$ απέδειξε το θεώρημά του και στη συνέχεια, βάσει του ίδιου επιχειρήματος, αποδεικνύονται τα θεωρήματα των Καραθεοδωρή και Helly.

Ας δώσουμε μια δική μας προσέγγιση για το σκεπτικό που πιθανόν τον οδήγησε στο βασικό εργαλείο των αποδείξεων του μέσω παραδειγμάτων.

Ας βρεθούμε κατ' αρχάς στον $\mathbb{R}$ όπου $d = 1$.



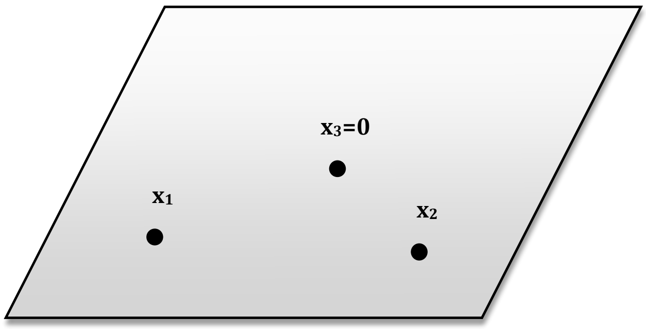
Σχήμα 2.1: Υπάρχουν μ_1, μ_2, μ_3 όχι όλα μηδέν με $\mu_1 + \mu_2 + \mu_3 = 0$ και $\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 = 0$.

Εάν σε ένα μόνο σημείο $x_1 \neq 0$ θελήσουμε να βάλουμε κάποιο βάρος λ_1, αποκλείεται να έχουμε ισορροπία ως προς το 0, εκτός αν το $\lambda_1 = 0$. Εάν πάρουμε σύνολο $A = \{x_1, x_2\}$ σε τυχαία θέση, επειδή ένα από τα δύο μπορεί να είναι το 0 θα πάρουμε και πάλι ανισορροπία ως προς το 0, αν προσπαθήσουμε να βάλουμε μη μηδενικό βάρος στο άλλο.

Ας πάρουμε $A = \{x_1, x_2, x_3\}$. Τότε για τα $x_1 - x_3$, $x_2 - x_3$ υπάρχουν λ_1, λ_2 διάφορα του μηδενός, ώστε $\lambda_1(x_1 - x_3) + \lambda_2(x_2 - x_3) = 0$ ή $\lambda_1 x_1 + \lambda_2 x_2 - (\lambda_1 + \lambda_2)x_3 = 0$, ή ισοδύναμα, υπάρχουν μ_1, μ_2, μ_3 όχι όλα μηδέν, με $\mu_1 + \mu_2 + \mu_3 = 0$ ώστε $\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 = 0$. Ακόμα και εάν ένα από τα x_1, x_2, x_3 είναι 0 (Σχ. 2.1) τα υπόλοιπα δύο μας είναι αρκετά για να δώσουν σίγουρη ισορροπία. Οπότε χρειαζόμαστε σύνολο με *τουλάχιστον* $d+2 = 3$ σημεία ώστε να υπάρχει ισορροπία ως προς το 0.

Τι γίνεται στον $\mathbb{R}^2$;

Επειδή το σύνολο A των σημείων μπορεί να είναι σε ευθεία ($\cong \mathbb{R}$) θα πρέπει να αρχίσουμε με σύνολο τριών σημείων. Όμως ένα από αυτά μπορεί να είναι το 0, άρα, αν τα x_1, x_2 δεν βρεθούν σε ευθεία με το 0, δεν θα έχουμε ισορροπία (Σχ. 2.2).



Σχήμα 2.2: Δεν υπάρχουν μ_1, μ_2, μ_3, όχι όλα μηδεν με $\mu_1 + \mu_2 + \mu_3 = 0$ και $\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 = 0$.

Ας πάρουμε $A = \{x_1, x_2, x_3, x_4\} \subseteq \mathbb{R}^2$. Το $B = \{x_1 - x_4, x_2 - x_4, x_3 - x_4\}$ είναι γραμμικά εξαρτημένο σύνολο μη μηδενικών σημείων, άρα υπάρχουν $\lambda_1, \lambda_2, \lambda_3$, όχι όλα μηδέν, ώστε $\lambda_1(x_1 - x_4) + \lambda_2(x_2 - x_4) + \lambda_3(x_3 - x_4) = 0$ ή υπάρχουν $\lambda_1, \lambda_2, \lambda_3$, όχι όλα μηδέν, ώστε $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + (-\lambda_1 - \lambda_2 - \lambda_3)x_4 = 0$ ή υπάρχουν $\mu_1, \mu_2, \mu_3, \mu_4$, όχι όλα μηδέν, με $\mu_1 + \mu_2 + \mu_3 + \mu_4 = 0$ και $\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 + \mu_4 x_4 = 0$. Δηλαδή, χρειαζόμαστε σύνολο με *τουλάχιστον* $d + 2 = 4$ σημεία για να πετύχουμε ισορροπία.

Τι συμβαίνει με τον αριθμό $n \geq d + 2$;

Ας πάρουμε τον $\mathbb{R}^d$ και $A = \{x_1, \ldots, x_n\}$ με $n \geq d + 2$. Τότε το $B = \{x_1 - x_n, x_2 - x_n, \ldots, x_{n-1} - x_n\}$ έχει πλήθος στοιχείων $n - 1 \geq d + 1$. Επειδή η διάσταση του $\mathbb{R}^d$ είναι d, το B είναι γραμμικά εξαρτημένο σύνολο, άρα υπάρχουν $\lambda_1, \ldots, \lambda_{n-1} \in \mathbb{R}$, όχι όλα μηδέν, ώστε $\lambda_1(x_1 - x_n) + \cdots + \lambda_{n-1}(x_{n-1} - x_n) = 0$ ή υπάρχουν $\mu_1, \ldots, \mu_{n-1}, \mu_n \in \mathbb{R}$, όχι όλα μηδέν, με $\mu_1 + \cdots + \mu_{n-1} + \mu_n = 0$ και $\mu_1 x_1 + \cdots + \mu_{n-1} x_{n-1} + \mu_n x_n = 0$.

Άρα καταλαβαίνουμε ότι δεν μπορούμε να επιτύχουμε το σκοπό μας με σύνολα με πλήθος στοιχείων μικρότερο του $d + 2$.

Έχουμε έτοιμο το βασικό εργαλείο του Radon:

Πρόταση R. Έστω $A = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^d$ και $B = \{x_1 - x_n, \dots, x_{n-1} - x_n\}$. Το B είναι γραμμικά εξαρτημένο, αν και μόνο αν υπάρχουν $\mu_1, \dots, \mu_n$ όχι όλα μηδέν με $\mu_1 + \cdots + \mu_n = 0$ και $\mu_1 x_1 + \cdots + \mu_n x_n = 0$.
Ιδιαιτέρως, αν $n \geq d + 2$, το B είναι γραμμικά εξαρτημένο.

Για να διατυπώσουμε τα τρία θεωρήματά μας χρειαζόμαστε δύο γνώριμους ορισμούς.

Ορισμός 2.1. Το $x \in \mathbb{R}^d$ είναι *κυρτός συνδυασμός* των σημείων $x_1, \dots, x_\nu$ του $\mathbb{R}^d$, αν υπάρχουν $\lambda_1, \dots, \lambda_\nu \geq 0$ με $\lambda_1 + \cdots \lambda_\nu = 1$ ώστε $x = \lambda_1 x_1 + \cdots \lambda_\nu x_\nu$. Εάν A ένα μη κενό υποσύνολο του $\mathbb{R}^d$, η *κυρτή θήκη*

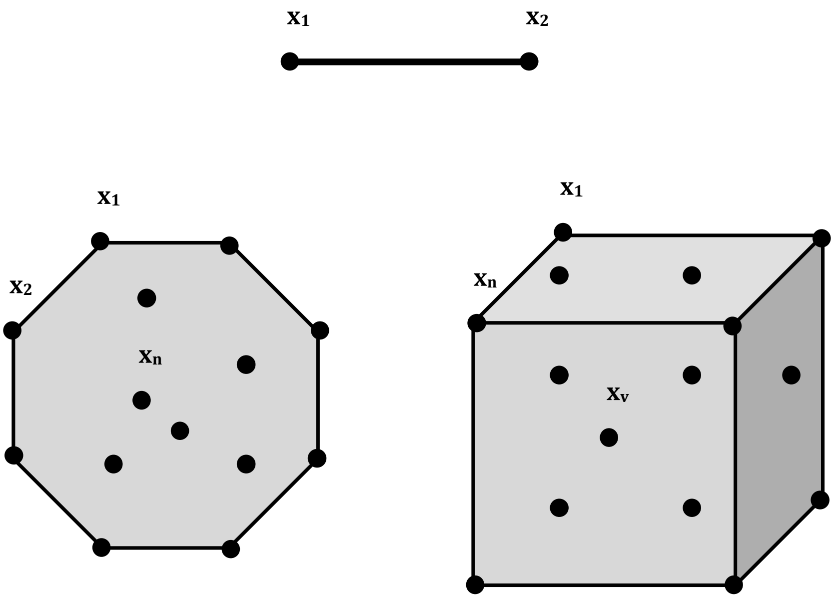
$$\begin{aligned}\operatorname{conv}(A) &= \Big\{x \in \mathbb{R}^d : \ x \text{ κυρτός συνδυασμός σημείων του } A\Big\} \\ &= \Bigg\{x \in \mathbb{R}^d : \ \text{υπάρχουν } x_1, \dots, x_\nu \in A \text{ και } \lambda_1, \dots, \lambda_\nu \in \mathbb{R} \\ &\qquad\qquad \text{με } \sum_{i=1}^{\nu} \lambda_i = 1 \text{ ώστε } x = \sum_{i=1}^{\nu} \lambda_i x_i\Bigg\}\end{aligned}$$

Καταλαβαίνουμε ότι ακόμα και για πεπερασμένο σύνολο A, με μεγάλο πλήθος σημείων, ο εντοπισμός όλων αυτών των κέντρων ισορροπίας είναι ανέφικτος (Σχ. 2.3). Επιπλέον, για τυχαίο A δεν δίνει πληροφορίες εάν η $\operatorname{conv}(A)$ κληρονομεί ιδιότητες του A (π.χ. συμπάγεια κ.ά.).

Το θεώρημα του Καραθεοδωρή ελαττώνει το πλήθος των συνδυασμών με βέλτιστο τρόπο και αυτό έχει ευεργετικά αποτελέσματα.

Ορισμός 2.2. Έστω K μη κενό υποσύνολο του $\mathbb{R}^d$. Το K είναι *κυρτό*, αν για οιανδήποτε $x, y \in K$ το ευθύγραμμο τμήμα $[x, y] = \{(1 - \lambda)x + \lambda y : \ \lambda \in [0, 1]\} \subseteq K$.

Αποδεικνύεται ότι η κυρτή θήκη ενός συνόλου A είναι κυρτό σύνολο και μάλιστα το μικρότερο κυρτό σύνολο που περιέχει το A. Ως συμπέρασμα έχουμε ότι το K είναι κυρτό, αν και μόνο αν $K = \operatorname{conv}(K)$.



Σχήμα 2.3: Κυρτή θήκη πεπερασμένου πλήθους σημείων.

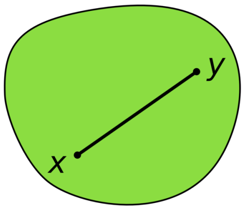
Δεν θα αναφέρουμε τη χρησιμότητα των κυρτών συνόλων στον $\mathbb{R}^d$ όσο ή σε απειροδιάστατους γραμμικούς τοπολογικούς χώρους, ας θυμηθούμε μόνο τον ρόλο τους στην Κυρτή Βελτιστοποίηση, στη Θεωρία Κυρτών Συναρτήσεων, στη Θεωρία Παιγνίων αλλά και στην Ευκλείδεια Γεωμετρία.

Έχουμε πλέον τα εφόδια να διατυπώσουμε και να αποδείξουμε τα τρία θεωρήματα, βασιζόμενοι στην Πρόταση R και χρησιμοποιώντας μαθηματική επαγωγή και απλή λογική.

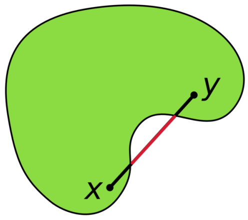
Θεώρημα 2.1 (Radon). *Έστω* $A = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$ *με πλήθος σημείων* $n \geq d + 2$. *Τότε υπάρχουν μη κενά υποσύνολα* B, C *του* A *ώστε* $B \cap C = \varnothing$, $A = B \cup C$ *και* $\mathrm{conv}\,(B) \cap \mathrm{conv}\,(C) = \varnothing$.

Απόδειξη Επειδή $n \geq d + 2$ από την Πρόταση R έχουμε ότι υπάρχουν $\mu_1, \ldots, \mu_n$, όχι όλα μηδέν, με

$$\mu_1 + \cdots \mu_n = 0 \quad \text{και} \quad \mu_1 x_1 + \cdots + \mu_n x_n = 0\,. \tag{2.1}$$



(a) κυρτό



(b) μη κυρτό

Σχήμα 2.4: Ένα κυρτό (α) και ένα μη κυρτό (β) σύνολο στον $\mathbb{R}^2$.

Εφ' όσον τα $\mu_1, \ldots, \mu_n$ δεν είναι όλα μηδεν, αλλά αθροίζουν στο μηδέν, τα $I = \{i \in 1, \ldots, n : \ \mu_i > 0\}$ και $J = \{j \in 1, \ldots, n : \ \mu_j \leq 0\}$ είναι μη κενά, $I \cap J = \varnothing$ και $I \cup J = \{1, \ldots, n\}$.

Ορίζουμε $B = \{x_i \in A : \quad i \in I\}$ και $C = \{x_j \in A : \quad j \in J\}$. Τα B, C είναι μη κενά, $B \cap C = \varnothing$ και $B \cup C = A$. Έχουμε $\sum_{i \in I} \mu_i x_i + \sum_{j \in J} \mu_j x_j = 0$, και $\mu \equiv \sum_{i \in I} \mu_i = + \sum_{j \in J} (-\mu_j) > 0$ (από την (2.1) και τον ορισμό των I, J).

Θεωρούμε το:

$$x = \sum_{i \in I} \frac{\mu_i}{\mu} x_i = \sum_{j \in J} \frac{(-\mu_j)}{\mu} x_j \ .$$

Όμως $\mu_i > 0$, $i \in I$, με $\sum_{i \in I} \frac{\mu_i}{\mu} = 1$ και $-\mu_j \geq 0$, $j \in J$ με $\sum_{j \in J} \frac{(-\mu_j)}{\mu} = 1$.

Από τον ορισμό των conv (B), conv (C) συμπεραίνουμε ότι το $x \in \operatorname{conv}(B) \cap \operatorname{conv}(C)$.

ο.ε.δ.

Θεώρημα 2.2 (Καραθεοδωρή). *Έστω A μη κενό υποσύνολο του $\mathbb{R}^d$. Τότε:*

$$\operatorname{conv}(A) = \left\{ x \in \mathbb{R}^d : \ \textit{υπάρχουν } x_1, \ldots, x_{d+1} \in A \textit{ και } \lambda_i \geq 0 \right.$$
$$\left. \textit{με } \sum_{i=1}^{d+1} \lambda_i = 1 \textit{ ώστε } x = \sum_{i=1}^{d+1} \lambda_i x_i \right\}$$

Απόδειξη

Έστω $x \in \operatorname{conv}(A)$. Τότε υπάρχουν $x_1, \ldots, x_n \in A$ και $\lambda_i > 0$ με $\sum_{i=1}^{n} \lambda_i = 1$ ώστε $x = \sum_{i=1}^{n} \lambda_i x_i$. Υποθέτουμε ότι το n είναι το ελάχιστο πλήθος στοιχείων του A που ο κυρτός συνδυασμός τους δίνει το x, δηλαδή αν $x = \sum_{i=1}^{k} \xi_i y_i$ για $\xi_i > 0$ με $\sum_{i=1}^{k} \xi_i = 1$ και $y_1, \ldots, y_k \in A$ τότε

$$k \geq n \, . \tag{2.2}$$

Έστω ότι $n > d+1$, δηλαδή $n \geq d+2$. Από την Πρόταση R έχουμε ότι υπάρχουν $\mu_1, \ldots, \mu_n \in \mathbb{R}$, όχι όλα μηδέν, με

$$\mu_1 + \cdots + \mu_n = 0 \quad \text{και} \quad \mu_1 x_1 + \cdots + \mu_n x_n = 0 \, . \tag{2.3}$$

Εφ' όσον τα $\mu_1, \ldots, \mu_n$ δεν είναι όλα μηδέν, αλλά αθροίζουν στο μηδέν, το $I = \{i \in 1, \ldots, n : \ \mu_i > 0\}$ και το $J = \{j \in 1, \ldots, n : \ \mu_j \leq 0\}$ είναι μη κενά, $I \cap J = \varnothing$ και $I \cup J = \{1, \ldots, n\}$. Έστω k ώστε $\lambda_k/\mu_k = \min\{\lambda_i/\mu_i : \ i \in I\} > 0$. Για $i \in I$ θα έχουμε $\lambda_i - (\lambda_k/\mu_k)\lambda_i \geq 0$ και $\lambda_k - (\lambda_k/\mu_k)\mu_k = 0$. Για $j \in J$ θα έχουμε $\lambda_j - (\lambda_k/\mu_k)\mu_j \geq 0$. Οπότε $\lambda_i - (\lambda_k/\mu_k)\mu_i \geq 0$ για κάθε $i \in \{1, \ldots, n\}$, $\lambda_k - (\lambda_k/\mu_k)\mu_k = 0$ και $\sum_{i=1}^{n} [\lambda_i - (\lambda_k/\mu_k)\mu_i] = \sum_{i=1}^{n} \lambda_i - (\lambda_k/\mu_k) \sum_{i=1}^{n} \mu_i = 1$ (λόγω της (2.3)).

Όμως:

$$\begin{aligned}
x &= \sum_{i=1}^{n} \lambda_i x_i \\
&= \sum_{i=1}^{n} \lambda_i x_i - \frac{\lambda_k}{\mu_k}\left(\sum_{i=1}^{n} \mu_i x_i\right) \qquad \text{(από (2.3))} \\
&= \sum_{i=1}^{n} \left(\lambda_i - \frac{\lambda_k}{\mu_k}\mu_i\right) x_i \\
&= \sum_{\substack{i=1 \\ i \neq k}}^{n} \left(\lambda_i - \frac{\lambda_k}{\mu_k}\mu_i\right) x_i \, ,
\end{aligned}$$

δηλαδή το x είναι κυρτός συνδυασμός το πολύ $n-1$ στοιχείων του A, το οποίο είναι άτοπο σύμφωνα με την υπόθεση (2.2) για το n. Άρα $n \leq d+1$ και έχουμε το ζητούμενο.

ο.ε.δ.

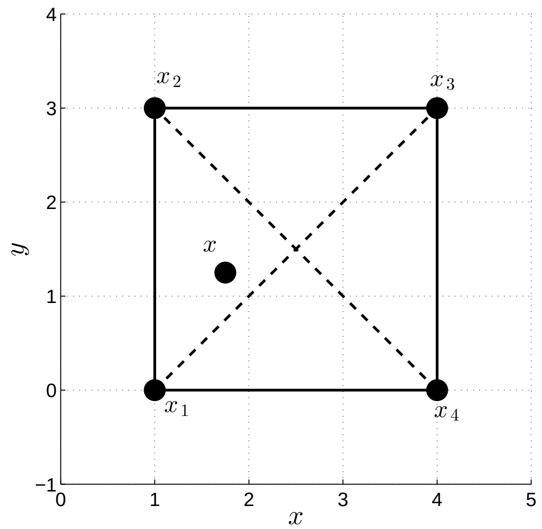
Τα βήματα της απόδειξης είναι πειστικά, αναλυτικά, άψογα, όμως ο φορμαλισμός τους δε μας αφήνει περιθώρια να διακρίνουμε το νόημα, ώστε να το μεταδώσουμε στους αμύητους.

Μήπως η απόδειξη του Θεωρήματος μπορεί να αναδείξει τη Φυσική ή Γεωμετρική πλευρά του;

Ας πάρουμε ένα απλό παράδειγμα στον $\mathbb{R}^2$.

Έστω $A = \{x_1 = (1,0), x_2 = (1,3), x_3 = (4,3), x_4 = (4,0)\}$ και $x = \frac{1}{2}x_1 + \frac{1}{4}x_2 + \frac{1}{6}x_3 + \frac{1}{12}x_4 = (\frac{7}{4}, \frac{5}{4}) \in \operatorname{conv}(A)$ με $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{1}{4}$, $\lambda_3 = \frac{1}{6}$, $\lambda_4 = \frac{1}{12}$.

Ο Gauss σε χειρόγραφες σημειώσεις έδωσε την εξής περιγραφή της $\operatorname{conv}(A)$ από την πλευρά της Φυσικής: Υποθέστε ότι στο σημείο $x \in \mathbb{R}^d$ έχουμε θετικό βάρος. Εάν υπάρχουν μη αρνητικά βάρη στα $x_1, \ldots, x_n$ ώστε το x να είναι το κέντρο βάρους αυτών, τότε $x \in \operatorname{conv}\{x_1, \ldots, x_n\}$ και αντίστροφα.



Σχήμα 2.5: Το x ως κέντρο βάρους των x_1, x_2, x_3, x_4 με βάρη $\lambda_1, \lambda_2, \lambda_3, \lambda_4$.

Αν δούμε το $x = (\frac{7}{4}, \frac{5}{4})$ με τη ματιά του Gauss. Το $x = (\frac{7}{4}, \frac{5}{4})$ είναι το κέντρο βάρους των x_1, x_2, x_3, x_4, όπου σε αυτά έχουν τοποθετηθεί βάρη $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{1}{4}$, $\lambda_3 = \frac{1}{6}$, $\lambda_4 = \frac{1}{12}$ αντίστοιχα, δηλαδή $x = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4$, $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$ (Σχ. 2.5). Ισοδύναμα, το $0 = (0,0)$ είναι το κέντρο βάρους των $x-x_1=(\frac{3}{4},\frac{5}{4})$, $x-x_2=(\frac{3}{4},-\frac{3}{4})$, $x-x_3=(-\frac{9}{4},-\frac{3}{4})$, $x-x_4=(-\frac{9}{4},\frac{5}{4})$, με βάρη ίσα με $\lambda_1,\lambda_2,\lambda_3,\lambda_4$ αντίστοιχα, δηλαδή:

$$\lambda_1(x-x_1)+\lambda_2(x-x_2)+\lambda_3(x-x_3)+\lambda_4(x-x_4)=0\ ,$$

(Σχ. 2.6 (a)). Παρατηρούμε ότι εάν επιλέξουμε $\mu_1=\mu_3=1$, $\mu_2=\mu_4=-1$, τότε έχουμε ότι $\mu_1x_1+\mu_2x_2+\mu_3x_3+\mu_4x_4=0$, με $\mu_1+\mu_2+\mu_3+\mu_4=0$, ή ισοδύναμα ότι το 0 είναι το κέντρο βάρους των $x_1-x, x-x_2, x_3-x, x-x_4$ με όλα τα βάρη ίσα με 1 (Σχ. 2.6 (b)).

Οπότε

$$\frac{1}{2}(x-x_1)+\frac{1}{4}(x-x_2)+\frac{1}{6}(x-x_3)+\frac{1}{12}(x-x_4)=0\ , \tag{2.4}$$
$$(x_1-x)+(x-x_2)+(x_3-x)+(x-x_4)=0\ . \tag{2.5}$$

Βάζοντας τα συστήματα των (2.4) και (2.5) το ένα πάνω στο άλλο έχουμε πάλι ότι το κέντρο βάρους του νέου συστήματος είναι το 0. Το ζητούμενο είναι να βρούμε σε ποιο από τα $x-x_1, x-x_2, x-x_3, x-x_4$ είναι δυνατόν να μηδενίσουμε (αφαιρέσουμε) το βάρος, μοιράζοντάς το στα υπόλοιπα, με τέτοιο τρόπο όμως ώστε να διατηρηθεί η ισορροπία.

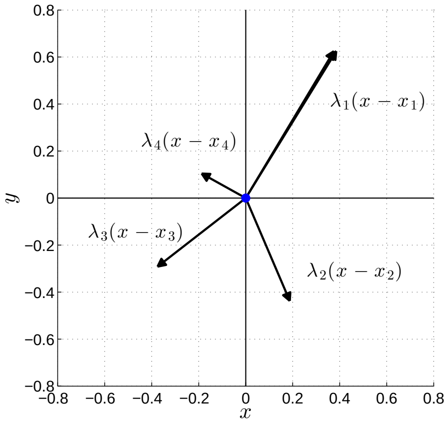
Δηλαδή ζητάμε $\alpha>0$ ώστε να ισχύει:

$$\begin{aligned}\left(\frac{1}{2}-\alpha\right)(x-x_1)+&\left(\frac{1}{4}+\alpha\right)(x-x_2)+\\ &+\left(\frac{1}{6}-\alpha\right)(x-x_3)+\left(\frac{1}{12}+\alpha\right)(x-x_4)=0\ ,\end{aligned} \tag{2.6}$$

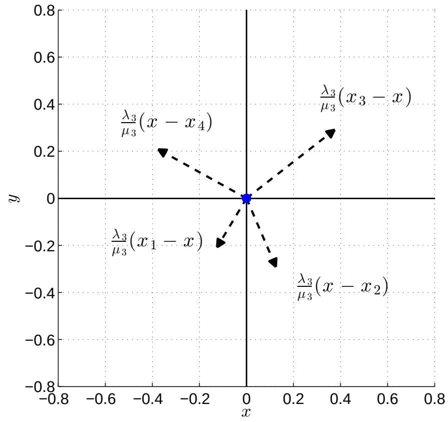
με $\frac{1}{2}-\alpha\geq 0$ και $\frac{1}{6}-\alpha\geq 0$, και το νέο σύστημα να μην περιέχει κάποιο από τα $x-x_1$, $x-x_3$. Η επιλογή είναι $\alpha=\frac{1}{6}=\lambda_3/\mu_3=\min\{\lambda_i/\mu_i,\ \mu_i>0\}$. Τότε η (2.6) γίνεται:

$$0=\frac{1}{3}(x-x_1)+\frac{5}{12}(x-x_2)+\frac{1}{4}(x-x_4)\ , \tag{2.7}$$

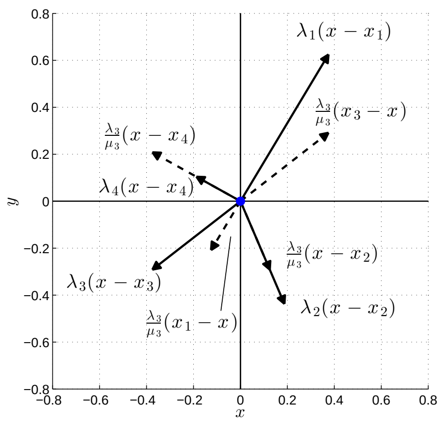
δηλαδή τα $x-x_1, x-x_2, x-x_4$ βρίσκονται σε ισορροπία με βάρη $1/3, 5/12$, $1/4$ αντίστοιχα (βλ. Σχ. 2.6).



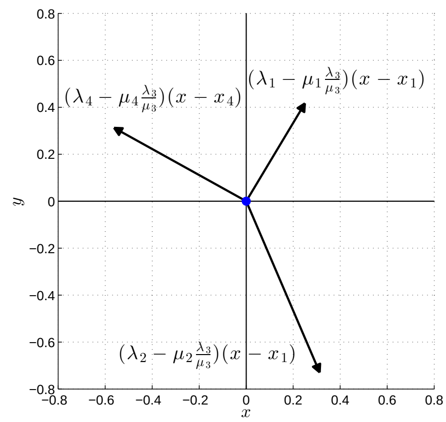
(a)



(b)



(c)



(d)

Σχήμα 2.6: (a) Το 0 ως κέντρο βάρους των $x-x_1, x-x_2, x-x_3, x-x_4$ με βάρη $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (βλ. (2.4)). (b) Το 0 ως κέντρο βάρους των $x_1-x, x-x_2, x_3-x, x-x_4$, όπου όλα τα διανύσματα έχουν πολλαπλασιαστεί με το λ_3/μ_3 (βλ. (2.5)). (c) Τα συστήματα του (a) και (b) το ένα πάνω στο άλλο (βλ. (2.6)). (d) Το 0 ως κέντρο βάρους των $x-x_1, x-x_2, x-x_4$ με βάρη $\lambda_1-\mu_1\frac{\lambda_3}{\mu_3}, \lambda_2-\mu_2\frac{\lambda_3}{\mu_3}, \lambda_4-\mu_4\frac{\lambda_3}{\mu_3}$ (βλ. (2.7)).

Βέβαια η επιλογή των μ_i ήταν αυθαίρετη. Μια άλλη κατάλληλη επιλογή είναι $\mu_1=\mu_3=-1, \mu_2=\mu_4=1$. Σε αυτήν την περίπτωση θα αφαιρεθεί το βάρος από το $x-x_4$.

Ας δούμε τώρα το παράδειγμα από την πλευρά της Γεωμετρίας.

Παρατηρούμε (Σχ. 2.5) ότι το x βρίσκεται στο τρίγωνο με κορυφές τα x_1, x_2, x_4 (και στο τρίγωνο με κορυφές τα x_1, x_2, x_3). Ακριβώς αυτός είναι ο ισχυρισμός του θεωρήματος Καραθεοδωρή για το επίπεδο.

Τέλος θα αποδείξουμε το Θεώρημα του Helly που αποτελεί την αρχή μιας μεγάλης κατηγορίας «θεωρημάτων τύπου Helly» που διατυπώνονται και αποδεικνύονται στη Συνδυαστική Γεωμετρία.

Θεώρημα 2.3 (Helly). *Έστω $\mathcal{F} = \{K_1, K_2, \ldots, K_n\}$ οικογένεια κυρτών συνόλων του $\mathbb{R}^d$ με $n \geq d+1$. Εάν κάθε $d+1$ από αυτά έχουν μη κενή τομή, τότε η τομή όλων είναι μη κενή.*

Απόδειξη

Θα γίνει επαγωγή στο πλήθος n των στοιχείων της $\mathcal{F}$. Εάν $n = d+1$ ισχύει από την υπόθεση.

Έστω ότι ισχύει για $n \geq d+1$.

Για την $\mathcal{F} = \{K_1, K_2, \ldots, K_n, K_{n+1}\}$ με την υπόθεση ότι κάθε $d+1$ από τα $K_1, \ldots, K_{n+1}$ έχουν μη κενή τομή θα αποδείξουμε ότι $\cap_{i=1}^{n+1} K_i \neq \varnothing$. Σύμφωνα με την επαγωγική υπόθεση υπάρχουν

$$\begin{aligned}
x_1 &\in K_2 \cap \cdots \cap K_{n+1} \\
&\vdots \\
x_i &\in K_1 \cap \cdots \cap K_{i-1} \cap K_{i+1} \cap \cdots \cap K_{n+1} \\
&\vdots \\
x_{n+1} &\in K_1 \cap \cdots \cap K_n
\end{aligned}$$

Θεωρούμε τα σημεία $x_1, \ldots, x_{n+1}$. Εάν δύο από αυτά συμπίμπτουν, π.χ. $x_1 = x_2$ τότε $x_1 = x_2 \in \bigcap_{i=1}^{n+1} K_i$, οπότε $\bigcap_{i=1}^{n+1} K_i \neq \varnothing$. Έστω ότι δεν συμβαίνει αυτό, οπότε το $A = \{x_1, \ldots, x_{n+1}\}$ έχει $n+1 \geq d+2$ στοιχεία. Τότε σύμφωνα με το Θεώρημα Radon που είναι άμεση απόρροια της Πρότασης R υπάρχουν μη κενά υποσύνολα B, C του A τέτοια ώστε $B \cap C = \varnothing$, $A = B \cup C$ και $x \in \operatorname{conv}(B) \cap \operatorname{conv}(C)$. Αναδιατάσσοντας, υποθέτουμε ότι $C = \{x_1, \ldots, x_k\}$

και $B = \{x_{k+1}, \dots, x_{n+1}\}$. Τότε

$$C \subseteq K_{k+1} \cap \cdots \cap K_{n+1} \quad \text{και} \quad B \subseteq K_1 \cap \cdots \cap K_k \, .$$

Επειδή η τομή κυρτών συνόλων είναι κυρτό σύνολο και η κυρτή θήκη ενός συνόλου είναι το μικρότερο κυρτό που το περιέχει, θα έχουμε ότι

$$x \in \operatorname{conv}(B) \cap \operatorname{conv}(C) \subseteq K_1 \cap \cdots \cap K_{n+1} \, ,$$

επομένως $\bigcap_{i=1}^{n+1} K_i \neq \varnothing$. Άρα το θεώρημα ισχύει για κάθε $n \geq d+1$.

ο.ε.δ.

Δεν θα αποτολμήσουμε να καταγράψουμε έστω και λίγες από τις συνέπειες που έχουν τα τρία αυτά θεμελιώδη θεωρήματα. Για τους ενδιαφερόμενους παραπέμπουμε σε ενδεικτική βιβλιογραφία [1], όπου υπάρχει μέγα πλήθος αναφορών.

Αναφορές

[1] P. M. Gruber, *Convex and Discrete Geometry*, Springer, 2007.

[2] C. Carathéodory, Über das lineare Mass von Punktmengen–eine Verallgemeinerung des Längenbegriffs, *Nachr. Gesell. Wiss. Göttingen* (1914), 406-426, *Ges. Math. Schr. IV* 249-275.

[3] E. Helly, Über mengen konvexer Körper mit gemeinschaftlichen Punkten, *Jahresbericht Deutsch. Math. Verein* **32** (1923), 175-176.

[4] J. Radon, Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten, *Mathematische Annalen.* **83** (1921), 113-115, doi:10.1007/BF01464231.