Διαπανεπιστημιακό - Διατμηματικό Πρόγραμμα
Μεταπτυχιακών Σπουδών
«Διδακτική και Μεθοδολογία των Μαθηματικών»

Εισαγωγή στην Θεωρία των Ανισοτήτων

Κωνσταντίνος Γιαννακάς

Αθήνα, 2007

Η παρούσα Διπλωματική Εργασία εκπονήθηκε στα πλαίσια των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης που απονέμει το Διαπανεπιστημιακό - Διατμηματικό Πρόγραμμα Μεπταπτυχιακών Σπουδών «Διδακτική και Μεθοδολογία των Μαθηματικών».

Εγκρίθηκε την 30ήν Απριλίου 2007 από Εξεταστική Επιτροπή αποτελούμενη από τους: Απόστολο Γιαννόπουλο (επιβλέποντα), Θεοδόσιο Ζαχαριάδη και Νικόλαο Κλαουδάτο.

Περιεχόμενα

Εισαγωγή

Οι ανισότητες παίζουν κεντρικό ρόλο στα Μαθηματικά, ιδιαίτερα στην Ανάλυση. Αφετηρία αυτής της εργασίας είναι η πεποίθησή μας ότι, πέρα από τη θεμελιώδη σημασία των ανισοτήτων και τις σημαντικές εφαρμογές τους, η ενασχόληση με αυτές προσφέρει ιδιαίτερη ευχαρίστηση: η διαδικασία της επίλυσης προβλημάτων είναι θεμελιώδης ανθρώπινη δραστηριότητα, μια δραστηριότητα γεμάτη συγκινήσεις, δημιουργικότητα και αγάπη για τη ζωή.

Τα τελευταία χρόνια, οι μαθηματικές ανισότητες είναι παραγκωνισμένες, θα έλεγε κανείς ότι απουσιάζουν από τα σχολικά μαθηματικά δρώμενα. Πιστεύουμε ότι η τάση αυτή πρέπει και μπορεί να ανατραπεί. Ένα κεφάλαιο αφιερωμένο στις θεμελιώδεις ανισότητες, με την κατάλληλη παρουσίαση και με την προσθήκη μιας συλλογής προβλημάτων, θα μπορούσε να αποτελέσει τη βάση για ένα σχολικό «σεμινάριο προβλημάτων» το οποίο θα προσέφερε πολλά στους μαθητές που αγαπούν τα Μαθηματικά. Με την εργασία αυτή ελπίζουμε να προσφέρουμε κάποιες ιδέες για την πραγματοποίηση αυτής της ιδέας.

Η εργασία χτίζεται γύρω από την ιστορία και τις ποικίλες αποδείξεις θεμελιωδών ανισοτήτων. Βασίζεται στο κλασικό βιβλίο των Hardy– Littlewood–Pólya και στο εξαιρετικό πρόσφατο βιβλίο του Steele.

Στο πρώτο μέρος της εργασίας μελετάμε αναλυτικά τις τέσσερις θεμελιώδεις ανισότητες: την ανισότητα Cauchy-Schwarz, την ανισότητα αριθμητικού–γεωμετρικού μέσου, την ανισότητα του Jensen και την ανισότητα του Hölder:

Στο πρώτο Κεφάλαιο μελετάμε την ανισότητα Cauchy–Schwarz: αν $a_1, \ldots, a_n$ και $b_1, \ldots, b_n$ είναι δύο n–άδες πραγματικών αριθμών, τότε

$$|a_1b_1 + a_2b_2 + \cdots + a_nb_n| \leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}.$$

Η κλασική αυτή ανισότητα μας δίνει την ευκαιρία να συζητήσουμε κάποιες συγκεκριμένες γενικές αρχές ή τεχνικές πάνω στις οποίες οικοδομείται η «θεωρία των ανισοτήτων». Περιγράφουμε τέσσερις αποδείξεις της ανισότητας: (α) απόδειξη με τη μέθοδο της μαθηματικής επαγωγής, (β) απόδειξη μέσω της προσθετικής ανισότητας

$$|a_1b_1 + \cdots + a_nb_n| \leq \frac{1}{2}\sum_{i=1}^{n} a_i^2 + \frac{1}{2}\sum_{i=1}^{n} b_i^2,$$

η οποία οδηγεί στην ανισότητα Cauchy αν εκμεταλλευτούμε σωστά την παρατήρηση ότι η τελευταία είναι ομογενής, (γ) απόδειξη μέσω της ταυτότητας του

Lagrange

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 = \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2.$$

(δ) απόδειξη με το «τέχνασμα» του τριωνύμου: η πολύ κομψή και σύντομη αυτή απόδειξη οφείλεται στον Schwarz, ο οποίος χρειάστηκε το συνεχές ανάλογο της ανισότητας Cauchy και την πλήρη διερεύνηση των περιπτώσεων ισότητας σε αυτό.

Στο δεύτερο Κεφάλαιο μελετάμε τη «δεύτερη ανισότητα» του Cauchy, την ανισότητα αριθμητικού–γεωμετρικού μέσου: για κάθε ακολουθία μη αρνητικών πραγματικών αριθμών $a_1, a_2, \ldots, a_n$ ισχύει η ανισότητα

$$(a_1 a_2 \cdots a_n)^{1/n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Περιγράφουμε το επιχείρημα του Cauchy με επαγωγή «προς τα εμπρός και προς τα πίσω»: η ζητούμενη ανισότητα αποδεικνύεται πρώτα για κάθε $n = 2^k$ και αυτό που μένει είναι να βρούμε κάποιον τρόπο για να γεμίσουμε τα κενά μεταξύ των δυνάμεων του 2. Η ιδέα του Cauchy είναι, για δοθείσα ακολουθία $a_1, a_2, \ldots, a_n$, να την επεκτείνει σε μια μακρύτερη ακολουθία $\alpha_1, \alpha_2, \ldots, \alpha_{2^k}$, όπου $2^k > n$, και να εφαρμόσει σε αυτήν την ήδη αποδεδειγμένη ανισότητα. Περιγράφουμε μια δεύτερη απόδειξη της ανισότητας με την κλασική μαθηματική επαγωγή. Παράλληλα, διερευνούμε τις περιπτώσεις ισότητας. Δίνουμε επίσης την απόδειξη της «γενίκευσης της ανισότητας αριθμητικού–γεωμετρικού μέσου»: Αν $p_1, p_2, \ldots, p_n$ είναι μη αρνητικοί πραγματικοί αριθμοί με άθροισμα 1 και αν $a_1, a_2, \ldots, a_n$ είναι μη αρνητικοί πραγματικοί αριθμοί, τότε

$$a_1^{p_1} a_2^{p_2} \cdots a_n^{p_n} \leq p_1 a_1 + p_2 a_2 + \cdots + p_n a_n.$$

Στην περίπτωση των ρητών εκθετών, η ανισότητα αυτή προκύπτει σχετικά απλά από την ανισότητα αριθμητικού–γεωμετρικού μέσου. Κατόπιν, μπορούμε να χρησιμοποιήσουμε την τεχνική της «προσέγγισης πραγματικών εκθετών από ρητούς». Η μέθοδος αυτή έχει το μειονέκτημα ότι δεν μπορεί να διευκρινίσει τις περιπτώσεις ισότητας. Περιγράφουμε μια άλλη απόδειξη, του G. Pólya, η οποία δίνει αυτή την πρόσθετη πληροφορία. Μια τρίτη απόδειξη δίνεται στο τρίτο Κεφάλαιο, μέσω της ανισότητας του Jensen και της παρατήρησης ότι η λογαριθμική συνάρτηση είναι γνησίως κοίλη.

Την τετράδα συμπληρώνει η ανισότητα του Hölder. Στη σύγχρονη γλώσσα, η ανισότητα ισχυρίζεται ότι για όλους τους μη αρνητικούς a_k και b_k, $k = 1, 2, \ldots, n$, ισχύει η ανισότητα

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q},$$

υπό τον όρο ότι οι εκθέτες $p > 1$ και $q > 1$ ικανοποιούν την σχέση $\frac{1}{p} + \frac{1}{q} = 1$.

Η παρουσίαση αυτών των ανισοτήτων μας δίνει την αφορμή να συζητήσουμε κάποια κεντρικά θέματα της «θεωρίας»:

1. Τη σημασία της πλήρους περιγραφής των περιπτώσεων στις οποίες μια ανισότητα ισχύει σαν ισότητα, και την έννοια της ευστάθειας μιας ανισότητας.

2. Την παράλληλη πορεία των συνεχών ανισοτήτων (για ολοκληρώματα συναρτήσεων) με τις διακριτές ανισότητες (για πεπερασμένα αθροίσματα ή αθροίσματα σειρών) και την αναγωγή των μεν στις δε.

3. Τη διαδικασία γενίκευσης μιας ανισότητας σε ευρύτερο πλαίσιο και το ρόλο που παίζει η «κατάλληλη απόδειξη» της αρχικής ανισότητας τόσο στο να σκεφτούμε όσο και στο να υλοποιήσουμε αυτή τη γενίκευση.

Στο δεύτερο μέρος της εργασίας παρουσιάζουμε κάποια ειδικά θέματα: τα περισσότερα αφορούν κλασικές ανισότητες οι οποίες αποδεικνύονται με τη βοήθεια των θεμελιωδών ανισοτήτων και των μεθόδων που αναπτύχθηκν στο πρώτο μέρος.

Στο πέμπτο Κεφάλαιο παρουσιάζουμε διάφορες αποδείξεις της ανισότητας του Torsten Carleman (1923): Αν $\sum_{n=1}^{\infty} a_n$ είναι μια σειρά μη αρνητικών πραγματικών αριθμών, και αν συμβολίσουμε με γ_n το γεωμετρικό μέσο των $a_1, a_2, \ldots, a_n$, τότε,

$$\sum_{n=1}^{\infty} \gamma_n \leq e \sum_{n=1}^{\infty} a_n.$$

Μεταξύ των αποδείξεων που μελετάμε είναι η αρχική απόδειξη του Carleman, η «μαγική» απόδειξη του Pólya, και η συνεχής έκδοση της ανισότητας που οφείλεται στον Carleson.

Στο έκτο Κεφάλαιο περιγράφουμε την αρνητική απάντηση στο εξής ερώτημα του Minkowski: Είναι σωστό ότι κάθε μη αρνητικό πολυώνυμο γράφεται σαν άθροισμα τετραγώνων πραγματικών πολυωνύμων; Ο Minkowski έκανε την εικασία ότι η απάντηση είναι αρνητική. Το 1888, ο Hilbert απέδειξε ότι η απάντηση είναι καταφατική μόνο στις εξής περιπτώσεις: όταν το πολυώνυμο είναι μιας μεταβλητής και ο βαθμός του P αυθαίρετος, όταν ο βαθμός του P είναι 2 και το πλήθος των μεταβλητών αυθαίρετο, και τέλος, όταν το P είναι πολυώνυμο με δύο μεταβλητές και έχει βαθμό 4. Η απόδειξη του Hilbert ήταν μακροσκελής, πολύπλοκη και έμμεση. Τα πρώτα συγκεκριμένα παραδείγματα μη αρνητικών πολυωνύμων που δεν γράφονται στη μορφή αθροίσματος τετραγώνων πραγματικών πολυωνύμων δόθηκαν το 1967 από τον T. S. Motzkin και το 1969 από τον R. M. Robinson, σχεδόν ογδόντα χρόνια μετά την απόδειξη της ύπαρξης τέτοιων πολυωνύμων από τον Hilbert. Τα συγκεκριμένα αυτά παραδείγματα χρησιμοποιούσαν την «τεχνική της ανισότητας αριθμητικού–γεωμετρικού μέσου» την οποία περιγράφουμε.

Στα δύο τελευταία Κεφάλαια αποδεικνύουμε κλασικές ανισότητες αναδιάταξης, τις ανισότητες των Newton και McLaurin για τις στοιχειώδεις συμμετρικές συναρτήσεις, και την ανισότητα του Muirhead.

Σε όλα τα Κεφάλαια δίνουμε ιστορικά στοιχεία για τα αποτελέσματα που παρουσιάζονται.

Μέρος I

Οι θεμελιώδεις ανισότητες

Κεφάλαιο 1

Η ανισότητα του Cauchy

1.1 Η ανισότητα Cauchy

Η ανισότητα Cauchy μας λέει ότι αν $a_1, \ldots, a_n$ και $b_1, \ldots, b_n$ είναι δύο n–άδες πραγματικών αριθμών, τότε

$$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \leq (a_1^2 + a_2^2 + \cdots + a_n^2)(b_1^2 + b_2^2 + \cdots + b_n^2)$$

ή, ισοδύναμα,

$$|a_1b_1 + a_2b_2 + \cdots + a_nb_n| \leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}.$$

Αφού $a \leq |a|$ για κάθε πραγματικό αριθμό a, έπεται ότι

$$a_1b_1 + a_2b_2 + \cdots + a_nb_n \leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}.$$

Η ανισότητα Cauchy είναι αναμφίβολα μία από τις ευρύτερα χρησιμοποιούμενες και ενδιαφέρουσες ανισότητες σε όλα τα μαθηματικά. Ένας από τους βασικούς στόχους αυτής της εργασίας είναι να ανακαλύψουμε μονοπάτια για την κατανόηση αυτής της ανισότητας, των πολλών επεκτάσεων και εφαρμογών της.

1.2 Επαγωγική απόδειξη

Ένα βασικό ερώτημα που θα μας απασχολήσει στην εργασία είναι αν υπάρχουν κάποιες συγκεκριμένες γενικές αρχές ή τεχνικές πάνω στις οποίες μπορούμε να οικοδομήσουμε την «θεωρία των ανισοτήτων». Αν κάποιος ρίξει μια ψύχραιμη ματιά στην ανισότητα Cauchy, βλέπει πως υπάρχει μια βασική αρχή που μπορεί να φανεί χρήσιμη.

- Κάθε φορά που βλέπουμε μια πρόταση που εξαρτάται από τον φυσικό αριθμό n, είναι αρκετά πιθανόν η μαθηματική επαγωγή να οδηγήσει στην απόδειξή της.

Αξίζει να σημειώσουμε ότι κανένα από τα βασικά εγχειρίδια της άλγεβρας ή της ανάλυσης δεν δίνει επαγωγική απόδειξη της ανισότητας Cauchy: ο λόγος είναι ότι

υπάρχουν άλλες, ιδιαίτερα ευρηματικές αποδείξεις της ανισότητας, οι οποίες είναι πολύ σύντομες και κομψές. Το ερώτημα λοιπόν είναι: υπάρχει κάποιο εμπόδιο στην εφαρμογή της γνωστής μεθόδου της επαγωγής;

Για $n = 1$ βλέπουμε ότι η ανισότητα Cauchy είναι τετριμμένα αληθής. Δηλαδή έχουμε

$$ab \leq |ab| = |a|\,|b| = \sqrt{a^2}\sqrt{b^2}.$$

- Αυτό μας αρκεί για να ξεκινήσουμε την επαγωγή μας, δεν προσφέρει όμως καμία υπόδειξη για τον τρόπο με τον οποίο θα συνεχίσουμε. Αν θέλουμε να ελπίζουμε ότι θα ανακαλύψουμε κάποια αξιόλογη ιδέα στο πρώτο μας βήμα, πρέπει να εξετάσουμε την περίπτωση $n = 2$.

Τότε, η μορφή που παίρνει η ανισότητα Cauchy είναι η εξής:

$$(a_1b_1 + a_2b_2)^2 \leq (a_1^2 + a_2^2)(b_1^2 + b_2^2). \tag{1.1}$$

Αυτή η ανισότητα είναι απλή, και με μια ματιά μπορούμε να δούμε γιατί αληθεύει. Για χάρη της απόδειξης όμως, ας προσποιηθούμε ότι η (1.1) δεν είναι προφανής.

- Πώς θα μπορούσαμε να δουλέψουμε συστηματικά για να την αποδείξουμε;

Προφανώς, δεν υπάρχει τίποτε πιο συστηματικό από το να αναπτύξουμε τα δύο μέλη και να καταλήξουμε στην ισοδύναμη ανισότητα

$$a_1^2b_1^2 + 2a_1b_1a_2b_2 + a_2^2b_2^2 \leq a_1^2b_1^2 + a_1^2b_2^2 + a_2^2b_1^2 + a_2^2b_2^2.$$

Έτσι, αφού κάνουμε τις φυσιολογικές απαλοιφές και συγκεντρώσουμε όλους τους όρους στο ίδιο μέλος, βλέπουμε ότι η ανισότητα (1.1) είναι ισοδύναμη με την

$$0 \leq (a_1b_2)^2 - 2(a_1b_2)(a_2b_1) + (a_2b_1)^2. \tag{1.2}$$

Αυτή η ισοδύναμη ανισότητα πραγματικά μας φέρνει στη λύση του προβλήματός μας. Από την γνωστή ταυτότητα $x^2 - 2xy + y^2 = (x - y)^2$ βλέπουμε ότι

$$(a_1b_2)^2 - 2(a_1b_2)(a_2b_1) + (a_2b_1)^2 = (a_1b_2 - a_2b_1)^2. \tag{1.3}$$

Αφού η τελευταία ποσότητα είναι μη αρνητική, έχουμε αποδείξει ότι η (1.2) είναι αληθής. Διαμέσου των ισοδυναμιών μας, συμπεραίνουμε ότι η ανισότητα (1.1) είναι επίσης αληθής, και έτσι αποδείξαμε την ανισότητα του Cauchy για $n = 2$.

Τώρα που αποδείξαμε μια μη τετριμμένη περίπτωση της ανισότητας Cauchy, είμαστε έτοιμοι να επιχειρήσουμε το επαγωγικό βήμα.

- Αν ονομάσουμε $P(n)$ την πρόταση ότι η ανισότητα του Cauchy αληθεύει για τον φυσικό n, πρέπει να δείξουμε ότι οι $P(2)$ και $P(n)$ έχουν ως επακόλουθο την $P(n + 1)$. Με αυτό το σχέδιο στο μυαλό, δεν αργούμε να σκεφτούμε ότι πρέπει πρώτα να εφαρμόσουμε την υπόθεση $P(n)$ και μετά να χρησιμοποιήσουμε την $P(2)$ για να συνδυάσουμε τους δύο όρους που θα προκύψουν.

Συγκεκριμένα, έχουμε

$$\begin{aligned} a_1b_1 + a_2b_2 + \cdots + a_nb_n + a_{n+1}b_{n+1} &= (a_1b_1 + a_2b_2 + \cdots + a_nb_n) + a_{n+1}b_{n+1} \\ &\le \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2} + a_{n+1}b_{n+1} \\ &\le \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2 + a_{n+1}^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2 + b_{n+1}^2}, \end{aligned}$$

όπου στην πρώτη ανισότητα χρησιμοποιήσαμε την επαγωγική υπόθεση $P(n)$, και στην δεύτερη ανισότητα χρησιμοποιήσαμε την $P(2)$ στη μορφή

$$\alpha\beta + a_{n+1}b_{n+1} \le \sqrt{\alpha^2 + a_{n+1}^2}\sqrt{\beta^2 + b_{n+1}^2}$$

με νέες μεταβλητές τις

$$\alpha = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \quad \text{και} \quad \beta = \sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}.$$

Η μόνη δυσκολία που μπορεί να συναντήσει κανείς σε αυτήν την απόδειξη εμφανίζεται στο τελευταίο βήμα, όπου χρειάζεται να δούμε πώς να χρησιμοποιήσουμε την $P(2)$. Στη συγκεκριμένη περίπτωση, το τέχνασμα που απαιτείται ήταν μάλλον μέτριας δυσκολίας, δίνει όμως μια ιδέα για την φύση των προκλήσεων που θα συναντήσουμε σε περισσότερο πολύπλοκα προβλήματα.

1.3 Από το ποιοτικό στο ποσοτικό συμπέρασμα

Σχεδόν κάθε σημαντικό ποσοτικό αποτέλεσμα έχει διάφορα άμεσα ποιοτικά πορίσματα και, σε πολλές περιπτώσεις, αυτά τα πορίσματα μπορούν να παραχθούν ανεξάρτητα, χωρίς προσφυγή στο αποτέλεσμα που πρώτο τα έφερε στο φως. Οι εναλλακτικές αποδείξεις που παίρνουμε με αυτόν τον τρόπο μας βοηθάνε να κατανοήσουμε βαθύτερα την φύση των προβλημάτων μας. Σε αυτήν την παράγραφο, θα χρησιμοποιήσουμε την ανισότητα Cauchy σαν παράδειγμα για την αρχή της

- «κατανόησης ενός **ποσοτικού αποτελέσματος** μέσω της εστίασης πάνω στις **ποιοτικές συνέπειές** του».

Μια άμεση ποιοτική συνέπεια της ανισότητας Cauchy είναι το γεγονός ότι

$$\text{αν } \sum_{k=1}^{\infty} a_k^2 < \infty \text{ και } \sum_{k=1}^{\infty} b_k^2 < \infty, \text{ τότε } \sum_{k=1}^{\infty} |a_kb_k| < \infty. \tag{1.4}$$

Ας υποθέσουμε ότι θέλουμε να δώσουμε απόδειξη αυτού του ισχυρισμού χωρίς να χρησιμοποιήσουμε την ανισότητα Cauchy.

Προσπαθώντας να αντιμετωπίσουμε αυτό το πρόβλημα, γρήγορα συνειδητοποιούμε πως αυτό που χρειάζεται να αποδείξουμε είναι ότι το γινόμενο a_kb_k είναι μικρό όταν τα a_k^2 και b_k^2 είναι μικρά. Θα μπορούσαμε να βεβαιωθούμε γι' αυτό αν αποδεικνύαμε την ύπαρξη μιας σταθεράς C τέτοιας ώστε

$$xy \le C(x^2 + y^2) \quad \text{για κάθε } x, y \in \mathbb{R}.$$

Ευτυχώς, μόλις γράψουμε αυτή την ανισότητα, βλέπουμε αμέσως γιατί αληθεύει. Αρκεί να θυμηθούμε την γνωστή παραγοντοποίηση

$$0 \leq (x-y)^2 = x^2 - 2xy + y^2,$$

και αμέσως θα παρατηρήσουμε ότι

$$xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2 \quad \text{για κάθε } x, y \in \mathbb{R}. \tag{1.5}$$

Τώρα, αν εφαρμόσουμε αυτή την ανισότητα για $x = |a_k|$ και $y = |b_k|$, και αν προσθέσουμε για όλα τα k, βρίσκουμε την *προσθετική* ανισότητα

$$\sum_{k=1}^{\infty} |a_k b_k| \leq \frac{1}{2}\sum_{k=1}^{\infty} a_k^2 + \frac{1}{2}\sum_{k=1}^{\infty} b_k^2, \tag{1.6}$$

η οποία έχει ως προφανή συνέπεια την (1.4).

Το ερώτημα είναι πόσο ισχυρή είναι η ποιοτική προσέγγιση της προηγούμενης παραγράφου.

- Κάθε φορά που συναντάμε μια καινούργια ανισότητα, είναι «καθήκον» μας να ελέγξουμε την ισχύ της.

Στη συγκεκριμένη περίπτωση, αυτό που παρουσιάζει ενδιαφέρον είναι να συγκρίνουμε τη νέα προσθετική ανισότητα με τις ποσοτικές εκτιμήσεις που προκύπτουν από την ανισότητα Cauchy.

Η προσθετική ανισότητα (1.6) έχει δύο όρους στο δεξιό της μέλος ενώ η ανισότητα του Cauchy μόνο έναν. Μια πρώτη ιδέα θα ήταν να συνδυάσουμε τους δύο όρους της (1.6), και ένας φυσιολογικός τρόπος για να υλοποιήσουμε αυτήν την ιδέα είναι να κανονικοποιήσουμε τις ακολουθίες $\{a_k\}$ και $\{b_k\}$ ώστε καθένα από τα αθροίσματα στο δεξιό μέλος να ισούται με 1.

Έτσι, αν καμία από τις δύο ακολουθίες δεν είναι η μηδενική, μπορούμε να θεωρήσουμε τις νέες μεταβλητές

$$x_k = a_k \Big/ \left(\sum_j a_j^2\right)^{1/2} \quad \text{και} \quad y_k = b_k \Big/ \left(\sum_j b_j^2\right)^{1/2},$$

οι οποίες είναι *κανονικοποιημένες* με την έννοια ότι

$$\sum_{k=1}^{\infty} x_k^2 = \sum_{k=1}^{\infty} \left\{ a_k / \left(\sum_j a_j^2\right)\right\} = 1$$

και

$$\sum_{k=1}^{\infty} y_k^2 = \sum_{k=1}^{\infty} \left\{ b_k / \left(\sum_j b_j^2\right)\right\} = 1.$$

Τώρα, αν εφαρμόσουμε την ανισότητα (1.6) για τις ακολουθίες $\{x_k\}$ και $\{y_k\}$, παίρνουμε το «αθώο» άνω φράγμα

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \frac{1}{2}\sum_{k=1}^{\infty} x_k^2 + \frac{1}{2}\sum_{k=1}^{\infty} y_k^2 = 1.$$

Αν τακτοποιήσουμε τους παρονομαστές, ανακαλύπτουμε την ανισότητα του Cauchy – τώρα μάλιστα, καλύπτει και την περίπτωση που οι ακολουθίες μας είναι άπειρες:

$$\sum_{k=1}^{\infty} |a_k b_k| \leq \left(\sum_{j=1}^{\infty} a_j^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} b_j^2 \right)^{1/2}. \tag{1.7}$$

Η προσθετική ανισότητα (1.6) μας οδήγησε σε μια απόδειξη της ανισότητας του Cauchy η οποία είναι σύντομη και απλή. Την ίδια όμως στιγμή, μας εισάγει σε ένα πολύ σημαντικό θέμα:

- *Η κανονικοποίηση μας προσφέρει ένα συστηματικό τρόπο για να περνάμε από μια προσθετική σε μια πολλαπλασιαστική ανισότητα, κι αυτό το πέρασμα απαιτείται συχνά για την απόδειξη σημαντικών ανισοτήτων.*

Μια άλλη βασική αρχή που προκύπτει από τη μελέτη του τρόπου με τον οποίο αναπτύσσονται και εφαρμόζονται οι ανισότητες είναι ότι

- *μπορούμε να αποκομίσουμε πολλά οφέλη προσπαθώντας να κατανοήσουμε πότε μια ανισότητα ισχύει σαν ισότητα και κατ' επέκταση, πότε είναι ακριβής ή σχεδόν ακριβής.*

Στην περίπτωση της ανισότητας Cauchy, αυτό που πρέπει να δούμε είναι η σχέση που πρέπει να υπάρχει μεταξύ των ακολουθιών $\{a_k\}$ και $\{b_k\}$ ώστε να ισχύει η ισότητα

$$\sum_{k=1}^{\infty} a_k b_k = \left(\sum_{j=1}^{\infty} a_j^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} b_j^2 \right)^{1/2}. \tag{1.8}$$

Αν περιοριστούμε στην μη τετριμμένη περίπτωση όπου καμία από τις δύο ακολουθίες δεν είναι ταυτοτικά μηδενική και, ταυτόχρονα, τα δύο αθροίσματα του δεξιού μέλους της (1.8) είναι πεπερασμένα, τότε βλέπουμε ότι όλα τα βήματα που χρησιμοποιήσαμε στην απόδειξη της (1.7) αντιστρέφονται. Έτσι, βλέπουμε ότι η ισότητα στην (1.8) έχει σαν συνέπεια την

$$\sum_{k=1}^{\infty} x_k y_k = \frac{1}{2} \sum_{k=1}^{\infty} x_k^2 + \frac{1}{2} \sum_{k=1}^{\infty} y_k^2 = 1. \tag{1.9}$$

Από την ανισότητα $xy \leq (x^2 + y^2)/2$ ξέρουμε επίσης ότι

$$x_k y_k \leq \frac{1}{2} x_k^2 + \frac{1}{2} y_k^2 \quad \text{για κάθε } k = 1, 2, \ldots, \tag{1.10}$$

και συνεπώς, αν, έστω και για μία τιμή του k, είχαμε γνήσια ανισότητα στην (1.10), τότε δεν θα μπορούσαμε να έχουμε ισότητα στην (1.9). Βλέπουμε λοιπόν ότι ισότητα στην (1.8) μπορεί να ισχύει μόνο αν $x_k = y_k$ για κάθε $k = 1, 2, \ldots$. Από τον τρόπο ορισμού των x_k και y_k, αυτό σημαίνει ότι τότε

$$a_k = \lambda b_k \quad \text{για κάθε } k = 1, 2, \ldots, \tag{1.11}$$

όπου η σταθερά λ δίνεται από τον λόγο

$$\lambda = \left(\sum_{j=1}^{\infty} a_j^2\right)^{1/2} \Big/ \left(\sum_{j=1}^{\infty} b_j^2\right)^{1/2}.$$

Παρατηρήστε ότι το επιχείρημα που χρησιμοποιήσαμε για την διερεύνηση της περίπτωσης της ισότητας δεν ήταν δύσκολο. Παρόλα αυτά, το συμπέρασμα περιγράφει ένα «*μικρό θαύμα*»:

- *Μία μόνο ισότητα – η* (1.8) *– είναι ικανή να επιβάλλει άπειρες το πλήθος ισότητες, μία για κάθε τιμή του* k*, στην* (1.11).

1.4 Η αξία της γενίκευσης

Αθροίσματα σαν αυτά που εμφανίζονται στην ανισότητα Cauchy δεν είναι ιδιαίτερα εύχρηστα από την άποψη του συμβολισμού. Για το λόγο αυτό, είναι πολύ χρήσιμη η εισαγωγή της σύντμησης

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{j=1}^{n} a_j b_j, \tag{1.12}$$

όπου $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ και $\mathbf{b} = (b_1, b_2, \ldots, b_n)$. Αυτή η σύντμηση μας επιτρέπει να γράψουμε την ανισότητα Cauchy, εντελώς λακωνικά, στη μορφή

$$\langle \mathbf{a}, \mathbf{b} \rangle \leq \langle \mathbf{a}, \mathbf{a} \rangle^{\frac{1}{2}} \langle \mathbf{b}, \mathbf{b} \rangle^{\frac{1}{2}}. \tag{1.13}$$

Η συντομία είναι πάντοτε ωραία, αλλά υπάρχουν ακόμη βαθύτερα οφέλη με αυτόν τον συμβολισμό, τα οποία προκύπτουν αν τον εφοδιάσουμε με μια πιο αφηρημένη ερμηνεία. Συγκεκριμένα, αν V είναι ένας πραγματικός διανυσματικός χώρος (για παράδειγμα, ο $\mathbb{R}^d$), τότε λέμε ότι μια συνάρτηση $V \times V \to \mathbb{R}$ με $(\mathbf{a}, \mathbf{b}) \mapsto \langle \mathbf{a}, \mathbf{b} \rangle$ είναι *εσωτερικό γινόμενο* και λέμε ότι ο $(V, \langle \cdot, \cdot \rangle)$ είναι ένας *πραγματικός χώρος με εσωτερικό γινομένο*, αν το ζευγάρι $(V, \langle \cdot, \cdot \rangle)$ έχει τις ακόλουθες πέντε ιδιότητες :

(i) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ για κάθε $\mathbf{v} \in V$,

(ii) $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ αν και μόνο αν $\mathbf{v} = 0$,

(iii) $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle$ για κάθε $\alpha \in \mathbb{R}$ και για κάθε $\mathbf{v}, \mathbf{w} \in V$,

(iv) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$ για κάθε $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, και τέλος,

(v) $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ για κάθε $\mathbf{v}, \mathbf{w} \in V$.

Μπορούμε εύκολα να ελέγξουμε ότι η συντομογραφία που εισάγεται στην (1.12) έχει όλες αυτές τις ιδιότητες, αλλά υπάρχουν πολλά ακόμα παραδείγματα χρήσιμων εσωτερικών γινομένων. Για παράδειγμα, αν θεωρήσουμε τυχόν σύνολο $\{w_j : j = 1, 2, \ldots, n\}$ θετικών πραγματικών αριθμών, τότε μπορούμε πολύ εύκολα να ορίσουμε ένα εσωτερικό γινόμενο στον $\mathbb{R}^n$, θέτοντας

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{j=1}^{n} a_j b_j w_j. \tag{1.14}$$

Με αυτόν τον ορισμό, μπορεί κάποιος να ελέγξει όπως προηγούμενα ότι το $\langle \mathbf{a}, \mathbf{b} \rangle$ ικανοποιεί όλες τις ιδιότητες που απαιτούνται από ένα εσωτερικό γινόμενο. Επιπλέον, αυτό το παράδειγμα απλώς αποκαλύπτει την κορυφή ενός παγόβουνου. Υπάρχουν πολλά χρήσιμα εσωτερικά γινόμενα, και εμφανίζονται σε διάφορα μαθηματικά πλαίσια.

Ένα ιδιαίτερα χρήσιμο παράδειγμα εσωτερικού γινομένου προκύπτει αν θεωρήσουμε το σύνολο $V = C[a, b]$ των συνεχών πραγματικών συναρτήσεων με πεδίο ορισμού το κλειστό και φραγμένο διάστημα $[a, b]$ και ορίσουμε $\langle \cdot, \cdot \rangle$ στο V θέτοντας

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx. \tag{1.15}$$

Γενικότερα, αν $w : [a, b] \to \mathbb{R}$ είναι μια συνεχής συνάρτηση τέτοια ώστε $w(x) > 0$ για κάθε $x \in [a, b]$, τότε μπορούμε να ορίσουμε ένα εσωτερικό γινόμενο στο $C[a, b]$ θέτοντας

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)dx.$$

Αντιμετωπίζουμε τώρα μία από τις ευχάριστες στιγμές όπου ένας καλός συμβολισμός μας οδηγεί σε ένα καλό θεώρημα. Παρουσιάσαμε την έννοια του εσωτερικού γινομένου για να διατυπώσουμε την ανισότητα Cauchy με απλό τρόπο, και τώρα βλέπουμε ότι ο συμβολισμός του μας οδηγεί προς μια ενδιαφέρουσα εικασία:

- Είναι αλήθεια ότι σε κάθε χώρο εσωτερικού γινόμενου ισχύει η ανισότητα $\langle \mathbf{v}, \mathbf{w} \rangle \leq \langle \mathbf{v}, \mathbf{v} \rangle^{\frac{1}{2}} \langle \mathbf{w}, \mathbf{w} \rangle^{\frac{1}{2}}$;

Αυτή η εικασία είναι πραγματικά σωστή, και θα την διατυπώσουμε με ακρίβεια στο επόμενο πρόβλημά μας.

Πρόβλημα 1.4.1. *Σε κάθε πραγματικό χώρο με εσωτερικό γινόμενο $(V, \langle \cdot, \cdot \rangle)$ ισχύει ότι: για κάθε $\mathbf{v}$ και $\mathbf{w}$ στον V ισχύει*

$$\langle \mathbf{v}, \mathbf{w} \rangle \leq \langle \mathbf{v}, \mathbf{v} \rangle^{\frac{1}{2}} \langle \mathbf{w}, \mathbf{w} \rangle^{\frac{1}{2}}. \tag{1.16}$$

Επιπλέον, για μη μηδενικά διανύσματα $\mathbf{v}$ και $\mathbf{w}$, έχουμε

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle^{\frac{1}{2}} \langle \mathbf{w}, \mathbf{w} \rangle^{\frac{1}{2}} \quad \text{αν και μόνο αν } \mathbf{v} = \lambda \mathbf{w}$$

για κάποιον μη μηδενικό πραγματικό αριθμό λ.

Μια ιδέα είναι να χρησιμοποιήσουμε κάποια παραλλαγή της προσθετικής μεθόδου με την βοήθεια της οποίας δώσαμε την δεύτερη απόδειξη της ανισότητας Cauchy.

- Είχαμε παρατηρήσει ότι η $(x-y)^2 \geq 0$ δίνει την $xy \leq x^2/2 + y^2/2$. Προσπαθούμε λοιπόν να δούμε αν κάποια ανάλογη ιδέα θα μπορούσε να δουλέψει στο αφηρημένο πλαίσιο.

Εδώ, φυσικά, μπορούμε να χρησιμοποιήσουμε μόνο τις ιδιότητες του εσωτερικού γινομένου, και, αν ανατρέξουμε στη λίστα (i)–(v) κοιτάζοντας για ένα ανάλογο της $(x-y)^2 \geq 0$, είναι αρκετά λογικό να σκεφτούμε να χρησιμοποιήσουμε την ιδιότητα (i) στη μορφή

$$\langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w}\rangle \geq 0.$$

Τώρα, αν αναπτύξουμε αυτήν την ανισότητα με τη βοήθεια των υπολοίπων ιδιοτήτων του εσωτερικού γινομένου, βλέπουμε ότι

$$\langle \mathbf{v}, \mathbf{w}\rangle \leq \frac{1}{2}\langle \mathbf{v}, \mathbf{v}\rangle + \frac{1}{2}\langle \mathbf{w}, \mathbf{w}\rangle. \tag{1.17}$$

Αυτό είναι ακριβώς το ανάλογο της προσθετικής ανισότητας που οδήγησε στην δεύτερή μας απόδειξη της ανισότητας Cauchy. Από δώ και πέρα, «ακολουθούμε την πεπατημένη».

Εδώ είμαστε υπερβολικά τυχεροί, αφού έχουμε ήδη ανακαλύψει μια τεχνική που ταιριάζει στην περίπτωση – την μέθοδο της κανονικοποίησης, που μετατρέπει μια προσθετική ανισότητα σε μια πολλαπλασιαστική ανισότητα. Ο όρος «κανονικοποίηση» δεν σημαίνει παντού το ίδιο πράγμα, αν όμως πάρουμε την προγενέστερη ανάλυση σαν οδηγό μας, αυτό που θέλουμε εδώ είναι να κανονικοποιήσουμε τα $\mathbf{v}$ και $\mathbf{w}$ ώστε το δεξιό μέλος της (1.17) να ισούται με 1.

Αφού η ανισότητα (1.16) ισχύει τετριμμένα αν κάποιο από τα $\mathbf{v}$ και $\mathbf{w}$ ισούται με μηδέν, μπορούμε να υποθέσουμε, χωρίς βλάβη της γενικότητας, ότι τα $\langle \mathbf{v}, \mathbf{v}\rangle$ και $\langle \mathbf{w}, \mathbf{w}\rangle$ είναι και τα δύο μη μηδενικά, οπότε οι κανονικοποιημένες μεταβλητές

$$\hat{\mathbf{v}} = v/\langle \mathbf{v}, \mathbf{v}\rangle^{\frac{1}{2}} \quad \text{και} \quad \hat{\mathbf{w}} = w/\langle \mathbf{w}, \mathbf{w}\rangle^{\frac{1}{2}} \tag{1.18}$$

είναι καλά ορισμένες. Αν αντικαταστήσουμε αυτές τις τιμές για $\mathbf{v}$ και $\mathbf{w}$ στην (1.17), βρίσκουμε $\langle \hat{\mathbf{v}}, \hat{\mathbf{w}}\rangle \leq 1$. Αυτό μας λέγει ότι $\langle \mathbf{v}, \mathbf{w}\rangle \leq \langle \mathbf{v}, \mathbf{v}\rangle^{\frac{1}{2}}\langle \mathbf{w}, \mathbf{w}\rangle^{\frac{1}{2}}$. Δηλαδή, αυτό ακριβώς που θέλαμε να δείξουμε.

Τέλος, για να δούμε πότε ισχύει ισότητα, αρκεί να επαναλάβουμε το συλλογισμό μας στην αντίστροφη κατεύθυνση. Αν ισχύει ισότητα στην αφηρημένη ανισότητα Cauchy (1.16) για κάποια μη μηδενικά διανύσματα $\mathbf{v}$ και $\mathbf{w}$, τότε οι κανονικοποιημένες μεταβλητές $\hat{\mathbf{v}}$ και $\hat{\mathbf{w}}$ είναι καλά ορισμένες και η ισότητα $\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle^{\frac{1}{2}}\langle \mathbf{w}, \mathbf{w}\rangle^{\frac{1}{2}}$ μας λέει ότι $\langle \hat{\mathbf{v}}, \hat{\mathbf{w}}\rangle = 1$. Αυτό με τη σειρά του σημαίνει ότι $\langle \hat{\mathbf{v}}-\hat{\mathbf{w}}, \hat{\mathbf{v}}-\hat{\mathbf{w}}\rangle = 0$ (για να το δούμε, απλώς αναπτύσσουμε το εσωτερικό γινόμενο). Η τελευταία όμως σχέση σημαίνει ότι $\hat{\mathbf{v}}-\hat{\mathbf{w}} = 0$. Με άλλα λόγια, $\mathbf{v} = \lambda\mathbf{w}$ όπου θέσαμε $\lambda = \langle \mathbf{v}, \mathbf{v}\rangle^{\frac{1}{2}}/\langle \mathbf{w}, \mathbf{w}\rangle^{\frac{1}{2}}$.

1.5 Η ταυτότητα του Lagrange

Για την απόδειξη της ανισότητας Cauchy με επαγωγή χρησιμοποιήσαμε την ταυτότητα

$$(a_1^2+a_2^2)(b_1^2+b_2^2) = (a_1b_1+a_2b_2)^2 + (a_1b_2-a_2b_1)^2, \tag{1.19}$$

χωρίς να την εκμεταλλευτούμε πλήρως. Συγκεκριμένα, αγνοήσαμε τον όρο $(a_1b_2-a_2b_1)^2$, αφού το μόνο που χρειαζόταν να παρατηρήσουμε ήταν ότι είναι *μη αρνητικός*.

Φυσικά, σωστό είναι, ότι σε κάθε ανισότητα γίνεται ένας συμβιβασμός ανάμεσα στην ακρίβεια και την απλότητα, δεν θέλουμε όμως αυτό να έχει σαν συνέπεια την απώλεια πληροφορίας. Έτσι, αντιμετωπίζουμε ένα φυσιολογικό ερώτημα:

- μπορούμε να αποσπάσουμε κάποια χρήσιμα συμπεράσματα από τον όρο που πετάξαμε;

Δεν υπάρχει αμφιβολία ότι ο όρος $(a_1b_2 - a_2b_1)^2$ περιέχει κάποια πληροφορία. Το λιγότερο που μπορούμε να πούμε είναι ότι μετράει ακριβώς την διαφορά ανάμεσα στα τετράγωνα των δύο μελών της ανισότητας του Cauchy.

- Μας δίνει λοιπόν ένα χρήσιμο τρόπο για να υπολογίσουμε το *σφάλμα* που προκύπτει κάθε φορά που εφαρμόζουμε την ανισότητα του Cauchy.

Η βασική ταυτότητα (1.19) μας λέει επίσης ότι, στην περίπτωση $n = 2$, έχουμε ισότητα στην ανισότητα του Cauchy ακριβώς όταν $(a_1b_2 - a_2b_1)^2 = 0$. Αν λοιπόν υποθέσουμε ότι $(b_1, b_2) \neq (0, 0)$, βλέπουμε ότι

- έχουμε ισότητα αν και μόνο αν τα (a_1, a_2) και (b_1, b_2) είναι ανάλογα με την έννοια ότι

$$a_1 = \lambda b_1 \text{ και } a_2 = \lambda b_2 \quad \text{για κάποιον πραγματικό αριθμό } \lambda.$$

Αυτή η παρατήρηση έχει σημαντικές συνέπειες: το πρώτο μας πρόβλημα ζητάει να αποδείξουμε τον αντίστοιχο χαρακτηρισμό για την περίπτωση της ισότητας στην n–διάστατη ανισότητα του Cauchy.

Πρόβλημα 1.5.1 (ισότητα στην ανισότητα του Cauchy). *Δείξτε ότι αν $(b_1, \ldots, b_n) \neq \mathbf{0}$ τότε έχουμε ισότητα στην ανισότητα του Cauchy αν και μόνο αν υπάρχει σταθερά λ ώστε $a_i = \lambda b_i$ για κάθε $i = 1, 2, \ldots, n$. Αν ήδη γνωρίζετε μια απόδειξη αυτού του ισχυρισμού, προσπαθήστε να σκεφτείτε κάποια διαφορετική.*

Η ταυτότητα (1.19) μας δίνει μια σύντομη λύση για το Πρόβλημα 1.5.1 στην περίπτωση $n = 2$. Θα μπορούσαμε λοιπόν να προσπαθήσουμε να λύσουμε το γενικό πρόβλημα επεκτείνοντας κατάλληλα την ταυτότητα (1.19) στις n διαστάσεις. Έτσι, θεωρούμε το πολυώνυμο $Q_n = Q_n(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n)$ που προκύπτει ως διαφορά των τετραγώνων των δύο μελών της ανισότητας του Cauchy, δηλαδή

$$Q_n = (a_1^2 + a_2^2 + \cdots + a_n^2)(b_1^2 + b_2^2 + \cdots + b_n^2) - (a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2.$$

Το Q_n μετράει την «απόκλιση» της ανισότητας του Cauchy από το να είναι ισότητα στις n διαστάσεις, ακριβώς όπως το $Q_2 = (a_1b_2 - a_2b_1)^2$ μετράει την αντίστοιχη απόκλιση στις δύο διαστάσεις.

- Έχουμε ήδη δεί ότι το Q_2 μπορεί να γραφτεί σαν τετράγωνο ενός πολυωνύμου, και τώρα το πρόβλημα είναι αν υπάρχει ανάλογη αναπαράσταση του Q_n στη μορφή ενός τετραγώνου ή ενός αθροίσματος τετραγώνων.

Αναπτύσσοντας το Q_n, βλέπουμε ότι γράφεται στη μορφή

$$Q_n = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i^2 b_j^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_i a_j b_j. \tag{1.20}$$

Η παράσταση αυτή, ως έχει, δεν μοιάζει να εξυπηρετεί το σκοπό μας. Μπορούμε όμως εδώ να χρησιμοποιήσουμε μια γενική αρχή που είναι συχνά χρήσιμη: την *αρχή της συμμετρίας*.

Πρακτικά, η ιδέα είναι να προσπαθήσουμε να γράψουμε την παράσταση μας με τέτοιο τρόπο που να φανερωθούν όλες οι συμμετρίες που υπάρχουν σε αυτήν. Εδώ, η συμμετρία μεταξύ i και j στο δεύτερο διπλό άθροισμα είναι ισχυρή και φανερή, ο συμμετρικός όμως ρόλος των i και j στο πρώτο διπλό άθροισμα δεν είναι αρκετά προφανής. Υπάρχει όμως κι εδώ συμμετρία, κάτι που μπορούμε να δούμε αν ξαναγράψουμε το Q_n στη μορφή

$$Q_n = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i^2 b_j^2 + a_j^2 b_i^2) - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_i a_j b_j. \tag{1.21}$$

Τώρα, τα δύο διπλά αθροίσματα είναι αμφότερα συμμετρικά ως προς i και j, και η νέα αναπαράσταση «φωνάζει» ότι πρέπει να συνδυάσουμε τα δύο αθροίσματα σε ένα και να κάνουμε την παραγοντοποίηση

$$Q_n = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\{a_i^2 b_j^2 - 2a_i b_j a_j b_i + a_j^2 b_i^2\right\} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2.$$

Όλα όσα είπαμε συνοψίζονται σε μια γραμμή. Έχουμε αποδείξει την *ταυτότητα του Lagrange*:

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 = \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2. \tag{1.22}$$

Οδηγηθήκαμε σε αυτήν την ταυτότητα από την επιθυμία μας να κατανοήσουμε το μη αρνητικό πολυώνυμο Q_n, από τη στιγμή όμως που θα μας δώσουν την (1.22), είναι πολύ εύκολο να την επαληθεύσουμε με κάποιους πολλαπλασιασμούς. Αυτό είναι ένα από τα παράδοξα των πολυωνυμικών ταυτοτήτων.

- Αξίζει να σημειώσουμε ότι η ανισότητα του Cauchy είναι άμεση συνέπεια της ταυτότητας του Lagrange, και, πραγματικά, η απόδειξη που ο ίδιος ο Cauchy επέλεξε να συμπεριλάβει στο βιβλίο που εξέδωσε το 1821 ήταν βασισμένη σε αυτήν την παρατήρηση.

- Εμείς, καταλήξαμε στην ταυτότητα του Lagrange (1.22) ξεκινώντας από την επιθυμία να κατανοήσουμε πλήρως την περίπτωση της ισότητας στην ανισότητα του Cauchy. Στην πορεία, ανακαλύψαμε μια ανεξάρτητη απόδειξη της ανισότητας του Cauchy, πρέπει όμως τώρα να κλείσουμε τον κύκλο δίνοντας την απάντηση στο Πρόβλημα 1.5.1.

Αν $(b_1, b_2, \ldots, b_n) \neq 0$, τότε υπάρχει κάποιος $b_k \neq 0$, και αν ισχύει ισότητα στην ανισότητα του Cauchy, τότε όλοι οι όροι στο δεξιό μέλος της ταυτότητας του Lagrange (1.22) πρέπει να είναι ίσοι με μηδέν. Αν θεωρήσουμε μόνο τους όρους που περιέχουν τον b_k, τότε βλέπουμε ότι

$$a_i b_k = a_k b_i \quad \text{για κάθε } 1 \leq i \leq n,$$

και, αν θέσουμε $\lambda = a_k/b_k$, τότε παίρνουμε

$$a_i = \lambda b_i \quad \text{για κάθε } 1 \leq i \leq n.$$

Δηλαδή, η ταυτότητα του Lagrange μας λέει ότι, για μη μηδενικές ακολουθίες, μπορούμε να έχουμε ισότητα στην ανισότητα του Cauchy αν και μόνο αν οι δύο ακολουθίες είναι ανάλογες. Έτσι, έχουμε μια πλήρη και ακριβή απάντηση στο πρώτο μας πρόβλημα.

Αυτή η ανάλυση της περίπτωσης της ισότητας υπογραμμίζει το γεγονός ότι η συμμετρική μορφή

$$Q_n = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i b_j - a_j b_i)^2$$

επιδέχεται δύο χρήσιμες ερμηνείες.

- Αρχικά την χρησιμοποιήσαμε σαν ένα μέτρο της διαφοράς ανάμεσα στα δύο μέλη της ανισότητας του Cauchy.
- Βλέπουμε όμως τώρα ότι ταυτόχρονα μετράει σε ποιό βαθμό είναι ανάλογα τα διανύσματα $(a_1, a_2, \ldots, a_n)$ και $(b_1, b_2, \ldots, b_n)$.

Ο Joseph Louis de Lagrange (1736-1813) ανακάλυψε και χρησιμοποίησε την ταυτότητα (1.22) στην περίπτωση $n = 3$, στη διάρκεια μιας μελέτης που έκανε για τη γεωμετρία των πυραμίδων. Αυτή η μελέτη εστιαζόταν σε προβλήματα στον τριδιάστατο χώρο, και ο Lagrange δεν ανέφερε το γεγονός ότι τα αντίστοιχα αποτελέσματα στην περίπτωση $n = 2$ ήταν γνωστά, ακόμα και στους μαθηματικούς της αρχαιότητας. Ειδικότερα, η διδιάστατη έκδοση της ταυτότητας (1.22) ήταν γνωστή στον Διόφαντο, κάτι που μπορεί κανείς να συμπεράνει από ένα πρόβλημα στα *Αριθμητικά* του. Ο Lagrange και ο προγενέστερος αυτού Pierre de Fermat (1601-1665) γνώριζαν καλά τα γραπτά του Διόφαντου.

1.6 Η σύντομη απόδειξη

Ο Augustin-Louis Cauchy (1789-1857) δημοσίευσε την περίφημη ανισότητά του το 1821, στην δεύτερη από τις δύο εργασίες του επί της θεωρίας των ανισοτήτων, οι οποίες απετέλεσαν το τελευταίο κομμάτι του βιβλίου του Cours d'Analyse Algébrique, ενός βιβλίου που ήταν το πρώτο αυστηρό σύγγραμμα Απειροστικού Λογισμού στην Ιστορία. Ο Cauchy δεν χρησιμοποίησε την ανισότητα στο βιβλίο του, με την εξαίρεση κάποιων ασκήσεων που χρησιμοποιούνταν σαν παραδείγματα. Η πρώτη φορά που η ανισότητα του Cauchy χρησιμοποιήθηκε ουσιαστικά ήταν το 1829, όταν ο Cauchy χρησιμοποίησε την ανισότητά του μελετώντας τη μέθοδο

του Newton για τον υπολογισμό των ριζών αλγεβρικών και υπερβατικών εξισώσεων. Αυτό το κενό των οκτώ ετών δίνει μια ενδιαφέρουσα εικόνα για τον ρυθμό ανάπτυξης της επιστήμης. Σήμερα, κάθε μήνα, υπάρχουν εκατοντάδες – ίσως χιλιάδες – νέες επιστημονικές δημοσιεύσεις στις οποίες η ανισότητα του Cauchy χρησιμοποιείται με τον έναν ή τον άλλον τρόπο.

Πολλές από αυτές τις εφαρμογές βασίζονται στο ολοκληρωτικό ανάλογο της ανισότητας του Cauchy, όπου τα αθροίσματα αντικαθίστανται από ολοκληρώματα:

$$\int_a^b f(x)g(x)dx \leq \left(\int_a^b f^2(x)dx\right)^{\frac{1}{2}} \left(\int_a^b g^2(x)dx\right)^{\frac{1}{2}}. \tag{1.23}$$

Η ανισότητα αυτή εμφανίστηκε για πρώτη φορά σε ένα Mémoire του Victor Yacovlevich Bunyakovsky το οποίο εκδόθηκε από την Ακαδημία Επιστημών της St. Petersburg το 1859. Ο Bunyakovsky (1804-1889) είχε σπουδάσει στο Παρίσι με τον Cauchy, και γνώριζε τη δουλειά του Cauchy στις ανισότητες. Την ήξερε μάλιστα τόσο καλά, που αναφερόταν στην κλασσική μορφή της ανισότητας του Cauchy για τα πεπερασμένα αθροίσματα χρησιμοποιώντας την έκφραση *διάσημη*. Επιπλέον, ο Bunyakovsky δεν δίσταζε μπροστά στη διαδικασία του ορίου: μέσα σε μια γραμμή, περνούσε από την ανισότητα του Cauchy για πεπερασμένα αθροίσματα στο συνεχές της ανάλογο (1.23). Εντελώς συμπτωματικά, αυτό το συνεχές ανάλογο ονομάζεται «ανισότητα **(C)**» στο Mémoire του Bunyakovsky, σαν να είχε στο μυαλό του τον Cauchy.

Το Mémoire του Bunyakovsky ήταν γραμμένο στα Γαλλικά, φαίνεται όμως ότι δεν κυκλοφόρησε ευρέως στη Δυτική Ευρώπη. Ειδικότερα, φαίνεται ότι δεν είχε γίνει γνωστό στο Göttingen το 1885, την περίοδο που ο Hermann Amandus Schwarz (1843-1921) ολοκλήρωνε τη θεμελιώδη μελέτη του για τις ελάχιστες επιφάνειες.

Στην πορεία της μελέτης του, ο Schwarz χρειάστηκε ένα διδιάστατο ολοκληρωτικό ανάλογο της ανισότητας του Cauchy. Πιο συγκεκριμένα, χρειαζόταν να δείξει ότι αν $S \subset \mathbb{R}^2$ και αν $f : S \to \mathbb{R}$ και $g : S \to \mathbb{R}$, τότε τα διπλά ολοκληρώματα

$$A = \iint_S f^2 dxdy, \quad B = \iint_S fg dxdy, \quad C = \iint_S g^2 dxdy$$

ικανοποιούν την ανισότητα

$$|B| \leq \sqrt{A} \cdot \sqrt{C}, \tag{1.24}$$

και, επίσης, ο Schwarz χρειαζόταν να ξέρει ότι η ανισότητα είναι γνήσια εκτός αν οι συναρτήσεις f και g είναι ανάλογες.

Η χρήση της ανισότητας του Cauchy για την απόδειξη αυτού του αποτελέσματος είναι προβληματική, για διάφορους λόγους. Ένας από αυτούς είναι ότι η πληροφορία ότι μια διακριτή ανισότητα είναι γνήσια χάνεται όταν περνάμε στα ολοκληρώματα μέσω ορίων. Έτσι, ο Schwarz έπρεπε να ψάξει για μια εναλλακτική οδό, και, αντιμετωπίζοντας αυτήν την ανάγκη, ανακάλυψε μια απόδειξη που η γοητεία της έχει αντέξει στο χρόνο.

Ο Schwarz βάσισε την απόδειξή του σε μια έκπληκτική παρατήρηση. Συγκεκριμένα, παρατήρησε ότι το πραγματικό πολυώνυμο

$$p(t) = \iint_S (tf(x,y) + g(x,y))^2 \, dxdy = At^2 + 2Bt + C$$

είναι πάντα μη αρνητικό και, επιπλέον, το $p(t)$ είναι γνήσια θετικό εκτός αν οι f και g είναι ανάλογες. Αυτό δείχνει ότι η διακρίνουσά του είναι μικρότερη ή ίση του μηδενός, άρα οι συντελεστές του ικανοποιούν την $B^2 \leq AC$, και αν οι f και g δεν είναι ανάλογες, τότε έχουμε την γνήσια ανισότητα $B^2 < AC$. Έτσι, κάνοντας χρήση μιάς απλής αλγεβρικής παρατήρησης, ο Schwarz εξασφάλισε όλα όσα χρειαζόταν να ξέρει.

Η απόδειξη του Schwarz απαιτεί τη σοφία να θεωρήσουμε το πολυώνυμο $p(t)$, αν όμως προσπεράσουμε αυτό το βήμα, τότε η απόδειξη είναι εκπληκτικά σύντομη. Επιπλέον, το επιχείρημα του Schwarz, χωρίς καμία απολύτως τροποποίηση, δίνει μια απόδειξη της μορφής (1.16) της ανισότητας του Cauchy για τα εσωτερικά γινόμενα, διευκρινίζοντας κι εκεί την περίπτωση της ισότητας. Δεν είναι λοιπόν παράξενο το γεγονός ότι όλοι οι συγγραφείς μαθηματικών διδακτικών βιβλίων προτιμούν να αποδεικνύουν την ανισότητα του Cauchy με το επιχείρημα του Schwarz, παρόλο που θυμίζει λίγο τον ταχυδακτυλουργό που βγάζει το λαγό (εδώ, το πολυώνυμο) από το καπέλο. Ας δούμε την απόδειξη στην περίπτωση των n–άδων $a_1, \ldots, a_n$ και $b_1, \ldots, b_n$: Θεωρούμε το τριώνυμο

$$f(x) = (a_1x + b_1)^2 + (a_2x + b_2)^2 + \cdots + (a_nx + b_n)^2. \tag{1.25}$$

Το τριώνυμο αυτό είναι μη αρνητικό, ως άθροισμα τετραγώνων, και γράφεται στη μορφή

$$f(x) = (a_1^2 + a_2^2 + \cdots + a_n^2)x^2 + 2(a_1b_1 + a_2b_2 + \cdots + a_nb_n)x + (b_1^2 + b_2^2 + \cdots + b_n^2). \tag{1.26}$$

Η διακρίνουσα του τριωνύμου ισούται με

$$\Delta = 4(a_1b_1 + a_2b_2 + \cdots + a_nb_n) - 4(a_1^2 + a_2^2 + \cdots + a_n^2)(b_1^2 + b_2^2 + \cdots + b_n^2). \tag{1.27}$$

Το τριώνυμο, ως μη αρνητικό, θα πρέπει να έχει διακρίνουσα $\Delta \leq 0$. Αυτό όμως – από την (1.27) – μας δίνει ακριβώς εκείνο που θέλαμε να δείξουμε, δηλαδή την ανισότητα Cauchy!

Αν ισχύει ισότητα στην ανισότητα Cauchy, τότε θα είναι $f(x) = 0$ για κάποιον x. Λόγω της (1.25), συμπεραίνουμε τότε ότι

$$a_1x + b_1 = a_2x + b_2 = \cdots = a_nx + b_n = 0.$$

Έπεται ότι ισχύει ισότητα στην ανισότητα Cauchy αν και μόνο αν οι n–άδες $a_1, \ldots, a_n$ και $b_1, \ldots, b_n$ είναι ανάλογες.

Κεφάλαιο 2

Η δεύτερη ανισότητα του Cauchy

2.1 Η ανισότητα αριθμητικού–γεωμετρικού μέσου

Για την απόδειξη της ανισότητας Cauchy βασιστήκαμε στη στοιχειώδη ανισότητα

$$xy \leq \frac{x^2}{2} + \frac{y^2}{2} \quad \text{για κάθε } x, y \in \mathbb{R}. \tag{2.1}$$

- Για μη αρνητικούς x και y, η (2.1) απλά λέει ότι το εμβαδόν ενός ορθογωνίου με πλευρές x και y δεν είναι ποτέ μεγαλύτερο από τον μέσο όρο των εμβαδών των δύο τετραγώνων με πλευρές x και y.

Αν αντικαταστήσουμε τους x και y με τις τετραγωνικές τους ρίζες, τότε η ανισότητα (2.1) μας δίνει

$$4\sqrt{xy} < 2x + 2y \quad \text{για όλους τους μη αρνητικούς } x \neq y, \tag{2.2}$$

και αυτή η ανισότητα έχει μια πολύ πιό πλούσια ερμηνεία. Θεωρούμε το σύνολο όλων των ορθογωνίων με εμβαδόν A και μήκη πλευρών x και y. Αφού $A = xy$, η ανισότητα (2.2) μας λέει ότι

- ένα τετράγωνο με μήκος πλευράς $s = \sqrt{A}$ πρέπει να έχει τη μικρότερη δυνατή περίμετρο μεταξύ όλων των ορθογωνίων με εμβαδόν A.

Ισοδύναμα, η ανισότητα μας λέει ότι

- μεταξύ όλων των ορθογωνίων με περίμετρο p, το τετράγωνο με πλευρά $s = p/4$, και μόνον αυτό, επιτυγχάνει το μέγιστο εμβαδόν.

Έτσι, η ανισότητα (2.2) δεν είναι τίποτε άλλο παρά η «ορθογώνια έκδοση» της φημισμένης *ισοπεριμετρικής ανισότητας*, που λέει ότι

- μεταξύ όλων των επίπεδων χωρίων με περίμετρο p, ο δίσκος με περιφέρεια p έχει το μέγιστο εμβαδόν.

Βλέπουμε τώρα καθαρά γιατί η $xy \leq x^2/2 + y^2/2$ μπορεί να είναι τόσο ισχυρή. Είναι μέρος μιας μεγάλης αλυσίδας αποτελεσμάτων που συνδέουν την συμμετρία με τα προβλήματα μεγίστου–ελαχίστου.

Ένα πλεονέκτημα της ισοπεριμετρικής ερμηνείας της ανισότητας $\sqrt{xy} \leq (x + y)/2$ είναι ότι δίνει ώθηση στη διαίσθησή μας. Μπορούμε εύκολα να προχωρήσουμε σε εικασίες για γενικεύσεις της ανισότητας στις τρείς ή περισσότερες διαστάσεις. Η πλέον φυσιολογική τέτοια γενίκευση είναι ο ισχυρισμός ότι ο κύβος στον $\mathbb{R}^3$ έχει τον μεγαλύτερο όγκο ανάμεσα μεταξύ όλων των ορθογωνίων παραλληλεπιπέδων που έχουν δοσμένο εμβαδόν επιφάνειας. Ο άμεσος όμως στόχος μας είναι μια άλλη γενίκευση – που έχει πολλές εφαρμογές. Ένα ορθογώνιο στον $\mathbb{R}^n$ έχει 2^n κορυφές, και καθεμία από αυτές τις κορυφές ανήκει σε n ακμές του ορθογωνίου. Αν ονομάσουμε $a_1, a_2, \ldots, a_n$ τα μήκη αυτών των ακμών, τότε η ίδια ισοπεριμετρική διαίσθηση που χρησιμοποιήσαμε για τα τετράγωνα και τους κύβους, υποδεικνύει ότι ο n-κύβος με ακμή μήκους S/n θα έχει τον μεγαλύτερο όγκο μεταξύ των ορθογωνίων για τα οποία $a_1 + a_2 + \cdots + a_n = S$. Η ανισότητα αριθμητικού–γεωμετρικού μέσου μεταφράζει αυτήν την γεωμετρική εικασία στη γλώσσα των αριθμητικών και γεωμετρικών μέσων.

Θεώρημα 2.1.1 (ανισότητα αριθμητικού–γεωμετρικού μέσου). *Δείξτε ότι για κάθε ακολουθία μη αρνητικών πραγματικών αριθμών $a_1, a_2, \ldots, a_n$ ισχύει η ανισότητα*

$$(a_1 a_2 \cdots a_n)^{1/n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}. \tag{2.3}$$

Για $n = 2$, η ανισότητα (2.3) προκύπτει κατ' ευθείαν από την στοιχειώδη ανισότητα $\sqrt{xy} \leq (x + y)/2$ που μόλις συζητήσαμε. Η παρατήρηση του Cauchy) ήταν ότι αν εφαρμόσουμε την ίδια ανισότητα δύο φορές, παίρνουμε

$$(a_1 a_2 a_3 a_4)^{1/4} \leq \frac{(a_1 a_2)^{1/2} + (a_3 a_4)^{1/2}}{2} \leq \frac{a_1 + a_2 + a_3 + a_4}{4}. \tag{2.4}$$

Αυτή η ανισότητα επιβεβαιώνει την εικασία (2.3) στην περίπτωση $n = 4$, και εφαρμόζοντας τη νέα ανισότητα (2.4) μαζί με την $\sqrt{xy} \leq (x + y)/2$ βλέπουμε ότι

$$(a_1 a_2 \cdots a_8)^{1/8} \leq \frac{(a_1 a_2 a_3 a_4)^{1/4} + (a_5 a_6 a_7 a_8)^{1/4}}{2} \leq \frac{a_1 + a_2 + \cdots + a_8}{8},$$

δηλαδή την εικασία (2.3) για $n = 8$.

Επαναλαμβάνοντας το ίδιο επιχείρημα k φορές (ή χρησιμοποιώντας επαγωγή) δείχνουμε ότι

$$(a_1 a_2 \cdots a_{2^k})^{1/2^k} \leq (a_1 + a_2 + \cdots + a_{2^k})/2^k \quad \text{για κάθε } k \geq 1. \tag{2.5}$$

Με άλλα λόγια, έχουμε αποδείξει τη ζητούμενη ανισότητα για κάθε $n = 2^k$. Αυτό που μένει είναι να βρούμε κάποιον τρόπο για να γεμίσουμε τα κενά μεταξύ των δυνάμεων του 2.

Το φυσιολογικό σχέδιο είναι να πάρουμε κάποιον $n < 2^k$ και να βρούμε κάποιον τρόπο ώστε χρησιμοποιώντας n δοθέντες αριθμούς $a_1, a_2, \ldots, a_n$ να ορίσουμε μια

μακρύτερη ακολουθία $\alpha_1, \alpha_2, \ldots, \alpha_{2^k}$ για την οποία θα μπορούμε να εφαρμόσουμε την ανισότητα (2.5). Η αποτελεσματική επιλογή των όρων της νέας ακολουθίας απαιτεί έμπνευση: μια ιδέα που τελικά δουλεύει είναι να θέσουμε $\alpha_i = a_i$ για $1 \le i \le n$ και

$$\alpha_i = \frac{a_1 + a_2 + \cdots + a_n}{n} \equiv A \quad \text{για } n < i \le 2^k.$$

Με άλλα λόγια, απλώς επεκτείνουμε την δοθείσα ακολουθία $\{a_i : 1 \le i \le n\}$ με αρκετά αντίγραφα του αριθμητικού μέσου A ώστε να πάρουμε μια ακολουθία $\{\alpha_i : 1 \le i \le 2^k\}$ που έχει μήκος ίσο με 2^k.

Ο αριθμητικός μέσος A χρησιμοποιείται $2^k - n$ φορές στην διευρυμένη ακολουθία $\{\alpha_i\}$, αν λοιπόν εφαρμόσουμε την ανισότητα (2.5) για την $\{\alpha_i\}$, προκύπτει η

$$\left\{a_1 a_2 \cdots a_n \cdot A^{2^k - n}\right\}^{1/2^k} \le \frac{a_1 + a_2 + \cdots + a_n + (2^k - n)A}{2^k} = \frac{2^k A}{2^k} = A.$$

Τώρα, αν φέρουμε όλες τις δυνάμεις του A στο δεξιό μέλος, βλέπουμε ότι

$$(a_1 a_2 \cdots a_n)^{1/2^k} \le A^{n/2^k},$$

και, αν υψώσουμε τα δύο μέλη στη δύναμη $2^k/n$, πετυχαίνουμε ακριβώς τον στόχο μας, την ανισότητα

$$(a_1 a_2 \cdots a_n)^{1/n} \le \frac{a_1 + a_2 + \cdots + a_n}{n}. \tag{2.6}$$

Μια δεύτερη απόδειξη με επαγωγή

Δίνουμε τώρα μια δεύτερη απόδειξη της (2.3) με επαγωγή «από το n στο $n+1$».

Βοηθητική πρόταση: *Για κάθε $n \in \mathbb{N}$ και $a, b > 0$ ισχύει*

$$b^{n+1} \ge (n+1)a^n b - n a^{n+1}. \tag{2.7}$$

Για την απόδειξη θα χρησιμοποιήσουμε την αρχή της μαθηματικής επαγωγής. Για $n = 1$ ζητάμε

$$b^2 \ge 2ab - a^2,$$

που ισχύει λόγω της $(a-b)^2 \ge 0$.

Υποθέτουμε ότι η (2.7) ισχύει για τον φυσικό n και θα δείξουμε ότι ισχύει και για τον φυσικό $n+1$, δηλαδή ότι ισχύει

$$b^{n+2} \ge (n+2)a^{n+1}b - (n+1)a^{n+2}. \tag{2.8}$$

Επειδή $b^{n+2} = b \cdot b^{n+1}$, πολλαπλασιάζουμε με b την (2.7) και έχουμε

$$b^{n+2} \ge (n+1)a^n b^2 - nba^{n+1}. \tag{2.9}$$

Από τις (2.8) και (2.9) καταλαβαίνουμε ότι για να ισχύει η (2.9) αρκεί να ισχύει η

$$(n+1)a^n b^2 - nba^{n+1} \ge (n+2)a^{n+1}b - (n+1)a^{n+2}. \tag{2.10}$$

Η σχέση αυτή – αν απαλείψουμε τον $a^n > 0$ – ισοδύναμα γράφεται

$$(n+1)b^2 - nab \geq (n+2)ab - (n+1)a^2, \tag{2.11}$$

ή, αλλιώς,

$$(n+1)(a^2+b^2) \geq 2(n+1)ab, \tag{2.12}$$

που ισχύει. Άρα, σύμφωνα με την αρχή της μαθηματικής επαγωγής η βοηθητική μας πρόταση (2.7) ισχύει για όλα τα $n \in \mathbb{N}$.

Δεύτερη απόδειξη της ανισότητας αριθμητικού–γεωμετρικού μέσου. Έχουμε δεί ότι η ανισότητα (2.3) για $n = 2$. Υποθέτουμε ότι η ανισότητα ισχύει για τον φυσικό n, δηλαδή ισχύει

$$a_1 + a_2 + \cdots + a_n \geq n(a_1 a_2 \cdots a_n)^{1/n}, \tag{2.13}$$

και θα δείξουμε ότι ισχύει για τον φυσικό $n+1$. Η βοηθητική μας πρόταση (2.7) για $b = (a_{n+1})^{1/(n+1)}$ και $a = (a_1 a_2 \cdots a_n)^{1/n(n+1)}$ γίνεται

$$a_{n+1} \geq (n+1)(a_1 a_2 \cdots a_n)^{1/(n+1)}(a_{n+1})^{1/(n+1)} - n(a_1 a_2 \cdots a_n)^{1/n}, \tag{2.14}$$

δηλαδή

$$a_{n+1} \geq (n+1)(a_1 a_2 \cdots a_{n+1})^{1/(n+1)} - n(a_1 a_2 \cdots a_n)^{1/n}. \tag{2.15}$$

Προσθέτοντας κατά μέλη τις (2.13) και (2.15) έχουμε

$$a_1 + a_2 + \cdots + a_{n+1} \geq (n+1)(a_1 a_2 \cdots a_{n+1})^{1/(n+1)}, \tag{2.16}$$

δηλαδή η (2.3) ισχύει και για τον φυσικό $n+1$. Σύμφωνα με την αρχή της μαθηματικής επαγωγής, η ανισότητα ισχύει για κάθε $n \in \mathbb{N}$.

2.2 Η γενική ανισότητα αριθμητικού–γεωμετρικού μέσου

Η ανισότητα αριθμητικού–γεωμετρικού μέσου (2.6) έχει το προσόν της *αυτογενίκευσης*. Υποδεικνύει σχεδόν αυτόματα ένα νέο αποτέλεσμα, το οποίο καλύπτει περιπτώσεις που η ίδια άφηνε άθικτες.

Θεώρημα 2.2.1 (ανισότητα αριθμητικού–γεωμετρικού μέσου με ρητά βάρη). *Έστω $p_1, p_2, \ldots, p_n$ μη αρνητικοί ρητοί αριθμοί με άθροισμα 1. Αν $a_1, a_2, \ldots, a_n$ είναι μη αρνητικοί πραγματικοί αριθμοί, τότε*

$$a_1^{p_1} a_2^{p_2} \cdots a_n^{p_n} \leq p_1 a_1 + p_2 a_2 + \cdots + p_n a_n. \tag{2.17}$$

Από τη στιγμή που θα αναρωτηθούμε ποιόν ρόλο παίζει η υπόθεση ότι οι p_j είναι ρητοί, βρισκόμαστε ήδη πολύ κοντά στη λύση του προβλήματος. Αν θεωρήσουμε έναν φυσικό αριθμό M ώστε για κάθε j να μπορούμε να γράψουμε $p_j = k_j/M$ για κάποιον φυσικό αριθμό k_j, τότε βλέπουμε ότι η φαινομενικά πολύ

πιό γενική έκδοση (2.17) της ανισότητας αριθμητικού–γεωμετρικού μέσου προκύπτει από την αρχική έκδοση (2.3) αν την εφαρμόσουμε για μια ακολουθία που έχει μήκος M και εμφανίζει πολλές επαναλήψεις. Θεωρούμε την ακολουθία που περιέχει k_j αντίγραφα του a_j για κάθε $1 \leq j \leq n$ και μετά εφαρμόζουμε την κλασική ανισότητα αριθμητικού–γεωμετρικού μέσου (2.3).

Υπάρχει όμως ακόμα μία παρατήρηση που μπορούμε να κάνουμε. Από τη στιγμή που η (2.17) έχει αποδειχτεί για ρητά p_j, η ίδια ανισότητα προκύπτει για γενικές τιμές των p_j, «παίρνοντας όρια». Για να δώσουμε λεπτομερή απόδειξη, πρώτα επιλέγουμε μια ακολουθία ρητών αριθμών $p_j(t)$, $j = 1, 2, \ldots, n$ και $t = 1, 2, \ldots$ για την οποία έχουμε

$$p_j(t) \geq 0, \quad \sum_{j=1}^{n} p_j(t) = 1, \quad \text{και} \quad \lim_{t\to\infty} p_j(t) = p_j.$$

Κατόπιν, εφαρμόζουμε την (2.7) για τις n-άδες $(p_1(t), p_2(t), \ldots, p_n(t))$, και, τέλος, αφήνουμε το n να πάει στο άπειρο και παίρνουμε το ζητούμενο.

- Η τεχνική της απόδειξης μιας ανισότητας πρώτα για ρητούς και μετά για πραγματικούς με «πέρασμα στο όριο» είναι πολύ συχνά χρήσιμη, έχει όμως κάποια μειονεκτήματα. Για παράδειγμα, το γνήσιο μιας ανισότητας μπορεί να χαθεί όταν περνάμε στο όριο, χρησιμοποιώντας λοιπόν αυτήν την τεχνική δεν έχουμε πλήρη κατανόηση για τις περιπτώσεις ισότητας. Μερικές φορές αυτή η απώλεια δεν είναι σημαντική, για ένα εργαλείο όμως τόσο σημαντικό όσο η ανισότητα αριθμητικού–γεωμετρικού μέσου, οι συνθήκες κάτω από τις οποίες γίνεται ισότητα είναι σημαντικές. Θα προτιμούσαμε μια απόδειξη η οποία να αντιμετωπίζει όλους τους όρους της ανισότητας με ενιαίο τρόπο, κάτι που η μέθοδος της προσέγγισης με ρητούς δεν μπορεί να επιτύχει.

Το όνειρο του Pólya

Υπάρχουν πολλές αποδείξεις της ανισότητας αριθμητικού–γεωμετρικού μέσου. Μια ιδιαίτερα γοητευτική απόδειξη είναι αυτή του George Pólya, που ανέφερε κάποτε πως η απόδειξη τον επισκέφτηκε σε κάποιο όνειρο. Πραγματικά, όταν ρωτήθηκε γι' αυτήν την απόδειξη, χρόνια αργότερα, ο Pólya απάντησε ότι

> «*ήταν τα καλύτερα Μαθηματικά που είχε ποτέ ονειρευτεί*».

Όπως ο Cauchy, ο Pólya αρχίζει την απόδειξή του με μια απλή παρατήρηση για μια μη αρνητική συνάρτηση, μόνο που ο Pólya θεωρεί την συνάρτηση $x \mapsto e^x$ αντί για την συνάρτηση $x \mapsto x^2$. Η γραφική παράσταση της $y = e^x$ δείχνει την ιδιότητα της $y = e^x$ που είναι το κλειδί της απόδειξης του Pólya. Συγκεκριμένα, δείχνει ότι η εφαπτόμενη ευθεία $y = 1 + x$ βρίσκεται κάτω από την καμπύλη $y = e^x$. Με άλλα λόγια, έχουμε την ανισότητα

$$1 + x \leq e^x \quad \text{για κάθε } x \in \mathbb{R}. \tag{2.18}$$

Θεώρημα 2.2.2 (η γενική ανισότητα αριθμητικού–γεωμετρικού μέσου). *Αν $p_1, p_2, \ldots, p_n$ είναι μη αρνητικοί πραγματικοί αριθμοί με άθροισμα*

1 *και αν* $a_1, a_2, \ldots, a_n$ *είναι μη αρνητικοί πραγματικοί αριθμοί, τότε*

$$a_1^{p_1} a_2^{p_2} \cdots a_n^{p_n} \leq p_1 a_1 + p_2 a_2 + \cdots + p_n a_n. \tag{2.19}$$

Στην ανισότητα αριθμητικού–γεωμετρικού μέσου (2.19), το αριστερό μέλος περιέχει ένα γινόμενο όρων, και η αναλυτική ανισότητα $1+x \leq e^x$ φράσσει αμέσως ένα τέτοιο γινόμενο από το εκθετικό ενός αθροίσματος. Επιπλέον, υπάρχουν δύο τρόποι για να εκμεταλλευτούμε αυτήν την δυνατότητα. Θα μπορούσαμε να γράψουμε τους a_k στη μορφή $1 + x_k$ και μετά να εφαρμόσουμε την (2.18), ή θα μπορούσαμε να τροποποιήσουμε την (2.18) ώστε να εφαρμόζεται απευθείας για τους a_k. Στην πράξη, θα μπορούσαμε να διερευνήσουμε και τους δύο τρόπους, για την ώρα όμως θα επικεντρωθούμε στο δεύτερο σχέδιο.

Αν κάνουμε την αλλαγή μεταβλητής $x \mapsto x-1$, τότε η ανισότητα (2.18) παίρνει τη μορφή

$$x \leq e^{x-1} \quad \text{για κάθε } x \in \mathbb{R}, \tag{2.20}$$

και αν εφαρμόσουμε αυτήν την ανισότητα για τους a_k, $k = 1, 2, \ldots, n$, παίρνουμε

$$a_k \leq e^{a_k - 1} \quad \text{και} \quad a_k^{p_k} \leq e^{p_k a_k - p_k}.$$

Πολλαπλασιάζοντας, βλέπουμε ότι ο γεωμετρικός μέσος $G = a_1^{p_1} a_2^{p_2} \cdots a_n^{p_n}$ φράσσεται από την ποσότητα

$$R(a_1, a_2, \ldots, a_n) = \exp\left(\left\{\sum_{k=1}^{n} p_k a_k\right\} - 1\right). \tag{2.21}$$

Η ανισότητα $G \leq R$ είναι ενδιαφέρουσα, αυτό όμως που θα θέλαμε να δούμε είναι πώς συγκρίνεται η ποσότητα R με τον αριθμητικό μέσο

$$A = p_1 a_1 + p_2 a_2 + \cdots + p_n a_n.$$

Αυτό είναι το σημείο στο οποίο το πρόβλημα γίνεται ενδιαφέρον.

Το πρώτο πράγμα που παρατηρούμε είναι ότι, από την ανισότητα $A \leq e^{A-1}$, ο R είναι *άνω φράγμα* και για τον αριθμητικό μέσο A. Δηλαδή, έχουμε τη διπλή ανισότητα

$$\max\{a_1^{p_1} \cdots a_n^{p_n}, p_1 a_1 + \cdots + p_n a_n\} \leq \exp\left(\left\{\sum_{k=1}^{n} p_k a_k\right\} - 1\right). \tag{2.22}$$

Οδηγούμαστε έτσι σε ένα εκ πρώτης όψεως παράδοξο ερώτημα:

- Πώς μπορούμε να συγκρίνουμε δύο ποσότητες όταν το μόνο που γνωρίζουμε είναι ένα κοινό άνω φράγμα τους;

Θα μπορούσαμε να αποθαρρυνθούμε σε αυτό το σημείο, παρατηρούμε όμως το εξής. Η ανισότητα (2.22) δίνει μια σχέση ανάμεσα στα A και G στην ειδική περίπτωση που ένας από αυτούς τους δύο αριθμούς ισούται με R! Παρατηρήστε ότι αν $\max\{A, G\} = R$ και $R = A$, τότε $G \leq A$.

Αν κάνουμε αυτήν την παρατήρηση, τότε η γνωστή πλέον τεχνική της κανονικοποίησης μας έρχεται στο νού. Αν θεωρήσουμε τις νέες μεταβλητές α_k, $k = 1, 2, \ldots, n$, που ορίζονται από τους λόγους

$$\alpha_k = \frac{a_k}{A} \quad \text{όπου } A = p_1 a_1 + p_2 a_2 + \cdots + p_n a_n,$$

και αν εφαρμόσουμε την ανισότητα (2.21) γι' αυτές τις νέες μεταβλητές, τότε παίρνουμε

$$\max\left\{1, \left(\frac{a_1}{A}\right)^{p_1} \left(\frac{a_2}{A}\right)^{p_2} \cdots \left(\frac{a_n}{A}\right)^{p_n}\right\} \leq \exp\left(\left\{\sum_{k=1}^{n} p_k \frac{a_k}{A}\right\} - 1\right) = 1,$$

και συνεπώς,

$$\left(\frac{a_1}{A}\right)^{p_1} \left(\frac{a_2}{A}\right)^{p_2} \cdots \left(\frac{a_n}{A}\right)^{p_n} \leq 1.$$

Φέρνοντας στο δεξιό μέλος τις δυνάμεις του A και χρησιμοποιώντας την υπόθεση ότι $p_1 + p_2 + \cdots + p_n = 1$, καταλήγουμε στην γενική ανισότητα αριθμητικού–γεωμετρικού μέσου (2.19).

Κοιτάζοντας πίσω στα βήματα αυτής της απόδειξης της ανισότητας (2.19), βλέπουμε ότι μπορούμε εύκολα να προσδιορίσουμε τις προϋποθέσεις κάτω από τις οποίες ισχύει σαν ισότητα. Εξετάζοντας το πρώτο βήμα βλέπουμε ότι έχουμε

$$\frac{a_k}{A} < e^{(a_k/A)-1} \quad \text{εκτός αν } \frac{a_k}{A} = 1, \tag{2.23}$$

και έχουμε πάντα

$$\frac{a_k}{A} \leq e^{(a_k/A)-1},$$

άρα

$$\left(\frac{a_1}{A}\right)^{p_1} \left(\frac{a_2}{A}\right)^{p_2} \cdots \left(\frac{a_n}{A}\right)^{p_n} < \exp\left(\left\{\sum_{k=1}^{n} p_k \frac{a_k}{A}\right\} - 1\right) = 1, \tag{2.24}$$

εκτός αν $a_k = A$ για κάθε $k = 1, 2, \ldots, n$. Με άλλα λόγια, βλέπουμε ότι ισχύει ισότητα στην ανισότητα αριθμητικού–γεωμετρικού μέσου (2.19) αν και μόνο αν

$$a_1 = a_2 = \cdots = a_n.$$

Ποιός ήταν ο George Pólya

Ο George Pólya (1887-1985) ήταν ένας από τους πιό σημαντικούς Μαθηματικούς του 20ου αιώνα, ο οποίος είχε ιδιαίτερη επιρροή στον τρόπο με τον οποίο βλέπουμε σήμερα τις διαδικασίες της διδασκαλίας και της μάθησης. Ο Pólya θεωρούσε ότι η διαδικασία της επίλυσης προβλημάτων είναι θεμελιώδης ανθρώπινη δραστηριότητα – μια δραστηριότητα γεμάτη συγκινήσεις, δημιουργικότητα και αγάπη για τη ζωή. Είχε επίσης σκεφτεί πολύ πάνω στον τρόπο με τον οποίο θα μπορούσε κανείς να γίνει πιό ικανός λύτης μαθηματικών προβλημάτων και στον τρόπο με τον οποίο θα μπορούσε κανείς να διδάξει άλλους να κάνουν το ίδιο.

Ο Pólya συνόψισε τις σκέψεις του σε άρκετά βιβλία, με πιό γνωστό το «**Πώς να το λύσεις**». Το κεντρικό σημείο στα γραπτά του Pólya είναι ότι συχνά

μπορούμε να κάνουμε βήματα για την επίλυση ενός μαθηματικού προβλήματος ρωτώντας μερικές *γενικές ερωτήσεις κοινής λογικής*. Πολλές από τις ερωτήσεις του Pólya φαίνονται προφανείς σε έναν «γεννημένο λύτη προβλημάτων», όμως, ο χρόνος δείχνει ότι περιέχουν αρκετή σοφία.

- Μπορείς να λύσεις το πρόβλημά σου σε μια ειδική περίπτωση;
- Μπορείς να συσχετίσεις το πρόβλημά σου με κάποιο παρόμοιο στο οποίο η απάντηση είναι ήδη γνωστή;
- Μπορείς να υπολογίσεις κάτι που να σχετίζεται με αυτό που θέλεις πραγματικά να υπολογίσεις;

Κεφάλαιο 3

Κυρτότητα

3.1 Η ανισότητα του Jensen

Μια συνάρτηση $f : [a, b] \to \mathbb{R}$ λέγεται *κυρτή* αν για κάθε $x, y \in [a, b]$ και για κάθε $0 \leq p \leq 1$ έχουμε

$$f(px + (1-p)y) \leq pf(x) + (1-p)f(y). \tag{3.1}$$

Η ανισότητα του Jensen συνδέει κατά θεμελιώδη τρόπο την έννοια της κυρτότητας με την θεωρία των ανισοτήτων.

Θεώρημα 3.1.1 (ανισότητα του Jensen). *Υποθέτουμε ότι η $f : [a, b] \to \mathbb{R}$ είναι κυρτή συνάρτηση και ότι οι μη αρνητικοί αριθμοί p_j, $j = 1, 2, \ldots, n$ ικανοποιούν την σχέση*

$$p_1 + p_2 + \cdots + p_n = 1.$$

Τότε, για κάθε $x_j \in [a, b]$, $j = 1, 2, \ldots, n$ έχουμε

$$f\left(\sum_{j=1}^{n} p_j x_j\right) \leq \sum_{j=1}^{n} p_j f(x_j). \tag{3.2}$$

Στην περίπτωση $n = 2$ η ανισότητα του Jensen (3.2) είναι ο ορισμός της κυρτής συνάρτησης, το ένστικτό μας λοιπόν υποδεικνύει να ψάξουμε για μια απόδειξη με επαγωγή. Αυτή η προσέγγιση απαιτεί να συσχετίσουμε μέσους όρους μήκους $n-1$ με μέσους όρους μήκους n, και αυτό μπορεί να γίνει με διάφορους τρόπους.

Μια φυσιολογική ιδέα είναι να απομονώσουμε τον τελευταίο προσθετέο και να κανονικοποιήσουμε ότι απομένει. Πιο συγκεκριμένα, παρατηρούμε πρώτα ότι δεν χάνουμε σε γενικότητα αν υποθέσουμε ότι $0 < p_n < 1$ και, σε αυτήν την περίπτωση, μπορούμε να γράψουμε

$$\sum_{j=1}^{n} p_j x_j = p_n x_n + (1 - p_n) \sum_{j=1}^{n-1} \frac{p_j}{1 - p_n} x_j.$$

Τώρα, από τον ορισμό της κυρτής συνάρτησης και από την επαγωγική υπόθεση – με αυτή τη σειρά – βλέπουμε ότι

$$\begin{aligned} f\left(\sum_{j=1}^{n} p_j x_j\right) &\leq p_n f(x_n) + (1-p_n) f\left(\sum_{j=1}^{n} \frac{p_j}{1-p_n} x_j\right) \\ &\leq p_n f(x_n) + (1-p_n) \sum_{j=1}^{n-1} \frac{p_j}{1-p_n} f(x_j) \\ &= \sum_{j=1}^{n} p_j f(x_j). \end{aligned}$$

Αυτή η ανισότητα ολοκληρώνει το επαγωγικό βήμα.

Υπάρχουν πολλές εφαρμογές της ανισότητας του Jensen, και κάποιες από αυτές βασίζονται στην κατανόηση των συνθηκών κάτω από τις οποίες έχουμε ισότητα. Εδώ, θα μας φανεί χρήσιμο να περιοριστούμε σε εκείνες τις συναρτήσεις $f : [a, b] \to \mathbb{R}$ με την ιδιότητα: για κάθε $x \neq y \in [a, b]$ και για κάθε $0 < p < 1$ ισχύει η γνήσια ανισότητα

$$f(px + (1-p)y) < pf(x) + (1-p)f(y). \tag{3.3}$$

Αυτές οι συναρτήσεις ονομάζονται *γνησίως κυρτές*.

Θεώρημα 3.1.2 (η περίπτωση της ισότητας στην ανισότητα του Jensen). *Έστω $f : [a, b] \to \mathbb{R}$ μια γνησίως κυρτή συνάρτηση και ας υποθέσουμε ότι*

$$f\left(\sum_{j=1}^{n} p_j x_j\right) = \sum_{j=1}^{n} p_j f(x_j) \tag{3.4}$$

όπου οι θετικοί αριθμοί p_j, $j = 1, 2, \ldots, n$ έχουν άθροισμα $p_1 + p_2 + \cdots + p_n = 1$. Τότε, πρέπει να ισχύει

$$x_1 = x_2 = \cdots = x_n. \tag{3.5}$$

Η απόδειξη είναι πάλι απλή, όπως και η απόδειξη της ανισότητας του Jensen, η σημασία όμως του αποτελέσματος είναι μεγάλη.

- *Για πολλές ανισότητες, ανακαλύπτουμε πότε ισχύει ισότητα παίρνοντας την απόδειξη της ανισότητας και ακολουθώντας την από το τέλος προς την αρχή.* Αυτή η προσέγγιση δουλεύει τέλεια στην περίπτωση της ανισότητας του Jensen, η λογική όμως του επιχειρήματος που θα χρησιμοποιήσουμε είναι κάπως διαφορετική.

Αν το συμπέρασμα (3.5) δεν ισχύει, τότε το σύνολο

$$S = \{j : x_j \neq \max_{1 \leq k \leq n} x_k\}$$

είναι γνήσιο υποσύνολο του $\{1, 2, \ldots, n\}$, και από αυτήν την υπόθεση θα οδηγηθούμε σε αντίφαση. Για το σκοπό αυτό, θέτουμε

$$p = \sum_{j \in S} p_j, \quad x = \sum_{j \in S} \frac{p_j}{p} x_j, \quad \text{και } y = \sum_{j \notin S} \frac{p_j}{1-p} x_j,$$

και, χρησιμοποιώντας την γνήσια κυρτότητα της f, συμπεραίνουμε ότι

$$f\left(\sum_{j=1}^{n} p_j x_j\right) = f(px + (1-p)y) < pf(x) + (1-p)f(y). \tag{3.6}$$

Επιπλέον, εφαρμόζοντας την κυρτότητα της f χωριστά για τα x και y, παίρνουμε την ανισότητα

$$\begin{aligned} pf(x) + (1-p)f(y) &\leq p\sum_{j \in S} \frac{p_j}{p} f(x_j) + (1-p)\sum_{j \notin S} \frac{p_j}{1-p} f(x_j) \\ &= \sum_{j=1}^{n} p_j f(x_j). \end{aligned}$$

Τέλος, από αυτήν την ανισότητα και από την γνήσια ανισότητα (3.6), βλέπουμε ότι

$$f\left(\sum_{j=1}^{n} p_j x_j\right) < \sum_{j=1}^{n} p_j f(x_j),$$

και αυτό αντιφάσκει προς την υπόθεση (3.4). Έτσι, έχουμε το ζητούμενο.

Κριτήριο της παραγώγου για την κυρτότητα

Ένα βασικό πλεονέκτημα της ανισότητας του Jensen είναι η γενικότητά της, πριν όμως χρησιμοποιήσουμε την ανισότητα του Jensen σε κάποιο πρόβλημα, πρέπει πρώτα να εξασφαλίσουμε την κυρτότητα κάποιας (κατάλληλης για το πρόβλημα) συνάρτησης. Σε μερικές περιπτώσεις αυτό επιτυγχάνεται με επαλήθευση του ορισμού (3.1), συνήθως όμως, επαληθεύουμε την κυρτότητα χρησιμοποιώντας το ακόλουθο κριτήριο παραγώγου.

Θεώρημα 3.1.3 (κριτήριο παραγώγου για την κυρτότητα). *Αν η $f : (a, b) \to \mathbb{R}$ είναι δύο φορές παραγωγίσιμη και $f''(x) \geq 0$ για κάθε $x \in (a, b)$, τότε η f είναι κυρτή στο (a, b). Αν $f''(x) > 0$ για κάθε $x \in (a, b)$, τότε η f είναι γνησίως κυρτή στο (a, b).*

Για την απόδειξη, αφού θέλουμε να συσχετίσουμε τη συνάρτηση f με τις παραγώγους της, είναι πολύ φυσικό να ξεκινήσουμε με την αναπαράσταση της f που μας δίνει το θεμελιώδες θεώρημα του Απειροστικού Λογισμού. Συγκεκριμένα, αν σταθεροποιήσουμε μια τιμή $x_0 \in [a, b]$, έχουμε την αναπαράσταση

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(u)\, du \quad \text{για κάθε } x \in [a, b], \tag{3.7}$$

και αυτό μας δίνει την ιδέα να χρησιμοποιήσουμε το γεγονός ότι, αφού $f'' \geq 0$, η συνάρτηση f' είναι αύξουσα. Για την ακρίβεια, η υπόθεσή μας δεν περιέχει καμία άλλη πληροφορία, *πρέπει* λοιπόν η αναπαράσταση (3.7), η μονοτονία της f' και κάποια αριθμητική να μας οδηγήσουν στο συμπέρασμα.

Για να προχωρήσουμε, θεωρούμε $a \leq x < y \leq b$ και $0 < p < 1$. Θέτοντας $q = 1 - p$ και εφαρμόζοντας την αναπαράσταση (3.7) για τα x, y και $x_0 = px + qy$, βλέπουμε ότι η ποσότητα

$$\Delta = pf(x) + qf(y) - f(px + qy)$$

γράφεται στη μορφή

$$\Delta = q \int_{px+qy}^{y} f'(u)\, du - p \int_{x}^{px+qy} f'(u)\, du. \tag{3.8}$$

Για κάθε $u \in [x, px + qy]$ έχουμε $f'(u) \leq f'(px + qy)$, οπότε ισχύει η ανισότητα

$$p \int_{x}^{px+qy} f'(u)\, du \leq qp(y - x) f'(px + qy), \tag{3.9}$$

ενώ για κάθε $u \in [px+qy, y]$ έχουμε $f'(u) \geq f'(px+qy)$, οπότε ισχύει η ανισότητα

$$q \int_{px+qy}^{y} f'(u)\, du \geq qp(y - x) f'(px + qy). \tag{3.10}$$

Συνεπώς, από την ολοκληρωτική αναπαράσταση (3.8) για την Δ και από τις ανισότητες (3.9) και (3.10), βλέπουμε ότι $\Delta \geq 0$.

Για τον δεύτερο ισχυρισμό του προβλήματος, αρκεί να παρατηρήσουμε ότι αν $f''(x) > 0$ για κάθε $x \in (a, b)$, τότε οι ανισότητες (3.9) και (3.10) είναι γνήσιες. Έτσι, η αναπαράσταση (3.8) για την Δ δίνει $\Delta > 0$, το οποίο αποδεικνύει την γνήσια κυρτότητα της f.

Η ανισότητα αριθμητικού–γεωμετρικού μέσου μέσω της εκθετικής συνάρτησης

Το κριτήριο της παραγώγου μας εξασφαλίζει ότι η συνάρτηση $x \mapsto e^x$ είναι κυρτή, οπότε η ανισότητα του Jensen μας λέει ότι για κάθε n-άδα πραγματικών αριθμών $y_1, y_2, \ldots, y_n$ και για κάθε επιλογή θετικών αριθμών p_j, $j = 1, \ldots, n$ με $p_1 + p_2 + \cdots + p_n = 1$, ισχύει η ανισότητα

$$\exp\left(\sum_{j=1}^{n} p_j y_j\right) \leq \sum_{j=1}^{n} p_j e^{y_j}.$$

Θέτοντας $x_j = e^{y_j}$, ξαναβρίσκουμε τη γνωστή ανισότητα

$$\prod_{j=1}^{n} x_j^{p_j} \leq \sum_{j=1}^{n} p_j x_j.$$

Βλέπουμε λοιπόν ότι, με εκπληκτικά σύντομο και καθαρό τρόπο, η ανισότητα του Jensen μας οδηγεί στη γενικευμένη ανισότητα αριθμητικού–γεωμετρικού μέσου.

- Αυτή η οπτική για την ανισότητα αριθμητικού-γεωμετρικού μέσου είναι αποκαλυπτική για την φύση και τη σημασία της. *Η κεντρική θέση που κατέχει η ανισότητα αριθμητικού–γεωμετρικού μέσου στη θεωρία των ανισοτήτων αντανακλά το θεμελιώδη ρόλο της εκθετικής συνάρτησης, η οποία δρά σαν ισομορφισμός ανάμεσα στις δύο πιό σημαντικές ομάδες: την προσθετική ομάδα των πραγματικών αριθμών και την πολλαπλασιαστική ομάδα των θετικών πραγματικών αριθμών.*

3.2 Η χρήση της κυρτότητας σε κάποια τυπικά προβλήματα

Πολλές από τις γνωστές συναρτήσεις που συναντάμε στην τριγωνομετρία και την γεωμετρία είναι κυρτές, και, πολύ συχνά, αυτή τους η ιδιότητα έχει πολύ χρήσιμες συνέπειες. Ένα παράδειγμα μας δίνει το εξής πρόβλημα:

Πρόβλημα 3.2.1 (το μέγιστο γινόμενο δύο ακμών). *Σε ένα ισόπλευρο τρίγωνο εμβαδού A, το γινόμενο οποιωνδήποτε δύο πλευρών ισούται με $(4/\sqrt{3})A$. Δείξτε ότι αυτή η περίπτωση είναι ακραία: κάθε τρίγωνο εμβαδού A έχει δύο πλευρές με μήκη που το γινόμενό τους είναι μεγαλύτερο ή ίσο από $(4/\sqrt{3})A$.*

Για να ξεκινήσουμε χρειαζόμαστε τύπους οι οποίοι να συνδέουν το εμβαδόν A ενός τριγώνου με τα μήκη a, b, c των πλευρών του και τις αντίστοιχες γωνίες α, β, γ. Γνωρίζουμε ότι

$$A = \frac{1}{2}ab\sin\gamma = \frac{1}{2}ac\sin\beta = \frac{1}{2}bc\sin\alpha.$$

Αν πάρουμε τον μέσο όρο των τριών ισοτήτων, μπορούμε να γράψουμε

$$\frac{1}{3}(ab+ac+bc) = \frac{2A}{3}\left\{\frac{1}{\sin\alpha}+\frac{1}{\sin\beta}+\frac{1}{\sin\gamma}\right\}. \tag{3.11}$$

Αυτή η σχέση μας παρακινεί να εξετάσουμε την κυρτότητα της συνάρτησης $1/\sin x$. Η γραφική παράσταση της $x \mapsto 1/\sin x$ για $x \in (0,\pi)$ την υποδεικνύει, και αυτό επιβεβαιώνεται από το κριτήριο της παραγώγου: έχουμε

$$\left(\frac{1}{\sin x}\right)'' = \frac{1}{\sin x} + 2\frac{\cos^2 x}{\sin^3 x} > 0 \quad \text{για κάθε } x \in (0,\pi). \tag{3.12}$$

Συνεπώς, αφού $(\alpha+\beta+\gamma)/3 = \pi/3$, από την ανισότητα του Jensen βλέπουμε ότι

$$\frac{1}{3}\left\{\frac{1}{\sin\alpha}+\frac{1}{\sin\beta}+\frac{1}{\sin\gamma}\right\} \geq \frac{1}{\sin(\pi/3)} = \frac{2}{\sqrt{3}},$$

και, από την (3.11), παίρνουμε τη ζητούμενη ανισότητα

$$\max\{ab, ac, bc\} \geq \frac{1}{3}(ab+ac+bc) \geq \frac{4}{\sqrt{3}}A. \tag{3.13}$$

Το πρόβλημα που συζητήσαμε συνδέεται στενά με μια γνωστή *ανισότητα του Weitzenböck*, η οποία ισχυρίζεται ότι σε κάθε τρίγωνο έχουμε

$$a^2+b^2+c^2 \geq \frac{4}{\sqrt{3}}A. \tag{3.14}$$

Πράγματι, για να περάσουμε από την (3.13) στην ανισότητα του Weitzenböck, αρκεί να θυμηθούμε ότι

$$ab + ac + bc \leq a^2 + b^2 + c^2,$$

κάτι που μπορούμε να αποδείξουμε με δύο τουλάχιστον τρόπους: η ανισότητα του Cauchy και η ανισότητα αριθμητικού–γεωμετρικού μέσου δουλεύουν εξίσου καλά σε αυτήν την περίπτωση.

3.3 Βελτιώνοντας την ανισότητα του Jensen

Υπάρχουν κάποιες μαθηματικές μέθοδοι που, με ημιαυτόματο τρόπο, γενικεύουν μια ταυτότητα, ισχυροποιούν μια ανισότητα, βελτιώνουν κάποιο γνωστό αποτέλεσμα. Το επόμενο αποτέλεσμα μας δίνει ένα παράδειγμα μιας τέτοιας μεθόδου. Δείχνει πώς θα μπορούσαμε να προσπαθήσουμε να ισχυροποιήσουμε κάθε ανισότητα που προκύπτει με εφαρμογή της ανισότητας του Jensen.

Θεώρημα 3.3.1 (ο τύπος σφάλματος του Hölder). *Έστω $f : [a, b] \to \mathbb{R}$ δύο φορές παραγωγίσιμη συνάρτηση με την ιδιότητα*

$$0 \leq m \leq f''(x) \leq M \quad \text{για κάθε } x \in [a, b]. \tag{3.15}$$

Αν $a \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq b$ και αν p_k, $k = 1, 2, \ldots, n$ είναι μη αρνητικοί αριθμοί με $p_1 + p_2 + \cdots + p_n = 1$, υπάρχει πραγματικός αριθμός $\mu \in [m, M]$ για τον οποίο ισχύει η ισότητα

$$\sum_{k=1}^{n} p_k f(x_k) - f\left(\sum_{k=1}^{n} p_k x_k\right) = \frac{1}{4}\mu \sum_{j=1}^{n}\sum_{k=1}^{n} p_j p_k (x_j - x_k)^2. \tag{3.16}$$

Αυτό το αποτέλεσμα περιέχεται στο φημισμένο άρθρο (που δημοσιεύτηκε το 1885) του Otto Ludwig Hölder (1859-1937) στο οποίο αποδεικνύεται και η περίφημη ανισότητα που έμεινε γνωστή ως «ανισότητα του Hölder». Η ανισότητα του σφάλματος είναι λιγότερο γνωστή αλλά πολύ σημαντική. Δίνει έναν πολύ φυσιολογικό τρόπο για να μετρήσουμε τη διαφορά ανάμεσα στα δύο μέλη της ανισότητας του Jensen, και μας λέει πώς να πετύχουμε καλύτερα αποτελέσματα από αυτά που δίνει η απλή εφαρμογή της ανισότητας του Jensen στις περιπτώσεις που έχουμε μη τετριμμένες εκτιμήσεις στην (3.15). Πολύ συχνά, η μεγαλύτερη ακρίβεια των εκτιμήσεων που πετυχαίνουμε δεν δικαιολογεί τον πρόσθετο κόπο που απαιτείται, μερικά όμως λεπτά προβλήματα απαιτούν αυτήν ακριβώς τη μεγαλύτερη ακρίβεια.

Η ανισότητα σφάλματος του Hölder (3.16) μας βοηθάει επίσης να καταλάβουμε καλύτερα τη σχέση των κυρτών συναρτήσεων με τις απλούστερες γραμμικές ή τετραγωνικές συναρτήσεις. Για παράδειγμα, αν η διαφορά $M - m$ είναι μικρή, τότε η ανισότητα (3.16) μας λέει ότι η f συμπεριφέρεται περισότερο σαν τετραγωνική συνάρτηση στο $[a, b]$. Επιπλέον, στην ακραία περίπτωση όπου $m = M$, βλέπουμε ότι η f είναι ακριβώς τετραγωνική, δηλαδή $f(x) = \alpha + \beta x + \gamma x^2$ με $m = M = 2\gamma$, και η (3.16) παίρνει τη μορφή μιας απλής τετραγωνικής ταυτότητας.

Όμοια, αν το M είναι μικρό, ας πούμε $0 \leq M \leq \epsilon$, τότε η ανισότητα (3.16) μας λέει ότι η f συμπεριφέρεται περισότερο σαν μια γραμμική συνάρτηση $f(x) = \alpha + \beta x$.

Για μια γραμμική συνάρτηση, το αριστερό μέλος της (3.16) ισούται με μηδέν, γενικά όμως η (3.16) μας λέει ότι το αριστερό μέλος είναι ένα μικρό πολλαπλάσιο ενός μέτρου του βαθμού διάχυσης των x_j, $j = 1, 2, \ldots, n$ στο $[a, b]$.

Η διατύπωση του θεωρήματος μας οδηγεί αμέσως στο εξής ερώτημα: πώς μπορούμε να χρησιμοποιήσουμε το γεγονός ότι $0 \leq m \leq f''(x) \leq M$; Παρατηρούμε τότε ότι οι συναρτήσεις

$$g(x) = \frac{1}{2}Mx^2 - f(x) \quad \text{και} \quad h(x) = f(x) - \frac{1}{2}mx^2$$

οι οποίες συνδέονται στενά με αυτό το ερώτημα, είναι επίσης κυρτές. Φυσιολογικά τότε, κοιτάμε τη μορφή που παίρνει η ανισότητα του Jensen αν εφαρμοστεί γι' αυτές τις συναρτήσεις.

Στην περίπτωση της g, η ανισότητα του Jensen δίνει την

$$\frac{1}{2}M\overline{x}^2 - f(\overline{x}) \leq \sum_{k=1}^{n} p_k \left\{ \frac{1}{2}Mx_k^2 - f(x_k) \right\}$$

όπου έχουμε θέσει $\overline{x} = p_1x_1 + p_2x_2 + \cdots + p_nx_n$, και αναδιατάσσοντας τους όρους σε αυτήν την ανισότητα, παίρνουμε την

$$\begin{aligned} \left\{ \sum_{k=1}^{n} p_k f(x_k) \right\} - f(\overline{x}) &\leq \frac{1}{2}M \left\{ \left(\sum_{k=1}^{n} p_k x_k^2 \right) - \overline{x}^2 \right\} \\ &= \frac{1}{2}M \sum_{k=1}^{n} p_k (x_k - \overline{x})^2. \end{aligned}$$

Κάνοντας εντελώς ανάλογο υπολογισμό για την h, βλέπουμε ότι

$$\left\{ \sum_{k=1}^{n} p_k f(x_k) \right\} - f(\overline{x}) \geq \frac{1}{2}m \sum_{k=1}^{n} p_k (x_k - \overline{x})^2.$$

Οι δύο αυτές ανισότητες ολοκληρώνουν την απόδειξη της (3.16). Το μόνο που λείπει είναι να παρατηρήσουμε ότι ισχύει η ταυτότητα

$$\sum_{k=1}^{n} p_k (x_k - \overline{x})^2 = \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} p_j p_k (x_j - x_k)^2,$$

κάτι που ελέγχεται εύκολα αν αναπτύξουμε τα δύο μέλη και χρησιμοποιήσουμε τον ορισμό του $\overline{x}$.

Κεφάλαιο 4

Η ανισότητα του Hölder

Έχουμε δεί ως τώρα τρία από τα τέσσερα κεντρικά αποτελέσματα της κλασικής θεωρίας των ανισοτήτων: την ανισότητα Cauchy–Schwarz, την ανισότητα αριθμητικού–γεωμετρικού μέσου, και την ανισότητα του Jensen. Την τετράδα συμπληρώνει ένα αποτέλεσμα που αποδείχθηκε πρώτα από τον L. C. Rogers το 1888 και, με διαφορετικό τρόπο, έναν χρόνο αργότερα από τον Otto Hölder. Στη σύγχρονη γλώσσα, η ανισότητα ισχυρίζεται ότι για όλους τους μη αρνητικούς a_k και b_k, $k = 1, 2, \ldots, n$, ισχύει η ανισότητα

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q}, \tag{4.1}$$

υπό τον όρο ότι οι εκθέτες $p > 1$ και $q > 1$ ικανοποιούν την σχέση

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{4.2}$$

Αξίζει να σημειωθεί ότι τα άρθρα των Rogers και Hölder δίνουν την εντύπωση ότι οι συγγραφείς ενδιαφέρονταν κυρίως για τις πιθανές επεκτάσεις και εφαρμογές της ανισότητας αριθμητικού–γεωμετρικού μέσου. Ειδικότερα, δεν έμοιαζαν να θεωρούν την ανισότητα (4.1) – στη μορφή που ο καθένας την απέδειξε – κάτι το ιδιαίτερα σημαντικό, παρόλο που ο Rogers την εκτιμούσε αρκετά ώστε να παρουσιάσει δύο αποδείξεις της. Αυτός που πρώτος διατύπωσε την ανισότητα στη μορφή (4.1) και αναγνώρισε τον θεμελιώδη ρόλο της, ήταν ο Frigyes Riesz. Μπορούμε λοιπόν να ισχυριστούμε ότι η ανισότητα (4.1) θα έπρεπε να ονομάζεται «ανισότητα του Rogers» ή «ανισότητα Rogers–Hölder–Riesz». Όπως όμως συμβαίνει συχνά, σε κάποια αρχική χρονική στιγμή επιβλήθηκε το όνομα «ανισότητα του Hölder», και πολύ δύσκολα θα μπορούσε κανείς να επιμείνει στη χρήση άλλου ονόματος.

4.1 Ιστορικά σχόλια

Η ανισότητα που απέδειξε ο Rogers το 1888, ισχυρίζεται ότι για κάθε $0 < r < s < t < \infty$ και για κάθε μη αρνητικούς αριθμούς a_k, b_k, $k = 1, 2, \ldots, n$, έχουμε

$$\left(\sum_{k=1}^{n} a_k b_k^s\right)^{t-r} \leq \left(\sum_{k=1}^{n} a_k b_k^r\right)^{t-s} \left(\sum_{k=1}^{n} a_k b_k^t\right)^{s-r}, \tag{4.3}$$

που για συντομία γράφεται στη μορφή

$$(S_s)^{t-r} \leq (S_r)^{t-s}(S_t)^{s-r}, \quad \text{όπου } S_p = \sum_{k=1}^{n} a_k b_k^p,\ p > 0. \tag{4.4}$$

Ο Rogers έδωσε δύο αποδείξεις γι' αυτήν την ανισότητα: στην πρώτη χρησιμοποίούσε την «ταυτότητα των τεσσάρων γραμμάτων», ενώ στην δεύτερη χρησιμοποιούσε την ανισότητα αριθμητικού–γεωμετρικού μέσου, την οποία έγραφε στη μορφή

$$x_1^{w_1} x_2^{w_2} \cdots x_n^{w_n} \leq \left(\frac{w_1 x_1 + w_2 x_2 + \cdots + w_n x_n}{w_1 + w_2 + \cdots + w_n}\right)^{w_1 + w_2 + \cdots + w_n}. \tag{4.5}$$

Χρησιμοποιώντας τις αντικαταστάσεις $w_k = a_k b_k^s$ και $x_k = b_k^{t-s}$, ο Rogers κατέληγε στην

$$\left(b_1^{a_1 b_1^s} b_2^{a_2 b_2^s} \cdots b_n^{a_n b_n^s}\right)^{t-s} \leq (S_t/S_s)^{S_s}, \tag{4.6}$$

και χρησιμοποιώντας τις αντικαταστάσεις $w_k = a_k b_k^s$ και $x_k = b_k^{r-s}$, κατέληγε στην

$$\left(b_1^{a_1 b_1^s} b_2^{a_2 b_2^s} \cdots b_n^{a_n b_n^s}\right)^{r-s} \leq (S_r/S_s)^{S_s}. \tag{4.7}$$

Συνδυάζοντας την (4.6) με την αντιστροφή της (4.7), παίρνουμε

$$(S_s/S_r)^{S_s/(s-r)} \leq b_1^{a_1 b_1^s} b_2^{a_2 b_2^s} \cdots b_n^{a_n b_n^s} \leq (S_t/S_s)^{S_s/(t-s)}. \tag{4.8}$$

Έπεται ότι

$$(S_s/S_r)^{1/(s-r)} \leq (S_t/S_s)^{1/(t-s)}, \tag{4.9}$$

δηλαδή,

$$S_s^{t-r} \leq S_r^{t-s} S_t^{s-r}. \tag{4.10}$$

Η ακριβής ανισότητα που απέδειξε ο Hölder το 1889, έλεγε ότι για κάθε $w_k \geq 0$, $y_k \geq 0$ και $p > 0$ ισχύει

$$\sum_{k=1}^{n} w_k y_k \leq \left(\sum_{k=1}^{n} w_k\right)^{(p-1)/p} \left(\sum_{k=1}^{n} w_k y_k^p\right)^{1/p}. \tag{4.11}$$

Ο Hólder ξεκινούσε από την ανισότητα του Jensen: αν $\phi : [0, +\infty) \to \mathbb{R}$ είναι μια κυρτή συνάρτηση, τότε για κάθε $y_1, \ldots, y_n \geq 0$ και για κάθε $w_1, \ldots, w_n > 0$ έχουμε

(4.12)
$$\phi\left(\frac{w_1 y_1 + w_2 x_2 + \cdots + w_n y_n}{w_1 + w_2 + \cdots + w_n}\right) \leq \frac{w_1\phi(y_1) + w_2\phi(y_2) + \cdots + w_n\phi(y_n)}{w_1 + w_2 + \cdots + w_n}.$$

Εφαρμόζοντας αυτήν την ανισότητα για την κυρτή συνάρτηση $\phi(x) = x^p$, παίρνουμε

$$\left(\frac{w_1 y_1 + w_2 x_2 + \cdots + w_n y_n}{w_1 + w_2 + \cdots + w_n}\right)^p \leq \frac{w_1 y_1^p + w_2 y_2^p + \cdots + w_n y_n^p}{w_1 + w_2 + \cdots + w_n}. \tag{4.13}$$

Με άλλα λόγια,

$$\left(\sum_{k=1}^n w_k y_k\right)^p \leq \left(\sum_{k=1}^n w_k\right)^{p-1} \left(\sum_{k=1}^n w_k y_k^p\right), \tag{4.14}$$

απ' όπου προκύπτει η (4.11).

Μπορούμε τώρα να περάσουμε από την (4.11) στην (4.1) θέτοντας $w_k = b_k^p$ και $y_k = a_k/b_k^{p-1}$.

Για την αντίστροφη κατεύθυνση – αν θέλουμε δηλαδή να ξεκινήσουμε από την (4.1) και να καταλήξουμε στην (4.11) – εφαρμόζουμε την (4.1) με

$$a_k = w_k^{1/p} y_k \quad \text{και} \quad b_k = w_k^{1/q}. \tag{4.15}$$

4.2 Απόδειξη της ανισότητας του Hölder

Δίνουμε τώρα την απόδειξη της ανισότητας (4.1) του Riesz και προσδιορίζουμε τις προϋποθέσεις κάτω από τις οποίες ισχύει σαν ισότητα: για δοσμένη μη μηδενική ακολουθία $a_1, a_2, \ldots, a_n$ ισχύει ισότητα στην (4.1) αν και μόνο αν υπάρχει σταθερά $\lambda \in \mathbb{R}$ ώστε

$$\lambda a_k^{1/p} = b^{1/q} \quad \text{για κάθε } 1 \leq k \leq n. \tag{4.16}$$

Η πρώτη σκέψη που μας έρχεται στο μυαλό είναι ασφαλώς να προσπαθήσουμε να προσαρμόσουμε κάποια από τις αποδείξεις της ανισότητας του Cauchy: είναι μάλιστα αρκετά χρήσιμο να δούμε γιατί κάποιες από αυτές δεν μπορούν να μας βοηθήσουν. Για παράδειγμα, αν $p \neq 2$, το επιχείρημα του Schwarz δεν μπορεί να δουλέψει αφού δεν υπάρχει κάποιο τριώνυμο που να σχετίζεται φυσιολογικά με το πρόβλημα. Όμοια, η απουσία ενός τέτοιου τριωνύμου σημαίνει ότι είναι απίθανο να βρούμε κάποια ταυτότητα τύπου Lagrange που να οδηγεί στην ανισότητα.

Ερχόμαστε έτσι στην πιό ισχυρή απόδειξη της ανισότητας του Cauchy, αυτήν που ξεκινάει με την «ταπεινή ανισότητα»

$$xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2 \qquad \text{για κάθε } x, y \in \mathbb{R}. \tag{4.17}$$

Αυτή η ανισότητα μας θυμίζει ότι, από την γενική ανισότητα αριθμητικού–γεωμετρικού μέσου (2.9) έπεται ότι

$$x^\alpha y^\beta \le \frac{\alpha}{\alpha+\beta} x^{\alpha+\beta} + \frac{\beta}{\alpha+\beta} y^{\alpha+\beta} \tag{4.18}$$

για κάθε $x \ge 0$, $y \ge 0$, $\alpha > 0$ και $\beta > 0$. Αν λοιπόν θέσουμε $u = x^\alpha$, $v = y^\beta$, $p = (\alpha+\beta)/\alpha$ και $q = (\alpha+\beta)/\beta$, τότε βλέπουμε ότι για κάθε $p > 1$ ισχύει η συνεπαγωγή

$$\frac{1}{p} + \frac{1}{q} = 1 \qquad \Longrightarrow \qquad uv \le \frac{1}{p}u^p + \frac{1}{q}v^q \qquad \text{για κάθε } u, v \in \mathbb{R}^+. \tag{4.19}$$

Αυτή ιδανική γενίκευση της (4.17) είναι γνωστή ως «ανισότητα του Young».

Η συνέχεια της απόδειξης της ανισότητας του Hölder είναι η συνηθισμένη. Αν κάνουμε τις αντικαταστάσεις $u \mapsto a_k$ και $y \mapsto b_k$ στην (4.19) και αν αθροίσουμε ως προς $1 \le k \le n$, παίρνουμε την

$$\sum_{k=1}^n a_k b_k \le \frac{1}{p}\sum_{k=1}^n a_k^p + \frac{1}{q}\sum_{k=1}^n b_k^q, \tag{4.20}$$

και, για να περάσουμε από αυτήν την προσθετική ανισότητα σε μια πολλαπλασιαστική ανισότητα, μπορούμε να εφαρμόσουμε το *τέχνασμα της κανονικοποίησης* το οποίο έχουμε ήδη χρησιμοποιήσει. Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι καμία από τις δύο ακολουθίες δεν είναι ταυτοτικά μηδενική, οπότε μπορούμε να ορίσουμε τις κανονικοποιημένες μεταβλητές

$$\hat{a}_k = a_k \Big/ \left(\sum_{k=1}^n a_k^p\right)^{1/p} \qquad \text{και} \qquad \hat{b}_k = b_k \Big/ \left(\sum_{k=1}^n b_k^q\right)^{1/q}.$$

Τώρα, αν αντικαταστήσουμε αυτούς τις τιμές στην προσθετική ανισότητα (4.20) παίρνουμε το ζητούμενο με απλή αριθμητική.

Το επιχείρημα του Riesz είναι άμεσο, όμως ο *τρόπος διατύπωσης* της ανισότητας από τον Riesz ήταν αυτός που απλούστευσε πολύ την απόδειξή της: είχε την σοφία να αντιληφθεί τον ρόλο των ζευγαριών εκθετών p και q που ικανοποιούν την $1/p + 1/q = 1$. Η ισορροπία των *συζυγών εκθετών* φαίνεται ήδη στην $p-q$ γενίκευση (4.19) της (4.17), και παίζει πολύ σημαντικό ρόλο στην Ανάλυση.

Για να ολοκληρώσουμε τη λύση του προβλήματος πρέπει να προσδιορίσουμε τις περιπτώσεις στις οποίες ισχύει ισότητα. Παρατηρήστε πρώτα ότι ισχύει τετριμμένα ισότητα αν $b_k = 0$ για κάθε $1 \le k \le n$, όμως σε αυτήν την περίπτωση η ισότητα (4.16) ικανοποιείται με $\lambda = 0$. Έτσι, χωρίς περιορισμό της γενικότητας υποθέτουμε ότι και οι δύο ακολουθίες είναι μη μηδενικές.

Στη συνέχεια, παρατηρούμε ότι ισχύει ισότητα στην ανισότητα του Hölder (4.1) αν και μόνο αν ισχύει ισότητα στην προσθετική ανισότητα (4.20) όταν αυτή εφαρμόζεται για τις κανονικοποιημένες μεταβλητές $\hat{a}_k$ και $\hat{b}_k$. Από την κατά σημείο ανισότητα (4.19) βλέπουμε τώρα ότι ισχύει ισότητα στην προσθετική ανισότητα (4.20) αν και μόνο αν έχουμε

$$\hat{a}_k \hat{b}_k = \frac{1}{p}\hat{a}_k^p + \frac{1}{q}\hat{b}_k^q \qquad \text{για κάθε } k = 1, 2, \ldots, n.$$

Τότε, από τον χαρακτηρισμό της ισότητας στην ανισότητα αριθμητικού–γεωμετρικού μέσου (4.18), βλέπουμε ότι για κάθε $1 \le k \le n$ πρέπει να ισχύει $\hat{a}_k^p = \hat{b}_k^q$. Τέλος, επιστρέφοντας στις αρχικές μεταβλητές a_k και b_k, βλέπουμε ότι $\lambda a_k^p = b_k^q$ για κάθε $1 \le k \le n$, όπου η σταθερά λ ισούται με

$$\lambda = \left(\sum_{k=1}^n b_k^q\right)^{1/q} \Big/ \left(\sum_{k=1}^n a_k^p\right)^{1/p}.$$

Έχουμε έτσι δώσει χαρακτηρισμό για τις περιπτώσεις ισότητας στην ανισότητα του Hölder.

4.3 Αντίστροφη της ανισότητας του Hölder

Το επόμενο αποτέλεσμα θα μπορούσε να ιδωθεί ως η *αντίστροφη της ανισότητας του Hölder*.

Θεώρημα 4.3.1 (η αντίστροφη της Hölder – η έννοια του δυϊσμού). *Αν $1 < p < \infty$ και αν C είναι μια σταθερά με την ιδιότητα*

$$\sum_{k=1}^n a_k x_k \le C \left\{\sum_{k=1}^n |x_k|^p\right\}^{1/p} \tag{4.21}$$

για κάθε x_k, $1 \le k \le n$, τότε, αν θέσουμε $q = p/(p-1)$, ισχύει η ανισότητα

$$\left\{\sum_{k=1}^n |a_k|^q\right\}^{1/q} \le C. \tag{4.22}$$

Αυτό το αποτέλεσμα μας βοηθάει να εξηγήσουμε γιατί η εμφάνιση των συζυγών εκθετών (p, q) του Riesz είναι αναπόφευκτη. Οι μη γραμμικοί περιορισμοί είναι πάντοτε δύσχρηστοι, και εδώ βλέπουμε ότι οι x–μεταβλητές εμφανίζονται και στα δύο μέλη της υπόθεσης (4.21). Για να τις απαλείψουμε χρειαζόμαστε ένα τέχνασμα.

Μια ιδέα που δουλεύει αρκετά συχνά όταν έχουμε ελεύθερες μεταβλητές και στα δύο μέλη μιας σχέσης είναι να «συνωμοτήσουμε» ώστε να κάνουμε τα δύο μέλη να μοιάζουν όσο γίνεται περισσότερο. Αυτή η «αρχή της εξίσωσης των δύο μελών» δεν είναι τελείως σαφής, στην περίπτωσή μας όμως υποδεικνύει ότι για κάθε $1 \le k \le n$ θα έπρεπε να επιλέξουμε το x_k έτσι ώστε να ισχύει η $a_k x_k = |x_k|^p$. Με άλλα λόγια, θέτουμε $x_k = \operatorname{sign}(a_k)|a_k|^{p/(p-1)}$, όπου $\operatorname{sign}(a_k) = 1$ αν $a_k \ge 0$ και $\operatorname{sign}(a_k) = -1$ αν $a_k < 0$. Με αυτήν την επιλογή, η συνθήκη (4.21) παίρνει τη μορφή

$$\sum_{k=1}^n |a_k|^{p/(p-1)} \le C \left\{\sum_{k=1}^n |a_k|^{p/(p-1)}\right\}^{1/p}. \tag{4.23}$$

Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι το άθροισμα στο δεξιό μέλος είναι μη μηδενικό, οπότε μπορούμε να διαιρέσουμε με αυτό το άθροισμα. Χρησιμοποιώντας τη σχέση $1/p + 1/q = 1$ παρατηρούμε ότι προκύπτει η ζητούμενη ανισότητα (4.22).

Η ανισότητα του Minkowski

Η ανισότητα του Hölder και η ανισότητα δυϊσμού (4.22) μπορούν να επαναδιατυπωθούν με πολλούς τρόπους, ο πιό όμορφος όμως απ' όλους απαιτεί την εισαγωγή κάποιου νέου συμβολισμού. Αν $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ είναι μια n–άδα πραγματικών αριθμών, για κάθε $1 \le p < \infty$ γράφουμε

$$\|\mathbf{a}\|_p = \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p}, \tag{4.24}$$

ενώ για $p = \infty$ θέτουμε $\|\mathbf{a}\|_\infty = \max_{1 \le k \le n} |a_k|$. Με αυτόν τον συμβολισμό, η ανισότητα του Hölder (4.1) για $1 \le p < \infty$ παίρνει την απλή μορφή

$$\left| \sum_{k=1}^{n} a_k b_k \right| \le \|\mathbf{a}\|_p \|\mathbf{b}\|_q,$$

όπου για $1 < p < \infty$ το ζευγάρι (p, q) είναι οι συνήθεις συζυγείς εκθέτες που προσδιορίζονται από την σχέση

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \text{όταν } 1 < p < \infty,$$

ενώ για $p = 1$ θέτουμε $q = \infty$.

Η ποσότητα $\|\mathbf{a}\|_p$ ονομάζεται p–νόρμα, ή ℓ^p–νόρμα, της n–άδας, για να δικαιολογήσουμε όμως αυτήν την ονομασία, πρέπει να ελέγξουμε ότι όντως ικανοποιεί όλες τις ιδιότητες που απαιτούνται από τον ορισμό της νόρμας. Πιο συγκεκριμένα, πρέπει να επαληθεύσουμε τις εξής τρεις ιδιότητες:

(i) $\|\mathbf{a}\|_p \ge 0$, με ισότητα αν και μόνο αν $\mathbf{a} = \mathbf{0}$.

(ii) $\|\alpha \mathbf{a}\|_p = |\alpha| \, \|\mathbf{a}\|_p$ για κάθε $\alpha \in \mathbb{R}$.

(iii) $\|\mathbf{a} + \mathbf{b}\|_p \le \|\mathbf{a}\|_p + \|\mathbf{b}\|_p$ για όλες τις n–άδες $\mathbf{a}$ και $\mathbf{b}$.

Οι πρώτες δύο ιδιότητες προκύπτουν άμεσα από τον ορισμό (4.24), η τρίτη όμως ιδιότητα είναι βαθύτερη. Είναι γνωστή ως «ανισότητα του Minkowski» και, παρόλο που δεν είναι δύσκολο να αποδειχθεί, είναι θεμελιώδης.

Θεώρημα 4.3.2 (η ανισότητα του Minkowski). *Για κάθε* $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ *και* $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ *ισχύει η ανισότητα*

$$\|\mathbf{a} + \mathbf{b}\|_p \le \|\mathbf{a}\|_p + \|\mathbf{b}\|_p, \tag{4.25}$$

ή, ισοδύναμα, για κάθε $p \ge 1$ *ισχύει η ανισότητα*

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p \right)^{1/p} \le \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p \right)^{1/p}. \tag{4.26}$$

Επιπλέον, αν $\|\mathbf{a}\|_p \ne 0$ *και αν* $p > 1$*, τότε ισχύει ισότητα στην ανισότητα* (4.25) *αν και μόνο αν* (1) *υπάρχει σταθερά* $\lambda \in \mathbb{R}$ *ώστε* $|b_k| = \lambda |a_k|$ *για κάθε* $k = 1, 2, \ldots, n$*, και* (2) *οι* a_k *και* b_k *έχουν το ίδιο πρόσημο για κάθε* $k = 1, 2, \ldots, n$.

Υπάρχουν πολλές διαφορετικές αποδείξεις της ανισότητας του Minkowski, η μέθοδος όμως που χρησιμοποιήθηκε από τον F. Riesz είναι η πιό γοητευτική – ιδιαίτερα αν κάποιος θέλει να παρουσιάσει την ανισότητα του Minkowski αμέσως μετά από την παρουσίαση της ανισότητας του Hölder. Το ερώτημα είναι: «θα μπορούσε να μας βοηθήσει η Hölder»; Αμέσως μετά, η άλγεβρα έρχεται να μας βοηθήσει.

Αφού το άνω φράγμα που ζητάμε είναι το άθροισμα δύο όρων, είναι λογικό να σπάσουμε το αριστερό μέλος σε δύο μέρη:

$$\sum_{k=1}^{n} |a_k + b_k|^p \leq \sum_{k=1}^{n} |a_k|\,|a_k + b_k|^{p-1} + \sum_{k=1}^{n} |b_k|\,|a_k + b_k|^{p-1}. \tag{4.27}$$

Αυτή η διάσπαση δίνει ήδη την ανισότητα του Minkowski (4.26) για $p = 1$, οπότε μπορούμε πλέον να υποθέσουμε ότι $p > 1$. Εφαρμόζοντας την ανισότητα του Hölder χωριστά σε καθένα από τα αθροίσματα της (4.27), βλέπουμε ότι

$$\sum_{k=1}^{n} |a_k|\,|a_k + b_k|^{p-1} \leq \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{(p-1)/p}$$

και

$$\sum_{k=1}^{n} |b_k|\,|a_k + b_k|^{p-1} \leq \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{(p-1)/p}.$$

Έτσι, με τον συμβολισμό της (4.24) παίρνουμε

$$\|\mathbf{a} + \mathbf{b}\|_p^p \leq \|\mathbf{a}\|_p \cdot \|\mathbf{a} + \mathbf{b}\|_p^{p-1} + \|\mathbf{b}\|_p \cdot \|\mathbf{a} + \mathbf{b}\|_p^{p-1}. \tag{4.28}$$

Αφού η ανισότητα του Minkowski ισχύει τετριμμένα στην περίπτωση που $\|\mathbf{a} + \mathbf{b}\|_p = 0$, μπορούμε να υποθέσουμε ότι $\|\mathbf{a}+\mathbf{b}\|_p \neq 0$. Διαιρώντας τα δύο μέλη της (4.28) με $\|\mathbf{a} + \mathbf{b}\|_p^{p-1}$ ολοκληρώνουμε την απόδειξη.

Ένα πλεονέκτημα της απόδειξης του Riesz για την ανισότητα του Minkowski είναι ότι εξετάζοντας προσεκτικά το επιχείρημά του μπορούμε να προσδιορίσουμε τις περιπτώσεις της ισότητας.

Κατ' αρχήν, παρατηρούμε ότι αν ισχύει ισότητα στην ανισότητα (4.25) τότε πρέπει να έχουμε ισότητα στο πρώτο μας βήμα, την (4.27), και επίσης να ισχύει η $|a_k + b_k| = |a_k| + |b_k|$ για κάθε $1 \leq k \leq n$. Έτσι, μπορούμε να υποθέσουμε ότι οι a_k και b_k είναι ομόσημοι για κάθε $1 \leq k \leq n$, και χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $a_k \geq 0$ και $b_k \geq 0$ για κάθε $1 \leq k \leq n$.

Από την ισότητα στην (4.25) συμπεραίνουμε επίσης ότι έχουμε ισότητα στις δύο περιπτώσεις όπου εφαρμόσαμε την ανισότητα του Hölder, οπότε, υποθέτοντας ότι $\|\mathbf{a} + \mathbf{b}\|_p \neq 0$, συμπεραίνουμε ότι υπάρχει $\lambda \geq 1$ τέτοιος ώστε

$$\lambda |a_k|^p = \{|a_k + b_k|^{p-1}\}^q = |a_k + b_k|^p$$

και ότι υπάρχει $\lambda' \geq 1$ τέτοιος ώστε

$$\lambda' |b_k|^p = \{|a_k + b_k|^{p-1}\}^q = |a_k + b_k|^p.$$

Από αυτές τις ισότητες βλέπουμε ότι αν θέσουμε $\lambda'' = \lambda/\lambda'$ τότε έχουμε $\lambda'' |a_k|^p = |b_k|^p$ για κάθε $k = 1, 2, \ldots, n$.

Αυτός είναι ακριβώς ο χαρακτηρισμός που ελπίζαμε να αποδείξουμε.

4.4 Υποπροσθετικότητα και ημιγραμμικοποίηση

Η ανισότητα του Minkowski μας λέει ότι η συνάρτηση $h : \mathbb{R}^n \to \mathbb{R}$ με $h(\mathbf{a}) = \|\mathbf{a}\|_p$ είναι *υποπροσθετική* με την έννοια ότι ικανοποιεί την ανισότητα

$$h(\mathbf{a}+\mathbf{b}) \leq h(\mathbf{a}) + h(\mathbf{b}) \qquad \text{για κάθε } \mathbf{a}, \mathbf{b} \in \mathbb{R}^n.$$

Η απόδειξη της υποπροσθετικότητας μιας συνάρτησης είναι συνήθως απλή, και αυτό μας κάνει να αναρωτηθούμε αν υπάρχει κάποιος άμεσος τρόπος απόδειξης της ανισότητας του Minkowski. Το επόμενο θεώρημα επιβεβαιώνει αυτήν την υποψία και ταυτόχρονα δίνει πρόσθετες πληροφορίες.

Θεώρημα 4.4.1 (ημιγραμμικοποίηση της ℓ^p–νόρμας). *Για κάθε $1 \leq p \leq \infty$ ισχύει η ταυτότητα*

$$\|\mathbf{a}\|_p = \max\left\{\sum_{k=1}^n a_k x_k : \|\mathbf{x}\|_q = 1\right\}, \tag{4.29}$$

όπου $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ και οι p και q είναι συζυγείς εκθέτες: $q = p/(p-1)$ αν $p > 1$ και $q = \infty$ αν $p = 1$, $q = 1$ αν $p = \infty$. Τέλος, αυτή η ισότητα έχει σαν άμεση συνέπεια την ανισότητα του Minkowski.

Πριν αποδείξουμε το θεώρημα, είναι χρήσιμο να δούμε την ημιγραμμικοποίηση σε ένα ευρύτερο πλαίσιο. Αν V είναι ένας γραμμικός χώρος (για παράδειγμα, ο $\mathbb{R}^n$) και αν $L : V \times W \to \mathbb{R}$ είναι μια συνάρτηση προσθετική ως προς την πρώτη της μεταβλητή, δηλαδή αν ικανοποιεί την $L(\mathbf{a}+\mathbf{b}, \mathbf{w}) = L(\mathbf{a}, \mathbf{w}) + L(\mathbf{b}, \mathbf{w})$, τότε η συνάρτηση $h : V \to \mathbb{R}$ που ορίζεται από την

$$h(\mathbf{a}) = \max_{\mathbf{w} \in W} L(\mathbf{a}, \mathbf{w}), \tag{4.30}$$

είναι πάντα υποπροσθετική, αφού

$$\begin{aligned} h(\mathbf{a}+\mathbf{b}) &= \max_{\mathbf{w}\in W} L(\mathbf{a}+\mathbf{b}, \mathbf{w}) = \max_{\mathbf{w}\in W} \big\{L(\mathbf{a}, \mathbf{w}) + L(\mathbf{b}, \mathbf{w})\big\} \\ &\leq \max_{\mathbf{w}_0\in W} L(\mathbf{a}, \mathbf{w}_0) + \max_{\mathbf{w}_1\in W} L(\mathbf{b}, \mathbf{w}_1) = h(\mathbf{a}) + h(\mathbf{b}). \end{aligned}$$

Η αναπαράσταση (4.30) λέγεται *ημιγραμμικοποίηση* της h, και πολλές από τις θεμελιώδεις ποσότητες στην θεωρία των ανισότητων επιδέχονται αναπαραστάσεις αυτού του τύπου.

Η ύπαρξη της ημιγραμμικής αναπαράστασης (4.29) για την συνάρτηση $h(\mathbf{a}) = \|\mathbf{a}\|_p$ είναι απλή συνέπεια της ανισότητας του Hölder και της αντίστροφής της. Θεωρούμε το σύνολο

$$S = \left\{\sum_{k=1}^n a_k x_k : \sum_{k=1}^n |x_k|^q \leq 1\right\},$$

και παρατηρούμε ότι από την ανισότητα του Hölder έχουμε $s \leq \|\mathbf{a}\|_p$ για κάθε $s \in S$. Αυτό αποδεικνύει ότι $\max\{s \in S\} \leq \|\mathbf{a}\|_p$. Τώρα, από τον ορισμό του S, κάνοντας κανονικοποίηση βλέπουμε ότι

$$\sum_{k=1}^n a_k y_k \leq \|\mathbf{y}\|_q \max\{s \in S\} \qquad \text{για κάθε } y \in \mathbb{R}^n. \tag{4.31}$$

Έτσι, από την αντίστροφη ανισότητα Hölder για το ζευγάρι συζυγών εκθετών (q, p) – προσέξτε την εναλλαγή των ρόλων των p και q – παίρνουμε την δεύτερη ανισότητα $\|\mathbf{a}\|_p \leq \max\{s \in S\}$. Συνδυάζοντας τις δύο ανισότητες παίρνουμε την ημιγραμμική αναπαράσταση (4.29) για την $h(\mathbf{a}) = \|\mathbf{a}\|_p$.

4.5 Ένα αποτέλεσμα ευστάθειας για την ανισότητα του Hölder

Σε πολλές περιοχές των Μαθηματικών συναντάμε *αποτελέσματα χαρακτηρισμού* και *αποτελέσματα ευστάθειας*. Συνήθως, ένα αποτέλεσμα χαρακτηρισμού δίνει πλήρη περιγραφή των λύσεων κάποιας εξίσωσης, ενώ το αντίστοιχο αποτέλεσμα ευστάθειας ισχυρίζεται ότι αν η εξίσωση «σχεδόν ισχύει» τότε ο χαρακτηρισμός «σχεδόν ικανοποιείται». Σε αυτήν την παράγραφο εξετάζουμε αν υπάρχει κάποιο αποτέλεσμα ευστάθειας σχετικό με την ανισότητα του Hölder.

Για να γίνουμε πιο συγκεκριμένοι, ας παρατηρήσουμε πρώτα ότι το «τέχνασμα της εισαγωγής μονάδων» και η ανισότητα του Hölder δείχνουν ότι για κάθε $p > 1$ και για κάθε ακολουθία μη αρνητικών πραγματικών αριθμών $a_1, a_2, \ldots, a_n$ ισχύει η ανισότητα

$$\sum_{j=1}^{n} a_j \leq n^{(p-1)/p} \left(\sum_{j=1}^{n} a_j^p \right)^{1/p}.$$

Αν λοιπόν ορίσουμε τον *συντελεστή απόκλισης* $\delta(\mathbf{a})$ θέτοντας

$$\delta(\mathbf{a}) := \sum_{j=1}^{n} a_j^p - n^{1-p} \left(\sum_{j=1}^{n} a_j \right)^p, \tag{4.32}$$

τότε έχουμε $\delta(\mathbf{a}) \geq 0$ και, για να περάσουμε στο ενδιαφέρον σημείο αυτού του ορισμού, το κριτήριο της ισότητας στην ανισότητα του Hölder μας λέει ότι $\delta(\mathbf{a}) = 0$ αν και μόνο αν υπάρχει κάποια σταθερά μ τέτοια ώστε $a_j = \mu$ για κάθε $j = 1, 2, \ldots, n$. Δηλαδή, η συνθήκη $\delta(\mathbf{a}) = 0$ ικανοποιείται αν και μόνο αν όλες οι συντεταγμένες του $\mathbf{a}$ είναι ίσες.

Αυτός ο χαρακτηρισμός οδηγεί με τη σειρά του σε μια σειρά αποτελεσμάτων ευστάθειας:

Θεώρημα 4.5.1 (ένα αποτέλεσμα ευστάθειας για την ανισότητα του Hölder). *Αν $p \geq 2$ και αν $a_j \geq 0$ για κάθε $1 \leq j \leq n$, τότε υπάρχει σταθερά $\lambda = \lambda(\mathbf{a}, p)$ τέτοια ώστε*

$$(\lambda - \sqrt{\delta})^{2/p} \leq a_j \leq (\lambda + \sqrt{\delta})^{2/p} \tag{4.33}$$

για κάθε $j = 1, 2, \ldots, n$. Με άλλα λόγια, αν ο συντελεστής απόκλισης $\delta = \delta(\mathbf{a})$ είναι μικρός, τότε η ακολουθία $a_1, a_2, \ldots, a_n$ είναι σχεδόν σταθερή.

Ο τρόπος με τον οποίο το ίδιο το θεώρημα μας ζητάει να εκφράσουμε το γεγονός ότι η ακολουθία $a_1, a_2, \ldots, a_n$ είναι σχεδόν σταθερή, μας δίνει μια υπόδειξη για το πώς να ξεκινήσουμε. Η σχέση (4.33) μπορεί να γραφτεί ισοδύναμα στη

μορφή $(a_j^{p/2} - \lambda)^2 \leq \delta(\mathbf{a})$, και μπορούμε να αποδείξουμε όλες (τις n) ζητούμενες ανισότητες ταυτόχρονα αν δείξουμε την ισχυρότερη εικασία ότι υπάρχει σταθερά λ τέτοια ώστε να ισχύει η ανισότητα

$$\sum_{j=1}^{n} (a_j^{p/2} - \lambda)^2 \leq \delta(\mathbf{a}). \tag{4.34}$$

Θα μπορούσε κάποιος να πεί ότι με αυτήν την εικασία ζητάμε πολλά, αξίζει όμως τον κόπο να την διερευνήσουμε.

Τι το ιδιαίτερο έχει η (4.34); Πρώτα απ' όλα, αν $p = 2$, με απευθείας υπολογισμό βλέπουμε από τον ορισμό του $\delta(\mathbf{a})$ ότι η (4.34) ισχύει στην πραγματικότητα σαν ταυτότητα, αρκεί να επιλέξουμε $\lambda = (a_1 + a_2 + \cdots + a_n)/n$. Αυτό είναι φυσικά πολύ καλό σημάδι.

Ένα λιγότερο προφανές προσόν της εικασίας (4.34) είναι ότι εμμέσως μας ρωτάει αν κάποιο συγκεκριμένο τριώνυμο έχει πραγματικές ρίζες. Αν η ανισότητα (4.34) ισχύει για κάποιον πραγματικό αριθμό λ, τότε λόγω συνέχειας θα υπάρχει επίσης πραγματικός αριθμός λ ο οποίος θα ικανοποιεί την *εξίσωση*

$$\sum_{j=1}^{n} (a_j^{p/2} - \lambda)^2 = \delta(\mathbf{a}) := \sum_{j=1}^{n} a_j^p - n^{1-p} \left(\sum_{j=1}^{n} a_j \right)^p .$$

Μετά από κάποιες αλγεβρικές πράξεις και απλοποιήσεις βλέπουμε λοιπόν ότι η εικασία (4.34) ισχύει αν και μόνο αν υπάρχει πραγματική ρίζα της εξίσωσης

$$n\lambda^2 - 2\lambda \sum_{j=1}^{n} a_j^{p/2} + n^{1-p} \left(\sum_{j=1}^{n} a_j \right)^p = 0. \tag{4.35}$$

Αφού μια δευτεροβάθμια εξίσωση $a\lambda^2 + 2B\lambda + C = 0$ έχει πραγματικές ρίζες αν και μόνο αν $AC \leq B^2$, βλέπουμε ότι η λύση του προβλήματός μας θα έχει ολοκληρωθεί αν μπορέσουμε να δείξουμε ότι

$$n^{2-p} \left(\sum_{j=1}^{n} a_j \right)^p \leq \left(\sum_{j=1}^{n} a_j^{p/2} \right)^2 . \tag{4.36}$$

Ευτυχώς, είναι εύκολο να διαπιστώσουμε ότι η τελευταία ανισότητα ισχύει. Για την ακρίβεια, είναι ένα απλό πόρισμα της ανισότητας του Hölder και του «τεχνάσματος της εισαγωγής μονάδων». Αρκεί να εφαρμόσουμε την ανισότητα του Hölder με $p' = p/2$ και $q' = p/(p-2)$ στο άθροισμα $a_1 \cdot 1 + a_2 \cdot 1 + \cdots + a_n \cdot 1$.

4.6 Παρεμβολή

Η ℓ^1–νόρμα και η ℓ^∞–νόρμα είναι οι δύο ακραίες μεταξύ των ℓ^p–νορμών, και σε αρκετές περιπτώσεις μπορούμε συνδυάζοντας μια ℓ^1 ανισότητα με μια ℓ^∞ ανισότητα να πάρουμε αντίστοιχη ανισότητα για την ℓ^p–νόρμα, όπου $1 < p < \infty$.

Το τελευταίο μας θεώρημα δίνει ένα σημαντικό παράδειγμα εφαρμογής αυτής της δυνατότητας. Έχει επίσης σκοπό να δώσει μια ιδέα για ένα από τα πλέον διαδεδομένα θέματα που συναντάμε στην θεωρία των ανισοτήτων – την *παρεμβολή*.

Θεώρημα 4.6.1 (ένα παράδειγμα $\ell^1 - \ell^\infty$ παρεμβολής). *Έστω c_{jk}, $1 \le j \le m$, $1 \le k \le n$ ένας πίνακας μη αρνητικών πραγματικών αριθμών οι οποίοι ικανοποιούν τις*

$$\sum_{j=1}^{m} \left| \sum_{k=1}^{n} c_{jk} x_k \right| \le A \sum_{k=1}^{n} |x_k| \quad \text{και} \quad \max_{1 \le j \le m} \left| \sum_{k=1}^{n} c_{jk} x_k \right| \le B \max_{1 \le k \le n} |x_k|$$

για κάθε x_k, $1 \le k \le n$. Αν $1 < p < \infty$ και $q = p/(p-1)$, τότε ισχύει η ανισότητα παρεμβολής

$$\left(\sum_{j=1}^{m} \left| \sum_{k=1}^{n} c_{jk} x_k \right|^p \right)^{1/p} \le A^{1/p} B^{1/q} \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} \tag{4.37}$$

για κάθε x_k, $1 \le k \le n$.

Το πιο ενοχλητικό στοιχείο της ανισότητας (4.37) είναι η παρουσία των p–ριζών. Θα θέλαμε λοιπόν να βρούμε έναν τρόπο για να τις εξαφανίσουμε. Η ρίζα στο δεξιό μέλος δεν δημιουργεί προβλήματα γιατί κανονικοποιώντας το $\mathbf{x}$ μπορούμε, χωρίς περιορισμό της γενικότητας, να υποθέσουμε ότι $\|\mathbf{x}\|_p \le 1$. Τι θα μπορούσαμε όμως να κάνουμε για την p–ρίζα στο αριστερό μέλος;

Ευτυχώς, διαθέτουμε ένα εργαλείο που είναι απολύτως κατάλληλο για την περίσταση. Η αντίστροφη της ανισότητας του Hölder μας λέει ότι, για να αποδείξουμε την ανισότητα (4.37), αρκεί να αποδείξουμε ότι, για όλα τα διανύσματα $\mathbf{x}$ και $\mathbf{y}$ που ικανοποιούν τις $\|\mathbf{x}\|_p \le 1$ και $\|\mathbf{y}\|_q \le 1$, ισχύει η ανισότητα

$$\sum_{j=1}^{m} \sum_{k=1}^{n} c_{jk} x_k y_j \le A^{1/p} B^{1/q}. \tag{4.38}$$

Επιπλέον, αφού υποθέτουμε ότι $c_{jk} \ge 0$ για κάθε j και k, αρκεί να αποδείξουμε την ανισότητα για $\|\mathbf{x}\|_p \le 1$ και $\|\mathbf{y}\|_q \le 1$ με $x_k \ge 0$ και $y_j \ge 0$ για κάθε j και k.

Διατυπώνοντας το πρόβλημα στη μορφή της (4.38) έχουμε κάνει πραγματική πρόοδο. Τώρα αντιμετωπίζουμε έναν τύπο προβλήματος που έχουμε ξανασυναντήσει: θέλουμε να εκτιμήσουμε ένα άθροισμα κάτω από κάποιους μη γραμμικούς περιορισμούς.

Εδώ, είναι λογικό να χρησιμοποιήσουμε το τέχνασμα της διάσπασης, εκμεταλλευόμενοι τη σχέση $1/p + 1/q = 1$. Κάνοντας διάσπαση και χρησιμοποιώντας την ανισότητα του Hölder, βλέπουμε ότι

$$\begin{aligned} \sum_{j=1}^{m} \sum_{k=1}^{n} c_{jk} x_k y_j &= \sum_{j=1}^{m} \sum_{k=1}^{n} (c_{jk} x_k^p)^{1/p} (c_{jk} y_j^q)^{1/q} \\ &\le \left(\sum_{j=1}^{m} \sum_{k=1}^{n} c_{jk} x_k^p \right)^{1/p} \left(\sum_{j=1}^{m} \sum_{k=1}^{n} c_{jk} y_j^q \right)^{1/q}, \end{aligned}$$

και τώρα χρειάζεται να εκτιμήσουμε τους τελευταίους δύο όρους.

Ο πρώτος όρος είναι εύκολος, αφού η πρώτη μας υπόθεση και η $\|\mathbf{x}\|_p \leq 1$ μας δίνουν το φράγμα

$$\sum_{j=1}^{m}\sum_{k=1}^{n} c_{jk}x_k^p \leq A\sum_{k=1}^{n} x_k^p \leq A. \tag{4.39}$$

Η εκτίμηση του δεύτερου όρου δεν είναι πολύ δυσκολότερη. Αρκεί να χρησιμοποιήσουμε την δεύτερη υπόθεση και την $\|\mathbf{y}\|_q \leq 1$, σε συνδυασμό με ένα χονδροειδές φράγμα:

$$\sum_{j=1}^{m}\sum_{k=1}^{n} c_{jk}y_j^q \leq \sum_{j=1}^{m} y_j^q \left\{\max_{1\leq k\leq n}\sum_{k=1}^{n} c_{jk}\right\} \leq B\sum_{j=1}^{m} y_j^q \leq B. \tag{4.40}$$

Τέλος, χρησιμοποιώντας τις (4.39) και (4.40) για να εκτιμήσουμε το γινόμενο $\left(\sum_{j=1}^{m}\sum_{k=1}^{n} c_{jk}x_k^p\right)^{1/p}\left(\sum_{j=1}^{m}\sum_{k=1}^{n} c_{jk}y_j^q\right)^{1/q}$, παίρνουμε την (4.38). Αυτό ολοκληρώνει την απόδειξη.

Μέρος **II**

Ειδικά θέματα

Κεφάλαιο 5

Η ανισότητα του Carleman

5.1 Εισαγωγή

Το 1923, ο Torsten Carleman απέδειξε την εξής ανισότητα: Αν $\sum_{n=1}^{\infty} a_n$ είναι μια σειρά μη αρνητικών πραγματικών αριθμών, και αν συμβολίσουμε με γ_n τον γεωμετρικό μέσο των $a_1, a_2, \ldots, a_n$, δηλαδή αν θέσουμε

$$\gamma_n = (a_1 a_2 \cdots a_n)^{1/n}, \tag{5.1}$$

τότε,

$$\sum_{n=1}^{\infty} \gamma_n \leq e \sum_{n=1}^{\infty} a_n. \tag{5.2}$$

Ο Carleman χρησιμοποίησε αυτήν την ανισότητα για να δώσει μια νέα απόδειξη ενός θεωρήματος του Denjoy για quasi–αναλυτικές συναρτήσεις. Σκοπός μας σε αυτό το Κεφάλαιο είναι να δούμε αρκετές διαφορετικές αποδείξεις αυτής της ανισότητας.

Υπάρχουν πολλά γνωστά αποτελέσματα για την σύγκλιση ή την απόκλιση σειρών που έχουν σαν όρους κάποιους γραμμικούς συνδυασμούς των όρων μιας ακολουθίας $\{a_n\}$ με μη αρνητικούς όρους. Για παράδειγμα, αν μ_n είναι ο αριθμητικός μέσος των $a_1, a_2, \ldots, a_n$, τότε η $\sum_{n=1}^{\infty} \mu_n$ αποκλίνει πάντοτε, εκτός αν η $\{a_n\}$ είναι η μηδενική ακολουθία: παρατηρήστε ότι αν $a_k > 0$ για κάποιον k, τότε

$$\mu_1 + \mu_2 + \cdots + \mu_k + \cdots + \mu_{n+k} \geq a_k \left(\frac{1`}{k} + \cdots + \frac{1}{n+k} \right), \tag{5.3}$$

και η τελευταία ποσότητα τείνει στο άπειρο όταν $n \to \infty$, διότι η αρμονική σειρά αποκλίνει. Από την ανισότητα αριθμητικού–γεωμετρικού μέσου έχουμε $\gamma_n \leq \mu_n$, με ισότητα (για κάθε n) αν και μόνο αν η ακολουθία είναι σταθερή. Αν λοιπόν η σειρά $\sum_{n=1}^{\infty} a_n$ συγκλίνει και είναι μη μηδενική, η ανισότητα του Carleman εξασφαλίζει την σύγκλιση της $\sum_{n=1}^{\infty} \gamma_n$ και από αυτήν συμπεραίνουμε ότι πρέπει να έχουμε ισχυρά γνήσια ανισότητα στην ανισότητα αριθμητικού–γεωμετρικού μέσου $\gamma_n \leq \mu_n$.

Η δυσκολία της ανισότητας του Carleman έγκειται στο γεγονός ότι ο μετασχηματισμός $\gamma_n = (a_1 a_2 \cdots a_n)^{1/n}$ είναι ισχυρά μη γραμμικός. Είναι πάντως ομογενής: για κάθε $t > 0$, η $\{ta_n\}$ μετασχηματίζεται στην $\{t\gamma_n\}$. Σημαντικό ρόλο σε όλες τις αποδείξεις που θα παρουσιάσουμε παίζει η στοιχειώδης ανισότητα

$$\left(1 + \frac{1}{n}\right)^n < e. \tag{5.4}$$

5.2 Η απόδειξη του Carleman

Το επιχείρημα του Carleman ξεκινάει με μια κλασική μέθοδο που χρησιμοποιείται σε πολλά παρόμοια προβλήματα μετασχηματισμών συγκλινουσών σειρών. Θεωρούμε το αντίστοιχο πρόβλημα για πεπερασμένα αθροίσματα και χρησιμοποιούμε τη μέθοδο των πολλαπλασιαστών Lagrange. Προσπαθούμε δηλαδή να υπολογίσουμε το μέγιστο λ_n της ποσότητας $\gamma_1 + \gamma_2 + \cdots + \gamma_n$ υπό τον περιορισμό $a_1 + a_2 + \cdots + a_n = 1$. Επειδή το πρόβλημά μας είναι ομογενές, έπεται ότι

$$\gamma_1 + \gamma_2 + \cdots + \gamma_n \leq \lambda_n (a_1 + a_2 + \cdots + a_n) \tag{5.5}$$

για κάθε $\{a_n\}$ και κάθε n, οπότε το πρόβλημα ανάγεται στο ερώτημα αν η $\{\lambda_n\}$ είναι άνω φραγμένη. Πιο συγκεκριμένα, πρέπει να δείξουμε ότι $\lambda_n \leq e$ για κάθε n.

Σύμφωνα με τη μέθοδο των πολλαπλασιαστών Lagrange, πρέπει να λύσουμε την εξίσωση $\nabla F = 0$, όπου

$$F = a_1 + (a_1 a_2)^{1/2} + \cdots + (a_1 a_2 \cdots a_n)^{1/n} - \lambda(a_1 + a_2 + \cdots + a_n - 1). \tag{5.6}$$

Έτσι, παίρνουμε το σύστημα εξισώσεων

$$\begin{aligned}
0 &= 1 + \frac{1}{2a_1}\gamma_2 + \cdots + \frac{1}{na_1}\gamma_n - \lambda \\
0 &= \frac{1}{2a_2}\gamma_2 + \cdots + \frac{1}{na_2}\gamma_n - \lambda \\
\cdots &= \cdots \\
0 &= \frac{1}{na_n}\gamma_n - \lambda \\
0 &= -(a_1 + a_2 + \cdots + a_n - 1).
\end{aligned}$$

Πολλαπλασιάζοντας την k-οστή εξίσωση με a_k και προσθέτοντας τις πρώτες n εξισώσεις, παίρνουμε

$$\gamma_1 + \gamma_2 + \cdots + \gamma_n = \lambda(a_1 + a_2 + \cdots + a_n) = \lambda. \tag{5.7}$$

Συνδυάζοντας τις διαδοχικές εξισώσεις ανά δύο μπορούμε να γράψουμε το σύστημα στην ισοδύναμη μορφή

$$\gamma_1 = \lambda_n(a_1 - a_2), \qquad \frac{1}{2}\gamma_2 = \lambda_n(a_2 - a_3), \cdots, \qquad \frac{1}{n}\gamma_n = \lambda_n a_n. \tag{5.8}$$

Αν τώρα θέσουμε

$$t_k = k\left(1 - \frac{a_{k+1}}{a_k}\right) \qquad (k = 1, 2, \ldots, n-1), \tag{5.9}$$

οι εξισώσεις Lagrange γράφονται

$$\lambda_n t_k a_k = \gamma_k \qquad (k = 1, 2, \ldots, n-1), \qquad \lambda_n n a_n = \gamma_n. \tag{5.10}$$

Για $k = 1, 2, \ldots, n-2$, χρησιμοποιώντας την (5.10) και την $\gamma_{k+1}^{k+1} = \gamma_k^k a_{k+1}$ ελέγχουμε εύκολα ότι

$$t_{k+1}^{k+1} = \left(\frac{\gamma_{k+1}}{\lambda_n a_{k+1}}\right)^{k+1} = \frac{1}{\lambda_n} t_k^k \left(\frac{a_k}{a_{k+1}}\right)^k = \frac{1}{\lambda_n} t_k^k \left(1 - \frac{t_k}{k}\right)^{-k}. \tag{5.11}$$

Το θέμα είναι με ποιόν τρόπο μπορεί κανείς να χρησιμοποιήσει αυτές τις αναδρομικές σχέσεις για να καταλήξει στην $\lambda_n \le e$. Ο Carleman ορίζει μια ακολουθία συναρτήσεων $T_k(\lambda)$ με $T_1(\lambda) = 1/\lambda$ και

$$T_{k+1}(\lambda)^{k+1} = \frac{1}{\lambda} T_k(\lambda)^k \left(1 - \frac{T_k(\lambda)}{k}\right)^{-k} \qquad (k = 1, 2, \ldots). \tag{5.12}$$

Από την αναδρομική σχέση (5.11) για τα t_k βλέπουμε ότι $T_k(\lambda_n) = t_k$ για $k = 1, 2, \ldots, n-1$. Επίσης,

$$T_n(\lambda_n)^n = \frac{1}{\lambda_n}\left(\frac{\gamma_{n-1}}{\lambda_n a_{n-1}}\right)^{n-1}\left(\frac{a_n}{a_{n-1}}\right)^{-(n-1)} = \frac{1}{\lambda_n^n}\frac{\gamma_n^n}{a_n^n} = n^n, \tag{5.13}$$

άρα, $T_n(\lambda_n) = n$. Τώρα, δείχνουμε με επαγωγή ότι

$$T_k(\lambda) < \frac{k}{k+1} \qquad (k = 1, 2, \ldots) \tag{5.14}$$

για κάθε $\lambda \ge e$. Αφού $T_n(\lambda_n) = n$, πρέπει αναγκαστικά να ισχύει $\lambda_n < e$, όπως ισχυριστήκαμε.

Είναι φανερό ότι $T_1(\lambda) < 1/2$ αν $\lambda \ge e$. Υποθέτουμε ότι η ανισότητα ισχύει για την T_k. Τότε,

$$T_{k+1}(\lambda)^{k+1} < \frac{1}{\lambda}\left(\frac{\frac{k}{k+1}}{1 - \frac{k}{k(k+1)}}\right)^k = \frac{1}{\lambda} \le \frac{1}{e} < \left(\frac{k+1}{k+2}\right)^{k+1}. \tag{5.15}$$

Συνεπώς,

$$T_{k+1}(\lambda) < \frac{k+1}{k+2} \tag{5.16}$$

και έχουμε πλέον ολοκληρώσει την απόδειξη.

Σημείωση: Ο deBruijn, αναλύοντας πλήρως τις εξισώσεις Lagrange, έδειξε ότι η βέλτιστη σταθερά λ_n ικανοποιεί την

$$\lambda_n = e - \frac{2\pi^2 e}{(\log n)^2} + O\left(\frac{1}{(\log n)^3}\right). \tag{5.17}$$

5.3 Η απόδειξη του Pólya

Το 1926, ο George Pólya έδωσε μια κομψή απόδειξη της ανισότητας του Carleman χρησιμοποιώντας την ανισότητα αριθμητικού–γεωμετρικού μέσου και (ουσιαστικά) τίποτε άλλο.

Η «απόδειξη»

Η απόδειξη του Pólya είναι πολύ σύντομη αλλά μυστηριώδης: ορίζουμε $c_1, c_2, c_3, \ldots$ επαγωγικά, έτσι ώστε να ικανοποιείται η σχέση

$$c_1 c_2 \cdots c_n = (n+1)^n \tag{5.18}$$

για κάθε n. Με άλλα λόγια,

$$c_k = \frac{(k+1)^k}{k^{k-1}}. \tag{5.19}$$

Τότε, μπορούμε να γράψουμε

$$\begin{aligned}
\sum_{n=1}^{\infty} (a_1 a_2 \cdots a_n)^{1/n} &= \sum_{n=1}^{\infty} \frac{(a_1 c_1 a_2 c_2 \cdots a_n c_n)^{1/n}}{n+1} \\
&\leq \sum_{n=1}^{\infty} \frac{a_1 c_1 + a_2 c_2 + \cdots + a_n c_n}{n(n+1)} \\
&= \sum_{k=1}^{\infty} a_k c_k \sum_{n=k}^{\infty} \frac{1}{n(n+1)} \\
&= \sum_{k=1}^{\infty} a_k c_k \sum_{k=n}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) \\
&= \sum_{k=1}^{\infty} a_k \frac{(k+1)^k}{k^{k-1}} \frac{1}{k} \\
&\leq e \sum_{k=1}^{\infty} a_k.
\end{aligned}$$

Το πρόβλημα της παρουσίασης

Αρκετά χρόνια αργότερα, ο Pólya επέλεξε αυτήν ακριβώς την απόδειξη για να θεμελιώσει την άποψή του σχετικά με τη δομή της «ιδανικής μαθηματικής διάλεξης». Αν κάποιο ακροατήριο παρακολουθήσει την απόδειξη της προηγούμενης παραγράφου, όλοι θα συμφωνήσουν ότι το κρίσιμο σημείο είναι ο τρόπος ορισμού των $c_1, c_2, c_3, \ldots$. Οι συντελεστές αυτοί δίνονται κατευθείαν στο ξεκίνημα του επιχειρήματος, χωρίς καμία προετοιμασία ή αιτιολόγηση. Ο Pólya αναφέρει τις πιθανές αντιδράσεις των ακροατών:

- «Μοιάζει με το λαγό που βγαίνει από το καπέλο».
- «Ξεπετάγονται από το πουθενά. Μοιάζουν τόσο αυθαίρετοι. Ποιό είναι το κίνητρο για την επιλογή τους;»
- «Δεν μου αρέσει να βαδίζω στο σκοτάδι. Δεν μου αρέσει να κάνω βήματα που δεν βλέπω γιατί θα με φέρουν πιο κοντά στο στόχο».
- «Εσύ μπορεί να γνωρίζεις γιατί κάνεις αυτό το βήμα, εγώ όμως όχι, και συνεπώς δεν μπορώ να σε παρακολουθήσω».

- «Κοίταξε, δεν βρίσκομαι εδώ μόνο και μόνο για να σε θαυμάζω. Θέλω να μάθω πώς να λύνω κι εγώ προβλήματα. Όμως, δεν μπορώ να δώ πώς είναι ανθρωπίνως δυνατόν να μου έρθει η ιδέα του ορισμού σου. Τι με μαθαίνεις με αυτό που μου έδειξες;»
- «Αυτό το βήμα δεν είναι τετριμμένο. Μοιάζει το πιο σημαντικό. Αν μπορούσα να πειστώ ότι έχει κάποιες πιθανότητες επιτυχίας ή αν μου έδινες κάποια αιτιολόγηση για την επιλογή σου, θα μπορούσα ίσως να φανταστώ πώς το σκέφτηκες και να παρακολουθήσω την συνέχεια με μεγαλύτερη αυτοπεποίθηση και ενδιαφέρον».

Οι πρώτες αντιδράσεις είναι απλώς θυμωμένες, οι επόμενες μπαίνουν κάπως περισσότερο στην ουσία, η δε τελευταία είναι η καλύτερη: φανερώνει έναν σκεπτόμενο ακροατή που ζητάει δύο βασικά πράγματα: να δεί ότι το συγκεκριμένο βήμα της απόδειξης είναι σωστό αλλά και να δεί ότι αυτό το βήμα είναι το ενδεδειγμένο.

Κατά τον Pólya, ένα βήμα μιας μαθηματικής απόδειξης είναι κατάλληλο αν συνδέεται ουσιαστικά με τον τελικό σκοπό και μας φέρνει πιο κοντά σ' αυτόν. Δεν αρκεί όμως να είναι κατάλληλο: θα πρέπει και *να φαίνεται κατάλληλο* στον ακροατή. Αν το βήμα είναι απλό, ένα στοιχειώδες βήμα ρουτίνας, ο ακροατής μπορεί εύκολα να φανταστεί πώς αυτό συνδέεται με τον στόχο της απόδειξης. Αν πάλι η παρουσίαση είναι πολύ προσεκτικά προετοιμασμένη, τότε μπορεί η σύνδεση του βήματος με τον στόχο να υποδεικνύεται επαρκώς. Αν όμως, παρόλο που το βήμα φαίνεται σημαντικό, η σχέση του με τον στόχο δεν είναι καθόλου ορατή, τότε ο ακροατής έχει κάθε δικαίωμα να αισθάνεται απογοητευμένος.

Στο παράδειγμα του Pólya, οι συντελεστές c_n εμφανίζονται σαν «από μηχανής θεός». Το επιχείρημα που αποδεικνύει την ανισότητα βασίζεται ακριβώς στην επιλογή αυτών των συντελεστών, και αποδεικνύει την ανισότητα πεντακάθαρα και γρήγορα. Το πρόβλημα όμως είναι ότι το βήμα δικαιώνεται στο τέλος, ενώ στην αρχή μοιάζει ακατανόητο.

Πώς λοιπόν θα μπορούσε ο ομιλητής να το εξηγήσει από την αρχή; Η πλήρης αιτιολόγηση απαιτεί αρκετό χρόνο και χώρο: θα δοθεί στις επόμενες παραγράφους. Κατά τον Pólya, αυτό που χρειάζεται είναι μια *μερική αιτιολόγηση*, κάποιες *πειστικές νύξεις* για τους λόγους που οδήγησαν φυσιολογικά στο συγκεκριμένο βήμα.

Η απόδειξη

Το μυστικό πίσω από την απόδειξη του Pólya ήταν η πίστη του στη γενική αρχή ότι πρέπει κανείς να χρησιμοποιεί μια ανισότητα στις περιπτώσεις όπου είναι πιό ακριβής. Θυμηθείτε ότι θέλουμε να δείξουμε ότι για κάθε ακολουθία θετικών πραγματικών αριθμών $a_1, a_2, \ldots$ ισχύει η ανισότητα

$$\sum_{k=1}^{\infty}(a_1a_2\cdots a_k)^{1/k} \le e\sum_{k=1}^{\infty}a_k. \tag{5.20}$$

Η εμπειρία μας από την ανισότητα του Cauchy για σειρές υποδεικνύει ότι ένας χρήσιμος τρόπος να προσεγγίσουμε ένα *ποσοτικό* αποτέλεσμα όπως η ανισότητα (5.20) είναι να μελετήσουμε πρώτα ένα απλούστερο *ποιοτικό* αποτέλεσμα: για

παράδειγμα, να δείξουμε ότι

$$\sum_{k=1}^{\infty} a_k < \infty \quad \Longrightarrow \quad \sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} < \infty. \tag{5.21}$$

Η **υπόθεση** είναι ότι τα μερικά αθροίσματα της $\sum_{k=1}^{\infty} a_k$ είναι φραγμένα, δηλαδή

τα αθροίσματα $a_1 + a_2 + \cdots + a_n$ δεν είναι μεγάλα.

Το **συμπέρασμα** είναι ότι η σειρά $\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k}$ συγκλίνει, δηλαδή

οι όροι $(a_1 a_2 \cdots a_n)^{1/n}$ είναι μικροί.

Γνωρίζουμε κάποιο θεώρημα που θα μπορούσε να φανεί χρήσιμο; Η φυσιολογική επιλογή θα ήταν να εφαρμόσουμε την ανισότητα αριθμητικού–γεωμετρικού μέσου στους προσθετέους του δεξιού μέλους της (5.21), να κάνουμε τους απαιτούμενους υπολογισμούς, και να ελπίζουμε ότι θα είμαστε τυχεροί. Με αυτό το σχέδιο, οδηγούμαστε στην ανισότητα

$$\sum_{k=1}^{n} (a_1 a_2 \cdots a_k)^{1/k} \le \sum_{k=1}^{n} \frac{1}{k} \sum_{j=1}^{k} a_j = \sum_{j=1}^{n} a_j \sum_{k=j}^{n} \frac{1}{k}, \tag{5.22}$$

και, όπως ίσως θα περιμέναμε, βλέπουμε ότι το σχέδιο δεν δουλεύει. Όταν το $n \to \infty$, το άνω φράγμα αποκλίνει, και βλέπουμε ότι η απλοϊκή εφαρμογή της ανισότητας αριθμητικού–γεωμετρικού μέσου δεν είναι αρκετή.

Για να φτάσουμε στην ουσία, θα πρέπει να αναρωτηθούμε γιατί απέτυχε η ανισότητα αριθμητικού–γεωμετρικού μέσου και τι θα μπορούσαμε να κάνουμε για να ξεπεράσουμε αυτήν την αποτυχία.

Μαθαίνοντας από την αποτυχία

Από την υπόθεση για το αριστερό μέλος της ανισότητας (5.21), η σειρά $a_1 + a_2 + \cdots$ συγκλίνει, και αυτό το γεγονός μπορεί να μας δείξει την αιτία της αποτυχίας μας. Η σύγκλιση της σειράς έχει σαν συνέπεια το ότι σε κάθε μακρύ αρχικό τμήμα $a_1, a_2, \ldots, a_n$ υπάρχουν όροι οι οποίοι είναι «εντελώς άνισοι», και γνωρίζουμε ότι σε αυτές τις περιπτώσεις η ανισότητα αριθμητικού–γεωμετρικού μέσου δεν είναι ιδιαίτερα αποτελεσματική. Μπορούμε να βρούμε κάποιον τρόπο για να εφαρμόσουμε την ανισότητα αριθμητικού–γεωμετρικού μέσου πάνω σε αριθμούς οι οποίοι να είναι «σχεδόν ίσοι»;

Αφού γνωρίζουμε πολύ λίγα πράγματα για τους αριθμούς a_k, δεν ξέρουμε ακριβώς τι πρέπει να κάνουμε, μια ιδέα όμως που έρχεται στο μυαλό είναι η εξής: να πολλαπλάσιάσουμε κάθε αριθμό a_k με κάποιον παράγοντα c_k, τον οποίο θα προσπαθήσουμε να επιλέξουμε όταν θα καταλάβουμε καλύτερα τι συμβαίνει, ώστε η ακολουθία των γινομένων $c_1 a_1, c_2 a_2, \ldots$ να έχει όρους «σχεδόν ίσους», ή τουλάχιστον πιό ισορροπημένους απ' ότι η αρχική μας ακολουθία. Φυσικά, ο μόνος τρόπος για να κάνουμε αυτήν την ιδέα πιό συγκεκριμένη είναι να προχωρήσουμε σε υπολογισμούς.

Φυσιολογικές δοκιμές

Θα μπορούσαμε να δοκιμάσουμε τους αριθμούς

$$1 \cdot a_1, 2a_2, 3a_3, \ldots, na_n, \ldots$$

ή, για να έχουμε μεγαλύτερη ελευθερία, τους

$$1^s a_1, 2^s a_2, 3^s a_3, \ldots, n^s a_n, \ldots,$$

όπου s μια *παράμετρος* την οποία θα επιλέξουμε με τον βέλτιστο τρόπο. Ακολουθώντας τα βήματα του αρχικού υπολογισμού μας, γράφουμε

$$\begin{aligned}
\sum_{k=1}^{\infty}(a_1 a_2 \cdots a_k)^{1/k} &= \sum_{k=1}^{\infty} \frac{(a_1 1^s \cdot a_2 2^s \cdots a_k k^s)^{1/k}}{(1 \cdot 2 \cdots k)^{s/k}} \\
&\leq \sum_{k=1}^{\infty} \frac{a_1 1^s + a_2 2^s + \cdots + a_k k^s}{k(k!)^{s/k}} \\
&= \sum_{k=1}^{\infty} a_k k^s \sum_{j=k}^{\infty} \frac{1}{j(j!)^{s/j}}.
\end{aligned}$$

Σε αυτό το σημείο αρχίζουν οι δυσκολίες: δεν μπορούμε να υπολογίσουμε ακριβώς το τελευταίο άθροισμα. Αναγκαζόμαστε λοιπόν να ανατρέξουμε στην εκτίμηση

$$(n!)^{1/n} \simeq n/e, \tag{5.23}$$

απ' όπου προκύπτει ότι

$$\begin{aligned}
\sum_{n=k}^{\infty} \frac{1}{n(n!)^{s/n}} &\simeq e^s \sum_{n=k}^{\infty} n^{-(s+1)} \simeq e^s \int_k^{\infty} x^{-(s+1)} dx \\
&= e^s s^{-1} k^{-s}.
\end{aligned}$$

Εισάγοντας αυτήν την εκτίμηση στο προηγούμενο φράγμα μας, βλέπουμε ότι βρισκόμαστε πολύ κοντά στο να δείξουμε ότι

$$\sum_{k=1}^{\infty}(a_1 a_2 \cdots a_k)^{1/k} \leq C \sum_{k=1}^{\infty} a_k, \tag{5.24}$$

όπου C είναι μια σταθερά περίπου ίση με $e^s s^{-1}$. Αυτή είναι φυσικά ακριβώς η ανισότητα του Carleman.

Κοιτάζοντας ξανά τον υπολογισμό που κάναμε, παρατηρούμε ότι η βέλτιστη επιλογή της παραμέτρου s είναι ίσως εκείνη για την οποία ελαχιστοποιείται η ποσότητα $e^s s^{-1}$. Με απλή μελέτη αυτής της συνάρτησης του s, βλέπουμε ότι η βέλτιστη τιμή είναι η $s = 1$. Αυτό σημαίνει ότι η βέλτιστη επιλογή του c_k είναι $c_k = k$ ή κάποια ποσότητα που συμπεριφέρεται περίπου σαν το k όταν το k είναι μεγάλο. Με μια τέτοια επιλογή μοιάζει εφικτό να αποδείξουμε την (5.24) με $C = e$.

Η αρχή της μέγιστης αποτελεσματικότητας

Η συζήτηση της προηγούμενης παραγράφου μας δίνει περισσότερη αυτοπεποίθηση. Θεωρούμε μεταβλητές c_k και επαναλαμβάνουμε τον υπολογισμό που κάναμε για την (ειδική) επιλογή $c_k = k^s$:

$$\begin{aligned}\sum_{k=1}^{\infty}(a_1a_2\cdots a_k)^{1/k} &= \sum_{k=1}^{\infty}\frac{(a_1c_1a_2c_2\cdots a_kc_k)^{1/k}}{(c_1c_2\cdots c_k)^{1/k}}\\ &\le \sum_{k=1}^{\infty}\frac{a_1c_1+a_2c_2+\cdots+a_kc_k}{k(c_1c_2\cdots c_k)^{1/k}}\\ &= \sum_{k=1}^{\infty}a_kc_k\sum_{j=k}^{\infty}\frac{1}{j(c_1c_2\cdots c_j)^{1/j}}.\end{aligned}$$

Έδώ, πρέπει φυσικά να σταθούμε και να σκεφτούμε. Από την σχέση στην οποία καταλήξαμε, βλέπουμε ότι για την απόδειξη της ανισότητας του Carleman αρκεί να επιλέξουμε τους παράγοντες c_k, $k = 1, 2, \ldots$ έτσι ώστε τα αθροίσματα

$$s_k = c_k\sum_{j=k}^{\infty}\frac{1}{j(c_1c_2\cdots c_j)^{1/j}}, \qquad k = 1, 2, \ldots \tag{5.25}$$

να ικανοποιουν την $s_k \le e$.

Όποιον δρόμο κι αν ακολουθήσουμε για την επιλογή των c_k, τελικά θα χρειαστεί να εκτιμήσουμε το άθροισμα s_k. Καλό είναι λοιπόν να κάνουμε αυτή τη δουλειά όσο γίνεται ευκολότερη. Η εμπειρία δείχνει ότι υπάρχουν πολύ λίγες σειρές με ουρές που να υπολογίζονται ακριβώς με εύκολο τρόπο. Για την ακρίβεια, όλες έχουν σχέση με την τηλεσκοπική ταυτότητα

$$\sum_{j=k}^{\infty}\left\{\frac{1}{b_j}-\frac{1}{b_{j+1}}\right\} = \frac{1}{b_k}. \tag{5.26}$$

Η απλούστερη ίσως επιλογή που μας παρέχει αυτή η ταυτότητα είναι η

$$\sum_{j=k}^{\infty}\frac{1}{j(j+1)} = \sum_{j=k}^{\infty}\left\{\frac{1}{j}-\frac{1}{j+1}\right\} = \frac{1}{k} \tag{5.27}$$

και, συγκρίνοντας τα αθροίσματα (5.25) και (5.26), βλέπουμε ότι το s_k παίρνει την απλούστερη δυνατή μορφή αν ορίσουμε τους συντελεστές c_k μέσω της πεπλεγμένης αναδρομικής σχέσης

$$(c_1c_2\cdots c_j)^{1/j} = j+1, \qquad j = 1, 2, \ldots. \tag{5.28}$$

Αυτή η επιλογή μας δίνει έναν σύντομο τύπο για το s_k:

$$s_k = c_k\sum_{j=k}^{\infty}\frac{1}{j(c_1c_2\cdots c_j)^{1/j}} = c_k\sum_{j=k}^{\infty}\frac{1}{j(j+1)} = \frac{c_k}{k}, \tag{5.29}$$

και το μόνο που χρειάζεται τώρα είναι να εκτιμήσουμε το μέγεθος των c_k.

Ευτυχώς, αυτή η εκτίμηση γίνεται εύκολα. Χρησιμοποιώντας δύο φορές την πεπλεγμένη αναδρομική σχέση (5.28) για τους c_j, βλέπουμε ότι

$$c_1 c_2 \cdots c_{j-1} = j^{j-1} \quad \text{και} \quad c_1 c_2 \cdots c_j = (j+1)^j.$$

Διαιρώντας, παίρνουμε

$$c_j = \frac{(j+1)^j}{j^{j-1}} = j\left(1+\frac{1}{j}\right)^j. \tag{5.30}$$

Από αυτόν τον τύπο και από την αρχική ανισότητα (5.25), συμπεραίνουμε ότι

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} \le \sum_{k=1}^{\infty} \left(1+\frac{1}{k}\right)^k a_k, \tag{5.31}$$

και αυτή η ανισότητα μας φέρνει στην ανισότητα του Carleman. Για την ακρίβεια, η (5.31) είναι ισχυρότερη από την ανισότητα του Carleman: αν θέσουμε $x = 1/k$ στην γνωστή ανισότητα $1 + x \le e^x$, βλέπουμε ότι

$$\left(1+\frac{1}{k}\right)^k < e \quad \text{για κάθε } k = 1, 2, \ldots.$$

Το επιχείρημα του Pólya που περιγράψαμε παραπάνω, πέρα από την κομψότητα που συχνά συναντάμε στα Μαθηματικά, περιέχει μια πολύ βαθιά ιδέα που αξίζει τον κόπο να προσέξουμε. Η ιδέα που πρέπει να συγκρατήσουμε είναι ότι μπορούμε να βελτιώσουμε σημαντικά την αποτελεσματικότητα μιας ανισότητας αν απλώς αναδιατυπώσουμε το πρόβλημα ώστε η ανισότητα να εφαρμοστεί κάτω από προϋποθέσεις που εξασφαλίζουν ότι ισχύει «περίπου» σαν ισότητα. Η απόδειξη του Pólya για την ανισότητα του Carleman δίνει ένα εξαιρετικά κομψό παράδειγμα εφαρμογής αυτής της ιδέας, υπάρχουν όμως πολλές άλλες περιπτώσεις στις οποίες η ίδια ιδέα μπορεί να αξιοποιηθεί με εξίσου θαυμάσια αποτελέσματα.

5.4 Η απόδειξη του Knopp

Η απόδειξη του Knopp (1928) είναι ίσως η συντομότερη. Ουσιαστικά, αξιοποιεί κι αυτή τις ιδέες της απόδειξης του Pólya. Ορίζουμε

$$c_n = \frac{a_1 + 2a_2 + \cdots + na_n}{n(n+1)}. \tag{5.32}$$

Αναλύοντας σε απλά κλάσματα και αθροίζοντας κατά μέρη παίρνουμε

$$\sum_{n=1}^{N} c_n = \sum_{n=1}^{N} a_n - N c_N < \sum_{n=1}^{N} a_n. \tag{5.33}$$

Εφαρμόζοντας τώρα την ανισότητα αριθμητικού–γεωμετρικού μέσου για τους αριθμούς

$$\frac{a_1}{n+1}, \frac{2a_2}{n+1}, \ldots, \frac{na_n}{n+1}, \tag{5.34}$$

βλέπουμε ότι

$$c_n \geq \left[\frac{n!a_1a_2\cdots a_n}{(n+1)^n}\right]^{1/n} = \frac{(n!)^{1/n}}{n+1}\gamma_n. \tag{5.35}$$

Για να ολοκληρώσουμε την απόδειξη χρησιμοποιούμε την γνωστή ανισότητα $(n+1)^n \leq e^n n!$. Έπεται ότι $\gamma_n \leq ec_n$, και από την (5.33) προκύπτει η ζητούμενη ανισότητα.

5.5 Η απόδειξη του Redheffer

Η απόδειξη του Redheffer (1967) μοιάζει ακόμα πιο απροσδόκητη, δείχνει όμως πόσο χρήσιμη είναι η εισαγωγή πρόσθετων μεταβλητών για τη μελέτη ενός προβλήματος. Για κάθε επιλογή θετικών πραγματικών αριθμών a_n και για κάθε επιλογή μη αρνητικών πραγματικών αριθμών b_n ισχύει (!) η ανισότητα

$$(b_1-1)\gamma_1+2(b_2-1)\gamma_2+\cdots+n(b_n-1)\gamma_n+n\gamma_n \leq a_1b_1+a_2b_2^2+\cdots+a_nb_n^n. \tag{5.36}$$

Για την απόδειξη αυτής της ανισότητας, απομονώνουμε πρώτα την μεταβλητή b_n και προσπαθούμε να βρούμε εκείνην την επιλογή της για την οποία είναι δυσκολότερο να ισχύει η ανισότητα. Με άλλα λόγια, ζητάμε την τιμή του x που μεγιστοποιεί την συνάρτηση $\phi(x) = n\gamma_n x - a_n x^n$ πάνω από όλους τους μη αρνητικούς x. Με στοιχειώδη απειροστικό λογισμό βλέπουμε ότι το μέγιστο συμβαίνει όταν $n\gamma_n = na_nx^{n-1}$, δηλαδή όταν $x^{n-1} = \gamma_n/a_n$. Γι' αυτήν την τιμή του x παρατηρούμε ότι

$$\phi(x) = x(n-1)\gamma_n = (n-1)[\gamma_n^n/a_n]^{1/(n-1)} = (n-1)\gamma_{n-1}. \tag{5.37}$$

Έτσι, η αρχική μας ανισότητα θα ισχύει για κάθε $b_n \geq 0$ αν ισχύει στην περίπτωση που, επιπλέον, ικανοποιείται η

$$nb_n\gamma_n - a_nb_n^n = (n-1)\gamma_{n-1}. \tag{5.38}$$

Με αυτήν την αντικατάσταση, η ανισότητα μετασχηματίζεται στην

$$\begin{aligned}(b_1-1)\gamma_1+2(b_2-1)\gamma_2+\cdots+(n-1)(b_{n-1}-1)\gamma_{n-1}+(n-1)\gamma_{n-1}\\ \leq a_1b_1+a_2b_2^2+\cdots+a_{n-1}b_{n-1}^{n-1}.\end{aligned}$$

Αυτή όμως είναι ακριβώς η ίδια ανισότητα με την αρχική, με $n-1$ όρους αντί για n. Έχουμε έτσι ένα *σχήμα αναδρομικής ανισότητας* (με την ορολογία του Redheffer), στο οποίο το πλήθος των παραμέτρων μειώνεται διαδοχικά με τέτοιον τρόπο που η αρχική ανισότητα θα έχει αποδειχθεί για κάθε n αν καταφέρουμε να αποδείξουμε ότι ισχύει για $n = 1$. Όμως, στην περίπτωση $n = 1$, η ανισότητα ισχυρίζεται απλώς ότι $a_1b_1 \leq a_1b_1$. Έτσι, έχουμε αποδείξει την ανισότητα (5.36).

Στην προηγούμενη απόδειξη δεν χρησιμοποιήσαμε την ανισότητα αριθμητικού–γεωμετρικού μέσου σε κανένα σημείο. Παρατηρήστε όμως ότι αν επιλέξουμε $b_k = 1$ για κάθε k, τότε η ανισότητα του Redheffer παίρνει τη μορφή

$$n\gamma_n \leq a_1 + a_2 + \cdots + a_n. \tag{5.39}$$

Δηλαδή, η (5.36) έχει σαν ειδική περίπτωση την ανισότητα αριθμητικού–γεωμετρικού μέσου.

Για την ανισότητα του Carleman, επιλέγουμε $b_k = 1 + \frac{1}{k}$ στην (5.36). Τότε, έχουμε

$$\gamma_1+\gamma_2+\cdots+\gamma_n+n\gamma_n \le (1+1)^1 a_1+\left(1+\frac{1}{2}\right)^2 a_2+\cdots+\left(1+\frac{1}{n}\right)^n a_n. \tag{5.40}$$

Για την ακρίβεια, η ανισότητα αυτή αποδείχθηκε με την επιπλέον υπόθεση ότι $a_k > 0$ για κάθε k, λόγω συνέχειας όμως εξακολουθεί να ισχύει αν απλώς υποθέσουμε ότι $a_k \ge 0$. Πετώντας τον όρο $n\gamma_n$ από το αριστερό μέλος, χρησιμοποιώντας την $\left(1+\frac{1}{k}\right)^k < e$ και αφήνοντας το $n \to \infty$, παίρνουμε την ανισότητα του Carleman.

Σημείωση: Μια προσεκτική ανάγνωση της απόδειξης του Redheffer δείχνει ότι ισότητα ισχύει στην (5.40) αν και μόνο αν

$$a_k = 2a_1 \frac{k^{k-1}}{(k+1)^k} = O\left(\frac{1}{k}\right). \tag{5.41}$$

5.6 Το συνεχές ανάλογο της ανισότητας του Carleman

Κλείνουμε αυτό το Κεφάλαιο με δύο ολοκληρωτικές ανισότητες που μπορούν να θεωρηθούν ως *συνεχή ανάλογα* της ανισότητας του Carleman. Όπως θα δούμε μάλιστα, μπορούμε να περάσουμε από αυτές στην «διακριτή» ανισότητα του Carleman.

Μια ανισότητα του Knopp

Οι Hardy, Littlewood και Pólya αποδίδουν στον Knopp την ανισότητα

$$\int_0^\infty \exp\left(\frac{1}{x}\int_0^x \log f(t)\,dt\right) dx < e\int_0^\infty f(x)\,dx, \tag{5.42}$$

όπου f είναι τυχούσα μη αρνητική ολοκληρώσιμη συνάρτηση στο $(0,\infty)$.

Για την απόδειξη χρειάζεται να θυμηθούμε τα εξής: (α) Αν ορίσουμε μια συνάρτηση g μέσω της

$$g(x) = \frac{1}{x^2}\int_0^x y f(y)\,dy, \tag{5.43}$$

τότε

$$\int_0^\infty f(x)\,dx = \int_0^\infty g(x)\,dx. \tag{5.44}$$

Αυτή η ισότητα είναι άμεση συνέπεια του θεωρήματος του Fubini, αφού

$$\begin{aligned}\int_0^\infty g(x)\,dx &= \int_0^\infty \frac{1}{x^2}\int_0^x y f(y)\,dy\,dx\\ &= \int_0^\infty y f(y)\int_y^\infty \frac{1}{x^2}\,dx\,dy\\ &= \int_0^\infty f(y)\,dy.\end{aligned}$$

(β) Χρειαζόμαστε επίσης την εξής μορφή της ανισότητας του Jensen: αν Φ είναι μια κυρτή συνάρτηση, ϕ είναι μια ολοκληρώσιμη συνάρτηση, και μ είναι ένα μέτρο πιθανότητας, τότε

$$\Phi\left(\int \phi\, d\mu\right) \leq \int \Phi(\phi)\, d\mu. \tag{5.45}$$

Θα χρειαστούμε αυτήν την ανισότητα μόνο στην περίπτωση που $\Phi = \exp$ και $d\mu = dx/(b-a)$ στο διάστημα $[a,b]$.

Παρατηρούμε ότι η $x\log x - x$ είναι αντιπαράγωγος της $\log x$. Ξεκινώντας από την $\log f(y) = \log(yf(y)) - \log y$, γράφουμε

$$\begin{aligned}
\exp\left(\frac{1}{x}\int_0^x \log\, f(y)\, dy\right) &= \exp\left(\frac{1}{x}\int_0^x \log(yf(y))\, dy - \frac{1}{x}\int_0^x \log y\, dy\right) \\
&= e^{1-\log x}\exp\left(\frac{1}{x}\int_0^x \log(yf(y))\, dy\right) \\
&\leq \frac{e}{x}\cdot\frac{1}{x}\int_0^x \exp(\log(yf(y)))\, dy \\
&= \frac{e}{x^2}\int_0^x yf(y)\, dy,
\end{aligned}$$

όπου – στο προτελευταίο βήμα – χρησιμοποιήσαμε την ανισότητα του Jensen. Ολοκληρώνοντας αυτήν την ανισότητα στο $(0,\infty)$ και εφαρμόζοντας την ταυτότητα (5.44), παίρνουμε την ανισότητα (5.42).

Μια ανισότητα του Carleson

Ο Carleson αποδεικνύει την εξής ανισότητα: αν $m(x)$ είναι μια κυρτή συνάρτηση στο $[0,\infty)$ με $m(0)=0$, τότε για κάθε $p > -1$ ισχύει η ανισότητα

$$I_1 := \int_0^\infty x^p e^{-m(x)/x} dx \leq e^{p+1}\int_0^\infty x^p e^{-m'(x)} dx := e^{p+1} I_2. \tag{5.46}$$

Για την απόδειξη, παρατηρούμε ότι – λόγω της κυρτότητας της m – για κάθε $k > 1$ ισχύει

$$m(kx) \geq m(x) + (k-1)m'(x), \tag{5.47}$$

και συνεπώς,

$$\begin{aligned}
k^{-p-1}\int_0^A x^p e^{-m(x)/x} dx &= \int_0^{A/k} x^p e^{-\frac{m(kx)}{kx}} dx \\
&\leq \int_0^A x^p e^{-\frac{m(kx)}{kx}} dx \\
&\leq \int_0^A x^p e^{-\frac{m(x)}{kx} - \frac{k-1}{k}m'(x)} dx \\
&\leq \left(\int_0^A x^p e^{-m(x)/x} dx\right)^{1/k} \left(\int_0^A x^p e^{-m'(x)} dx\right)^{(k-1)/k}
\end{aligned}$$

για κάθε $A > 0$. Αφήνοντας το $A \to \infty$ συμπεραίνουμε ότι

$$I_1 \leq k^{\frac{(p+1)k}{k-1}} I_2. \tag{5.48}$$

Παρατηρώντας ότι

$$\lim_{k \to 1} k^{k/(k-1)} = e, \tag{5.49}$$

παίρνουμε το ζητούμενο.

Η σχέση με την ανισότητα του Carleman

Αν θέσουμε $p = 0$ στην ανισότητα του Carleson, παίρνουμε την ανισότητα

$$\int_0^\infty e^{-m(x)/x} dx \leq e \int_0^\infty e^{-m'(x)} dx, \tag{5.50}$$

η οποία συμπίπτει με την ανισότητα του Knopp αν κάνουμε την επιπλέον υπόθεση ότι η f είναι φθίνουσα: αρκεί να θέσουμε $f(x) = e^{-m(x)}$.

Από την (5.50) μπορούμε να αποδείξουμε την «διακριτή» ανισότητα του Carleman, επιλέγοντας κατάλληλη συνάρτηση m. Κατ' αρχήν, ο Carleson παρατηρεί ότι προκειμένου να αποδείξουμε την ανισότητα του Carleman, μπορούμε να υποθέσουμε ότι η ακολουθία $\{a_n\}$ είναι φθίνουσα: αυτό είναι σωστό διότι, από όλες τις αναδιατάξεις της $\{a_n\}$, εκείνη που δίνει τη μεγαλύτερη σειρά γεωμετρικών μέσων γ_n είναι αυτή για την οποία η $\{a_n\}$ είναι φθίνουσα.

Αν μας δώσουν μια τέτοια φθίνουσα ακολουθία $\{a_n\}$, θεωρούμε την συνάρτηση m που είναι κατά τμήματα γραμμική και ενώνει τα σημεία

$$(0,0), \left(1, \log \frac{1}{a_1}\right), \left(2, \log \frac{1}{a_1} + \log \frac{1}{a_2}\right), \ldots.$$

Τότε, μπορούμε να ελέγξουμε ότι

$$\int_0^\infty e^{m'(x)} dx = \sum_{n=1}^\infty a_n \tag{5.51}$$

και

$$\int_n^{n+1} e^{-m(x)/x} dx \geq \int_n^{n+1} e^{-m(x)/n} dx = \gamma_n, \tag{5.52}$$

όπου στο τελευταίο βήμα χρησιμοποιούμε το γεγονός ότι η m είναι κυρτή με $m(0) = 0$, οπότε οι κλίσεις των χορδών που περνούν από το $(0,0)$ αυξάνουν με το x. Έπεται ότι

$$\sum_{n=1}^\infty \gamma_n \leq e \sum_{n=1}^\infty a_n. \tag{5.53}$$

Κεφάλαιο 6

Αθροίσματα τετραγώνων πολυωνύμων

6.1 Η εικασία του Minkowski

Για την απόδειξη της ταυτότητας του Lagrange ξεκινήσαμε με ένα πολυώνυμο που γνωρίζαμε ότι είναι μη αρνητικό, και, χρησιμοποιώντας στοιχειώδη άλγεβρα, δείξαμε ότι μπορούσε να γραφτεί σαν άθροισμα τετραγώνων. Η ταυτότητα που προέκυψε μας έδωσε μια ανεξάρτητη απόδειξη της ανισότητας του Cauchy και, ταυτόχρονα, μια διάφανη ερμηνεία για τις ικανές και αναγκαίες συνθήκες ώστε να ισχύει ισότητα. Σε αυτό το Κεφάλαιο θα συζητήσουμε το εξής πρόβλημα:

Πρόβλημα. *Είναι σωστό ότι κάθε μη αρνητικό πολυώνυμο γράφεται σαν άθροισμα τετραγώνων; Δηλαδή, αν το πραγματικό πολυώνυμο $P(x_1, x_2, \ldots, x_n)$ ικανοποιεί την*

$$P(x_1, x_2, \ldots, x_n) \geq 0 \quad \textit{για κάθε } (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n,$$

μπορούμε να βρούμε $s \in \mathbb{N}$ και πραγματικά πολυώνυμα $Q_k(x_1, x_2, \ldots, x_n)$, $1 \leq k \leq s$, ώστε

$$P(x_1, x_2, \ldots, x_n) = Q_1^2 + Q_2^2 + \cdots + Q_s^2;$$

Το ερώτημα αυτό τέθηκε από τον Minkowski, ο οποίος έκανε την εικασία ότι η απάντηση είναι αρνητική. Αυτό που οδήγησε τον Minkowski στο πρόβλημα δεν είναι εντελώς καθαρό: κατά πάσα πιθανότητα, ενδιαφερόταν για κάποια αριθμοθεωρητικά αποτελέσματα όπως το κλασσικό θεώρημα του Lagrange το οποίο ισχυρίζεται ότι κάθε φυσικός αριθμός είναι το άθροισμα τεσσάρων ή λιγότερων τέλειων τετραγώνων. Όπως και να είναι, ο Minkowski ανακοίνωσε την εικασία του στον David Hilbert, και το 1888, ο Hilbert δημοσίευσε μια απόδειξη για την ύπαρξη μη αρνητικών πολυωνύμων τα οποία δεν γράφονται σαν άθροισμα τετραγώνων πραγματικών πολυωνύμων. Πιο συγκεκριμένα, ο Hilbert απέδειξε ότι η απάντηση είναι καταφατική μόνο στις εξής περιπτώσεις: όταν το πολυώνυμο είναι

μιας μεταβλητής και ο βαθμός του P αυθαίρετος, όταν ο βαθμός του P είναι 2 και το πλήθος των μεταβλητών αυθαίρετο, και τέλος, όταν το P είναι πολυώνυμο με δύο μεταβλητές και έχει βαθμό 4. Η απόδειξη του Hilbert ήταν μακροσκελής, πολύπλοκη και έμμεση.

Τα πρώτα συγκεκριμένα παραδείγματα μη αρνητικών πολυωνύμων που δεν γράφονται στη μορφή αθροίσματος τετραγώνων πραγματικών πολυωνύμων δόθηκαν το 1967 από τον T. S. Motzkin και το 1969 από τον R. M. Robinson, σχεδόν ογδόντα χρόνια μετά την απόδειξη της ύπαρξης τέτοιων πολυωνύμων από τον Hilbert. Τα συγκεκριμένα αυτά παραδείγματα χρησιμοποιούσαν την «τεχνική της ανισότητας αριθμητικού–γεωμετρικού μέσου» την οποία θα περιγράψουμε στην τελευταία παράγραφο.

Το 1900, ο David Hilbert έδωσε μια ομιλία στο Παρίσι, στο Δεύτερο Διεθνές Συνέδριο των Μαθηματικών, όπου περιέγραψε 23 προβλήματα τα οποία πίστευε ότι άξιζαν της προσοχής των μαθηματικών όλου του κόσμου στην αυγή του 20ου αιώνα. Τα προβλήματα ήταν διαλεγμένα με σοφία, και άσκησαν μεγάλη επιρροή στην εξέλιξη των Μαθηματικών τον τελευταίο αιώνα.

Το 17ο πρόβλημα στη λίστα του Hilbert ήταν ένας άμεσος απόγονος της εικασίας του Minkowski, και σε αυτό το πρόβλημα ο Hilbert ρωτούσε αν κάθε μη αρνητικό πολυώνυμο n μεταβλητών αναπαρίσταται σαν το άθροισμα των τετραγώνων κάποιων *πηλίκων* πολυωνύμων. Αυτή η τροποποίηση του προβλήματος του Minkowski το καθιστά τελείως διαφορετικό: το ερώτημα του Hilbert απαντήθηκε καταφατικά το 1927, από τον Emil Artin. Η λύση του 17ου προβλήματος του Hilbert από τον Artin θεωρείται στις μέρες μας ένα από τα πιο όμορφα και σημαντικά αποτελέσματα της σύγχρονης Άλγεβρας.

6.2 Η περίπτωση της μιας μεταβλητής

Η πρώτη περίπτωση που δεν είναι εντελώς τετριμμένη εμφανίζεται όταν $n = 1$ και το πολυώνυμο $P(x)$ είναι της μορφής $ax^2 + bx + c$ με $a \neq 0$. Συμπληρώνοντας σε τέλειο τετράγωνο, όπως στη μελέτη του τριωνύμου, βλέπουμε ότι το $P(x)$ γράφεται στη μορφή

$$P(x) = ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a}, \tag{6.1}$$

και αυτή η αναπαράσταση σχεδόν απαντάει στο ερώτημά μας. Αρκεί να ελέγξουμε ότι οι τελευταίοι δύο προσθετέοι γράφονται σαν τετράγωνα πραγματικών πολυωνύμων.

Αν θεωρήσουμε μεγάλες τιμές του x, βλέπουμε ότι η $P(x) \geq 0$ έχει σαν συνέπεια την $a > 0$, και αν θέσουμε $x_0 = -b/2a$, τότε από το άθροισμα (6.1) βλέπουμε ότι η $P(x_0) \geq 0$ έχει σαν συνέπεια την $4ac - b^2 \geq 0$. Το συμπέρασμα είναι ότι οι δύο προσθετέοι στο δεξιό μέλος της ταυτότητας (6.1) είναι μη αρνητικοί, άρα το $P(x)$ γράφεται στη μορφή $Q_1^2 + Q_2^2$, όπου Q_1 και Q_2 είναι τα πραγματικά πολυώνυμα που ορίζονται από τις

$$Q_1(x) = \sqrt{a}\left(x + \frac{b}{2a}\right) \quad \text{και} \quad Q_2(x) = \frac{\sqrt{4ac - b^2}}{2\sqrt{a}}. \tag{6.2}$$

Έτσι, έχουμε λύσει το πρόβλημα για δευτεροβάθμια πολυώνυμα μιας μεταβλητής, η δε λύση, αν και απλή, δεν είναι τετριμμένη. Για παράδειγμα, δείχνει ότι το $P(x)$ ελαχιστοποιείται όταν $x = -b/2a$ και ότι η ελάχιστη τιμή του $P(x)$ ισούται με $(4ac - b^2)/4a$.

Χρησιμοποιώντας αυτά που γνωρίζουμε

Η απλούστερη μη τετριμμένη περίπτωση της ταυτότητας του Lagrange είναι η

$$(a_1^2 + a_2^2)(b_1^2 + b_2^2) = (a_1b_1 + a_2b_2)^2 + (a_1b_2 - a_2b_1)^2,$$

και αφού μπορούμε να αντικαταστήσουμε με πολυώνυμα τους αριθμούς σε αυτή τη σχέση, οδηγούμαστε σε ένα πολύ ισχυρό συμπέρασμα: το σύνολο των πολυωνύμων που γράφονται σαν άθροισμα τετραγώνων δύο πολυωνύμων είναι κλειστό ως προς τον πολλαπλασιασμό. Δηλαδή, αν $P(x) = Q(x)R(x)$ και τα $Q(x)$ και $R(x)$ έχουν τις αναπαραστάσεις

$$Q(x) = Q_1^2(x) + Q_2^2(x) \quad \text{και} \quad R(x) = R_1^2(x) + R_2^2(x),$$

τότε το $P(x)$ αναπαρίσταται κι αυτό σαν άθροισμα δύο τετραγώνων. Πιο συγκεκριμένα, αν έχουμε

$$P(x) = Q(x)R(x) = (Q_1^2(x) + Q_2^2(x))(R_1^2(x) + R_2^2(x)), \tag{6.3}$$

τότε το $P(x)$ γράφεται επίσης στη μορφή

$$\{Q_1(x)R_1(x) + Q_2(x)R_2(x)\}^2 + \{Q_1(x)R_2(x) - Q_2(x)R_1(x)\}^2. \tag{6.4}$$

Αυτή η ταυτότητα υποδεικνύει ότι η επαγωγή θα μπορούσε να φανεί χρήσιμη. Έχουμε ήδη δεί ότι κάθε δευτεροβάθμιο πολυώνυμο γράφεται σαν άθροισμα τετραγώνων, μπορούμε λοιπόν να ξεκινήσουμε μια επαγωγική απόδειξη. Από τη στιγμή που θα καταλάβουμε πώς να παραγοντοποιούμε μη αρνητικά πολυώνυμα, μπορούμε να χρησιμοποιήσουμε την αναπαράσταση (6.3) για να ολοκληρώσουμε το επαγωγικό βήμα.

Παραγοντοποίηση μη αρνητικών πολυωνύμων

Προσπαθώντας να παραγοντοποιήσουμε το $P(x)$, φυσιολογικά διακρίνουμε δύο περιπτώσεις: το $P(x)$ έχει πραγματική ρίζα ή όχι. Αν το $P(x)$ έχει μια πραγματική ρίζα r με πολλαπλότητα m, μπορούμε να γράψουμε

$$P(x) = (x - r)^m R(x), \quad \text{όπου } R(r) \neq 0,$$

επομένως, αν θέσουμε $x = r + \epsilon$, τότε έχουμε $P(r + \epsilon) = \epsilon^m R(r + \epsilon)$. Επίσης, από την συνέχεια της R, υπάρχει δ ώστε το $R(r + \epsilon)$ να έχει το ίδιο πρόσημο για κάθε ϵ με $|\epsilon| \leq \delta$. Αφού το $P(x)$ είναι πάντα μη αρνητικό, βλέπουμε ότι το ϵ^m έχει το ίδιο πρόσημο για κάθε $|\epsilon| \leq \delta$, άρα ο m πρέπει να είναι άρτιος. Αν θέσουμε $m = 2k$, βλέπουμε ότι

$$P(x) = Q^2(x)R(x) \quad \text{όπου } Q(x) = (x - r)^k, \tag{6.5}$$

και από αυτήν την αναπαράσταση, βλέπουμε ότι το $R(x)$ είναι επίσης μη αρνητικό πολυώνυμο. Έτσι, έχουμε βρεί μια χρήσιμη παραγοντοποίηση στην περίπτωση που το $P(x)$ έχει πραγματική ρίζα.

Ας υποθέσουμε τώρα ότι το $P(x)$ δεν έχει πραγματικές ρίζες. Από το θεμελιώδες θεώρημα της άλγεβρας, υπάρχει μια μιγαδική ρίζα r, και αφού

$$0 = P(r) \quad \Longrightarrow \quad 0 = \overline{P(r)} = P(\overline{r}), \tag{6.6}$$

βλέπουμε ότι ο μιγαδικός συζυγής $\overline{r}$ του r είναι επίσης ρίζα του P. Έτσι, το P έχει την παραγοντοποίηση

$$P(x) = (x - r)(x - \overline{r})R(x) = Q(x)R(x). \tag{6.7}$$

Το πραγματικό πολυώνυμο $Q(x) = (x - r)(x - \overline{r})$ είναι θετικό για μεγάλα x, και δεν έχει πραγματικές ρίζες, άρα πρέπει να είναι θετικό για κάθε πραγματικό x. Από την υπόθεση, το $P(x)$ είναι μη αρνητικό, συνεπώς το $R(x)$ είναι επίσης μη αρνητικό. Έτσι, βλέπουμε και πάλι ότι κάθε μη αρνητικό πολυώνυμο $P(x)$ με βαθμό μεγαλύτερο από δύο γράφεται σαν γινόμενο δύο μη σταθερών, μη αρνητικών πολυωνύμων.

Χρησιμοποιώντας τώρα επαγωγή που βασίζεται στην αναπαράσταση (6.3), βλέπουμε ότι κάθε μη αρνητικό πολυώνυμο μιας μεταβλητής είναι το άθροισμα των τετραγώνων δύο πραγματικών πολυωνύμων.

6.3 Περισσότερες μεταβλητές

Μετά την επιτυχία μας στην περίπτωση των πολυωνύμων μιας μεταβλητής, προσπαθούμε να εξετάσουμε μη αρνητικά πολυώνυμα δύο ή περισσοτέρων μεταβλητών. Δυστυχώς, η διαφορά δυσκολίας ανάμεσα σε ένα πρόβλημα μιας μεταβλητής και ένα πρόβλημα δύο μεταβλητών είναι συχνά πολύ μεγάλη. Για πολυώνυμα δύο μεταβλητών, το σύνολο $\{(x, y) : P(x, y) = 0\}$ των ριζών δεν είναι πιά ένα απλό διακριτό σύνολο σημείων. Τώρα, μπορούμε να έχουμε μια μεγάλη ποικιλία συνόλων ριζών, τα οποία δεν είναι εύκολο να ταξινομηθούν.

Υπάρχει μόνο μία περίπτωση στην οποία η απάντηση είναι απλή και καταφατική: όταν το πολυώνυμο έχει βαθμό 2. Η απόδειξη εδώ χρησιμοποιεί το γεγονός ότι κάθε συμμετρική τετραγωνική μορφή διαγωνοποιείται. Αν το P έχει βαθμό 2 και $P(x_1, x_2, \ldots, x_n) \geq 0$ για κάθε x_i, μπορούμε να βρούμε γραμμικό μετασχηματισμό των μεταβλητών που φέρνει το P στη μορφή

$$P(x_1, x_2, \ldots, x_n) = c + \sum_{i=1}^{n} \lambda_i x_i^2, \tag{6.8}$$

με $\lambda_i \geq 0$ για κάθε i, και $c \geq 0$.

Το πλεονέκτημα των περιορισμένων δυνατοτήτων

Μετά από κάποια σκέψη, τείνουμε να πιστέψουμε, όπως ο Minkowski, ότι υπάρχουν μη αρνητικά πολυώνυμα δύο μεταβλητών τα οποία *δεν μπορούν* να γραφτούν στη μορφή αθροίσματος τετραγώνων πραγματικών πολυωνύμων.

Για να κατασκευάσουμε ένα αντιπαράδειγμα πρέπει όμως πρώτα να μπορούμε – με συστηματικό τρόπο – να κατασκευάζουμε μη αρνητικά πολυώνυμα πολλών μεταβλητών. Τα πολυώνυμα που εκφράζονται σαν αθροίσματα τετραγώνων πραγματικών πολυωνύμων είναι πάντα μη αρνητικά, όμως αυτά τα πολυώνυμα δεν μπορούν να μας προσφέρουν κάτι σε σχέση με την εικασία του Minkowski. Θα μπορούσαμε επίσης να θεωρήσουμε τα μη αρνητικά πολυώνυμα που βρίσκουμε υψώνοντας στο τετράγωνο τα δύο μέλη της ανισότητας του Cauchy και παίρνοντας τη διαφορά τους, όμως η ταυτότητα του Lagrange μας λέει ότι κι αυτή η κατασκευή είναι καταδικασμένη να αποτύχει. Τέλος, μπορούμε να θεωρήσουμε εκείνα τα πολυώνυμα που είναι μη αρνητικά λόγω της ανισότητας αριθμητικού–γεωμετρικού μέσου. Αυτή είναι μια απλή (η μόνη ίσως) ιδέα που δεν έχουμε εξαντλήσει, αξίζει λοιπόν τον κόπο να την δοκιμάσουμε.

Χρησιμοποιώντας την ανισότητα αριθμητικού-γεωμετρικού μέσου

Είδαμε ότι για κάθε επιλογή μη αρνητικών αριθμών $a_1, a_2, \ldots, a_n$ ικανοποιείται η ανισότητα αριθμητικού–γεωμετρικού μέσου

$$(a_1 a_2 \cdots a_n)^{1/n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n},$$

και μπορούμε να χρησιμοποιήσουμε αυτήν την ανισότητα για να κατασκευάσουμε μια μεγάλη κλάση μη αρνητικών πολυωνύμων. Αν βέβαια δεν θέλουμε να εμπλακούμε σε ένα πλήθος πολύπλοκων παραδειγμάτων, θα πρέπει να περιορίσουμε την προσοχή μας στις απλούστερες περιπτώσεις. Η απλούστερη επιλογή μη αρνητικών a_1 και a_2 είναι να πάρουμε $a_1 = x^2$ και $a_2 = y^2$. Έτσι, αν θέλουμε να κάνουμε το γινόμενο $a_1 a_2 a_3$ όσο πιο απλό γίνεται, μπορούμε να πάρουμε $a_3 = 1/(x^2 y^2)$, οπότε το $a_1 a_2 a_3$ θα ισούται με 1. Τότε, η ανισότητα αριθμητικού-γεωμετρικού μέσου μας λέει ότι

$$1 \leq \frac{1}{3}(x^2 + y^2 + 1/(x^2 y^2))$$

και, μετά από τις φυσιολογικές απλοποιήσεις, βλέπουμε ότι το πολυώνυμο

$$P(x, y) = x^4 y^2 + x^2 y^4 - 3x^2 y^2 + 1$$

είναι μη αρνητικό για κάθε επιλογή των μεταβλητών x και y. Θα προσπαθήσουμε να αποδείξουμε ότι – γι' αυτό το πολυώνυμο – η αναπαράσταση

$$P(x, y) = Q_1^2(x, y) + Q_2^2(x, y) + \cdots + Q_s^2(x, y) \tag{6.9}$$

για κάποιον φυσικό s, είναι όντως αδύνατη. Τα εργαλεία που έχουμε στα χέρια μας είναι στοιχειώδη, είναι όμως αρκετά. Μια μικρή διερεύνηση δείχνει ότι η αναπαράσταση (6.9) είναι πολύ περιοριστική.

Για παράδειγμα, παρατηρήστε πρώτα ότι ο βαθμός του πολυωνύμου $P(x, y)$ είναι ίσος με 6, άρα, κανένα από τα πολυώνυμα Q_k δεν μπορεί να έχει βαθμό μεγαλύτερο από 3. Επιπλέον, αν θέσουμε $y = 0$, βλέπουμε ότι

$$1 = P(x, 0) = Q_1^2(x, 0) + Q_2^2(x, 0) + \cdots + Q_s^2(x, 0),$$

και αν θέσουμε $x = 0$, βλέπουμε ότι

$$1 = P(0, y) = Q_1^2(0, y) + Q_2^2(0, y) + \cdots + Q_s^2(0, y),$$

άρα, καθένα από τα πολυώνυμα $Q_k^2(x, 0)$ και $Q_k^2(0, y)$ πρέπει να είναι φραγμένο. Από αυτήν την παρατήρηση και από το γεγονός ότι κάθε πολυώνυμο $Q_k(x, y)$ έχει βαθμό το πολύ ίσο με 3, βλέπουμε ότι πρέπει να είναι της μορφής

$$Q_k(x, y) = a_k + b_k xy + c_k x^2 y + d_k xy^2 \tag{6.10}$$

για κάποιες σταθερές a_k, b_k, c_k και d_k.

Αν επιστρέψουμε στο πολυώνυμο $P(x, y)$, παρατηρούμε ότι έχει την εξής χαρακτηριστική ιδιότητα: όλοι οι συντελεστές του είναι μη αρνητικοί, εκτός από τον συντελεστή του x^2y^2 ο οποίος ισούται με -3. Με βάση αυτήν την παρατήρηση, προσπαθούμε να δούμε τι μπορούμε να πούμε για τον συντελεστή του x^2y^2 στο άθροισμα $Q_1^2(x, y) + Q_2^2(x, y) + \cdots + Q_s^2(x, y)$.

Σε αυτό το σημείο θα σταθούμε τυχεροί. Από τη μορφή (6.10) που έχουν οι όροι $Q_k(x, y)$, $1 \leq k \leq s$, μπορούμε να ελέγξουμε εύκολα ότι ο συντελεστής του x^2y^2 στο πολυώνυμο $Q_1^2(x, y) + Q_2^2(x, y) + \cdots + Q_s^2(x, y)$ είναι ίσος με $b_1^2 + b_2^2 + \cdots + b_s^2$. Αφού αυτό το άθροισμα είναι μη αρνητικό, δεν μπορεί να είναι ίσο με -3. Συνεπώς, το μη αρνητικό πολυώνυμο $P(x, y)$ δεν γράφεται σαν άθροισμα τετραγώνων πραγματικών πολυωνύμων. Μάλλον αναπάντεχα – κι αυτό δικαιολογεί την απόσταση των ογδόντα χρόνων από τον Hilbert ως τον Motzkin – η ανισότητα αριθμητικού–γεωμετρικού μέσου μας οδήγησε επιτυχώς στην απόδειξη της εικασίας του Minkowski.

Κεφάλαιο 7

Συνέπειες της διάταξης

7.1 Αντιστροφή της ανισότητας Cauchy

Ένα από τα φυσιολογικά ερωτήματα που συνοδεύουν κάθε ανισότητα είναι η δυνατότητα αντιστροφής της υπό προϋποθέσεις. Θα μελετήσουμε αυτό το ερώτημα για την ανισότητα του Cauchy. Η μελέτη αυτού του προβλήματος οδηγεί σε μία από τις πιό θεμελιώδεις αρχές στη θεωρία των ανισοτήτων – την συστηματική εκμετάλλευση ενδεχόμενων σχέσεων διάταξης.

Πρόβλημα. *Έστω $\rho \geq 1$. Κάτω από ποιές προϋποθέσεις οι μη αρνητικοί αριθμοί a_k, b_k, $k = 1, 2, \ldots, n$ ικανοποιούν μια ανισότητα της μορφής*

$$\left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \left(\sum_{k=1}^{n} b_k^2\right)^{1/2} \leq \rho \sum_{k=1}^{n} a_k b_k \,; \tag{7.1}$$

Μέρος της δυσκολίας εδώ είναι ότι το πρόβλημα δεν είναι πλήρως διατυπωμένο – οι προϋποθέσεις και οι συνθήκες κάτω από τις οποίες ισχύει η (7.1) πρέπει να προσδιοριστούν.

Σε αυτές τις περιπτώσεις, σχεδόν πάντα ξεκινάμε με κάποιους πειραματισμούς, και αφού η περίπτωση $n = 1$ είναι τετριμμένη, η απλούστερη περίπτωση που αξίζει τον κόπο να μελετήσουμε είναι να θεωρήσουμε τα διανύσματα $(1, a)$ και $(1, b)$ με $a > 0$ και $b > 0$. Τότε, τα δύο μέλη της (7.1) σχετίζονται με τις ποσότητες

$$(1 + a^2)^{1/2}(1 + b^2)^{1/2} \quad \text{και} \quad 1 + ab.$$

Παρατηρούμε ότι αν τα a και b επιλεγούν έτσι ώστε το γινόμενο ab να παραμένει σταθερό όταν το $a \to \infty$, τότε βλέπουμε ότι το δεξιό μέλος μένει φραγμένο, αλλά το αριστερό μέλος δεν είναι φραγμένο. Αυτή η παρατήρηση ουσιαστικά δείχνει ότι, για μια δεδομένη σταθερή τιμή του $\rho \geq 1$, η (7.1) δεν μπορεί να ισχύει εκτός αν επιβάλλουμε τη συνθήκη οι λόγοι a_k/b_k να είναι φραγμένοι από πάνω και από κάτω.

Το καλύτερο λοιπόν που μπορούμε να ελπίζουμε είναι ότι μια ανισότητα σαν την (7.1) θα ισχύει αν οι προσθετέοι ικανοποιούν κάποιον περιορισμό της μορφής

$$m \leq \frac{a_k}{b_k} \leq M \quad \text{για κάθε } k = 1, 2, \ldots, n, \tag{7.2}$$

για κάποιες σταθερές $0 < m < M < \infty$. Επίσης, η σταθερά ρ αναγκαστικά θα εξαρτάται από τις τιμές των m και M, και αυτό που ελπίζουμε να δείξουμε είναι ότι το ρ μπορεί να επιλεγεί ώστε να μην έχει καμία περαιτέρω εξάρτηση από τη συγκεκριμμένη επιλογή των a_k και b_k. Τώρα, το πρόβλημα είναι να βρούμε έναν τρόπο για να εκμεταλλευτούμε την υπόθεση (7.2).

Κοιτάζοντας την *ζητούμενη* ανισότητα (7.1) και την *δεδομένη* ανισότητα (7.2), θα μπορούσε κανείς να σκεφτεί ότι οι γνωστές αποδείξεις της ανισότητας του Cauchy μπορούν να βοηθήσουν. Πιο συγκεκριμένα, αν θυμηθούμε την απόδειξη που ξεκινούσε από την $(a-b)^2 \geq 0$, αρχίζουμε να υποψιαζόμαστε ότι μια παρόμοια ιδέα μπορεί να χρησιμοποιηθεί εδώ. Ψάχνουμε λοιπόν για κάποια «τετραγωνική ανισότητα» η οποία να προκύπτει από την (7.2).

Αν σκεφτούμε λίγο, παρατηρούμε ότι η διπλή ανισότητα (7.2) δίνει εύκολα την τετραγωνική ανισότητα

$$\left(M - \frac{a_k}{b_k}\right)\left(\frac{a_k}{b_k} - m\right) \geq 0. \tag{7.3}$$

Εκτελώντας τους απαραίτητους πολλαπλασιασμούς, από την (7.3) καταλήγουμε στην ισοδύναμη ανισότητα

$$a_k^2 + (mM)b_k^2 \leq (m+M)a_k b_k \quad \text{για κάθε } k = 1, 2, \ldots, n. \tag{7.4}$$

Αθροίζοντας τις ανισότητες (7.4) πάνω από όλα τα $1 \leq k \leq n$, καταλήγουμε στην προσθετική ανισότητα

$$\sum_{k=1}^{n} a_k^2 + (mM)\sum_{k=1}^{n} b_k^2 \leq (m+M)\sum_{k=1}^{n} a_k b_k. \tag{7.5}$$

Αυτό που μένει είναι να την μετατρέψουμε σε μια πολλαπλασιαστική ανισότητα.

Αν θυμηθούμε τι έχουμε κάνει σε ανάλογες περιπτώσεις, θα πρέπει τώρα να ορίσουμε κανονικοποιημένες μεταβλητές $\tilde{a}_k$ και $\tilde{b}_k$, αυτήν όμως τη φορά, η κανονικοποίηση παρουσιάζει δυσκολίες. Το πρόβλημα είναι ότι η ανισότητα (7.5) εφαρμόζεται για τους $\tilde{a}_k$ και $\tilde{b}_k$ μόνο αν ικανοποιούν τις $m \leq \tilde{a}_k/\tilde{b}_k \leq M$. Αυτοί οι περιορισμοί δεν ικανοποιούνται από τους προφανείς υποψηφίους $\tilde{a}_k$ και $\tilde{b}_k$. Χρειαζόμαστε λοιπόν μια καινούργια ιδέα για να περάσουμε σε γινόμενο.

Στο σημείο αυτό, μπορεί κάποιος να κολλήσει. Μετά από λίγη σκέψη όμως, βλέπουμε ότι η ανισότητα αριθμητικού–γεωμετρικού μέσου μπορεί να μας οδηγήσει στο γινόμενο που ζητάμε: εφαρμόζοντάς την στην (7.5) βλέπουμε ότι

$$\begin{aligned}\left(\sum_{k=1}^{n} a_k^2\right)^{1/2}\left(mM\sum_{k=1}^{n} b_k^2\right)^{1/2} &\leq \frac{1}{2}\left\{\sum_{k=1}^{n} a_k^2 + (mM)\sum_{k=1}^{n} b_k^2\right\} \\ &\leq \frac{1}{2}\left\{(m+M)\sum_{k=1}^{n} a_k b_k\right\}.\end{aligned}$$

Τώρα, φέρνοντας όλες τις σταθερές στο δεξιό μέλος, ερχόμαστε σε μια ανισότητα που ικανοποιεί τις απαιτήσεις μας. Αν θέσουμε

$$A = \frac{m+M}{2} \quad \text{και} \quad G = \sqrt{mM}, \tag{7.6}$$

τότε, για κάθε επιλογή μη αρνητικών a_k, b_k, $k = 1, 2, \ldots, n$ με

$$0 < m \leq a_k/b_k \leq M < \infty,$$

έχουμε αποδείξει την ανισότητα

$$\left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \left(\sum_{k=1}^{n} b_k^2\right)^{1/2} \leq \frac{A}{G} \sum_{k=1}^{n} a_k b_k. \tag{7.7}$$

Αυτή είναι μια ικανοποιητική αντιστροφή της ανισότητας Cauchy.

7.2 Μονοτονία και μια «ανισότητα διάταξης» του Chebyshev

Ένας τρόπος για να εκμεταλλευτούμε τις σχέσεις διάταξης είναι να περιορίσουμε την προσοχή μας σε μονότονες ακολουθίες και μονότονες συναρτήσεις. Αυτή η επιλογή οδηγεί σε ένα σημαντικό αποτέλεσμα που έχει πολλές εφαρμογές, ιδιαίτερα στη θεωρία πιθανοτήτων και την στατιστική.

Το αποτέλεσμα που θα περιγράψουμε αποδείχθηκε από τον Pafnuty Lvovich Chebyshev (1821-1894), ο οποίος κατά πάσα πιθανότητα εισήχθη στη θεωρία πιθανοτήτων από τον Victor Yacovlevich Bunyakovsky. Η θεωρία πιθανοτήτων ήταν ένας από τους μοντέρνους κλάδους των μαθηματικών που ο Bunyakovsky έφερε μαζί του στην St. Petersburg επιστρέφοντας από το Παρίσι και τη μαθητεία του κοντά στον Cauchy.

Θεώρημα 7.2.1 (η ανισότητα διάταξης του Chebyshev). *Έστω $f : \mathbb{R} \to \mathbb{R}$ και $g : \mathbb{R} \to \mathbb{R}$ δύο αύξουσες συναρτήσεις και έστω $p_j \geq 0$, $j = 1, 2, \ldots, n$, οι οποίοι ικανοποιούν την $p_1 + p_2 + \cdots + p_n = 1$. Για κάθε αύξουσα ακολουθία $x_1 \leq x_2 \leq \cdots \leq x_n$ ισχύει η ανισότητα*

$$\left\{\sum_{k=1}^{n} f(x_k) p_k\right\} \left\{\sum_{k=1}^{n} g(x_k) p_k\right\} \leq \sum_{k=1}^{n} f(x_k) g(x_k) p_k. \tag{7.8}$$

Μπορούμε εύκολα να κατανοήσουμε την ανισότητα (7.8) χωρίς να βασιστούμε στη διασύνδεσή της με τη θεωρία πιθανοτήτων. Οι εφαρμογές της σε άλλες περιοχές των Μαθηματικών είναι ούτως ή άλλως πολλές. Η πιθανοθεωρητική όμως ερμηνεία της ανισότητας (7.8) είναι εξαιρετικά ενδιαφέρουσα. Στη γλώσσα των πιθανοτήτων, λέει ότι αν X είναι μια τυχαία μεταβλητή με την ιδιότητα $P(X = x_k) = p_k$ για $k = 1, 2, \ldots, n$, τότε

$$\mathbb{E}\left[f(X)\right] \mathbb{E}\left[g(X)\right] \leq \mathbb{E}\left[f(X) g(X)\right], \tag{7.9}$$

όπου, ως συνήθως, με P συμβολίζουμε την πιθανότητα και με $\mathbb{E}$ τη μέση τιμή. Με άλλα λόγια, αν οι τυχαίες μεταβλητές Y και Z μπορούν να γραφτούν σαν αύξουσες συναρτήσεις της ίδιας τυχαίας μεταβλητής X, τότε οι Y και Z έχουν μη αρνητική συνδιακύμανση. Χωρίς την ανισότητα του Chebyshev, η διαίσθηση που έχουμε για την στατιστική έννοια της συνδιακύμανσης θα ήταν ασθενής.

Σημείωση. Υπάρχει μια άλλη ανισότητα του Chebyshev, η οποία είναι ακόμα πιο σημαντική για τη θεωρία πιθανοτήτων. Μας λέει ότι για κάθε τυχαία μεταβλητή X με πεπερασμένη μέση τιμή $\mu = \mathbb{E}(X)$, η ανισότητα

$$P(|X-\mu| \geq \lambda) \leq \frac{1}{\lambda^2}\mathbb{E}\left(|X-\mu|^2\right) \tag{7.10}$$

ισχύει για κάθε $\lambda > 0$. Η απόδειξη αυτής της ανισότητας είναι πολύ απλή, είναι όμως τόσο σημαντική για τη θεωρία πιθανοτήτων που η *ανισότητα διάταξης του Chebyshev* (7.8) συχνά αποκαλείται *η άλλη ανισότητα* του Chebyshev.

Η ανισότητα του Chebyshev (7.8) είναι τετραγωνική, και οι υποθέσεις μας φέρνουν στο νού την τεχνική της *μετατροπής της διάταξης σε τετραγωνική ανισότητα*. Εδώ, η μονοτονία των f και g έχει σαν συνέπεια την τετραγωνική ανισότητα

$$0 \leq \{f(x_k) - f(x_j)\}\{g(x_k) - g(x_j)\},$$

και αναπτύσσοντας την τελευταία, παίρνουμε

$$f(x_k)g(x_j) + f(x_j)g(x_k) \leq f(x_j)g(x_j) + f(x_k)g(x_k). \tag{7.11}$$

Αν έχουμε φτάσει σε αυτό το σημείο, αρκεί να εισάγουμε τα p_j και να κάνουμε τις απαραίτητες αριθμητικές πράξεις.

Έτσι, πολλαπλασιάζοντας την ανισότητα (7.11) με $p_j p_k$ και αθροίζοντας πάνω από όλα τα $1 \leq j \leq n$ και $1 \leq k \leq n$, βλέπουμε ότι το αριστερό μέλος μας δίνει

$$\sum_{j,k=1}^{n} \{f(x_k)g(x_j) + f(x_j)g(x_k)\}p_j p_k = 2\left\{\sum_{k=1}^{n} f(x_k)p_k\right\}\left\{\sum_{k=1}^{n} g(x_k)p_k\right\},$$

ενώ το δεξιό μέλος μας δίνει

$$\sum_{j,k=1}^{n} \{f(x_j)g(x_j) + f(x_k)g(x_k)\}p_j p_k = 2\left\{\sum_{k=1}^{n} f(x_k)g(x_k)p_k\right\}.$$

Έτσι, η ανισότητα (7.11) μας δίνει μια απόδειξη της ανισότητας του Chebyshev.

7.3 Μια ανισότητα αναδιάταξης

Η απόδειξη της ανισότητας του Chebyshev μας οδηγεί σε μια πολύ σημαντική *ανισότητα αναδιάταξης*. Στη συνέχεια, θα γράφουμε $[n]$ για το σύνολο $\{1, 2, \ldots, n\}$. Θυμηθείτε ότι με τον όρο *μετάθεση* του $[n]$ εννοούμε απλώς μια ένα-προς-ένα και επί απεικόνιση από το $[n]$ στο $[n]$.

Θεώρημα 7.3.1 (ανισότητα αναδιάταξης). *Για κάθε ζευγάρι διατεταγμένων πραγματικών ακολουθιών*

$$-\infty < a_1 \leq a_2 \leq \cdots \leq a_n < \infty \quad \textit{και} \quad -\infty < b_1 \leq b_2 \leq \cdots \leq b_n < \infty,$$

και για κάθε μετάθεση $\sigma : [n] \to [n]$, ισχύει η ανισότητα

$$\sum_{k=1}^{n} a_k b_{n-k+1} \leq \sum_{k=1}^{n} a_k b_{\sigma(k)} \leq \sum_{k=1}^{n} a_k b_k. \tag{7.12}$$

Οι υποθέσεις του προβλήματος δίνουν κάποιες σχέσεις διάταξης, και το συμπέρασμα είναι μια τετραγωνική ανισότητα. Αμέσως μας έρχεται στο νού να χρησιμοποιήσουμε τη μέθοδο των προηγούμενων παραγράφων, για να πάρουμε όμως μια ιδέα για την φύση του προβλήματος, ξεκινάμε από την απλούστερη περίπτωση $n = 2$.

Σε αυτήν την περίπτωση, θυμόμαστε πρώτα ότι από τις $a_1 \leq a_2$ και $b_1 \leq b_2$ έπεται η

$$0 \leq (a_2 - a_1)(b_2 - b_1),$$

και, εκτελώντας τις πράξεις, συμπεραίνουμε ότι

$$a_1 b_2 + a_2 b_1 \leq a_1 b_1 + a_2 b_2.$$

Αυτή είβαι ακριβώς η μορφή που παίρνει η (7.12) στην περίπτωση $n = 2$. Το πρόβλημα τώρα είναι να δούμε αν η ίδια ιδέα μπορεί να μας βοηθήσει να χειριστούμε το πιο γενικό άθροισμα

$$S(\sigma) = \sum_{k=1}^{n} a_k b_{\sigma(k)}.$$

Αντιστροφές και η αναίρεσή τους

Αν η σ δεν είναι η ταυτοτική μετάθεση, τότε πρέπει να υπάρχει κάποιο ζευγάρι δεικτών $j < k$ ώστε $\sigma(k) < \sigma(j)$. Ένα τέτοιο ζευγάρι ονομάζεται *αντιστροφή*, και αυτό που παρατηρούμε από τη μελέτη της περίπτωσης $n = 2$ είναι ότι αν εναλλάξουμε τις τιμές των $\sigma(k)$ και $\sigma(j)$, τότε η τιμή του αθροίσματος που μελετάμε θα αυξηθεί – ή, τουλάχιστον, δεν θα μειωθεί. Για να μετατρέψουμε αυτήν την ιδέα σε αυστηρό επιχείρημα, εισάγουμε πρώτα μια νέα μετάθεση τ, θέτοντας

(i) $\tau(i) = \sigma(i)$ αν $i \neq j$ και $i \neq k$,

(ii) $\tau(i) = \sigma(j)$ αν $i = k$,

(iii) $\tau(i) = \sigma(k)$ αν $i = j$.

Από τον ορισμό της τ, κάνοντας κατάλληλη παραγοντοποίηση, βλέπουμε ότι

$$\begin{aligned} S(\tau) - S(\sigma) &= a_j b_{\tau(j)} + a_k b_{\tau(k)} - a_j b_{\sigma(j)} - a_k b_{\sigma(k)} \\ &= a_j b_{\tau(j)} + a_k b_{\tau(k)} - a_j b_{\tau(k)} - a_k b_{\tau(j)} \\ &= (a_k - a_j)(b_{\tau(k)} - b_{\tau(j)}) \geq 0. \end{aligned}$$

Έτσι, με τον μετασχηματισμό $\sigma \mapsto \tau$ πετυχαίνουμε δύο στόχους. Πρώτον, μεγαλώνουμε την ποσότητα S: έχουμε $S(\sigma) \leq S(\tau)$. Δεύτερον, το πλήθος των αντιστροφών στην τ είναι γνησίως μικρότερο από το πλήθος των αντιστροφών στην σ.

Επανάληψη της διαδικασίας

Μια μετάθεση του $[n]$ εμφανίζει το πολύ $n(n-1)/2$ αντιστροφές και η μόνη μετάθεση που δεν έχει αντιστροφές είναι η ταυτοτική. Υπάρχει λοιπόν μια πεπερασμένη

ακολουθία μετασχηματισμών του είδους που περιγράψαμε πιο πάνω, η οποία μας οδηγεί από την σ στην ταυτοτική μετάθεση. Αν συμβολίσουμε αυτές τις διαδοχικές μεταθέσεις με $\sigma = \sigma_0, \sigma_1, \ldots, \sigma_m$, όπου σ_m είναι η ταυτοτική μετάθεση και $m \leq n(n-1)/2$, τότε, εφαρμόζοντας την ανισότητα $S(\sigma_{j-1}) \leq S(\sigma_j)$ για $j = 1, 2, \ldots, m$, συμπεραίνουμε ότι

$$S(\sigma) \leq \sum_{k=1}^{n} a_k b_k.$$

Έτσι, ολοκληρώνεται η απόδειξη του άνω φράγματος στην ανισότητα αναδιάταξης (7.12).

Ο εύκολος τρόπος για να πάρουμε το κάτω φράγμα είναι να παρατηρήσουμε ότι είναι άμεση συνέπεια του άνω φράγματος. Αν ορίσουμε $b'_1 = -b_n$, $b'_2 = -b_{n-1}, \ldots, b'_n = -b_1$, βλέπουμε ότι

$$b'_1 \leq b'_2 \leq \cdots \leq b'_n,$$

και, εφαρμόζοντας το άνω φράγμα της ανισότητας αναδιάταξης (7.12) για την ακολουθία $b'_1, b'_2, \ldots, b'_n$, παίρνουμε το κάτω φράγμα της ανισότητας (7.12) για την ακολουθία $b_1, b_2, \ldots, b_n$.

Κεφάλαιο 8

Συμμετρικά αθροίσματα

8.1 Οι κλασικές ανισότητες των Newton και Maclaurin

Η k-οστή στοιχειώδης συμμετρική συνάρτηση των n μεταβλητών $x_1, x_2, \ldots, x_n$ είναι το πολυώνυμο

$$e_k(x_1, x_2, \ldots, x_n) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

Τα πρώτα τρία συμμετρικά πολυώνυμα είναι τα

$$e_0(x_1, x_2, \ldots, x_n) = 1, \quad e_1(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n$$

και

$$e_2(x_1, x_2, \ldots, x_n) = \sum_{1 \le j < k \le n} x_j x_k,$$

ενώ η n-οστή στοιχειώδης συμμετρική συνάρτηση είναι το πλήρες γινόμενο

$$e_n(x_1, x_2, \ldots, x_n) = x_1 x_2 \cdots x_n.$$

Οι συναρτήσεις αυτές εμφανίζονται σε όλους κυριολεκτικά τους κλάδους των Μαθηματικών. Μια πολύ σημαντική ιδιότητά τους είναι ότι συνδέουν τους *συντελεστές* ενός πολυωνύμου με κάποιες συναρτήσεις των *ριζών* του: αν το πολυώνυμο $P(t)$ γράφεται στη μορφή $P(t) = (t - x_1)(t - x_2) \cdots (t - x_n)$, τότε έχει και την αναπαράσταση

$$P(t) = t^n - e_1(\mathbf{x}) t^{n-1} + \cdots + (-1)^k e_k(\mathbf{x}) + \cdots + (-1)^n e_n(\mathbf{x}), \tag{8.1}$$

όπου, για συντομία, γράφουμε $e_k(\mathbf{x})$ αντί για $e_k(x_1, x_2, \ldots, x_n)$.

Οι στοιχειώδεις συμμετρικές συναρτήσεις παίζουν σημαντικό ρόλο στην θεωρία των ανισοτήτων. Δύο από τις πιό διάσημες ανισότητες γι' αυτές μας οδηγούν πίσω στον Isaak Newton (1542-1727) και τον Colin Maclaurin (1696-1746). Μπορούμε να τις διατυπώσουμε πιό κομψά μέσω των κανονικοποιημένων συναρτήσεων

$$E_k(\mathbf{x}) = E_k(x_1, x_2, \ldots, x_n) = \frac{e_k(x_1, x_2, \ldots, x_n)}{\binom{n}{k}},$$

ως εξής:

Οι ανισότητες των Newton και Maclaurin: Για κάθε $\mathbf{x} \in \mathbb{R}^n$ ισχύουν οι ανισότητες του Newton

$$E_{k-1}(\mathbf{x}) \cdot E_{k+1}(\mathbf{x}) \leq E_k^2(\mathbf{x}) \quad \text{για } 0 < k < n, \tag{8.2}$$

και έπονται οι ανισότητες του Maclaurin:

$$E_n^{1/n}(\mathbf{x}) \leq E_{n-1}^{1/(n-1)}(\mathbf{x}) \leq \cdots \leq E_2^{1/2}(\mathbf{x}) \leq E_1(\mathbf{x}) \tag{8.3}$$

για κάθε $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ με $x_k \geq 0$ για κάθε $1 \leq k \leq n$.

Αν πάρουμε $n = 3$ και γράψουμε $\mathbf{x} = (x, y, z)$, τότε οι ανισότητες του Maclaurin μας λένε ότι, αν $x, y, z \geq 0$ τότε

$$(xyz)^{1/3} \leq \left(\frac{xy + xz + yz}{3}\right)^{1/2} \leq \frac{x + y + z}{3},$$

δηλαδή μας δίνουν μια εκλέπτυνση της ανισότητας αριθμητικού–γεωμετρικού μέσου. Στην γενική περίπτωση, οι ανισότητες του Maclaurin δίνουν μια ολόκληρη αύξουσα ακολουθία παραστάσεων ανάμεσα στον γεωμετρικό μέσο $(x_1 x_2 \cdots x_n)^{1/n}$ και τον αριθμητικό μέσο $(x_1 + x_2 + \cdots + x_n)/n$.

Από τον Newton στον Maclaurin μέσω γεωμετρίας

Για ένα διάνυσμα $\mathbf{x} \in \mathbb{R}^n$ με θετικές συντεταγμένες, οι τιμές $\{E_k(\mathbf{x}) : 0 \leq k \leq n\}$ είναι επίσης θετικές, και μπορούμε να πάρουμε λογαρίθμους στις ανισότητες του Newton: δηλαδή,

$$\frac{\log E_{k-1}(\mathbf{x}) + \log E_{k+1}(\mathbf{x})}{2} \leq \log E_k(\mathbf{x}) \tag{8.4}$$

για κάθε $1 \leq k \leq n$. Ειδικότερα, αν $\mathbf{x} \in [0, \infty)^n$ οι ανισότητες του Newton μας λένε ότι η κατά τμήματα γραμμική συνάρτηση που προσδιορίζεται από το σύνολο σημείων $\{(k, \log E_k(\mathbf{x})) : 0 \leq k \leq n\}$ είναι κοίλη.

Αν ονομάσουμε L_k την ευθεία που προσδιορίζεται από τα σημεία $(0, 0) = (0, \log E_0(\mathbf{x}))$ και $(k, \log E_k(\mathbf{x}))$, τότε η κλίση της ευθείας L_{k+1} είναι μικρότερη ή ίση από την κλίση της ευθείας L_k για κάθε $k = 1, 2, \ldots, n-1$. Αφού η κλίση της L_k είναι ίση με $\log E_k(\mathbf{x})/k$, συμπεραίνουμε ότι

$$\frac{\log E_k(\mathbf{x})}{k} \leq \frac{\log E_{k+1}(\mathbf{x})}{k+1},$$

και αυτή η ανισότητα είναι ακριβώς η k-οστή από τις ανισότητες του Maclaurin.

Το πραγματικό πρόβλημα είναι να αποδείξουμε τις ανισότητες του Newton. Όπως θα περίμενε κανείς για ένα θεμελιώδες και παλαιό αποτέλεσμα, υπάρχουν πολλές διαφορετικές προσεγγίσεις. Οι περισσότερες από αυτές κάνουν κάποια στιγμή χρήση Απειροστικού Λογισμού, όμως ο Newton δεν δημοσίευσε ποτέ κάποια απόδειξη των ανισοτήτων του, οπότε δεν γνωρίζουμε ποιά ακριβώς ήταν η μέθοδος που χρησιμοποίησε.

Πολυώνυμα και οι παράγωγοί τους

Ακόμα κι αν ο Newton ακολούθησε διαφορετικό δρόμο, έχει νόημα να εξετάσουμε αν η παράγωγος $P'(t)$ μπορεί να μας δώσει κάποιες πληροφορίες για τα πολυώνυμα $E_k(x_1, x_2, \ldots, x_n)$, $1 \le k \le n$. Αν γράψουμε την ταυτότητα (8.1) στη μορφή

$$P(t) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} E_k(x_1, x_2, \ldots, x_n) t^{n-k}, \tag{8.5}$$

τότε η παράγωγος $P'(t)$ παίρνει μια πολύ ενδιαφέρουσα μορφή. Συγκεκριμένα, έχουμε

$$\begin{aligned} Q(t) = \frac{1}{n} P'(t) &= \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} \frac{n-k}{n} E_k(x_1, x_2, \ldots, x_n) t^{n-k-1} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} E_k(x_1, x_2, \ldots, x_n) t^{n-k-1}, \end{aligned}$$

όπου, στην δεύτερη γραμμή, χρησιμοποιήσαμε την γνωστή ταυτότητα

$$\binom{n}{k} \frac{n-k}{n} = \frac{n!}{k!(n-k)!} \frac{n-k}{n} = \frac{(n-1)!}{k!(n-k-1)!} = \binom{n-1}{k}.$$

Αν οι x_k, $k = 1, 2, \ldots, n$ ανήκουν στο διάστημα $[a, b]$, τότε το πολυώνυμο $P(t)$ έχει n πραγματικές ρίζες στο $[a, b]$, και το θεώρημα του Rolle μας λέει ότι η παράγωγος $P'(t)$ πρέπει να έχει $n-1$ πραγματικές ρίζες στο $[a, b]$. Αν συμβολίσουμε αυτές τις ρίζες με $y_1, y_2, \ldots, y_{n-1}$, τότε προκύπτει η ταυτότητα

$$\begin{aligned} Q(t) = \frac{1}{n} P'(t) &= (t - y_1)(t - y_2) \cdots (t - y_{n-1}) \\ &= \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} E_k(y_1, y_2, \ldots, y_{n-1}) t^{n-k-1}. \end{aligned}$$

Αν τώρα εξισώσουμε τους συντελεστές των δύο πολυωνυμικών εκφράσεων που βρήκαμε για το $Q(t)$, καταλήγουμε στην αξιοσημείωτη ταυτότητα

$$E_k(x_1, x_2, \ldots, x_n) = E_k(y_1, y_2, \ldots, y_{n-1}) \qquad (0 \le k \le n-1). \tag{8.6}$$

Το αριστερό μέλος της ταυτότητας (8.6) είναι μια συνάρτηση του n–διάστατου διανύσματος $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ ενώ το δεξιό μέλος είναι μια συνάρτηση του $(n-1)$–διάστατου διανύσματος $\mathbf{y} = (y_1, y_2, \ldots, y_{n-1})$. Αν λοιπόν αποδείξουμε μια ανισότητα της μορφής

$$0 \le F(E_0(\mathbf{y}), E_1(\mathbf{y}), \ldots, E_{n-1}(\mathbf{y})) \quad \text{για κάθε } \mathbf{y} \in [a, b]^{n-1},$$

τότε έχουμε και την ανισότητα

$$0 \le F(E_0(\mathbf{x}), E_1(\mathbf{x}), \ldots, E_{n-1}(\mathbf{x})) \quad \text{για κάθε } \mathbf{x} \in [a, b]^{n}.$$

Με άλλα λόγια, κάθε ανισότητα – ή ταυτότητα – που δίνει μια σχέση ανάμεσα στις $(n-1)$ ποσότητες $E_0(\mathbf{y}), E_1(\mathbf{y}), \ldots, E_{n-1}(\mathbf{y})$ και ισχύει για κάθε $\mathbf{y} \in$

$[a,b]^{n-1}$, εξασφαλίζει αυτόματα ότι η αντίστοιχη σχέση για τις $(n-1)$ ποσότητες $E_0(\mathbf{x}), E_1(\mathbf{x}), \ldots, E_{n-1}(\mathbf{x})$ ισχύει για κάθε $\mathbf{x} \in [a,b]^n$.

Αντιμετωπίζουμε λοιπόν ένα πρόβλημα της εξής μορφής: θέλουμε να αποδείξουμε μια σχέση για συναρτήσεις n μεταβλητών και αρκεί να αποδείξουμε μια ανάλογη σχέση για συναρτήσεις $n-1$ (μόνον!) μεταβλητών. Αυτή η παρατήρηση μπορεί να χρησιμοποιηθεί για την «παραγωγή» πολυάριθμων ειδικών ταυτοτήτων οι οποίες φαίνονται αρκετά εντυπωσιακές. Μπορεί επίσης να χρησιμοποιηθεί για την «οργάνωση» επαγωγικών (ως προς το πλήθος των μεταβλητών) αποδείξεων για αποτελέσματα όπως οι ανισότητες του Newton.

Επαγωγή ως προς το πλήθος των μεταβλητών

Ονομάζουμε H_n την πρόταση η οποία ισχυρίζεται ότι

$$E_{j-1}(x_1, x_2, \ldots, x_n)E_{j+1}(x_1, x_2, \ldots, x_n) \leq E_j^2(x_1, x_2, \ldots, x_n) \tag{8.7}$$

για κάθε $\mathbf{x} \in \mathbb{R}^n$ και για κάθε $1 \leq j < n$. Για $n = 1$ η πρόταση δεν έχει νόημα, οπότε το επαγωγικό μας επιχείρημα ξεκινάει με την H_2. Σε αυτήν την περίπτωση, αρκεί να ελέγξουμε μία μόνο ανισότητα, την

$$E_0(x_1, x_2)E_2(x_1, x_2) \leq E_1^2(x_1, x_2), \tag{8.8}$$

ή, ισοδύναμα, την γνωστή ανισότητα $x_1x_2 \leq \left(\frac{x_1+x_2}{2}\right)^2$.

Λογικά, θα μπορούσαμε τώρα να επιχειρήσουμε το επαγωγικό βήμα, προκειμένου όμως να καταλάβουμε καλύτερα τη λογική των πραγμάτων, θεωρούμε πρώτα την H_3, η οποία περιγράφεται από δύο ισχυρισμούς:

$$E_0(x_1, x_2, x_3)E_2(x_1, x_2, x_3) \leq E_1^2(x_1, x_2, x_3), \tag{8.9}$$

$$E_1(x_1, x_2, x_3)E_3(x_1, x_2, x_3) \leq E_2^2(x_1, x_2, x_3). \tag{8.10}$$

Σε αυτό το σημείο αρχίζουμε να χρησιμοποιούμε πραγματικά την «αξιοσημείωτη ταυτότητα» (8.6). Ο ισχυρισμός (8.9) μας λέει για τρείς μεταβλητές αυτό που η ανισότητα (8.8) μας λέει για δύο, άρα η (8.6) μας λέει ότι η πρώτη ανισότητα (8.9) ισχύει. Έχουμε ολοκληρώσει τη μισή απόδειξη χωρίς να κάνουμε τίποτα.

Για να συμπληρώσουμε την απόδειξη της H_3 πρέπει τώρα να αποδείξουμε την δεύτερη ανισότητα (8.10). Για να δούμε ποιός ακριβώς είναι ο στόχος μας, ξαναγράφουμε την (8.10) στην πλήρη μορφή της:

$$\left\{\frac{x_1 + x_2 + x_3}{3}\right\}(x_1x_2x_3) \leq \left\{\frac{x_1x_2 + x_1x_3 + x_2x_3}{3}\right\}^2. \tag{8.11}$$

Αυτή η ανισότητα ισχύει τετριμμένα αν $x_1x_2x_3 = 0$, υποθέτουμε λοιπόν ότι $x_1x_2x_3 \neq 0$. Μπορούμε τότε να διαιρέσουμε τα δύο μέλη με $(x_1x_2x_3)^2$ και να αναχθούμε στην

$$\frac{1}{3}\left\{\frac{1}{x_1x_2} + \frac{1}{x_1x_3} + \frac{1}{x_2x_3}\right\} \leq \frac{1}{9}\left\{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}\right\}^2,$$

η οποία, αν αναπτύξουμε το τετράγωνο και κάνουμε απλοποιήσεις, γράφεται ισοδύναμα στη μορφή

$$\frac{1}{x_1x_2}+\frac{1}{x_1x_3}+\frac{1}{x_2x_3}\le\frac{1}{x_1^2}+\frac{1}{x_2^2}+\frac{1}{x_3^2}.$$

Με όσα έχουμε δεί ως τώρα, είναι πολύ εύκολο να αποδείξουμε αυτήν την ανισότητα. Μπορούμε να εφαρμόσουμε την ανισότητα του Cauchy στα διανύσματα $(1/x_1, 1/x_3, 1/x_2)$ και $(1/x_2, 1/x_1, 1/x_3)$, ή να αθροίσουμε τις τρείς ανισότητες

$$\frac{1}{x_jx_k}\le\frac{1}{2}\left\{\frac{1}{x_j^2}+\frac{1}{x_k^2}\right\},\qquad 1\le j<k\le 3.$$

Έτσι, έχουμε αποδείξει την H_3 και έχουμε πάρει μια ιδέα για τον τρόπο με τον οποίο θα δουλέψουμε στο επαγωγικό βήμα.

Το επαγωγικό βήμα

Η H_n αποτελείται από $n-1$ ανισότητες οι οποίες χωρίζονται σε δύο ομάδες. Πρώτον, για το $\mathbf{x}=(x_1,x_2,\dots,x_n)$ έχουμε τις $n-2$ ανισότητες στις οποίες εμφανίζονται οι ποσότητες $E_j(\mathbf{x})$, $0\le j<n$,

$$E_{k-1}(\mathbf{x})E_{k+1}(\mathbf{x})\le E_k^2(\mathbf{x})\quad \text{για } 1\le k<n-1, \tag{8.12}$$

και, δεύτερον, έχουμε μία μόνο ανισότητα στην οποία εμφανίζεται η ποσότητα $E_n(\mathbf{x})$, την

$$E_{n-2}(\mathbf{x})E_n(\mathbf{x})\le E_{n-1}^2(\mathbf{x}). \tag{8.13}$$

Όπως είδαμε και στη συζήτηση για την H_3, όλες οι ανισότητες της πρώτης ομάδας (8.12) προκύπτουν από την επαγωγική υπόθεση H_{n-1} και από την ταυτότητα (8.6).

Αν τώρα αναπτύξουμε την (8.13) και αν με το σύμβολο $\hat{x}_j$ εννοούμε ότι ο x_j παραλείπεται, τότε βλέπουμε ότι η ανισότητα που μένει να αποδειχτεί είναι η

$$\begin{aligned}\frac{2}{n(n-1)}&\left\{\sum_{1\le j<k\le n}x_1\cdots\hat{x}_j\cdots\hat{x}_k\cdots x_n\right\}x_1x_2\cdots x_n\\ &\le\left\{\frac{1}{n}\sum_{j=1}^{n}x_1x_2\cdots\hat{x}_j\cdots x_n\right\}^2.\end{aligned} \tag{8.14}$$

Όπως και στην προηγούμενη παράγραφο, μπορούμε να υποθέσουμε ότι $x_1x_2\cdots x_n\neq 0$. Διαιρώντας με $(x_1x_2\cdots x_n)^2$ και κάνοντας κάποιες απλοποιήσεις, βλέπουμε ότι η ανισότητα (8.14) είναι ισοδύναμη με την

$$\frac{1}{\binom{n}{2}}\sum_{1\le j<k\le n}\frac{1}{x_jx_k}\le\left\{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right\}^2. \tag{8.15}$$

Μπορούμε τώρα να επαναλάβουμε το επιχείρημα που χρησιμοποιήσαμε για την H_3, θα προτιμήσουμε όμως έναν διαφορετικό τρόπο, ο οποίος έρχεται σχεδόν

αμέσως στο νού. Αν υιοθετήσουμε τη γλώσσα των συμμετρικών συναρτήσεων, η ανισότητα (8.15) μπορεί να γραφτεί στη μορφή

$$E_0(1/x_1, 1/x_2, \ldots, 1/x_n)E_2(1/x_1, 1/x_2, \ldots, 1/x_n)$$
$$\leq E_1^2(1/x_1, 1/x_2, \ldots, 1/x_n),$$

και αυτή η ανισότητα ισχύει, αφού καλύπτεται από την πρώτη ομάδα (8.12). Έτσι, η απόδειξη των ανισοτήτων του Newton είναι πλήρης.

8.2 Η ανισότητα του Muirhead

Ο David Hilbert είπε κάποτε ότι «η τέχνη του να κάνεις μαθηματικά είναι το να βρίσκεις εκείνη την ειδική περίπτωση που θα σου δώσει τη δυνατότητα να γενικεύσεις». Εφαρμόζοντας αυτή τη φιλοσοφία, θα ξεκινήσουμε από μια στοιχειώδη ανισότητα, απ' όπου θα οδηγηθούμε στην πολύ γενικότερη ανισότητα του Muirhead.

Αν x, y και z είναι μη αρνητικοί πραγματικοί αριθμοί, τότε

$$x^2y^3 + x^2z^3 + y^2x^3 + y^2z^3 + z^2x^3 + z^2y^3$$

$$\leq xy^4 + xz^4 + yx^4 + yz^4 + zx^4 + zy^4. \tag{8.16}$$

Έχουμε ήδη συναντήσει αρκετά προβλήματα όπου η ανισότητα αριθμητικού–γεωμετρικού μέσου μας βοήθησε να καταλάβουμε τη σχέση ανάμεσα σε δύο ομογενή πολυώνυμα, και ελπίζουμε ότι η ίδια ιδέα θα μας βοηθήσει να δείξουμε ότι κάθε προσθετέος στο αριστερό μέλος μπορεί να γραφτεί σαν γεωμετρικός μέσος των προσθετέων του δεξιού μέλους. Μετά από κάποιους πειραματισμούς, θα παρατηρήσετε ότι αν a και b είναι δύο μη αρνητικοί αριθμοί, τότε $a^2b^3 = (ab^4)^{\frac{2}{3}}(a^4b)^{\frac{1}{3}}$, και, τότε, η ανισότητα αριθμητικού–γεωμετρικού μέσου (2.9) μας δίνει την ανισότητα

$$a^2b^3 = (ab^4)^{\frac{2}{3}}(a^4b)^{\frac{1}{3}} \leq \frac{2}{3}ab^4 + \frac{1}{3}a^4b. \tag{8.17}$$

Μένει να δούμε με ποιόν τρόπο θα την χρησιμοποιήσουμε. Αν στη θέση του (a, b) πάρουμε διαδοχικά τα διατεταγμένα ζεύγη (x, y) και (y, x), τότε το άθροισμα των δύο ανισοτήτων που θα προκύψουν μας δίνει την

$$x^2y^3 + y^2x^3 \leq xy^4 + x^4y.$$

Με τον ίδιο ακριβώς τρόπο, παίρνουμε ανάλογες ανισότητες προσθέτοντας την (8.17) για τα ζεύγη (x, z), (z, x) και τα ζεύγη (y, z), (z, y). Τέλος, προσθέτοντας τις τρείς ανισότητες που θα προκύψουν, καταλήγουμε στην ζητούμενη ανισότητα (8.16).

Το επιχείρημα που περιγράψαμε εφαρμόζεται, σχεδόν χωρίς τροποποιήσεις, σε κάθε συμμετρικό άθροισμα όρων της μορφής x^ay^b, αν όμως επιχειρήσουμε να χειριστούμε αθροίσματα που περιέχουν όρους της μορφής $x^ay^bz^c$, βλέπουμε ότι

η πολυπλοκότητά τους αυξάνεται τόσο που μπορεί να γίνει απαγορευτική αν δεν αναπτύξουμε μια συστηματική προσέγγιση στο πρόβλημα.

Ευτυχώς, η γεωμετρία έρχεται να μας δείξει το δρόμο. Ένα απλό σχήμα δείχνει ότι $(2,3) = \frac{2}{3}(1,4) + \frac{1}{3}(4,1)$, και, με δοκιμές, βλέπουμε ότι αυτό οδηγεί στην διάσπαση $a^2b^3 = (ab^4)^{\frac{2}{3}}(a^4b)^{\frac{1}{3}}$. Με τη γεωμετρία μπορούμε εύκολα να χειριστούμε όρους της μορφής $x^a y^b$, το πραγματικό όμως πλεονέκτημα της γεωμετρικής οπτικής γωνίας είναι ότι μας υποδεικνύει χρήσιμες αναπαραστάσεις για γινόμενα τριών ή περισσότερων μεταβλητών. Το κλειδί γι' αυτό είναι να βρούμε την σωστή γεωμετρική γενίκευση.

Μιλώντας αφηρημένα, για να λύσουμε το πρώτο μας πρόβλημα χρειάστηκε να παρατηρήσουμε ότι το $(2,3)$ ανήκει στην κυρτή θήκη του $(1,4)$ και της μετάθεσής του $(4,1)$. Πιό γενικά, αν μας δώσουν δύο n-διάστατα διανύσματα $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ και $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$, μπορούμε να θεωρήσουμε την ανάλογη κατάσταση όπου το α ανήκει στην κυρτή θήκη $H(\beta)$ του συνόλου των σημείων $(\beta_{\tau(1)}, \beta_{\tau(2)}, \ldots, \beta_{\tau(n)})$ τα οποία προσδιορίζονται αν αφήσουμε την τ να διατρέχει το σύνολο S_n όλων των $n!$ μεταθέσεων του $\{1, 2, \ldots, n\}$.

Αυτή η ιδέα οδηγεί στην ανισότητα του Robert Franklin Muirhead (1860-1941). Αποδείχτηκε το 1903, και ίσως φαίνεται πολύπλοκη αρχικά. Έχοντας όμως αποκτήσει κάποια εμπειρία, μπορούμε να εκτιμήσουμε τόσο την απλότητα όσο και την κομψότητά της.

Θεώρημα 8.2.1 (η ανισότητα του Muirhead). *Αν $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ και $\alpha \in H(\beta)$, τότε για κάθε n-άδα $x_1, x_2, \ldots, x_n$ θετικών αριθμών ισχύει η ανισότητα*

$$\sum_{\sigma \in S_n} x_{\sigma(1)}^{\alpha_1} x_{\sigma(2)}^{\alpha_2} \cdots x_{\sigma(n)}^{\alpha_n} \leq \sum_{\sigma \in S_n} x_{\sigma(1)}^{\beta_1} x_{\sigma(2)}^{\beta_2} \cdots x_{\sigma(n)}^{\beta_n}. \tag{8.18}$$

Για να εξοικειωθούμε με το συμβολισμό, ελέγχουμε πρώτα ότι η ανισότητα του πρώτου μας προβλήματος είναι ειδική περίπτωση της ανισότητας του Muirhead. Σε αυτήν την περίπτωση, S_3 είναι το σύνολο των έξι μεταθέσεων του συνόλου $\{1, 2, 3\}$, και έχουμε $(x_1, x_2, x_3) = (x, y, z)$. Επίσης, έχουμε

$$(\alpha_1, \alpha_2, \alpha_3) = (2, 3, 0) \quad \text{και} \quad (\beta_1, \beta_2, \beta_3) = (1, 4, 0).$$

Αφού $(2,3,0) = \frac{2}{3}(1,4,0) + \frac{1}{3}(4,1,0)$, βλέπουμε ότι $\alpha \in H(\beta)$. Τέλος, έχουμε το α-άθροισμα

$$\sum_{\sigma \in S_n} x_{\sigma(1)}^{\alpha_1} x_{\sigma(2)}^{\alpha_2} x_{\sigma(3)}^{\alpha_3} = x^2y^3 + x^2z^3 + y^2x^3 + y^2z^3 + z^2x^3 + z^2y^3,$$

ενώ το β-άθροισμα ισούται με

$$\sum_{\sigma \in S_n} x_{\sigma(1)}^{\beta_1} x_{\sigma(2)}^{\beta_2} x_{\sigma(3)}^{\beta_3} = xy^4 + xz^4 + yx^4 + yz^4 + zx^4 + zy^4.$$

Συνεπώς, η ανισότητα του Muirhead (8.18) όντως γενικεύει την ανισότητα (8.16).

Τέλος, πριν προχωρήσουμε στην απόδειξη, θα έπρεπε να παρατηρήσουμε ότι δεν υπάρχει περιορισμός σχετικά με το πρόσημο των συντεταγμένων των α και β στην

ανισότητα του Muirhead. Έτσι, για παράδειγμα, αν πάρουμε $\alpha = (1/2, 1/2, 0)$ και $\beta = (-1, 2, 0)$, τότε η ανισότητα του Muirhead μας λέει ότι, για κάθε τριάδα θετικών πραγματικών αριθμών x, y και z, ισχύει η ανισότητα

$$2(\sqrt{xy} + \sqrt{xz} + \sqrt{yz}) \leq \frac{x^2}{y} + \frac{x^2}{z} + \frac{y^2}{x} + \frac{y^2}{z} + \frac{z^2}{x} + \frac{z^2}{y}. \tag{8.19}$$

Αυτή η ανισότητα μπορεί να αποδειχτεί με πολλούς τρόπους: για παράδειγμα, θα μπορούσαμε να βασίσουμε την απόδειξη στην ανισότητα του Cauchy ή στην ανισότητα αριθμητικού–γεωμετρικού μέσου. Όμως, η ανισότητα του Muirhead δίνει μια άμεση ερμηνεία της ανισότητας και την εντάσσει στο ευρύτερο δυνατό πλαίσιο.

Απόδειξη της ανισότητας του Muirhead

Η υπόθεση ότι $\alpha \in H(\beta)$ είναι ισοδύναμη με τον ισχυρισμό ότι υπάρχουν $p_\tau \geq 0$ με $\sum_{\tau \in S_n} p_\tau = 1$, ώστε

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{\tau \in S_n} p_\tau (\beta_{\tau(1)}, \beta_{\tau(2)}, \ldots, \beta_{\tau(n)}).$$

Τώρα, αν χρησιμοποιήσουμε την j-οστή συντεταγμένη αυτής της ισότητας για να εκφράσουμε τον $x_{\sigma(j)}^{\alpha_j}$ σαν γινόμενο, μπορούμε να πολλαπλασιάσουμε ως προς j και να καταλήξουμε στην ισότητα

$$x_{\sigma(1)}^{\alpha_1} x_{\sigma(2)}^{\alpha_2} \cdots x_{\sigma(n)}^{\alpha_n} = \prod_{\tau \in S_n} \left(x_{\sigma(1)}^{\beta_{\tau(1)}} x_{\sigma(2)}^{\beta_{\tau(2)}} \cdots x_{\sigma(n)}^{\beta_{\tau(n)}} \right)^{p_\tau}.$$

Από εδώ και πέρα, η ανισότητα αριθμητικού–γεωμετρικού μέσου και κάποια αριθμητική θα δουλέψουν για μάς. Πιό συγκεκριμένα, έχουμε

$$\begin{aligned}
\sum_{\sigma \in S_n} x_{\sigma(1)}^{\alpha_1} x_{\sigma(2)}^{\alpha_2} \cdots x_{\sigma(n)}^{\alpha_n} &\leq \sum_{\sigma \in S_n} \prod_{\tau \in S_n} \left(x_{\sigma(1)}^{\beta_{\tau(1)}} x_{\sigma(2)}^{\beta_{\tau(2)}} \cdots x_{\sigma(n)}^{\beta_{\tau(n)}} \right)^{p_\tau} \\
&= \sum_{\sigma \in S_n} \sum_{\tau \in S_n} p_\tau x_{\sigma(1)}^{\beta_{\tau(1)}} x_{\sigma(2)}^{\beta_{\tau(2)}} \cdots x_{\sigma(n)}^{\beta_{\tau(n)}} \\
&= \sum_{\tau \in S_n} p_\tau \sum_{\sigma \in S_n} x_{\sigma(1)}^{\beta_{\tau(1)}} x_{\sigma(2)}^{\beta_{\tau(2)}} \cdots x_{\sigma(n)}^{\beta_{\tau(n)}} \\
&= \sum_{\tau \in S_n} p_\tau \sum_{\sigma \in S_n} x_{\sigma(1)}^{\beta_1} x_{\sigma(2)}^{\beta_2} \cdots x_{\sigma(n)}^{\beta_n} \\
&= \sum_{\sigma \in S_n} x_{\sigma(1)}^{\beta_1} x_{\sigma(2)}^{\beta_2} \cdots x_{\sigma(n)}^{\beta_n},
\end{aligned}$$

δηλαδή, την ανισότητα του Muirhead.

Βιβλιογραφία

[1] L. Carleson, A proof of an inequality of Carleman, *Proc. Amer. Math. Soc.* **5** (1954), 932–933.

[2] A. Cauchy, *Cours d'Analyse* (1821).

[3] J. Duncan and M. A. McGregor, Carleman's inequality, *Amer. Math. Monthly* **101** (2003), 424–431.

[4] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press (1952).

[5] G. Pólya, Proof of an inequality, *Proc. London Math. Soc.* **24** (1926), 57.

[6] G. Pólya, With, or without, motivation, *Amer. Math. Monthly* **57** (1949), 684–691.

[7] S. Rosset, Normalized symmetric functions, Newton's inequalities, and a new set of stronger inequalities, *Amer. Math. Monthly* **96** (1989), 815–819.

[8] J. M. Steele, *The Cauchy–Schwarz Master Class*, Cambridge University Press (2004).