ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

Μέθοδος Τυχαίων Προβολών και Εφαρμογές

Διπλωματική Εργασία

Άγγελος-Κίμων Μακρυνάκης

Εφαρμοσμένες Μαθηματικές Επιστήμες
Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών
Αθήνα – 2024

Περιεχόμενα

1 Εισαγωγή **1**

2 Εργαλεία από την θεωρία πιθανοτήτων **9**

2.1 Η πιθανοθεωρητική μέθοδος 9
2.2 Ανισότητα Markov 12
2.3 Η μέθοδος Cramér-Chernoff 12
2.4 Υποκανονικές τυχαίες μεταβλητές 15
2.5 Χώροι Orlicz 18
2.6 Υποεκθετικές τυχαίες μεταβλητές 22
2.7 Ανισότητες τύπου Bernstein 26
2.8 Ανισότητα Sudakov και η δυϊκή της 29

3 Η μέθοδος των τυχαίων προβολών **35**

3.1 Λήμμα Johnson-Lindenstrauss 35
3.2 Απόδειξη μέσω συγκέντρωσης του μέτρου στη σφαίρα 36
3.3 Απόδειξη με τυχαίους πίνακες 42
3.4 Το βέλτιστο του λήμματος 45
3.5 Αναγωγή της διάστασης στον L_1 52
3.6 ϵ-ισομετρική αναγωγή της διάστασης στον ℓ_p 61

4 Ανάκτηση σήματος **69**

4.1 Το πρόβλημα της ακριβούς ανακατασκευής 69
4.2 Η ιδιότητα της περιορισμένης ισομετρίας 76
4.3 Η γεωμετρία του πυρήνα 78
4.4 Πλάτος Gelfand 81
4.5 Gaussian τυχαίοι πίνακες 83
4.6 Σχεδόν αραιά διανύσματα 87
4.7 Ένας άλλος δείκτης πολυπλοκότητας 96

Κεφάλαιο 1

Εισαγωγή

Το θέμα αυτής της εργασίας είναι η εφαρμογή πιθανοθεωρητικών μεθόδων στο πρόβλημα της προσεγγιστικής και ακριβούς ανάκτησης αραιών σημάτων από λίγες μετρήσεις. Πιο συγκεκριμένα, έστω $T \subset \mathbb{R}^N$ και A ένας $n \times N$ πραγματικός πίνακας με γραμμές $Y_1, \dots, Y_n \in \mathbb{R}^N$. Θεωρούμε το πρόβλημα της ανακατασκευής τυχόντος διανύσματος $x \in T$ από τα δεδομένα $Ax \in \mathbb{R}^n$, δηλαδή από τις μετρήσεις

$$\langle Y_1, x\rangle, \dots, \langle Y_n, x\rangle.$$

Αυτό που ίσως θα περίμενε κανείς είναι ότι το πλήθος n των μετρήσεων που απαιτούνται ώστε να πετύχουμε ακριβή ανάκτηση του x πρέπει να είναι της τάξης της διάστασης N του x. Οι μέθοδοι που θα περιγράψουμε μας επιτρέπουν να ανακτήσουμε το αρχικό σήμα x από τις μετρήσεις Ax χρησιμοποιώντας πολύ μικρότερο πλήθος μετρήσεων $n \ll N$. Για τον σκοπό αυτό πρέπει φυσικά να έχουμε κάποιες υποθέσεις για το σύνολο T των σημάτων και επίσης να επιλέξουμε κατάλληλα τον πίνακα A. Επομένως, φυσιολογικά, χρειάζεται να ασχοληθούμε με τα εξής δύο σημεία:

(i) Την *πολυπλοκότητα* του συνόλου T. Για να «μετρήσουμε» την πολυπλοκότητα ενός συνόλου μελετάμε διάφορα εργαλεία από την θεωρία προσέγγισης και την γεωμετρία των χώρων πεπερασμένης διάστασης με νόρμα.

(ii) Την επιλογή του *πίνακα μετρήσεων* A. Θα χρησιμοποιήσουμε τυχαίους πίνακες που προκύπτουν αν επιλέξουμε τυχαία διανύσματα $Y_1, \dots, Y_n \in \mathbb{R}^N$ με κατάλληλο τρόπο. Σε αυτήν την εργασία, οι τυχαίοι πίνακες είναι συνήθως Gaussian ή Bernoulli (± 1) πίνακες.

Για τη μελέτη αυτών των τυχαίων πινάκων θα χρησιμοποιήσουμε αρκετά εργαλεία από την θεωρία πιθανοτήτων: κλασσικές πιθανοθεωρητικές ανισότητες για αθροίσματα ανεξάρτητων τυχαίων μεταβλητών, αποτελέσματα συγκέντρωσης του μέτρου, στοιχεία από την θεωρία των εμπειρικών διαδικασιών και την θεωρία τυχαίων πινάκων. Στο Κεφάλαιο 2 παρουσιάζουμε αυτά τα εργαλεία. Μεταξύ άλλων, μελετάμε τις βασικές ιδιότητες των υποκανονικών και υποεκθετικών τυχαίων μεταβλητών, τη συμπεριφορά των αθροισμάτων ανεξάρτητων υποεκθετικών τυχαίων μεταβλητών, τη διακριτοποίηση με τη βοήθεια εκτιμήσεων για αριθμούς κάλυψης.

Στο Κεφάλαιο 3 μελετάμε αρχικά το κλασσικό λήμμα Johnson-Lindenstrauss [22]: Αν V είναι ένα σύνολο n-σημείων σε κάποιον Ευκλείδειο χώρο $(H, \|\cdot\|)$ τότε για κάθε $\epsilon \in (0,1]$ υπάρχει $(1+\epsilon)$-εμφύτευση του V στον ℓ_2^k για κάποιον $k = O(\epsilon^{-2}\log n)$. Λέγοντας $(1+\epsilon)$-εμφύτευση εννοούμε μια απεικόνιση $T : V \to \ell_2^k$ τέτοια ώστε, για κάθε $x, y \in V$ να ισχύει ότι

$$\frac{1}{\sqrt{1+\epsilon}}\|x-y\| \leqslant \|T(x)-T(y)\|_{\ell_2^k} \leqslant \sqrt{1+\epsilon}\|x-y\|.$$

Παρουσιάζουμε δύο αποδείξεις αυτού του θεωρήματος. Η πρώτη εκμεταλλεύεται τη συγκέντρωση του μέτρου στην Ευκλείδεια σφαίρα S^{n-1} του $\mathbb{R}^n$, η οποία έχει (όπως θα δούμε) ως συνέπεια το ακόλουθο θεώρημα για Lipschitz συναρτήσεις ορισμένες στη σφαίρα, γνωστό ως «λήμμα του Lévy»: Αν $f : S^{n-1} \to \mathbb{R}$ είναι μια 1-Lipschitz συνάρτηση τότε, για κάθε $t > 0$,

$$\mathbb{P}[f < \mathrm{med}(f) - t] \leqslant e^{-t^2 n/2} \quad \text{και} \quad \mathbb{P}[f > \mathrm{med}(f) + t] \leqslant e^{-t^2 n/2}.$$

Για την απόδειξη του λήμματος Johnson-Lindenstrauss, χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι $H = \ell_2^n$ και εφαρμόζουμε το λήμμα του Lévy για τη συνάρτηση $f_k : S^{n-1} \to \mathbb{R}$ που απεικονίζει κάθε $x \in S^{n-1}$ στο μήκος $f_k(x)$ της προβολής του πάνω στον υπόχωρο H_k που παράγεται από τα πρώτα k διανύσματα της συνήθους ορθοκανονικής βάσης του $\mathbb{R}^n$. Δηλαδή,

$$f_k(x) = \sqrt{x_1^2 + \cdots + x_k^2}$$

όπου $x = (x_1, \ldots, x_n)$. Τον ρόλο της $T : \ell_2^n \to \ell_2^k$ παίζει η (τυχαία) ορθογώνια προβολή $p_F : \ell_2^n \to F$, όπου F είναι τυχαίος k-διάστατος γραμμικός υπόχωρος του ℓ_2^n ως προς το μέτρο Haar στην πολλαπλότητα Grassman $G_{n,k}$ των k-διάστατων υποχώρων του $\mathbb{R}^n$. Στην δεύτερη απόδειξη που παρουσιάζουμε, θεωρούμε $k \times n$ τυχαίους πίνακες $A = (a_{ij})$, τα στοιχεία των οποίων είναι ανεξάρτητες υποκανονικές τυχαίες μεταβλητές με μέση τιμή 0 και διασπορά 1. Ο ζητούμενος τελεστής T επάγεται από έναν τυχαίο πίνακα της μορφής λA, για κατάλληλη επιλογή της παραμέτρου $\lambda \in \mathbb{R}$ ώστε η τυχαία μεταβλητή $\|T(u)\|_2$ (για σταθερό $u \in S^{n-1}$) να έχει μέση τιμή 1.

Η εξάρτηση του k από το πλήθος n των σημείων του V είναι βέλτιστη. Υπάρχει σύνολο $V \subset H$ με πληθάριθμο n τέτοιο ώστε για κάθε τελεστή $T : H \to \ell_2^k$ που $(1+\epsilon)$-διατηρεί τις αποστάσεις για $\epsilon = 1/2$ να έχουμε $k > c \cdot \log n$. Πιο συγκεκριμένα, με ένα απλό ογκομετρικό επιχείρημα μπορούμε να δούμε ότι αν $V = \{u_1, u_2, \ldots, u_n\}$ είναι μια ορθοκανονική βάση του $\mathbb{R}^n$, τότε για κάθε τελεστή $T : \mathbb{R}^n \to \mathbb{R}^k$ με την ιδιότητα ότι

$$\|x_i - x_j\|_2 \leqslant \|T(x_i) - T(x_j)\|_2 \leqslant 3\|x_i - x_j\|_2, \quad \text{για κάθε } i, j \leqslant n,$$

ισχύει ότι $k = \Omega(\log |V|)$. Το ερώτημα αν το λήμμα Johnson-Lindenstrauss είναι βέλτιστο τόσο ως προς το πλήθος n των σημείων όσο και ως προς το ϵ είναι πολύ πιο απαιτητικό. Περιγράφουμε τη δουλειά των K. Green-Larsen και J. Nelson, οι οποίοι απέδειξαν στο [20] ότι για κάθε $n > 1$ και $\epsilon \in (0, 1/2)$ υπάρχει $V \subset \mathbb{R}^n$ με $|V| = N = n^3 + n$ τέτοιο ώστε αν $T : \mathbb{R}^n \to \mathbb{R}^k$

είναι γραμμικός τελεστής με την ιδιότητα ότι

$$(1-\epsilon)\|v_i\|_2^2 \leqslant \|T(v_i)\|_2^2 \leqslant (1+\epsilon)\|v_i\|_2^2$$

για κάθε $v_i \in V$, τότε $k = \Omega(\epsilon^{-2}\log n) = \Omega(\epsilon^{-2}\log N)$. Στη μεταγενέστερη εργασία [21] οι ίδιοι συγγραφείς απέδειξαν το εξής ισχυρότερο αποτέλεσμα: Για κάθε $d, N \geqslant 2$ και $1/(\min\{N,d\})^{0.4999} < \varepsilon < 1$, υπάρχει σύνολο N διανυσμάτων $V \subset \mathbb{R}^d$ τέτοιο ώστε αν $T : V \to \mathbb{R}^k$ είναι εμφύτευση με την ιδιότητα

$$(1-\epsilon)\|x-y\|_2^2 \leqslant \|T(x)-T(y)\|_2^2 \leqslant (1+\epsilon)\|x-y\|_2^2$$

για κάθε $x, y \in V$, τότε $k = \Omega(\epsilon^{-2}\log N)$.

Στο τελευταίο μέρος του Κεφαλαίου 3 παρουσιάζουμε κάποια αποτελέσματα σχετικά με τη δυνατότητα επέκτασης του λήμματος Johnson-Lindenstrauss για υποσύνολα άλλων χώρων Banach πέραν των χώρων Hilbert H. Γενικά, μια απεικόνιση $f : (X,d) \to (Y,\sigma)$ μεταξύ δύο μετρικών χώρων λέγεται αμφι-Lipschitz εμφύτευση με παραμόρφωση $\alpha \geqslant 1$ αν υπάρχει σταθερά $c > 0$ τέτοια ώστε

$$c\,d(x,y) \leqslant \sigma(f(x),f(y)) \leqslant \alpha c\,d(x,y)$$

για κάθε $x, y \in X$. Με αυτήν την ορολογία, το λήμμα Johnson-Lindenstrauss μας λέει ότι για κάθε $\epsilon \in (0,1)$, $n \in \mathbb{N}$ και $k \geqslant \frac{C\log n}{\epsilon^2}$, κάθε σύνολο n σημείων ενός χώρου Hilbert επιδέχεται μια αμφι-Lipschitz εμφύτευση στον ℓ_2^k με παραμόρφωση το πολύ ίση με $1+\epsilon$.

Λέμε ότι ένας απειροδιάστατος χώρος Banach $(X, \|\cdot\|)$ επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης αν υπάρχει σταθερά $\alpha = \alpha(X) \geqslant 1$ τέτοια ώστε για κάθε $n \in \mathbb{N}$ να υπάρχει $k_n = k_n(X,\alpha) \in \mathbb{N}$ που ικανοποιεί την

$$\lim_{n\to\infty} \frac{\log k_n}{\log n} = 0$$

ώστε κάθε σύνολο $S \subseteq X$ με n σημεία να επιδέχεται μια αμφι-Lipschitz εμφύτευση με παραμόρφωση το πολύ ίση με α σε έναν υπόχωρο Y του X που έχει πεπερασμένη διάσταση $\dim(Y) \leqslant k_n$. Είναι γνωστό ότι ο χώρος ℓ_∞ δεν επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης και ότι υπάρχει σύνολο $S \subseteq \ell_1$ με n σημεία το οποίο δεν επιδέχεται αμφι-Lipschitz εμφύτευση σε κανέναν $n^{o(1)}$-διάστατο υπόχωρο του ℓ_1. Το τελευταίο αποτέλεσμα αποδείχθηκε από τους Lee και Naor, στην εργασία [25] την οποία περιγράφουμε, με χρήση των λεγόμενων diamond graphs. Πιο συγκεκριμένα, έδειξαν ότι για κάθε $n \in \mathbb{N}$ υπάρχει $X \subset L_1$ με n σημεία, τέτοιο ώστε για κάθε $D > 1$, αν ο X εμφυτεύεται με παραμόρφωση D στον ℓ_1^k τότε $k \geqslant n^{\Omega(1/D^2)}$.

Το επόμενο όμως ερώτημα παραμένει ανοικτό:

> *Για ποιές τιμές του $p \notin \{1,2,\infty\}$ είναι σωστό ότι ο ℓ_p επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης;*

Ο Εσκενάζης [13] μελέτησε μια παραλλαγή αυτού του ερωτήματος. Έστω μ ένα μέτρο πιθα-

νότητας. Αν S είναι ένα υποσύνολο του $L_p(\mu)$, ορίζουμε

$$J(S) := \left\| \max_{x \in S} |x| \right\|_{L_p(\mu)}$$

και λέμε ότι το S είναι K-μη-συμπιέσιμο αν $J(S) \leqslant K$. Παρουσιάζουμε ένα θεώρημα αναγωγής της διάστασης που απέδειξε για μη-συμπιέσιμα υποσύνολα του $L_p(\mu)$, το οποίο ισχύει για κάθε $p \in [1, \infty)$: Έστω $1 \leqslant p < \infty$, $n \in \mathbb{N}$, $0 < K < \infty$ και $\{x_1, \ldots, x_n\}$ μια K-μη-συμπιέσιμη οικογένεια διανυσμάτων στον $L_p(\mu)$, όπου μ είναι ένα μέτρο πιθανότητας. Τότε, για κάθε $\epsilon \in (0,1)$ υπάρχουν $d \in \mathbb{N}$ με $d \leqslant \frac{32e^2(2K)^{2p}\log n}{\epsilon^2}$ και σημεία $y_1, \ldots, y_n \in \ell_p^d$ τέτοια ώστε, για κάθε $i, j = 1, \ldots, n$,

$$\|x_i - x_j\|_{L_p(\mu)}^p - \epsilon \leqslant \|y_i - y_j\|_{\ell_p^d}^p \leqslant \|x_i - x_j\|_{L_p(\mu)}^p + \epsilon.$$

Πέρα από την παράμετρο K της μη-συμπιεσιμότητας που υπεισέρχεται στο φράγμα για τη διάσταση d του χώρου ℓ_p^d στον οποίο επιτυγχάνεται η εμφύτευση, το αποτέλεσμα αυτό διαφέρει από το λήμμα Johnson-Lindenstrauss στο ότι το σφάλμα στην τελευταία ανισότητα είναι προσθετικό και όχι πολλαπλασιαστικό. Λέμε ότι μια απεικόνιση $f : (X, d) \to (Y, \sigma)$ μεταξύ μετρικών χώρων είναι ϵ-ισομετρική εμφύτευση αν

$$|\sigma(f(x), f(y)) - d(x, y)| \leqslant \epsilon$$

για κάθε $x, y \in X$. Χρησιμοποιώντας τη στοιχειώδη ανισότητα $|a - b| \leqslant |a^p - b^p|^{1/p}$ που ισχύει για κάθε $a, b > 0$, βλέπουμε ότι το παραπάνω θεώρημα μπορεί να διατυπωθεί ως εξής: κάθε K-μη-συμπιέσιμο σύνολο n σημείων στον $L_p(\mu)$ επιδέχεται $\epsilon^{1/p}$-ισομετρική εμφύτευση στον ℓ_p^d για κάποιον $d \leqslant \frac{32e^2(2K)^{2p}\log n}{\epsilon^2}$.

Στο Κεφάλαιο 4 παρουσιάζουμε σημαντικές μαθηματικές ιδέες και τεχνικές που έχουν αναπτυχθεί για το πρόβλημα της ανάκτησης σήματος. Αρχικά, μελετάμε το πρόβλημα της ακριβούς ανακατασκευής. Η υπόθεση που κάνουμε για το σύνολο T των σημάτων είναι ότι αυτά είναι *αραιά*. Λέμε ότι ένα διάνυσμα είναι m-αραιό αν το πλήθος των μη μηδενικών συντεταγμένων του είναι το πολύ ίσο με m. Το σύνολο όλων των m-αραιών διανυσμάτων έχει την ιδιότητα ότι η τομή του με την Ευκλείδεια μοναδιαία σφαίρα είναι η ένωση μοναδιαίων σφαιρών που περιέχονται στους m-διάστατους υπόχωρους συντεταγμένων, και συνεπώς είναι ισχυρά συμπαγές αν θεωρήσουμε ως μέτρο της συμπάγειας τους αριθμούς κάλυψης: είναι *μικρό* υποσύνολο της σφαίρας αν $m \ll N$.

Η συνήθης τεχνική ανάκτησης σήματος βασίζεται στη χρήση αποτελεσματικών αλγορίθμων όπως η μέθοδος της ℓ_1-ελαχιστοποίησης: για δοθέντα δεδομένα $y = Ax$, λύνουμε το πρόβλημα

$$\min_{t \in \mathbb{R}^N} \sum_{i==1}^{N} |t_i| \qquad \text{υπό τον περιορισμό} \qquad At = y.$$

Τότε, το πρόβλημα μεταφράζεται στο να επιλέξουμε πίνακες για τους οποίους αυτός ο αλγόριθμος θα εξασφαλίζει ότι κάθε m-αραιό διάνυσμα ανακτάται ακριβώς με τη μέθοδο της ℓ_1-ελαχιστοποίησης για m αρκετά μεγάλο. Μελετώντας τον κώνο των περιορισμών ώστε να

βεβαιωθούμε ότι τα m-αραιά διανύσματα ανακτώνται ακριβώς, οδηγούμαστε σε μια αναγκαία και ικανή συνθήκη, η οποία είναι γνωστή ως *ιδιότητα πυρήνα* τάξης m: για κάθε $w \in \ker(A)$, $w \neq 0$ και για κάθε $I \subseteq [N] = \{1, \dots, N\}$ με $|I| \leqslant m$ πρέπει να ισχύει ότι

$$\sum_{i \in I} |w_i| < \sum_{i \in I^c} |w_i|.$$

Αυτή η ιδιότητα έχει μια ωραία γεωμετρική ερμηνεία που αφορά τη δομή των εδρών τυχαίων πολυτόπων. Αν P είναι το πολύτοπο που προκύπτει αν θεωρήσουμε τη συμμετρική κυρτή θήκη των στηλών του A, τότε η ιδιότητα πυρήνα τάξης m για τον A είναι ισοδύναμη με το να έχουμε ότι το P είναι *m-γειτνιάζον*. Αυτό σημαίνει ότι ο πίνακας A που απεικονίζει τις κορυφές του cross-polytope $B_1^N = \left\{ t \in \mathbb{R}^N : \sum_{i=1}^N |t_i| \leqslant 1 \right\}$ στις κορυφές του P διατηρεί τη δομή των k-διάστατων εδρών για όλες τις διαστάσεις $k \leqslant m$. Αυτή η σύνδεση της ανάκτησης σήματος με τη γεωμετρία στις μεγάλες διαστάσεις παρατηρήθηκε για πρώτη φορά από τον Donoho στο [11].

Ένα μειονέκτημα της ιδιότητας πυρήνα είναι ότι δεν είναι εύκολο να την επαληθεύσουμε. Μια ικανή όμως συνθήκη, η λεγόμενη *ιδιότητα περιορισμένης ισομετρίας* (RIP) είναι πολύ πιο εύχρηστη. Η ιδιότητα περιορισμένης ισομετρίας τάξης m απαιτεί όλοι οι υποπίνακες στηλών τάξης m του πίνακα να έχουν καλό δείκτη κατάστασης. Ακριβέστερα, λέμε ότι ο A ικανοποιεί την ιδιότητα περιορισμένης ισομετρίας τάξης p με παράμετρο $\delta = \delta_p$ αν

$$1 - \delta_p \leqslant \|Ax\|_2^2 \leqslant 1 + \delta_p$$

για όλα τα p-αραιά μοναδιαία διανύσματα $x \in \mathbb{R}^N$. Όπως θα δούμε, αν ο A έχει την ιδιότητα περιορισμένης ισομετρίας τάξης $2m$ με παράμετρο δ αρκετά μικρή, τότε μπορούμε να ανακτήσουμε κάθε m-αραιό διάνυσμα με τη μέθοδο της ℓ_1-ελαχιστοποίησης. Παρόλο που, πάλι, είναι δύσκολο να ελέγξουμε την (RIP) για δοθέντα πίνακα, μπορούμε να δείξουμε ότι ισχύει με μεγάλη πιθανότητα για κάποια μοντέλα τυχαίων πινάκων και ελέγχεται εύκολα για κάποια από αυτά. Σε αυτό το σημείο χρησιμοποιούμε πιθανοθεωρητικές μεθόδους: ελέγχουμε την ιδιότητα περιορισμένης ισομετρίας για Gaussian και Bernoulli τυχαίους πίνακες, χρησιμοποιώντας τεχνικές διακριτοποίησης που αναπτύσσουμε στο Κεφάλαιο 2.

Ένας διαφορετικός τρόπος για να αντιμετωπίσουμε το πρόβλημα της ανάκτησης σήματος με τη μέθοδο της ℓ_1-ελαχιστοποίηση είναι να αναλύσουμε την Εθκλείδεια διάμετρο των τομών του cross-polytope B_1^N με τον πυρήνα του A. Οδηγούμαστε έτσι στη μελέτη του *πλάτους Gelfand*, ειδικά για την B_1^N. Τα πλάτη Gelfand της B_1^N ορίζονται ως εξής:

$$d_n(B_1^N, \ell_2^N) = \inf_{\operatorname{codim}(S) \leqslant n} \operatorname{rad}(S \cap B_1^N), \qquad 1 \leqslant n \leqslant N$$

όπου

$$\operatorname{rad}(S \cap B_1^N) = \max\{\|x\|_2 : x \in S \cap B_1^N\}$$

είναι το μισό της Ευκλείδειας διαμέτρου της τομής της B_1^N με τον S. Σε αυτό το σημείο εκμεταλλευόμαστε σημαντικά αποτελέσματα από την θεωρία προσέγγισης και την ασυμπτωτική

γεωμετρική ανάλυση, με αφετηρία το θεώρημα του Kashin [23] το οποίο εξασφαλίζει ότι

$$d_n(B_1^N, \ell_2^N) \leqslant \frac{C}{\sqrt{n}} \log^{O(1)}(N/n)$$

για κάποια απόλυτη σταθερά $C > 0$. Η σχέση αυτού του αποτελέσματος με την ανάκτηση σήματος δίνεται από το εξής θεώρημα: Αν $1 \leqslant m \leqslant n$ και

$$\mathrm{rad}(\ker(A) \cap B_1^N) \leqslant \frac{1}{2\sqrt{m}},$$

τότε κάθε m-αραιό διάνυσμα ανακτάται ακριβώς με τη μέθοδο της ℓ_1-ελαχιστοποίησης. Σύμφωνα με αυτό το αποτέλεσμα, στόχος μας είναι να δώσουμε άνω φράγμα για τη διάμετρο $\mathrm{rad}(\ker(A) \cap B_1^N)$. Συζητάμε αυτό το ερώτημα για κάποια μοντέλα τυχαίων πινάκων, αξιοποιώντας το εξής θεώρημα: Αν ο A έχει την ιδιότητα περιορισμένης ισομετρίας τάξης p με παράμετρο δ, τότε

$$\mathrm{rad}(\ker(A) \cap B_1^N) \leqslant \frac{C}{\sqrt{p}} \frac{1}{1-\delta},$$

όπου $C > 0$ είναι μια απόλυτη σταθερά. Συνεπώς, έχουμε την επιθυμητή εκτίμηση $\mathrm{rad}(\ker(A) \cap B_1^N) \leqslant \frac{1}{2\sqrt{m}}$ για $m = O(p)$.

Μπορούμε να επεκτείνουμε την εφαρμογή της μεθόδου ℓ_1-ελαχιστοποίησης στο πρόβλημα της προσεγγιστικής ανάκτησης διανυσμάτων που δεν απέχουν πολύ από το να είναι αραιά. Έστω $x \in \mathbb{R}^N$ και x^* μια λύση του προβλήματος

$$\min_{t \in \mathbb{R}^N} \sum_{i=1}^N |t_i| \qquad \text{υπό τον περιορισμό} \quad At = Ax.$$

Πάλι, αξιοποιούμε την έννοια του πλάτους Gelfand δείχνοντας το εξής θεώρημα: Αν

$$\mathrm{rad}(\ker(A) \cap B_1^N) < \frac{1}{4\sqrt{m}}$$

τότε για κάθε $I \subseteq [N]$ με $|I| \leqslant m$ και για κάθε $x \in \mathbb{R}^N$ ισχύει ότι

$$\|x - x^*\|_2 \leqslant \frac{1}{\sqrt{m}} \sum_{i \notin I} |x_i|.$$

Αυτό ισχύει ειδικότερα για τα μοναδιαία διανύσματα του $\ell_{p,\infty}^N$, $0 < p < 1$, για τα οποία ισχύει ότι $\min_{|I| \leqslant m} \sum_{i \notin I} |x_i| = O(m^{1-1/p})$.

Στο τελευταίο μέρος της εργασίας εισάγουμε την παράμετρο πολυπλοκότητας $\ell_*(T)$ ενός $T \subset \mathbb{R}^N$ που ορίζεται ως εξής:

$$\ell_*(T) = \mathbb{E} \sup_{t \in T} X_t,$$

όπου (X_t) είναι η Gaussian διαδικασία $X_t = \sum_{i=1}^N t_i g_i$, με δείκτες $t = (t_i)_{i=1}^N \in T$, όπου $g_1, \ldots, g_N$ ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές. Η παράμετρος αυτή παίζει βασικό ρόλο

στις εμπειρικές διαδικασίες και στην ασυμπτωτική γεωμετρική ανάλυση. Συγκεκριμένα, αποδεικνύουμε το εξής θεώρημα: Υπάρχουν απόλυτες σταθερές $c, c' > 0$ τέτοιες ώστε να ισχύει το ακόλουθο: Έστω $1 \leqslant n \leqslant N$ και A ένας Gaussian τυχαίος πίνακας με ανεξάρτητες τυπικές κανονικές συντεταγμένες. Έστω $T \subset S^{n-1}$ ένα αστρόμορφο σύνολο. Τότε, με πιθανότητα μεγαλύτερη ή ίση από $1 - \exp(-c'n)$,

$$\mathrm{rad}(\ker(A) \cap T) \leqslant c\ell_*(T)/\sqrt{n}.$$

Ως πόρισμα έχουμε το εξής φράγμα για την $\mathrm{rad}(\ker(A) \cap T)$: αν $1 \leqslant n \leqslant N$, $\lambda > 0$ και A είναι ένας Gaussian τυχαίος πίνακας με ανεξάρτητες τυπικές κανονικές συντεταγμένες, τότε αν για κάποιο $T \subset S^{n-1}$ έχουμε $T \subset 2\mathrm{conv}(\Lambda)$ όπου $\Lambda \subset B_2^N$ με $|\Lambda| \leqslant \exp(\lambda^2 n)$, με πιθανότητα μεγαλύτερη ή ίση από $1 - \exp(-c'n)$ ισχύει ότι

$$\mathrm{rad}(\ker(A) \cap T) \leqslant c\lambda.$$

Κεφάλαιο 2

Εργαλεία από την θεωρία πιθανοτήτων

2.1 Η πιθανοθεωρητική μέθοδος

Η πιθανοθεωρητική μέθοδος έχει αναπτυχθεί ιδιαίτερα τα τελευταία πενήντα χρόνια και έχει εξελιχθεί σε ένα από τα πιο σημαντικά εργαλεία που χρησιμοποιούνται στη Συνδυαστική, την Θεωρητική Πληροφορική αλλά και στα Μαθηματικά γενικότερα, στην Ανάλυση, τη Γεωμετρία και τη Θεωρία Αριθμών. Η βασική ιδέα πίσω από τη μέθοδο μπορεί να περιγραφεί ως εξής. Προκειμένου να αποδείξουμε την ύπαρξη κάποιας συνδυαστικής δομής με έναν αριθμό από ιδιότητες, κατασκευάζουμε κατάλληλο χώρο πιθανότητας και αποδεικνύουμε ότι ένα στοιχείο που επιλέγεται τυχαία σε αυτόν τον χώρο έχει τις επιθυμητές ιδιότητες με θετική πιθανότητα.

Σε αυτήν την εισαγωγική ενότητα περιγράφουμε την χρησιμότητα της πιθανοθεωρητικής μεθόδου μέσα από ένα παράδειγμα που προέρχεται από τη Γεωμετρία. Υπενθυμίζουμε ότι κυρτός συνδυασμός των σημείων $z_1, \ldots, z_m \in \mathbb{R}^n$ είναι ένας γραμμικός συνδυασμός τους με συντελεστές που είναι μη αρνητικοί και έχουν άθροισμα 1, δηλαδή ένα άθροισμα της μορφής

$$\sum_{i=1}^{m} \lambda_i z_i \quad \text{όπου} \quad \lambda_i \geqslant 0 \quad \text{και} \quad \sum_{i-1}^{m} \lambda_i = 1. \tag{2.1}$$

Η κυρτή θήκη $\mathrm{conv}(T)$ ενός συνόλου $T \subset \mathbb{R}^n$ είναι το σύνολο όλων των κυρτών συνδυασμών όλων των πεπερασμένων συνόλων σημείων του T. Το πλήθος m των σημείων του T που ορίζουν έναν κυρτό συνδυασμό στο $\mathrm{conv}(T)$ δεν είναι εξ αρχής περιορισμένο. Όμως, το κλασσικό θεώρημα του Καραθεοδωρή μας εξασφαλίζει ότι μπορούμε πάντα να έχουμε $m \leqslant n+1$.

Θεώρημα 2.1.1 (Καραθεοδωρή). *Κάθε σημείο στην κυρτή θήκη* $\mathrm{conv}(T)$ *ενός συνόλου* $T \subset \mathbb{R}^n$ *αναπαρίσταται ως κυρτός συνδυασμός* $n+1$ *το πολύ σημείων του* T.

Το φράγμα $n+1$ δεν μπορεί να βελτιωθεί: θεωρήστε για παράδειγμα ένα simplex $S = \mathrm{conv}(T)$ (την κυρτή θήκη ενός συνόλου T που αποτελείται από $n+1$ αφινικά ανεξάρτητα σημεία

στον $\mathbb{R}^n$). Ας υποθέσουμε όμως ότι θέλουμε απλώς να προσεγγίσουμε ένα σημείο $x \in \mathrm{conv}(T)$ και όχι να το αναπαραστήσουμε ακριβώς ως κυρτό συνδυασμό. Το επόμενο θεώρημα δείχνει ότι μπορούμε να το επιτύχουμε χρησιμοποιώντας κυρτούς συνδυασμούς από ένα μικρότερο πλήθος σημείων του T, μπορούμε μάλιστα να επιτύχουμε αυτό το πλήθος να μην εξαρτάται καθόλου από τη διάσταση n.

Θεώρημα 2.1.2 (προσεγγιστικό θεώρημα Καραθεοδωρή). *Έστω T ένα υποσύνολο του $\mathbb{R}^n$ με διάμετρο*

$$\mathrm{diam}(T) = \sup\{\|s-t\|_2 : s,t \in T\} \leqslant 1.$$

Τότε, για κάθε $x \in \mathrm{conv}(T)$ και κάθε ακέραιο $k \geqslant 1$, μπορούμε να βρούμε σημεία $x_1,\dots,x_k \in T$ ώστε

$$\Big\| x - \frac{1}{k}\sum_{j=1}^{k} x_j \Big\|_2 \leqslant \frac{1}{\sqrt{k}}.$$

Δύο σημαντικές παρατηρήσεις γι᾽ αυτό το θεώρημα, οι οποίες προκαλούν κάποια έκπληξη, είναι ότι το πλήθος k των σημείων στον κυρτό συνδυασμό δεν εξαρτάται από τη διάσταση και ότι όλοι οι συντελεστές στον κυρτό συνδυασμό είναι ίσοι με $1/k$ (παρατηρήστε όμως ότι επιτρέπουμε επαναλήψεις μεταξύ των x_j).

Απόδειξη. Η απόδειξη γίνεται με τη λεγόμενη εμπειρική μέθοδο του B. Maurey (βλ. [31]). Μεταφέροντας το T αν χρειαστεί, μπορούμε να υποθέσουμε ότι όχι μόνο η διάμετρος αλλά και η περιγεγραμμένη ακτίνα του T είναι φραγμένη από 1, δηλαδή

$$\|t\|_2 \leqslant 1 \quad \text{για κάθε} \quad t \in T.$$

Σταθεροποιούμε ένα σημείο $x \in \mathrm{conv}(T)$ και αρχικά το εκφράζουμε ως κυρτό συνδυασμό κάποιων σημείων $z_1,\dots,z_m \in T$ όπως στην (2.1). Μπορούμε τώρα να ερμηνεύσουμε τον ορισμό του κυρτού συνδυασμού στην (2.1) πιθανοθεωρητικά, με τους συντελεστές λ_i να παίζουν τον ρόλο των πιθανοτήτων. Πιο συγκεκριμένα, μπορούμε να θεωρήσουμε ένα τυχαίο διάνυσμα Z που παίρνει την τιμή z_i με πιθανότητα λ_i:

$$\mathbb{P}(Z = z_i) = \lambda_i, \qquad i = 1,\dots,m.$$

Αυτό είναι συνεπές διότι τα βάρη λ_i είναι μη αρνητικοί αριθμοί και το άθροισμά τους είναι ίσο με 1. Παρατηρήστε ότι

$$\mathbb{E}(Z) = \sum_{i=1}^{m} \lambda_i z_i = x.$$

Θεωρούμε ανεξάρτητα αντίγραφα $Z_1, Z_2, \dots$ του Z. Από τον ισχυρό νόμο των μεγάλων αριθμών,

$$\frac{1}{k}\sum_{j=1}^{k} Z_j \longrightarrow x \quad \text{σχεδόν βεβαίως καθώς} \quad k \to \infty.$$

Για να πάρουμε μια ποσοτική εκδοχή αυτού του αποτελέσματος, υπολογίζουμε τη διασπορά

της $\frac{1}{k}\sum_{j=1}^{k} Z_j$. Παίρνοντας υπ' όψιν την $\mathbb{E}(Z_i - x) = 0$ βλέπουμε ότι

$$\mathbb{E}\left\|x - \frac{1}{k}\sum_{j=1}^{k} Z_j\right\|_2^2 = \frac{1}{k^2}\mathbb{E}\left\|\sum_{j=1}^{k}(Z_j - x)\right\|_2^2 = \frac{1}{k^2}\sum_{j=1}^{k}\mathbb{E}\|Z_j - x\|_2^2.$$

Παρατηρούμε τώρα ότι

$$\begin{aligned}\mathbb{E}\|Z_j - x\|_2^2 &= \mathbb{E}\|Z - \mathbb{E}(Z)\|_2^2 = \mathbb{E}\|Z\|_2^2 - \|\mathbb{E}(Z)\|_2^2 \\ &\leqslant \mathbb{E}\|Z\|_2^2 \leqslant 1,\end{aligned}$$

διότι $Z \in T$, άρα $\|Z\|_2 \leqslant 1$. Έτσι, έχουμε αποδείξει ότι

$$\mathbb{E}\left\|x - \frac{1}{k}\sum_{j=1}^{k} Z_j\right\|_2^2 \leqslant \frac{1}{k}.$$

Υπάρχει λοιπόν κάποια υλοποίηση των τυχαίων διανυσμάτων $Z_1, \ldots, Z_k$ που ικανοποιεί την

$$\left\|x - \frac{1}{k}\sum_{j=1}^{k} Z_j\right\|_2^2 \leqslant \frac{1}{k}.$$

Από την κατασκευή, κάθε Z_j παίρνει τιμές στο T, άρα η απόδειξη είναι πλήρης. □

Δίνουμε τώρα μια εφαρμογή του Θεωρήματος 2.1.2 στην υπολογιστική γεωμετρία. Ας υποθέσουμε ότι θέλουμε να καλύψουμε ένα φραγμένο υποσύνολο P του $\mathbb{R}^n$ με μπάλες δοθείσης ακτίνας ϵ. Το ερώτημα είναι να εκτιμήσουμε το ελάχιστο πλήθος μπαλών με τις οποίες μπορούμε να το επιτύχουμε. Το θεώρημα που ακολουθεί δίνει ένα άνω φράγμα.

Θεώρημα 2.1.3. *Έστω P ένα πολύτοπο στον $\mathbb{R}^n$ με N κορυφές και διάμετρο φραγμένη από 1. Τότε, μπορούμε να καλύψουμε το P με $N^{\lceil 1/\epsilon^2\rceil}$ Ευκλείδειες μπάλες ακτίνας $\epsilon > 0$.*

Απόδειξη. Ορίζουμε τα κέντρα των μπαλών με τον ακόλουθο τρόπο. Θέτουμε $k = \lceil 1/\epsilon^2\rceil$ και θεωρούμε το σύνολο

$$A = \left\{\frac{1}{k}\sum_{j=1}^{k} x_j : x_j \text{ κορυφή του } P\right\}.$$

Θα δείξουμε ότι η ένωση των ϵ-μπαλών με κέντρα τα σημεία του A ικανοποιεί το συμπέρασμα του θεωρήματος. Πράγματι, θυμηθείτε ότι το πολύτοπο P είναι εξ ορισμού η κυρτή θήκη του συνόλου των κορυφών του, το οποίο συμβολίζουμε με T. Μπορούμε λοιπόν να εφαρμόσουμε το Θεώρημα 2.1.2 για κάθε σημείο $x \in P = \mathrm{conv}(T)$ και να συμπεράνουμε ότι το x απέχει το πολύ $1/\sqrt{k} \leqslant \epsilon$ από κάποιο σημείο του A. Αυτό δείχνει ότι οι ϵ-μπάλες με κέντρα τα σημεία του A καλύπτουν το P.

Τέλος, για να φράξουμε τον πληθάριθμο του A παρατηρούμε ότι υπάρχουν N^k τρόποι για

να επιλέξουμε k από τις N κορυφές του P, επιτρέποντας επαναλήψεις. Συνεπώς,

$$|A| \leqslant N^k = N^{\lceil 1/\epsilon^2 \rceil},$$

και η απόδειξη είναι πλήρης. □

2.2 Ανισότητα Markov

Έστω $(\Omega, \mathcal{F}, \mathbb{P})$ χώρος πιθανότητας και $X : \Omega \to \mathbb{R}$ τυχαία μεταβλητή. Ενδιαφερόμαστε για την οριακή συμπεριφορά των ουρών της τυχαίας μεταβλητής X και ένας τρόπος για να εκτιμήσουμε την πιθανότητα $\mathbb{P}(X \geqslant t)$ είναι μέσω των ροπών της.

Ξεκινώντας από την ανισότητα Markov, αν X είναι μια μη αρνητική τυχαία μεταβλητή με πεπερασμένη μέση τιμή $\mu := \mathbb{E}(X)$ τότε

$$\mathbb{P}(X \geqslant t) \leqslant \frac{\mu}{t} \text{ , για κάθε } t > 0.$$

Πράγματι, για $t > 0$, θεωρούμε το ενδεχόμενο $\{X \geqslant t\} = \{\omega \in \Omega : X(\omega) \geqslant t\}$ και παρατηρούμε ότι ικανοποιείται η ανισότητα $X \geqslant t \cdot \mathbb{1}_{\{X \geqslant t\}}$. Λαμβάνοντας μέσες τιμές βλέπουμε ότι

$$\mathbb{E}(X) \geqslant t \cdot \mathbb{E}(\mathbb{1}_{\{X \geqslant t\}}) \implies \mathbb{E}(X) \geqslant t \cdot \mathbb{P}(\{X \geqslant t\}).$$

Αν, επιπλέον, η τυχαία μεταβλητή X έχει και ροπή δεύτερης τάξης, δηλαδή πεπερασμένη διασπορά, τότε θέτοντας $Y = (X-\mu)^2$ στην ανισότητα Markov, προκύπτει η ανισότητα Chebyshev, ήτοι

$$\mathbb{P}[|X - \mu| \geqslant t] \leqslant \frac{\mathrm{Var}(X)}{t^2} \text{ , για κάθε } t > 0.$$

Η ανισότητα Chebyshev είναι μια απλή μορφή ανισότητας συγκέντρωσης, η οποία εγγυάται ότι η τυχαία μεταβλητή είναι αρκετά κοντά στη μέση τιμή της.

Για τυχαίες μεταβλητές με ροπές ανώτερης τάξης, η παραπάνω εκτίμηση μπορεί να γενικευτεί με τον ίδιο τρόπο. Για παράδειγμα, αν η τυχαία μεταβλητή X έχει ροπή k τάξης τότε θέτοντας $Y = |X - \mu|^k$ στην ανισότητα Markov, παίρνουμε

$$\mathbb{P}(|X - \mu| \geqslant t] \leqslant \frac{\mathbb{E}[|X - \mu|]^k}{t^k} \text{ , για κάθε } t > 0.$$

2.3 Η μέθοδος Cramér-Chernoff

Σε αυτήν την ενότητα αναπτύσσουμε μια χρήσιμη μέθοδο για την εκτίμηση της πιθανότητας $\mathbb{P}(X \geqslant t)$, χρησιμοποιώντας ροπογεννήτριες. Πιο συγκεκριμένα, έστω X τυχαία μεταβητή σε έναν χώρο πιθανότητας $(\Omega, \mathbb{F}, \mathbb{P})$ της οποίας η ροπογεννήτρια υπάρχει σε μια περιοχή γύρω

από το 0. Θέτουμε

$$M_X(t) := \mathbb{E}[e^{tX}] \quad \text{και} \quad \Lambda_X(t) := \log\, M_X(t).$$

Ορισμός 2.3.1 (Μετασχηματισμός Cramér). Έστω $\delta > 0$ και $f : (-\delta, \delta) \to \mathbb{R}$ πραγματική συνάρτηση. Ορίζουμε τον μετασχηματισμό Cramér της f ως $f^* : \mathbb{R} \to \mathbb{R}$ με

$$f^*(\lambda) = \sup_{-\delta<x<\delta} \{x\lambda - f(x)\}.$$

Ειδικότερα, αν X είναι μια τυχαία μεταβλητή και $\delta > 0$ ώστε η $L_X(t) = \log M_X(t)$ να υπάρχει για κάθε $t \in (-\delta, \delta)$ τότε ορίζουμε τον μετασχηματισμό Cramér της Λ_X ως

$$\Lambda_X^*(s) = \sup_{t\in(-\delta,\delta)} \{ts - \Lambda_X(t)\}.$$

Στα επόμενα δύο λήμματα, καταγράφουμε κάποιες βασικές ιδιότητες της συνάρτησης του λογαρίθμου της ροπογεννήτριας Λ_X και του μετασχηματισμού Cramér.

Λήμμα 2.3.2. *Έστω X τυχαία μεταβλητή σε έναν χώρο πιθανότητας $(\Omega, \mathbb{F}, \mathbb{P})$. Τότε:*

(i) *Ο λογάριθμος Λ_X της ροπογεννήτριας M_X είναι κυρτή συνάρτηση .*

(ii) *Αν $a > 0$ και $b \in \mathbb{R}$, τότε $\Lambda_{aX+b}(t) = \Lambda_X(at) + bt$ για κάθε $t \in \mathbb{R}$.*

(iii) *Αν οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες τότε*

$$\Lambda_{\sum_{i=1}^n X_i}(t) = \sum_{i=1}^{n} \Lambda_{X_i}(t) \text{ για κάθε } t \in \mathbb{R}.$$

Απόδειξη. Η απόδειξή των (i) και (ii) είναι άμεση από τους ορισμούς. Από την ανεξαρτησία των τυχαίων μεταβλητών έπεται το (iii). □

Λήμμα 2.3.1. *Έστω X τυχαία μεταβλητή σε έναν χώρο πιθανότητας $(\Omega, \mathbb{F}, \mathbb{P})$. Τότε:*

(i) *Ο μετασχηματισμός Cramér είναι κυρτή συνάρτηση.*

(ii) *Αν για τις τυχαίες μεταβλητές X, Y ισχύει ότι $\Lambda_X \leqslant \Lambda_Y$ σχεδόν βεβαίως, τότε $\Lambda_X^* \geqslant \Lambda_Y^*$.*

(iii) *Αν $a > 0$ και $b \in \mathbb{R}$, τότε $\Lambda_{aX+b}^*(s) = \Lambda_X^*\left(\frac{s-b}{a}\right)$ για κάθε $s \in \mathbb{R}$.*

Απόδειξη. (i) Ο μετασχηματισμός Cramér $\Lambda_X^*(s) = \sup_t\{ts - \Lambda_X(t)\}$, είναι το supremum μιας οικογένειας γραμμικών συναρτήσεων και άρα κυρτή συνάρτηση.

(ii) Παρατηρούμε ότι, για $s \in \mathbb{R}$ μπορούμε να γράψουμε $st - \Lambda_X(t) \geqslant st - \Lambda_Y(t)$, από όπου λαμβάνοντας supremum ως προς $t > 0$ έπεται το ζητούμενο.

(iii) Για $s \in \mathbb{R}$ υπολογίζουμε τον μετασχηματισμό Cramér της τυχαίας μεταβλητής $aX + b$ ως εξής

$$\begin{aligned}\Lambda^*_{aX+b}(s) &= \sup_t\{ts - \Lambda_{aX+b}(t)\}\\ &= \sup_t\{t(s-b) - \Lambda_X(at)\}\\ &= \sup_\lambda\{\lambda\frac{s-b}{a} - \Lambda_X(\lambda)\}\\ &= \Lambda^*_X\left(\frac{s-b}{a}\right).\end{aligned}$$

□

Πρόταση 2.3.3. *Έστω X τυχαία μεταβλητή της οποίας η ροπογεννήτρια υπάρχει σε μια περιοχή γύρω από το 0. Τότε,*

$$\mathbb{P}(X > t) \leqslant e^{-\Lambda^*_X(t)},$$

όπου $\Lambda_X(\lambda) = \log \mathbb{E}[e^{\lambda X}]$ και Λ^*_X είναι ο μετασχηματισμός Cramér της Λ_X.

Απόδειξη. Χρησιμοποιώντας την ανισότητα Markov για την πιθανότητα $\mathbb{P}(X > t)$ υπολογίζουμε

$$\mathbb{P}(X > t) = \mathbb{P}(e^{\lambda X} > e^{\lambda t}) \leqslant \frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda t}},$$

και άρα η προς εκτίμηση πιθανότητα είναι κάτω φράγμα της ποσότητας $\frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda t}}$, συνεπώς

$$\begin{aligned}\mathbb{P}(X > t) &\leqslant \inf\{e^{\log \mathbb{E}[e^{\lambda X}] - \lambda t}\}\\ &= e^{\inf\{\log \mathbb{E}[e^{\lambda X}] - \lambda t\}}\\ &= e^{-\sup\{\lambda t - \log \mathbb{E}[e^{\lambda X}]\}}\\ &= e^{-\Lambda^*_X(t)},\end{aligned}$$

από όπου έπεται το ζητούμενο. □

Παράδειγμα 2.3.4. Έστω $X \sim N(\mu, \sigma^2)$ κανονική τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2. Χρησιμοποιώντας τη συνάρτηση πυκνότητας πιθανότητας της X, υπολογίζουμε τη ροπογεννήτρια της τυχαίας μεταβλητής ως εξής:

$$\begin{aligned}\mathbb{E}(e^{\lambda X}) &= \frac{1}{\sqrt{(2\pi\sigma^2)}} \cdot \int_{-\infty}^{+\infty} e^{\lambda x - \frac{(x-\mu)^2}{2\sigma^2}}\, dx\\ &= \frac{1}{\sqrt{(2\pi\sigma^2)}} \cdot \int_{-\infty}^{+\infty} e^{\frac{\lambda^2\sigma^2}{2} + \lambda\mu} \cdot e^{-\frac{(x-\mu+\lambda\sigma^2)^2}{2\sigma^2}}\, dx\\ &= e^{\frac{\lambda^2\sigma^2}{2} + \lambda\mu},\end{aligned}$$

όπου στη δεύτερη ισότητα εφαρμόσαμε τη μέθοδο συμπλήρωσης τετραγώνων στους εκθέτες.

Αντικαθιστώντας την παραπάνω έκφραση στην ανισότητα του Chernoff βλέπουμε ότι

$$\inf_{\lambda>0}\{\log \mathbb{E}\,(e^{\lambda(X-\mu)}) - \lambda t\} = \inf_{\lambda>0}\left\{\frac{\lambda^2\sigma^2}{2} - \lambda t\right\}.$$

Παρατηρούμε ότι η παραβολή $p(\lambda) := \frac{\sigma^2}{2}\lambda^2 - \lambda t$, παρουσιάζει ελάχιστο στην κορυφή της $\lambda_0 = -\frac{t}{\sigma^2}$, με τιμή $p(\lambda_0) = \frac{-t^2}{2\sigma^2}$, και άρα $\inf_{\lambda>0}\{\frac{\lambda^2\sigma^2}{2} - \lambda t\} = \frac{-t^2}{2\sigma^2}$.

Συνεπώς, για τη συγκέντρωση της τυχαίας μεταβλητής X γύρω από την μέση τιμή της έπεται ότι

$$\mathbb{P}(X - \mu \geqslant t) \leqslant e^{-\frac{t^2}{2\sigma^2}}.$$

2.4 Υποκανονικές τυχαίες μεταβλητές

Ορμώμενοι από το προηγούμενο παράδειγμα θα ορίσουμε μια κλάση κατανομών των οποίων οι ουρές φθίνουν τουλάχιστον τόσο γρήγορα όσο αυτές της κανονικής.

Ορισμός 2.4.1 (υποκανονική τυχαία μεταβλητή). Μια τυχαία μεταβλητή X λέγεται υποκανονική με σταθερά $\alpha > 0$ αν για κάθε $t > 0$ ισχύει

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t^2/\alpha^2}.$$

Τυπικό παράδειγμα υποκανονικής τυχαίας μεταβλητής είναι η $g \sim N(0,1)$. Παρατηρήστε ότι:

(i) Για κάθε $t > 0$,

$$\mathbb{P}(|g| \geqslant t) = \frac{2}{\sqrt{2\pi}}\int_t^\infty e^{-s^2/2}ds \leqslant 2e^{-t^2/2}.$$

(ii) Για κάθε $p \geqslant 1$ έχουμε

$$\begin{aligned}\|g\|_p^p &= \frac{2}{\sqrt{2\pi}}\int_0^\infty s^p e^{-s^2/2}ds \\ &= \frac{2}{\sqrt{2\pi}}\int_0^\infty (2y)^{\frac{p-1}{2}}e^{-y}dy \\ &= \frac{2^{p/2}}{\sqrt{2\pi}}\int_0^\infty y^{\frac{p+1}{2}-1}e^{-y}dy \\ &= \frac{2^{p/2}}{\sqrt{2\pi}}\Gamma\left(\frac{p+1}{2}\right),\end{aligned}$$

και από τον τύπο του Stirling ($\Gamma(x+1) \sim (x/e)^x\sqrt{2\pi x}$ καθώς το $x \to \infty$) παίρνουμε

$$\|g\|_p = \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\right)^{1/p} \leqslant c\sqrt{p},$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

(iii) Για κάθε $\lambda \in \mathbb{R}$ έχουμε

$$\mathbb{E}(e^{\lambda g}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\lambda s - \frac{s^2}{2}} ds = e^{\lambda^2/2}.$$

Η επόμενη πρόταση δείχνει ότι οι υποκανονικές τυχαίες μεταβλητές «χαρακτηρίζονται» από οποιαδήποτε από τις παραπάνω ιδιότητες σε σχέση με τις ουρές, τις ροπές τάξης p και τη συμπεριφορά της ροπογεννήτριας. Ακριβέστερα, οι ιδιότητες αυτές είναι ισοδύναμες.

Πρόταση 2.4.2. *Έστω X μια τυχαία μεταβλητή. Οι παρακάτω ιδιότητες είναι ισοδύναμες (με τις σταθερές α, β, γ να διαφέρουν το πολύ κατά μία απόλυτη σταθερά).*

(i) *Υπάρχει $\alpha > 0$ ώστε: για κάθε $t > 0$,*

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t^2/\alpha^2}.$$

(ii) *Υπάρχει $\beta > 0$ ώστε: για κάθε $p \geqslant 1$,*

$$\|X\|_p = (E|X|^p)^{1/p} \leqslant \beta\sqrt{p}.$$

(iii) *Υπάρχει $\gamma > 0$ ώστε*

$$\mathbb{E}(e^{X^2/\gamma^2}) \leqslant 2.$$

Αν επιπλέον $\mathbb{E}(X) = 0$ τότε οι παραπάνω τρεις ιδιότητες είναι ισοδύναμες με την ακόλουθη:

(iv) *Υπάρχει $\delta > 0$ ώστε: για κάθε $\lambda \in \mathbb{R}$,*

$$\mathbb{E}(e^{\lambda X}) \leqslant e^{\delta^2\lambda^2}.$$

Απόδειξη. (i) $\Longrightarrow$ (ii): Έστω $p \geqslant 1$. Γράφουμε

$$\begin{aligned}
\|X\|_p^p &= \int_0^\infty pt^{p-1}\mathbb{P}(|X| \geqslant t)\, dt \leqslant 2\int_0^\infty pt^{p-1}e^{-t^2/\alpha^2} dt \\
&= \alpha^2 \int_0^\infty p(\alpha^2 s)^{\frac{p}{2}-1} e^{-s} ds = \alpha^p p \int_0^\infty s^{\frac{p}{2}-1} e^{-s} ds = \alpha^p p\Gamma\left(\frac{p}{2}\right) \\
&\leqslant (c\alpha\sqrt{p})^p,
\end{aligned}$$

όπου $c > 0$ είναι μια απόλυτη σταθερά. Δηλαδή, $\|X\|_p \leqslant \beta\sqrt{p}$, με $\beta = c\alpha$.
(ii) $\Longrightarrow$ (iii): Έστω $t > 0$. Γράφουμε

$$\begin{aligned}
\mathbb{E}(e^{t^2X^2}) &= \mathbb{E}\left(1 + \sum_{k=1}^\infty \frac{t^{2k}X^{2k}}{k!}\right) = 1 + \sum_{k=1}^\infty \frac{t^{2k}}{k!}\mathbb{E}(X^{2k}) \\
&\leqslant 1 + \sum_{k=1}^\infty \frac{t^{2k}}{k!}\beta^{2k}(2k)^k \leqslant \sum_{k=0}^\infty (2et^2\beta^2)^k = \frac{1}{1-2et^2\beta^2} \\
&\leqslant e^{4et^2\beta^2} \leqslant \sqrt{e} \leqslant 2,
\end{aligned}$$

αν $t \leqslant \frac{1}{\sqrt{8e}\beta}$, χρησιμοποιώντας στο τέλος την ανισότητα $\frac{1}{1-x} \leqslant e^{2x}$ για $0 \leqslant x \leqslant 1/2$. Επομένως, αν επιλέξουμε $\gamma = \sqrt{8e}\beta$ έχουμε

$$\mathbb{E}(e^{X^2/\gamma^2}) \leqslant 2.$$

(iii) $\Longrightarrow$ (i): Για κάθε $t > 0$, από την ανισότητα Markov, έχουμε

$$\mathbb{P}(|X| \geqslant t) = \mathbb{P}(e^{X^2/\gamma^2} \geqslant e^{t^2/\gamma^2}) \leqslant 2e^{-t^2/\gamma^2}.$$

Δείχνουμε τώρα ότι αν $\mathbb{E}(X) = 0$ τότε (ii) $\Longrightarrow$ (iv). Θα χρησιμοποιήσουμε την ανισότητα $e^x \leqslant x + e^{x^2}$, η οποία ισχύει για κάθε $x \in \mathbb{R}$. Μπορούμε λοιπόν να γράψουμε

$$\mathbb{E}(e^{\lambda X}) \leqslant \mathbb{E}(\lambda X + e^{\lambda^2 X^2}) = \mathbb{E}(e^{\lambda^2 X^2}).$$

Επαναλαμβάνοντας τον υπολογισμό που κάναμε παραπάνω, βλέπουμε ότι

$$\begin{aligned}\mathbb{E}(e^{\lambda^2 X^2}) &= \mathbb{E}\left(1 + \sum_{k=1}^{\infty} \frac{\lambda^{2k} X^{2k}}{k!}\right) = 1 + \sum_{k=1}^{\infty} \frac{\lambda^{2k}}{k!}\mathbb{E}(X^{2k}) \\ &\leqslant 1 + \sum_{k=1}^{\infty} \frac{\lambda^{2k}}{k!}\beta^{2k}(2k)^k \leqslant \sum_{k=0}^{\infty}(2e\lambda^2\beta^2)^k = \frac{1}{1-2e\lambda^2\beta^2} \\ &\leqslant e^{4e\beta^2\lambda^2},\end{aligned}$$

αν $|\lambda| \leqslant 1/(\sqrt{8e}\beta)$. Από την άλλη πλευρά, αν $|\lambda| > 1/(\sqrt{8e}\beta)$ τότε θέτοντας $\gamma = \sqrt{4e}\beta$ από την ανισότητα $2\lambda X \leqslant \lambda^2\gamma^2 + \frac{X^2}{\gamma^2}$ παίρνουμε

$$\begin{aligned}\mathbb{E}(e^{\lambda X}) &\leqslant e^{\gamma^2\lambda^2/2}\mathbb{E}(e^{X^2/2\gamma^2}) \leqslant e^{2e^2\beta^2\lambda^2}\mathbb{E}(e^{X^2/(8e\beta^2)}) \\ &\leqslant e^{2e^2\beta^2\lambda^2}e^{1/2} \leqslant e^{2e^2\beta^2\lambda^2}e^{8e\beta^2\lambda^2} \leqslant e^{5e^2\beta^2\lambda^2}.\end{aligned}$$

Άρα, η (iv) ισχύει με $\delta = \sqrt{5}e\gamma$.

Τέλος αποδεικνύουμε ότι (iv) $\Longrightarrow$ (i). Για κάθε $t > 0$ και για κάθε $\lambda > 0$ έχουμε

$$\mathbb{P}(X \geqslant t) = \mathbb{P}(e^{\lambda X} \geqslant e^{\lambda t}) \leqslant e^{-\lambda t}\mathbb{E}(e^{\lambda X}) \leqslant e^{-\lambda t + \delta^2\lambda^2}.$$

Επιλέγοντας $\lambda = \frac{t^2}{2\delta^2}$ βλέπουμε ότι

$$\mathbb{P}(X \geqslant t) \leqslant e^{-\frac{t^2}{4\delta^2}}.$$

Εφαρμόζοντας το ίδιο επιχείρημα για την $-X$ ολοκληρώνουμε την απόδειξη με $\alpha = 2\delta$. □

Παράδειγμα 2.4.1. Μια τυχαία μεταβλητή X ονομάζεται Rademacher αν λαμβάνει ισοπίθανα τις τιμές $\{-1, 1\}$ δηλαδή

$$\mathbb{P}(X = 1) = \frac{1}{2} \text{ και } \mathbb{P}(X = -1) = \frac{1}{2}.$$

Μια Rademacher τυχαία μεταβλητή είναι υποκανονική, με σταθερά $\alpha = 1$. Πράγματι, χρησιμοποιώντας το ανάπτυγμα της εκθετικής συνάρτησης υπολογίζουμε

$$\begin{aligned}\mathbb{E}(e^{\lambda X}) = \frac{1}{2}\cdot(e^{\lambda}+e^{-\lambda}) &= \frac{1}{2}\cdot\left(\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}+\sum_{k=0}^{\infty}\frac{(-\lambda)^k}{k!}\right)\\ &= \sum_{k=0}^{\infty}\frac{\lambda^{2k}}{(2k)!}\\ &\leqslant 1+\sum_{k=1}^{\infty}\frac{\lambda^{2k}}{2^k k!} = e^{\lambda^2/2},\end{aligned}$$

εφόσον $2^n n! \leqslant (2n)!$ για $n \geqslant 0$.

Παράδειγμα 2.4.2. Μια τυχαία μεταβλητή X ομοιόμορφα κατανεμημένη στο διάστημα $[-\alpha,\alpha]$ είναι υποκανονική, με παράμετρο $\delta = \alpha$. Πράγματι, υπολογίζουμε

$$\begin{aligned}\mathbb{E}(e^{\lambda X}) = \int_{-\alpha}^{\alpha}\frac{1}{2\alpha}e^{\lambda x}dx &= \frac{1}{2\lambda\alpha}\cdot(e^{\lambda\alpha}-e^{-\lambda\alpha})\\ &= \frac{1}{2\lambda\alpha}\cdot\left(\sum_{k=0}^{\infty}\frac{(\lambda\alpha)^k}{k!}-\sum_{k=0}^{\infty}\frac{(-\lambda\alpha)^k}{k!}\right)\\ &= \frac{1}{2\lambda\alpha}\cdot\sum_{k=0}^{\infty}\frac{(\lambda\alpha)^{2k+1}}{(2k+1)!}\\ &= \sum_{k=0}^{\infty}\frac{(\lambda\alpha)^{2k}}{(k+1)!}\leqslant\sum_{k=0}^{\infty}\frac{(\lambda\alpha)^{2k}}{2^k k!} = e^{\frac{\lambda^2\alpha^2}{2}}\end{aligned}$$

εφόσον $2^n n! \leqslant (2n+1)!$ για $n \geqslant 0$.

2.5 Χώροι Orlicz

Ορισμός 2.5.1 (χώροι Orlicz). Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $\psi : \mathbb{R} \to [0,+\infty)$ μια άρτια κυρτή συνάρτηση που ικανοποιεί τα ακόλουθα:

$$\psi(0) = 0 \text{ και } \lim_{x\to\infty}\psi(x) = +\infty.$$

Τότε λέμε ότι η ψ είναι *συνάρτηση Orlicz*. Ο χώρος Orlicz $L_\psi(\mu)$ που αντιστοιχεί στη συνάρτηση ψ αποτελείται από όλες τις $\mathcal{A}$-μετρήσιμες συναρτήσεις f για τις οποίες υπάρχει $\kappa > 0$ τέτοιος ώστε

$$\int_X \psi(f/\kappa)d\mu < \infty.$$

Σταθεροποιώντας μια Orlicz *συνάρτηση* ψ, ορίζουμε την νόρμα Orlicz μιας $f \in L_\psi(\mu)$ ως εξής:

$$\|f\|_{\psi(\mu)} := \inf\left\{\kappa > 0 : \int_X \psi(f/\kappa)d\mu \leqslant 1\right\}.$$

Ισχύει ότι $L_\psi(\mu) \subseteq L_1(\mu)$: αν μια μετρήσιμη συνάρτηση f έχει πεπερασμένη $\psi(\mu)$-νόρμα τότε η f είναι ολοκληρώσιμη ως προς μ. Για να το δούμε αυτό, παρατηρούμε πρώτα ότι, αφού η ψ είναι κυρτή και $\psi(0) = (0)$, η συνάρτηση $t \mapsto \frac{\psi(t)}{t}$ είναι αύξουσα. Έπεται ότι $\psi(t) \geqslant \frac{\psi(t_0)}{t_0} \cdot t$ για κάθε $t > t_0$, όπου t_0 είναι οποιοσδήποτε θετικός αριθμός για τον οποίο $\psi(t_0) > 0$. Τότε, για κάθε $\kappa > \|f\|_{\psi(\mu)}$ μπορούμε να γράψουμε

$$\begin{aligned}\frac{1}{\kappa}\mathbb{E}_\mu(|f|) &= \mathbb{E}_\mu\Big(\frac{|f|}{\kappa} \cdot \mathbb{1}_{\{|f| \leqslant t_0\kappa\}}\Big) + \mathbb{E}_\mu\Big(\frac{|f|}{\kappa} \cdot \mathbb{1}_{\{|f| > t_0\kappa\}}\Big) \\ &\leqslant t_0 + \frac{t_0}{\psi(t_0)}\mathbb{E}_\mu(\psi(|f|/\kappa)) \leqslant t_0 \cdot [1 + (\psi(t_0))^{-1}] < +\infty.\end{aligned}$$

Εφαρμόζοντας την ανισότητα Jensen για την κυρτή συνάρτηση ψ παίρνουμε

$$\psi(\mathbb{E}_\mu(|f|/\kappa)) \leqslant \mathbb{E}_\mu(\psi(|f|/\kappa)) \leqslant 1$$

για κάθε $\kappa > \|f\|_{\psi(\mu)}$. Συνεπώς,

$$\mathbb{E}_\mu(|f|) \leqslant \psi_*^{-1}(1) \cdot \|f\|_{\psi(\mu)},$$

όπου $\psi_*^{-1}(1) = \inf\{s > 0 : \psi(t) > 1 \text{ για κάθε } t \geqslant s\}$.

Ιδιαίτερα σημαντική είναι η οικογένεια των ψ_α-νορμών, η οποία είναι υποοικογένεια των νορμών Orlicz.

Ορισμός 2.5.2 (ψ_α-νόρμα). Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και έστω $f : X \to \mathbb{R}$ μια $\mathcal{A}$-μετρήσιμη συνάρτηση. Για κάθε $\alpha \in [1, 2]$ ορίζουμε την ψ_α-νόρμα της f ως εξής:

$$\|f\|_{\psi_\alpha} := \inf\left\{t > 0 : \int_X \exp\left(\frac{|f|}{t}\right)^\alpha d\mu \leqslant 2\right\},$$

αρκεί το σύνολο στο δεξιό μέλος να είναι μη κενό. Παρατηρήστε ότι η ψ_α-νόρμα είναι ακριβώς η νόρμα Orlicz που αντιστοιχεί στην κυρτή συνάρτηση $t \in \mathbb{R} \to e^{|t|^\alpha} - 1$.

Με βάση τον ορισμό της ψ_2-νόρμας, μπορούμε να επαναδιατυπώσουμε τα συμπεράσματα της Πρότασης 2.4.2 ως εξής:

Πρόταση 2.5.3. *Έστω X μια τυχαία μεταβλητή. Αν $X \in L_{\psi_2}$ τότε:*

(i) *Για κάθε $t > 0$,*

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-c_1 t^2/\|X\|_{\psi_2}^2}$$

όπου $c_1 > 0$ είναι μια απόλυτη σταθερά.

(ii) *Για κάθε $p \geqslant 1$,*

$$\|X\|_p \leqslant c_2 \sqrt{p}\|X\|_{\psi_2}$$

όπου $c_2 > 0$ είναι μια απόλυτη σταθερά.

(iii) *Ισχύει η ανισότητα*

$$\mathbb{E}(e^{X^2/\|X\|_{\psi_2}^2}) \leqslant 2.$$

(iv) *Αν επιπλέον* $\mathbb{E}(X) = 0$ *τότε, για κάθε* $\lambda \in \mathbb{R}$,

$$\mathbb{E}(e^{\lambda X}) \leqslant e^{c_3 \|X\|_{\psi_2}^2 \lambda^2}$$

όπου $c_3 > 0$ *είναι μια απόλυτη σταθερά.*

Τυπικά παραδείγματα υποκανονικών τυχαίων μεταβλητών είναι τα παρακάτω:

(i) *Τυπικές κανονικές τυχαίες μεταβλητές.* Αν $X \sim N(0,1)$ έχουμε ήδη δει ότι $\|X\|_{\psi_2} \leqslant C$. Γενικότερα, αν $X \sim N(0,\sigma^2)$ τότε $\|X\|_{\psi_2} \leqslant C\sigma$.

(ii) *Συμμετρικές Bernoulli τυχαίες μεταβλητές.* Στην περίπτωση αυτή έχουμε $|X| = 1$, άρα $\|X\|_{\psi_2} = \frac{1}{\sqrt{\ln 2}}$.

(iii) *Φραγμένες τυχαίες μεταβλητές.* Αν $\|X\|_\infty < \infty$, από τον ορισμό της ψ_2-νόρμας είναι φανερό ότι $\|X\|_{\psi_2} \leqslant C\|X\|_\infty$, όπου $C = \frac{1}{\sqrt{\ln 2}}$.

Κλείνουμε αυτήν την ενότητα με μια χρήσιμη εκτίμηση για τη μέση τιμή του maximum πεπερασμένων το πλήθος υποκανονικών (ψ_2) τυχαίων μεταβλητών.

Πρόταση 2.5.4. *Έστω* $X_1, \dots, X_N$, $N \geqslant 2$, *υποκανονικές τυχαίες μεταβλητές τέτοιες ώστε*

$$\|X_i\|_{\psi_2} \leqslant b$$

για κάθε $i = 1, \dots, N$. *Τότε,*

$$\mathbb{E} \max_{1 \leqslant i \leqslant N} |X_i| \leqslant Cb\sqrt{\log N},$$

όπου $C > 0$ *είναι μια απόλυτη σταθερά.*

Απόδειξη. Από τον ορισμό της ψ_2-νόρμας και την ανισότητα Markov, για κάθε $t > 0$ έχουμε

$$\mathbb{P}\left(\max_{1 \leqslant i \leqslant N} |X_i| \geqslant t \right) \leqslant \sum_{i=1}^N \mathbb{P}\left(|X_i| \geqslant t \right) \leqslant 2Ne^{-t^2/b^2}.$$

Συνεπώς, για κάθε $\alpha > 0$ μπορούμε να γράψουμε

$$\begin{aligned} \mathbb{E} \max_{1 \leqslant i \leqslant N} |X_i| &= \int_0^\infty \mathbb{P}\left(\max_{1 \leqslant i \leqslant N} |X_i| \geqslant t \right) dt \\ &\leqslant \alpha + \int_\alpha^\infty \mathbb{P}\left(\max_{1 \leqslant i \leqslant N} |X_i| \geqslant t \right) dt \\ &\leqslant \alpha + 2N \int_\alpha^\infty e^{-t^2/b^2} dt. \end{aligned}$$

Αν επιλέξουμε $\alpha = 2b\sqrt{\log N}$ παίρνουμε

$$\begin{aligned}\int_\alpha^\infty e^{-t^2/b^2}dt &= 2b\sqrt{\log N}\int_1^\infty \exp(-4s^2\log N)ds \leqslant 2b\sqrt{\log N}\int_1^\infty \exp(-4s\log N)ds\\ &\leqslant 2b\sqrt{\log N}\exp(-2\log N)\int_1^\infty e^{-s}ds\\ &\leqslant 2bN^{-2}\sqrt{\log N},\end{aligned}$$

όπου χρησιμοποιήσαμε την

$$\exp(-4s\log N) \leqslant \exp(-2\log N)\cdot e^{-s}$$

που ισχύει για κάθε $s \geqslant 1$. Έπεται ότι

$$\mathbb{E}\max_{1\leqslant i\leqslant N}|X_i| \leqslant Cb\sqrt{\log N}$$

με $C = 4$. □

Η επόμενη πρόταση δείχνει ότι το φράγμα της Πρότασης 2.5.4 είναι βέλτιστο. Αν οι X_i είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές τότε ισχύει και η αντίστροφη ανισότητα.

Πρόταση 2.5.5. *Έστω $N \geqslant 2$ και $g_1,\dots,g_N$ ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές. Τότε,*

$$\mathbb{E}\max_{1\leqslant j\leqslant N}|g_j| \geqslant c\sqrt{\log N},$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Πρώτα, παρατηρούμε ότι η $\mathbb{E}(\max\{g_1,g_2\})$ υπολογίζεται και είναι ίση με $1/\sqrt{\pi}$. Μπορούμε λοιπόν να υποθέσουμε ότι ο N είναι αρκετά μεγάλος (ας πούμε μεγαλύτερος από 10^4). Σε αυτή την περίπτωση, παρατηρούμε ότι

$$2\mathbb{E}\left(\max_{1\leqslant k\leqslant N} g_k\right) \geqslant \mathbb{E}\left(\max_{1\leqslant k\leqslant N}|g_k|\right) - \mathbb{E}|g_1|.$$

Πράγματι,

$$\mathbb{E}\left(\max_{1\leqslant k\leqslant N} g_k\right) = \mathbb{E}\left(\max_{1\leqslant k\leqslant N}(g_k - g_1)\right) = \mathbb{E}\left(\max_{1\leqslant k\leqslant N}\max\{g_k - g_1, 0\}\right),$$

και, λόγω της συμμετρίας των g_i,

$$\begin{aligned}\mathbb{E}\left(\max_{1\leqslant k\leqslant N}|g_k - g_1|\right) &\leqslant \mathbb{E}\left(\max_{1\leqslant k\leqslant N}\max\{g_k - g_1, 0\}\right) + \mathbb{E}\left(\max_{1\leqslant k\leqslant N}\max\{g_1 - g_k, 0\}\right)\\ &= 2\mathbb{E}\left(\max_{1\leqslant k\leqslant N}(g_k - g_1)\right) = 2\mathbb{E}\left(\max_{1\leqslant k\leqslant N} g_k\right).\end{aligned}$$

Όμως, από την ανεξαρτησία των g_i,

$$\mathbb{E}\left(\max_{1\leqslant k\leqslant M} g_k\right) = \int_0^\infty \mathbb{P}\left(\max_{1\leqslant k\leqslant M} |g_k| > t\right) dt = \int_0^\infty \left(1 - \mathbb{P}\left(\max_{1\leqslant k\leqslant M} |g_k| \leqslant t\right)\right) dt$$
$$= \int_0^\infty \left(1 - \left(1 - \sqrt{2/\pi}\int_t^\infty e^{-u^2/2}du\right)^M\right) dt.$$

Είναι εύκολο να δούμε ότι, για κάθε $t > 0$,

$$\int_t^\infty e^{-u^2/2}du \geqslant e^{-(t+1)^2/2}.$$

Θέτοντας $t_0 + 1 = \sqrt{2\log M}$, έχουμε

$$\mathbb{E}\left(\max_{1\leqslant k\leqslant M} g_k\right) \geqslant \int_0^{t_0} \left(1 - \left(1 - \sqrt{2/\pi}\int_t^\infty e^{-u^2/2}du\right)^M\right) dt$$
$$\geqslant t_0\left(1 - \left(1 - \frac{\sqrt{2}}{M\sqrt{\pi}}\right)^M\right) \geqslant t_0(1 - e^{-\sqrt{2/\pi}}),$$

το οποίο ολοκληρώνει την απόδειξη. □

2.6 Υποεκθετικές τυχαίες μεταβλητές

Η κλάση των υποκανονικών κατανομών είναι αρκετά ευρεία και περιγράφει ικανοποιητικά τις ουρές της κανονικής, της ομοιόμορφης αλλά και γενικότερα κάθε κατανομής με φραγμένο φορέα. Ωστόσο, αφήνει έξω ορισμένες σημαντικές κατανομές των οποίων οι ουρές είναι πιο βαριές από της κανονικής.

Για παράδειγμα, ας θεωρήσουμε ένα τυχαίο διάνυσμα $X \in \mathbb{R}^N$ του οποίου οι συντεταγμένες είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές. Σε πολλές εφαρμογές ενδιαφερόμαστε για τη συγκέντρωση της τυχαίας μεταβλητής

$$\|X\|_2^2 = \sum_{i=1}^N X_i^2.$$

Παρατηρούμε ότι αυτή η τυχαία μεταβλητή είναι άθροισμα των ανεξάρτητων τυχαίων μεταβλητών X_i^2 οι οποίες όμως δεν ειναι υποκανονικές. Πράγματι,

$$\mathbb{P}(X_i^2 > t) = \mathbb{P}(|X_i| > \sqrt{t}) \approx e^{-t/2},$$

εφόσον οι X_i έχουν υποτεθεί υποκανονικές.

Ορισμός 2.6.1 (υποεκθετική τυχαία μεταβλητή). Μια τυχαία μεταβλητή X λέγεται υποεκθε-

τική με σταθερά $\alpha > 0$ αν για κάθε $t > 0$ ισχύει

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t/\alpha}.$$

Η επόμενη πρόταση δίνει χαρακτηρισμούς σε σχέση με τη συμπεριφορά των ουρών, των p-ροπών και της ροπογεννήτριας των υποεκθετικών τυχαίων μεταβλητών. Η απόδειξη είναι ανάλογη με αυτήν της Πρότασης 2.4.2.

Πρόταση 2.6.2. *Έστω X μια τυχαία μεταβλητή. Οι παρακάτω ιδιότητες είναι ισοδύναμες (με τις σταθερές α, β, γ να «διαφέρουν» το πολύ κατά μία απόλυτη σταθερά).*

(i) *Υπάρχει $\alpha > 0$ ώστε: για κάθε $t > 0$,*

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t/\alpha}.$$

(ii) *Υπάρχει $\beta > 0$ ώστε: για κάθε $p \geqslant 1$,*

$$\|X\|_p \leqslant \beta p.$$

(iii) *Υπάρχει $\gamma > 0$ ώστε*

$$\mathbb{E}(e^{|X|/\gamma}) \leqslant 2.$$

Αν επιπλέον $\mathbb{E}(X) = 0$ τότε οι παραπάνω τρεις ιδιότητες είναι ισοδύναμες με την ακόλουθη:

(iv) *Υπάρχει $\delta > 0$ ώστε: για κάθε $0 < \lambda < \frac{1}{\sqrt{\delta}}$,*

$$\mathbb{E}(e^{\lambda X}) \leqslant e^{\delta\lambda^2}.$$

Απόδειξη. (i) $\Longrightarrow$ (ii): Για $p \geqslant 1$ έχουμε

$$\begin{aligned}
\mathbb{E}|X|^p &= p\int_0^{+\infty} t^{p-1}\cdot \mathbb{P}(|X| > t)dt \leqslant 2p\int_0^{+\infty} t^{p-1}e^{-\frac{t}{\alpha}}dt \\
&= 2p\alpha^p \int_0^{+\infty} s^{p-1}\cdot e^{-s}ds \\
&= 2p\alpha^p \cdot \Gamma(p) = 2\alpha^p \cdot \Gamma(p+1) \leqslant 4\alpha^p \cdot p^p,
\end{aligned}$$

χρησιμοποιώντας την $\Gamma(p+1) \leqslant 2 \cdot p^p$. Άρα, έχουμε $\|X\|_p \leqslant \beta \cdot p$, με $\beta \leqslant 4\alpha$.

(ii) $\Longrightarrow$ (iii): Έστω $\lambda > 0$. Από το ανάπτυγμα Taylor υπολογίζουμε

$$\begin{aligned}
\mathbb{E}e^{\lambda|X|} &= 1 + \sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot \mathbb{E}|X|^k \leqslant 1 + \sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot (\beta k)^k \\
&\leqslant 1 + \sum_{k=1}^{\infty} (\lambda\beta e)^k \leqslant 1 + \frac{\lambda\beta e}{1-\lambda\beta e} \leqslant 1 + 2\lambda\beta e \leqslant e^{2\lambda\beta e},
\end{aligned}$$

για $0 < \lambda < \frac{1}{2\beta e}$, και χρησιμοποιώντας την ανισότητα $k! \geqslant (\frac{k}{e})^k$, για $k \geqslant 1$. Άρα επιλέγοντας $\gamma \leqslant \frac{1}{2\beta e}$ έχουμε ότι $\mathbb{E}(e^{|X|/2\gamma}) < e^{1/2} < 2$ και το ζητούμενο έπεται.

(ii) $\Longrightarrow$ (iii): Για $t > 0$, υπολογίζουμε

$$\mathbb{P}(|X| > t) = \mathbb{P}(e^{X/\gamma} > e^{t/\gamma}) \leqslant 2e^{-t/2\gamma}$$

από την ανισότητα Markov. Το ζητούμενο έπεται με $\alpha \leqslant \gamma$.

Τέλος, αν $\mathbb{E}(X) = 0$ αποδεικνύουμε ότι (ii) $\Longrightarrow$ (iii) επαναλαμβάνοντας τον υπολογισμό στο ανάπτυγμα Taylor ως εξής:

$$\begin{aligned}\mathbb{E}(e^{\lambda X}) &= 1 + \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot \mathbb{E}(X^k) = 1 + \sum_{k=2}^{\infty} \frac{\lambda^k}{k!} \cdot \mathbb{E}(X^k) \\ &\leqslant 1 + \sum_{k=2}^{\infty} (\lambda\beta e)^k \leqslant 1 + 2\lambda^2\beta^2 e^2\end{aligned}$$

για $\lambda \leqslant \frac{1}{2\beta e.}$ Άρα, επιλέγοντας $\delta \leqslant 2\beta^2 e^2$, έχουμε ότι $\mathbb{E}(e^{\lambda X}) \leqslant e^{\delta\lambda^2}$. Χρησιμοποιώντας και πάλι την ανισότητα Markov δείχνουμε ότι (iv) $\Longrightarrow$ (i). □

Με βάση τον ορισμό της ψ_1-νόρμας, μπορούμε να επαναδιατυπώσουμε τα συμπεράσματα της Πρότασης 2.6.2 ως εξής:

Πρόταση 2.6.3. *Έστω X μια τυχαία μεταβλητή. Αν $X \in L_{\psi_1}$ τότε:*

(i) *Για κάθε $t > 0$,*

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-c_1 t/\|X\|_{\psi_1}}$$

όπου $c_1 > 0$ είναι μια απόλυτη σταθερά.

(ii) *Για κάθε $p \geqslant 1$,*

$$\|X\|_p \leqslant c_2 p \|X\|_{\psi_1}$$

όπου $c_2 > 0$ είναι μια απόλυτη σταθερά.

(iii) *Ισχύει η ανισότητα*

$$\mathbb{E}(e^{|X|/\|X\|_{\psi_1}}) \leqslant 2.$$

Τα επόμενα δύο λήμματα δείχνουν ότι το γινόμενο δύο υποκανονικών τυχαίων μεταβλητών είναι υποεκθετική τυχαία μεταβλητή. Ειδικότερα, το τετράγωνο μιας υποκανονικής τυχαίας μεταβλητής είναι υποεκθετική τυχαία μεταβλητή.

Λήμμα 2.6.4. *Μια τυχαία μεταβλητή X είναι υποκανονική αν και μόνο αν η X^2 είναι υποεκθετική. Επιπλέον,*

$$\|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2.$$

Απόδειξη. Αρκεί να παρατηρήσουμε ότι η $\|X^2\|_{\psi_1}$ είναι το infimum των $\alpha > 0$ για τους οποίους $\mathbb{E}\left(\exp(X^2/\alpha)\right) \leqslant 2$ ενώ η $\|X\|_{\psi_2}$ είναι το infimum των $\beta > 0$ για τους οποίους $\mathbb{E}\left(\exp(X^2/\beta^2)\right) \leqslant 2$. Οι δύο ορισμοί συμπίπτουν αν θέσουμε $\alpha = \beta^2$. □

Λήμμα 2.6.5. *Έστω X και Y υποκανονικές τυχαίες μεταβλητές. Τότε, η XY είναι υποεκθετική και*

$$\|XY\|_{\psi_1} \leqslant \|X\|_{\psi_2}\|Y\|_{\psi_2}.$$

Απόδειξη. Θέτουμε $\alpha = \|X\|_{\psi_2}$ και $\beta = \|Y\|_{\psi_2}$. Παρατηρούμε ότι

$$\begin{aligned}
\mathbb{E}\left(e^{\frac{|XY|}{\alpha\beta}}\right) &\leqslant \mathbb{E}\left(e^{\frac{X^2}{2\alpha^2}+\frac{Y^2}{2\beta^2}}\right) \\
&= \mathbb{E}\left(e^{\frac{X^2}{2\alpha^2}} e^{\frac{Y^2}{2\beta^2}}\right) \\
&\leqslant \frac{1}{2}\mathbb{E}\left(e^{\frac{X^2}{\alpha^2}} + e^{\frac{Y^2}{\beta^2}}\right) \\
&= \frac{1}{2}\left(2+2\right) = 2,
\end{aligned}$$

χρησιμοποιώντας δύο φορές την $ab \leqslant \frac{1}{2}(a^2+b^2)$. Αυτό αποδεικνύει ότι $\|XY\|_{\psi_1} \leqslant \alpha\beta$. □

Η επόμενη πρόταση δείχνει ότι η ροπογεννήτρια μιας υποεκθετικής τυχαίας μεταβλητής είναι πεπερασμένη σε μια περιοχή του 0.

Πρόταση 2.6.6. *Έστω X μια υποεκθετική τυχαία μεταβλητή με $\mathbb{E}(X) = 0$. Αν $|\lambda| \leqslant c_1/\|X\|_{\psi_1}$, όπου $c_1 > 0$ είναι μια απόλυτη σταθερά, τότε*

$$\mathbb{E}\left(e^{\lambda X}\right) \leqslant e^{c_2\lambda^2\|X\|_{\psi_1}^2},$$

όπου $c_2 > 0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $\|X\|_{\psi_1} = 1$. Γράφουμε

$$\mathbb{E}\left(e^{\lambda X}\right) = \mathbb{E}\left(1 + \lambda X + \sum_{k=2}^{\infty}\frac{(\lambda X)^k}{k!}\right) = 1 + \sum_{k=2}^{\infty}\frac{\lambda^k\mathbb{E}(X^k)}{k!},$$

χρησιμοποιώντας την υπόθεση ότι $\mathbb{E}(X) = 0$. Γνωρίζουμε ότι $\mathbb{E}(X^k) \leqslant (c_3 k)^k$, από την Πρόταση 2.6.3. Παίρνοντας υπ' όψιν μας και την $k! \geqslant (k/e)^k$ βλέπουμε ότι

$$\mathbb{E}\left(e^{\lambda X}\right) \leqslant 1 + \sum_{k=2}^{\infty}\frac{(c_3 k\lambda)^k}{(k/e)^k} = 1 + \sum_{k=2}^{\infty}(c_4\lambda)^k.$$

Αν $|c_4\lambda| < \frac{1}{2}$ τότε η γεωμετρική αυτή σειρά συγκλίνει και έχουμε

$$\mathbb{E}\left(e^{\lambda X}\right) \leqslant 1 + 2(c_4\lambda)^2 \leqslant e^{2(c_4\lambda)^2}.$$

Αυτό αποδεικνύει το ζητούμενο. □

2.7 Ανισότητες τύπου Bernstein

Σε αυτήν την ενότητα συζητάμε ανισότητες τύπου Bernstein για αθροίσματα ανεξάρτητων τυχαίων μεταβλητών $X_1, \dots, X_N$, αρχίζοντας από την υποεκθετική περίπτωση.

Θεώρημα 2.7.1. *Έστω $X_1, \dots, X_N$ ανεξάρτητες υποεκθετικές τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Τότε, για κάθε $t > 0$,*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^N X_i\Big| \geqslant t\Big\}\Big) \leqslant 2\exp\Big(-c\min\Big\{\frac{t^2}{\sum_{i=1}^N \|X_i\|_{\psi_1}^2}, \frac{t}{\max_{1\leqslant i\leqslant N}\|X_i\|_{\psi_1}}\Big\}\Big),$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Θέτουμε $S_N = X_1 + \cdots + X_N$. Από την ανισότητα Markov, για κάθε $\lambda > 0$ έχουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant e^{-\lambda t}\prod_{i=1}^N \mathbb{E}\left(e^{\lambda X_i}\right). \tag{2.2}$$

Αν επιλέξουμε

$$|\lambda| \leqslant \frac{c_1}{\max_{1\leqslant i\leqslant N}\|X_i\|_{\psi_1}}, \tag{2.3}$$

από την Πρόταση 2.6.6 έχουμε

$$\mathbb{E}\left(e^{\lambda X_i}\right) \leqslant e^{c_2\lambda^2\|X_i\|_{\psi_1}^2}, \qquad i = 1, \dots, N.$$

Αντικαθιστώντας στην (2.2) παίρνουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant \exp(-\lambda t + C^2\lambda^2\sigma^2),$$

όπου

$$\sigma^2 = \sum_{i=1}^N \|X_i\|_{\psi_1}^2.$$

Τώρα, ελαχιστοποιούμε αυτή την ποσότητα ως προς λ, υπό τον περιορισμό (2.3). Η βέλτιστη επιλογή είναι

$$\lambda = \min\left\{\frac{t}{2c_2\sigma^2}, \frac{c_1}{\max_{1\leqslant i\leqslant N}\|X_i\|_{\psi_1}}\right\},$$

η οποία μας δίνει

$$\mathbb{P}(S_N \geqslant t) \leqslant \exp\Big(-\min\Big\{\frac{t^2}{4c_2\sum_{i=1}^N \|X_i\|_{\psi_1}^2}, \frac{c_1 t}{2\max_{1\leqslant i\leqslant N}\|X_i\|_{\psi_1}}\Big\}\Big).$$

Επαναλαμβάνοντας το ίδιο επιχείρημα με τις $-X_i$ στη θέση των X_i παίρνουμε το ίδιο φράγμα για την $\mathbb{P}(-S_n \geqslant t)$. Συνδυάζοντας τα δύο φράγματα έχουμε το θεώρημα. □

Εφαρμόζοντας το προηγούμενο θεώρημα για τις $a_i X_i$ αντί των X_i έχουμε το εξής.

Θεώρημα 2.7.2. *Έστω $X_1, \dots, X_N$ ανεξάρτητες υποεκθετικές τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Υποθέτουμε ότι, για κάποιον $M > 0$, ισχύει $\|X_i\|_{L_{\psi_1}} \leqslant M$ για κάθε $i = 1, \dots, N$. Τότε, για κάθε $a = (a_1, \dots, a_N) \in \mathbb{R}^N$ και για κάθε $t > 0$,*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^N a_i X_i\Big| \geqslant t\Big\}\Big) \leqslant 2\exp\Big(-c\min\Big\{\frac{t^2}{M^2|a|^2}, \frac{t}{M\|a\|_\infty}\Big\}\Big),$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Μπορούμε επίσης να διατυπώσουμε την ανισότητα Bernstein ως ποσοτική μορφή του νόμου των μεγάλων αριθμών.

Θεώρημα 2.7.3. *Έστω $X_1, \dots, X_N$ ανεξάρτητες υποεκθετικές τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Υποθέτουμε ότι, για κάποιον $M > 0$, ισχύει $\|X_i\|_{L_{\psi_1}} \leqslant M$ για κάθε $i = 1, \dots, N$. Τότε, για κάθε $t > 0$,*

$$\mathbb{P}\Big(\Big\{\Big|\frac{1}{N}\sum_{i=1}^N X_i\Big| \geqslant t\Big\}\Big) \leqslant 2\exp\Big(-cN\,\min\Big\{\frac{t^2}{M^2}, \frac{t}{M}\Big\}\Big),$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Δίνουμε δύο ακόμα παραδείγματα ανισοτήτων τύπου Bernstein, με διαφορετικές υποθέσεις για τις X_i.

Θεώρημα 2.7.4. *Έστω $X_1, \dots, X_N$ ανεξάρτητες τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Υποθέτουμε ότι για κάποιον $M > 0$ ισχύει $\|X_i\|_\infty \leqslant M$ για κάθε $i = 1, \dots, N$. Θέτουμε $\sigma^2 = \frac{1}{N}\sum_{i=1}^N \mathbb{E}(X_i^2)$. Τότε, για κάθε $t > 0$,*

$$\mathbb{P}\Big(\Big\{\sum_{j=1}^N X_j \geqslant tN\Big\}\Big) \leqslant \exp\Big(-\frac{\sigma^2 N}{M^2} F\Big(\frac{Mt}{\sigma^2}\Big)\Big),$$

όπου $F(u) = (1+u)\log(1+u) - u$, $u > 0$.

Απόδειξη. Έστω $t > 0$. Από την ανισότητα Markov και από την ανεξαρτησία των X_i έχουμε

$$\begin{aligned}
\mathbb{P}\Big(\sum_{j=1}^N X_j \geqslant tN\Big) &= \mathbb{P}\Big(\Big\{\exp\Big(\frac{\lambda}{N}\sum_{j=1}^N X_j\Big) \geqslant e^{\lambda t}\Big\}\Big) \\
&\leqslant e^{-\lambda t}\mathbb{E}\,\exp\Big(\frac{\lambda}{N}\sum_{j=1}^N X_j\Big) \\
&= e^{-\lambda t}\prod_{j=1}^N \mathbb{E}\,\exp(\lambda X_j/N)
\end{aligned}$$

για κάθε $\lambda > 0$. Στη συνέχεια παρατηρούμε ότι, αφού $\mathbb{E}(X_j) = 0$,

$$\begin{aligned}\mathbb{E}\,\exp(\lambda X_j/N) &= 1 + \sum_{k=2}^{\infty} \frac{\lambda^k \mathbb{E}(X_j^k)}{N^k k!} \leqslant 1 + \mathbb{E}(X_j^2) \sum_{k=2}^{\infty} \frac{\lambda^k M^{k-2}}{N^k k!} \\ &= 1 + \frac{\mathbb{E}(X_j^2)}{M^2}\Big(e^{\frac{\lambda M}{N}} - \frac{\lambda M}{N} - 1\Big).\end{aligned}$$

Αφού $e^u \geqslant 1 + u$ για κάθε $u \in \mathbb{R}$, παίρνουμε

$$\prod_{j=1}^{N} \mathbb{E}\,\exp(\lambda X_j/N) \leqslant \exp\Big(\frac{\sum_{j=1}^{N} \mathbb{E}(X_j)^2}{M^2}\Big(e^{\frac{\lambda M}{N}} - \frac{\lambda M}{N} - 1\Big)\Big).$$

Από τον ορισμό του σ^2 έπεται ότι

$$\mathbb{P}\Big(\sum_{j=1}^{N} X_j \geqslant tN\Big) \leqslant \exp\Big(\frac{\sigma^2 N}{M^2}\Big(e^{\frac{\lambda M}{N}} - \frac{\lambda M}{N} - 1\Big) - \lambda t\Big)$$

για κάθε $\lambda > 0$. Επιλέγουμε το λ έτσι ώστε να ικανοποιείται η $\exp(\lambda M/N) = 1 + tM/\sigma^2$ και έπεται το ζητούμενο. □

Στο επόμενο θεώρημα υποθέτουμε ότι έχουμε κάποιο ομοιόμορφο φράγμα για την L_1 και την L_∞-νόρμα των X_i.

Θεώρημα 2.7.5. *Έστω $X_1, \dots, X_N$ ανεξάρτητες τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Υποθέτουμε ότι οι X_i είναι φραγμένες και ότι $\mathbb{E}|X_i| \leqslant 2$ και $\|X_i\|_\infty \leqslant M$ για κάθε $i = 1, \dots, N$ και κάποια σταθερά $M > 0$. Τότε, για κάθε $0 < t < 1$,*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^{N} X_i\Big| \geqslant tN\Big\}\Big) \leqslant 2\exp(-t^2 N/(8M)).$$

Απόδειξη. Χρησιμοποιώντας την $e^x \leqslant 1 + x + x^2$, που ισχύει για $0 \leqslant x \leqslant 1$, μαζί με τις υποθέσεις $\mathbb{E}(X_i) = 0$ και $|X_i| \leqslant M$, βλέπουμε ότι αν $0 < \lambda \leqslant 1/M$ τότε

$$\mathbb{E}\exp(\lambda X_i) \leqslant 1 + \lambda^2 \mathbb{E}(X_i^2) \leqslant 1 + \lambda^2 \|X_i\|_1 \|X_i\|_\infty \leqslant \exp(2\lambda^2 M).$$

Από την ανεξαρτησία των X_i έπεται ότι

$$\mathbb{E}\exp(\sum_{i=1}^{N} \lambda X_i) = \prod_{i=1}^{N} \mathbb{E}\exp(\lambda X_i) \leqslant \exp(2\lambda^2 MN).$$

Τέλος, επιλέγοντας $\lambda = t/(4M)$ παίρνουμε

$$\begin{aligned}\mathbb{P}\Big(\Big\{\sum_{i=1}^N X_i > tN\Big\}\Big) &\leqslant \exp(-\lambda tN)\mathbb{E}\exp\Big(\lambda\sum_{i=1}^N X_i\Big)\\ &\leqslant \exp(-\lambda tN)\exp(2\lambda^2 MN)\\ &= \exp(2\lambda^2 MN - \lambda tN) = \exp(-t^2N/(8M)).\end{aligned}$$

Αντικαθιστώντας τις X_i με τις $-X_i$ στο προηγούμενο επιχείρημα, παίρνουμε τελικά το ζητούμενο. □

2.8 Ανισότητα Sudakov και η δυϊκή της

Αν U και V είναι δύο φραγμένα υποσύνολα του $\mathbb{R}^n$ και το V έχει μη κενό εσωτερικό, ορίζουμε τον αριθμό κάλυψης $N(U,V)$ του U από το V να είναι το ελάχιστο πλήθος μεταφορών του V που χρειάζονται για να καλύψουμε το U. Τυπικά,

$$N(U,V) = \inf\left\{N : \exists\, x_1,\dots,x_N \in \mathbb{R}^n, U \subseteq \bigcup_{i=1}^N (x_i + V)\right\}.$$

Αν, επιπλέον, το V είναι συμμετρικό κυρτό σώμα, ο αριθμός στοίχισης $M(U,V)$ είναι το μέγιστο πλήθος σημείων του U τα οποία είναι 1-διαχωρισμένα ως προς τη νόρμα που επάγεται από το κυρτό σώμα V. Τυπικά, για κάθε ζεύγος συνόλων $U, V \subset \mathbb{R}^n$,

$$M(U,V) = \sup\left\{N : \exists\, x_1,\dots,x_N \in U, \forall\, i \neq j, x_i - x_j \notin V\right\}.$$

Αν το V είναι συμμετρικό κυρτό σώμα (γράφοντας συμμετρικό εννοούμε πάντα ως προς την αρχή των αξόνων), μπορούμε να ορίσουμε τη νόρμα που επάγεται από το V: για κάθε $x \in \mathbb{R}^n$,

$$\|x\|_V = \inf\{t > 0 : x \in tV\}.$$

Άρα, η σχέση $x_i - x_j \notin V$ είναι ισοδύναμη με την $\|x_i - x_j\|_V > 1$. Για κάθε θετικό αριθμό ϵ χρησιμοποιούμε επίσης τον συμβολισμό

$$N(U, \epsilon, \|\cdot\|_V)$$

αντί του $N(U, \epsilon V)$. Λέμε επίσης ότι η οικογένεια $x_1,\dots,x_N$ είναι ένα ϵ-δίκτυο αν $U \subseteq \bigcup_{i=1}^N (x_i + \epsilon V)$. Τέλος ορίζουμε το πολικό του V ως εξής:

$$V^\circ = \{y \in \mathbb{R}^n : \forall\, x \in V, \langle x, y\rangle \leqslant 1\}.$$

Δηλαδή, ο δυϊκός του χώρου με νόρμα $(\mathbb{R}^n, \|\cdot\|_V)$ είναι ισομετρικά ισόμορφος με τον $(\mathbb{R}^n, \|\cdot\|_{V^\circ})$.

Στην περίπτωση που το V είναι συμμετρικό κυρτό σώμα, οι έννοιες των αριθμών κάλυψης

και αριθμών στοίχισης συνδέονται στενά.

Πρόταση 2.8.1. (α) *Αν $U, V \subset \mathbb{R}^n$ και $0 \in V$, τότε $N(U,V) \leqslant M(U,V)$.*

(β) *Αν το U είναι κυρτό σώμα και το V είναι συμμετρικό κυρτό σώμα, τότε $M(U,V) \leqslant N(U,V/2)$.*

Απόδειξη. (α) Έστω $N = M(U,V)$ το μέγιστο πλήθος σημείων $x_1,\dots,x_N \in U$ ώστε για κάθε $i \neq j$ να έχουμε $x_i - x_j \notin V$. Έστω $u \in U \setminus \{x_1,\dots,x_N\}$. Τότε, το σύνολο $\{x_1,\dots,x_N,u\}$ δεν είναι 1-διαχωρισμένο ως προς V, άρα υπάρχει $i \in \{1,\dots,N\}$ τέτοιο ώστε $u - x_i \in V$. Αφού $0 \in V$ έπεται ότι $U \subseteq \bigcup_{i=1}^N (x_i + V)$, άρα $N(U,V) \leqslant M(U,V)$.

(β) Έστω $\{x_1,\dots,x_M\}$ μια οικογένεια σημείων του U τα οποία είναι 1-διαχωρισμένα. Έστω $\{z_1,\dots,z_N\}$ μια οικογένεια σημείων του $\mathbb{R}^n$ ώστε $U \subseteq \bigcup_{i=1}^N (z_i + V/2)$. Εφόσον $x_i \in U$ για κάθε $i = 1,\dots,M$, μπορούμε να ορίσουμε μια απεικόνιση $j : \{1,\dots,M\} \to \{1,\dots,N\}$, όπου το $j(i)$ επιλέγεται έτσι ώστε $x_i \in z_{j(i)} + V/2$. Αν $j(i_1) = j(i_2)$ τότε $x_{i_1} - x_{i_2} \in V/2 - V/2$. Από την κυρτότητα και τη συμμετρία του V έχουμε $V/2 - V/2 = V$, άρα $x_{i_1} - x_{i_2} \in V$. Όμως τα $x_1,\dots,x_M$ είναι 1-διαχωρισμένα ως προς V, οπότε αναγκαστικά έχουμε $i_1 = i_2$. Αυτό αποδεικνύει ότι η απεικόνιση j είναι 1-1, και τώρα έπεται ότι $M(U,V) \leqslant N(U,V/2)$. □

Επιπλέον, δεν είναι δύσκολο να ελέγξουμε ότι για κάθε τριάδα κυρτών σωμάτων U, V, W ισχύει η ανισότητα

$$N(U,W) \leqslant N(U,V)N(V,W).$$

Θα χρησιμοποιούμε συχνά την ακόλουθη απλή και σημαντική ογκομετρική εκτίμηση.

Λήμμα 2.8.2. *Έστω V ένα συμμετρικό κυρτό σώμα στον $\mathbb{R}^n$. Για κάθε $\epsilon > 0$,*

$$N(V, \epsilon V) \leqslant \left(1 + \frac{2}{\epsilon}\right)^n.$$

Απόδειξη. Από την Πρόταση 2.8.1 έχουμε ότι $N(V,\epsilon V) \leqslant M(V,\epsilon V)$. Έστω $M = M(V,\epsilon V)$ το μέγιστο πλήθος σημείων $x_1,\dots,x_M \in V$ ώστε για κάθε $i \neq j$ να έχουμε $x_i - x_j \notin \epsilon V$. Εφόσον το V είναι συμμετρικό κυρτό σώμα, τα σύνολα $x_i + \epsilon V/2$ είναι ξένα και

$$\bigcup_{i=1}^M (x_i + \epsilon V/2) \subseteq V + \epsilon V/2 = \left(1 + \tfrac{\epsilon}{2}\right) V.$$

Παίρνοντας όγκους βλέπουμε ότι

$$M\left(\tfrac{\epsilon}{2}\right)^n \leqslant \left(1 + \tfrac{\epsilon}{2}\right)^n,$$

το οποίο δίνει το άνω φράγμα που θέλαμε. □

Μια άλλη σημαντική παράμετρος που χρησιμοποιούμε για να μετρήσουμε το μέγεθος ενός υποσυνόλου T του $\mathbb{R}^n$ είναι η πολυπλοκότητα $\ell_*(T)$ που ορίζεται από την

$$\ell_*(T) = \mathbb{E}\left(\sup_{t \in T} \langle G, t\rangle\right),$$

όπου G είναι ένα τυπικό Gaussian τυχαίο διάνυσμα στον $\mathbb{R}^n$. Από τον ορισμό είναι φανερό ότι $\ell_*(T) = \ell_*(\mathrm{conv}(T))$, όπου $\mathrm{conv}(T)$ είναι η κυρτή θήκη του T.

Η ανισότητα Sudakov και η δυϊκή της είναι κλασσικά εργαλεία που βοηθούν στον υπολογισμό των αριθμών κάλυψης ενός συνόλου από Ευκλείδειες μπάλες ή, αντίστοιχα, των αριθμών κάλυψης μιας Ευκλείδειας μπάλας από μεταφορές ενός συμμετρικού κυρτού σώματος. Σε αυτές τις εκτιμήσεις υπεισέρχεται η πολυπλοκότητα ℓ_* του συνόλου.

Θεώρημα 2.8.3. *Έστω T ένα υποσύνολο του $\mathbb{R}^N$ και V ένα συμμετρικό κυρτό σώμα στον $\mathbb{R}^N$. Τότε, ισχύουν οι ανισότητες*

$$\sup_{\epsilon>0} \epsilon \sqrt{\log N(T, \epsilon B_2^N)} \leqslant c\ell_*(T) \tag{2.4}$$

και

$$\sup_{\epsilon>0} \epsilon \sqrt{\log N(B_2^N, \epsilon V)} \leqslant c\ell_*(V^\circ), \tag{2.5}$$

όπου για ένα τυπικό Gaussian τυχαίο διάνυσμα G στον $\mathbb{R}^N$,

$$\ell_*(T) = \mathbb{E}\left(\sup_{t\in T}\langle G, t\rangle\right) \quad \text{και} \quad \ell_*(V^\circ) = \mathbb{E}\left(\sup_{t\in V^\circ}\langle G, t\rangle\right) = \mathbb{E}\,\|G\|_V.$$

Η απόδειξη της ανισότητας Sudakov (2.4) βασίζεται στο λήμμα του Slepian, ένα κλασσικό λήμμα σύγκρισης Gaussian ανελίξεων.

Λήμμα 2.8.4 (Slepian). *Έστω $X_1, \dots, X_M$ και $Y_1, \dots, Y_M$ Gaussian τυχαίες μεταβλητές τέτοιες ώστε, για κάθε $i, j = 1, \dots, M$,*

$$\mathbb{E}\,|Y_i - Y_j|^2 \leqslant \mathbb{E}\,|X_i - X_j|^2.$$

Τότε,

$$\mathbb{E}\max_{1\leqslant k\leqslant M} Y_k \leqslant 2\mathbb{E}\max_{1\leqslant k\leqslant M} X_k.$$

Απόδειξη του Θεωρήματος 2.8.3. Δίνουμε πρώτα την απόδειξη της (2.4). Θεωρούμε M σημεία $x_1, \dots, x_M$ του T τα οποία είναι ϵ-διαχωρισμένα ως προς την Ευκλείδεια νόρμα $\|\cdot\|_2$ και για κάθε $i = 1, \dots, M$ ορίζουμε την Gaussian τυχαία μεταβλητή $X_i = \langle x_i, G\rangle$, όπου G είναι ένα τυπικό Gaussian τυχαίο διάνυσμα στον $\mathbb{R}^N$. Για κάθε $i \neq j$ έχουμε

$$\mathbb{E}\,|X_i - X_j|^2 = \|x_i - x_j\|_2^2 \geqslant \epsilon^2.$$

Θεωρούμε τώρα M ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές $g_1, \dots, g_M$ και για κάθε $i = 1, \dots, M$ ορίζουμε την τυχαία μεταβλητή $Y_i = \frac{\epsilon}{\sqrt{2}} g_i$. Για κάθε $i \neq j$ έχουμε

$$\mathbb{E}\,|Y_i - Y_j|^2 = \epsilon^2.$$

Από το Λήμμα 2.8.4 συμπεραίνουμε ότι

$$\frac{\epsilon}{\sqrt{2}}\,\mathbb{E}\left(\max_{1\leqslant k\leqslant M} g_k\right) \leqslant 2\mathbb{E}\left(\max_{1\leqslant k\leqslant M}\langle x_k, G\rangle\right) \leqslant 2\ell(T).$$

Από την Πρόταση 2.5.5 γνωρίζουμε ότι υπάρχει σταθερά $c>0$ τέτοια ώστε, για κάθε ακέραιο M,

$$\mathbb{E}\left(\max_{1\leqslant k\leqslant M} g_k\right) \geqslant \frac{1}{c}\sqrt{\log M}, \tag{2.6}$$

απ' όπου έπεται ότι

$$\epsilon\sqrt{\log M} \leqslant 2c\sqrt{2}\ell(T).$$

Από την Πρόταση 2.8.1 έχουμε τότε την (2.4).

Περνάμε τώρα στην απόδειξη της δυϊκής ανισότητας Sudakov (2.5). Το επιχείρημα που θα χρησιμοποιήσουμε μοιάζει πολύ με το ογκομετρικό επιχείρημα του Λήμματος 2.8.2. Αντικαθιστούμε το μέτρο Lebesgue με το μέτρο Gauss. Έστω $r>0$ το οποίο θα επιλέξουμε αργότερα. Παρατηρούμε ότι $N(B_2^N, \epsilon V) = N(rB_2^N, r\epsilon V)$ και θεωρούμε M σημεία $x_1,\dots,x_M \in rB_2^N$ τα οποία είναι $r\epsilon$-διαχωρισμένα ως προς τη νόρμα που επάγεται από το συμμετρικό κυρτό σώμα V. Από την Πρόταση 2.8.1 αρκεί να δείξουμε ότι

$$\epsilon\sqrt{\log M} \leqslant c\ell_*(V^\circ).$$

Οι μπάλες ακτίνας $\epsilon/2$ με κέντρα τα σημεία x_i είναι ξένες, και θεωρώντας το μέτρο Gauss της ένωσης αυτών των συνόλων, βλέπουμε ότι

$$\gamma_N\left(\bigcup_{i=1}^M (x_i + r\epsilon/2V)\right) = \sum_{i=1}^M \frac{1}{(2\pi)^{N/2}}\int_{\|z-x_i\|_V\leqslant r\epsilon/2} e^{-\|z\|_2^2}dz \leqslant 1.$$

Κάνοντας την αλλαγή μεταβλητής $z-x_i=u_i$ έχουμε

$$\frac{1}{(2\pi)^{N/2}}\int_{\|z-x_i\|_V\leqslant r\epsilon/2} e^{-\|z\|_2^2}dz = \frac{1}{(2\pi)^{N/2}}e^{-\|x_i\|_2^2/2}\int_{\|u_i\|_V\leqslant r\epsilon/2} e^{-\|z\|_2^2}e^{-\langle u_i,x_i\rangle}du_i,$$

και από την ανισότητα Jensen και το γεγονός ότι το V έχει βαρύκεντρο την αρχή των αξόνων,

$$\frac{1}{\gamma_N\left(\frac{r\epsilon}{2}V\right)}\frac{1}{(2\pi)^{N/2}}\int_{\|z-x_i\|_V\leqslant r\epsilon/2} e^{-\|z\|_2^2}dz \geqslant e^{-\|x_i\|_2^2/2}.$$

Εφόσον $x_i\in rB_2^N$, αυτό αποδεικνύει ότι

$$Me^{-r^2/2}\gamma_N\left(\tfrac{r\epsilon}{2}V\right) \leqslant 1.$$

Επιλέγουμε τώρα το r έτσι ώστε $r\epsilon/2 = 2\ell_*(V^\circ)$. Από την ανισότητα Markov έχουμε $\gamma_N\left(\frac{r\epsilon}{2}V\right) \geqslant$

$\frac{1}{2}$, και αυτό αποδεικνύει ότι $M \leqslant 2e^{r^2/2}$, το οποίο σημαίνει ότι

$$\epsilon \sqrt{\log M} \leqslant c\ell_*(V^\circ)$$

για κάποια σταθερά $c > 0$. □

Κλείνουμε αυτό το κεφάλαιο με μία άλλη σημαντική ιδιότητα των Gaussian ανελίξεων, μια ανισότητα συγκέντρωσης που θα χρησιμοποιήσουμε στο τρίτο κεφάλαιο της εργασίας.

Θεώρημα 2.8.5. *Έστω G ένα τυπικό Gaussian τυχαίο διάνυσμα στον $\mathbb{R}^n$. Έστω T υποσύνολο του $\mathbb{R}^n$ και έστω*

$$\sigma(T) = \sup_{t\in T} \left(\mathbb{E}\,|\langle G, t\rangle|^2\right)^{1/2}.$$

Τότε, για κάθε $u > 0$,

$$\mathbb{P}\left(\left|\sup_{t\in T}\langle G, t\rangle - \mathbb{E}\sup_{t\in T}\langle G, t\rangle\right| > u\right) \leqslant 2\exp(-cu^2/\sigma^2(T)), \tag{2.7}$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Κεφάλαιο 3

Η μέθοδος των τυχαίων προβολών

3.1 Λήμμα Johnson-Lindenstrauss

Υποθέτουμε ότι έχουμε ένα σύνολο δεδομένων που αποτελείται από ένα σύνολο N διανυσμάτων στον $\mathbb{R}^n$ και μας ενδιαφέρουν οι ανά δύο μεταξύ τους αποστάσεις ως προς την Ευκλείδεια νόρμα. Για την αποθήκευση αυτών των δεδομένων θα χρειαζόμασταν έναν $N \times n$ πίνακα ο οποίος θα είναι αρκετά μεγάλος στην περίπτωση που τα δεδομένα είναι πολλά. Θα μας ενδιέφερε λοιπόν να προβάλουμε τα δεδομένα σε έναν χώρο χαμηλότερης διάστασης, μέσω ενός τελεστή που θα δρούσε σαν ισομετρία στα διανύσματα αυτά διατηρώντας τις μεταξύ τους αποστάσεις.

Σύμφωνα με το λήμμα των Johnson–Lindenstrauss, αν μας επιτραπεί ένα σφάλμα $\epsilon > 0$, τότε υπάρχει γραμμικος τελεστής μέσω του οποίου η γεωμετρία του συνόλου των διανυσμάτων, θα μείνει αναλλοίωτη αν προβάλουμε τα διανύσματα σε κατάλληλο χώρο μικρότερης διάστασης. Μάλιστα, ο χώρος εικόνα του τελεστή έχει διάσταση της τάξης του λογαρίθμου του πλήθους των σημείων. Πιο συγκεκριμένα:

Λήμμα Johnson-Lindenstrauss. Εστω $\epsilon > 0$ και ένα σύνολο $V = \{x_1, \dots, x_N\}$ σε κάποιον χώρο Hilbert $\mathcal{H}$. Τότε, υπάρχει $k \leqslant c \cdot \epsilon^{-2} \cdot \log N$, όπου $c > 0$ είναι μια απόλυτη σταθερά, και ένας γραμμικός τελεστής $T : V \to l_2^k = (\mathbb{R}^k, \|\cdot\|_2)$ τέτοιος ώστε

$$(1-\epsilon)\|x_i - x_j\|_{\mathcal{H}} \leqslant \|T(x_i) - T(x_j)\|_2 \leqslant (1+\epsilon)\|x_i - x_j\|_{\mathcal{H}}$$

για κάθε $1 \leqslant i, j \leqslant N$.

Παρατηρήστε ότι, για τον πυρήνα μιας τέτοιας γραμμικής απεικόνισης θα έχουμε

$$\dim(\mathrm{Ker}(T)) = N - \dim(\mathrm{Im}(T)) \geqslant (N-k) = N\Big(1 - \frac{k}{N}\Big),$$

με τη διάσταση του χώρου εικόνα να είναι $k \sim \log N$. Με άλλα λόγια, από τη σκοπιά της γραμμικής άλγεβρας ο τελεστής είναι σχεδόν μηδενικός. Παρ' όλα αυτά, περιορισμένος στο σύνολο V, δρα σχεδόν σαν ισομετρία.

3.2 Απόδειξη μέσω συγκέντρωσης του μέτρου στη σφαίρα

Έστω $(X, \mathcal{A}, d, \mu)$ ένας μετρικός χώρος πιθανότητας. Δηλαδή, ο (X, d) είναι μετρικός χώρος και το μ είναι ένα μέτρο πιθανότητας στη σ-άλγεβρα $\mathcal{A}$ των Borel υποσυνόλων του (X, d). Αν $A \in \mathcal{A}$ και $t > 0$, η *t-περιοχή* του A είναι το σύνολο

$$A_t = \{x \in X : d(x, A) < t\}.$$

Η *συνάρτηση συγκέντρωσης* του $(X, \mathcal{A}, d, \mu)$ ορίζεται στο $(0, \infty)$ από την

$$\alpha(X, t) := 1 - \inf\{\mu(A_t) : \mu(A) \geqslant 1/2\}.$$

Λέμε ότι υπάρχει «συγκέντρωση μέτρου» στον χώρο αν η $\alpha(X, t)$ φθίνει γρήγορα (για παράδειγμα, εκθετικά ως προς t). Πολλές από τις εφαρμογές της συγκέντρωσης του μέτρου βασίζονται στο εξής θεώρημα.

Θεώρημα 3.2.1. *Έστω $(X, \mathcal{A}, d, \mu)$ μετρικός χώρος πιθανότητας. Αν $f : X \to \mathbb{R}$ είναι μια συνάρτηση Lipschitz με σταθερά 1, δηλαδή αν $|f(x) - f(y)| \leqslant d(x, y)$ για κάθε $x, y \in X$, τότε*

$$\mu\left(\{x \in X : |f(x) - \mathrm{med}(f)| > t\}\right) \leqslant 2\alpha(X, t)$$

όπου $\mathrm{med}(f)$ είναι ο μέσος Lévy της f.

Σημείωση. Ο μέσος Lévy $\mathrm{med}(f)$ της f είναι ένας αριθμός για τον οποίο

$$\mu(\{f \geqslant \mathrm{med}(f)\}) \geqslant 1/2 \text{ και } \mu(\{f \leqslant \mathrm{med}(f)\}) \geqslant 1/2.$$

Απόδειξη του Θεωρήματος 3.2.1. Θέτουμε $A = \{x : f(x) \geqslant \mathrm{med}(f)\}$ και $B = \{x : f(x) \leqslant \mathrm{med}(f)\}$. Αν $y \in A_t$ τότε υπάρχει $x \in A$ με $d(x, y) \leqslant t$, οπότε

$$f(y) = f(y) - f(x) + f(x) \geqslant -d(y, x) + \mathrm{med}(f) \geqslant \mathrm{med}(f) - t$$

αφού η f είναι 1-Lipschitz. Ομοίως, αν $y \in B_t$ τότε υπάρχει $x \in B$ με $d(x, y) \leqslant t$, οπότε

$$f(y) = f(y) - f(x) + f(x) \leqslant d(y, x) + \mathrm{med}(f) \leqslant \mathrm{med}(f) + t.$$

Δηλαδή, αν $y \in A_t \cap B_t$ τότε $|f(x) - \mathrm{med}(f)| \leqslant t$. Με άλλα λόγια,

$$\{x \in X : |f(x) - \mathrm{med}(f)| > t\} \subseteq (A_t \cap B_t)^c = A_t^c \cup B_t^c. \tag{3.1}$$

Όμως, από τον ορισμό της συνάρτησης συγκέντρωσης έχουμε $\mu(A_t) \geqslant 1 - \alpha(X, t)$ και $\mu(B_t) \geqslant 1 - \alpha(X, t)$. Επιστρέφοντας στην (3.1) βλέπουμε ότι

$$\mu\left(\{|f - \mathrm{med}(f)| > t\}\right) \leqslant (1 - \mu(A_t)) + (1 - \mu(B_t)) \leqslant 2\alpha(X, t).$$

□

Στην περίπτωση που η συνάρτηση συγκέντρωσης φθίνει πολύ γρήγορα, το Θεώρημα 3.2.1 δείχνει ότι οι 1-Lipschitz συνεχείς συναρτήσεις είναι «σχεδόν σταθερές» σε «σχεδόν ολόκληρο τον χώρο». Ισχύει μάλιστα και το αντίστροφο.

Πρόταση 3.2.2. *Έστω $(X, \mathcal{A}, d, \mu)$ μετρικός χώρος πιθανότητας. Αν για κάποιο $t > 0$ και για κάθε 1-Lipschitz συνάρτηση $f : X \to \mathbb{R}$ έχουμε*

$$\mu\left(\{x \in X : |f(x) - \mathrm{med}(f)| > t\}\right) \leqslant \eta,$$

τότε $\alpha(X, t) \leqslant \eta$.

Απόδειξη. Έστω A Borel υποσύνολο του X με $\mu(A) \geqslant 1/2$. Θεωρούμε τη συνάρτηση $f(x) = d(x, A)$. Η f είναι 1-Lipschitz και $\mathrm{med}(f) = 0$ γιατί η f παίρνει μη αρνητικές τιμές και $\mu(\{x : f(x) = 0\}) \geqslant 1/2$. Από την υπόθεση παίρνουμε

$$\mu(\{x \in X : d(x, A) > t\}) \leqslant \eta,$$

δηλαδή $1 - \mu(A_t) \leqslant \eta$. Έπεται ότι $\alpha(X, t) \leqslant \eta$. □

Για την απόδειξη του λήμματος Johnson-Lindenstrauss θα εκμεταλλευτούμε τη συγκέντρωση του μέτρου στη σφαίρα. Θεωρούμε τη μοναδιαία σφαίρα S^{n-1} στον $\mathbb{R}^n$ εφοδιασμένη με τη γεωδαισιακή μετρική ρ: η απόσταση $\rho(x, y)$ δύο σημείων $x, y \in S^{n-1}$ είναι η κυρτή γωνία xoy στο επίπεδο που ορίζεται από την αρχή των αξόνων o και τα x, y. Η S^{n-1} γίνεται χώρος πιθανότητας με το μοναδικό αναλλοίωτο ως προς στροφές μέτρο σ: για κάθε Borel σύνολο $A \subseteq S^{n-1}$ θέτουμε

$$\sigma(A) := \frac{|\tilde{A}|}{|B_2^n|},$$

όπου B_2^n είναι η μοναδιαία Ευκλείδεια μπάλα και

$$\tilde{A} := \{sx : x \in A \text{ και } 0 \leqslant s \leqslant 1\}.$$

Είναι εύκολο να δεί κανείς ότι αν $\rho(x, y) = \theta$ τότε

$$\|x - y\|_2 = 2 \sin \frac{\theta}{2},$$

συνεπώς η γεωδαισιακή και η Ευκλείδεια απόσταση των $x, y \in S^{n-1}$ συγκρίνονται μέσω της

$$\frac{2}{\pi}\rho(x, y) \leqslant \|x - y\|_2 \leqslant \rho(x, y).$$

Το ισοπεριμετρικό πρόβλημα στη σφαίρα διατυπώνεται ως εξής:

> Δίνονται $\alpha \in (0, 1)$ και $t > 0$. Ανάμεσα σε όλα τα Borel υποσύνολα A της σφαίρας για τα οποία $\sigma(A) = \alpha$, να βρεθούν εκείνα για τα οποία ελαχιστοποιείται η

επιφάνεια $\sigma(A_t)$ της t-περιοχής του A.

Η απάντηση δίνεται από το ακόλουθο θεώρημα:

Ισοπεριμετρική ανισότητα στη σφαίρα. *Έστω $\alpha \in (0,1)$ και*

$$B(x,r) = \{y \in S^{n-1} : \rho(x,y) \leqslant r\}$$

μια μπάλα στην S^{n-1} με ακτίνα $r > 0$ που επιλέγεται ώστε $\sigma(B(x,r)) = \alpha$. Τότε, για κάθε $A \subseteq S^{n-1}$ με $\sigma(A) = \alpha$ και για κάθε $t > 0$ έχουμε

$$\sigma(A_t) \geqslant \sigma\big(B(x,r)_t\big) = \sigma\big(B(x,r+t)\big).$$

Δηλαδή, για οποιοδήποτε δοσμένο μέτρο α και οποιοδήποτε $t > 0$ οι μπάλες μέτρου α δίνουν τη λύση του ισοπεριμετρικού προβλήματος.

Η απόδειξη της ισοπεριμετρικής ανισότητας γίνεται με σφαιρική συμμετρικοποίηση και επαγωγή ως προς τη διάσταση. Ας θεωρήσουμε την ειδική περίπτωση $\alpha = 1/2$. Αν $\sigma(A) = 1/2$ και $t > 0$, τότε μπορούμε να εκτιμήσουμε το μέγεθος του A_t χρησιμοποιώντας την ισοπεριμετρική ανισότητα:

$$\sigma(A_t) \geqslant \sigma\left(B(x, \tfrac{\pi}{2} + t)\right)$$

για κάθε $t > 0$ και $x \in S^{n-1}$. Εκτιμώντας από κάτω το δεξιό μέλος αυτής της ανισότητας οδηγούμαστε στην ακόλουθη ανισότητα.

Θεώρημα 3.2.3. *Έστω $A \subseteq S^{n+1}$ με $\sigma(A) = 1/2$ και έστω $t > 0$. Τότε,*

$$\sigma(A_t) \geqslant 1 - \sqrt{\pi/8}\exp(-t^2 n/2). \tag{3.2}$$

Αυτό που έχει σημασία σε σχέση με την παραπάνω εκτίμηση είναι ότι, όσο μικρό $t > 0$ κι αν διαλέξουμε, η ακολουθία $\exp(-t^2 n/2)$ τείνει στο 0 καθώς $n \to \infty$ και μάλιστα με πολύ ταχύ ρυθμό (εκθετικά ως προς n). Επομένως, το ποσοστό της σφαίρας που μένει έξω από την t-περιοχή οποιουδήποτε υποσυνόλου A της S^{n+1} με $\sigma(A) = 1/2$ είναι «σχεδόν μηδενικό» αν η διάσταση n είναι αρκετά μεγάλη.

Η απόδειξη του Θεωρήματος 3.2.3 βασίζεται πολύ ισχυρά στη σφαιρική ισοπεριμετρική ανισότητα. Για τις περισσότερες όμως εφαρμογές που έχουμε στο νού μας είναι αρκετή μια ανισότητα σαν την (3.2) και όχι η ακριβής λύση του ισοπεριμετρικού προβλήματος. Θα δώσουμε μια απλή απόδειξη της (3.2) χωρίς να περάσουμε μέσα από την ισοπεριμετρική ανισότητα, χρησιμοποιώντας την ανισότητα Brunn-Minkowski.

Ορισμός 3.2.4. *Έστω A και B μη κενά υποσύνολα του $\mathbb{R}^n$. Ορίζουμε*

$$A + B := \{a + b \mid a \in A,\, b \in B\}$$

και για κάθε $t \geqslant 0$,

$$tA = \{ta \mid a \in A\}.$$

Η ανισότητα Brunn-Minkowski συνδέει το άθροισμα Minkowski με τον όγκο $|\cdot|$ στον $\mathbb{R}^n$:

Θεώρημα 3.2.5 (ανισότητα Brunn-Minkowski). *Έστω K και T δύο μη κενά συμπαγή υποσύνολα του $\mathbb{R}^n$. Τότε,*

$$|K+T|^{1/n} \geqslant |K|^{1/n} + |T|^{1/n}.$$

Στην περίπτωση που τα K και T είναι κυρτά σώματα (κυρτά συμπαγή με μη κενό εσωτερικό), ισότητα στην ανισότητα Brunn-Minkowski μπορεί να ισχύει μόνο αν τα K και T είναι ομοιοθετικά (δηλαδή, αν $K = aT + x$ για κάποιον $a \geqslant 0$ και κάποιο $x \in \mathbb{R}^n$). Η ανισότητα Brunn-Minkowski εκφράζει με μια έννοια το γεγονός ότι ο όγκος είναι **κοίλη** συνάρτηση ως προς την πρόσθεση κατά Minkowski. Για το λόγο αυτό συχνά γράφεται στην ακόλουθη μορφή: Αν K, T είναι μη κενά συμπαγή υποσύνολα του $\mathbb{R}^n$ και $\lambda \in (0,1)$, τότε

$$|\lambda K + (1-\lambda)T|^{1/n} \geqslant \lambda |K|^{1/n} + (1-\lambda)|T|^{1/n}.$$

Χρησιμοποιώντας αυτήν την ανισότητα και την ανισότητα αριθμητικού-γεωμετρικού μέσου, μπορούμε ακόμα να γράψουμε:

$$|\lambda K + (1-\lambda)T| \geqslant |K|^{\lambda}|T|^{1-\lambda}.$$

Η ασθενέστερη αυτή μορφή της ανισότητας Brunn-Minkowski έχει το πλεονέκτημα ότι είναι ανεξάρτητη της διάστασης.

Λήμμα 3.2.6. *Θεωρούμε το ομοιόμορφο μέτρο πιθανότητας μ στην Ευκλείδεια μοναδιαία μπάλα B_2^n. Δηλαδή, $\mu(A) = |A|/|B_2^n|$ για κάθε Borel $A \subseteq B_2^n$. Αν $A, C \subseteq B_2^n$ συμπαγή, και*

$$d(A,C) := \min\{\|a-c\|_2 : a \in A, c \in C\} = \rho > 0,$$

τότε

$$\min\{\mu(A), \mu(C)\} \leqslant \exp(-\rho^2 n/8).$$

Απόδειξη. Για το σύνολο $\frac{A+C}{2}$, από την ανισότητα Brunn–Minkowski παίρνουμε $\left|\frac{A+C}{2}\right| \geqslant \min\{|A|, |C|\}$. Συνεπώς,

$$\mu\left(\frac{A+C}{2}\right) \geqslant \min\{\mu(A), \mu(C)\}.$$

Από την άλλη πλευρά, αν $a \in A$ και $c \in C$, ο κανόνας του παραλληλογράμμου δίνει

$$\|a+c\|_2^2 = 2\|a\|_2^2 + 2\|c\|_2^2 - \|a-c\|_2^2 \leqslant 4 - \rho^2,$$

επομένως

$$\frac{A+C}{2} \subseteq \sqrt{1-\frac{\rho^2}{4}}\, B_2^n.$$

Συνδυάζοντας αυτές τις ανισότητες βλέπουμε ότι

$$\min\{\mu(A), \mu(C)\} \leqslant \left(1-\frac{\rho^2}{4}\right)^{n/2} \leqslant \exp(-\rho^2 n/8)$$

□

Απόδειξη του Θεωρήματος 3.2.3. Έστω $A \subseteq S^{n-1}$ με $\sigma(A) = 1/2$ και έστω $t > 0$. Θέτουμε $C = S^{n-1} \setminus A_t$ και θεωρούμε τα υποσύνολα

$$A_1 = \{\rho a : a \in A, \tfrac{1}{2} \leqslant \rho \leqslant 1\} \text{ και } C_1 = \{\rho a : a \in C, \tfrac{1}{2} \leqslant \rho \leqslant 1\}$$

της B_2^n. Εύκολα ελέγχουμε ότι

$$d(A_1, C_1) \geqslant \sin\frac{t}{2} \geqslant \frac{t}{\pi}.$$

Από το Λήμμα 3.2.6 συμπεραίνουμε ότι

$$|C_1| \leqslant \exp(-d^2 n/8)|B_2^n| \leqslant \exp\left(-t^2 n/(8\pi^2)\right)|B_2^n|.$$

Όμως, από τον ορισμό του σ έχουμε $|B_2^n|\sigma(C) = |\tilde{C}|$ και $|C_1| = (1 - 2^{-n})|\tilde{C}|$. Έτσι, βλέπουμε ότι

$$\sigma(A_t^c) = \sigma(C) \leqslant \frac{1}{1 - 2^{-n}} \exp\left(-t^2 n/(8\pi^2)\right).$$

Δηλαδή,

$$\sigma(A_t) \geqslant 1 - c_1 \exp(-c_2 t^2 n)$$

όπου $c_1 = 2$ και $c_2 = 1/(8\pi^2)$. Η ανισότητα αυτή είναι εντελώς ανάλογη με την ανισότητα του Θεωρήματος 3.2.3 αν εξαιρέσουμε τις ακριβείς τιμές των σταθερών c_1 και c_2. □

Θεώρημα 3.2.7 (Johnson-Lindenstrauss). *Έστω V ένα σύνολο N-σημείων σε κάποιον χώρο Hilbert $\mathcal{H}$. Για κάθε $\epsilon \in (0, 1]$ υπάρχει $(1 + \epsilon)$-εμφύτευση του V στον ℓ_2^k για κάποιον $k = O(\epsilon^{-2} \log n)$. Δηλαδή, υπάρχει $T : V \to \ell_2^k$ τέτοια ώστε*

$$\|x - y\|_{\mathcal{H}} \leqslant \|T(x) - T(y)\|_2 \leqslant (1 + \epsilon)\|x - y\|_{\mathcal{H}}$$

για κάθε $x, y \in V$.

Για την απόδειξη χρησιμοποιούμε το λήμμα του Lévy που είναι άμεση συνέπεια των προηγούμενων αποτελεσμάτων.

Θεώρημα 3.2.8 (λήμμα του Lévy). *Αν $f : S^{N-1} \to \mathbb{R}$ είναι μια 1-Lipschitz συνάρτηση τότε, για κάθε $t > 0$,*

$$\mathbb{P}[f < \mathrm{med}(f) - t] \leqslant e^{-t^2 N/2} \quad \textit{και} \quad \mathbb{P}[f > \mathrm{med}(f) + t] \leqslant e^{-t^2 N/2}.$$

Έστω $1 \leqslant k \leqslant N$. Θεωρούμε τη συνάρτηση $f_k : S^{N-1} \to \mathbb{R}$ που απεικονίζει κάθε $x \in S^{N-1}$ στο μήκος $f_k(x)$ της προβολής του πάνω στον υπόχωρο E_k που παράγεται από τα πρώτα k διανύσματα της συνήθους ορθοκανονικής βάσης του $\mathbb{R}^N$. Δηλαδή,

$$f_k(x) = \sqrt{x_1^2 + \cdots + x_k^2}$$

όπου $x = (x_1, \ldots, x_N)$.

Λήμμα 3.2.9. *Υπάρχει πραγματικός αριθμός $m = m(N,k)$ ώστε*

$$\mathbb{P}[f_k < m - t] \leqslant e^{-t^2N/2} \quad \text{και} \quad \mathbb{P}[f_k > m + t] \leqslant e^{-t^2N/2}$$

για κάθε $t > 0$. Αν το N είναι μεγαλύτερο από κατάλληλη απόλυτη σταθερά και αν $k \geqslant 10 \log N$, τότε $m \geqslant \frac{1}{2}\sqrt{\frac{k}{N}}$.

Απόδειξη. Η ανισότητα ισχύει με $m = \mathrm{med}(f)$. Αυτό είναι άμεση συνέπεια του Θεωρήματος 3.2.8, αρκεί να παρατηρήσουμε ότι η f_k είναι 1-Lipschitz.

Μένει να δούμε το κάτω φράγμα για το m. Παρατηρούμε ότι

$$\mathbb{E}[f_k^2] = k\mathbb{E}[x_1^2] = \frac{k}{N}\mathbb{E}[\|x\|_2^2] = \frac{k}{N}.$$

Για κάθε $t > 0$ μπορούμε να γράψουμε

$$\begin{aligned} \frac{k}{N} &= \mathbb{E}[f_k^2] \leqslant \mathbb{P}[f \leqslant m + t] \cdot (m + t)^2 + \mathbb{P}[f > m + t] \cdot \|f_k\|_\infty^2 \\ &\leqslant (m + t)^2 + e^{-t^2N/2}. \end{aligned}$$

Επιλέγουμε $t = \sqrt{k/5N}$. Έχουμε υποθέσει ότι $k \geqslant 10 \log N$, απ' όπου έπεται ότι $\exp(-t^2N/2) \leqslant \frac{1}{N}$. Άρα,

$$\frac{k}{N} \leqslant (m + t)^2 + \frac{1}{N},$$

δηλαδή,

$$m \geqslant \sqrt{\frac{k-1}{N}} - t \geqslant \frac{1}{2}\sqrt{\frac{k}{N}}$$

αν το N (οπότε και το k) είναι αρκετά μεγάλο. Ακριβέστερος υπολογισμός δείχνει ότι $m = \sqrt{k/N} + O(1/\sqrt{N})$ για κάθε k. □

Απόδειξη του Θεωρήματος 3.2.7. Μπορούμε να υποθέσουμε ότι το N είναι αρκετά μεγάλο. Έστω V ένα σύνολο N σημείων σε κάποιον χώρο Hilbert $\mathcal{H}$. Χωρίς περιορισμό της γενικότητας, θεωρώντας έναν N-διάστατο υπόχωρο του $\mathcal{H}$ τον οποίο περιέχεται το V, μπορούμε να υποθέσουμε ότι $\mathcal{H} = \ell_2^N$. Θέτουμε $k = 200\epsilon^{-2}\log N$. Αν $k \geqslant N$ τότε δεν έχουμε τίποτα να δείξουμε, οπότε υποθέτουμε ότι $k < N$.

Έστω L ο τυχαίος k-διάστατος γραμμικός υπόχωρος του ℓ_2^N. Έστω $p_L : \ell_2^N \to L$ η ορθογώνια προβολή επί του L και έστω $m = m(N,k)$.

Έστω x και y στο V. Θα δείξουμε ότι η πιθανότητα (ως προς L) να μην ισχύει η

$$\left(1 - \frac{\epsilon}{3}\right)m\|x - y\|_2 \leqslant \|p_L(x) - p_L(y)\|_2 \leqslant \left(1 + \frac{\epsilon}{3}\right)m\|x - y\|_2 \tag{3.3}$$

είναι μικρότερη από $\frac{1}{N^2}$.

Θέτουμε $u = x - y$. Αφού κάθε p_L είναι γραμμική απεικόνιση, η (3.3) γράφεται

$$\left(1 - \frac{\epsilon}{3}\right) m\|u\|_2 \leqslant \|p_L(u)\|_2 \leqslant \left(1 + \frac{\epsilon}{3}\right) m\|u\|_2,$$

και, επειδή όλοι οι όροι της ανισότητας είναι θετικά ομογενείς, μπορούμε να υποθέσουμε ότι $\|u\|_2 = 1$. Δηλαδή, για σταθερό $u \in S^{N-1}$ ζητάμε να ισχύει η

$$\mathbb{P}\left(L : \ \big|\|p_L(u)\|_2 - m\big| > \frac{\epsilon m}{3}\right) \leqslant \frac{1}{N^2}.$$

Όμως, η πιθανότητα αυτή είναι ακριβώς ίση με την

$$\mathbb{P}\left(x \in S^{N-1} : \ \big|f_k(x) - m\big| > \frac{\epsilon m}{3}\right).$$

Από το Λήμμα 3.2.9 είναι μικρότερη από

$$2\exp\left(-\frac{\epsilon^2 m^2 N}{18}\right) \leqslant \exp\left(-\frac{\epsilon^2 k}{72}\right) < N^{-2}.$$

Αφού υπάρχουν λιγότερα από N^2 ζευγάρια σημείων $x, y \in V$, υπάρχει κάποιος L ώστε η (3.3) να ισχύει για κάθε $x, y \in V$. Τότε, η απεικόνιση $p_L : V \to L$ ορίζει μια εμφύτευση του V στον ℓ_2^k με παραμόρφωση $D \leqslant (1 + \epsilon/3)/(1 - \epsilon/3) < 1 + \epsilon$ (αν $0 < \epsilon < 1$). □

3.3 Απόδειξη με τυχαίους πίνακες

Για δοθέν $\epsilon > 0$, η κατασκευή του γραμμικού τελεστή $T : \mathbb{R}^N \to \mathbb{R}^k$ θα γίνει πιθανοθεωρητικά, μέσω ενός τυχαίου $k \times N$ πίνακα A, ο οποίος θα επάγει μια γραμμική απεικόνιση που δρα σχεδόν σαν ισομετρία στα στοιχεία του συνόλου V. Συνήθως στη βιβλιογραφία, τα στοιχεία του πίνακα επιλέγονται ανεξάρτητα να ακολουθούν την τυπική κανονική κατανομή όμως, όπως θα δούμε, μπορούν να επιλεγούν να έχουν οποιαδήποτε υποκανονική κατανομή με μέση τιμή 0 και διασπορά 1. Θα ξεκινήσουμε με ένα βοηθητικό λήμμα.

Λήμμα 3.3.1. *Έστω $u_1, u_2, \ldots, u_N$ ανεξάρτητες υποκανονικές τυχαίες μεταβλητές με $\mathbb{E}(u_i) = 0$ και $a_1, \ldots, a_N$ με $\sum_{i=1}^N a_i^2 = 1$. Υποθέτουμε ότι υπάρχει $K > 0$, ώστε $\|u_i\|_{L_{\psi_1}} \leqslant M$ για κάθε $i = 1, \ldots, N$. Τότε, η τυχαία μεταβλητή $S = \sum_{i=1}^N a_i \cdot u_i$ είναι υποκανονική με σταθερά $\leqslant K$.*

Απόδειξη. Θα χρησιμοποιήσουμε τον χαρακτηρισμό μιας υποκανονικής τυχαίας μεταβλητής

μέσω της ροπογεννήτριας. Για κάθε $\lambda \in \mathbb{R}$, υπολογίζουμε

$$\begin{aligned}\mathbb{E}(e^{\lambda S}) &= \mathbb{E}\left(\exp\left(\lambda \sum_{i=1}^{N} a_i \cdot u_i\right)\right) \\ &= \prod_{i=1}^{N} \mathbb{E}(e^{\lambda a_i u_i}) \\ &\leqslant \prod_{i=1}^{N} (e^{(\lambda a_i)^2 K^2}) \\ &= \exp\left(\lambda^2 K^2 \sum_{i=1}^{N} a_i\right) = e^{\lambda^2 K^2},\end{aligned}$$

εφόσον οι τυχαίες μεταβλητές έχουν υποτεθεί ανεξάρτητες υποκανονικές και η έκφραση για την ροπογεννήτρια ισχύει για κάθε $\lambda \in \mathbb{R}$. □

Θεώρημα 3.3.2 (Johnson-Lindenstrauss). *Έστω $\epsilon > 0$ και $V = \{x_1, \ldots, x_N\}$ ένα σύνολο N σημείων σε έναν χώρο Hilbert $\mathcal{H}$. Τότε υπάρχει $k \leqslant c \cdot \epsilon^{-2} \cdot \log N$ και τελεστής $T : V \to \ell_2^k$ τέτοιος ώστε*

$$(1-\epsilon)\|x_i - x_j\|_{\mathcal{H}} \leqslant \|T(x_i) - T(x_j)\|_2 \leqslant (1+\epsilon)\|x_i - x_j\|_{\mathcal{H}}$$

για κάθε $1 \leqslant i, j \leqslant N$.

Απόδειξη. Μπορούμε να υποθέσουμε ότι $\mathcal{H} = \ell_2^N$. Για δοθέν $\epsilon > 0$, θα αναζητήσουμε γραμμικό τελεστή $T : \mathbb{R}^N \to \mathbb{R}^k$ ώστε, για κάθε $1 \leqslant i \neq j \leqslant N$,

$$(1-\epsilon) \leqslant \left\| T\left(\frac{x_i - x_j}{\|x_i - x_j\|_{\mathcal{H}}}\right)\right\|_2 \leqslant (1+\epsilon).$$

Κανονικοποιώντας τα διανύσματα του συνόλου V, θεωρούμε το σύνολο Δ των μοναδιαίων διανυσμάτων στην σφαίρα του $\mathbb{R}^N$ ως

$$\Delta = \left\{\frac{x_i - x_j}{\|x_i - x_j\|_{\mathcal{H}}}, i, j = 1, 2, \ldots, N, i \neq j\right\}.$$

Χρησιμοποιώντας την πιθανοθεωρητική μέθοδο θα αποδείξουμε ότι με μεγάλη πιθανότητα, υπάρχει γραμμικός τελεστής T με $\|T(u)\|_2 \sim 1$ ή πιο συγκεκριμένα ότι $|\|T(u)\|_2 - 1| \leqslant \epsilon$ για κάθε $u \in \Delta$.

Εισάγουμε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$, με τα στοιχεία του δειγματικού χώρου Ω να αποτελούνται από $k \times N$ τυχαίους πίνακες $A = (a_{ij})$, τα στοιχεία των οποίων είναι ανεξάρτητες υποκανονικές τυχαίες μεταβλητές με μέση τιμή 0 και διασπορά 1.

Ο ζητούμενος τελεστής T θα επάγεται από έναν τυχαίο πίνακα της μορφής λA, για κατάλληλη επιλογή του $\lambda \in \mathbb{R}$ ώστε η τυχαία μεταβλητή $\|T(u)\|_2$ (για σταθερό $u \in S^{N-1}$) να έχει μέση τιμή 1.

Για την επιλογή του $\lambda \in \mathbb{R}$ θεωρούμε έναν $k \times N$ πίνακα στο δειγματικό χώρο Ω, με

$$A = \begin{pmatrix} - & \Gamma_1 & - \\ - & \cdots & - \\ - & \Gamma_m & - \end{pmatrix} \text{όπου } \Gamma_i = (a_{i1}, \ldots, a_{iN})$$

και υπολογίζουμε τη μέση τιμή της τυχαίας μεταβλητής $\|A(u)\|_2^2$ ως

$$\mathbb{E}\|A(u)\|_2^2 = \mathbb{E}\Big[\sum_{i=1}^{k} \langle \Gamma_i, u\rangle^2\Big] = \sum_{i=1}^{k} \mathbb{E}\langle \Gamma_i, u\rangle^2 = k\|u\|_2^2$$

διότι για κάθε $i = 1, 2, \ldots, N$ είναι

$$\begin{aligned} \mathbb{E}\langle \Gamma_i, u\rangle^2 = \mathbb{E}\left[\sum_{j=1}^{N} a_{ij} \cdot u_j\right]^2 &= \mathbb{E}\left[\sum_{j=1}^{N} a_{ij}^2 \cdot u_j^2 + \sum_{j\neq s} a_{ij} \cdot u_j \cdot a_{is} \cdot u_s\right] \\ &= \sum_{j=1}^{N} \mathbb{E}\left[a_{ij}^2 \cdot u_j^2\right] + \sum_{j\neq s} \mathbb{E}\left[a_{ij} \cdot u_j \cdot a_{is} \cdot u_s\right] \\ &= \sum_{j=1}^{N} u_j^2 + 0 = \|u\|_2^2. \end{aligned}$$

Χρησιμοποιήσαμε το γεγονός ότι $\mathbb{E}[a_{ij}^2] = 1$, $\mathbb{E}[a_{ij}] = 0$ αλλά και την ανεξαρτησία των υποκανονικών τυχαίων μεταβλητών. Τελικά έπεται ότι

$$\mathbb{E}\|Au\|_2^2 = k\|u\|_2^2 \iff \mathbb{E}\|(\frac{1}{\sqrt{k}}A) \cdot u\|_2^2 = \|u\|_2^2,$$

δηλαδή ο τελεστής $T = \frac{1}{\sqrt{k}} \cdot A$, δρά σαν ισομετρία ως προς τη μέση τιμή. Στη συνεχεια αποδεικνύουμε ότι ενας τέτοιος τελεστής ικανοποιεί το ζητούμενο με μεγάλη πιθανότητα.

Θεωρούμε το ενδεχόμενο $G = \{\forall\, u \in \Delta : |\|T(u)\|_2 - 1| \leqslant \epsilon\}$ και θα δείξουμε ότι $\mathbb{P}(G) > 0$ και μάλιστα πραγματοποιείται με μεγάλη πιθανότητα. Παρατηρούμε ότι για το συμπληρωματικό του ενδεχόμενο έχουμε

$$B = \{\exists\, u \in \Delta : |\|T(u)\|_2 - 1| > \epsilon\} = \bigcup_{u\in\Delta} B_u,$$

όπου $B_u = \{u \in \Delta : |\|Tu\|_2 - 1| > \epsilon\}$. Άρα, ισοδύναμα, αρκεί να δείξουμε ότι $\mathbb{P}(B) < 1$, αποδεικνύοντας ότι κάθε πιθανότητα $\mathbb{P}(B_u)$ είναι αρκούντως «μικρή» για κάθε $u \in \Delta$.

Ισχυρισμός 3.3.3. *Για σταθερό $u \in \Delta$ για την τυχαία μεταβλητή $\frac{1}{\sqrt{k}} \cdot \|Au\|_2^2$ ικανοποιείται η ανισότητα*

$$\mathbb{P}(B_u) = \mathbb{P}\left(|\frac{1}{k}\|Au\|_2^2 - 1| \geqslant \epsilon\right) \leqslant 2e^{-c\cdot\epsilon^2 k}.$$

Για την απόδειξη του ισχυρισμού, μπορούμε να υποθέσουμε ότι $0 < \epsilon < 1$. Σταθεροποιούμε

ένα $u \in \Delta$ και για κάθε $i = 1, 2, \ldots, N$ θεωρούμε την τυχαία μεταβλητή

$$X_i = \langle \Gamma_i, u \rangle^2 - 1.$$

Η τυχαία μεταβλητή X_i είναι υποεκθετική με $\mathbb{E}(X_i) = 0$, διότι η τυχαία μεταβλητή $\langle \Gamma_i, u \rangle$ είναι υποκανονική με σταθερά K, άρα η $\langle \Gamma_i, u \rangle^2$ είναι υποεκθετική με σταθερά K^2 και συνεπώς η X_i είναι υποεκθετική τυχαία μεταβλητή με σταθερά $K^2 + 1$.
Παρατηρούμε τώρα, ότι το ενδεχόμενο $\{|\frac{1}{k}\|Au\|_2^2 - 1| \geqslant \epsilon\}$ μπορεί να γραφτεί ως

$$\{|\frac{1}{k}\|Au\|_2^2 - 1| \geqslant \epsilon\} = \{|\frac{1}{k}\sum_{i=1}^{k}\langle \Gamma_i, u \rangle^2 - 1| \geqslant \epsilon\} = \{|\frac{1}{k}\sum_{i=1}^{k} X_i\Big| \geqslant \epsilon\},$$

και έπεται ότι

$$\begin{aligned}
\mathbb{P}\Big(\Big\{\Big|\frac{1}{k}\sum_{i=1}^{k} X_i\Big| \geqslant \epsilon\Big\}\Big) &\leqslant 2\exp\Big(- ck \cdot \min\Big\{\frac{\epsilon^2}{(K+1)^2}, \frac{\epsilon}{K+1}\Big\}\Big) \\
&\leqslant 2exp\left(-ck\frac{\epsilon^2}{(K+1)^2}\right) \\
&\leqslant 2e^{-c\cdot\epsilon^2 k}.
\end{aligned}$$

Ολοκληρώνουμε τώρα την απόδειξη του λήμματος. Από την υποπροσθετικότητα του μέτρου $\mathbb{P}$, υπολογίζουμε

$$\mathbb{P}(B) = \mathbb{P}(\bigcup_{u\in\Delta} B_u) \leqslant \sum_{u\in\Delta}\mathbb{P}(B_u) \leqslant 2N^2 \cdot e^{-c\epsilon^2 k} < 1,$$

για $k \geqslant c\epsilon^{-2}\log N$. Επιπλέον, είναι και αρκούντως μικρή, ήτοι $\mathbb{P}(B) < e^{-\frac{c}{2}\epsilon^2 k}$, αφού

$$2N^2 e^{-c\epsilon^2 k} < e^{-\frac{c}{2}\epsilon^2 k} \qquad \Longleftrightarrow \qquad N^3 < e^{\frac{c}{2}\epsilon^2 k} \qquad \Longleftrightarrow \qquad k > \frac{6}{c}\epsilon^{-2}\log N.$$

Άρα, για κατάλληλη επιλογή της απόλυτης σταθεράς $c > 0$, και για $k = c\epsilon^{-2}\log N$, η πιθανότητα του ενδεχομένου

$$G = \{\forall\, u \in \Delta : |\|Tu\|_2 - 1| \leqslant \epsilon\}$$

είναι τουλάχιστον $1 - e^{-c\epsilon^2 k}$ και άρα η ζητούμενη γραμμική απεικόνιση υπάρχει και μπορεί να προσδιοριστεί από έναν τυχαίο αλγόριθμο. □

3.4 Το βέλτιστο του λήμματος

Σε αυτήν την ενότητα αποδεικνύουμε ότι η εξάρτηση της διάστασης k του χώρου εικόνα από το πλήθος N των σημείων του συνόλου V αλλά και το ϵ στο λήμμα Johnson-Lindenstrauss είναι βέλτιστη. Αρχικά αποδεικνύουμε ότι αν για το σφάλμα απαιτηθεί $\epsilon = 0$, δηλαδή ο τελεστής να είναι ισομετρία, τότε $k \geqslant N$.

Πρόταση 3.4.1. *Έστω $V = \{0, e_1, \dots, e_N\}$ η συνήθης ορθοκανονική βάση του $\mathbb{R}^N$ με την προσθήκη του μηδενικού διανύσματος. Υποθέτουμε ότι υπάρχει τελεστής $T : V \to \mathbb{R}^k$, ο οποίος να είναι ισομετρία, δηλαδή*

$$\|T(e_i) - T(e_j)\|_2 = \|e_i - e_j\|_2 \text{ για κάθε } i \neq j.$$

Τότε, $k \geqslant N$.

Απόδειξη. Για κάθε $i \leqslant N$, θέτουμε $w_i = T(e_i) - T(0)$ και παρατηρούμε ότι ισχύουν οι $\|w_i\|_2 = 1$ και $\|w_i - w_j\|_2 = \sqrt{2}$ αν $i \neq j$. Γράφουμε τώρα

$$\|w_i - w_j\|_2^2 = \|w_i\|_2^2 + \|w_j\|_2^2 - 2\langle w_i, w_j\rangle,$$

απ' όπου λαμβάνουμε $\langle w_i, w_j\rangle = 0$ δηλαδή έχουμε N γραμμικώς ανεξάρτητα διανύσματα στον $\mathbb{R}^k$ και άρα $k \geqslant N$. □

Παρατηρούμε ότι το λήμμα Johnson-Lindenstrauss προσδιορίζει έναν γραμμικό τελεστή $T : \mathbb{R}^n \to \mathbb{R}^k$ με $k = c\epsilon^2 \log N$. Μπορούμε όμως να επιτύχουμε κάτι καλύτερο από αυτό; Η απάντηση είναι αρνητική καθώς όπως θα αποδείξουμε υπάρχει ένα σύνολο $V \subset \mathbb{R}^d$ ώστε για κάθε τελεστή $T : \mathbb{R}^d \to \mathbb{R}^k$ που διατηρεί τις αποστάσεις για $\epsilon = 1/2$, να έχουμε $k > c \cdot \log N$. Συνεπώς, η επιλογή του k της τάξης του $\log N$ είναι βέλτιστη. Πιο συγκεκριμένα, ισχύει το εξής.

Πρόταση 3.4.1. *Αν $V = \{u_1, u_2, \dots, u_d\}$ είναι μια ορθοκανονική βάση του $\mathbb{R}^d$, τότε για κάθε τελεστή $T : V \to \mathbb{R}^k$ με την ιδιότητα ότι*

$$\|u_i - u_j\|_2 \leqslant \|T(u_i) - T(u_j)\|_2 \leqslant 3\|u_i - u_j\|_2, \quad \text{για κάθε } i, j \leqslant d,$$

έπεται ότι $k = \Omega(\log |V|)$.

Απόδειξη. Η απόδειξη βασίζεται σε ένα ογκομετρικό επιχείρημα. Αρχικά παρατηρούμε ότι εφόσον τα διανύσματα της βάσης V έχουν υποτεθεί ορθογώνια ανά δύο αλλά και μοναδιαία, θα είναι $\|u_i - u_j\|_2 = \sqrt{2}$, για κάθε $u_i, u_j \in V$. Σταθεροποιούμε ένα διάνυσμα στη βάση V, έστω το u_1, και θεωρούμε τα διανύσματα $y_i = T(u_i) - T(u_1)$. Χρησιμοποιώντας την υπόθεση θα έχουμε $\|y_i\|_2 \leqslant 3\sqrt{2}$ αλλά και

$$\|y_i - y_j\|_2 = \|T(u_i) - T(u_j)\|_2 \geqslant \|u_i - u_j\|_2 = \sqrt{2}.$$

Παρατηρούμε τώρα ότι οι μπάλες με κέντρα τα y_i και ακτίνα $1/2$ είναι ανά δυό ξένες και περιέχονται στη μπάλα

$$B\Big(0, 3\sqrt{2} + \frac{1}{2}\Big) \subset B(0, 6),$$

και άρα αν $B_2^k := \{x \in \mathbb{R}^m : \|x\|_2 \leqslant 1\}$, θέτοντας $B_i := B(y_i, \frac{1}{2}) = y_i + \frac{1}{2}B_2^k$ έχουμε

$$\mathrm{vol}_k\Big(\bigcup_{i=1}^{d} B\Big(y_i, \frac{1}{2}\Big)\Big) \leqslant \mathrm{vol}_k(B(0,6)),$$

άρα

$$d \cdot \mathrm{vol}_k\Big(\frac{1}{2}B_2^k\Big) = \sum_{i=1}^{d} \mathrm{vol}_k(B_i) \leqslant \mathrm{vol}_k(6B_2^k),$$

δηλαδή

$$d \cdot 2^{-k}\mathrm{vol}_k(B_2^k) \leqslant 6^k \mathrm{vol}_k(B_2^k),$$

απ' όπου προκύπτει ότι $d \leqslant 12^k$. Συνεπώς, αν θέσουμε $c = 1/\log 12$ βλέπουμε ότι $k \geqslant c \cdot \log d$ απ' όπου έπεται το συμπέρασμα. □

Το ερώτημα αν το λήμμα Johnson-Lindenstrauss είναι βέλτιστο τόσο ως προς το πλήθος N των σημείων όσο και ως προς το ϵ, μελετήθηκε από τους K. Green-Larsen και J. Nelson, οι οποίοι απέδειξαν το εξής.

Θεώρημα 3.4.2. *Για κάθε $n > 1$ και $\epsilon \in (0, 1/2)$ υπάρχει $V \subset \mathbb{R}^n$ με $|V| = N = n^3 + n$ τέτοιο ώστε αν $T : \mathbb{R}^n \to \mathbb{R}^m$ είναι γραμμικός τελεστής με την ιδιότητα ότι*

$$(1-\epsilon)\|v_i\|_2^2 \leqslant \|T(v_i)\|_2^2 \leqslant (1+\epsilon)\|v_i\|_2^2$$

για κάθε $v_i \in V$, τότε $m = \Omega(\epsilon^{-2}\log n) = \Omega(\epsilon^{-2}\log N)$.

Σε ό,τι ακολουθεί, συμβολίζουμε με $g \in \mathbb{R}$ μια τυπική κανονική τυχαία μεταβλητή με τιμές στο $\mathbb{R}$, μέση τιμή 0 και διασπορά 1, και με $G \in \mathbb{R}^n$ ένα τυπικό τυχαίο διάνυσμα με τιμές στον $\mathbb{R}^n$ και πίνακα συνδιακυμάνσεων τον ταυτοτικό πίνακα (δηλαδή, $G = (g_1, \ldots, g_n)$ όπου οι $g_j \in \mathbb{R}$ είναι ανεξάρτητες). Γράφουμε $\pm B$ για οποιαδήποτε τιμή στο $[-B, B]$. Αν $T : \mathbb{R}^n \to \mathbb{R}^n$ είναι ένας γραμμικός τελεστής με $T_{i,j} = \langle T(e_i), e_j\rangle$, γράφουμε $\|T\|$ για τη νόρμα τελεστή του T (από τον ℓ_2^n στον ℓ_2^m) και $\|T\|_F = \left(\sum_{i,j} T_{i,j}^2\right)^{1/2}$ για τη νόρμα Frobenius του T.

Θα χρησιμοποιήσουμε το ακόλουθο θεώρημα του Latała, το οποίο δείχνει ότι η ανισότητα Hanson-Wright είναι ακριβής για κανονικές τυχαίες μεταβλητές.

Θεώρημα 3.4.3. *Υπάρχει απόλυτη σταθερά $c > 0$ τέτοια ώστε, αν $G \in \mathbb{R}^n$ τότε για κάθε $n \times n$ πραγματικό συμμετρικό πίνακα $T = (T_{i,j})$ με μηδενικά στη διαγώνιο ισχύει ότι*

$$\mathbb{P}\left(|\langle T(G), G\rangle| > c(\sqrt{t}\,\|T\|_F + t\,\|T\|)\right) \geqslant \min\{c, e^{-t}\}$$

για κάθε $t \geqslant 1$.

Από το Θεώρημα 3.4.3 παίρνουμε το εξής.

Πόρισμα 3.4.4. *Έστω G και T όπως στο Θεώρημα 3.4.3, μόνο που τώρα δεν υποθέτουμε ότι ο T έχει μηδενικά στη διαγώνιο. Τότε, για κάθε $t \geqslant 1$ ισχύει ότι*

$$\mathbb{P}\left(|\langle T(G), G\rangle - \operatorname{tr}(T)| > c(\sqrt{t}\,\|T\|_F + t\,\|T\|)\right) \geqslant \min\{c, e^{-t}\}.$$

Απόδειξη. Έστω N ένας θετικός ακέραιος. Θεωρούμε το τυπικό κανονικό τυχαίο διάνυσμα

$$G = (g_{1,1}, g_{1,2}, \dots, g_{1,N}, \dots, g_{n,1}, g_{n,2}, \dots, g_{n,N}).$$

Γράφουμε $G = (g_1, \dots, g_n)$ και παρατηρούμε ότι κάθε g_i έχει την ίδια κατανομή με την $N^{-1/2}\sum_{j=1}^N g_{i,j}$. Θεωρούμε επίσης τον $(nN)\times(nN)$ πίνακα T_N που σχηματίζεται αν μετατρέψουμε κάθε συντεταγμένη $T_{i,j}$ του T σε έναν $N\times N$ τετραγωνικό πίνακα με συντεταγμένες $T_{i,j}/N$. Τότε,

$$\begin{aligned} G^tTG - \operatorname{tr}(T) = \sum_{i=1}^n\sum_{j=1}^n T_{i,j}g_ig_j - \operatorname{tr}(T) &\overset{d}{=} \sum_{i=1}^n\sum_{j=1}^n\sum_{r=1}^N\sum_{s=1}^N \frac{T_{i,j}}{N} g_{i,r}g_{j,s} - \operatorname{tr}(T) \\ &\overset{\text{def}}{=} G^tT_NG - \operatorname{tr}(T_N). \end{aligned}$$

[Παρατηρήστε ότι $\operatorname{tr}(T) = \operatorname{tr}(T_N)$.] Από τον ασθενή νόμο των μεγάλων αριθμών, για κάθε $\lambda > 0$ έχουμε ότι

$$\lim_{N\to\infty} \mathbb{P}\left(\left|G^tT_NG - \operatorname{tr}(T_N)\right| > \lambda\right) = \lim_{N\to\infty} \mathbb{P}\left(\left|G^t(T_N - D_N)G\right| > \lambda\right), \tag{3.4}$$

όπου D_N είναι ο διαγώνιος πίνακας που συμπίπτει στη διαγώνιο με τον T_N. Παρατηρούμε ότι $\|T_N\| = \|T\|$. Πράγματι, αν θεωρήσουμε την αναπαράσταση $T = \sum_i \sigma_i u_i v_i^t$, όπου $\{u_i\}$ και $\{v_i\}$ είναι ορθοκανονικές βάσεις, $\sigma_i > 0$ και $\|T\| = \max_i \sigma_i$, τότε

$$T_N = \sum_i \sigma_i u_i^{(N)} (v_i^{(N)})^t,$$

όπου το $u_i^{(N)}$ προκύπτει από το u_i αν επαναλάβουμε κάθε συντεταγμένη N φορές και διαιρέσουμε με $\sqrt{N}$. Έπεται ότι

$$\left|\,\|T_N - D_N\| - \|T\|\,\right| \leqslant \|D_N\| = \max_i \frac{|T_{i,i}|}{N} = o_N(1)$$

από την τριγωνική ανισότητα. Συνεπώς,

$$\lim_{N\to\infty} \|T_N - D_N\| = \|T\|.$$

Επίσης,

$$\lim_{N\to\infty} \|T_N - D_N\|_F = \|T\|_F.$$

Εφαρμόζοντας το Θεώρημα 3.4.3 στο δεξιό μέλος της (3.4) παίρνουμε το ζητούμενο. □

Το επόμενο λήμμα είναι συνέπεια της Gaussian συγκέντρωσης για Lipschitz συναρτήσεις. Προκύπτει επίσης από την ανισότητα Hanson-Wright.

Λήμμα 3.4.5. *Υπάρχει απόλυτη σταθερά $c > 0$ τέτοια ώστε για κάθε τυπικό κανονικό τυχαίο διάνυσμα $G \in \mathbb{R}^n$ και κάθε $t > 0$,*

$$\mathbb{P}(\,|\,\|G_j\|_2^2 - n\,|\,> c\,\sqrt{nt}) < e^{-t}.$$

Συνδυάζοντας τα παραπάνω καταλήγουμε στο εξής.

Πόρισμα 3.4.6. *Έστω $T \in \mathbb{R}^{d\times n}$ και $\lambda_1 \geqslant \cdots \geqslant \lambda_n \geqslant 0$ οι ιδιοτιμές του T^tT. Έστω $G_1, \ldots, G_N \in \mathbb{R}^n$ ανεξάρτητα τυπικά κανονικά τυχαία διανύσματα. Για κάποιες απόλυτες σταθερές $c_1, c_2, \delta_0 > 0$ και για κάθε $0 < \delta < \delta_0$,*

$$\mathbb{P}(\nexists j \in n : I_{\delta,j} \wedge II_{\delta,j}) \leqslant e^{\delta N}, \tag{3.5}$$

όπου $I_{\delta,j}$ και $II_{\delta,j}$ είναι οι ανισότητες

$$\Big|\,\|TG_j\|_2^2 - \sum_{i=1}^n \lambda_i\Big| \geqslant c_1\,\sqrt{\log(1/\delta)}\Big(\sum_{i=1}^n \lambda_i^2\Big)^{1/2}$$

και

$$|\,\|G_j\|_2^2\,| \leqslant c_2\,\sqrt{n\log(1/\delta)}$$

αντίστοιχα.

Απόδειξη. Θα δείξουμε ότι για κάθε σταθερό $j \in [N]$ ισχύει ότι

$$\mathbb{P}(I_{\delta,j} \wedge II_{\delta,j}) > \delta. \tag{3.6}$$

Τότε, αφού τα G_j είναι ανεξάρτητα, το αριστερό μέλος της (3.5) φράσσεται από $(1-\delta)^N \leqslant e^{-\delta N}$.

Για την απόδειξη της (3.6) αρκεί να δείξουμε ότι

$$\mathbb{P}(|\,\|G_j\|_2^2\,| \leqslant c_2\,\sqrt{n\log(1/\delta)}) > 1 - \frac{\delta}{2} \tag{3.7}$$

και

$$\mathbb{P}\left(\Big|\,\|TG_j\|_2^2 - \sum_{i=1}^n \lambda_i\Big| \geqslant c_1\,\sqrt{\log(1/\delta)}\Big(\sum_{i=1}^n \lambda_i^2\Big)^{1/2}\right) > \frac{3\delta}{2}. \tag{3.8}$$

Στη συνέχεια η (3.6) θα προκύψει με εφαρμογή του «union bound». Η (3.7) προκύπτει άμεσα από το Λήμμα 3.4.5 αν επιλέξουμε τη σταθερά c_2 αρκετά μεγάλη. Για την (3.8) παρατηρούμε ότι $\|TG_j\|_2^2 = G_j^tT^tTG_j$. Επίσης, $\sum_i \lambda_i = \mathrm{tr}(T^tT)$ και $\left(\sum_i \lambda_i^2\right)^{1/2} = \|T^tT\|_F$, οπότε η (3.8) προκύπτει από το Πόρισμα 3.4.4 για $\delta > 0$ μικρότερο από κατάλληλα μικρή σταθερά δ_0. □

Θα χρειαστούμε επίσης μια στοιχειώδη ανισότητα για αριθμούς κάλυψης.

Λήμμα 3.4.7. *Για κάθε $0 < \alpha < 1$ υπάρχει οικογένεια $\mathcal{F}_\alpha \subset \bigcup_{m=1}^n \mathbb{R}^{m\times n}$ πινάκων με τις ακόλουθες ιδιότητες:*

$$|\mathcal{F}_\alpha| = e^{O(n^2 \log(n/\alpha))}$$

και, για κάθε πίνακα $T \in \bigcup_{m=1}^n \mathbb{R}^{m\times n}$ που έχει όλες τις συντεταγμένες του απολύτως φραγμένες από 2, υπάρχει πίνακας $T' \in \mathcal{F}_\alpha$ τέτοιος ώστε οι T, T' να έχουν το ίδιο πλήθος γραμμών και ο $S = T - T'$ να ικανοποιεί την $\mathrm{tr}(S^t S) \leqslant \alpha/100$.

Απόδειξη. Ορίζουμε την οικογένεια $\mathcal{F}_\alpha$ ως εξής: Για κάθε ακέραιο $1 \leqslant m \leqslant n$ συμπεριλαμβάνουμε όλους τους $m \times n$ πίνακες που έχουν συντεταγμένες της μορφής $i\frac{\sqrt{\alpha}}{10n}$ για ακεραίους $i \in \{-20n/\sqrt{\alpha}, \ldots, 20n/\sqrt{\alpha}\}$. Τότε, για κάθε πίνακα $T \in \bigcup_{m=1}^n \mathbb{R}^{m\times n}$ υπάρχει πίνακας T' τέτοιος ώστε οι T και T' να έχουν το ίδιο πλήθος ακμών και κάθε συντεταγμένη του $S = T - T'$ να είναι απολύτως φραγμένη από $\frac{\sqrt{\alpha}}{10n}$. Αυτό σημαίνει ότι κάθε διαγώνια συντεταγμένη του $S^t S$ φράσσεται από $n\alpha/(100n^2)$, επομένως $\mathrm{tr}(S^t S) \leqslant \alpha/100$. Τέλος, παρατηρούμε ότι ο πληθάριθμος της οικογένειας $\mathcal{F}_\alpha$ φράσσεται από $n(40n/\sqrt{\alpha})^{n^2} = e^{O(n^2 \log(n/\alpha))}$. □

Λήμμα 3.4.8. *Έστω $\mathcal{F}_\alpha$ όπως στο Λήμμα 3.4.7 με $1/poly(n) \leqslant \alpha < 1$. Τότε, υπάρχει σύνολο $N = O(n^3)$ σημείων $v_1, \ldots, v_N \in \mathbb{R}^n$ τέτοιο ώστε για κάθε πίνακα $T \in \mathcal{F}_\alpha$ να υπάρχει $j \in [N]$ ώστε*

$$\left| \|T(v_j)\|_2^2 - \sum_i \lambda_i \right| = \Omega\left(\sqrt{\log n \cdot \sum_i \lambda_i^2} \right)$$

και

$$\left| \|v_j\|_2^2 - n \right| = O(\sqrt{n \log n}).$$

Απόδειξη. Έστω $G_1, \ldots, G_N \in \mathbb{R}^n$ ανεξάρτητα τυπικά κανονικά τυχαία διανύσματα. Θεωρούμε $T \in \mathcal{F}_\alpha$ και εφαρμόζουμε το Πόρισμα 3.4.6 με $\delta = n^{-1/4} = N^{-1/12}$. Με πιθανότητα $1 - e^{-\Omega(n^{3-1/4})}$, κάποιο από τα G_j ικανοποιεί το συμπέρασμα του λήμματος για τον συγκεκριμένο T. Αφού $|\mathcal{F}_\alpha| = e^{O(n^2 \log(n/\alpha))}$, ο ισχυρισμός του λήμματος προκύπτει με εφαρμογή του «union bound» πάνω από όλους τους $T \in \mathcal{F}_\alpha$. □

Απόδειξη του Θεωρήματος 3.4.2. Μπορούμε να υποθέσουμε ότι $\epsilon > 1/\sqrt{n}$ (αλλιώς ζητάμε φράγμα της μορφής $m = \Omega(n)$, το οποίο εξασφαλίζεται εύκολα). Για να κατασκευάσουμε το V, χρησιμοποιούμε αρχικά το Λήμμα 3.4.8 με $\alpha = \epsilon^2/n^2$ για να βρούμε n^3 διανύσματα $w_1, \ldots, w_{n^3}$ τέτοια ώστε για κάθε πίνακα $\tilde{T} \in \mathcal{F}_{\epsilon^2/n^2}$ να υπάρχει δείκτης $j \in [n^3]$ που ικανοποιεί τις

1. $\left| \|\tilde{T} w_j\|_2^2 - \sum_i \tilde{\lambda}_i \right| \geqslant \Omega\left(\sqrt{(\log n) \sum_i \tilde{\lambda}_i^2} \right)$,
2. $\left| \|w_j\|_2^2 - n \right| = O\left(\sqrt{n \log n} \right)$,

όπου $\tilde{\lambda}_1 \geqslant \cdots \geqslant \tilde{\lambda}_n \geqslant 0$ είναι οι ιδιοτιμές του $\tilde{T}^t \tilde{T}$. Ορίζουμε

$$V = \{e_1, \ldots, e_n, w_1, \ldots, w_{n^3}\}$$

και ισχυριζόμαστε ότι αυτό το σύνολο $N = n^3 + n$ διανυσμάτων ικανοποιεί το συμπέρασμα του θεωρήματος (τα e_i είναι τα διανύσματα της συνήθους ορθοκανονικής βάσης του $\mathbb{R}^n$).

Για να αποδείξουμε τον ισχυρισμό, θεωρούμε έναν πίνακα $T \in \mathbb{R}^{m\times n}$, όπου $m \leqslant n$, που ικανοποιεί την $\|Tv\|_2^2 \in (1\pm\epsilon)\|v\|_2^2$ για κάθε $v \in V$. Παρατηρούμε ότι, αφού $e_1, \dots, e_n \in V$, ο T ικανοποιεί την

$$\|Te_i\|_2^2 \in (1\pm\epsilon)\|e_i\|_2^2 = 1\pm\epsilon$$

για όλα τα e_i. Συνεπώς, όλες οι συντεταγμένες $T_{i,j}$ του T πρέπει να ικανοποιούν την

$$T_{i,j}^2 \leqslant 1+\epsilon < 2,$$

και μάλιστα όλες οι στήλες του T έχουν ℓ_2-νόρμα το πολύ ίση με $\sqrt{2}$. Έπεται ότι υπάρχει ένας $m\times n$ πίνακας $\widehat{T} \in \mathcal{F}_{\epsilon^2/n^2}$ τέτοιος ώστε ο $B = T - \widehat{T} = (S_{i,j})$ να ικανοποιεί την

$$\operatorname{tr}(S^tS) \leqslant \epsilon^2/(100n^2).$$

Αφού $\operatorname{tr}(S^tS) = \|S\|_F^2$, έπεται επίσης ότι $\|S\|_F \leqslant \epsilon/(10n)$. Έτσι, από την ανισότητα Cauchy-Schwarz,

$$\begin{aligned}
\sum_{i=1}^n \widehat{\lambda_i} &= \operatorname{tr}(\widehat{T}^t\widehat{T}) \\
&= \operatorname{tr}((T-S)^t(T-S)) \\
&= \operatorname{tr}(T^tT) + \operatorname{tr}(S^tS) - \operatorname{tr}(T^tS) - \operatorname{tr}(S^tT) \\
&= \sum_{i=1}^n \|T(e_i)\|_2^2 + \operatorname{tr}(S^tS) - \operatorname{tr}(T^tS) - \operatorname{tr}(S^tT) \\
&= n \pm \left(O(\epsilon n) + 2n\cdot \max_j \Big(\sum_i S_{i,j}^2\Big)^{1/2} \cdot \max_k \Big(\sum_i T_{i,k}^2\Big)^{1/2} \right) \\
&= n \pm \left(O(\epsilon n) + 2n\cdot \|S\|_F \cdot \sqrt{2} \right) \\
&= n \pm O(\epsilon n).
\end{aligned}$$

Από την επιλογή του V υπάρχει διάνυσμα $v^* \in V$ που ικανοποιεί τις

(i) $\Big| \|\widehat{T}v^*\|_2^2 - n \Big| \geqslant \Omega\Big(\sqrt{(\log n)\sum_i \widehat{\lambda_i}^2} \Big) - O(\epsilon n)$,

(ii) $\Big| \|v^*\|_2^2 - n \Big| = O\Big(\sqrt{n\log n} \Big)$.

Παρατηρούμε ότι

$$\|S\|^2 \leqslant \|S\|_F^2 = \operatorname{tr}(S^tS) \leqslant \epsilon^2/(100n^2)$$

και

$$\|\widehat{T}\|^2 \leqslant \|\widehat{T}\|_F^2 \leqslant (\|T\|_F + \|S\|_F)^2 = O(n^2).$$

Τότε, από την (i) παίρνουμε την (iii):

$$\begin{aligned}\left|\,\|Tv^*\|_2^2 - n\,\right| &= \left|\,\|\widehat{T}v^*\|_2^2 + \|Sv^*\|_2^2 + 2\langle \widehat{T}v^*, Sv^*\rangle - n\,\right| \\ &\geqslant \Omega\left(\sqrt{(\log n)\sum_i \widehat{\lambda}_i^2}\right) - \|Sv^*\|_2^2 - 2|\langle \widehat{T}v^*, Sv^*\rangle| - O(\epsilon n) \\ &\geqslant \Omega\left(\sqrt{(\log n)\sum_i \widehat{\lambda}_i^2}\right) - \|S\|^2\cdot\|v^*\|_2^2 - 2\|\widehat{T}\|\cdot\|S\|\cdot\|v^*\|_2^2 - O(\epsilon n) \\ &= \Omega\left(\sqrt{(\log n)\sum_i \widehat{\lambda}_i^2}\right) - O(\epsilon n).\end{aligned}$$

Υποθέσαμε ότι

$$\left|\,\|Tv^*\|_2^2 - \|v^*\|_2^2\,\right| = O(\epsilon\|v^*\|_2^2) = O(\epsilon n).$$

Επομένως, από την (ii),

$$\left|\,\|Tv^*\|_2^2 - n\,\right| \leqslant \left|\,\|Tv^*\|_2^2 - \|v^*\|_2^2\,\right| + \left|\,\|v^*\|_2^2 - n\,\right| = O(\epsilon n + \sqrt{n\log n}),$$

και συνδυάζοντας αυτήν την ανισότητα με την (iii) παίρνουμε

$$\sum_{i=1}^n \widehat{\lambda}_i^2 = O\left(\frac{\epsilon^2 n^2}{\log n} + n\right).$$

Από την ανισότητα Cauchy-Schwarz, και αφού ακριβώς $\operatorname{rank}(\widehat{T}^t\widehat{T})$ από τους $\widehat{\lambda}_i$ είναι μη μηδενικοί, βλέπουμε ότι

$$\frac{n^2}{2} \leqslant \left(\sum_{i=1}^n \widehat{\lambda}_i\right)^2 \leqslant \operatorname{rank}(\widehat{T}^t\widehat{T})\cdot\left(\sum_{i=1}^n \widehat{\lambda}_i^2\right) \leqslant m\cdot O\left(\frac{\epsilon^2 n^2}{\log n} + n\right).$$

Λύνοντας ως προς m συμπεραίνουμε ότι

$$m = \Omega(\min\{n, \epsilon^{-2}\log n\}) = \Omega(\min\{n, \epsilon^{-2}\log n\}),$$

όπως θέλαμε. □

3.5 Αναγωγή της διάστασης στον L_1

Σε αυτήν την ενότητα αποδεικνύουμε ότι δεν υπάρχει ένα ανάλογο του λήμματος των Johnson και Lindenstrauss για τον L_1. Για τον σκοπό αυτό χρησιμοποιούμε τα λεγόμενα diamond graphs. Βασικό ρόλο θα παίξει επίσης η ανισότητα της Πρότασης 3.5.1 για συναρτήσεις του L_p $(1 < p < 2)$.

Το λήμμα της ομοιόμορφης κυρτότητας

Στα επόμενα σταθεροποιούμε ένα $p \in (1,2)$.

Πρόταση 3.5.1. *Για κάθε $a, b \in L_p$ ισχύει*

$$\|a+b\|_p^2 + (p-1)\|a-b\|_p^2 \leqslant 2\|a\|_p^2 + 2\|b\|_p^2.$$

Παρατήρηση. Το αποτέλεσμα ισχύει και για $p = 1$, ενώ για $p = 2$ ισχύει ως ισότητα και είναι ο κανόνας του παραλληλογράμμου.

Λήμμα 3.5.2. *Θεωρούμε τις συναρτήσεις $\alpha, \beta : [0,1] \to \mathbb{R}$ με*

$$\alpha(r) = (1+r)^{p-1} + (1-r)^{p-1}$$

και

$$\beta(r) = \frac{(1+r)^{p-1} - (1-r)^{p-1}}{r^{p-1}}.$$

Τότε, για κάθε $A, B \in \mathbb{R}$ έχουμε

$$\alpha(r)|A|^p + \beta(r)|B|^p \leqslant |A+B|^p + |A-B|^p. \tag{3.9}$$

Απόδειξη. Κατ' αρχάς, θα έχουμε

$$\begin{aligned}
\beta(0) &= \lim_{r\to 0^+} \frac{(1+r)^{p-1} - (1-r)^{p-1}}{r^{p-1}} \\
&= \lim_{r\to 0^+} \frac{(p-1)(1+r)^{p-2} - (p-1)(1-r)^{p-2}}{(p-1)r^{p-2}} \\
&= \lim_{r\to 0^+} r^{2-p}\left[\frac{1}{(1+r)^{2-p}} + \frac{1}{(1-r)^{2-p}}\right] = 0 \qquad \text{(διότι } p < 2\text{)}.
\end{aligned}$$

Επίσης,

$$\alpha'(r) = (p-1)\left[\frac{1}{(1+r)^{2-p}} - \frac{1}{(1-r)^{2-p}}\right] \leqslant 0$$

για κάθε r. Έτσι, η α είναι φθίνουσα. Επομένως,

$$\alpha(1) \leqslant \alpha(r) \leqslant \alpha(0), \qquad \text{δηλαδή} \qquad 2^{p-1} \leqslant \alpha(r) \leqslant 2.$$

Ακόμη,

$$\beta'(r) = (p-1)\frac{1}{r^p}\left[\frac{1}{(1-r)^{2-p}} - \frac{1}{(1+r)^{2-p}}\right] \geqslant 0$$

για κάθε r. Επομένως, η β είναι αύξουσα. Συνεπώς,

$$\beta(0) \leqslant \beta(r) \leqslant \beta(1), \qquad \text{δηλαδή} \qquad 0 \leqslant \beta(r) \leqslant 2^{p-1}.$$

Η (3.9) ισχύει προφανώς, όταν $A = B = 0$. Για $A = 0$ γίνεται $\beta(r)|B|^p \leqslant 2|B|^p$, που αληθεύει.

Για $B = 0$ γίνεται $\alpha(r)|A|^p \leqslant 2|A|^p$, που επίσης αληθεύει.

Μπορούμε λοιπόν να υποθέσουμε ότι $A \neq 0$ και $B \neq 0$. Μάλιστα, μπορούμε να υποθέσουμε ότι $A > 0$ και $B > 0$ (παρατηρήστε ότι $||A| + |B||^p + ||A| - |B||^p = |A + B|^p + |A - B|^p$).

Για να συνεχίσουμε παρατηρούμε ότι $\alpha(r) \geqslant \beta(r)$ για κάθε r. Συνεπώς, αν $0 < A < B$ τότε $\alpha(r)(B^p - A^p) \geqslant \beta(r)(B^p - A^p)$, δηλαδή

$$\alpha(r)A^p + \beta(r)B^p \leqslant \alpha(r)B^p + \beta(r)A^p.$$

Έτσι, λόγω και της συμμετρίας του δεύτερου μέλους, αρκεί να αποδείξουμε την (3.9) όταν ο συντελεστής του $\alpha(r)$ είναι μεγαλύτερος ή ίσος από τον συντελεστή του $\beta(r)$. Δηλαδή, θα δείξουμε ότι αν $0 < B \leqslant A$ τότε

$$\alpha(r)A^p + \beta(r)B^p \leqslant (A + B)^p + (A - B)^p$$

ισοδύναμα,

$$\alpha(r) + \beta(r)\left(\frac{B}{A}\right)^p \leqslant \left(1 + \frac{B}{A}\right)^p + \left(1 - \frac{B}{A}\right)^p.$$

Θέτοντας $R = B/A$ βλέπουμε ότι αρκεί να δείξουμε την

$$\alpha(r) + \beta(r)R^p \leqslant (1 + R)^p + (1 - R)^p$$

για κάθε $R \in [0, 1]$.

Θέτουμε $F(r) = \alpha(r) + \beta(r)R^p$ και παρατηρούμε ότι

$$F(R) = \alpha(R) + \beta(R)R^p = (1 + R)^p - (1 - R)^p.$$

Αρκεί λοιπόν να δείξουμε ότι η συνάρτηση $F(r)$ $(0 \leqslant r \leqslant 1)$ έχει ολικό μέγιστο για $r = R$. Έχουμε:

$$F'(r) = (p - 1)\left[\frac{1}{(1 - r)^{2-p}} - \frac{1}{(1 + r)^{2-p}}\right]\left(\frac{R^p}{r^p} - 1\right)$$

και $F'(r) \geqslant 0$ αν και μόνο αν $r \leqslant R$. Έτσι, η F είναι αύξουσα στο $[0, R]$ και φθίνουσα στο $[R, 1]$. Στο R έχει ολικό μέγιστο και η απόδειξη ολοκληρώθηκε. □

Λήμμα 3.5.3. *Για κάθε $f, g \in L_p$ ισχύει:*

$$(\|f\|_p + \|g\|_p)^p + \left|\, \|f\|_p - \|g\|_p \,\right|^p \leqslant \|f + g\|_p^p + \|f - g\|_p^p.$$

Απόδειξη. Αν $f = 0$ ή $g = 0$, τότε ισχύει το αποτέλεσμα. Έστω $f \neq 0$ και $g \neq 0$. Θέτουμε

$$r = \frac{\|g\|_p}{\|f\|_p}.$$

Λόγω συμμετρίας, μπορούμε να υποθέσουμε ότι $r \leqslant 1$. Από το Λήμμα 3.5.2 προκύπτει ότι

$$\alpha(r)|f|^p + \beta(r)|g|^p \leqslant |f+g|^p + |f-g|^p,$$

άρα

$$\alpha(r)\int |f|^p + \beta(r)\int |g|^p \leqslant \int |f+g|^p + \int |f-g|^p,$$

δηλαδή,

$$\alpha(r)\|f\|_p^p + \beta(r)\|g\|_p^p \leqslant \|f+g\|_p^p + \|f-g\|_p^p.$$

Από τον ορισμό του r,

$$\alpha\left(\frac{\|g\|_p}{\|f\|_p}\right)\|f\|_p^p + \beta\left(\frac{\|g\|_p}{\|f\|_p}\right)\|g\|_p^p \leqslant \|f+g\|_p^p + \|f-g\|_p^p.$$

Από τον ορισμό των α, β και την τελευταία ανισότητα, καταλήγουμε στην

$$(\|f\|_p + \|g\|_p)^p + (\|f\|_p - \|g\|_p)^p \leqslant \|f+g\|_p^p + \|f-g\|_p^p.$$

□

Λήμμα 3.5.4. *Για κάθε $a, b \in \mathbb{R}$ ισχύει*

$$[a^2 + (p-1)b^2]^{1/2} \leqslant \left(\frac{|a+b|^p + |a-b|^p}{2}\right)^{1/p}.$$

Απόδειξη. Αν $a = 0$ και $b = 0$, προφανώς ισχύει. Αν $a = 0$, τότε γίνεται $\sqrt{p-1}|b| \leqslant |b|$ που ισχύει διότι $1 < p < 2$. Αν $b = 0$, τότε έχουμε $|a| \leqslant |a|$ που ισχύει.

Έτσι, μπορούμε να θεωρήσουμε $a, b \neq 0$ και μάλιστα $a, b > 0$ (χρησιμοποιώντας την $||a| + |b||^p + ||a| - |b||^p = |a+b|^p + |a-b|^p$).

Παρατηρούμε επίσης ότι αν $0 < a < b$ τότε $a^2 + (p-1)b^2 \leqslant b^2 + (p-1)a^2$. Έτσι, αρκεί να δείξουμε την ανισότητα όταν $0 < b \leqslant a$. Τότε έχουμε

$$[a^2 + (p-1)b^2]^{1/2} \leqslant \left(\frac{(a+b)^p + (a-b)^p}{2}\right)^{1/p}$$

αν και μόνο αν

$$\left(1 + (p-1)\left(\frac{b}{a}\right)^2\right)^{1/2} \leqslant \left(\frac{(1+\frac{b}{a})^p + (1-\frac{b}{a})^p}{2}\right)^{1/p}$$

δηλαδή, αν και μόνο αν

$$(1 + (p-1)x^2)^{1/2} \leqslant \left(\frac{(1+x)^p + (1-x)^p}{2}\right)^{1/p}$$

για $0 < x \leqslant 1$. Ισοδύναμα, ζητάμε

$$(1+(p-1)x^2)^{p/2} \leqslant \frac{(1+x)^p + (1-x)^p}{2}$$

για $0 < x \leqslant 1$. Τώρα,

$$\begin{aligned} \frac{(1+x)^p + (1-x)^p}{2} &= \frac{1}{2}\left(\sum_{k=0}^{\infty}\binom{p}{k}x^k + \sum_{k=0}^{\infty}\binom{p}{k}(-x)^k\right) \\ &= \sum_{k=0}^{\infty}\binom{p}{2k}x^{2k} \geqslant \binom{p}{0} + \binom{p}{2}x^2 \\ &= 1 + \frac{p(p-1)}{2}x^2, \end{aligned}$$

διότι $\binom{p}{2k} > 0$ για κάθε $k \geqslant 2$, αφού $p < 2$.

Τέλος,

$$1 + \frac{p(p-1)}{2}x^2 \geqslant (1+(p-1)x^2)^{p/2}$$

άρα

$$(1+(p-1)x^2)^{p/2} \leqslant \frac{(1+x)^p + (1-x)^p}{2}$$

όπως ζητούσαμε. □

Σημείωση. Η ανίσωση $1 + \frac{p(p-1)}{2}x^2 \geqslant (1+(p-1)x^2)^{p/2}$ προκύπτει από την $1 + at \geqslant (1+t)^a$ που ισχύει για $a, t \in [0,1]$.

Απόδειξη της Πρότασης 3.5.1. Έστω $x, y \in L_p$. Τότε,

$$\begin{aligned} \left(\frac{\|x+y\|_p^2 + \|x-y\|_p^2}{2}\right)^{1/2} &\geqslant \left(\frac{\|x+y\|_p^p + \|x-y\|_p^p}{2}\right)^{1/p} \\ &\geqslant \left[\frac{(\|x\|_p + \|y\|_p)^p + |\,\|x\|_p - \|y\|_p\,|^p}{2}\right]^{1/p} \\ &\geqslant \left(\|x\|_p^2 + (p-1)\|y\|_p^2\right)^{1/2}. \end{aligned}$$

Η πρώτη ανισότητα προκύπτει διότι $p < 2$, η δεύτερη από το Λήμμα 3.5.3 και η τρίτη από το Λήμμα 3.5.4.

Άρα,

$$\|x\|_p^2 + (p-1)\|y\|_p^2 \leqslant \frac{\|x+y\|_p^2 + \|x-y\|_p^2}{2}.$$

Θέτουμε $x = a + b$ και $y = a - b$. Τότε παίρνουμε

$$\|a+b\|_p^2 + (p-1)\|a-b\|_p^2 \leqslant 2\|a\|_p^2 + 2\|b\|_p^2,$$

όπως ζητούσαμε. □

Σημείωση. Η πρώτη ανίσωση στην απόδειξη της Πρότασης 3.5.1 είναι αποτέλεσμα της

$$(*) \qquad (\theta a^p + (1-\theta)b^p)^{1/p} \leqslant (\theta a^q + (1-\theta)b^q)^{1/q}$$

όταν $0 < p < q$, $a, b > 0$ και $0 < \theta < 1$.

Απόδειξη της $(*)$. Θεωρούμε την $f(x) = x^{q/p}$, με $x > 0$. Από την $p < q$ έπεται ότι η f είναι κυρτή, οπότε

$$f(\theta x + (1-\theta)y) \leqslant \theta f(x) + (1-\theta)f(y),$$

δηλαδή

$$(\theta x + (1-\theta)y)^{q/p} \leqslant \theta x^{q/p} + (1-\theta)y^{q/p}$$

για κάθε $x, y > 0$. Παίρνουμε $x = a^p$, $y = b^p$ και έχουμε

$$(\theta a^p + (1-\theta)b^p)^{q/p} \leqslant \theta a^q + (1-\theta)b^q,$$

άρα

$$(\theta a^p + (1-\theta)b^p)^{1/p} \leqslant (\theta a^q + (1-\theta)b^q)^{1/q}.$$

□

Diamond Graphs

Πρόκειται για μια ακολουθία γραφημάτων $\{G_k\}_{k=0}^{\infty}$.

(α) Το G_0 αποτελείται από δύο κορυφές s και t και μια ακμή μήκους 1. Η $\{s, t\}$ λέγεται και ακμή επιπέδου 0.

(β) Το G_1 προκύπτει από το G_0 αν αντικαταστήσουμε την ακμή $\{s, t\}$ με ένα τετράπλευρο (ρόμβο) s, a, t, b με μήκος ακμών 1/2. Οι ακμές $\{s, a\}$, $\{a, t\}$, $\{b, t\}$, $\{s, b\}$ λέγονται ακμές επιπέδου 1. Το σύνολο $\{a, b\}$ λέγεται αντιακμή επιπέδου 1 που αντιστοιχεί στην ακμή $\{s, t\}$.

(β) Το G_2 προκύπτει από το G_1 αν αντικαταστήσουμε κάθε ακμή του G_1 με ένα τετράπλευρο (ρόμβο) με μήκος ακμών 1/4. Τα σύνολα $\{a_i, b_i\}$ $(i = 1, 2, 3, 4)$ λέγονται αντιακμές επιπέδου 2 που αντιστοιχούν στις $\{s, a\}$, $\{a, t\}$, $\{t, b\}$, $\{s, b\}$. Οι ακμές του γραφήματος λέγονται ακμές επιπέδου 2.

Επαγωγικά, αν έχουμε το G_{i-1}, τότε κατασκευάζουμε το G_i αντικαθιστώντας κάθε ακμή $\{u, v\}$ του G_{i-1} με ένα τετράπλευρο u, a, v, b του οποίου οι ακμές έχουν μήκος $1/2^i$. Το σύνολο $\{a, b\}$ λέγεται αντιακμή επιπέδου i που αντιστοιχεί στην ακμή $\{u, v\}$. Η $\{u, v\}$ είναι ακμή επιπέδου $i - 1$.

Εμφύτευση των diamond graphs σε χώρους με νόρμα

Το ακόλουθο θεώρημα μας δίνει ένα κάτω φράγμα για την παραμόρφωση με την οποία το G_k εμφυτεύεται στον L_p για $1 < p \leqslant 2$.

Θεώρημα 3.5.5. *Έστω $1 < p \leqslant 2$ και $f : G_k \to L_p$ μια εμφύτευση του G_k στον L_p. Τότε, η παραμόρφωση της f είναι τουλάχιστον*

$$\sqrt{1 + (p-1)k}.$$

Λήμμα 3.5.6. *Έστω $x, y, z, w \in L_p$ $(1 < p \leqslant 2)$. Τότε,*

$$\|y-z\|_p^2 + (p-1)\|x-w\|_p^2 \leqslant \|x-y\|_p^2 + \|y-w\|_p^2 + \|w-z\|_p^2 + \|z-x\|_p^2.$$

Απόδειξη. Από την Πρόταση 3.5.1 έχουμε ότι

$$\|y-z\|_p^2 + (p-1)\|y-2w+z\|_p^2 \leqslant 2\|y-w\|_p^2 + 2\|w-z\|_p^2$$

και

$$\|y-z\|_p^2 + (p-1)\|y-2x+z\|_p^2 \leqslant 2\|x-y\|_p^2 + 2\|z-x\|_p^2.$$

Προσθέτοντας κατά μέλη και διαιρώντας με το 2 έχουμε

$$\begin{aligned}\|y-z\|_p^2 + (p-1)&\frac{\|y-2w+z\|_p^2 + \|y-2x+z\|_p^2}{2} \\ &\leqslant \|x-y\|_p^2 + \|y-w\|_p^2 + \|w-z\|_p^2 + \|z-x\|_p^2.\end{aligned}$$

Η απόδειξη ολοκληρώνεται αν παρατηρήσουμε ότι

$$\begin{aligned}\|x-w\|_p^2 &= \left(\frac{\|2x-2w\|_p}{2}\right)^2 = \frac{\|2x-y-z+y+z-2w\|_p^2}{4} \\ &\leqslant \frac{(\|y-2x+z\|_p + \|y-2w+z\|_p)^2}{4} \\ &\leqslant \frac{\|y-2x+z\|_p^2 + \|y-2w+z\|_p^2}{2}.\end{aligned}$$

□

Λήμμα 3.5.7. *Έστω $V(G_0) = \{s,t\}$ και A_i το σύνολο των αντιακμών επιπέδου i. Τότε, για κάθε $f : G_k \to L_p$ $(1 < p \leqslant 2)$ ισχύει:*

$$\|f(s)-f(t)\|_p^2 + (p-1)\sum_{i=1}^{k}\sum_{\{x,y\}\in A_i}\|f(x)-f(y)\|_p^2 \leqslant \sum_{\{x,y\}\in E(G_k)}\|f(x)-f(y)\|_p^2.$$

Απόδειξη. Έστω $\{u,v\}$ μια ακμή επιπέδου i και έστω $\{a,b\}$ η αντίστοιχη αντιακμή. Από το Λήμμα 3.5.6 ισχύει:

$$\begin{aligned}\|f(u)-f(v)\|_p^2 + (p-1)\|f(a)-f(b)\|_p^2 \leqslant\ & \|f(u)-f(a)\|_p^2 + \|f(a)-f(v)\|_p^2 \\ & + \|f(v)-f(b)\|_p^2 + \|f(b)-f(u)\|_p^2.\end{aligned}$$

Γράφουμε την παραπάνω σχέση για όλες τις ακμές επιπέδου $i = 0, 1, \ldots, k-1$ και τις αντίστοιχες αντιακμές τους. Με άθροιση κατά μέρη προκύπτει το αποτέλεσμα, αν παρατηρήσουμε ότι οι όροι που αντιστοιχούν στις ακμές επιπέδου $i = 1, 2, \ldots, k-1$ διαγράφονται από τα δύο μέλη. □

Απόδειξη του Θεωρήματος 3.5.5. Έστω $f : G_k \to L_p$ μια εμφύτευση του G_k στον L_p. Υποθέτουμε ότι η f έχει παραμόρφωση γ. Τότε,

$$\frac{1}{\gamma} d_k(x,y) \leqslant \|f(x) - f(y)\|_p \leqslant d_k(x,y),$$

όπου d_k είναι η μετρική (του συντομότερου μονοπατιού) στο G_k. Από το Λήμμα 3.5.7 έχουμε ότι

$$\|f(s) - f(t)\|_p^2 + (p-1) \sum_{i=1}^{k} \sum_{\{x,y\} \in A_i} \|f(x) - f(y)\|_p^2 \leqslant \sum_{\{x,y\} \in E(G_k)} \|f(x) - f(y)\|_p^2.$$

Αλλά

$$\sum_{\{x,y\} \in E(G_k)} \|f(x) - f(y)\|_p^2 \leqslant \sum_{\{x,y\} \in E(G_k)} (d_k(x,y))^2 = \left(\frac{1}{2^k}\right)^2 |E(G_k)| = \frac{1}{4^k} 4^k = 1.$$

Ακόμη,

$$\|f(s) - f(t)\|_p^2 \geqslant \frac{1}{\gamma^2}$$

και

$$\begin{aligned} \sum_{\{x,y\} \in A_i} \|f(x) - f(y)\|_p^2 &\geqslant \sum_{\{x,y\} \in A_i} \frac{1}{\gamma^2} (d_k(x,y))^2 \\ &= \frac{1}{\gamma^2} \sum_{\{x,y\} \in A_i} \left(\frac{1}{2^{i-1}}\right)^2 = \frac{1}{\gamma^2} \frac{1}{4^{i-1}} |A_i| \\ &= \frac{1}{\gamma^2} \frac{1}{4^{i-1}} 4^{i-1} = \frac{1}{\gamma^2}. \end{aligned}$$

Άρα,

$$\frac{1}{\gamma^2} + (p-1)k \frac{1}{\gamma^2} \leqslant 1,$$

απ' όπου βλέπουμε ότι

$$\gamma \geqslant \sqrt{1 + (p-1)k}.$$

□

Για το πόρισμα που ακολουθεί, θεωρούμε γνωστό το εξής αποτέλεσμα:

> Υπάρχει μια σταθερά $c \geqslant 1$ ώστε κάθε G_k εμφυτεύεται με παραμόρφωση c στον L_1.

Πόρισμα 3.5.8. *Για κάθε $n \in \mathbb{N}$ υπάρχει $X \subset L_1$ με n σημεία, τέτοιο ώστε για κάθε $D > 1$, αν ο X εμφυτεύεται με παραμόρφωση D στον ℓ_1^d τότε $d \geqslant n^{\Omega(1/D^2)}$.*

Απόδειξη. Γνωρίζουμε ότι το G_k εμφυτεύεται με παραμόρφωση c στον L_1. Θέτουμε X την εικόνα του G_k στον L_1 μέσω μιας τέτοιας εμφύτευσης H. Θα δείξουμε ότι για κατάλληλο k, το X ικανοποιεί το συμπέρασμα.

Θεωρούμε μια D-εμφύτευση F του X στον ℓ_1^d. Παίρνουμε και την ταυτοτική απεικόνιση $I : \ell_1^d \to \ell_p^d$, όπου

$$p = 1 + \frac{1}{\log d}.$$

Τότε,

$$\|x\|_p \leqslant \|x\|_1 \leqslant e\|x\|_p \quad \text{για κάθε } x \in \ell_1^d.$$

Πράγματι,

$$\|x\|_p \leqslant \|x\|_1 \leqslant d^{1-\frac{1}{p}}\|x\|_p$$

από την ανισότητα του Hölder, και

$$d^{1-\frac{1}{p}} = d^{\frac{1/\log d}{1+1/\log d}} \leqslant d^{1/\log d} = e.$$

Άρα, μέσω της $I \circ F \circ H$, το G_k εμφυτεύεται στον L_p με παραμόρφωση ceD. Από το Θεώρημα 3.5.5 ισχύει

$$(ce)^2 D^2 \geqslant 1 + (p-1)k.$$

Από την $p = 1 + \frac{1}{\log d}$ παίρνουμε $(ce)^2 D^2 \geqslant \frac{k}{\log d}$, δηλαδή,

$$\log d \geqslant c_1 \frac{k}{D^2}.$$

Το πλήθος των κορυφών του G_k δίνεται από την ακολουθία $\alpha_0 = 2$, $\alpha_{k+1} = \alpha_k + 2 \cdot 4^k$. Άρα,

$$\alpha_k \geqslant 2 \cdot 4^{k-1} = \frac{1}{2} \cdot 4^k.$$

Για $k = \Omega(\log n)$ θα έχουμε

$$\alpha_k = \frac{1}{2} \cdot 4^{\log n} > \frac{n}{2}$$

και

$$\log d \geqslant c_1 \frac{\log n}{D^2},$$

άρα

$$d \geqslant n^{\Omega(1/D^2)}.$$

Έτσι, για συγκεκριμένο k το X έχει πλήθος σημείων της τάξης του n και ικανοποιεί το συμπέρασμα. □

Παρατήρηση. Σύμφωνα με το Πόρισμα 3.5.8 τα G_k δεν μπορούν να εμφυτευτούν σε ℓ_1-χώρους χαμηλής διάστασης, χωρίς να επέλθει σημαντική μεταβολή στην παραμόρφωση. Συνεπώς, δεν ισχύει ένα L_1-ανάλογο του Λήμματος των Johnson-Lindenstrauss.

3.6 ϵ-ισομετρική αναγωγή της διάστασης στον ℓ_p

Μια απεικόνιση $f : (X, d) \to (Y, \sigma)$ μεταξύ δύο μετρικών χώρων λέγεται αμφι-Lipschitz εμφύτευση με παραμόρφωση $\alpha \geqslant 1$ αν υπάρχει σταθερά $c > 0$ τέτοια ώστε

$$c\, d(x, y) \leqslant \sigma(f(x), f(y)) \leqslant \alpha c\, d(x, y)$$

για κάθε $x, y \in X$. Υπενθυμίζουμε ότι, από το λήμμα Johnson-Lindenstrauss, αν $\mathcal{H}$ είναι ένας χώρος Hilbert και $x_1, \dots, x_n \in \mathcal{H}$ τότε για κάθε $\epsilon \in (0, 1)$ υπάρχουν $k \leqslant \frac{C \log n}{\epsilon^2}$ και $y_1, \dots, y_n \in \ell_2^k$ τέτοια ώστε

$$\|x_i - x_j\|_{\mathcal{H}} \leqslant \|y_i - y_j\|_{\ell_2^k} \leqslant (1 + \epsilon)\|x_i - x_j\|_{\mathcal{H}}$$

για κάθε $i, j = 1, \dots, n$, όπου $C > 0$ είναι μια απόλυτη σταθερά. Με την ορολογία των εμφυτεύσεων, το λήμμα Johnson-Lindenstrauss μας λέει ότι για κάθε $\epsilon \in (0, 1)$, $n \in \mathbb{N}$ και $k \geqslant \frac{C \log n}{\epsilon^2}$, κάθε σύνολο n σημείων ενός χώρου Hilbert επιδέχεται μια αμφι-Lipschitz εμφύτευση στον ℓ_2^k με παραμόρφωση το πολύ ίση με $1 + \epsilon$.

Λέμε ότι ένας απειροδιάστατος χώρος Banach $(X, \|\cdot\|)$ επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης αν υπάρχει σταθερά $\alpha = \alpha(X) \geqslant 1$ τέτοια ώστε για κάθε $n \in \mathbb{N}$ να υπάρχει $k_n = k_n(X, \alpha) \in \mathbb{N}$ που ικανοποιεί την

$$\lim_{n \to \infty} \frac{\log k_n}{\log n} = 0$$

ώστε κάθε σύνολο $S \subseteq X$ με n σημεία να επιδέχεται μια αμφι-Lipschitz εμφύτευση με παραμόρφωση το πολύ ίση με α σε έναν υπόχωρο Y του X που έχει πεπερασμένη διάσταση $\dim(Y) \leqslant k_n$. Ο μόνος απειροδιάστατος χώρος Banach που δεν είναι χώρος Hilbert για τον οποίο είναι γνωστό ότι επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης είναι η 2-κυρτοποίηση του κλασσικού χώρου Tsirelson. Είναι επίσης γνωστό ότι ο χώρος ℓ_∞ δεν επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης και, όπως είδαμε στην προηγούμενη ενότητα, υπάρχει σύνολο $S \subseteq \ell_1$ με n σημεία το οποίο δεν επιδέχεται αμφι-Lipschitz εμφύτευση σε κανέναν $n^{o(1)}$-διάστατο υπόχωρο του ℓ_1. Το ακόλουθο ερώτημα παραμένει ανοικτό:

> *Για ποιές τιμές του $p \notin \{1, 2, \infty\}$ είναι σωστό ότι ο ℓ_p επιδέχεται αμφι-Lipschitz αναγωγή της διάστασης;*

Σε αυτή την ενότητα μελετάμε μια παραλλαγή αυτού του ερωτήματος. Έστω μ ένα μέτρο πιθανότητας. Αν S είναι ένα υποσύνολο του $L_p(\mu)$, ορίζουμε

$$J(S) := \left\| \max_{x \in S} |x| \right\|_{L_p(\mu)} \tag{3.10}$$

και λέμε ότι το S είναι K-μη-συμπιέσιμο αν $J(S) \leqslant K$. Θα δείξουμε ένα θεώρημα αναγωγής της διάστασης για μη-συμπιέσιμα υποσύνολα του $L_p(\mu)$, το οποίο ισχύει για κάθε $p \in [1, \infty)$.

Θεώρημα 3.6.1. *Έστω $1 \leqslant p < \infty$, $n \in \mathbb{N}$, $0 < K < \infty$ και $\{x_1, \dots, x_n\}$ μια K-μη-συμπιέσιμη οικογένεια διανυσμάτων στον $L_p(\mu)$, όπου μ είναι ένα μέτρο πιθανότητας. Τότε, για κάθε*

$\epsilon \in (0,1)$ υπάρχουν $d \in \mathbb{N}$ με $d \leqslant \frac{32e^2(2K)^{2p}\log n}{\epsilon^2}$ και σημεία $y_1,\dots,y_n \in \ell_p^d$ τέτοια ώστε, για κάθε $i,j=1,\dots,n$,

$$\|x_i - x_j\|_{L_p(\mu)}^p - \epsilon \leqslant \|y_i - y_j\|_{\ell_p^d}^p \leqslant \|x_i - x_j\|_{L_p(\mu)}^p + \epsilon. \tag{3.11}$$

Πέρα από την παράμετρο K της μη-συμπιεσιμότητας που υπεισέρχεται στο φράγμα για τη διάσταση d του χώρου ℓ_p^d στον οποίο επιτυγχάνεται η εμφύτευση, το Θεώρημα 3.6.1 διαφέρει από το λήμμα Johnson-Lindenstrauss στο ότι το σφάλμα στην (3.11) είναι προσθετικό και όχι πολλαπλασιαστικό. Λέμε ότι μια απεικόνιση $f:(X,d)\to(Y,\sigma)$ μεταξύ μετρικών χώρων είναι ϵ-ισομετρική εμφύτευση αν

$$|\sigma(f(x),f(y)) - d(x,y)| \leqslant \epsilon \tag{3.12}$$

για κάθε $x,y\in X$. Χρησιμοποιώντας τη στοιχειώδη ανισότητα $|a-b|\leqslant |a^p-b^p|^{1/p}$ που ισχύει για κάθε $a,b>0$, βλέπουμε ότι το Θεώρημα 3.6.1 μπορεί να διατυπωθεί ως εξής: κάθε K-μη-συμπιέσιμο σύνολο n σημείων στον $L_p(\mu)$ επιδέχεται $\epsilon^{1/p}$-ισομετρική εμφύτευση στον ℓ_p^d για κάποιον $d\leqslant \frac{32e^2(2K)^{2p}\log n}{\epsilon^2}$.

Για την απόδειξη του Θεωρήματος 3.6.1 θα χρησιμοποιήσουμε το γεγονός ότι ο $L_p(\mu)$ έχει Rademacher τύπο p. Γενικά, λέμε ότι ένας χώρος με νόρμα $(X,\|\cdot\|)$ έχει Rademacher τύπο p αν υπάρχει σταθερά $0<T<\infty$ τέτοια ώστε, για κάθε $n\in\mathbb{N}$ και κάθε $x_1,\dots,x_n\in X$,

$$\frac{1}{2^n}\sum_{\epsilon\in\{-1,1\}^n}\left\|\sum_{i=1}^n \epsilon_i x_i\right\|^p \leqslant T^p\sum_{i=1}^n \|x_i\|^p. \tag{3.13}$$

Η μικρότερη θετική σταθερά T για την οποία ικανοποιείται η (3.13) συμβολίζεται με $T_p(X)$. Με ένα απλό επιχείρημα συμμετρικοποίησης βλέπουμε ότι αν $Z_1,\dots,Z_n$ είναι ανεξάρτητες τυχαίες μεταβλητές με τιμές στον X και $\mathbb{E}(Z_i)=0$ για κάθε $i=1,\dots n$ τότε

$$\mathbb{E}\left\|\sum_{i=1}^n Z_i\right\|^p \leqslant (2T_p(X))^p\sum_{i=1}^n \mathbb{E}\|Z_i\|^p. \tag{3.14}$$

Θα χρειαστούμε την ακόλουθη παραλλαγή του Θεωρήματος 2.1.2, η οποία οφείλεται κι αυτή στον Maurey.

Λήμμα 3.6.2. *Έστω $(X,\|\cdot\|)$ χώρος Banach και $T\subseteq RB_X$. Για κάθε $z\in \mathrm{conv}(T)$ και κάθε $d\geqslant 1$ υπάρχουν $x_1,\dots,x_d\in T$ τέτοια ώστε*

$$\left\|\frac{1}{d}\sum_{s=1}^d x_s - z\right\| \leqslant \frac{4RT_p(X)}{d^{1-1/p}}. \tag{3.15}$$

Απόδειξη. Το z είναι κυρτός συνδυασμός σημείων του T, δηλαδή υπάρχουν $m\in\mathbb{N}$, $\lambda_k>0$ και

$t_k \in T$ ώστε

$$\sum_{k=1}^{m} \lambda_k = 1 \quad \text{και} \quad z = \sum_{k=1}^{m} \lambda_k t_k. \tag{3.16}$$

εωρούμε μια διακριτή τυχαία μεταβλητή Z με τιμές στον X, που παίρνει τις τιμές t_k με $\mathbb{P}(Z = t_k) = \lambda_k$ για κάθε $k = 1, \dots, m$. Στη συνέχεια θεωρούμε ανεξάρτητα αντίγραφα $X_1, \dots, X_d$ της X. Από την (3.16) βλέπουμε ότι η Z είναι καλά ορισμένη και $\mathbb{E}(Z) = z$. Εφαρμόζοντας την (3.14) για τις κεντραρισμένες τυχαίες μεταβλητές $\{Z_i - z\}_{s=1}^{d}$ και κανονικοποιώντας, παίρνουμε

$$\mathbb{E}\left\| \frac{1}{d} \sum_{s=1}^{d} Z_s - z \right\|^p \leqslant \frac{(2T_p(X))^p}{d^{p-1}} \mathbb{E}\|X - z\|^p.$$

Αφού η X παίρνει τιμές στο $T \subseteq RB_X$, συμπεραίνουμε ότι υπάρχουν $x_1, \dots, x_d \in T$ ώστε

$$\left\| \frac{1}{d} \sum_{s=1}^{d} x_s - z \right\| \leqslant \frac{4RT_p(X)}{d^{1-1/p}},$$

όπως θέλαμε. □

Έχοντας το Λήμμα 3.6.2 στη διάθεσή μας, αποδεικνύουμε μια ειδική περίπτωση του θεωρήματος, δηλαδή μια ϵ-ισομετρική αναγωγή της διάστασης κάνοντας την ισχυρότερη υπόθεση ότι το σύνολο διανυσμάτων που μας δίνεται αποτελείται από συναρτήσεις που είναι φραγμένες στον $L_\infty(\mu)$.

Πρόταση 3.6.3. *Έστω $1 \leqslant p < \infty$, $n \in \mathbb{N}$ και $\{x_1, \dots, x_n\}$ μια οικογένεια διανυσμάτων στον $L_p(\mu)$, όπου μ είναι ένα μέτρο πιθανότητας. Θέτουμε*

$$L = \max_{1 \leqslant i \leqslant n} \|x_i\|_{L_\infty(\mu)}.$$

Τότε, για κάθε $\epsilon \in (0,1)$ υπάρχουν $d \in \mathbb{N}$ με $d \leqslant \frac{32e^2(2L)^{2p} \log n}{\epsilon^2}$ και σημεία $y_1, \dots, y_n \in \ell_p^d$ τέτοια ώστε, για κάθε $i, j = 1, \dots, n$,

$$\|x_i - x_j\|_{L_p(\mu)}^p - \epsilon \leqslant \|y_i - y_j\|_{\ell_p^d}^p \leqslant \|x_i - x_j\|_{L_p(\mu)}^p + \epsilon.$$

Απόδειξη. Ταυτίζουμε τον $\ell_\infty^{\binom{n}{2}}$ με τον διανυσματικό χώρο όλων των συμμετρικών $n \times n$ πραγματικών πινάκων που έχουν 0 στη διαγώνιο, εφοδιασμένο με την supremum νόρμα. Θεωρούμε το σύνολο

$$C_p = \left\{ \left(\|z_i - z_j\|_{L_p(\rho)}^p \right)_{i,j=1,\dots,n} : \rho \text{ μέτρο πιθανότητας και } z_1, \dots, z_n \in L_p(\rho) \right\} \subseteq \ell_\infty^{\binom{n}{2}}.$$

Το σύνολο C_p είναι κώνος, δηλαδή $C_p = \lambda C_p$ για κάθε $\lambda > 0$, και επιπλέον είναι κυρτό σύνολο. Για να το ελέγξουμε αυτό, θεωρούμε $A, B \in C_p$, χώρους πιθανότητας $(\Omega_1, \rho_1), (\Omega_2, \rho_2)$

και διανύσματα $\{z_i\}_{i=1}^n, \{w_i\}_{i=1}^n$ στους $L_p(\rho_1)$ και $L_p(\rho_2)$ αντίστοιχα, ώστε

$$A_{ij} = \|z_i - z_j\|^p_{L_p(\rho_1)} \quad \text{και} \quad B_{ij} = \|w_i - w_j\|^p_{L_p(\rho_2)} \tag{3.17}$$

για κάθε $i, j = 1, \dots, n$. Σταθεροποιούμε $\lambda \in (0,1)$ και θεωρούμε την ξένη ένωση $\Omega = \Omega_1 \sqcup \Omega_2$ των Ω_1 και Ω_2 εφοδιασμένη με το μέτρο πιθανότητας $\rho(\lambda) = \lambda\rho_1 + (1-\lambda)\rho_2$. Τότε, από την (3.17), οι συναρτήσεις $\zeta_i : \Omega_1 \sqcup \Omega_2 \to \mathbb{R}$ που ορίζονται από τις $\zeta_i\big|_{\Omega_1} = z_i$ και $\zeta_i\big|_{\Omega_2} = w_i$, $i = 1, \dots, n$ ανήκουν στον $L_p(\rho(\lambda))$ και ικανοποιούν τις

$$\|\zeta_i - \zeta_j\|^p_{L_p(\rho(\lambda))} = \lambda\|z_i - z_j\|^p_{L_p(\rho_1)} + (1-\lambda)\|w_i - w_j\|^p_{L_p(\rho_2)} = \lambda A_{ij} + (1-\lambda)B_{ij}$$

για κάθε $i, j = 1, \dots n$, το οποίο δείχνει ότι $\lambda A + (1-\lambda)B \in C_p$, συνεπώς το C_p είναι κυρτός κώνος. Θεωρούμε την εμφύτευση $M : L_p(\mu)^n \to C_p$ που απεικονίζει ένα διάνυσμα $z = (z_1, \dots, z_n)$ στον αντίστοιχο πίνακα αποστάσεων, δηλαδή

$$M(z)_{ij} = \|z_i - z_j\|^p_{L_p(\mu)}, \qquad i, j = 1, \dots, n.$$

Το θεώρημα ισομετρικής εμφύτευσης του K. Ball δείχνει ότι τα $x_1, \dots, x_n$ έχουν ισομετρικές εικόνες στον ℓ_p^N με $N = \binom{n}{2} + 1$. Χωρίς περιορισμό της γενικότητας μπορούμε λοιπόν να υποθέσουμε ότι οι δοσμένες συναρτήσεις $x_1, \dots, x_n \in L_p(\mu)$ είναι απλές συναρτήσεις (δηλαδή, καθεμία από αυτές παίρνει πεπερασμένες το πλήθος τιμές) με $\|x_i\|_{L_\infty(\mu)} \leqslant L$.

Έστω $\{S_1, \dots, S_m\}$ μια διαμέριση του υποκείμενου χώρου πιθανότητας, τέτοια ώστε κάθε συνάρτηση x_i να είναι σταθερή σε κάθε S_k και έστω ότι $x_i\big|_{S_k} = a(i,k)$ για κάθε $i = 1, \dots, n$ και $k = 1, \dots, m$. Τότε, για κάθε $i, j = 1, \dots, n$ έχουμε

$$M(x)_{ij} = \sum_{k=1}^m \int_{S_k} |x_i - x_j|^p d\mu = \sum_{k=1}^m \mu(S_k) \cdot |a(i,k) - a(j,k)|^p = \sum_{k=1}^m \mu(S_k) M(y(k))_{ij}, \tag{3.18}$$

όπου $y(k) = (a(1,k), \dots, a(n,k)) \in L_p(\mu)^n$ είναι ένα διάνυσμα που οι συντεταγμένες του είναι σταθερές συναρτήσεις. Εφόσον το μ είναι μέτρο πιθανότητας και το σύνολο $\{S_1, \dots, S_m\}$ είναι διαμέριση, από την (3.18) συμπεραίνουμε ότι

$$M(x) \in \operatorname{conv}\{M(y(k)) : k = 1, \dots, m\} \subseteq \ell_\infty^{\binom{n}{2}}.$$

Δεδομένου ότι $a(i,k) \in [-L, L]$ για κάθε $i = 1, \dots n$ και $k = 1, \dots, m$, έχουμε

$$\|M(y(k))\|_{\ell_\infty^{\binom{n}{2}}} = \max_{1 \leqslant i,j \leqslant n} |a(i,k) - a(j,k)|^p \leqslant (2L)^p \tag{3.19}$$

για κάθε $k = 1, \dots, m$. Επιπλέον, ο $\ell_\infty^{\binom{n}{2}}$ είναι e-ισομορφικός με τον $\ell_{p_n}^{\binom{n}{2}}$, όπου $p_n = \log\binom{n}{2}$. Τώρα χρησιμοποιούμε το γεγονός ότι $T_2(\ell_p) \leqslant \sqrt{p-1}$ για κάθε $p \geqslant 2$, άρα

$$T_2(\ell_\infty^{\binom{n}{2}}) \leqslant e\sqrt{p_n - 1} \leqslant \sqrt{2e^2 \log n}. \tag{3.20}$$

Εφαρμόζοντας το Λήμμα 3.6.2 και παίρνοντας υπ' όψιν μας τις (3.19) και (3.20) συμπεραίνουμε ότι για κάθε $d \geqslant 1$ υπάρχουν $k_1, \dots, k_d \in \{1, \dots, m\}$ τέτοιοι ώστε

$$\left\| \frac{1}{d} \sum_{s=1}^{d} M(y(k_s)) - M(x) \right\|_{\ell_\infty^{\binom{d}{2}}} \leqslant \frac{2^{p+\frac{5}{2}} e L^p \sqrt{\log n}}{\sqrt{d}}. \tag{3.21}$$

Συνεπώς, αν ο $\epsilon \in (0,1)$ ικανοποιεί την $d \geqslant \frac{32e^2(2L)^{2p}\log n}{\epsilon^2}$ έχουμε ότι

$$\left| \frac{1}{d} \sum_{s=1}^{d} |a(i,k_s) - a(j,k_s)|^p - \|x_i - x_j\|_{L_p(\mu)}^p \right| \leqslant \epsilon \tag{3.22}$$

για κάθε $i, j = 1, \dots, n$. Τώρα, για κάθε $i = 1, \dots, n$ θεωρούμε το διάνυσμα $y_i = (y_i(1), \dots, y_i(d)) \in \ell_p^d$ με συντεταγμένες

$$y_i(s) = \frac{a(i,k_s)}{d^{1/p}}, \qquad s = 1, \dots, d$$

και παρατηρούμε ότι μπορούμε να γράψουμε την (3.22) ισοδύναμα ως

$$\|x_i - x_j\|_{L_p(\mu)}^p - \epsilon \leqslant \|y_i - y_j\|_{\ell_p^d}^p \leqslant \|x_i - x_j\|_{L_p(\mu)}^p + \epsilon, \qquad i, j = 1, \dots, n$$

το οποίο ολοκληρώνει την απόδειξη της πρότασης. □

Το συμπέρασμα της Πρότασης 3.6.3 εξαρτάται από τις ανά δύο αποστάσεις των σημείων x_i στον $L_p(\mu)$ ενώ το φράγμα για την διάσταση εξαρτάται από την παράμετρο $L = \max\limits_{1 \leqslant i \leqslant n} \|x_i\|_{L_\infty(\mu)}$. Για να πάρουμε το Θεώρημα 3.6.1 από αυτήν την ασθενέστερη πρόταση, θα εκμεταλλευτούμε το γεγονός ότι η Πρόταση 3.6.3 ισχύει για κάθε μέτρο πιθανότητας μ βελτιστοποιώντας αυτήν την παράμετρο L πάνω από όλες τις lattice-ισομορφικές εικόνες του $\{x_1, \dots, x_n\}$. Θα βρούμε μια βέλτιστη αλλαγή μέτρου που θα μας επιτρέψει να αντικαταστήσουμε την παράμετρο L με την $K = \| \max_i |x_i| \|_{L_p(\mu)}$. Αυτή η αλλαγή μέτρου επιτυγχάνεται με τη βοήθεια ενός κλασσικού θεωρήματος παραγοντοποίησης του Maurey.

Πρόταση 3.6.4. *Έστω $n \in \mathbb{N}$, $0 < p < \infty$ και (Ω, μ) ένας χώρος πιθανότητας. Αν $x_1, \dots, x_n \in L_p(\mu)$ τότε υπάρχει μια μη αρνητική συνάρτηση πυκνότητας $f : \Omega \to \mathbb{R}$ που ο φορέας της περιέχεται στον φορέα της $\max_i |x_i|$, τέτοια ώστε: αν ν είναι το μέτρο πιθανότητας στον Ω με $\frac{d\nu}{d\mu} = f$ να ισχύει ότι*

$$\max_{1 \leqslant i \leqslant n} \left\| x_i f^{-1/p} \right\|_{L_\infty(\nu)} \leqslant \left\| \max_{1 \leqslant i \leqslant n} |x_i| \right\|_{L_p(\mu)}. \tag{3.23}$$

Απόδειξη. Θέτουμε $V = \operatorname{supp}(\max_i |x_i|) \subseteq \Omega$ και ορίζουμε $f : \Omega \to \mathbb{R}$ με

$$f(\omega) := \frac{\max_i |x_i(\omega)|^p}{\int_\Omega \max_i |x_i(\theta)|^p \, d\theta}, \qquad \omega \in \Omega. \tag{3.24}$$

Είναι τότε εύκολο να ελέγξουμε ότι ισχύει η (3.23). □

Απόδειξη του Θεωρήματος 3.6.1. Θεωρούμε μια K-μη-συμπιέσιμη οικογένεια διανυσμάτων

$$x_1, \dots, x_n \in L_p(\Omega, \mu)$$

και θέτουμε $V = \operatorname{supp}(\max_i |x_i|) \subseteq \Omega$. Θεωρούμε επίσης τη συνάρτηση αλλαγής πυκνότητας $f : \Omega \to \mathbb{R}$ που μας εξασφαλίζει η Πρόταση 3.6.4. Αν $\frac{d\nu}{d\mu} = f$, τότε ελέγχουμε εύκολα ότι ο γραμμικός τελεστής $T : L_p(V, \mu) \to L_p(\Omega, \nu)$ που ορίζεται από την $T(g) = f^{-1/p} g$ είναι γραμμική ισομετρία. Από την Πρόταση 3.6.3 και την (3.23) έχουμε ότι υπάρχουν $d \in \mathbb{N}$ με $d \leqslant \frac{32e^2(2K)^{2p}\log n}{\epsilon^2}$ και σημεία $y_1, \dots, y_n \in \ell_p^d$ ώστε να ικανοποιείται η

$$\begin{aligned} \|x_i - x_j\|_{L_p(\mu)}^p - \epsilon = \|Tx_i - Tx_j\|_{L_p(\nu)}^p - \epsilon &\leqslant \|y_i - y_j\|_{\ell_p^d}^p \\ &\leqslant \|Tx_i - Tx_j\|_{L_p(\nu)}^p + \epsilon = \|x_i - x_j\|_{L_p(\mu)}^p + \epsilon \end{aligned}$$

για κάθε $i, j = 1, \dots, n$. Αυτό ολοκληρώνει την απόδειξη του Θεωρήματος 3.6.1. □

Όπως φαίνεται από την απόδειξη του Θεωρήματος 3.6.1, τα σημεία $y_1, \dots, y_n$ του ℓ_p^d επιλέγονται ως εικόνες των δοσμένων σημείων $x_1, \dots, x_n$ μέσω ενός γραμμικού μετασχηματισμού. Πράγματι, ξεκινώντας από μία K-μη-συμπιέσιμη οικογένεια συναρτήσεων $x_1, \dots, x_n$ στον $L_p(\Omega, \mu)$, χρησιμοποιούμε την Πρόταση 3.6.4 για να βρούμε μια αλλαγή μέτρου $T : L_p(V, \mu) \to L_p(\Omega, \nu)$ τέτοια ώστε οι $Tx_1, \dots, Tx_n$ να ικανοποιούν την ισχυρότερη υπόθεση της Πρότασης 3.6.3. Τότε, για κάποιον $d \in \mathbb{N}$ με $d \leqslant \frac{32e^2(2K)^{2p}\log n}{\epsilon^2}$ βρίσκουμε ξένα μετρήσιμα υποσύνολα $S_1, \dots, S_d$ του Ω, καθένα με θετικό μέτρο, ώστε αν $G : L_p(\Omega, \nu) \to \ell_p^d$ είναι η γραμμική απεικόνιση με

$$G(z) = \frac{1}{d^{1/p}} \left(\frac{1}{\mu(S_1)} \int_{S_1} z \, d\nu, \dots, \frac{1}{\mu(S_d)} \int_{S_d} x \, d\nu \right) \in \ell_p^d$$

για κάθε $z \in L_p(\Omega, \nu)$, τότε τα σημεία $y_i = (G \circ T)(x_i)$ του ℓ_p^d να ικανοποιούν την (3.11).

Το επόμενο όμως θεώρημα δείχνει ότι για κάθε $p \neq 2$ και n αρκετά μεγάλο, υπάρχει σύνολο n σημείων της B_{ℓ_p} που η εικόνα του κάτω από οποιαδήποτε ϵ-ισομετρική εμφύτευση έχει τάξη γραμμική ως προς n.

Θεώρημα 3.6.5. *Έστω $p \neq 2$ και $\omega, \Omega : [0, \infty) \to [0, \infty)$ δύο συναρτήσεις με $\omega(1) > 0$. Για η οσοδήποτε μεγάλο, υπάρχει ένα υποσύνολο $S_{n,p}$ της B_{ℓ_p} με η σημεία τέτοιο ώστε να ισχύει το εξής: Αν $T : \operatorname{span}(S_{n,p}) \to \ell_p^d$ είναι ένας γραμμικός τελεστής τέτοιος ώστε*

$$\omega(\|x - y\|_{\ell_p}) \leqslant \|Tx - Ty\|_{\ell_p^d} \leqslant \Omega(\|x - y\|_{\ell_p})$$

για κάθε $x, y \in S_{n,p}$, τότε

$$d \geqslant \left(\frac{\omega(1)}{\Omega(1)} \right)^{\frac{2p}{|p-2|}} \cdot \frac{n-1}{2}.$$

Απόδειξη. Σταθεροποιούμε $m \in \mathbb{N}$ και συμβολίζουμε με $\{w_i\}_{i=1}^{2^m}$ τις γραμμές του $2^m \times 2^m$ πίνακα Walsh και με $\{e_i\}_{i=1}^{2^m}$ τα διανύσματα της συνήθους βάσης του $\mathbb{R}^{2^m}$. Θεωρούμε το σύνολο n

σημείων

$$S_{n,p} = \{0\} \cup \{e_1, \dots, e_{2^m} \cup \left\{ \frac{w_1}{2^{m/p}}, \dots, \frac{w_{2^m}}{2^{m/p}} \right\}$$

στην $B_{\ell_p^{2^m}}$, όπου $n = 2^{m+1} + 1$, και υποθέτουμε ότι $T : \ell_p^{2^m} \to \ell_p^d$ είναι ένας γραμμικός τελεστής τέτοιος ώστε

$$\omega(\|x-y\|_{\ell_p^{2^m}}) \leqslant \|Tx - Ty\|_{\ell_p^d} \leqslant \Omega(\|x-y\|_{\ell_p^{2^m}}) \tag{3.25}$$

για κάθε $x, y \in S_{n,p}$.

Υποθέτουμε πρώτα ότι $1 \leqslant p < 2$. Αν γράψουμε $w_i = \sum_{j=1}^{2^m} w_i(j) e_j$ τότε από την ορθογωνιότητα του $\{w_i\}_{i=1}^{2^m}$ έχουμε ότι

$$\sum_{i=1}^{2^m} \|Tw_i\|_{\ell_2^d}^2 = \sum_{i=1}^{2^m} \Big\| \sum_{j=1}^{2^m} w_i(j) Te_j \Big\|_{\ell_2^d}^2 = \sum_{j,k=1}^{2^m} \langle w_j, w_k \rangle \langle Te_j, Te_k \rangle = 2^m \sum_{j=1}^{2^m} \|Te_j\|_{\ell_2^d}^2. \tag{3.26}$$

Αφού ο T ικανοποιεί την (3.25), έχουμε

$$\|Te_j\|_{\ell_2^d}^2 \leqslant \|Te_j\|_{\ell_p^d}^2 \leqslant \Omega(1)^2 \tag{3.27}$$

για κάθε $j = 1, \dots, 2^m$ και

$$\|Tw_j\|_{\ell_2^d}^2 \geqslant 2^{\frac{2m}{p}} d^{-\frac{2-p}{p}} \Big\| T\Big(\frac{w_j}{2^{m/p}} \Big) \Big\|_{\ell_p^d}^2 \geqslant \omega(1)^2 \tag{3.28}$$

για κάθε $j = 1, \dots, 2^m$. Συνδυάζοντας τις (3.26), (3.27) και (3.28) συμπεραίνουμε ότι

$$2^{m\left(1+\frac{2}{p}\right)} d^{-\frac{2-p}{p}} \omega(1)^2 \leqslant 4^m \Omega(1)^2,$$

η οποία είναι ισοδύναμη με την

$$d \geqslant \left(\frac{\omega(1)}{\Omega(1)} \right)^{\frac{2p}{2-p}} 2^m = \left(\frac{\omega(1)}{\Omega(1)} \right)^{\frac{2p}{|p-2|}} \cdot \frac{n-1}{2}.$$

Για την περίπτωση $p > 2$ δουλεύουμε με τον ίδιο τρόπο. □

Κεφάλαιο 4

Ανάκτηση σήματος

Για απλότητα θα θεωρήσουμε την πραγματική περίπτωση, πραγματικά διανύσματα και πραγματικούς πίνακες. Έστω $1 < n \leqslant N$ ακέραιοι και A ένας $n \times N$ πραγματικός πίνακας. Συνήθως σκεφτόμαστε ότι $N \gg n$. Αυτό που θέλουμε είναι να «συμπιέσουμε» κάποια διανύσματα από τον $\mathbb{R}^N$, όπου το N είναι μεγάλο, σε διανύσματα στον $\mathbb{R}^n$. Έστω $X_1, \ldots, X_N \in \mathbb{R}^n$ οι στήλες του A και έστω $Y_1, \ldots, Y_n \in \mathbb{R}^N$ οι γραμμές του. Δηλαδή, έχουμε

$$A = (X_1\ X_2\ \cdots\ X_N) = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}.$$

Μας δίνεται επίσης ένα υποσύνολο $T \subset \mathbb{R}^n$ διανυσμάτων. Έστω τώρα $x \in T$ ένα *άγνωστο* διάνυσμα. Τα δεδομένα που έχουμε είναι n γραμμικές μετρήσεις του x (πάλι, σκεφτόμαστε ότι $N \gg n$)

$$\langle Y_1, x\rangle, \ldots, \langle Y_n, x\rangle$$

ή, ισοδύναμα, μας δίνεται το

$$y = Ax.$$

Θέλουμε να *επανακτήσουμε* το x ή καλύτερα να *ανακατασκευάσουμε* το x, ακριβώς ή προσεγγιστικά, με προκαθορισμένη ακρίβεια και με αποτελεσματικό τρόπο (να βρούμε κάποιον γρήγορο αλγόριθμο).

4.1 Το πρόβλημα της ακριβούς ανακατασκευής

Θα συζητήσουμε αρχικά το πρόβλημα της ακριβούς ανακατασκευής. Έστω $x \in T$ το άγνωστο διάνυσμα και έστω ότι τα δεδομένα είναι το $y = Ax$. Όταν $N \gg n$, το πρόβλημα δεν είναι καλά τοποθετημένο γιατί το σύστημα $At = y$, $t \in \mathbb{R}^n$ είναι ισχυρά απροσδιόριστο. Αν λοιπόν θέλουμε να ανακτήσουμε το x, χρειαζόμαστε περισσότερες πληροφορίες για την *φύση* του. Επιπλέον, αν θέλουμε να ανακτήσουμε οποιοδήποτε x από το T, θα πρέπει να έχουμε κάποιες

αρχικές πληροφορίες για το T και την *πολυπλοκότητά* του. Στη συνέχεια θα μιλήσουμε πιο συγκεκριμένα για διάφορες παραμέτρους πολυπλοκότητας. Η αρχική υπόθεση που θα μελετήσουμε τώρα είναι η *αραιότητα*.

Εισάγουμε πρώτα κάποιον συμβολισμό. Εφοδιάζουμε τον $\mathbb{R}^n$ και τον $\mathbb{R}^N$ με το σύνηθες εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ και την επαγόμενη Ευκλείδεια νόρμα $\| \cdot \|_2$. Συμβολίζουμε με B_2^n τη μοναδιαία Ευκλείδεια μπάλα στον $\mathbb{R}^n$ και με S^{n-1} τη μοναδιαία της σφαίρα. Χρησιμοποιούμε τον συμβολισμό $|A|$ για τον πληθάριθμο ενός πεπερασμένου συνόλου A.

Ορισμός 4.1.1. Έστω $0 \leqslant m \leqslant N$ ακέραιοι. Για κάθε $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ θέτουμε $\mathrm{supp}(x) = \{1 \leqslant k \leqslant N : x_k \neq 0\}$, το σύνολο δεικτών των μη μηδενικών συντεταγμένων του x. Λέμε ότι το x είναι m-αραιό αν $|\mathrm{supp}(x)| \leqslant m$. Το σύνολο των m-αραιών διανυσμάτων του $\mathbb{R}^N$ συμβολίζεται με $\Sigma_m = \Sigma_m(\mathbb{R}^N)$ και η μοναδιαία του σφαίρα με

$$S_2(\Sigma_m) = \{x \in \mathbb{R}^N : \ \|x\|_2 = 1 \quad \text{και} \quad |\mathrm{supp}(x)| \leqslant m\}.$$

Ανάλογα, θέτουμε

$$B_2(\Sigma_m) = \{x \in \mathbb{R}^N : \ \|x\|_2 \leqslant 1 \quad \text{και} \quad |\mathrm{supp}(x)| \leqslant m\}.$$

Σημειώνουμε ότι το σύνολο Σ_m δεν είναι διανυσματικός υπόχωρος του $\mathbb{R}^N$ και ότι το σύνολο $B_2(\Sigma_m)$ δεν είναι κυρτό.

Πρόβλημα 4.1.2 (Το πρόβλημα της ακριβούς ανακατασκευής). Θέλουμε να ανακτήσουμε ακριβώς κάθε m-αραιό διάνυσμα $x \in \Sigma_m$ από τα δεδομένα $y = Ax$. Δηλαδή, ζητάμε έναν αποκωδικοποιητή Δ τέτοιον ώστε, για κάθε $x \in \Sigma_m$,

$$\Delta(A, Ax) = x.$$

Ισχυρισμός 4.1.3. Από τη Γραμμική Άλγεβρα γνωρίζουμε ότι ένας τέτοιος αποκωδικοποιητής υπάρχει αν και μόνο αν

$$\ker(A) \cap \Sigma_{2m} = \{0\}.$$

Όπως έχουμε ήδη αναφέρει, υπάρχουν πολλοί διαφορετικοί τρόποι για να μετρήσουμε την πολυπλοκότητα ενός μετρικού χώρου. Ο απλούστερος τρόπος είναι ίσως να εκτιμήσουμε πόσο συμπαγής είναι, μέσω των λεγόμενων αριθμών κάλυψης και στοίχισης.

Δεδομένου ότι όλοι οι μετρικοί χώροι που θα θεωρήσουμε είναι υποσύνολα χώρων με νόρμα, περιοριζόμαστε σε αυτό το πλαίσιο.

Ορισμός 4.1.4. Έστω B και C υποσύνολα ενός διανυσματικού χώρου και έστω $\epsilon > 0$. Ένα ϵ-δίκτυο με μεταφορές του ϵC είναι ένα υποσύνολο Λ του B τέτοιο ώστε: για κάθε $x \in B$ υπάρχουν $y \in \Lambda$ και $z \in C$ ώστε $x = y + \epsilon z$. Με άλλα λόγια,

$$B \subseteq \Lambda + \epsilon C = \bigcup_{x \in \Lambda} (x + \epsilon C),$$

όπου $\Lambda + \epsilon C := \{a + \epsilon c : a \in \Lambda, c \in C\}$ είναι το άθροισμα Minkowski των συνόλων Λ και ϵC.

Ο αριθμός κάλυψης του B από το ϵC είναι ο ελάχιστος πληθάριθμος ενός τέτοιου ϵ-δικτύου και συμβολίζεται με $N(B, \epsilon C)$. Η συνάρτηση $\epsilon \mapsto \log B(B, \epsilon C)$ λέγεται συνάρτηση μετρικής εντροπίας του B από το C.

Παρατήρηση 4.1.5. Αν (B, d) είναι ένας μετρικός χώρος, τότε ένα ϵ-δίκτυο του (B, d) είναι μια κάλυψη του B με μπάλες ακτίνας ϵ ως προς τη μετρική d. Ο αριθμός κάλυψης είναι ο ελάχιστος πληθάριθμος ενός ϵ-δικτύου και συμβολίζεται με $N(B, d, \epsilon)$. Στο δικό μας πλαίσιο, η μετρική d επάγεται από μια νόρμα με μοναδιαία μπάλα C. Τότε, το $x + \epsilon C$ είναι η μπάλα ακτίνας ϵ με κέντρο το x.

Έστω $C \subset \mathbb{R}^N$ ένα συμμετρικό κυρτό σώμα, δηλαδή ένα συμμετρικό κυρτό συμπαγές υποσύνολο του $\mathbb{R}^N$ με μη κενό εσωτερικό (δηλαδή, η μοναδιαία μπάλα μιας νόρμας στον $\mathbb{R}^N$). Θεωρούμε ένα υποσύνολο $\Lambda \subset C$ με τον μέγιστο δυνατό πληθάριθμο έτσι ώστε τα σημεία του Λ να βρίσκονται ϵC-μακριά, με την έννοια ότι αν $x, y \in \Lambda$ και $x \neq y$ τότε $x - y \notin \epsilon C$ (θυμηθείτε ότι $-C = C$). Τότε, το Λ είναι ϵ-δίκτυο του C ως προς ϵC. Επιπλέον, οι μπάλες

$$(x + (\epsilon/2)C)_{x \in \Lambda}$$

ακτίνας $\epsilon/2$ με κέντρα τα σημεία του Λ είναι ανά δύο ξένες (ακριβέστερα, τα εσωτερικά τους είναι ξένα) και η ένωσή τους είναι υποσύνολο του $(1 + (\epsilon/2))C$ (στο σημείο αυτό χρησιμοποιείται η κυρτότητα του C). Θεωρώντας τον όγκο αυτής της ένωσης, βλέπουμε ότι

$$N(C, \epsilon C) \leqslant (1 + (2/\epsilon))^N.$$

Έχουμε λοιπόν δείξει το εξής.

Πρόταση 4.1.6. *Έστω $\epsilon \in (0, 1)$. Έστω $C \subset \mathbb{R}^N$ ένα συμμετρικό κυρτό σώμα, δηλαδή η μοναδιαία μπάλα μιας νόρμας. Υπάρχει ϵ-δίκτυο Λ του C με μεταφορές του ϵC τέτοιο ώστε $|\Lambda| \leqslant (1 + 2/\epsilon)^N$. Επιπλέον,*

$$\Lambda \subset (1 - \epsilon)^{-1}\mathrm{conv}(\Lambda).$$

Το «επιπλέον» μέρος της Πρότασης 4.1.6 αποδεικνύεται με διαδοχική προσέγγιση.

Απόδειξη. Αφού το Λ είναι ϵ-δίκτυο του C με μεταφορές του ϵC, κάθε $z \in C$ γράφεται ως $z = x_0 + \epsilon z_1$, όπου $x_0 \in \Lambda$ και $z_1 \in C$. Επαναλαμβάνοντας αυτό το επιχείρημα, βλέπουμε ότι

$$z = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots,$$

όπου $x_i \in \Lambda$, το οποίο συνεπάγεται, λόγω κυρτότητας, ότι $C \subset (1 - \epsilon)^{-1}\mathrm{conv}(\Lambda)$. □

Αυτό αποδεικνύει τον πρώτο ισχυρισμό μας.

Ισχυρισμός 4.1.7 (κάλυψη της Ευκλείδειας σφαίρας με Ευκλείδειες μπάλες ακτίνας ϵ). Για κάθε $\epsilon \in (0, 1)$ ισχύει ότι

$$N(S^{N-1}, \epsilon B_2^N) \leqslant \left(\frac{3}{\epsilon}\right)^N.$$

Τώρα, αφού η $S_2(\Sigma_m)$ είναι η ένωση σφαιρών διάστασης m, βλέπουμε ότι

$$N(S_2(\Sigma_m), \epsilon B_2^N) \leqslant \binom{N}{m} N(S^{m-1}, \epsilon B_2^m).$$

Χρησιμοποιώντας και την ανισότητα

$$\binom{N}{m} \leqslant \left(\frac{eN}{m}\right)^m,$$

παίρνουμε τον επόμενο ισχυρισμό μας.

Ισχυρισμός 4.1.8 (κάλυψη του συνόλου των αραιών μοναδιαίων διανυσμάτων με Ευκλείδειες μπάλες ακτίνας ϵ). Για κάθε $1 \leqslant m \leqslant N$ και $\epsilon \in (0,1)$ ισχύει ότι

$$N(S_2(\Sigma_m), \epsilon B_2^N) \leqslant \left(\frac{3eN}{m\epsilon}\right)^m.$$

Η μέθοδος της ℓ_1-ελαχιστοποίησης. Επιστρέφοντας στο πρόβλημα της ακριβούς ανακατασκευής, αν θέλουμε να λύσουμε το σύστημα

$$At = y$$

όπου μας έχει δοθεί το $y = Ax$ και το x είναι m-αραιό, μπορούμε να σκεφτούμε να δοκιμάσουμε όλους τους πιθανούς φορείς του άγνωστου διανύσματος x. Αυτή είναι η λεγόμενη ℓ_0-μέθοδος. Υπάρχουν όμως $\binom{N}{m}$ πιθανοί φορείς, οι οποίοι είναι πάρα πολλοί αν πράγματι επιδιώκουμε έναν γρήγορο αλγόριθμο. Έχει προταθεί μια πιο έξυπνη προσέγγιση, η «κυρτή χαλάρωση» της ℓ_0-μεθόδου. Τα δεδομένα είναι $y = Ax$. Για κάθε $t = (t_i) \in \mathbb{R}^N$ συμβολίζουμε με

$$\|t\|_1 = \sum_{i=1}^N |t_i|$$

την ℓ_1-νόρμα του. Η μέθοδος της ℓ_1-ελαχιστοποίησης (η οποία λέγεται και *αναζήτηση βάσης*) είναι το ακόλουθο πρόβλημα:

$$\text{(P)} \qquad \min_{t \in \mathbb{R}^N} \|t\|_1 \quad \text{με τον περιορισμό} \quad At = y.$$

Το πρόβλημα αυτό μπορεί να αναδιατυπωθεί ως πρόβλημα γραμμικού προγραμματισμού:

$$\min \sum_{i=1}^N s_i \quad \text{με τους περιορισμούς} \quad s \geqslant 0, -s \leqslant t \leqslant s, At = y.$$

Ορισμός 4.1.9 (ακριβής ανακατασκευή με ℓ_1-ελαχιστοποίηση). Λέμε ότι ο πίνακας A έχει την ιδιότητα ακριβούς ανακατασκευής τάξης m με ℓ_1-ελαχιστοποίηση αν για κάθε $x \in \Sigma_m$ το

πρόβλημα

$$\text{(P)} \qquad \min_{t\in\mathbb{R}^N} \|t\|_1 \quad \text{με τον περιορισμό} \quad At = Ax \tag{4.1}$$

έχει μοναδική λύση ίση με x.

Σημειώνουμε ότι αυτή η ιδιότητα είναι περισσότερο μια ιδιότητα του πυρήνα του A παρά του ίδιου του A. Για να τονίσουμε αυτό το γεγονός, εισάγουμε κάποιον συμβολισμό.

Για κάθε υποσύνολο $I \subset [N]$, όπου $[N] = \{1, \dots, N\}$, γράφουμε I^c για το συμπλήρωμα του I, και για κάθε $x \in \mathbb{R}^N$ συμβολίζουμε με x_I το διάνυσμα στον $\mathbb{R}^N$ που έχει τις ίδιες συντεταγμένες με το x για τους δείκτες στο I και 0 για τους δείκτες στο I^c. Μπορούμε τώρα να διατυπώσουμε ένα κριτήριο για τον πυρήνα.

Πρόταση 4.1.10 (ιδιότητα πυρήνα). *Τα εξής είναι ισοδύναμα:*

(i) *Για κάθε $x \in \Sigma_m$ το πρόβλημα*

$$\text{(P)} \qquad \min_{t\in\mathbb{R}^N} \|t\|_1 \quad \textit{με τον περιορισμό} \quad At = Ax$$

έχει μοναδική λύση ίση με x. Δηλαδή, ο A έχει την ιδιότητα ακριβούς ανακατασκευής τάξης m με ℓ_1-ελαχιστοποίηση.

(ii) *Για κάθε $w \in \ker(A)$, $w \neq 0$ και για κάθε $I \subset [N]$ με $|I| \leqslant m$,*

$$\|w_I\|_1 < \|w_{I^c}\|_1. \tag{4.2}$$

Απόδειξη. Για να δείξουμε ότι (i) $\Longrightarrow$ (ii), θεωρούμε $w \in \ker(A)$, $w \neq 0$ και $I \subset [N]$ με $|I| \leqslant m$. Θέτουμε $x = -w_I$. Τότε $x \in \Sigma_m$ και από την (4.1) βλέπουμε ότι $\|x + w\|_1 > \|x\|_1$, δηλαδή

$$\|w_{I^c}\|_1 > \|w_I\|_1.$$

Για την αντίστροφη συνεπαγωγή, υποθέτουμε ότι για κάθε $w \in \ker(A)$, $w \neq 0$ και για κάθε $I \subset [N]$ με $|I| \leqslant m$, ισχύει ότι $\|w_I\|_1 < \|w_{I^c}\|_1$. Θεωρούμε $x \in \Sigma_m$ και θέτουμε $I = \operatorname{supp}(x)$. Τότε $|I| \leqslant m$ και

$$\|x + w\|_1 = \|x + w_I\|_1 + \|w_{I^c}\|_1 > \|x_I + w_I\|_1 + \|w_I\|_1 \geqslant \|x\|_1.$$

□

Ορισμός 4.1.11. Λέμε ότι ο A έχει την ιδιότητα πυρήνα τάξης m αν ικανοποιείται η (4.2).

Αυτή η ιδότητα έχει μια ωραία γεωμετρική ερμηνεία. Για να την εισαγάγουμε, χρειαζόμαστε περαιτέρω συμβολισμό. Έστω $(e_i)_{1\leqslant i\leqslant N}$ η κανονική βάση του $\mathbb{R}^N$. Έστω ℓ_1^N ο N-διάστατος χώρος $\mathbb{R}^N$ εφοδιασμένος με την ℓ_1-νόρμα και B_1^N η μοναδιαία μπάλα του. Θέτουμε επίσης

$$S_1(\Sigma_m) = \{x \in \Sigma_m : \|x\|_1 = 1\} \quad \text{και} \quad B_1(\Sigma_m) = \{x \in \Sigma_m : \|x\|_1 \leqslant 1\}.$$

Έστω $1 \leqslant m \leqslant N$. Κάθε $(m-1)$-διάστατη έδρα της B_1^N είναι της μορφής $\mathrm{conv}(\{\epsilon_i e_i : i \in I\})$, όπου $I \subset [N]$ με $|I| = m$ και $(\epsilon_i) \in \{-1,1\}^I$, και $\mathrm{conv}(M)$ είναι η κυρτή θήκη ενός συνόλου $M \subset \mathbb{R}^N$. Από τη γεωμετρική σκοπιά, το $S_1(\Sigma_m)$ είναι η ένωση όλων των $(m-1)$-διάστατων εδρών της B_1^N.

Έστω A ένας $n \times N$ πίνακας και $X_1, \dots, X_N \in \mathbb{R}^n$ οι στήλες του. Ένα πολύτοπο $P \subset \mathbb{R}^n$ λέγεται κεντρικά συμμετρικό αν $-P = P$. Παρατηρούμε ότι

$$A(B_1^N) = \mathrm{conv}(\pm X_1, \dots, \pm X_N).$$

Μπορούμε τώρα να επαναδιατυπώσουμε την Πρόταση 4.1.10 σε γεωμετρική γλώσσα.

Πρόταση 4.1.12 (γεωμετρία των εδρών του $A(B_1^N)$). *Έστω $1 \leqslant m \leqslant n \leqslant N$. Έστω A ένας $n \times N$ πίνακας με στήλες $X_1, \dots, X_N \in \mathbb{R}^n$. Τότε ο A έχει την ιδιότητα πυρήνα τάξης m αν και μόνο αν για κάθε $I \subset [N]$ με $1 \leqslant |I| \leqslant m$ και για κάθε $(\epsilon_i) \in \{-1,1\}^I$ ισχύει ότι*

$$\mathrm{conv}(\{\epsilon_i X_i : i \in I\}) \cap \mathrm{conv}(\{\pm X_j : j \notin I\}) = \varnothing. \tag{4.3}$$

Εκμεταλλευόμενοι τις συμμετρίες, βλέπουμε ότι η ιδιότητα (4.3) είναι ισοδύναμη με την εξής: για κάθε $I \subset [N]$ με $1 \leqslant |I| \leqslant m$ και για κάθε $(\epsilon_i) \in \{-1,1\}^I$ ισχύει ότι

$$\mathrm{Aff}(\{\epsilon_i X_i : i \in I\}) \cap \mathrm{conv}(\{\pm X_j : j \notin I\}) = \varnothing, \tag{4.4}$$

όπου $\mathrm{Aff}(\{\epsilon_i X_i : i \in I\})$ είναι ο αφφινικός υπόχωρος που παράγεται από το $\{\epsilon_i X_i : i \in I\}$.

Η απόδειξη αυτής της ισοδυναμίας είναι μια απλή άσκηση.

Ορισμός 4.1.13. Έστω $1 \leqslant m \leqslant n$. Λέμε ότι ένα κεντρικά συμμετρικό πολύτοπο $P \subset \mathbb{R}^n$ είναι m-γειτνιάζον αν κάθε σύνολο m κορυφών του, το οποίο δεν περιέχει ζεύγος αντιποδικών σημείων, είναι το σύνολο όλων των κορυφών κάποιας έδρας του P.

Παρατηρούμε ότι κάθε κεντρικά συμμετρικό πολύτοπο είναι 1-γειτνιάζον. Η ιδιότητα της γειτνίασης γίνεται μη τετριμμένη όταν $m \geqslant 2$. Σε αυτήν την περίπτωση, παρατηρούμε ότι το $A(B_1^N) = \mathrm{conv}(\pm X_1, \dots, \pm X_N)$ έχει τα $\pm X_1, \dots, \pm X_N$ ως κορυφές και είναι m-γειτνιάζον αν και μόνο αν ο A απεικονίζει κάθε $(m-1)$-διάστατη έδρα της B_1^N σε μια $(m-1)$-διάστατη έδρα του $\mathrm{conv}(\pm X_1, \dots, \pm X_N)$. Από τις (4.3) και (4.4) παίρνουμε το ακόλουθο κριτήριο.

Πρόταση 4.1.14. *Έστω $1 \leqslant m \leqslant n$. Ο πίνακας A έχει την ιδιότητα πυρήνα τάξης m αν και μόνο αν οι στήλες του $\pm X_1, \dots, \pm X_N$ είναι οι $2N$ κορυφές του $A(B_1^N)$ και το $A(B_1^N)$ είναι m-γειτνιάζον.*

Θεωρούμε την απεικόνιση πηλίκο

$$Q : \ell_1^N \longrightarrow \ell_1^N / \ker(A).$$

Αν ο A έχει μέγιστη τάξη n, τότε ο $\ell_1^N/\ker(A)$ είναι n-διάστατος. Συμβολίζουμε με $\|\cdot\|$ τη νόρμα

του χώρου πηλίκου $\ell_1^N/\ker(A)$ που ορίζεται από την

$$\|Qx\| = \min_{w\in\ker(A)} \|x+w\|_1.$$

Όπως γίνεται συχνά στη θεωρία προσέγγισης, μπορούμε να επωφεληθούμε από μια τέτοια ισομετρική εμφύτευση για να φράξουμε την πολυπλοκότητα, συγκρίνοντας τη μετρική εντροπία του αρχικού χώρου $(\Sigma_{\lfloor m/2\rfloor}, \ell_1^N)$ με αυτήν της εικόνας, η οποία ζει σε πολύ χαμηλότερη διάσταση.

Το επόμενο λήμμα στοίχισης είναι πολύ γνωστό.

Λήμμα 4.1.15. *Υπάρχει μια οικογένεια Λ υποσυνόλων του $[N]$ με πληθάριθμο $m \leqslant N/2$ τέτοια ώστε για για κάθε $I, J \in \Lambda$ με $I \neq J$ ισχύει ότι $|I\cap J| \leqslant \lfloor m/2\rfloor$ και*

$$|\Lambda| \geqslant \Big\lfloor \frac{N}{32em}\Big\rfloor^{\lfloor m/2\rfloor}.$$

Απόδειξη. Χωρίς περιορισμό της γενικότητας, υποθέτουμε ότι ο $m/2$ είναι ακέραιος. Επιλέγουμε τυχόν υποσύνολο I_1 του $[N]$ με πληθάριθμο m και πετάμε όλα τα υποσύνολα J του $[N]$ που έχουν πληθάριθμο m και απόσταση Hamming $|I_1 \triangle J| \leqslant m$, όπου $\triangle$ είναι η συμμετρική διαφορά. Υπάρχουν το πολύ

$$\sum_{k=m/2}^{m} \binom{m}{k}\binom{N-m}{m-k}$$

τέτοια υποσύνολα, και αφού $m \leqslant N/2$ έχουμε

$$\sum_{k=m/2}^{m} \binom{m}{k}\binom{N-m}{m-k} \leqslant 2^m \max_{m/2\leqslant k\leqslant m} \binom{N-m}{m-k} \leqslant 2^m \binom{N}{m/2}.$$

Τώρα, επιλέγουμε ένα νέο υποσύνολο I_2 με πληθάριθμο m, από τα υποσύνολα που έχουν απομείνει. Επαναλαμβάνοντας αυτό το επιχείρημα, παίρνουμε μια οικογένεια $\Lambda = \{I_1, I_2, \dots, I_p\}$, όπου $p = |\Lambda|$, από υποσύνολα πληθαρίθμου m τα οποία είναι $(m/2)$-διαχωρισμένα ως προς την απόσταση Hamming και τέτοια ώστε

$$|\Lambda| \geqslant \Big\lfloor \binom{N}{m} \Big/ 2^m \binom{N}{m/2} \Big\rfloor.$$

Εφόσον για $m \leqslant N/2$ έχουμε

$$\Big(\frac{N}{2m}\Big)^m \leqslant \binom{N}{m} \leqslant \Big(\frac{eN}{m}\Big)^m,$$

συμπεραίνουμε ότι

$$|\Lambda| \geqslant \Big\lfloor \frac{(N/2m)^m}{2^m(eN/(m/2))^{m/2}}\Big\rfloor \geqslant \Big\lfloor \Big(\frac{N}{32em}\Big)^{m/2}\Big\rfloor \geqslant \Big\lfloor \frac{N}{32em}\Big\rfloor^{\lfloor m/2\rfloor},$$

και έχουμε το ζητούμενο. □

Έστω Λ η οικογένεια που κατασκευάσαμε στο προηγούμενο λήμμα. Για κάθε $I \in \Lambda$ ορίζουμε

$$x(I) = \frac{1}{m} \sum_{i \in I} e_i.$$

Τότε $x(I) \in S_1(\Sigma_m)$ και για κάθε $I, J \in \Lambda$ με $I \neq J$ έχουμε

$$\|x(I) - x(J)\|_1 = 2\left(1 - \frac{|I \cap J|}{m}\right) \geqslant 2\left(1 - \frac{\lfloor m/2 \rfloor}{m}\right) \geqslant 1.$$

Αν ο πίνακας A έχει την ιδιότητα ακριβούς ανακατασκευής τάξης m, τότε για κάθε $I, J \in \Lambda$ με $I \neq J$ έχουμε

$$\|Q(x(I)) - Q(x(J))\| = \|Q(x(I) - x(J))\| = \|x(I) - x(J)\|_1 \geqslant 1.$$

Από τη μία πλευρά έχουμε $|\Lambda| \geqslant \left\lfloor \frac{CN}{\lfloor m/2 \rfloor} \right\rfloor^{\lfloor m/2 \rfloor}$, όμως από την άλλη πλευρά ο πληθάριθμος του συνόλου $(Q(x(I)))_{I \in \Lambda}$ δεν μπορεί να είναι πολύ μεγάλος. Πράγματι, είναι ένα υποσύνολο της μοναδιαίας μπάλας $Q(B_1^N)$ του χώρου πηλίκου και είδαμε ότι ο πληθάριθμος ενός 1-διαχωρισμένου υποσυνόλου μιας μοναδιαίας μπάλας είναι το πολύ ίσος με 3^n. Έπεται ότι

$$\lfloor N/(32em) \rfloor^{\lfloor m/2 \rfloor} \leqslant 3^n.$$

Πρόταση 4.1.16. *Αν ο πίνακας A έχει την ιδιότητα ελάχιστης ανακατασκευής τάξης m με ℓ_1-ελαχιστοποίηση, τότε*

$$m \log(cN/m) \leqslant Cn,$$

όπου $C, c > 0$ είναι απόλυτες σταθερές.

Για οποιονδήποτε πίνακα A, αυτή η πρόταση μας δίνει άνω φράγμα για την τάξη m της αραιότητας που μας επιτρέπει να ανακατασκευάσουμε ακριβώς οποιοδήποτε διάνυσμα από το Σ_m με τη μέθοδο της ℓ_1-ελαχιστοποίησης.

Σχετική Βιβλιογραφία. Περισσότερες πληροφορίες για την προϊστορία της ανάκτησης σήματος και τη μέθοδο της ℓ_1-ελαχιστοποίησης υπάρχουν στα άρθρα των Donoho [12], Candes, Romberg και Tao [8], Candes και Tao [7]. Το βιβλίο [17] είναι αφιερωμένο σε αυτό το θέμα.

Η Πρόταση 4.1.14 περιέχεται στο άρθρο του Donoho [11]. Η Πρόταση 4.1.16, με αυτήν την απόδειξη, είναι ειδική περίπτωση ενός πιο γενικού αποτελέσματος από το [16]. Στο [26] γίνεται μελέτη του αντίστοιχου προβλήματος για την γειτνίαση.

4.2 Η ιδιότητα της περιορισμένης ισομετρίας

Ως τώρα, δεν γνωρίζουμε καμία «απλή» συνθήκη με την οποία να ελέγχουμε αν ένας πίνακας A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής (4.1). Ξεκινάμε με τον επόμενο ορισμό, ο οποίος παίζει σημαντικό ρόλο στην επαναφορά σήματος.

Ορισμός 4.2.1. Έστω A ένας $n \times N$ πίνακας. Για κάθε $0 \leqslant p \leqslant N$, η σταθερά περιορισμένης ισομετρίας του A τάξης p είναι ο μικρότερος αριθμός $\delta_p = \delta_p(A)$ με την ιδιότητα ότι

$$(1-\delta_p)\|x\|_2^2 \leqslant \|Ax\|_2^2 \leqslant (1+\delta_p)\|x\|_2^2$$

για όλα τα p-αραιά διανύσματα $x \in \mathbb{R}^N$. Έστω $\delta \in (0,1)$. Λέμε ότι ο πίνακας A ικανοποιεί την ιδιότητα περιορισμένης ισομετρίας τάξης p με σταθερά δ (για συντομία γράφουμε $\mathrm{RIP}_p(\delta)$, αν $0 \leqslant \delta_p(A) < \delta$.

Η σημασία της παραμέτρου της περιορισμένης ισομετρίας αποκαλύπτεται από το επόμενο θεώρημα.

Θεώρημα 4.2.2. *Έστω $1 \leqslant m \leqslant N/2$ και A ένας $n \times N$ πίνακας. Αν*

$$\delta_{2m}(A) < \sqrt{2}-1,$$

τότε ο A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης.

Για απλότητα, θα συζητήσουμε μια άλλη παράμετρο στην οποία υπεισέρχεται μια πιο γενική έννοια. Στόχος μας είναι να εξασθενίσουμε τον περιορισμό $\delta_{2m}(A) < \sqrt{2}-1$ στο Θεώρημα 4.2.2 και να εξακολουθούμε να έχουμε μια ιδιότητα ακριβούς ανακατασκευής κάποιας τάξης με τη μέθοδο της ℓ_1-ελαχιστοποίησης.

Ορισμός 4.2.3. Έστω $0 \leqslant p \leqslant n$ ακέραιοι και A ένας $n \times N$ πίνακας. Θέτουμε $\alpha_p = \alpha_p(A)$ και $\beta_p = \beta_p(A)$ τις καλύτερες σταθερές για τις οποίες ισχύει ότι

$$\alpha_p\|x\|_2 \leqslant \|Ax\|_2 \leqslant \beta_p\|x\|_2 \quad \text{για κάθε} \quad x \in \Sigma_p.$$

Δηλαδή,

$$\beta_p = \max\{\|Ax\|_2 : x \in \Sigma_p, \|x\|_2 = 1\} \quad \text{και} \quad \alpha_p = \min\{\|Ax\|_2 : x \in \Sigma_p, \|x\|_2 = 1\}.$$

Ορίζουμε επίσης την παράμετρο $\gamma_p = \gamma_p(A)$ θέτοντας

$$\gamma_p(A) := \frac{\beta_p(A)}{\alpha_p(A)}.$$

Με άλλα λόγια, έστω $I \subset [N]$ με $|I| = p$. Συμβολίζουμε με A^I τον $n \times p$ πίνακα με στήλες $(X_i)_{i\in I}$ ο οποίος προκύπτει αν από τον A κρατήσουμε τις στήλες X_i με δείκτη $i \in I$. Τότε, η σταθερά α_p είναι η μικρότερη ιδιάζουσα τιμή ανάμεσα σε όλους τους υποπίνακες A^I με $|I| = p$, και η σταθερά β_p είναι η μεγαλύτερη. Με άλλα λόγια, αν συμβολίσουμε με B^T τον ανάστροφο πίνακα ενός πίνακα B και με $\lambda_{\min}((A^I)^T A^I)$, αντίστοιχα $\lambda_{\max}((A^I)^T A^I)$, τη μικρότερη και τη μεγαλύτερη ιδιοτιμή του $(A^I)^T A^I$, τότε

$$\alpha_p^2 = \alpha_p^2(A) = \min_{I\subset[N],\,|I|=p} \lambda_{\min}((A^I)^T A^I)$$

και

$$\beta_p^2 = \beta_p^2(A) = \max_{I\subset[N],|I|=p} \lambda_{\max}((A^I)^T A^I).$$

Παρατηρήστε ότι αν ο A ικανοποιεί την $\mathrm{RIP}_p(\delta)$ τότε $\gamma_p(A)^2 \leqslant \frac{1+\delta}{1-\delta}$. Η έννοια της RIP δεν είναι ομογενής, με την έννοια ότι ο A μπορεί να ικανοποιεί την $\mathrm{RIP}_p(\delta)$ αλλά ένα πολλαπλάσιο του A όχι. Μπορούμε να «αλλάξουμε κλίμακα» έτσι ώστε ο πίνακας να ικανοποιεί την ιδιότητα περιορισμένης ισομετρίας. Αυτό δεν εξασφαλίζει ότι ο νέος πίνακας, ας τον πούμε A', θα ικανοποιεί την $\delta_{2m}(A') < \sqrt{2}-1$ και έτσι δεν μας επιτρέπει να συμπεράνουμε ακριβή ανακατασκευή από το Θεώρημα 4.2.2 (κάνετε τη σύγκριση με το Πόρισμα 4.3.3 στην επόμενη ενότητα). Σημειώνουμε επίσης ότι η ιδιότητα περιορισμένης ισομετρίας για τον A μπορεί να γραφτεί στη μορφή

$$\max_{x\in S_2(\Sigma_p)} |\,\|Ax\|_2^2 - 1| \leqslant \delta,$$

εκφράζοντας κάποιο είδος ιδιότητας συγκέντρωσης για την $\|Ax\|_2$. Μια τέτοια ιδιότητα μπορεί να μην ικανοποιείται ακόμα κι αν ο A πράγματι ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης p με τη μέθοδο της ℓ_1-ελαχιστοποίησης (δείτε το Παράδειγμα 4.5.6).

Σχετική Βιβλιογραφία. Ο Ορισμός 4.2.1 της ιδιότητας περιορισμένης ισομετρίας δόθηκε στο [6] και παίζει σημαντικό ρόλο στην ανάκτηση σήματος. Η σημασία της παραμέτρου περιορισμένης ισομετρίας για την ιδιότητα ανακατασκευής έγινε φανερή στα [7] και [6], όπου αποδείχθηκε ότι αν

$$\delta_m(A) + \delta_{2m}(A) + \delta_{3m}(A) < 1$$

τότε ο πίνακας κωδικοποίησης A έχει την ιδιότητα ακριβούς ανακατασκευής τάξης m. Αυτή η εκτίμηση βελτιώθηκε στο [5] σε $\delta_{2m}(A) < \sqrt{2}-1$, όπως διατυπώνεται στο Θεώρημα 4.2.2. Η σταθερά $\sqrt{2}-1$ βελτιώθηκε στο [15], όπου εισήχθη και η παράμετρος γ_p του Ορισμού 4.2.3.

4.3 Η γεωμετρία του πυρήνα

Έστω $1 \leqslant m \leqslant p \leqslant N$. Έστω $w \in \mathbb{R}^N$ και $\varphi = \varphi_w : [N] \to [N]$ μια μετάθεση που ορίζει μια φθίνουσα αναδιάταξη της $(|w_i|)$, δηλαδή $|w_{\varphi(1)}| \geqslant |w_{\varphi(2)}| \geqslant \cdots \geqslant |w_{\varphi(N)}|$. Θέτουμε $I_1 = \varphi_w(\{1,\dots,m\})$ (το σύνολο των δεικτών των m μεγαλύτερων συντεταγμένων του $(|w_i|)$), στη συνέχεια θέτουμε $I_2 = \varphi_w(\{m+1,\dots,m+p\})$ (το σύνολο των δεικτών των επόμενων p μεγαλύτερων συντεταγμένων του $(|w_i|)$), και διαδοχικά ορίζουμε $I_{k+1} = \varphi_w(\{m+(k-1)p+1,\dots,m+kp\})$ όσο ισχύει ότι $m+kp \leqslant N$, φτιάχνοντας μια διαμέριση του $[N]$ σε υποσύνολα πληθαρίθμου p, εκτός από το I_1 που έχει πληθάριθμο m και το τελευταίο σύνολο δεικτών που μπορεί να έχει πληθάριθμο το πολύ ίσο με p.

Ισχυρισμός 4.3.1. Έστω $w \in \mathbb{R}^N$. Υποθέτουμε ότι $1 \leqslant m \leqslant p \leqslant N$ και $N \geqslant m+p$. Με τον παραπάνω συμβολισμό ισχύει ότι

$$\|w_{I_{k+1}}\|_2 \leqslant \frac{1}{\sqrt{p}}\|w_{I_k}\|_1$$

για κάθε $k \geqslant 2$, και

$$\sum_{k\geqslant 3} \|w_{I_k}\|_2 \leqslant \frac{1}{\sqrt{p}} \|w_{I_1^c}\|_1.$$

Απόδειξη. Έστω $k \geqslant 1$. Έχουμε

$$\|w_{I_{k+1}}\|_2 \leqslant \sqrt{|I_{k+1}|}\, \max\{|w_i| : i \in I_{k+1}\}$$

και

$$\max\{|w_i| : i \in I_{k+1}\} \leqslant \min\{|w_i| : i \in I_k\} \leqslant \|w_{I_k}\|_1/|I_k|.$$

Έπεται ότι, για κάθε $k \geqslant 1$,

$$\|w_{I_{k+1}}\|_2 \leqslant \frac{\sqrt{|I_{k+1}|}}{|I_k|}\, \|w_{I_k}\|_1.$$

Προσθέτοντας όλες αυτές τις ανισότητες για όλα τα $k \geqslant 2$, και χρησιμοποιώντας το γεγονός ότι $\sqrt{|I_{k+1}|}/|I_k| \leqslant 1/\sqrt{p}$ από τον ορισμό των I_k, παίρνουμε το ζητούμενο. □

Μπορούμε τώρα να δείξουμε το κεντρικό θεώρημα αυτής της ενότητας.

Θεώρημα 4.3.2. *Έστω $1 \leqslant m \leqslant p \leqslant N$ και $N \geqslant m + p$. Έστω A ένας $n \times N$ πίνακας. Τότε, για κάθε $w \in \ker(A)$, $w \neq 0$ και για κάθε $I \subset [N]$ με $|I| \leqslant m$,*

$$\|w_I\|_1 \leqslant \sqrt{\frac{m}{p}}\, \gamma_{2p}(A)\, \|w_{I^c}\|_1, \tag{4.5}$$

και για κάθε $w \in \ker(A)$, $w \neq 0$ και για κάθε $I \subset [N]$ με $|I| \leqslant m$,

$$\|w\|_2 \leqslant \sqrt{\frac{1+\gamma_{2p}^2(A)}{p}}\, \|w_{I^c}\|_1 \leqslant \sqrt{\frac{1+\gamma_{2p}^2(A)}{p}}\, \|w\|_1. \tag{4.6}$$

Ειδικότερα,

$$\mathrm{rad}(\ker(A) \cap B_1^N) \leqslant \sqrt{\frac{1+\gamma_{2p}^2(A)}{p}},$$

όπου $\mathrm{rad}(B) = \sup_{x\in B} \|x\|_2$.

Απόδειξη. Θεωρούμε $w \in \ker(A)$, $w \neq 0$ και οργανώνουμε τις συντεταγμένες του w όπως παραπάνω. Από τον ορισμό του α_{2p} (βλέπε Ορισμό 4.2.3), έχουμε

$$\|w_{I_1} + w_{I_2}\|_2 \leqslant \frac{1}{\alpha_{2p}} \|A(w_{I_1} + w_{I_2})\|_2.$$

Εφόσον $w \in \ker(A)$, παίρνουμε

$$\|w_{I_1} + w_{I_2}\|_2 \leqslant \frac{1}{\alpha_{2p}} \|A(w_{I_1} + w_{I_2} - w)\|_2 = \frac{1}{\alpha_{2p}} \Big\|A\Big(-\sum_{k\geqslant 3} w_{I_k}\Big)\Big\|_2.$$

Τότε, από τον ορισμό των β_p και γ_p (βλέπε Ορισμό 4.2.3), χρησιμοποιώντας τον Ισχυρισμό 4.3.1, παίρνουμε

$$\|w_{I_1}\|_2 < \|w_{I_1} + w_{I_2}\|_2 \leqslant \frac{\beta_p}{\alpha_{2p}} \sum_{k\geqslant 3} \|w_{I_k}\|_2 \leqslant \frac{\gamma_{2p}(A)}{\sqrt{p}} \|w_{I^c}\|_1. \tag{4.7}$$

Η πρώτη ανισότητα είναι γνήσια γιατί αν $w_{I_2} = 0$ τότε $w_{I_1^c} = 0$ και συνεπώς $w_{I_1} = 0$. Για να ολοκληρώσουμε την απόδειξη της (4.5) παρατηρούμε ότι για κάθε υποσύνολο $I \subset [N]$ με $|I| \leqslant m$ έχουμε $\|w_{I_1^c}\|_1 \leqslant \|w_{I^c}\|_1$ και $\|w_I\|_1 \leqslant \|w_{I_1}\|_1 \leqslant \sqrt{m}\|w_{I_1}\|_2$.

Για να αποδείξουμε την (4.6) ξεκινάμε από την

$$\|w\|_2^2 = \|w - w_{I_1} - w_{I_2}\|_2^2 + \|w_{I_1} + w_{I_2}\|_2^2.$$

Χρησιμοποιώντας τον Ισχυρισμό 4.3.1 βλέπουμε ότι ο πρώτος όρος ικανοποιεί την

$$\|w - w_{I_1} - w_{I_2}\|_2 \leqslant \sum_{k\geqslant 3} \|w_{I_k}\|_2 \leqslant \frac{1}{\sqrt{p}} \|w_{I^c}\|_1.$$

Από την (4.7) έχουμε $\|w_{I_1}+w_{I_2}\|_2 \leqslant \frac{\gamma_{2p}(A)}{\sqrt{p}}\|w_{I_1^c}\|_1$, και συνδυάζοντας όλα τα παραπάνω παίρνουμε

$$\|w\|_2 \leqslant \sqrt{\frac{1+\gamma_{2p}^2(A)}{p}}\, \|w_{I_1^c}\|_1 \leqslant \sqrt{\frac{1+\gamma_{2p}^2(A)}{p}}\, \|w\|_1.$$

□

Από την (4.5) και την ιδιότητα του πυρήνα (Πρόταση 4.1.10) παίρνουμε το ακόλουθο πόρισμα.

Πόρισμα 4.3.3. *Έστω $1 \leqslant p \leqslant N/2$. Έστω A ένας $n\times N$ πίνακας. Αν $\gamma_{2p}(A) \leqslant \sqrt{p}$ τότε ο A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης, όπου*

$$m = \lfloor p/\gamma_{2p}^2(A) \rfloor.$$

Βασικός μας στόχος είναι να βρούμε p τέτοιον ώστε η γ_{2p} να φράσσεται από μια απόλυτη σταθερά. Αυτό σημαίνει ότι χρειαζόμαστε ομοιόμορφο έλεγχο για τη μικρότερη και τη μεγαλύτερη ιδιάζουσα τιμή όλων των υποπινάκων του A με $2p$ στήλες. Από το Πόρισμα 4.3.3 αυτή είναι μια ικανή συνθήκη για την ακριβή ανακατασκευή των m-αραιών διανυσμάτων με τη μέθοδο της ℓ_1-ελαχιστοποίησης με $m \sim p$. Όταν η $\|Ax\|_2$ έχει καλές ιδιότητες συγκέντρωσης, η ιδιότητα περιορισμένης ισομετρίας είναι καλύτερα προσαρμοσμένη. Σε αυτήν την κατάσταση έχουμε $\gamma_{2p} \sim 1$. Όταν η σταθερά ισομετρίας δ_{2p} είναι αρκετά μικρή, τότε ο A κανοποιεί την ακριβή ανακατασκευή των m-αραιών διανυσμάτων με $m = p$ (βλέπε Θεώρημα 4.2.2).

Όμοια, μια εκτίμηση για την $\mathrm{rad}(\ker(A) \cap B_1^N)$ δίνει μια εκτίμηση για το μέγεθος της αραιότητας των διανυσμάτων που μπορούμε να ανακατασκευάσουμε ακριβώς με τη μέθοδο της ℓ_1-ελαχιστοποίησης.

Πρόταση 4.3.4. *Έστω A ένας $n \times N$ πίνακας και $m \geqslant 1$. Αν*

$$\mathrm{rad}(\ker(A) \cap B_1^N) < \frac{1}{2\sqrt{m}},$$

τότε ο A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης.

Απόδειξη. Έστω $w \in \ker(A)$ και $I \subset [N]$ με $|I| \leqslant m$. Από την υπόθεση έχουμε ότι αν $w \in \ker(A)$ και $w \neq 0$ τότε

$$\|w\|_2 < \frac{1}{2\sqrt{m}}\|w\|_1.$$

Άρα,

$$\|w_I\|_1 \leqslant \sqrt{m}\|w_I\|_2 \leqslant \sqrt{m}\|w\|_2 < \frac{1}{2}\|w\|_1$$

και

$$\|w_I\|_1 < \|w_{I^c}\|_1.$$

Το συμπέρασμα έπεται από την ιδιότητα του πυρήνα (Πρόταση 4.1.10). □

Για να ολοκληρώσουμε αυτήν την ενότητα, σημειώνουμε ότι η (4.6) συνεπάγεται ότι αν ένας $n \times N$ πίνακας A ικανοποιεί την ιδιότητα περιορισμένης ισομετρίας τάξης $m \geqslant 1$, τότε $\mathrm{rad}(\ker(A) \cap B_1^N) = \frac{O(1)}{\sqrt{m}}$.

Σχετική Βιβλιογραφία. Οι αποδείξεις των αποτελεσμάτων που παρουσιάζονται σε αυτήν την ενότητα ακολουθούν τα [6], [10], [15], [16] ανδ [24]. Η σχέση (4.5) αποδείχθηκε στο [15] με ακριβέστερη αριθμητική σταθερά.

4.4 Πλάτος Gelfand

Η μελέτη που κάναμε στην προηγούμενη ενότητα μας οδηγεί στην έννοια του πλάτους Gelfand.

Ορισμός 4.4.1. Έστω T ένα φραγμένο υποσύνολο ενός χώρου E με νόρμα και έστω $k \geqslant 0$ ένας ακέραιος. Το k-οστό πλάτος Gelfand του T είναι η ποσότητα

$$d^k(T, E) := \inf_S \sup_{x \in S \cap T} \|x\|_E,$$

όπου $\|\cdot\|_E$ είναι η νόρμα του E και το infimum είναι πάνω από όλους τους διανυσματικούς υπόχωρους S του E που έχουν συνδιάσταση το πολύ ίση με k.

Στη θεωρία χώρων Banach και τη θεωρία τελεστών χρησιμοποιείται ένας διαφορετικός συμβολισμός. Έστω $u : X \to Y$ ένας γραμμικός τελεστής μεταξύ δύο χώρων με νόρμα X και Y. Ο k-οστός αριθμός Gelfand του u ορίζεται ως εξής:

$$c_k(u) = \inf\{\|u\big|_S\| : S \subset X, \mathrm{codim}(S) < k\},$$

όπου $u\big|_S$ είναι ο περιορισμός του u στον διανυσματικό υπόχωρο S. Ισοδύναμα,

$$c_k(u) = \inf_S \sup_{x\in S\cap B_X} \|u(x)\|_Y,$$

όπου B_X είναι η μοναδιαία μπάλα του X και το infimum είναι πάνω από όλους τους υποχώρους S του X συνδιάστασης μικρότερης από k. Οι δύο συμβολισμοί συνδέονται με την

$$c_{k+1}(u) = d^k(u(B_X), Y).$$

Αν F είναι ένας γραμμικός χώρος, για παράδειγμα ο $\mathbb{R}^N$, εφοδιασμένος με δύο νόρμες που ορίζουν δύο χώρους με νόρμα X και Y, και αν $id : X \to Y$ είναι ο ταυτοτικός τελεστής του F, τον οποίο θεωρούμε σαν τελεστή από τον X στον Y, τότε

$$d^k(B_X, Y) = c_{k+1}(if : X \to Y).$$

Ως σημαντική ειδική περίπτωση παίρνουμε την

$$d^k(B_1^N, \ell_2^N) = c_{k+1}(id : \ell_1^N \to \ell_2^N) = \inf_{\operatorname{codim}(S)\leqslant k} \operatorname{rad}(S \cap B_1^N).$$

Ένα σημαντικό αποτέλεσμα για τους αριθμούς Gelfand είναι το εξής.

Θεώρημα 4.4.2. *Υπάρχουν σταθερές $c, C > 0$ τέτοιες ώστε, για οποιουσδήποτε ακεραίους $1 \leqslant k \leqslant N$,*

$$c\,\min\left\{1, \sqrt{\frac{\log(N/k)}{k}}\right\} \leqslant c_k(id : \ell_1^N \to \ell_1^N) \leqslant C\,\min\left\{1, \sqrt{\frac{\log(N/k)}{k}}\right\}.$$

Επιπλέον, αν $\mathbb{P}$ είναι το αναλλοίωτο ως προς στροφές μέτρο πιθανότητας στην πολλαπλότητα Grassmann των υποχώρων S του $\mathbb{R}^N$ με $\operatorname{codim}(S) = k - 1$, τότε

$$\mathbb{P}\left(\operatorname{rad}(S \cap B_1^N) \leqslant C\,\min\left\{1, \sqrt{\frac{\log(N/k)}{k}}\right\}\right) \geqslant 1 - \exp(-ck).$$

Επιστρέφοντας στην επαναφορά σήματος, έστω $1 \leqslant m \leqslant n$ και ας υποθέσουμε ότι

$$d^n(B_1^N, \ell_2^N) < \frac{1}{2\sqrt{m}}.$$

Με άλλα λόγια, υποθέτουμε ότι υπάρχει ένας υπόχωρος $S \subset \mathbb{R}^N$ συνδιάστασης $\leqslant n$ τέτοιος ώστε

$$\operatorname{rad}(S \cap B_1^N) < \frac{1}{2\sqrt{m}}.$$

Επιλέγουμε έναν $n \times N$ πίνακα A τέτοιον ώστε $\ker(A) = S$. Τότε,

$$\mathrm{rad}(\ker(A) \cap B_1^N) < \frac{1}{2\sqrt{m}}.$$

Από την Πρόταση 4.3.4 έπεται ότι ο A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης.

Συμπεραίνουμε έτσι από το Θεώρημα 4.4.2 ότι υπάρχει ένας πίνακας A που ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης

$$\lfloor c_1 n / \log(c_2 N/n) \rfloor,$$

όπου $c_1, c_2 > 0$ είναι απόλυτες σταθερές. Η Πρόταση 4.1.16 δείχνει ότι αυτή είναι η βέλτιστη τάξη.

Σχετική Βιβλιογραφία. Το Θεώρημα 4.4.2 από το [18] δίνει βέλτιστες εκτιμήσεις για τα πλάτη Gelfand του cross-polytope. Έτσι παίρνει τελική μορφή ένα κλασικό αποτέλεσμα από το [23]. Στο [23] δίνονται αρχικά εκτιμήσεις για τα πλάτη Kolmogorov, που είναι δυϊκά των πλατών Gelfand, και με όχι βέλτιστη δύναμη του λογαρίθμου (3/2 αντί του 1/2 που επιτυγχάνεται στο [18]). Το άνω φράγμα για τα πλάτη Kolmogorov παίρνεται με χρήση τυχαίων πινάκων που έχουν ανεξάρτητες και ισόνομες Bernoulli συντεταγμένες, ενώ στα [19] και [18] χρησιμοποιούνται ιδιότητες Gaussian τυχαίων πινάκων.

4.5 Gaussian τυχαίοι πίνακες

Σε αυτήν την ενότητα δείχνουμε ότι οι Gaussian τυχαίοι πίνακες ικανοποιούν με μεγάλη πιθανότητα την ιδιότητα ακριβούς ανακατασκευής τάξης m με το m μεγάλο.

Θεωρούμε ένα μέτρο πιθανότητας $\mathbb{P}$ στον χώρο $M(n,N)$ των πραγματικών $n \times N$ πινάκων που ικανοποιεί την ακόλουθη ανισότητα συγκέντρωσης: υπάρχει απόλυτη σταθερά c_0 τέτοια ώστε για κάθε $x \in \mathbb{R}^N$ να ισχύει

$$\mathbb{P}(\{A : |\,\|Ax\|_2^2 - \|x\|_2^2\,| \geqslant t\|x\|_2^2\}) \leqslant 2e^{-c_0 t^2 n} \quad \text{για κάθε } \ 0 < t \leqslant 1. \tag{4.8}$$

Ορισμός 4.5.1. Υπενθυμίζουμε ότι η ψ_2-νόρμα μιας πραγματικής τυχαίας μεταβλητής Z ορίζεται ως εξής:

$$\|Z\|_{\psi_2} = \inf\{s > 0 : \mathbb{E}(\exp((|Z|/s)^2)) \leqslant 2\}.$$

Λέμε ότι ένα τυχαίο διάνυσμα $Y \in \mathbb{R}^N$ είναι ισοτροπικό αν $\mathbb{E}(Y) = 0$ και για κάθε $y \in \mathbb{R}^N$ ισχύει ότι

$$\mathbb{E}\,|\langle Y, y\rangle|^2 = \|y\|_2^2.$$

Λέμε ότι ένα τυχαίο διάνυσμα $Y \in \mathbb{R}^N$ ικανοποιεί ψ_2-εκτίμηση με σταθερά α (για συντομία,

είναι ψ_2 με σταθερά α) αν για κάθε $y \in \mathbb{R}^N$ ισχύει ότι

$$\|\langle Y, y\rangle\|_{\psi_2} \leqslant \alpha \|y\|_2.$$

Στο Κεφάλαιο 2 είδαμε ότι μια πραγματική τυχαία μεταβλητή Z είναι ψ_2, με κάποια σταθερά, αν και μόνο αν ικανοποιεί υποκανονικές εκτιμήσεις για τις ουρές της. Ειδικότερα, αν η Z είναι πραγματική τυχαία μεταβλητή με $\|Z\|_{\psi_2} \leqslant \alpha$, τότε

$$\mathbb{P}(|Z| \geqslant t) \leqslant 2e^{-(t/\alpha)^2}$$

για κάθε $t > 0$. Είδαμε επίσης ότι η ψ_2-ιδιότητα χαρακτηρίζεται από την $\sqrt{p}$-αυξητικότητα των p-ροπών της Z.

Έστω $Y_1, \dots, Y_n \in \mathbb{R}^N$ ανεξάρτητα ισοκατανεμημένα ισοτροπικά τυχαία διανύσματα τα οποία είναι ψ_2 με σταθερά α. Έστω A ο πίνακας που έχει γραμμές τα $Y_1, \dots, Y_n$. Θεωρούμε το μέτρο πιθανότητας $\mathbb{P}$ στον χώρο πινάκων $M(n, N)$ που επάγεται από την απεικόνιση $(Y_1, \dots, Y_n) \to A$.

Θα χρειαστούμε μια ανισότητα Bernstein που προκύπτει άμεσα από το Θεώρημα 2.7.3. Για $y \in S^{n-1}$ θεωρούμε τον μέσο n ανεξάρτητων αντιγράφων της τυχαίας μεταβλητής $\langle Y_1, y\rangle^2$. Τότε, για κάθε $t > 0$,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n \langle Y_i, y\rangle^2 - 1\right| > t\right) \leqslant 2\exp\left(-cn\min\left\{\frac{t^2}{\alpha^4}, \frac{t}{\alpha^2}\right\}\right),$$

όπου $c > 0$ είναι μια απόλυτη σταθερά. Δεδομένου ότι $\mathbb{E}\langle Y_1, y\rangle^2 = 1$, έχουμε $\alpha \geqslant 1$ και

$$\left\|\frac{1}{\sqrt{n}}Ay\right\|_2^2 = \frac{1}{n}\sum_{i=1}^n \langle Y_i, y\rangle^2.$$

Έτσι, έχουμε δείξει τον ακόλουθο ισχυρισμό.

Ισχυρισμός 4.5.2. Έστω $Y_1, \dots, Y_n \in \mathbb{R}^N$ ανεξάρτητα ισοκατανεμημένα τυχαία διανύσματα τα οποία είναι ψ_2 με σταθερά α. Έστω $\mathbb{P}$ το μέτρο πιθανότητας που επάγεται στον $M(n, N)$. Τότε, για κάθε $x \in \mathbb{R}^N$ ισχύει ότι

$$\mathbb{P}\left(\left|\left\|\frac{1}{\sqrt{n}}Ax\right\|_2^2 - \|x\|_2^2\right| \geqslant t\|x\|_2^2\right) \leqslant 2e^{-\frac{ct^2 n}{\alpha^4}} \quad \text{για κάθε } 0 < t \leqslant 1$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Τα πιο σημαντικά παραδείγματα μοντέλων τυχαίων πινάκων που ικανοποιούν την (4.8) είναι οι τυχαίοι πίνακες με ανεξάρτητες υποκανονικές γραμμικές, κατάλληλα κανονικοποιημένες.

Παράδειγμα 4.5.3. Κλασσικά παραδείγματα:

- Τα $Y_1, \dots, Y_n \in \mathbb{R}^n$ είναι ανεξάρτητα αντίγραφα του Gaussian τυχαίου διανύσματος $Y = (g_1, \dots, g_N)$, όπου οι g_i είναι ανεξάρτητες $N(0,1)$ κανονικές τυχαίες μεταβλητές.
- Τα $Y_1, \dots, Y_n \in \mathbb{R}^n$ είναι ανεξάρτητα αντίγραφα του τυχαίου διανύσματος προσήμων $Y = (\epsilon_1, \dots, \epsilon_N)$, όπου οι ϵ_i είναι συμμετρικές ± 1 τυχαίες μεταβλητές.
- Τα $Y_1, \dots, Y_n \in \mathbb{R}^n$ είναι ανεξάρτητα αντίγραφα ενός τυχαίου διανύσματος που είναι ομοιόμορφα κατανεμημένο στην Ευκλείδεια σφαίρα ακτίνας $\sqrt{N}$.

Σε όλες αυτές τις περιπτώσεις, τα τυχαία διανύσματα Y_i είναι ισοτροπικά, με ψ_2-σταθερά α για κατάλληλη σταθερά $\alpha \geqslant 1$.

Οι υποκανονικοί πίνακες σχεδόν διατηρούν τη νόρμα στο Σ_m. Ένα σημαντικό χαρακτηριστικό του Σ_m, και των υποσυνόλων του $S_2(\Sigma_m)$ και $B_2(\Sigma_m)$, είναι η ιδιαίτερη δομή τους: τα δύο τελευταία σύνολα είναι ενώσεις των μοναδιαίων σφαιρών, ή μπαλών αντίστοιχα, που φέρονται από τους m-διάστατους υπόχωρους συντεταγμένων του $\mathbb{R}^N$.

Υπενθυμίζουμε το ακόλουθο γνωστό λήμμα που επιτρέπει να περάσουμε από ένα ϵ-δίκτυο σε ολόκληρη τη σφαίρα.

Λήμμα 4.5.4. *Έστω $m \geqslant 1$ ένας ακέραιος, $\|\cdot\|$ μια νόρμα στον $\mathbb{R}^m$ και $\epsilon \in \left(0, \frac{1}{3}\right)$. Έστω $\Lambda \subset S^{m-1}$ ένα ϵ-δίκτυο της S^{m-1} με την ϵB_2^m. Αν για κάθε $y \in \Lambda$ ισχύει ότι*

$$1 - \epsilon \leqslant \|y\| \leqslant 1 + \epsilon,$$

τότε για κάθε $y \in S^{m-1}$ ισχύει ότι

$$\frac{1 - 3\epsilon}{1 - \epsilon} \leqslant \|y\| \leqslant \frac{1 + \epsilon}{1 - \epsilon}.$$

Απόδειξη. Από την Πρόταση 4.1.6 έχουμε ότι $S^{m-1} \subset (1-\epsilon)^{-1}\mathrm{conv}(\Lambda)$. Συνεπώς,

$$\sup_{y \in S^{m-1}} \|y\| \leqslant (1+\epsilon)(1-\epsilon)^{-1}.$$

Για να πάρουμε το κάτω φράγμα, γράφουμε το τυχόν $y \in S^{m-1}$ ως $y = y_1 + \epsilon y_2$, όπου $y_1 \in \Lambda$ και $y_2 \in B_2^m$. Τότε,

$$\|y\| \geqslant \|y_1\| - \epsilon\|y_2\| \geqslant (1-\epsilon) - \epsilon(1+\epsilon)(1-\epsilon)^{-1} = (1-3\epsilon)/(1-\epsilon),$$

και έχουμε αποδείξει το λήμμα. □

Μπορούμε τώρα να δώσουμε μια απλή απόδειξη για το γεγονός ότι οι υποκανονικοί πίνακες ικανοποιούν την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης για μεγάλο m.

Θεώρημα 4.5.5. *Έστω $\mathbb{P}$ ένα μέτρο πιθανότητας στον $M(n,N)$ το οποίο ικανοποιεί την (4.8). Τότε υπάρχουν θετικές σταθερές c_1, c_2 και c_3 που εξαρτώνται μόνο από την c_0 της (4.8), για τις*

οποίες ισχύει το ακόλουθο: με πιθανότητα μεγαλύτερη ή ίση από $1-2\exp(-c_3 n)$ *ο* A *ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης* m *με τη μέθοδο της* ℓ_1*-ελαχιστοποίησης, όπου*

$$m = \Big\lfloor \frac{c_1 n}{\log(c_2 N/n)} \Big\rfloor.$$

Επιπλέον, ο A *ικανοποιεί την* $\mathrm{RIP}_m(\delta)$ *για κάθε* $\delta \in (0,1)$*, με* $m \sim c\delta^2 n/\log(CN/\delta^3 n)$*, όπου οι σταθερές* c *και* C *εξαρτώνται μόνο από την* c_0*.*

Απόδειξη. Θεωρούμε $\epsilon \in \left(0, \frac{1}{3}\right)$ το οποίο θα επιλέξουμε κατάλληλα αργότερα. Έστω $1 \leqslant p \leqslant N/2$. Έστω $y_1, \dots, y_n$ οι γραμμές του A. Για κάθε υποσύνολο I του $[N]$ με πληθάριθμο $2p$, έστω Λ_I ένα ϵ-δίκτυο της μοναδιαίας σφαίρας του $\mathbb{R}^I$ με την ϵB_2^I, το οποίο ικανοποιεί τον Ισχυρισμό 4.1.7, δηλαδή έχει πληθάριθμο $|\Lambda_I| \leqslant \left(\frac{3}{\epsilon}\right)^{2p}$. Εφαρμόζουμε το Λήμμα 4.5.4 για την ημινόρμα

$$\|y\| := \left(\frac{1}{n}\sum_{i=1}^n \langle y_i, y\rangle^2\right)^{1/2}$$

στη μοναδιαία σφαίρα του $\mathbb{R}^I$. Έστω $\Lambda \subset \mathbb{R}^N$ η ένωση όλων αυτών των Λ_I για $|I| = 2p$.

Αν υποθέσουμε ότι

$$\sup_{y\in\Lambda} \left|\frac{1}{n}\sum_{i=1}^n (\langle y_i, y\rangle^2 - 1)\right| \leqslant \epsilon,$$

τότε για κάθε $y \in S_2(\Sigma_{2p})$ έχουμε ότι

$$\frac{1-3\epsilon}{1-\epsilon} \leqslant \frac{1}{\sqrt{n}}\left(\sum_{i=1}^n \langle y_i, y\rangle^2\right) \leqslant \frac{1+\epsilon}{1-\epsilon}.$$

Σημειώνουμε ότι σε αυτήν την ανισότητα δεν υπάρχει τυχαιότητα, γι' αυτό και συμβολίζουμε τις γραμμές του A με y_i αντί για Y_i. Η ανισότητα δείχνει ότι για να ελέγξουμε πόσο καλή είναι η δράση του πίνακα A με γραμμές τα y_i στο Σ_{2p} αρκεί να ελέγξουμε τη δράση του στο πεπερασμένο σύνολο Λ. Υπενθυμίζουμε ότι

$$|\Lambda| \leqslant \binom{N}{2p}\left(\frac{3}{\epsilon}\right)^{2p} \leqslant \exp\Big(2p\log\Big(\frac{3eN}{2p\epsilon}\Big)\Big).$$

Θεωρούμε ένα μέτρο πιθανότητας $\mathbb{P}$ στον $M(n,N)$ που ικανοποιεί την (4.8). Τότε, οι ανισότητες

$$\frac{1-3\epsilon}{1-\epsilon} \leqslant \|Ax\|_2 \leqslant \frac{1+\epsilon}{1-\epsilon}, \qquad x \in S_2(\Sigma_{2p})$$

ισχύουν με πιθανότητα μεγαλύτερη ή ίση από

$$1 - 2|\Lambda|e^{-c_0\epsilon^2 n} \geqslant 1 - 2\exp\Big(2p\log\Big(\frac{3eN}{2p\epsilon}\Big)\Big)e^{-c_0\epsilon^2 n} \geqslant 1 - 2e^{-c_0\epsilon^2 n/2}$$

αν ικανοποιείται η

$$2p\log\Big(\frac{3eN}{2p\epsilon}\Big) \leqslant c_0\epsilon^2 n/2.$$

Υποθέτοντας τα παραπάνω συμπεραίνουμε ότι

$$\gamma_{2p}(A) \leqslant (1+\epsilon)/(1-3\epsilon)$$

με πιθανότητα μεγαλύτερη από $1-\exp(-c_0^2\epsilon^2 n/2)$. Από το Πόρισμα 4.3.3 βλέπουμε ότι ο A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης για

$$m = \lfloor p/\gamma_{2p}^2(A) \rfloor.$$

Αν επιλέξουμε $\epsilon = \frac{1}{4}$ και λύσουμε την $2p\log\left(\frac{3eN}{2p\epsilon}\right) \leqslant c_0\epsilon^2 n/2$ παίρνουμε το ζητούμενο. □

Ολοκληρώνουμε αυτήν την ενότητα με ένα παράδειγμα ενός $n \times N$ πίνακα A που είναι καλός πίνακας ανάκτησης σήματος παρόλο που κανένας από τους $n \times N$ πίνακες που έχουν τον ίδιο πυρήνα με τον A δεν ικανοποιεί την ιδιότητα περιορισμένης ισομετρίας οποιασδήποτε τάξης $\geqslant 1$ με καλή παράμετρο. Όπως έχουμε ήδη παρατηρήσει, αν ο A έχει παράμετρο γ_p, μπορούμε να βρούμε $t_0 > 0$ και να αλλάξουμε κλίμακα στον A έτσι ώστε

$$\delta_p(t_0 A) = \gamma_p^2 - \frac{1}{\gamma_p^2} + 1 \in [0,1).$$

Σε αυτό το παράδειγμα, η παράμετρος γ_p είναι μεγάλη, $\delta_p(t_0 A) \sim 1$, και το Θεώρημα 4.2.2 δεν μας δίνει κανένα συμπέρασμα για την ακριβή ανακατασκευή.

Παράδειγμα 4.5.6. Έστω $1 \leqslant n \leqslant N$ και $\delta \in (0,1)$. Υπάρχει $n \times N$ πίνακας A τέτοιος ώστε για κάθε $p \leqslant cn/\log(CN/n)$ να ισχύει ότι

$$\gamma_{2p}^2(A) \leqslant c'(1-\delta)^{-1}.$$

Συνεπώς, για κάθε $m \leqslant c''(1-\delta)n/\log(CN/n)$, ο πίνακας A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής για m-αραιά διανύσματα με τη μέθοδο της ℓ_1-ελαχιστοποίησης. Όμως, για κάθε $n \times n$ πίνακα U, η σταθερά περιορισμένης ισομετρίας τάξης 1 του UA ικανοποιεί την $\delta_1(UA) \geqslant \delta$ (θεωρήστε $\delta \geqslant 1/2$). Στα παραπάνω, οι $C, c, c', c'' > 0$ είναι απόλυτες σταθερές.

Σχετική Βιβλιογραφία. Η απόδειξη του Θεωρήματος 4.5.5 που δείχνει ότι οι υποκανονικοί πίνακες έχουν την ιδιότητα ακριβούς ανακατασκευής τάξης m με τη μέθοδο της ℓ_1-ελαχιστοποίησης για μεγάλα m, προέρχεται από τα [4] και [28]. Η ιδέα της απόδειξης εμφανίζεται συχνά στη θεωρία προσέγγισης, π.χ. στο [23], και στη θεωρία χώρων Banach όπου έχει παίξει σημαντικό ρόλο σε ποσοτικές εκδοχές του θεωρήματος Dvoretzky (βλ. [14] και [29]).

4.6 Σχεδόν αραιά διανύσματα

Όπως έχουμε αναφέρει, διάφοροι τελεστές «τυχαίων προβολών» σχεδόν διατηρούν τη νόρμα σε «λεπτά» υποσύνολα της σφαίρας. Θα μελετήσουμε συνθήκες για τη δομή της μετρικής εντροπίας ενός συνόλου $T \subset \mathbb{R}^N$ έτσι ώστε, με μεγάλη πιθανότητα, οι Gaussian ή υποκανονικοί

τυχαίοι πίνακες να δρουν σχεδόν ως ισομετρίες στο T. Αυτές οι συνθήκες θα μας επιτρέψουν να δώσουμε εφαρμογές σε γενικότερα σύνολα από αυτά των αραιών διανυσμάτων.

Θεώρημα 4.6.1. *Θεωρούμε ένα μέτρο πιθανότητας $\mathbb{P}$ στον χώρο των $n \times N$ πινάκων, το οποίο για κάθε $x \in \mathbb{R}^N$ ικανοποιεί την*

$$\mathbb{P}(|\,\|Ax\|_2^2 - \|x\|_2^2\,| \geqslant t\|x\|_2^2) \leqslant 2e^{-c_0 t^2 n} \quad \text{για κάθε } \ 0 < t \leqslant 1.$$

Έστω $T \subset S^{N-1}$ και $\epsilon \in \left(0, \frac{1}{15}\right)$. Υποθέτουμε ότι:

(i) *Υπάρχει ϵ-δίκτυο $\Lambda \subset S^{N-1}$ του T με $|\Lambda| \leqslant \exp(c_0\epsilon^2 n/2)$.*

(ii) *Υπάρχει υποσύνολο Λ' της ϵB_2^N τέτοιο ώστε $(T-T) \cap \epsilon B_2^N \subset 2\mathrm{conv}(\Lambda')$ και $|\Lambda'| \leqslant \exp(c_0 n/2)$.*

Τότε, με πιθανότητα μεγαλύτερη ή ίση από $1 - 3\exp(-c_0\epsilon^2 n/2)$, ισχύει ότι, για κάθε $x \in T$,

$$1 - 15\epsilon \leqslant \|Ax\|_2^2 \leqslant 1 + 15\epsilon. \tag{4.9}$$

Απόδειξη. Θεωρούμε το σύνολο Ω των πινάκων A που ικανοποιούν τις

$$|\,\|Ax_0\|_2 - 1\,| \leqslant |\,\|Ax_0\|_2^2 - 1\,| \leqslant \epsilon \quad \text{για κάθε } \ x_0 \in \Lambda \tag{4.10}$$

και

$$\|Az\|_2 \leqslant 2\epsilon \quad \text{για κάθε } \ z \in \Lambda'. \tag{4.11}$$

Από την (4.8) και τις υποθέσεις (i) και (ii),

$$\mathbb{P}(\Omega) \geqslant 1 - 2\exp(-c_0\epsilon^2 n/2) - \exp(-c_0 n/2) \geqslant 1 - 3\exp(-c_0\epsilon^2 n/2).$$

Έστω $x \in T$. Υπάρχει $x_0 \in \Lambda$ τέτοιο ώστε $\|x - x_0\|_2 \leqslant \epsilon$. Τότε, για κάθε $A \in \Omega$,

$$\|Ax_0\|_2 - \|A(x-x_0)\|_2 \leqslant \|Ax\|_2 \leqslant \|Ax_0\|_2 + \|A(x-x_0)\|_2.$$

Εφόσον $x - x_0 \in (T-T) \cap \epsilon B_2^N$, από την υπόθεση (ii) και την (4.11) παίρνουμε

$$\|A(x-x_0)\|_2 \leqslant 2 \sup_{z \in \mathrm{conv}(\Lambda')} \|Az\|_2 = 2 \sup_{z \in \Lambda'} \|Az\|_2 \leqslant 4\epsilon. \tag{4.12}$$

Συνδυάζοντας αυτήν την ανισότητα με την (4.10) συμπεραίνουμε ότι

$$1 - 5\epsilon \leqslant \|Ax\|_2 \leqslant 1 + 5\epsilon.$$

Υψώνοντας στο τετράγωνο έχουμε το ζητούμενο. □

Προσεγγιστική ανακατασκευή σχεδόν αραιών διανυσμάτων. Έχοντας αναλύσει την ιδιότητα περιορισμένης ισομετρίας για λεπτά σύνολα όπως το Σ_m, εξετάζουμε πάλι τη μέθοδο

της ℓ_1-ελαχιστοποίησης για να επιτύχουμε προσεγγιστική ανακατασκευή διανυσμάτων που δεν απέχουν πολύ από αραιά διανύσματα. Όπως και η ακριβής ανακατασκευή, η προσεγγιστική ανακατασκευή βασίζεται σε μια ιδιότητα πυρήνα.

Πρόταση 4.6.2. *Έστω A ένας $n\times N$ πίνακας και $\lambda \in (0,1)$. Υποθέτουμε ότι: για κάθε $w \in \ker(A)$, $w \neq 0$ και για κάθε $I \subset [N]$ με $|I| \leqslant m$,*

$$\|w_I\|_1 \leqslant \lambda \|w_{I^c}\|_1. \tag{4.13}$$

Έστω $x \in \mathbb{R}^N$ και x^* ένας ελαχιστοποιητής του

$$\min_{t\in\mathbb{R}^N} \|t\|_1 \quad \text{με τον περιορισμό} \quad At = Ax. \tag{P}$$

Τότε, για κάθε $I \subset [N]$ με $|I| \leqslant m$,

$$\|x - x^*\|_1 \leqslant 2\frac{1+\lambda}{1-\lambda}\|x - x_I\|_1.$$

Απόδειξη. Έστω x^* ένας ελαχιστοποιητής του (P). Θέτουμε $w = x^* - x \in \ker(A)$. Έστω $m \geqslant 1$ και $I \subset [N]$ με $|I| \leqslant m$. Παρατηρούμε ότι

$$\|x\|_1 \geqslant \|x + w\|_1 = \|x_I + w_I\|_1 + \|x_{I^c} + w_{I^c}\|_1 \geqslant \|x_I\|_1 - \|w_I\|_1 + \|w_{I^c}\|_1 - \|x_{I^c}\|_1,$$

και συνεπώς,

$$\|w_{I^c}\|_1 \leqslant \|w_I\|_1 + 2\|x_{I^c}\|_1.$$

Από την άλλη πλευρά, από την υπόθεση για τον πυρήνα, παίρνουμε

$$\|w_{I^c}\|_1 \leqslant \|w_I\|_1 + 2\|x_{I^c}\|_1 \leqslant \lambda\|w_{I^c}\|_1 + 2\|x_{I^c}\|_1.$$

Έπεται ότι

$$\|w_{I^c}\|_1 \leqslant \frac{2}{1-\lambda}\|x_{I^c}\|_1.$$

Παρατηρώντας ότι η υπόθεση για τον πυρήνα είναι ισοδύναμη με την $\|w\|_1 \leqslant (1+\lambda)\|w_{I^c}\|_1$, ολοκληρώνουμε την απόδειξη. □

Σημειώνουμε ότι το minimum της $\|x - x_I\|_1$ πάνω από όλα τα υποσύνολα I με $|I| \leqslant m$ πιάνεται όταν το I είναι ο φορέας των m μεγαλύτερων συντεταγμένων του x. Άρα, το διάνυσμα x_I είναι η βέλτιστη m-αραιή προσέγγιση του x ως προς την ℓ_1-νόρμα. Σημειώνουμε επίσης ότι αν το x είναι m-αραιό τότε επιστρέφουμε στο σχήμα της ακριβούς ανακατασκευής.

Η ιδιότητα (4.13), η οποία είναι μια ισχυρή μορφή της ιδιότητας πυρήνα, μπορεί να μελετηθεί μέσω παραμέτρων όπως οι διάμετροι Gelfand. Αυτό μας οδηγεί στην επόμενη πρόταση.

Πρόταση 4.6.3. *Έστω A ένας $n \times N$ πίνακας και $1 \leqslant m \leqslant n$. Έστω $x \in \mathbb{R}^N$ και x^* ένας*

ελαχιστοποιητής του

$$\text{(P)} \qquad \min_{t\in\mathbb{R}^N} \|t\|_1 \quad \textit{με τον περιορισμό} \quad At = Ax.$$

Υποθέτουμε ότι

$$\rho = \operatorname{rad}(B_1^N \cap \ker(A)) = \sup_{x\in B_1^N\cap\ker(A)} \|x\|_2 \leqslant \frac{1}{4\sqrt{m}}.$$

Τότε, για κάθε $I \subset [N]$ με $|I| \leqslant m$,

$$\|x - x^*\|_1 \leqslant 4\|x - x_I\|_1$$

και

$$\|x - x^*\|_2 \leqslant \frac{1}{\sqrt{m}}\|x - x_I\|_1$$

Απόδειξη. Έστω $w \in \ker(A)$. Έχουμε

$$\|w_I\|_1 \leqslant \sqrt{m}\|w_I\|_2 \leqslant \sqrt{m}\|w\|_2 \leqslant \sqrt{m}\rho\|w\|_1.$$

Συνεπώς, αν $\rho\sqrt{m} < 1$ έχουμε ότι

$$\|w_I\|_1 \leqslant \frac{\rho\sqrt{m}}{1-\rho\sqrt{m}}\|w_{I^c}\|_1.$$

Συμπεραίνουμε έτσι ότι η ιδιότητα (4.13) ικανοποιείται με $\lambda = \frac{\rho\sqrt{m}}{1-\rho\sqrt{m}}$. Η ανισότητα

$$\|x - x^*\|_1 \leqslant 4\|x - x_I\|_1$$

προκύπτει άμεσα από την Πρόταση 4.6.2 και την υπόθεση ότι $\rho \leqslant \frac{1}{4\sqrt{m}}$. Χρησιμοποιώντας την

$$\|w\|_2 \leqslant \rho\|w\|_1 \leqslant 4\rho\|x - x_I\|_1$$

ολοκληρώνουμε την απόδειξη της τελευταίας ανισότητας. □

Έστω $1 \leqslant m \leqslant p \leqslant n$ και $N \geqslant m+p$. Μπορούμε επίσης να αναδιατυπώσουμε την τελευταία πρόταση συναρτήσει της σταθεράς της ιδιότητας περιορισμένης ισομετρίας ή συναρτήσει της παραμέτρου γ_p, δεδομένου ότι, από την (4.6),

$$\rho \leqslant \sqrt{\frac{1+\gamma_{2p}^2(A)}{p}}.$$

Παρατήρηση 4.6.4. Συνοψίζοντας, το Θεώρημα 4.3.2 δείχνει ότι αν ένας $n \times N$ πίνακας A ικανοποιεί την ιδιότητα περιορισμένης ισομετρίας τάξης $m \geqslant 1$, τότε

$$\operatorname{rad}(\ker(A) \cap B_1^N) = \frac{O(1)}{\sqrt{m}}. \tag{4.14}$$

Από την άλλη πλευρά, οι Προτάσεις 4.3.4 και 4.6.3 δείχνουν ότι αν ένας $n \times N$ πίνακας A ικανοποιεί την (4.14) τότε ο A ικανοποιεί την ιδιότητα ακριβούς ανακατασκευής τάξης $O(m)$ με τη μέθοδο της ℓ_1-ελαχιστοποίησης καθώς και την ιδιότητα προσεγγιστικής ανακατασκευής.

Ξεκινώντας από αυτήν την παρατήρηση, θα μπορούσαμε να εστιάσουμε σε εκτιμήσεις των διαμέτρων, όμως το παράδειγμα των Gaussian τυχαίων πινάκων δείχνει ότι ίσως είναι ευκολότερο να αποδείξουμε μια ιδιότητα περιορισμένης ισομετρίας από το να υπολογίσουμε πλάτη. Ολοκληρώνουμε αυτήν την ενότητα με μια εφαρμογή της Πρότασης 4.6.3.

Πόρισμα 4.6.5. *Έστω* $0 < p < 1$ *και*

$$T = B_{p,\infty}^N = \{x = (x_1, \dots, x_N) \in \mathbb{R}^N : |\{i : |x_i| \geqslant s\}| \leqslant s^{-p} \ \text{ για κάθε } \ s > 0\}$$

η μοναδιαία μπάλα του $\ell_{p,\infty}^N$. *Έστω* A *ένας* $n \times N$ *πίνακας και* $1 \leqslant m \leqslant n$. *Έστω* $x \in T$ *και* x^* *ένας ελαχιστοποιητής του*

$$\text{(P)} \qquad \min_{t \in \mathbb{R}^N} \|t\|_1 \quad \text{με τον περιορισμό} \quad At = Ax.$$

Αν

$$\rho = \mathrm{rad}(B_1^N \cap \ker(A)) = \sup_{x \in B_1^N \cap \ker(A)} \|x\|_2 \leqslant \frac{1}{4\sqrt{m}},$$

τότε

$$\|x - x^*\|_2 \leqslant ((1/p) - 1)^{-1} m^{1/2 - 1/p}.$$

Απόδειξη. Παρατηρούμε ότι για κάθε $x \in B_{p,\infty}^N$ έχουμε $x_i^* \leqslant 1/i^{1/p}$ για κάθε $i \geqslant 1$, όπου $(x_i^*)_{i=1}^N$ είναι μια φθίνουσα αναδιάταξη της $(|x_i|)_{i=1}^N$. Έστω $I \subset [N]$ με $|I| = m$ και έστω x_I η βέλτιστη m-αραιή προσέγγιση του x. Σημειώνουμε ότι

$$\sum_{i>m} i^{-1/p} \leqslant (1/p - 1)^{-1} m^{1-1/p}.$$

Από την Πρόταση 4.6.3 βλέπουμε ότι αν $\rho \leqslant \frac{1}{4\sqrt{m}}$ και αν x^* είναι ένας ελαχιστοποιητής του (P), τότε

$$\|x - x^*\|_2 \leqslant \frac{1}{\sqrt{m}} \|x - x_I\|_1 \leqslant ((1/p) - 1)^{-1} m^{1/2 - 1/p},$$

όπως θέλαμε. □

Αναγωγή του υπολογισμού των πλατών Gelfand με περικοπή. Ξεκινάμε με μια απλή αρχή που μας επιτρέπει να υπολογίζουμε πλάτη Gelfand υπολογίζοντας τα αντίστοιχα πλάτη ενός περικεκομμένου συνόλου.

Ορισμός 4.6.6. Λέμε ότι ένα υποσύνολο $T \subset \mathbb{R}^N$ είναι αστρόμορφο ως προς το 0, ή απλώς αστρόμορφο, αν $\lambda T \subset T$ για κάθε $0 \leqslant \lambda \leqslant 1$. Αν $\rho > 0$ και $T \subset \mathbb{R}^N$ αστρόμορφο σύνολο, συμβολίζουμε με T_ρ το σύνολο

$$T_\rho = T \cap \rho S^{N-1}.$$

Υπενθυμίζουμε ότι $\mathrm{rad}(S) = \sup_{x\in S} \|x\|_2$.

Λήμμα 4.6.7. *Έστω $\rho > 0$ και $T \subset \mathbb{R}^N$ αστρόμορφο. Τότε, για κάθε διανυσματικό υπόχωρο E του $\mathbb{R}^N$ με $E \cap T_\rho = \varnothing$ έχουμε ότι $\mathrm{rad}(E \cap T) < \rho$.*

Απόδειξη. Αν $\mathrm{rad}(E \cap T) \geqslant \rho$, τότε υπάρχει $x \in E \cap T$ με νόρμα μεγαλύτερη ή ίση από ρ. Εφόσον το T είναι αστρόμορφο, το ίδιο ισχύει για το $E \cap T$, επομένως $x/\|x\|_2 \in E \cap T_\rho$, το οποίο είναι άτοπο. □

Αυτό το απλό λήμμα μας δίνει ένα χρήσιμο εργαλείο. Ο υπόχωρος E θα είναι ο πυρήνας του πίνακα A, η σταθερά ρ θα είναι μια παράμετρος που προσπαθούμε να επιλέξουμε όσο γίνεται μικρότερη έτσι ώστε $\ker(A) \cap T_\rho = \varnothing$, δηλαδή έτσι ώστε $Ax \neq 0$ για κάθε $x \in T$ με $\|x\|_2 = \rho$. Ειδικότερα, αυτό θα συμβαίνει αν ο A δρα στο T_ρ σχεδόν διατηρώντας τη νόρμα.

Έχοντας κατά νου το Θεώρημα 4.6.1, εφαρμόζουμε αυτή τη στρατηγική σε υποσύνολα T όπως το Σ_m.

Πόρισμα 4.6.8. *Έστω $\mathbb{P}$ ένα μέτρο πιθανότητας στον $M(n, N)$ το οποίο ικανοποιεί την (4.8). Έστω $T \subset \mathbb{R}^N$ αστρόμορφο σύνολο και $\rho > 0$. Υποθέτουμε ότι το $\frac{1}{\rho}T_\rho \subset S^{N-1}$ ικανοποιεί την υπόθεση του Θεωρήματος 4.6.1 για κάποιο $\epsilon \in \left(0, \frac{1}{15}\right)$. Τότε,*

$$\mathrm{rad}(\ker(A) \cap T) < \rho$$

με πιθανότητα μεγαλύτερη ή ίση από $1 - 2\exp(-cn)$, όπου $c > 0$ είναι μια απόλυτη σταθερά.

Εφαρμογή σε υποσύνολα που σχετίζονται με ℓ_p-μπάλες. Για να περιγράψουμε τη μέθοδο, θεωρούμε κάποια παραδείγματα συνόλων T, για $0 < p < 2$:

1. Τη μοναδιαία μπάλα B_1^N του ℓ_1^N.
2. Τη μοναδιαία μπάλα $B_p^N = \{x \in \mathbb{R}^N : \sum_{i=1}^N |x_i|^p \leqslant 1\}$ του ℓ_p^N για $0 < p < 1$.
3. Τη μοναδιαία μπάλα $B_{p,\infty}^N = \{x \in \mathbb{R}^N : |\{i : |x_i| \geqslant s\}| \leqslant s^{-p} \text{ για κάθε } s > 0\}$ του $\ell_{p,\infty}^N$, του ασθενούς ℓ_p^N, για $0 < p < 1$.

Σημειώνουμε ότι αν $0 < p < 1$ τότε η «μοναδιαία μπάλα» B_p^N δεν είναι κυρτό σύνολο, και ότι $B_p^N \subset B_{p,\infty}^N$. Άρα, για να εκτιμήσουμε τα πλάτη Gelfand αυτών των συνόλων, μπορούμε να περιοριστούμε στις μπάλες $B_{p,\infty}^N$.

Χρειαζόμαστε δύο λήμματα. Το πρώτο χρησιμοποιεί τον ακόλουθο κλασσικό ισχυρισμό.

Ισχυρισμός 4.6.9. Έστω (a_i), (b_i) δύο ακολουθίες θετικών αριθμών, και έστω ότι η (a_i) είναι φθίνουσα. Τότε, το άθροισμα $\sum a_i b_{\pi(i)}$ ελαχιστοποιείται πάνω από όλες τις μεταθέσεις π του συνόλου δεικτών αν $b_{\pi(1)} \geqslant b_{\pi(2)} \geqslant \cdots$.

Λήμμα 4.6.10. *Έστω $0 < p < 1$, $1 \leqslant m \leqslant N$ και $r = (1/p - 1)m^{1/p-1/2}$. Τότε, για κάθε $x \in \mathbb{R}^N$,*

$$\sup_{z \in rB_{p,\infty}^N \cap B_2^N} \langle x, z\rangle \leqslant 2\left(\sum_{i=1}^m |x_i^*|^2\right)^{1/2},$$

όπου $(x_i^*)_{i=1}^N$ *είναι μια φθίνουσα αναδιάταξη της* $(|x_i|)_{i=1}^N$. *Ισοδύναμα,*

$$rB_{p,\infty}^N \cap B_2^N \subset 2\mathrm{conv}(S_2(\Sigma_m)). \tag{4.15}$$

Επιπλέον, ισχύει ότι

$$\sqrt{m}B_1^N \cap B_2^N \subset 2\mathrm{conv}(S_2(\Sigma_m)). \tag{4.16}$$

Απόδειξη. Θα ασχοληθούμε μόνο με την περίπτωση της $B_{p,\infty}^N$, $0 < p < 1$. Η περίπτωση της B_1^N είναι όμοια. Σημειώνουμε αρχικά ότι αν $z \in B_{p,\infty}^N$ τότε για κάθε $i \geqslant 1$ έχουμε $z_i^* \leqslant 1/i^{1/p}$, όπου $(z_i^*)_{i=1}^N$ είναι μια φθίνουσα αναδιάταξη της $(|z_i|)_{i=1}^N$. Χρησιμοποιώντας τον Ισχυρισμό 4.6.9 βλέπουμε ότι για κάθε $r > 0$, $m \geqslant 1$ και $z \in rB_{p,\infty}^N \cap B_2^N$,

$$\begin{aligned}
\langle x, z\rangle &\leqslant \left(\sum_{i=1}^m |x_i^*|^2\right)^{1/2} + \sum_{i>m} \frac{rx_i^*}{i^{1/p}} \\
&\leqslant \left(\sum_{i=1}^m |x_i^*|^2\right)^{1/2} \left(1 + \frac{r}{\sqrt{m}} \sum_{i>m} \frac{1}{i^{1/p}}\right) \\
&\leqslant \left(\sum_{i=1}^m |x_i^*|^2\right)^{1/2} \left(1 + \left(\tfrac{1}{p} - 1\right)^{-1} \frac{r}{m^{1/p-1/2}}\right).
\end{aligned}$$

Από τον ορισμό του r, αυτό ολοκληρώνει την απόδειξη. □

Το δεύτερο λήμμα δείχνει ότι το $m^{1/p-1/2}B_{p,\infty}^N \cap S^{N-1}$ προσεγγίζεται καλά από διανύσματα στη σφαίρα που έχουν μικρό φορέα.

Λήμμα 4.6.11. *Έστω* $0 < p < 2$, $\delta > 0$ *και* $\epsilon = 2(2/p-1)^{-1/2}\delta^{1/p-1/2}$. *Έστω* $1 \leqslant m \leqslant N$. *Τότε, το* $S_2(\Sigma_{\lceil m/\delta\rceil})$ *είναι* ϵ*-δίκτυο του* $m^{1/p-1/2}B_{p,\infty}^N \cap S^{N-1}$ *ως προς την Ευκλείδεια μετρική.*

Απόδειξη. Έστω $x \in m^{1/p-1/2}B_{p,\infty}^N \cap S^{N-1}$. Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_N \geqslant 0$. Ορίζουμε z' με $z_i' = x_i$ αν $1 \leqslant i \leqslant \lceil m/\delta\rceil$ και $z_i' = 0$ διαφορετικά. Τότε,

$$\|x - z'\|_2^2 = \sum_{i>m/\delta} |x_i|^2 \leqslant m^{2/p-1} \sum_{i>m/\delta} \frac{1}{i^{2/p}} < (2/p-1)^{-1}\delta^{2/p-1}.$$

Συνεπώς,

$$1 \geqslant \|z'\|_2 \geqslant 1 - (2/p-1)^{-1/2}\delta^{1/p-1/2}.$$

Θέτουμε $z = z'/\|z'\|_2$. Τότε, $z \in S_2(\Sigma_{\lceil m/\delta\rceil})$ και

$$\|z - z'\|_2 = 1 - \|z'\|_2 \leqslant (2/p-1)^{-1/2}\delta^{1/p-1/2}.$$

Από την τριγωνική ανισότητα παίρνουμε $\|x - z\|_2 < \epsilon$, και αυτό ολοκληρώνει την απόδειξη. □

Τα προηγούμενα λήμματα θα χρησιμοποιηθούν για να δείξουμε ότι οι υποθέσεις του Θεωρήματος 4.6.1 ικανοποιούνται για κατάλληλη επιλογή των T και ρ. Πριν από αυτό, δίνουμε

έναν ορισμό στον οποίο οδηγούμαστε από την ιδιότητα (ii) του Θεωρήματος 4.6.1.

Ορισμός 4.6.12. Λέμε ότι ένα υποσύνολο T του $\mathbb{R}^N$ είναι quasi-κυρτό με σταθερά $a \geqslant 1$ αν το T είναι αστρόμορφο και $T + T \subset 2aT$.

Σημειώνουμε την εξής απλή παρατήρηση.

Ισχυρισμός 4.6.13. Αν $0 < p < 1$ τότε τα $B_{p,\infty}^N$ και B_p^N είναι quasi-κυρτά σύνολα με σταθερά $2^{(1/p)-1}$.

Ο βασικός μας ισχυρισμός είναι ο ακόλουθος:

Ισχυρισμός 4.6.14. Έστω $0 < p < 1$ και $T = B_{p,\infty}^N$. Τότε, το $\frac{1}{\rho}T_\rho$ ικανοποιεί τις ιδιότητες (i) και (ii) του Θεωρήματος 4.6.1 με

$$\rho = C_p \left(\frac{n}{\log(cN/n)} \right)^{1/p-1/2},$$

όπου η σταθερά C_p εξαρτάται μόνο από το p και $c > 0$ είναι μια απόλυτη σταθερά. Επιπλέον, αν $T = B_1^N$, τότε το $\frac{1}{\rho}T_\rho$ ικανοποιεί τις ιδιότητες (i) και (ii) του Θεωρήματος 4.6.1 με

$$\rho = C_p \left(\frac{c_1 n}{\log(c_2 N/n)} \right)^{1/2},$$

όπου $c_1, c_2 > 0$ είναι απόλυτες σταθερές.

Απόδειξη. Εξετάζουμε μόνο την περίπτωση του $T = B_{p,\infty}^N$, $0 < p < 1$. Η περίπτωση της B_1^N είναι όμοια. Δεδομένου ότι έχουμε ήδη εξηγήσει λεπτομερώς τον μηχανισμό, θα παρουσιάσουμε μόνο τα βασικά βήματα. Θέτουμε $\epsilon_0 = 1/20$. Για να δείξουμε το (i) χρησιμοποιούμε το Λήμμα 4.6.11 με $\epsilon = \epsilon_0/2$ και δ το οποίο ορίζεται από την εξίσωση $\epsilon_0/2 = 2(2/p-1)^{-1/2}\delta^{1/p-1/2}$. Έστω $1 \leqslant m \leqslant N$. Έχουμε ότι το $S_2(\Sigma_{\lceil m/\delta \rceil})$ είναι $(\epsilon_0/2)$-δίκτυο του $m^{1/p-1/2}B_{p,\infty}^N$ ως προς την Ευκλείδεια μετρική. Θέτουμε $m' = \lceil m/\delta \rceil$. Από τον Ισχυρισμό 4.1.8, έχουμε

$$N\left(S_2(\Sigma_{m'}), \tfrac{\epsilon_0}{2}B_2^N\right) \leqslant \left(\frac{3eN}{m'(\epsilon_0/2)}\right)^{m'} = \left(\frac{6eN}{m'\epsilon_0}\right)^{m'}.$$

Έτσι, από την τριγωνική ανισότητα, έχουμε ότι

$$N(m^{1/p-1/2}B_{p,\infty}^N \cap S^{N-1}, \epsilon_0 B_2^N) \leqslant \left(\frac{6eN}{m'\epsilon_0}\right)^{m'},$$

απ' όπου έπεται ότι

$$N(m^{1/p-1/2}B_{p,\infty}^N \cap S^{N-1}, \epsilon_0 B_2^N) \leqslant \exp(c_0 n/2)$$

αν ικανοποιείται η

$$\left(\frac{6eN}{m'\epsilon_0}\right)^{m'} \leqslant \exp(c_0 n/2).$$

Κάτω από αυτόν τον περιορισμό για το m', δηλαδή για το m, το σύνολο $m^{1/p-1/2}B_{p,\infty}^N \cap S^{N-1}$ ικανοποιεί το (i).

Περνάμε στην απόδειξη του (ii). Υπενθυμίζουμε ότι το σύνολο $B_{p,\infty}^N$ είναι quasi-κυρτό με σταθερά $2^{1/p-1}$ (Ισχυρισμός 4.6.13). Λόγω συμμετρίας, έχουμε

$$B_{p,\infty}^N - B_{p,\infty}^N \subset 2^{1/p}B_{p,\infty}^N.$$

Θέτουμε $r = (1/p-1)m^{1/p-1/2}$. Από το Λήμμα 4.6.10 έχουμε

$$rB_{p,\infty}^N \cap B_2^N \subset 2\mathrm{conv}(S_2(\Sigma_m)).$$

Όπως είδαμε προηγουμένως,

$$N\left(S_2(\Sigma_m), \tfrac{1}{2}B_2^N\right) \leqslant \left(\frac{3eN}{m/2}\right)^m = \left(\frac{6eN}{m}\right)^m,$$

και από τον Ισχυρισμό 4.1.8 υπάρχει υποσύνολο $\Lambda' \subset S^{N-1}$ με $|\Lambda'| \leqslant N\left(S_2(\Sigma_m), \frac{1}{2}B_2^N\right)$ τέτοιο ώστε $S_2(\Sigma_m) \subset 2\mathrm{conv}(\Lambda')$. Έτσι, έχουμε

$$\epsilon_0 2^{-1/p}(rB_{p,\infty}^N - rB_{p,\infty}^N) \subset \epsilon_0(rB_{p,\infty}^N \cap B_2^N) \subset 4\epsilon_0\mathrm{conv}(\Lambda')) \subset 2\mathrm{conv}(\Lambda' \cup (-\Lambda')).$$

Αυτό δείχνει ότι το $\epsilon_0 2^{-1/p}rB_{p,\infty}^N \cap S^{N-1}$ ικανοποιεί το (ii) εάν

$$2(6eN/m)^m \leqslant \exp(c_0 n/2).$$

Τέλος, το $\epsilon_0 2^{-1/p}rB_{p,\infty}^N \cap S^{N=1}$ ικανοποιεί τα (i) και (ii) αν οι δύο συνθήκες για το m ικανοποιούνται ταυτόχρονα, δηλαδή εάν

$$cm\log(CN/m) \leqslant c_0 n/2,$$

όπου $c, C > 0$ είναι απόλυτες σταθερές. Υπολογίζουμε τα m και r, και θέτοντας $\rho = \epsilon_0 2^{-1/p}r$ ολοκληρώνουμε την απόδειξη. □

Εφαρμόζοντας το Πόρισμα 4.6.8 έχουμε το ακόλουθο θεώρημα.

Θεώρημα 4.6.15. *Έστω $\mathbb{P}$ ένα μέτρο πιθανότητας που ικανοποιεί την (4.8) στον χώρο των $n \times N$ πινάκων και έστω $0 < p < 1$. Υπάρχουν σταθερές c_p που εξαρτάται μόνο από το p, c' που εξαρτάται μόνο από την c_0 και μια απόλυτη σταθερά c έτσι ώστε το σύνολο Ω των $n \times N$ πινάκων A που ικανοποιούν την*

$$\mathrm{rad}(\ker(A) \cap B_p^N) \leqslant \mathrm{rad}(\ker(A) \cap B_{p,\infty}^N) \leqslant c_p\left(\frac{\log(cN/n)}{n}\right)^{1/p-1/2}$$

να έχει πιθανότητα μεγαλύτερη ή ίση από $1 - \exp(-c'n)$.

Ειδικότερα, αν $A \in \Omega$ και αν τα $x', x \in B_{p,\infty}^N$ ικανοποιούν την $Ax' = Ax$, τότε

$$\|x' - x\|_2 \leqslant c'_p \left(\frac{\log(c_1 N/n)}{n} \right)^{1/p-1/2}.$$

Αντίστοιχο αποτέλεσμα ισχύει για την B_1^N.

Σχετική Βιβλιογραφία. Σε αυτήν την ενότητα ακολουθούμε το [28]. Η Πρόταση 4.6.3 διατυπώνεται και αποδεικνύεται στο [24] για τα πλάτη Gelfand αντί για τις σταθερές ισομετρίας που εμφανίζονται στα [5] και [10]. Η μέθοδος που μας βοηθάει να υπολογίσουμε τα πλάτη Gelfand υπολογίζοντας τα αντίστοιχα πλάτη ενός περικεκομμένου συνόλου προέρχεται από το [19].

4.7 Ένας άλλος δείκτης πολυπλοκότητας

Σε αυτήν την ενότητα, εισάγουμε μια νέα παράμετρο $\ell_*(T)$, η οποία είναι ένας δείκτης της πολυπλοκότητας ενός συνόλου $T \subset \mathbb{R}^N$. Ορίζουμε

$$\ell_*(T) = \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^N g_i t_i \right|, \tag{4.17}$$

όπου $t = (t_1, \dots, t_N) \in \mathbb{R}^N$ και $g_1, \dots, g_N$ είναι ανεξάρτητες $N(0,1)$ Gaussian τυχαίες μεταβλητές. Αυτός ο δείκτης παίζει σημαντικό ρόλο στις εμπειρικές διαδικασίες και στη γεωμετρία των χώρων Banach.

Θεώρημα 4.7.1. *Υπάρχουν απόλυτες σταθερές $c, c' > 0$ τέτοιες ώστε να ισχύει το ακόλουθο: Έστω $1 \leqslant n \leqslant N$ και A ένας Gaussian τυχαίος πίνακας με ανεξάρτητες τυπικές κανονικές συντεταγμένες. Έστω $T \subset S^{n-1}$ ένα αστρόμορφο σύνολο. Τότε, με πιθανότητα μεγαλύτερη ή ίση από $1 - \exp(-c'n)$,*

$$\operatorname{rad}(\ker(A) \cap T) \leqslant c\ell_*(T)/\sqrt{n}.$$

Απόδειξη. Η στρατηγική μας είναι να δείξουμε μια ιδιότητα περιορισμένης ισομετρίας και στη συνέχεια να επιχειρηματολογήσουμε όπως στο Λήμμα 4.6.7. Έστω $\delta \in (0,1)$. Η απόδειξη της ιδιότητας περιορισμένης ισομετρίας γίνεται με συνδυασμό ενός επιχειρήματος διακριτοποίησης με δίκτυα και ενός επιχειρήματος προσέγγισης. Για κάθε $\theta > 0$, έστω $\Lambda(\theta) \subset T$ ένα θ-δίκτυο του T ως προς την Ευκλείδεια μετρική. Έστω $\pi_\theta : T \to \Lambda(\theta)$ μια απεικόνιση τέτοια ώστε

$$\|t - \pi_\theta(t)\|_2 \leqslant \theta \quad \text{για κάθε } t \in T.$$

Από την τριγωνική ανισότητα έχουμε

$$\sup_{t \in T} \left| \frac{\|At\|_2^2}{n} - \mathbb{E}|\langle Y, t\rangle|^2 \right| \leqslant \sup_{s \in \Lambda(\theta)} \left| \frac{\|As\|_2^2}{n} - \mathbb{E}|\langle Y, s\rangle|^2 \right| + \sup_{t \in T} \left| \frac{\|At\|_2^2}{n} - \frac{\|A\pi_\theta(t)\|_2^2}{n} \right|.$$

Το πρώτο μας βήμα είναι να αποδείξουμε μια ανισότητα εντροπίας μέσω της ανισότητας Sudakov. Έστω $s \in T$. Έστω (Y_i) οι γραμμές του A και Y ένα Gaussian τυχαίο διάνυσμα που έχει πίνακα συνδιακυμάνσεων τον ταυτοτικό πίνακα. Τότε, η $\langle Y, s\rangle$ είναι ψ_1 με κάποια απόλυτη σταθερά. Από την ανισότητα Bernstein του Θεωρήματος 2.7.3 συμπεραίνουμε ότι

$$\left| \frac{\|As\|_2^2}{n} - \mathbb{E}|\langle Y, s\rangle|^2 \right| = \left| \frac{1}{n} \sum_{i=1}^{n} (\langle Y_i, s\rangle^2 - \mathbb{E}|\langle Y, s\rangle|^2) \right| \leqslant \frac{\delta}{2}$$

με πιθανότητα μεγαλύτερη από $1 - 2\exp(-c\delta^2 n)$, όπου $c > 0$ είναι μια απόλυτη σταθερά. Έπεται ότι, για κάθε $\theta > 0$,

$$\sup_{s \in \Lambda(\theta)} \left| \frac{\|As\|_2^2}{n} - \mathbb{E}|\langle Y, s\rangle|^2 \right| \leqslant \frac{\delta}{2}$$

με πιθανότητα μεγαλύτερη από $1 - 2\exp(-c\delta^2 n + \log|\Lambda(\theta)|)$.

Από την ανισότητα Sudakov (βλέπε Θεώρημα 2.8.3) υπάρχει $c' > 0$ ώστε, αν

$$\theta = c' \frac{\ell_*(T)}{\delta\sqrt{n}}$$

να έχουμε $\log|\Lambda(\theta)| \leqslant c\delta^2 n/2$. Συμπεραίνουμε τότε ότι η

$$\sup_{s \in \Lambda(\theta)} \left| \frac{\|As\|_2^2}{n} - 1 \right| \leqslant \frac{\delta}{2}$$

ισχύει με πιθανότητα μεγαλύτερη από $1 - 2\exp(-c\delta^2 n/2)$.

Στο δεύτερο βήμα της απόδειξης φράσσουμε τον όρο

$$\sup_{t \in T} \left| \|At\|_2^2 - \|A\pi_\theta(t)\|_2^2 \right|.$$

Παρατηρήστε αρχικά ότι για κάθε $s, t \in T$ έχουμε

$$|\,\|As\|_2^2 - \|At\|_2^2\,| \leqslant \|A(s-t)\|_2 \|A(s+t)\| - 2.$$

Συνεπώς,

$$\sup_{t \in T} \left| \|At\|_2^2 - \|A\pi_\theta(t)\|_2^2 \right| \leqslant 2 \sup_{t \in T(\theta)} \|At\|_2 \sup_{t \in T} \|At\|_2,$$

όπου $T(\theta) = \{s-t : s, t \in T, \|t-s\|_2 \leqslant \theta\}$. Για να εκτιμήσουμε αυτές τις δύο νόρμες του πίνακα A, θεωρούμε ένα $(1/2)$-δίκτυο της Ευκλείδειας σφαίρας του $\mathbb{R}^n$. Από την Πρόταση 4.1.6 υπάρχει ένα τέτοιο δίκτυο $\mathcal{N}$ με πληθάριθμο το πολύ 5^n και τέτοιο ώστε $B_2^n \subset 2\mathrm{conv}(\mathcal{N})$. Έπεται ότι

$$\sup_{t \in T} \|At\|_2 = \sup_{t \in T} \sup_{\|u\|_2 \leqslant 1} |\langle At, u\rangle| \leqslant 2 \sup_{u \in \mathcal{N}} \sup_{t \in T} |\langle t, A^T u\rangle|.$$

Εφόσον ο A είναι ένας τυπικός Gaussian τυχαίος πίνακας με ανεξάρτητες $N(0,1)$ συντεταγ-

μένες, για κάθε $u \in \mathcal{N}$ το $A^T u$ είναι ένα τυπικό Gaussian τυχαίο διάνυσμα και

$$\mathbb{E} \sup_{t \in T} \langle t, A^T u \rangle = \ell_*(T).$$

Έπεται από το Θεώρημα 2.8.5 ότι για κάθε σταθερό $u \in \mathcal{N}$ και για κάθε $z > 0$

$$\mathbb{P}\left(\left| \sup_{t \in T} \langle t, A^Y u \rangle - \mathbb{E} \sup_{t \in T} \langle t, A^T u \rangle \right| > z \right) \leqslant 2 \exp(-c'' z^2 / \sigma^2(T))$$

για κάποια απόλυτη σταθερά $c'' > 0$, όπου

$$\sigma(T) = \sup_{t \in T} \left(\mathbb{E} \langle t, A^T u \rangle^2 \right)^{1/2}.$$

Συνδυάζοντας αυτήν την εκτίμηση με το άνω φράγμα για τον πληθάριθμο του δικτύου, βλέπουμε ότι για κάθε $z > 0$

$$\mathbb{P}\left(\sup_{u \in \mathcal{N}} \sup_{t \in T} \langle t, A^Y u \rangle \geqslant \ell_*(T) + z\sigma(T)\sqrt{n} \right) \leqslant 2 \exp(-c'' n(z^2 - \log 5)).$$

Συμπεραίνουμε έτσι ότι

$$\sup_{t \in T} \|At\|_2 \leqslant 2(\ell_*(T) + z\sigma(T)\sqrt{n})$$

με πιθανότητα μεγαλύτερη από $1 - 2\exp(-c'' n(z^2 - \log 5))$.

Ο ίδιος συλλογισμός εφαρμόζεται και για το $T(\theta)$. Στην περίπτωσή μας, $\sigma(T) = 1$ και $\sigma(T(\theta)) \leqslant \theta$. Άρα,

$$\sup_{t \in T} | \|At\|_2^2 - \|A\pi_\theta(t)\|_2^2 | \leqslant 8(\ell_*(T) + z\sqrt{n})(\ell_*(T) + z\theta\sqrt{n})$$

με πιθανότητα μεγαλύτερη από $1 - 4\exp(-c'' n(z^2 - \log 5))$.

Ερχόμαστε τώρα στην ιδιότητα περιορισμένης ισομετρίας. Επιλέγουμε $z = \sqrt{2\log 5}$. Εισάγοντας την τιμή του θ, βλέπουμε ότι με μεγάλη πιθανότητα ισχύει ότι

$$\sup_{t \in T} | \|At\|_2^2 - \|A\pi_\theta(t)\|_2^2 | \leqslant 8(\ell_*(T) + z\sqrt{n}) \left(\ell_*(T) + zc' \frac{\ell_*(T)}{\delta} \right)$$

και

$$\sup_{t \in T} \left| \frac{\|At\|_2^2}{n} - 1 \right| \leqslant \frac{\delta}{2} + c''' \left(\frac{\ell_*(T)^2}{n} + z^2 \frac{\ell_*(T)}{\delta\sqrt{n}} \right)$$

για κάποια νέα σταθερά c'''. Είναι φανερό ότι μπορούμε να επιλέξουμε αυτή τη σταθερά έτσι ώστε, αν

$$\ell_*(T) \leqslant c''' \delta^2 \sqrt{n}$$

τότε να έχουμε

$$\sup_{t \in T} \left| \frac{\|At\|_2^2}{n} - 1 \right| \leqslant \delta$$

με πιθανότητα μεγαλύτερη από $1-2\exp(-c\delta^2 n/2)-4\exp(-c''z^2 n/2)$.

Τέλος, δίνουμε μια εκτίμηση για το πλάτος. Εφαρμόζοντας την τελευταία ανισότητα με $\delta = 1/2$ για το υποσύνολο $\frac{1}{\rho}T\cap S^{N-1}$ παίρνουμε ότι, με μεγάλη πιθανότητα, $\ker(A)\cap\left(\frac{1}{\rho}T\cap S^{N-1}\right)=\varnothing$ εάν ικανοποιείται η

$$\ell_*\left(\tfrac{1}{\rho}T\cap S^{N-1}\right)\leqslant \frac{\ell_*(T)}{\rho} < c'''\delta^2\sqrt{n}.$$

Το συμπέρασμα έπεται τώρα από το Λήμμα 4.6.7. □

Παρατήρηση 4.7.2. Η απόδειξη του Θεωρήματος 4.7.1 γενικεύεται στην περίπτωση ενός πίνακα με ανεξάρτητες υποκανονικές γραμμές. Πρέπει να τροποποιήσουμε μόνο το δεύτερο βήμα, χρησιμοποιώντας το θεώρημα του κυριαρχούντος μέτρου το οποίο μας επιτρέπει ακριβώς να συγκρίνουμε ανισότητες απόκλισης για το supremum υποκανονικών ανελίξεων με τις αντίστοιχες στο Gaussian πλαίσιο.

Δείχνουμε τώρα με ποιον τρόπο μπορούμε να εφαρμόσουμε το Θεώρημα 4.7.1 σε κάποια σύνολα T.

Πόρισμα 4.7.3. *Υπάρχουν απόλυτες σταθερές $c,c'>0$ τέτοιες ώστε να ισχύει το ακόλουθο: Έστω $1\leqslant n\leqslant N$ και πίνακας A όπως στο Θεώρημα 4.7.1. Έστω $\lambda>0$. Έστω $T\subset S^{n-1}$ και ας υποθέσουμε ότι $T\subset 2\mathrm{conv}(\Lambda)$ για κάποιο $\Lambda\subset B_2^N$ με $|\Lambda|\leqslant\exp(\lambda^2 n)$. Τότε, με πιθανότητα μεγαλύτερη ή ίση από $1-\exp(-c'n)$,*

$$\mathrm{rad}(\ker(A)\cap T)\leqslant c\lambda.$$

Παρατήρηση 4.7.4. Η σταθερά 2 στον εγκλεισμό $T\subset 2\mathrm{conv}(\Lambda)$ δεν παίζει ουσιαστικό ρόλο.

Απόδειξη. Το βασικό σημείο στην απόδειξη είναι ότι αν $T\subset 2\mathrm{conv}(\Lambda)$, $\Lambda\subset B_2^N$ και έχουμε κάποιον έλεγχο για τον πληθάριθμο $|\Lambda|$ του Λ, τότε μπορούμε να φράξουμε την $\ell_*(T)$ από πάνω. Τα υπόλοιπα είναι άμεση εφαρμογή του Θεωρήματος 4.7.1. Έστω $c,c'>0$ οι σταθερές στο Θεώρημα 4.7.1. Είναι γνωστό ότι υπάρχει απόλυτη σταθερά $c''>0$ τέτοια ώστε, για κάθε $\Lambda\subset B_2^N$,

$$\ell_*(\mathrm{conv}(\Lambda))=\ell_*(\Lambda)\leqslant c''\sqrt{\log(|\Lambda|)},$$

και αφού $T\subset 2\mathrm{conv}(\Lambda)$ συμπεραίνουμε ότι

$$\ell_*(T)\leqslant 2\ell_*(\mathrm{conv}(\Lambda))\leqslant 2c''\sqrt{\lambda^2 n}.$$

Το συμπέρασμα έπεται από το Θεώρημα 4.7.1. □

Σχετική Βιβλιογραφία. Η παράμετρος $\ell_*(T)$ που ορίζεται σε αυτήν την ενότητα παίζει σημαντικό ρόλο στη γεωμετρία χώρων Banach (βλ. [32]). Το Θεώρημα 4.7.1 προέρχεται από το [30].

Βιβλιογραφία

[1] R. Adamczak, A. E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling*, preprint arXiv:0904.4723.

[2] R. Adamczak, A. E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles*, J. Amer. Math. Soc. 23 (2010), no. 2, 535–561.

[3] S. Artstein-Avidan, A. Giannopoulos and V. D. Milman, *Asymptotic Geometric Analysis, Vol. I*, Mathematical Surveys and Monographs, 202. American Mathematical Society, Providence, RI, 2015. xx+451 pp.

[4] R. Baraniuk, M. Davenport, R. DeVore and M. Wakin, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx. 28 (2008), no. 3, 253–263.

[5] E. J. Candés, *The restricted isometry property and its implications for compressed sensing*, C. R. Math. Acad. Sci. Paris 346 (2008), no. 9-10, 589–592.

[6] E. J. Candés and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory 51 (2005), no. 12, 4203–4215.

[7] E. J. Candés and T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies?*, IEEE Trans. Inform. Theory 52 (2006), no. 12, 5406–5425.

[8] E. J. Candés, J. K. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. 59 (2006), no. 8, p. 1207–1223.

[9] D. Chafaï, O. Guédon, G. Lecué and A. Pajor, *Interactions between compressed sensing random matrices and high dimensional geometry*, Panoramas et Synthèses, 37. Société Mathématique de France, Paris, 2012.

[10] A. Cohen, W. Dahmen and R. DeVore, *Compressed sensing and best k-term approximation*, J. Amer. Math. Soc. 22 (2009), no. 1, 211–231.

[11] D. L. Donoho, *Neighborly polytopes and sparse solutions of underdetermined linear equations*, Department of Statistics, Stanford University, 2005.

[12] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), no. 4, 1289–1306.

[13] A. Eskenazis, *ϵ-Isometric dimension reduction for incompressible subsets of ℓ_p*, Discrete Comput. Geom. 71 (2024), no. 1, 160–176.

[14] T. Figiel, J. Lindenstrauss and V. D. Milman, *The dimension of almost spherical sections of convex bodies*, Acta Math. 139 (1977), no. 1-2, 53–94.

[15] S. Foucart and M.-J. Li, *Sparsest solutions of underdetermined linear systems via q-minimization for $0 < q \leqslant 1$*, Applied and Computational Harmonic Analysis 26 (2009), no. 3, 395–407.

[16] S. Foucart, A. Pajor, H. Rauhut and T. Ullrich, *The Gelfand widths of p-balls for $0 < p \leqslant 1$*, preprint, 2010.

[17] S. Foucart and O. Rauhut, *A mathematical introduction to compressive sensing*, Appl. Numer. Harmon. Anal., Birkhäuser, Boston, New York, 2011.

[18] A. Y. Garnaev and E. D. Gluskin, *The widths of a Euclidean ball*, Dokl. Akad. Nauk SSSR 277 (1984), no. 5, 1048–1052.

[19] E. D. Gluskin, *Norms of random matrices and diameters of finite-dimensional sets*, Mat. Sb. (N.S.) 120(162) (1983), no. 2, 180–189.

[20] K. Green Larsen and J. Nelson, *The Johnson-Lindenstrauss lemma is optimal for linear dimensionality reduction*, 43rd International Colloquium on Automata, Languages, and Programming, Art. No. 82, 11 pp., LIPIcs. Leibniz Int. Proc. Inform., 55, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.

[21] K. Green Larsen and J. Nelson, *Optimality of the Johnson-Lindenstrauss lemma*, 58th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2017, 633–638, IEEE Computer Soc., Los Alamitos, CA, 2017.

[22] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, Conference in modern analysis and probability (New Haven, Conn., 1982), 189–206, Contemp. Math., 26, Amer. Math. Soc., Providence, RI, 1984.

[23] B. S. Kashin, *The widths of certain finite-dimensional sets and classes of smooth functions*, Izv. Akad. Nauk SSSR Ser. Mat. 41 (1977), no. 2, 334–351.

[24] B. S. Kashin and V. N. Temlyakov, *A remark on the problem of compressed sensing*, Mat. Zametki 82 (2007), no. 6, 829–837.

[25] J. R. Lee and A. Naor, *Embedding the diamond graph in L_p and dimension reduction in L_1*, Geom. Funct. Anal. 14 (2004), no. 4, 745–747.

[26] N. Linial and I. Novik, *How neighborly can a centrally symmetric polytope be?*, Discrete Comput. Geom. 36 (2006), no. 2, 273–281.

[27] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, *Reconstruction and subgaussian operators in asymptotic geometric analysis*, Geom. Funct. Anal. 17 (2007), no. 4, 1248–1282.

[28] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, *Uniform uncertainty principle for Bernoulli and subgaussian ensembles*, Constr. Approx. 28 (2008), no. 3, 277–289.

[29] V. D. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces*, Lecture Notes in Mathematics, vol. 1200, Springer-Verlag, Berlin, 1986, With an appendix by M. Gromov.

[30] A. Pajor and N. Tomczak-Jaegermann, *Subspaces of small codimension of finite-dimensional Banach spaces*, Proc. Amer. Math. Soc. 97 (1986), no. 4, 637–642.

[31] G. Pisier, *Remarques sur un résultat non publié de B. Maurey*, Seminar on Functional Analysis, 1980–1981, Exp. No. V, 13 pp., Ecole Polytech., Palaiseau, 1981.

[32] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Tracts in Mathematics, vol. 94, Cambridge University Press, Cambridge, 1989.