Διακριτοί τυχαίοι πίνακες - Ορίζουσα και αντιστρεψιμότητα

Ιωάννης Κωνσταντούλας

2 Ιουλίου 2008

1 Εισαγωγή

Η παρούσα εργασία επικεντρώνεται στην παρουσίαση δύο προσφάτων αποτελεσμάτων σχετικά με ασυμπτωτικές στατιστικές ιδιότητες των τυχαίων $n \times n$ πινάκων με στοιχεία στο $\{+1, -1\}$. Ο τυχαίος πίνακας παράγεται από ± 1 στοιχεία ανεξάρτητα μεταξύ τους με ίση πιθανότητα για την κάθε τιμή.

Το πρώτο αποτέλεσμα λέει ότι η ορίζουσα ενός τυχαίου πίνακα $n \times n$ είναι με μεγάλη πιθανότητα πολύ κοντά στο $\sqrt{n!}$ και το δεύτερο είναι μια βελτίωση ενός αποτελέσματος των Kahn, Komlos και Szemeredi που λέει ότι η πιθανότητα να είναι ένας τυχαίος πίνακας μη αντιστρέψιμος τείνει εκθετικά γρήγορα στο μηδέν καθώς η διάσταση μεγαλώνει προς το άπειρο. Το φράγμα που δίνεται από το αποτέλεσμα που θα παρουσιάσουμε είναι κοντά στο $(\frac{3}{4} + o(1))^n$, ενώ η πραγματική τιμή αυτής της πιθανότητας εικάζεται ότι είναι $(\frac{1}{2} + o(1))^n$. Παρατηρούμε ότι στην περίπτωση όπου τα στοιχεία των τυχαίων πινάκων είναι συνεχείς τυχαίες μεταβλητές, για παράδειγμα Γκαουσιανές, το σύνολο $\{A_n \in \mathbb{R}^{n^2} | \det(A_n) = 0\}$ είναι μια υπερεπιφάνεια που προφανώς έχει μέτρο μηδέν στον $\mathbb{R}^{n^2}$, οπότε $P(\det(A_n) = 0) = 0$ για κάθε n. Στη διακριτή περίπτωση, όμως, κάθε ενδεχόμενο έχει θετική πιθανότητα, και επομένως δεν ισχύει κάτι όπως το παραπάνω. Το δεύτερο αποτέλεσμα μας λέει ότι ασυμπτωτικά, η κατάσταση στη διακριτή περίπτωση είναι ίδια με αυτή στη συνεχή περίπτωση.

Η ιστορία του προβλήματος της εκτίμησης της $P(\det(M_n) = 0)$ για Bernoulli πίνακες πηγαίνει μέχρι το πρώτο μισό του 20ού αιώνα. Τότε με έκπληξη είχε διαπιστωθεί ότι δεν ήταν τετριμμένο το να δειχθεί ότι $P(\det(M_n) = 0) \to 0$ καθώς $n \to \infty$. Ο Komlos πρώτος το έδειξε για Bernoulli πίνακες στο [2] και έπειτα το επέκτεινε για γενικότερους πίνακες στο [3]. Μια καινούρια απόδειξη δόθηκε το 1995 για πρώτη φορά με εκθετικό φράγμα στο [1] και από τότε άρχισε να διαφαίνεται μια διέξοδος προς την πλήρη απόδειξη της βασικής εικασίας (η οποία παρ'ολα αυτά παραμένει ανοιχτή). Τα αποτελέσματα που θα παρουσιαστούν εδώ ακολουθούν το [4].

2 Μια ανισότητα συγκέντρωσης για την ορίζουσα

2.1 Κύρια Αποτελέσματα

Το βασικό θεώρημα που θα δείξουμε είναι το εξής:

Θεώρημα 2.1. *Έστω M_n τυχαίος Bernoulli $n \times n$ πίνακας. Τότε*

$$P\left(\det(M_n) = \sqrt{n!}e^{O(\sqrt{n\ln n})}\right) = \frac{1}{2} - \mathrm{o}(1),$$

$$P\left(\det(M_n) = -\sqrt{n!}e^{O(\sqrt{n\ln n})}\right) = \frac{1}{2} - \mathrm{o}(1).$$

Το μη τετριμμένο στοιχείο του παραπάνω θεωρήματος δίνεται από το ακόλουθο

Θεώρημα 2.2. *Έστω M_n όπως παραπάνω. Έχουμε*

$$P\left(|\det(M_n)| \geq \sqrt{n!}e^{-29\sqrt{n\ln n}}\right) = 1 - \mathrm{o}(1). \tag{1}$$

Από το θεώρημα 2.2 έπεται εύκολα το 2.1. Παρατηρούμε ότι $Var(|\det(M_n)|) = E(\det(M_n)^2) = n!$ ως εξής: Αναπτύσσουμε στους $(n!)^2$ όρους, χρησιμοποιούμε τη γραμμικότητα της μέσης τιμής και παρατηρούμε ότι επιβιώνουν μόνο οι $n!$ όροι που αντιστοιχούν σε γινόμενο δύο ίδιων μεταθέσεων, με τον κάθε ένα να συνεισφέρει 1 στο άθροισμα. Επομένως η ανισότητα Tchebychev δίνει

$$P\left(|\det(M_n)| \leq \sqrt{n!}h(n)\right) = 1 - \mathrm{o}(1) \tag{2}$$

για κάθε $h(x)$ που τείνει στο άπειρο καθώς $x \to \infty$. Διαλέγουμε $h(x) = e^{\sqrt{x\ln x}}$ και με πιθανότητα $1 - o(1)$ ισχύουν ταυτόχρονα τα ενδεχόμενα στις (1) και (2) άρα και

$$P\left(|\det(M_n)| = \sqrt{n!}e^{O(\sqrt{n\ln n})}\right) = 1 - \mathrm{o}(1). \tag{3}$$

Έπειτα από το γεγονός ότι τα $\{M_n : |\det(M_n)| = \sqrt{n!}e^{O(\sqrt{n\ln n})} \ \wedge \ \det M_n > 0\}$ και $\{M_n : |\det(M_n)| = \sqrt{n!}e^{O(\sqrt{n\ln n})} \ \wedge \ \det M_n < 0\}$ είναι ισοπληθικά έπονται αμέσως οι εξισώσεις του θεωρήματος 2.1.

Εδώ πρέπει να παρατηρήσουμε ότι αφού $n! \sim e^{n\ln n - n + O(\ln n)}$ η απόκλιση από την πραγματική τιμή στον εκθέτη είναι της τάξης της ρίζας της κύριας έκφρασης, οπότε πράγματι υπάρχει μεγάλη συγκέντρωση γύρω από την τυπική απόκλιση $\sqrt{n!}$.

2.2 Λήμματα

Σε αυτήν την ενότητα θα απαριθμήσουμε τα λήμματα που θα χρησιμοποιηθούν για την απόδειξη του 2.1 και βάσει αυτών θα δείξουμε το θεώρημα. Τα λήμματα θα αποδειχθούν στην επόμενη ενότητα.

Οι εκτιμήσεις που θα γίνουν για την ορίζουσα βασίζονται όλες στον εξής απλό τύπο:

$$\det(M_n) = \prod_{j=1}^{n} dist(X_{j+1}, W_j). \tag{4}$$

Στην παραπάνω εξίσωση X_j είναι η $j-$οστή γραμμή του M_n, $W_j =< X_1, \cdots, X_j >$ και $dist$ η συνάρτηση απόστασης σημείου από τον κάθε υπόχωρο. Ο παραπάνω τύπος είναι απλά η έκφραση του όγκου του παραλληλεπιπέδου που σχηματίζουν τα X_j σαν βάση επί ύψος επαγωγικά: στη μία διάσταση προφανώς το παραπάνω ισχύει. Αν ο όγκος ενός $n-1-$ παραλληλεπιπέδου δίνεται από τον παραπάνω τύπο, το $n-$παραλληλεπίπεδο που σχηματίζεται από το προηγούμενο με προσθήκη μιας επιπλέον διάστασης (σε μορφή διανύσματος έξω από τον υπόχωρο που περιέχει τα προηγούμενα), ο καινούριος όγκος είναι ο προηγούμενος επί το ύψος του X_n σχετικά με το $n-1-$παραλληλεπίπεδο. Το τελευταίο είναι ακριβώς $dist(X_n, W_{n-1})$.

Από τον παραπάνω τύπο έπεται ότι αν καταφέρουμε να πετύχουμε ανισότητες συγκέντρωσης για τους παράγοντες $dist(X_j, W_{j-1})$ ξεχωριστά, μπορούμε να χρησιμοποιήσουμε την ανεξαρτησία των ενδεχομένων και να μεταφέρουμε την ανισότητα συγκέντρωσης στο γινόμενο. Με αυτό το σκεπτικό, στρεφόμαστε στις ποσότητες $dist(X_j, W_{j-1})$. Τα λήμματα που ακολουθούν αφορούν τη συμπεριφορά τους ως προς ενδεχόμενα που μας ενδιαφέρουν. Θα δώσουμε εδώ την απόδειξη του πρώτου λήμματος που αναφέρουμε μια και είναι πολύ απλή και δείχνει τι είδους αντικείμενα θα μεταχειριζόμαστε συνήθως στην πορεία των υπόλοιπων αποδείξεων.

Λήμμα 2.3. *Έστω W δεδομένος υπόχωρος του $\mathbb{R}^n$ με $dim(W) = d$ και X τυχαίο ± 1 διάνυσμα. Τότε $P(dist(X, W) = 0) \leq 2^{d-n}$.*

Απόδειξη. Έστω ότι $X \in W$. Αν θεωρήσουμε το W ως τον πυρήνα του $A \in M_{n\times n}$ μπορούμε με μια απαλοιφή Gauss να φέρουμε τον A σε τριγωνική μορφή, με ακριβώς d γραμμές μηδενικές. Επομένως από εκεί μπορούμε να εκφράσουμε τις $n-d$ συντεταγμένες του X ως συναρτήσεις των d ελεύθερων συντεταγμένων που αντιστοιχούν στα μηδενικά της τριγωνικής μορφής του A. Επειδή ο X έχει στοιχεία ± 1, θα υπάρχουν το πολύ 2^d επιλογές του X από τις δυνατές 2^n, άρα $P(X \in W) \leq 2^{d-n}$. □

Το επόμενο λήμμα λέει ότι η ποσότητα $dist(X, W)$ συγκεντρώνεται κοντά στο $\sqrt{n-d}$ με μεγάλη πιθανότητα για κάθε δεδομένο υπόχωρο W διάστασης $d \leq n-4$.

Λήμμα 2.4. *Έστω W δεδομένος υπόχωρος του $\mathbb{R}^n$ με $dim(W) = d \leq n-4$ και X τυχαίο ± 1 διάνυσμα. Τότε*

$$E(dist^2(X,W)) = n - d. \tag{5}$$

$$P(|dist(X,W) - \sqrt{n-d}| \geq 1 + t) \leq 4e^{\frac{-t^2}{16}}. \tag{6}$$

Το παραπάνω λήμμα δεν είναι αρκετό για να μας δώσει τις εκτιμήσεις που θέλουμε όταν η διάσταση του W είναι κοντά στο n. Σε εκείνη την περίπτωση ακόμη και η πληροφορία ότι το $dist(X,W) \neq 0$ με μεγάλη πιθανότητα δεν εξάγεται από το (6). Το επόμενο λήμμα μας επιτρέπει να χειριστούμε διαστάσεις κοντά στο n.

Λήμμα 2.5. *Έστω W υπόχωρος παραγόμενος από τυχαία ± 1 διανύσματα με $dimW = d \leq n-1$ και X τυχαίο. Τότε*

$$P\left(dist(X,W) \leq \frac{1}{4n}\right) = O\left(\frac{1}{\sqrt{\ln n}}\right). \tag{7}$$

Το παραπάνω λήμμα ισχύει για οποιασδήποτε διάστασης υπόχωρο, όμως δε μπορεί να χρησιμοποιηθεί από μόνο του για να δώσει το θεώρημα 2.2 καθώς το φράγμα δεν είναι αρκετά ισχυρό ώστε να υπερνικήσει το μεγάλο πλήθος των όρων που συμβάλλουν. Γιάυτό θα χρησιμοποιηθεί το 2.4 για τις περισσότερες μικρές διαστάσεις, ενώ το 2.5 θα καλύψει τις υπόλοιπες μεγάλες διαστάσεις που δε μπορούν να αντιμετωπιστούν από το προηγούμενο.

Προτού χρησιμοποιήσουμε τα λήμματα για να αποδείξουμε το κεντρικό θεώρημα, θα χρησιμοποιήσουμε το τελευταίο λήμμα για να δείξουμε ότι

$$P(\det M_n = 0) = o(1)$$

ή ισοδύναμα

$$P(\exists\ j : dist(X_{j+1}, W_j) = 0) = o(1).$$

Καταρχήν, από το 2.3 έχουμε ότι $P(dist(X,W_j) = 0) \leq 2^{j-n}$ για τυχαίο X. Η ποσότητα $dist(X,W_j)$ είναι φθίνουσα ως προς j γιατί $W_{j-1} \subset W_j$ και η απόσταση από ένα υποσύνολο είναι μεγαλύτερη ή ίση από την απόσταση από το όλο σύνολο. Επομένως η πιθανότητα $P(dist(X,W_j) = 0)$ είναι αύξουσα με το j. Άρα

$$P(\exists\ j : dist(X_{j+1}, W_j) = 0) \leq \sum P(dist(X,W_j) = 0)$$
$$\leq 2^{-k} + kP(dist(X,W_{n-1}) = 0),$$

όπου στην τελευταία σχέση αθροίσαμε τις εκτιμήσεις απο 2^{1-n} μέχρι 2^{-k+1} παρατηρώντας ότι το άθροισμα είναι το πολύ 2^{-k} και κάναμε τις υπόλοιπες k πιθανότητες $P(dist(X,W_{n-1}) = 0)$. Το 2.5 λέει ότι

$$P(dist(X,W_{n-1}) = 0) = O\left(\frac{1}{\sqrt{\ln n}}\right)$$

και επομένως αν βάλουμε αυτήν την εκτίμηση στην παραπάνω σχέση και διαλέξουμε $k = (\ln n)^\rho$ με $0 < \rho < \frac{1}{2}$, η ποσότητα στη δεξιά μεριά είναι $o(1)$.

Τώρα, βασιζόμενοι στα λήμματα που διατυπώθηκαν παραπάνω, μπορούμε εύκολα να αποδείξουμε το θεώρημα 2.2.

2.3 Απόδειξη του Θεωρήματος 2.2.

Όπως είπαμε στην εισαγωγή, για την απόδειξη του θεωρήματος αρκεί να ελέγξουμε κάθε παράγοντα $dist(X_{j+1}, W_j)$. Γιαυτό χωρίζουμε τις διαστάσεις σε μικρές και μεγάλες, με σημείο διαχωρισμού το $d_0 = n - (\ln n)^{\frac{1}{4}}$. Θέτουμε

$$\gamma_j = 7\sqrt{\frac{\ln(n-j)}{n-j}}$$

για $1 \le j \le d_0$. Τα γ_j γίνονται οσοδήποτε μικρά καθώς μεγαλώνει το n, επομένως από το λήμμα 2.4 έχουμε ότι η πιθανότητα να ισχύει $dist(X_{j+1}, W_j) \le (1-\gamma_j)\sqrt{n-j}$ όταν $dim(W_j) = j$ είναι το πολύ

$$4e^{-\gamma_j^2(n-j)/16} = 4e^{-49/16\ln(n-j)} \le \frac{1}{(n-j)^2}$$

όταν το j είναι στο εύρος που αναφέραμε πιο πάνω. Αυτό διότι από το λήμμα έχουμε

$$P(|dist(X,W) - \sqrt{n-d}| \ge t+1) \le 4e^{-\frac{t^2}{16}}$$

οπότε

$$P(dist(X,W) \le (1-\gamma_j)\sqrt{n-j}) = P(dist(X,W) - \sqrt{n-j} \le -\gamma_j\sqrt{n-j})$$

$$= P(\sqrt{n-j} - dist(X,W) \ge \gamma_j\sqrt{n-j}) \le P(|dist(X,W) - \sqrt{n-j}| \ge \gamma_j\sqrt{n-j})$$

$$= P(|dist(X,W) - \sqrt{n-j}| \ge 1 + \gamma_j\sqrt{n-j} - 1)$$

$$\le 4e^{-\frac{\gamma_j^2(n-j)-1}{16}} << \frac{1}{(n-j)^2}.$$

Επομένως κάθε $dist(X_{j+1}, W_j), j \le d_0$ είναι τουλάχιστον $(1-\gamma_j)\sqrt{n-j}$ με πιθανότητα $\ge 1 - \sum_j^{d_0} \frac{1}{(n-j)^2} = 1 - \mathrm{o}(1)$.

Εξετάζουμε τώρα τα $dist(X_{j+1}, W_j)$ όταν $d_0 < j \le n-1$. Από το 2.5, η ποσότητα $dist(X_{j+1}, W_j)$ είναι μεγαλύτερη του $\frac{1}{4n}$ για κάθε τέτοιο j με πιθανότητα $1 - \sum_{d_0<j<n} O\left(\frac{1}{\sqrt{\ln n}}\right) = 1 - \mathrm{o}(1)$.

Από τις παραπάνω εκτιμήσεις έχουμε με πιθανότητα $1 - \mathrm{o}(1)$ την ανισότητα

$$\prod_{j=1}^{n} dist(X_{j+1}, W_j) \ge \frac{\sqrt{n!}}{\sqrt{(n-d_0)!}} \left(\frac{1}{4n}\right)^{n-d_0} \prod_{j=0}^{d_0}(1-\gamma_j)$$

απλά πολλαπλασιάζοντας τις επί μέρους ανισότητες. Ο όρος $\frac{1}{\sqrt{(n-d_0)!}}\left(\frac{1}{4n}\right)^{n-d_0} = e^{-1/2(n-d_0)\ln(n-d_0)}e^{-4(n-d_0)\ln(n)} = e^{-\mathrm{o}(\ln(n)^2)}$ διότι $n-d_0 = \mathrm{o}(\ln n)$. Επομένως ο όρος αυτός δε συμβάλλει ουσιαστικά. Ο όρος που συμβάλλει περισσότερο είναι το γινόμενο των $1-\gamma_j$. Για αυτόν τον όρο έχουμε

$$\prod_{j=1}^{d_0}(1-\gamma_j) \geq e^{-2\sum_{j=0}^{d_0}\gamma_j} \geq e^{-14\sum_{j=0}^{d_0}\sqrt{\frac{\ln(n-j)}{n-j}}}.$$

Για το άθροισμα στον εκθέτη χρησιμοποιούμε την προσέγγιση από το αντίστοιχο ολοκλήρωμα και παίρνουμε $\sum_{j=0}^{d_0}\sqrt{\frac{\ln(n-j)}{n-j}} \leq \sqrt{\ln n}\int_0^n x^{-\frac{1}{2}} = 2\sqrt{n\ln n}$. Βάζοντας αυτήν την εκτίμηση στα παραπάνω παίρνουμε ότι με πιθανότητα $1-\mathrm{o}(1)$ ισχύει

$$\prod_{j=1}^{n} dist(X_{j+1}, W_j) \geq \sqrt{n!}e^{-28\sqrt{n\ln n}-\mathrm{o}(\ln^2 n)} \geq \sqrt{n!}e^{-29\sqrt{n\ln n}}.$$

2.4 Αποδείξεις των Λημμάτων

Σε αυτήν την ενότητα θα ολοκληρώσουμε την απόδειξη του 2.2 αποδεικνύοντας τα λήμματα που χρησιμοποιήθηκαν. Κατά πρώτοις αποδεικνύουμε την ανισότητα συγκέντρωσης (6).

Απόδειξη του 2.4. Θεωρούμε τον πίνακα της ορθογώνιας προβολής $P = (p_{ij})$ στον W, συμβολίζουμε με D τον πίνακα που αποτελείται από τη διαγώιο του P και $A = P - D$ το εκτός διαγωνίου κομμάτι του P. Ο $A = (a_{ij})$ είναι πραγματικός, συμμετρικός πίνακας με μηδενικά στη διαγώνιο και κάθε στοιχείο του P έχει απόλυτη τιμή το πολύ ένα. Έπειτα το πυθαγόρειο θεώρημα μας δίνει $dist(X,W)^2 = X^2 - P(X)^2$ για ένα οποιοδήποτε $X \in \mathbb{R}^n$. Επομένως όταν το $X = (\epsilon_1, \cdots, \epsilon_n)$ είναι τυχαίο ± 1 διάνυσμα έχουμε $|X|^2 = n$ και ο προηγούμενος τύπος γίνεται $dist(X,W)^2 = n - P(X)^2$. Θυμίζουμε ότι για έναν πίνακα προβολής έχουμε $P^2 = P$ και επομένως

$$P^2(X) = P(X) = \sum_{i,j\leq n}\epsilon_j\epsilon_j p_{ij}$$

$$= \sum_{i=1}^{n}\epsilon_i^2 p_{ii} + \sum_{i\neq j}\epsilon_i\epsilon_j p_{ij} = tr(P) + \sum_{i,j}\epsilon_i\epsilon_j a_{ij}.$$

Το ίχνος του P όμως είναι $d = dim(W)$ γιατί μένει αναλλοίωτο από αλλαγή βάσης και σε κατάλληλη βάση ο πίνακας P είναι προβολή στον $e_{d+1} = \cdots = e_n = 0$. Συνδυάζοντας τα παραπάνω έχουμε

$$dist(X,W)^2 = n - d - \sum_{i,j}\epsilon_i\epsilon_j a_{ij}.$$

Παίρνοντας μέσες τιμές και στα δύο μέλη παίρνουμε

$$E(dist(X,W)^2) = n - d - \sum_{i,j} E(\epsilon_i)E(\epsilon_j)a_{ij}$$

από την ανεξαρτησία των ϵ_i και αφού έχουν μέση τιμή μηδέν, συνεπάγεται ότι $E(dist(X,W)^2) = n - d$. Αυτή είναι η πρώτη εξίσωση στο λήμμα. Τώρα αν θέσουμε $Y = \sum_{i,j} \epsilon_i \epsilon_j a_{ij}$ έχουμε ότι

$$Y^2 = \sum_{i,j \leq n} \epsilon_i^2 \epsilon_j^2 a_{ij}^2 + \sum_{i \neq k \vee j \neq l} \epsilon_i \epsilon_j \epsilon_k \epsilon_l a_{ij} a_{kl}.$$

Παίρνοντας μέσες τιμές και παρατηρώντας ότι $a_{ii} = 0$ βλέπουμε ότι $E(Y^2) = \sum_{i,j \leq n} a_{ij}^2 = tr(A^2)$. Πάλι από το $P^2 = P$ παίρνουμε $d = tr(P) = tr(P^2) = \sum_{i,j \leq n} p_{ij}^2$ και φυσικά $tr(P) = \sum_{i=1}^n p_{ii}$. Η ανισότητα Cauchy-Schwarz εφαρμοσμένη στα $(1,1,\cdots,1),(p_{11},p_{22},\cdots,p_{nn})$ δίνει $d^2 = (\sum_{i=1}^n p_{ii})^2 \leq n \sum_{i=1}^n p_{ii}^2 \Rightarrow \sum_{i=1}^n p_{ii}^2 \geq \frac{d^2}{n}$. Επομένως $tr(A^2) = \sum_{i,j \leq n} p_{ij}^2 - \sum_{i=1}^n p_{ii}^2 \leq d - \frac{d^2}{n} \leq \min\{d, n-d\}$.

Τώρα θα χρησιμοποιήσουμε αυτά τα φράγματα για να φράξουμε το median της $dist(X,W)$ πάνω και κάτω, ώστε να χρησιμοποιήσουμε την ανισότητα του Talagrand για να πάρουμε το αποτέλεσμα. Πιο συγκεκριμένα, θα δείξουμε ότι το ενδεχόμενο $dist(X,W) \leq \sqrt{n-d} - 1$ έχει πιθανότητα μεγαλύτερη του 1/2 και ότι το ενδεχόμενο $dist(X,W) \geq \sqrt{n-d} + 1$ έχει πιθανότητα το πολύ 1/2, επομένως το median βρίσκεται ανάμεσα στα δύο. Για το $dist(X,W) \geq \sqrt{n-d} + 1$ έχουμε

$$\begin{aligned} P(dist(X,W) \geq \sqrt{n-d} + 1) &\leq P(dist^2(X,W) \geq n - d + 2\sqrt{n-d}) \\ &= P(Y \geq 2\sqrt{n-d}) \\ &\leq P(Y^2 \geq 4(n-d)) \leq \frac{E(Y^2)}{4(n-d)} = \frac{1}{4} \end{aligned}$$

όπου η τελευταία ανισότητα είναι η ανισότητα του Markov. Από την άλλη,

$$P(dist(X,W) \geq \sqrt{n-d} - 1) \leq P(Y \leq -2\sqrt{n-d} + 1))$$

$$\leq P(Y^2 \leq 4(n-d) - 4\sqrt{n-d} + 1) \leq \frac{E(Y^2)}{4(n-d) - \sqrt{n-d} + 1} \leq \frac{4}{9}$$

από όπου παίρνουμε το ζητούμενο, δηλαδή ότι το median M ικανοποιεί

$$\sqrt{n-d} - 1 \leq M \leq \sqrt{n-d} + 1.$$

Η ανισότητα του Talagrand δίνει, για την κυρτή, 1-Lipschitz συνάρτηση $dist(X,W)$ στο $\{-1,1\}^n$ ότι η απόσταση από το median ικανοποιεί

$$P(|dist(X,W) - M| \geq t) \leq 4e^{-\frac{t^2}{16}}.$$

Συνδυάζοντας αυτό με το $\sqrt{n-d} - 1 \leq M \leq \sqrt{n-d} + 1$ έπεται το λήμμα. $\square$

Περνάμε τώρα στην απόδειξη του επόμενου λήμματος, το οποίο μας επιτρέπει να ελέγξουμε τη συνάρτηση $dist(X,W)$ για υπόχωρους μεγάλης διάστασης(κοντά στο n) με μεγάλη πιθανότητα. Γιάυτό το σκοπό ορίζουμε την έννοια του l-τυπικού υπόχωρου W. Ένας υπόχωρος W παραγόμενος από τυχαία X_i λέγεται l-τυπικός αν κάθε κάθετο διάνυσμα $u \in W^{\perp}$ έχει τουλάχιστον l συντεταγμένες οι οποίες έχουν απόλυτες τιμές τουλάχιστον $\frac{1}{2n}$. Για τέτοιου είδους υπόχωρους έχουμε το εξής αποτέλεσμα:

Λήμμα 2.6. *Έστω W υπόχωρος παραγόμενος από τυχαία ± 1 διανύσματα και $l \leq n$ τέτοιο ώστε ο W να είναι l-τυπικός. Τότε*

$$P\left(dist(X,W) \leq \frac{1}{4n}\right) = O\left(\frac{1}{\sqrt{l}}\right). \tag{8}$$

Απόδειξη. Χωρίς βλάβη γενικότητας μπορούμε να υποθέσουμε ότι υπάρχει κάθετο $w = (w_1, \cdots, w_n)$ τέτοιο ώστε $|w_1|, \cdots, |w_l| \geq \frac{1}{2n}$. Γράφουμε $X = (x_1, \cdots, x_n)$. Αν εκφράσουμε την απόσταση $dist(X,W)$ συναρτήσει του καθέτου w έχουμε

$$P(dist(X,W) \leq \frac{1}{4n}) = P(|x_1 w_1 + \cdots x_n w_n| \leq \frac{1}{4n})$$

$$\leq \sup_{x \in \mathbb{R}} P(|x_1 w_1 + \cdots x_l w_l - x| \leq \frac{1}{4n}).$$

Κάνοντας την αντικατάσταση $y = 2nx - \frac{1}{2}$ το παραπάνω supremum γίνεται

$$\sup_{y \in \mathbb{R}} P(2n x_1 w_1 + \cdots 2n x_l w_l \in [y, y+1]).$$

Το λήμμα Littlewood-Offord λέει ότι αν $a_1, \cdots, a_n$ έχουν απόλυτες τιμές τουλάχιστον ένα, τότε για κάθε διάστημα I μήκους το πολύ ένα ισχύει $P(\sum_i a_i x_i \in I) \leq \frac{1}{\sqrt{n}}$. Εφαρμόζοντας αυτό το γεγονός στο παραπάνω supremum παίρνουμε το αποτέλεσμα. □

Απόδειξη του 2.5. Αν αποδείξουμε το λήμμα για υπερεπίπεδα, θα έχουμε τελειώσει, διότι όσο μικραίνει η διάσταση του υπόχωρου(αρχίζοντας από ένα υπερεπίπεδο και περιοριζόμενοι σε φωλιασμένους υπόχωρους), τόσο η ποσότητα $dist(X,W)$ μεγαλώνει, επομένως η πιθανότητα να είναι η απόσταση μικρή μειώνεται. Έστω λοιπόν ότι το W είναι τυχαίο υπερεπίπεδο. Αν θέσουμε $l = \frac{\ln n}{10}$ αρκεί τότε να δείξουμε ότι η πιθανότητα το W να μην είναι l-τυπικό είναι $O\left(\frac{1}{\sqrt{n}}\right)$. Αν το W δεν είναι l-τυπικό τότε ένα μοναδιαίο κάθετο διάνυσμα w στο W έχει τουλάχιστον $n-l$ συντεταγμένες μικρότερες κατάπόλυτη τιμή από ένα(έστω V το σύνολο αυτών των w). Αυτές κατανέμονται ανάμεσα στις n συντεταγμένες του w με $\binom{n}{l}$ τρόπους. Επειδή το X έχει ισοκατανεμημένες συντεταγμένες, έχουμε ότι η πιθανότητα το W να μην είναι l-τυπικό φράσσεται από το $\binom{n}{l} P(w \perp W)$ για κάποιο $w = (w_1, \cdots, w_n)$ τέτοιο ώστε

$|w_{l+1}|, \cdots, |w_n| \leq \frac{1}{2n}$. Έστω ότι $W =< X_1, \cdots, X_n >$. Τότε γράφουμε $X_i = (x_{i1}, \cdots, x_{in})$ και έχουμε $\sum_j x_{ij} w_j = 0$ για κάθε j. Επειδή $x_{ij} = \pm 1$ και όλα τα $|w_j| \leq \frac{1}{2n}$ για $j > l$ έχουμε

$$|\sum_{j=l+1}^{n} x_{ij} w_j| \leq \frac{n-l}{2n} \leq \frac{1}{2}.$$

Από την άλλη το άθροισμα των απολύτων τιμών μέχρι και το w_l ικανοποιεί

$$\sum_{j=1}^{l} |w_j| \geq \sum_{j=1}^{l} |w_j|^2 = 1 - \sum_{j=l+1}^{n} |w_j|^2 \geq 1 - \frac{n-l}{4n^2} \geq 1 - \frac{1}{4n}.$$

Επομένως για κάθε $i = 1, \cdots, n-1$ κάποιο $x_{ij} w_j, j \leq l$ πρέπει να είναι αρνητικό, αλλιώς το $\sum_{j=1}^{n} x_{ij} w_j$ θα ήταν θετικό από τις παραπάνω ανισότητες.

Άρα αν διαλέξουμε $y_i, i \leq l$ πρόσημα τέτοια ώστε το $y_i w_i > 0, i \leq l$, οι ακολουθίες $(x_{ij})_{j \leq l}$ και $(y_j)_{j \leq l}$ δε συμπίπτουν για κανένα $i \leq n-1$. Αυτό το τελευταίο ενδεχόμενο έπεται από το ενδεχόμενο το $w \perp W$ για w ορισμένο όπως παραπάνω, επομένως η αντίστοιχη πιθανότητα θα είναι μεγαλύτερη από εκείνη του $w \perp W$ για τέτοια w. Επομένως

$$P(w \perp W, w \in V) \leq \sum_{y_1, \cdots, y_n \in -1, 1} P((x_{ij})_{j \leq l} \neq (y_j)_{j \leq l}).$$

Επειδή τα x_{ij} είναι ισοκατανεμημένες, ανεξάρτητες Bernoulli τυχαίες μεταβλητές, έπεται ότι οι πιθανότητες μέσα στο άθροισμα είναι ίσες με $(1 - 2^{n-l})^{n-1}$, άρα τελικά το άθροισμα είναι $2^l (1 - 2^{n-l})^{n-1}$. Επομένως η πιθανότητα ο W να μην είναι l-τυπικός φράσσεται από $\binom{n}{l} 2^l (1 - 2^{n-l})^{n-1} \leq n^l 2^l e^{-2^l (n-1)}$ που είναι όχι μόνο ότι ζητούσαμε, αλλά και εκθετικά μικρό. $\square$

3 Η πιθανότητα ένας τυχαίος Bernoulli πίνακας να είναι μη αντιστρέψιμος

3.1 Κύριο Αποτέλεσμα

Μια αναφορά στην ιστορία των εκτιμήσεων που θα παρουσιαστούν εδώ γίνεται στην εισαγωγή. Τώρα θα απαριθμήσουμε τα βασικά αποτελέσματα με τα οποία θα ασχοληθούμε στην παρούσα ενότητα. Για ευκολία στο συμβολισμό θα θέσουμε $N = 2^n$. Επίσης ορίζουμε ϵ_1 να είναι η μοναδική λύση στο $(0, \frac{1}{2})$ της εξίσωσης $h(\epsilon_1) + \frac{\epsilon_1}{\log_2 \frac{16}{15}} = 1$ όπου $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ η συνάρτηση εντροπίας. Τότε έχουμε το εξής βασικό θεώρημα:

Θεώρημα 3.1. *Έστω M_n τυχαίος Bernoulli $n \times n$ πίνακας με γραμμές $X_1, \cdots, X_n$. Τότε*

$$P(\det(M_n) = 0) = N^{-\epsilon_1 + o(1)}. \tag{9}$$

3.2 Αναγωγές

Αυτό το θεώρημα θα προκύψει από μία σειρά αναγωγών σε όλο και πιο εύκολα ελεγχόμενα αντικείμενα. Θα δώσουμε τώρα τους ορισμούς και τα αντίστοιχα λήμματα που θα οδηγήσουν στην απόδειξη του θεωρήματος. Η πρώτη αναγωγή γίνεται παρατηρώντας ότι μπορούμε να υποθέσουμε ότι το $< X_1, \cdots, X_n >$ είναι υπερεπίπεδο, λόγω του ακόλουθου λήμματος:

Λήμμα 3.2.

$$P(\det(M_n) = 0) \leq N^{o(1)} P(\dim < X_1, \cdots, X_n >= n-1).$$

Φυσικά αν δείξουμε ότι

$$\sum_V P(< X_1, \cdots, X_n >= V) \leq N^{-\epsilon_1 + o(1)},$$

όπου όταν γράφουμε V θα εννοούμε πάντα υπερεπίπεδα, θα έχουμε τελειώσει και πράγματι έτσι θα κινηθούμε. Όμως για να χειριστούμε το παραπάνω άθροισμα, θα χρειαστούμε επιπλέον αναγωγές. Θα λέμε ότι ένα υπερεπίπεδο V είναι μη τετριμμένο αν παράγεται από ± 1 διανύσματα, ή ισοδύναμα ότι η $P(< X_1, \cdots, X_n >= V) \neq 0$. Από εδώ και πέρα θα μελετάμε μόνο μη τετριμμένα υπερεπίπεδα εκτός αν αναφέρεται αλλιώς. Επιπλέον ένα υπερεπίπεδο V θα λέγεται εκφυλισμένο αν υπάρχει κάθετο διάνυσμα w με το πολύ $\log\log n$ συντεταγμένες μη μηδενικές. Το παρακάτω λήμμα επιτρέπει την αναγωγή σε μη εκφυλισμένα(και φυσικά μη τετριμμένα) υπερεπίπεδα.

Λήμμα 3.3. *Το πλήθος των εκφυλισμένων υπερεπιπέδων φράσσεται από $N^{o(1)}$.*

Αν θυμηθούμε ότι για οποιοδήποτε V η πιθανότητα $P(X \in V) \leq \frac{1}{2}$ πχ. από το λήμμα 2.3, έπεται ότι το παραπάνω άθροισμα πάνω στα εκφυλισμένα υπερεπίπεδα φράσσεται από $N^{-1+o(1)}$, επομένως αρκεί να δείξουμε το θεώρημα για μη εκφυλισμένα υπερεπίπεδα.

Τώρα κάνουμε έναν πιο λεπτό διαχωρισμό των υπερεπιπέδων, βάσει της διακριτής συνδιάστασής τους. Λέμε ότι ένα V έχει διακριτή συνδιάσταση $d(V)$ αν το $d(V)$ είναι το μοναδικό πολλαπλάσιο του $\frac{1}{n}$ τέτοιο ώστε

$$N^{-\frac{d(V)}{n}-\frac{1}{n^2}} < P(X \in V) \leq N^{-\frac{d(V)}{n}}.$$

Η διακριτή συνδιάσταση έχει το σκοπό να περιγράψει κατά πόσο ένα υπερεπίπεδο συμπεριφέρεται σαν υπερεπίπεδο ως προς τον "όγκο" όταν το κοιτάζουμε μέσα στο $\{-1,1\}^n$. Ορίζουμε Ω_d να είναι το σύνολο των υπερεπιπέδων διακριτής συνδιάστασης d. Από τον ορισμό του $d(V)$ φαίνεται ότι $1 \leq d(V) \leq n$, και επειδή είναι πολλαπλάσιο του $\frac{1}{n}$ υπάρχουν $n^2 = N^{o(1)}$ το πλήθος κλάσεις Ω_d. Επομένως μπορούμε να περιορίσουμε το άθροισμα σε μία κλάση και να δείξουμε ότι:

Λήμμα 3.4.

$$\sum_{V \in \Omega_d} P(< X_1, \cdots, X_n >= V) \leq N^{-\epsilon_1 + o(1)}.$$

Δεν θα χειριστούμε όλα τα Ω_d με τον ίδιο τρόπο. Τα μεγάλα d, αυτά δηλαδή κοντά στο $n-1$, είναι εύκολα στη μεταχείριση. Συγκεκριμένα, έχουμε το εξής, του οποίου την απόδειξη παραθέτουμε άμεσα λόγω της ευκολίας της.

Λήμμα 3.5. *Έστω* $d \geq (\epsilon_1 - o(1))n$. *Τότε*

$$\sum_{V \in \Omega_d} P(< X_1, \cdots, X_n >= V) \leq N^{-\epsilon_1 + o(1)}.$$

Απόδειξη. Αν $V =< X_1, \cdots, X_n >$, έχουμε n τρόπους να επιλέξουμε $n-1$ από τα X_i που να παράγουν ήδη το V. Επομένως

$$\sum_{V \in \Omega_d} P(< X_1, \cdots, X_n >= V) \leq n \sum_{V \in \Omega_d} P(< X_1, \cdots, X_{n-1} >= V) P(X \in V)$$

$$\leq n N^{-\frac{d}{n}} \sum_{V \in \Omega_d} P(< X_1, \cdots, X_{n-1} >= V)$$

$$\leq n N^{-\frac{d}{n}} = N^{-\frac{d}{n} + o(1)}.$$

□

Τα κύρια συστατικά της παραπάνω απόδειξης ήταν το μέγεθος του d και το γεγονός ότι τα ενδεχόμενα $V =< X_1, \cdots, X_{n-1} >$ είναι ξένα για διαφορετικά V. Για τα μικρότερα d ένα τέτοιο απλό επιχείρημα δεν δουλεύει και θα χρειαστούν πιο περίπλοκοι μηχανισμοί. Παρακάτω παραθέτουμε τα λήμματα που συνιστούν αυτούς τους μηχανισμούς.

3.3 Μηχανισμοί για τον χειρισμό μικρών διακριτών συνδιαστάσεων

Από εδώ και πέρα το d θα βρίσκεται στο εύρος $1 \le d \le (\epsilon_1 - \mathrm{o}(1))n$. Το βασικό εργαλείο που χρησιμοποιούμε είναι η εξής τροποποίηση των τυχαίων μεταβλητών X_i:

Ορισμός 3.6. *Έστω $0 \le \mu \le 1$ και $x^{(\mu)}$ τυχαία μεταβλητή που παίρνει την τιμή μηδέν με πιθανότητα $1-\mu$ και τις τιμές ± 1 με ίση πιθανότητα $\frac{\mu}{2}$. Τότε συμβολίζουμε με $X^{(\mu)} = (x_1^{(\mu)}, \cdots, x_m^{(\mu)})$ όπου οι συντεταγμένες είναι ανεξάρτητες.*

Η σχέση που συνδέει τα δύο είδη τυχαίων μεταβλητών δίνεται από το ακόλουθο:

Λήμμα 3.7.

$$P(X \in V) \le (\frac{1}{2} + \mathrm{o}(1))P(X^{(\mu)} \in V).$$

Επιπλέον ορίζουμε

$$\gamma = \frac{d}{n \log_2 \frac{16}{15}}.$$

Από το εύρος στο οποίο βρίσκεται το d έπεται ότι $0 < \gamma < 1$. Θέτουμε $\epsilon_2 = \min(\epsilon_1, \gamma)$. Αν ονομάσουμε B_V το γεγονός ότι τα $X_1, \cdots, X_n$ παράγουν το V, το ζητούμενο θεώρημα παίρνει τη μορφή

$$\sum_{V \in \Omega_d} P(B_V) \le N^{-\epsilon_1 + \mathrm{o}(1)}.$$

Για την εκτίμηση αυτού του αθροίσματος θα παρεμβάλλουμε τα ενδεχόμενα A_V που ορίζονται για κάθε V να είναι το ενδεχόμενο τα $X_1^{(\frac{1}{16})}, \cdots, X_{(1-\gamma)n}^{(\frac{1}{16})}$, $X_1', \cdots, X_{(\gamma-\epsilon_2)n}'$ να είναι μέσα στο V και γραμμικά ανεξάρτητα. Εννοείται ότι οι μεταβλητές είναι όλες ανεξάρτητες μεταξύ τους και οι τονισμένες έχουν κανονικά Bernoulli συντεταγμένες. Η ιδιότητα του A_V την οποία δε μοιράζεται με το B_V που θα χρησιμοποιήσουμε είναι η εξής:

Λήμμα 3.8.

$$P(A_V) \ge N^{(1-\gamma)-(1-\epsilon_2)d - \mathrm{o}(1)}.$$

Με αυτά τα εργαλεία είμαστε τώρα έτοιμοι να αποδείξουμε το βασικό θεώρημα αυτής της ενότητας.

3.4 Απόδειξη του θεωρήματος

Από τα παραπάνω λήμματα αρκεί να αποδείξουμε ότι

$$\sum_{V \in \Omega_d} P(B_V) \le N^{-\epsilon_1 + \mathrm{o}(1)},$$

δηλαδή το λήμμα 3.4 για μικρά d. Παρατηρώντας ότι τα A_V και B_V είναι ανεξάρτητα, έχουμε

$$P(B_V) = \frac{P(B_V \wedge A_V)}{P(A_V)} \leq N^{-(1-\gamma)+(1-\epsilon_2)d+\mathrm{o}(1)} P(B_V \wedge A_V)$$

από το 3.8. Αν έχουμε λοιπόν διανύσματα $X_1^{(\frac{1}{16})}, \cdots, X_{(1-\gamma)n}^{(\frac{1}{16})}, X'_1, \cdots, X'_{(\gamma-\epsilon_2)n}$, $X_1, \cdots, X_n$ που ικανοποιούν και το A_V και το B_V τότε από τα $X_1, \cdots, X_n$ υπάρχουν $\epsilon_2 n - 1$ διανύσματα που μαζί με τα διανύσματα που εμφανίζονται στο A_V παράγουν το V. Αυτά τα $\epsilon_2 n - 1$ διανύσματα μοιράζονται ανάμεσα στα n με $\binom{n}{\epsilon_2 n-1} = N^{h(\epsilon_2)+\mathrm{o}(1)}$ τρόπους, οπότε μπορούμε να υποθέσουμε ότι τα διανύσματα αυτά είναι τα $X_1, \cdots, X_{\epsilon_2 n-1}$. Έστω λοιπόν C_V το ενδεχόμενο αυτή η συλλογή να παράγει το V. Τότε

$$P(B_V) \leq N^{-(1-\gamma)+(1-\epsilon_2)d+h(\epsilon_2)+\mathrm{o}(1)} P(C_V \wedge \{X_{\epsilon_2 n}, \cdots, X_n\} \in V).$$

Τα δύο ενδεχόμενα C_V και $\{X_{\epsilon_2 n}, \cdots, X_n\} \in V$ είναι ανεξάρτητα, αφού το πρώτο δεν εξαρτάται από X_i με $i \geq \epsilon_2 n$. Επομένως

$$P(C_V \wedge \{X_{\epsilon_2 n}, \cdots, X_n\} \in V) = P(C_V) P(X \in V)^{n-\epsilon_2 n+1},$$

από όπου παίρνουμε, λαμβάνοντας υπόψιν ότι $V \in \Omega_d$,

$$P(B_V) \leq N^{-(1-\gamma)+(1-\epsilon_2)d+h(\epsilon_2)+\mathrm{o}(1)} N^{-\frac{d}{n}(n-\epsilon_2 n+1)} P(C_V)$$

$$= N^{-(1-\gamma)+h(\epsilon_2)-\epsilon_1+\mathrm{o}(1)} P(C_V).$$

Αθροίζοντας ως προς όλα τα $V \in \Omega_d$ και παρατηρώντας ότι τα διάφορα C_V είναι ξένα ενδεχόμενα, έχουμε

$$\sum_{V \in \Omega_d} P(B_V) \leq N^{-(1-\gamma)+h(\epsilon_2)-\epsilon_1+\mathrm{o}(1)}.$$

Τώρα παρατηρούμε ότι $\frac{d}{n} \leq \epsilon_1, \epsilon_2 \leq \epsilon_1, \frac{1}{\log_2 \frac{16}{15}} - 1 \geq 0$ και η συνάρτηση εντροπίας είναι αύξουσα στο $(0, \frac{1}{2})$. Επομένως από τον ορισμό του γ και τις παραπάνω ανισότητες έχουμε

$$\sum_{V \in \Omega_d} P(B_V) \leq N^{h(\epsilon_1)+\epsilon_1(\frac{1}{\log_2 \frac{16}{15}}-1)-1+\mathrm{o}(1)}$$

και από τον ορισμό του ϵ_1 παίρνουμε ακριβώς αυτό που θέλαμε.

3.5 Αποδείξεις των Λημμάτων

Απόδειξη του 3.2. Έστω $d \leq n-1$ τέτοιο ώστε να υπάρχουν $X_1, \cdots, X_{d+1}$ τέτοια ώστε $\dim(< X_1, \cdots, X_{d+1} >) = d$. Εφόσον τα διανύσματα είναι γραμμικά εξαρτημένα, ένα τέτοιο d πάντα θα υπάρχει. Υπάρχουν το πολύ $n-1$ τέτοια d, επομένως αρκεί να δειχθεί ότι

$$P(dim(< X_1, \cdots, X_{d+1} >) = d) \leq CP(dim(< X_1, \cdots, X_{d+1} >) = n-1)$$

για κάποια σταθερά $C > 0$. Όμως

$$P(dim(< X_1, \cdots, X_{d+2} >) = d+1 | dim(< X_1, \cdots, X_{d+1} >) = d) \geq 1 - 2^{d-n}$$

από το λήμμα 2.3, επομένως με διαδοχικές εφαρμογές του νόμου της ολικής πιθανότητας παίρνουμε

$$\begin{aligned} & P(dim(< X_1, \cdots, X_{d+1} >) = n-1) \\ \geq & \prod_{k=2}^{n-d} (1 - 2^{-k}) P(dim(< X_1, \cdots, X_{d+1} >) = d) \\ \geq & \prod_{k=2}^{\infty} (1 - 2^{-k}) P(dim(< X_1, \cdots, X_{d+1} >) = d) \\ = & \frac{1}{C} P(dim(< X_1, \cdots, X_{d+1} >) = d) \end{aligned}$$

από όπου παίρνουμε αυτό που θέλουμε. □

Απόδειξη του 3.3. Έστω ένα w κάθετο στο V με $k \leq \log\log n$ συντεταγμένες διάφορες του μηδενός. Υπάρχουν το πολύ $\sum_{k \leq \log\log n} \binom{n}{k} \leq \log\log n n^{log \log n} \leq N^{o(1)}$ θέσεις για αυτές τις συντεταγμένες, επομένως υποθέτουμε ότι οι μη μηδενικές συντεταγμένες είναι οι $w_1, \cdots, w_k$. Επομένως το σύνολο των V που αντιστοιχούν σε αυτά τα διανύσματα καθορίζεται πλήρως από τις τιμές των X_i από τα οποία παράγονται στις πρώτες $k \leq \log\log n$ συντεταγμένες, διότι οποιεσδήποτε και να είναι οι τιμές των X_i στις υπόλοιπες συντεταγμένες, το εσωτερικό γινόμενο με το w δεν αλλάζει τιμή, επομένως καθορίζουν το ίδιο V. Το πλήθος αυτών των X_i λοιπόν φράσσεται από $2^{2^{\log\log n}} = N^{o(1)}$, και έπεται το αποτέλεσμα.

□

Μένει τώρα να αποδείξουμε τα λήμματα 3.7 και 3.8 που χρησιμοποιήθηκαν στην απόδειξη του θεωρήματος (το 3.7 θα χρησιμοποιηθεί μόνο στην απόδειξη του 3.8).

Απόδειξη του 3.7. Τα εργαλεία που θα χρησιμοποιηθούν για αυτήν την απόδειξη διαφέρουν από αυτά στις προηγούμενες ενότητες. Εδώ θα εκφράσουμε την πιθανότητα $P(X^{(\mu)} \in V)$ με όρους της ανάλυσης Fourier. Ξεκινάμε παρατηρώντας ότι αν το V ορίζεται από ένα μοναδιαίο κάθετο v, τότε

$$P(X^{(\mu)} \in V) = P(X^{(\mu)}v = 0) = E\left(\int_0^1 e^{2\pi i t X^{(\mu)}v} dt\right)$$

$$= \int_0^1 E(e^{2\pi i t(\eta_1^{(\mu)}v_1 + \cdots \eta_n^{(\mu)}v_n)})dt$$

$$= \int_0^1 \prod_{j=1}^n \left((1-\mu) + \mu\cos(2\pi t v_j)\right) dt.$$

Θα χρησιμοποιήσουμε αυτήν την παράσταση για $\mu = 1$ που αντιστοιχεί στην περίπτωση των unbiased Bernoulli μεταβλητών, και για $\mu = \frac{1}{16}$ που αντιστοιχεί σε μια βεβαρυμένη έκδοση. Παίρνουμε επομένως

$$P(X \in V) = \int_0^1 \prod_{j=1}^n \cos(2\pi t v_j)dt,$$

$$P(X^{(\frac{1}{16})} \in V) = \int_0^1 \prod_{j=1}^n \left(\frac{15}{16} + \frac{1}{16}\cos(2\pi t v_j)\right) dt.$$

Φράσσουμε το πρώτο ολοκλήρωμα ως εξής:

$$\int_0^1 \prod_{j=1}^n \cos(2\pi t v_j)dt \leq \int_0^1 \prod_{j=1}^n |\cos(2\pi t v_j)|dt$$

$$= \int_0^1 \prod_{j=1}^n |\cos(\pi t v_j)|dt$$

και έπειτα θέτουμε

$$F(t) = \prod_{j=1}^n |\cos(\pi t v_j)|,$$

$$G(t) = \prod_{j=1}^n \left(\frac{15}{16} + \frac{1}{16}\cos(2\pi t v_j)\right).$$

Αυτό που θέλουμε να δείξουμε γίνεται

$$\int_0^1 F(t)dt \leq \left(\frac{1}{2} + \mathrm{o}(1)\right)\int_0^1 G(t)dt.$$

Θα κάνουμε τις εξής εκτιμήσεις για τις F, G:

$$F(t) \leq G(t)^4 \tag{10}$$

$$F(t)F(t') \leq G(t+t')^2, t, t' \in [0,1] \tag{11}$$

$$\int_0^1 G(t)dt = \mathrm{o}(1) \tag{12}$$

η τελευταία ασυμπτωτική σχέση φυσικά θεωρείται ως προς n. Για την πρώτη εκτίμηση αρκεί να δείξουμε κατά σημείο ότι

$$|\cos(t)| \leq \left(\frac{15}{16} + \frac{1}{16}\cos(2t)\right)^4.$$

Αν γράψουμε $\cos(2t) = 1 - 2x, x \in [0,1]$, παίρνουμε $|\cos(t)| = (1-x)^2$ και η ανισότητα γίνεται

$$(1-x)^{\frac{1}{2}} \leq (1 - \frac{x}{8})^4.$$

Αυτή όμως η ανισότητα έπεται άμεσα από το γεγονός ότι η συνάρτηση

$$\log \frac{1}{1-x}$$

είναι κυρτή.

Για τη δεύτερη εκτίμηση προχωράμε πάλι κατά σημείο. Εδώ πρέπει να δείξουμε ότι

$$|\cos t||\cos t'| \leq (\frac{15}{16} + \frac{1}{16}\cos 2(t+t'))^2.$$

Μπορούμε φυσικά να υποθέσουμε ότι $t, t' \in (-\frac{\pi}{2}, \frac{\pi}{2})$ και χρησιμοποιούμε τώρα το γεγονός ότι η συνάρτηση $\log \cos t$ είναι κοίλη στο $(-\frac{\pi}{2}, \frac{\pi}{2})$, επομένως

$$\cos t \cos t' \leq \cos^2 \frac{t+t'}{2} = \frac{1}{2} + frac \cos(t+t')2.$$

Τώρα όπως παραπάνω θέτουμε $\cos(t+t') = 1 - 2x$ και με απλές αλγεβρικές πράξεις παίρνουμε το αποτέλεσμα.

Για την τρίτη εκτίμηση υποθέτουμε, χωρίς βλάβη της γενικότητας, ότι $w_1, \cdots, w_K \neq 0, K > \log \log n$. Τότε παίρνουμε από την ανισότητα του Holder

$$\int_0^1 G(t)dt \leq \int_0^1 \prod_{j=1}^{K} \left(\frac{15}{16} + \frac{1}{16}\cos(2\pi t v_j)\right) dt$$

$$\leq \prod_{j=1}^{K} (\int_0^1 \left(\frac{15}{16} + \frac{1}{16}\cos(2\pi t v_j)\right)^K dt)^{1/K}$$

$$= \int_0^1 \left(\frac{15}{16} + \frac{1}{16}\cos(2\pi t v_j)\right)^K dt = o(1)$$

επειδή $K > \log\log n$.

Τώρα θα χρησιμοποιήσουμε τις εκτιμήσεις μαζί με μια ανισότητα που αφορά σύνολα αθροισμάτων(sum sets). Η δεύτερη εκτίμηση που κάναμε (11) δίνει τον εγκλεισμό

$$\{F(t) > a\} + \{F(t) > a\} \subset \{G(t) > a\}, t \in [0,1].$$

Παίρνοντας μέτρα και χρησιμοποιώντας την ανισότητα των Mann, Kneser και Mcbeath $|A+B| > \min(|A| + |B|, 1)$ παίρνουμε $\min(2|\{F(t) > a\}|, 1) \leq |\{G(t) > a\}|$. Αν $a > o(1)$ η ασυμπτωτική φθορά της $\int G$ δίνει ότι $|\{G(t) > a\}| < 1$, άρα $|\{F(t) > a\}| \leq \frac{1}{2}|\{G(t) > a\}|, a > o(1)$. Ολοκληρώνοντας την ανισότητα ως προς a, παίρνουμε

$$\int_{F(t)>o(1)} F(t)dt \leq \frac{1}{2}\int_0^1 G(t)dt.$$

Για το υπόλοιπο εύρος ολοκλήρωσης, όπου $t < o(1)$, η πρώτη εκτίμηση μας δίνει $F \leq o(F^{1/4}) \leq G$ και επομένως

$$\int_{F(t)<o(1)} F(t)dt \leq o(1)\int_0^1 G(t)dt.$$

Αθροίζοντας αυτές τις δύο παίρνουμε το αποτέλεσμα. □

Απόδειξη του 3.8. Καταρχήν παρατηρούμε ότι η ποσότητα $N^{(1-\gamma)-(1-\epsilon_2)d-o(1)}$ είναι η πιθανότητα του ενδεχομένου A'_V τα $X_1^{(\frac{1}{16})}, \cdots, X_{(1-\gamma)n}^{(\frac{1}{16})}, X'_1, \cdots, X'_{(\gamma-\epsilon_2)n}$ να ανήκουν στο V. Επομένως αν δείξουμε ότι $P(A_V|A'_V) = N^{o(1)}$ θα έχουμε τελειώσει χρησιμοποιώντας το νόμο ολικής πιθανότητας όπως στο λήμμα (3.2). Το $V \in \Omega_d$ και επομένως $P(X \in V) = (1 + O(1/n))2^{-d} \Rightarrow P(X^{(\frac{1}{16})} \in V) = (2 + O(1/n))2^{-d}$. Όπως στο (2.3) βλέπουμε ότι $P(X^{(\frac{1}{16})} \in W) \leq (\frac{15}{16})^{n-dimW}$ για κάθε υπόχωρο W. Επομένως ο νόμος ολικής πιθανότητας δίνει $P(X^{(\frac{1}{16})} \in W | X^{(\frac{1}{16})} \in V) \leq (2 + O(1/n))^{-1}2^d(\frac{15}{16})^{n-dimW}$. Αν θέσουμε E_k να είναι το ενδεχόμενο τα $X_1^{(\frac{1}{16})}, \cdots, X_k^{(\frac{1}{16})}$ να είναι γραμμικά ανεξάρτητα, τα παραπάνω δίνουν

$$P(E_{k+1}|E_k \wedge A_V) \geq 1 - (2 + O(1/n))^{-1}2^d(\frac{15}{16})^{n-k}$$

όταν $k \leq (1-\gamma)n$. Τώρα αν εφαρμόσουμε το νόμο της ολικής πιθανότητας παίρνουμε από τον ορισμό του γ

$$P(E_{(1-\gamma)n}|A'_V) \geq N^{-o(1)}$$

. Τώρα υποθέτουμε ότι $\gamma > \epsilon_1$ γιατί αλλιώς έχουμε τελειώσει. Από τη σχέση $P(X \in V) = (1 + O(1/n))2^{-d}$ παίρνουμε $P(X \in W|X \in V) \leq (1 + O(1/n))^{-1}2^d(\frac{1}{2})^{n-dimW}$. Αν ισχύει το $E_{(1-\gamma)n}$ και θέσουμε $W =< X_1^{(\frac{1}{16})}, \cdots, X_{(1-\gamma)n}^{(\frac{1}{16})} >$ τότε αν U_k το ενδεχόμενο τα $X_1, \cdots, X_k, W$ να είναι ανεξάρτητα,

$$P(U_{k+1}|U_k \wedge A'_V) \geq 1 - (1 + O(1/n))2^d(1/2)^{n-k-(1-\gamma)n}$$

$$\geq 1 - \frac{1}{100}2^{k+(\epsilon_1-\gamma)n}, \quad k < (\gamma - \epsilon_1)n$$

και επομένως ο νόμος ολικής πιθανότητας δίνει

$$P(A_V|A'_V) \geq N^{o(1)} \prod_{k<(\gamma-\epsilon_1)n} P(U_{k+1}|U_k \wedge A'_V) = N^{o(1)}$$

.

□

Αναφορές

[1] J. Kahn, J. Komlos, E. Szemeredi, On the probability that a random ±1 matrix is singular, J. Amer. Math. Soc. 8 (1995), 223-240.

[2] J. Komlos, On the determinant of $(0, 1)$ matrices, Studia Sci. Math. Hungar. 2 (1967), 7-22.

[3] J. Komlos, On the determinant of random matrices, Studia Sci. Math. Hungar. 3 (1968), 387-399.

[4] T.Tao, V. Vu: On random ±1 matrices: singularity and determinant, preprint.