Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Dariusz Matlak
Nr albumu: 373447

Oszacowania norm macierzy losowych

Praca magisterska
na kierunku MATEMATYKA

Praca wykonana pod kierunkiem
prof. dra hab. Rafała Latały
Instytut Matematyki

Sierpień 2017

Oświadczenie kierującego pracą

Potwierdzam, że niniejsza praca została przygotowana pod moim kierunkiem i kwalifikuje się do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data | Podpis kierującego pracą

Oświadczenie autora (autorów) pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data | Podpis autora (autorów) pracy

Streszczenie

Niech $(X_{ij})_{ij}$ będą niezależnymi gaussowskimi zmiennymi losowymi o średniej zero, zaś $p, q \in [1, 2]$ liczbami rzeczywistymi. W pracy przedstawiono oszacowania wartości oczekiwanej normy operatorowej macierzy $X = (X_{ij})_{ij}$ traktowanej jako odwzorowania $\ell_p \to \ell_{q^*}$. Wyniki powstały przez uogólnienie niektórych wyników z ostatnich kilkunastu lat dla macierzy X traktowanej jako odwzorowanie $\ell_2 \to \ell_2$.

Słowa kluczowe

macierz losowa, norma operatorowa, koncentracja gaussowska, suprema procesów stochastycznych.

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.1 Matematyka

Klasyfikacja tematyczna

60 Probability theory and stochastic processes
60G Stochastic processes

Tytuł pracy w języku angielskim

Estimation of norms of random matrices

Spis treści

Rozdział 1

Wstęp

Będziemy rozważać problem oszacowania z góry wartości oczekiwanej normy operatorowej macierzy losowej o współczynnikach gaussowskich traktowanej jako odwzorowanie z przestrzeni ℓ_p w przestrzeń $\ell_{q^\star}$, gdzie $1 \leq p \leq 2 \leq q^\star \leq \infty$.

1.1. Notacja

Wprowadzimy teraz kilka oznaczeń, które będą obowiązywały w całej pracy. Jeśli $p > 1$ jest liczbą rzeczywistą, to pisząc $p^\star$ mamy na myśli liczbę spełniającą $\frac{1}{p} + \frac{1}{p^\star} = 1$.

Przez B_p i B_p^n będziemy oznaczać kule jednostkowe odpowiednio w ℓ_p i ℓ_p^n.

Pisząc $(g_i)_i$ bądź $(g_{ij})_{ij}$ zawsze będziemy mieli na myśli rodzinę niezależnych zmiennych losowych o rozkładzie $\mathcal{N}(0,1)$, zaś X zawsze będzie oznaczało macierz losową $(X_{ij})_{ij}$ gdzie $X_{ij} = a_{ij} g_{ij}$ dla pewnego układu liczb nieujemnych $(a_{ij})_{ij}$.

Literą K będziemy oznaczać stałe uniwersalne, których wartości mogą zmieniać się przy każdym pojawieniu.

1.2. Omówienie zawartości pracy

Do uzyskania oszacowań wykorzystamy metody z [7] i [5], które w cytowanych pracach posłużyły do uzyskania oszacowań normy operatorowej macierzy losowych traktowanych jako odwzorowania $\ell_2 \to \ell_2$. W twierdzeniu 1 uzyskamy wynik analogiczny do [7, twierdzenie 4.1] pozwalający na oszacowanie normy operatorowej macierzy $\|X\|_{\ell_p \to \ell_{p^\star}}$.

W podobny sposób jak z [7, twierdzenie 4.1] uzyskuje się [7, wniosek 4.2] uzyskamy wniosek 1 analogiczny do [7, wniosek 4.2]. Umożliwi on uzyskanie oszacowania $\|X\|_{\ell_p \to \ell_{p^\star}}$ w terminach $\max_i (\sum_j |a_{ij}|^{p^\star})^{\frac{1}{p^\star}}$, $\max_j \left(\sum_i |a_{ij}|^{p^\star}\right)^{\frac{1}{p^\star}}$ oraz $\max_{ij} |a_{ij}|$ ze stałą proporcjonalną do $\sqrt{\log(n+1)}$, gdzie n jest większym z wymiarów macierzy.

W dalszej części zawrzemy oszacowania w tych samych terminach dla macierzy blokowo-diagonalnych ze stałą proporcjonalną do pierwiastka z logarytmu z wymiaru największego bloku (twierdzenie 3) oraz dla macierzy spełniających $X_{ij} \equiv 0$ dla $|i-j| > k$, ze stałą proporcjonalną do $\sqrt{\log(k+1)}$ (wniosek 3).

W kolejnym rozdziale zastosujemy metody z [5], pozwalające na oszacowanie $\|X\|_{\ell_p \to \ell_{q^\star}}$ przez $K\log(n+1)\Big(\max_i (\sum_j |a_{ij}|^{p^\star})^{\frac{1}{p^\star}} + \max_j \big(\sum_i |a_{ij}|^{q^\star}\big)^{\frac{1}{q^\star}}\Big)$ gdzie K jest stałą uniwersalną (twierdzenie 4).

1.3. Znane wyniki

Problem oszacowania normy operatorowej macierzy traktowanych jako odwzorowania $\ell_2 \to \ell_2$ rozważany był w wielu pracach. W [2, twierdzenie 1.1] dla macierzy symetrycznych zostało oszacowanie:

$$\mathbb{E}\|X\|_{\ell_2\to\ell_2} \le K\Big(\max_i \sqrt{\sum_j a_{ij}^2} + \sqrt{\log n}\max_{ij}|a_{ij}|\Big),$$

gdzie n jest wymiarem macierzy. Oszacowanie to zostało uzyskane również w [7, twierdzenie 1.1]. W tej samej pracy, jako twierdzenie 3.1, udowodnione zostało, że:

$$\mathbb{E}\|X\|_{\ell_2\to\ell_2} \le K\sqrt{\log\log(n+1)}\Big(\max_i \sqrt{\sum_j a_{ij}^2} + \max_{ij}|a_{ij}^\star|\sqrt{\log i}\Big),$$

gdzie $a_{ij}^\star$ powstają przez permutację kolumn i wierszy macierzy $A = (a_{ij})$, tak aby $\max_j a_{1j}^\star \ge \max_j a_{2j}^\star \ge \ldots \ge \max_j a_{nj}^\star$. Z kolei [5, twierdzenie 2] głosi, że

$$\mathbb{E}\|X\|_{\ell_2\to\ell_2} \le K\Big(\max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + \big(\sum_{ij} a_{ij}^4\big)^{\frac{1}{4}}\Big).$$

Badaniem $\mathbb{E}\|X\|_{\ell_p\to\ell_{q^\star}}$ zajmowano się w [4]. Udowodnione tam, dla $X = (a_{ij}g_{ij})_{i\le n, j\le m}$ twierdzenie 1.1 mówi, że

$$\mathbb{E}\|X\|_{\ell_p\to\ell_{q^\star}} \le K(p^\star)^{\frac{5}{q^\star}}(\log n)^{\frac{1}{q^\star}}\Big(\gamma_{p^\star}\max_i \big(\sum_j a_{ij}^{p^\star}\big)^{\frac{1}{p^\star}} + \gamma_{q^\star}\mathbb{E}\max_{ij}|a_{ij}g_{ij}|\Big) + 2^{\frac{1}{q^\star}}\gamma_{q^\star}\max_j \big(\sum_i a_{ij}^{q^\star}\big)^{\frac{1}{q^\star}},$$

gdzie $\gamma_r = \mathbb{E}|g_{ij}|^r \approx \sqrt{r}$. Wynik ten, poza przypadkiem gdy $(\log n)^{1-\frac{1}{q^\star}}$ jest mniejszego rzędu niż $(p^\star)^{\frac{5}{q^\star}}$ i lepsze oszacowanie daje twierdzenie 4, jest silniejszy od wyników uzyskanych w niniejszej pracy. Jego uzyskanie wymaga jednak nieco bardziej skomplikowanych metod.

1.4. Hipotezy

W [4] postawiona została hipoteza

Hipoteza 1. *Dla $1 \le p \le 2 \le q^\star \le \infty$ zachodzi*

$$\mathbb{E}||(a_{ij}g_{ij})||_{\ell_p^m\to\ell_{q^\star}^n} \le K_{p,q}\Big(\max_{j\le m}\big(\sum_{i=1}^n |a_{ij}|^{q^\star}\big)^{\frac{1}{q^\star}} + \max_{i\le n}\big(\sum_{j=1}^m |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \mathbb{E}\max_{ij}|a_{ij}g_{ij}|\Big),$$

gdzie $K_{p,q}$ jest stałą zależącą wyłącznie od p i q.

Uwaga 1. Zachodzi szacowanie w drugą stronę

$$\mathbb{E}||(a_{ij}g_{ij})||_{\ell_p^m\to\ell_{q^\star}^n} \ge \frac{1}{K}\Big(\max_{j\le m}\big(\sum_{i=1}^n |a_{ij}|^{q^\star}\big)^{\frac{1}{q^\star}} + \max_{i\le n}\big(\sum_{j=1}^m |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \mathbb{E}\max_{ij}|a_{ij}g_{ij}|\Big),$$

gdzie K jest stałą uniwersalną.

Istotnie, wystarczy pokazać, że każdy ze składników po prawej stronie nierówności szacuje z dokładnością do stałej z dołu lewą stronę. Niech $X_{ij} = a_{ij}g_{ij}$. Zauważmy, że

$$\mathbb{E}\|(a_{ij}g_{ij})\|_{\ell_p^m \to \ell_{q^\star}^n} = \mathbb{E} \sup_{v\in B_q^n, u\in B_p^m} \Big| \sum_{ij} v_i a_{ij} g_{ij} u_j \Big| = \mathbb{E} \sup_{v\in B_q^n, u\in B_p^m} |\langle v, Xu\rangle|.$$

Ale to ostatnie wyrażenie szacuje się z dołu przez

$$\begin{aligned}\mathbb{E}\max_j \sup_{v\in B_q^n} |\langle v, Xe_j\rangle| &= \mathbb{E}\max_j \sup_{v\in B_q^n} \sum_i a_{ij}g_{ij}v_i = \mathbb{E}\max_j \Big(\sum_i |a_{ij}g_{ij}|^{q^\star}\Big)^{\frac{1}{q^\star}} \\ &\geq \max_j \mathbb{E}\Big(\sum_i |a_{ij}g_{ij}|^{q^\star}\Big)^{\frac{1}{q^\star}} \geq \max_j \Big(\mathbb{E}\sum_i |a_{ij}g_{ij}|^{q^\star}\Big)^{\frac{1}{q^\star}} \\ &\geq \sqrt{\frac{2}{\pi}} \max_{j\leq m} (\sum_{i=1}^n |a_{ij}|^{q^\star})^{\frac{1}{q^\star}},\end{aligned}$$

gdzie w przedostatnim przejściu skorzystaliśmy z nierówności Jensena a w ostatnim z nierówności:

$$\big(\mathbb{E}|g_{ij}|^{q^\star}\big)^{\frac{1}{q^\star}} \geq \mathbb{E}|g_{ij}| = \sqrt{\frac{2}{\pi}}.$$

Podobnie wykazujemy, że

$$\mathbb{E}\|(a_{ij}g_{ij})\|_{\ell_p^m \to \ell_{q^\star}^n} \geq \sqrt{\frac{2}{\pi}} \max_{i\leq n} (\sum_{j=1}^m |a_{ij}|^{p^\star})^{\frac{1}{p^\star}}.$$

Oczywiście zachodzi też nierówność

$$\mathbb{E} \sup_{v\in B_q^n, u\in B_p^m} |\langle v, Xu\rangle| \geq \mathbb{E}\max_{ij} |\langle e_i, Xe_j\rangle| = \mathbb{E}\max_{ij} |a_{ij}g_{ij}|.$$

Hipoteza 1 ma swoją deterministyczną wersję.

Hipoteza 2. *Dla $1 \leq p \leq 2 \leq q^\star \leq \infty$ zachodzi*

$$\mathbb{E}||(a_{ij}g_{ij})||_{\ell_p^m \to \ell_{q^\star}^n} \leq K_{p,q}\Big(\max_{j\leq m}(\sum_{i=1}^n |a_{ij}|^{q^\star})^{\frac{1}{q^\star}} + \max_{i\leq n}(\sum_{j=1}^m |a_{ij}|^{p^\star})^{\frac{1}{p^\star}} + \max_{ij} |a_{ij}^\star| \sqrt{\log(i+1)}\Big),$$

gdzie $a_{ij}^\star$ powstają przez permutację kolumn i wierszy macierzy $A = (a_{ij})$, tak aby $\max_j a_{1j}^\star \geq \max_j a_{2j}^\star \geq \ldots \geq \max_j a_{nj}^\star$ a $K_{p,q}$ jest stałą zależącą wyłącznie od p i q.

Udowodnimy równoważność hipotez 1 i 2. W tym celu wystarczy wykazać, że

$$\mathbb{E}\max_{ij} |a_{ij}g_{ij}| \leq K_{p,q}\Big(\max_{j\leq m}(\sum_{i=1}^n |a_{ij}|^{q^\star})^{\frac{1}{q^\star}} + \max_{i\leq n}(\sum_{j=1}^m |a_{ij}|^{p^\star})^{\frac{1}{p^\star}} + \max_{ij} |a_{ij}^\star| \sqrt{\log(i+1)}\Big) \quad (1.1)$$

oraz

$$\max_{ij} |a_{ij}^\star| \sqrt{\log(i+1)} \leq K\Big(\max_{j\leq m}(\sum_{i=1}^n |a_{ij}|^{q^\star})^{\frac{1}{q^\star}} + \max_{i\leq n}(\sum_{j=1}^m |a_{ij}|^{p^\star})^{\frac{1}{p^\star}} + \mathbb{E}\max_{ij} |a_{ij}g_{ij}|\Big) \quad (1.2)$$

dla pewnej stałej K. Oczywiście

$$\mathbb{E}\max_{ij} |a_{ij}g_{ij}| \leq \mathbb{E}\max_i \big(\sum_j |a_{ij}g_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}}.$$

Ale

$$\mathbb{E}\max_i(\sum_j |a_{ij}g_{ij}|^{p^\star})^{\frac{1}{p^\star}} \le K\Big(\max_i\Big(\sum_j a_{ij}^{p^\star}\Big)^{\frac{1}{p^\star}} + \max_i\sqrt{\log(i+1)}\max_j a_{ij}^\star\Big),$$

czego dowodzimy w dalszej części pracy - zob. (2.8). W drugą stronę korzystając z lematu 2.4 z [7] otrzymujemy

$$\max_{ij}|a_{ij}^\star|\sqrt{\log(i+1)} \le K\mathbb{E}\max_{ij}|a_{ij}g_{ij}|,$$

co kończy dowód.

1.5. Redukcja problemu

Uwaga 2. Jeśli $p = q$, to w oszacowaniach wartości oczekiwanej normy operatorowej przez wyrażenia postaci $\max_{j\le m}(\sum_{i=1}^n |a_{ij}|^{p^\star})^{\frac{1}{p^\star}}$, $\max_{i\le n}(\sum_{j=1}^m |a_{ij}|^{p^\star})^{\frac{1}{p^\star}}$, $\mathbb{E}\max_j\Big(\sum_i |a_{ij}g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}}$, $\mathbb{E}\max_i\Big(\sum_j |a_{ij}g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}}$, $\mathbb{E}\max_{ij}|a_{ij}g_{ij}|$ lub $\max_{ij}|a_{ij}|$ z dokładnością do stałej K zależnej być może od p, q i co najwyżej wielomianowo od wymiaru macierzy, wystarczy ograniczyć się do macierzy symetrycznych.

Istotnie, macierze X i

$$\begin{pmatrix} 0 & X^T \\ X & 0 \end{pmatrix}$$

mają tę samą normę operatorową (por. [2]). Lewa strona nierówności z hipotezy pozostaje niezmieniona a stała po prawej stronie jest $O(K)$.

Kolejną możliwość redukcji daje następująca obserwacja.

Uwaga 3. Dla $p = q$ i macierzy symetrycznych wystarczy ograniczyć się do badania $\sup_{v\in B_p}|\langle v, Xv\rangle|$.

Zachodzi bowiem

$$||X||_{\ell_p\to\ell_{p^\star}} = \sup_{u,v\in B_p}\langle v, Xu\rangle.$$

Ponadto, jeżeli X jest macierzą symetryczną, to

$$\begin{aligned}\langle v, Xu\rangle &= \frac{1}{4}[\langle u+v, X(u+v)\rangle - \langle u-v, X(u-v)\rangle] \\ &\le \frac{1}{4}\Big(||x+y||_p^2 + ||x-y||_p^2\Big)\sup_{z\colon ||z||_p\le 1}|\langle z, Xz\rangle| \le 2\sup_{z\colon ||z||_p\le 1}|\langle z, Xz\rangle|.\end{aligned}$$

Rozdział 2

Metody z [7]

W tej części zajmiemy się oszacowaniami w przypadku macierzy traktowanej jako odwzorowanie $\ell_p \to \ell_{p^\star}$. Udowodnimy twierdzenie, będące przeniesieniem wyników z twierdzenia 4.1 z pracy [7] na przypadek odwzorowań z ℓ_p w $\ell_{p^\star}$ oraz kilka wypływających z niego wniosków. Zanim jednak to nastąpi wprowadzimy kilka oznaczeń. Niech $B = (a_{ij}^2)_{i,j\le n}$ będzie macierzą wariancji współczynników macierzy X i niech w przypadku gdy B jest symetryczna

$$B = \sum_i \lambda_i \xi_i \xi_i^T$$

będzie jej rozkładem spektralnym. Będziemy oznaczać $B^+ = \sum_i (\lambda_i \vee 0)\xi_i\xi_i^T$ oraz $B^- = -\sum_i (\lambda_i \wedge 0)\xi_i\xi_i^T$.

2.1. Główne twierdzenie

Naszym głównym narzędziem w tym rozdziale, będzie następujące twierdzenie.

Twierdzenie 1. *Niech $1 < p \le 2$ i niech $X = (a_{ij}g_{ij})_{i,j\le n}$, będzie macierzą symetryczną, gdzie $g_{ij} \sim \mathcal{N}(0,1)$ są niezależne dla $i \ge j$. Niech $Y \sim \mathcal{N}(0, B^-)$. Wtedy*

$$\mathbb{E}||X||_{\ell_p\to\ell_{p^\star}} \le 8\Big(\mathbb{E}\max_i \big(\sum_j |a_{ij}g_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \mathbb{E}\max_i Y_i\Big).$$

Dowód: Rozważmy proces $\langle x(u,v), g\rangle :\, = \sum_i v_i ||u||_i g_i$, gdzie $g = (g_1, \ldots, g_n)$, g_i są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(0,1)$ a $||u||_i^2 = \sum_j a_{ij}^2 u_j^2$. Oczywiście

$$\mathbb{E}|\langle x(u,v), g\rangle|^2 = \sum_i v_i^2 ||u||_i^2 \mathbb{E} g_i^2 = \sum_{ij} v_i^2 a_{ij}^2 u_j^2.$$

Zachodzi nierówność

$$\sup_{u,v\in B_p} |\langle x(u,v), g\rangle| \le \max_j \Big(\sum_i |a_{ij}|^{p^\star} |g_i|^{p^\star}\Big)^{\frac{1}{p^\star}} \tag{2.1}$$

Istotnie, z nierówności Höldera, dla $u, v \in B_p$ zachodzi

$$|\langle x(u,v), g\rangle| = \big|\sum_i v_i ||u||_i g_i\big| \le \Big(\sum_i |v_i|^p\Big)^{\frac{1}{p}} \Big(\sum_i ||u||_i^{p^\star} |g_i|^{p^\star}\Big)^{\frac{1}{p^\star}} \le \Big(\sum_i ||u||_i^{p^\star} |g_i|^{p^\star}\Big)^{\frac{1}{p^\star}}$$

$$= \Big(\sum_i \big(\sum_j a_{ij}^2 u_j^2\big)^{\frac{p^\star}{2}} |g_i|^{p^\star}\Big)^{\frac{1}{p^\star}} \le \max_j \Big(\sum_i |a_{ij} g_i|^{p^\star}\Big)^{\frac{1}{p^\star}},$$

gdzie, ostatnia nierówność wynika z faktu, że $\sum_j u_j^2 \leq 1$ a $\sum_i \big(\sum_j a_{ij}^2 u_j^2\big)^{\frac{p^\star}{2}} |g_i|^{p^\star}$ jest wypukłą funkcją $(u_j^2)_j$ zatem swoje maksimum osiąga w punktach ekstremalnych. Lemat 4.6 z [7], zastosowany z $\gamma = 1$ daje

$$\begin{aligned}\mathbb{E}|\langle v, Xv\rangle - \langle w, Xw\rangle|^2 \\ \leq 4\mathbb{E}|\langle x(v,v), g\rangle - \langle x(w,w), g\rangle|^2 - \sum_{ij}(v_i^2 - w_i^2)a_{ij}^2(v_j^2 - w_j^2).\end{aligned} \tag{2.2}$$

Oszacujemy teraz normę operatorową symetrycznej macierzy losowej X, postępując jak w dowodzie [7, twierdzenie 4.1]. Wobec opisanych w 1.5 zależności jest ona co do stałej 2 ograniczona przez

$$\mathbb{E} \sup_{v\in B_p} |\langle v, Xv\rangle| \leq \mathbb{E}\Big(\sup_{v\in B_p}\langle v, Xv\rangle \vee \sup_{v\in B_p} -\langle v, Xv\rangle\Big).$$

Jako, że zmienne losowe $\sup_{v\in B_p}\langle v, Xv\rangle$ i $\sup_{v\in B_p} -\langle v, Xv\rangle$ są nieujemne a macierz X ma ten sam rozkład co $-X$, prawa strona nierówności szacuje się przez

$$\mathbb{E}\Big(\sup_{v\in B_p}\langle v, Xv\rangle + \sup_{v\in B_p} -\langle v, Xv\rangle\Big) \leq 2\mathbb{E}\sup_{v\in B_p}\langle v, Xv\rangle.$$

Określmy proces $(Z_v)_{v\in B_p}$ wzorem

$$Z_v = 2\langle x(v,v), g\rangle + \sum_i v_i^2 Y_i, \tag{2.3}$$

gdzie $Y \sim \mathcal{N}(0, B^-)$ jest niezależne od g. Wtedy wobec (2.2)

$$\begin{aligned}\mathbb{E}|Z_v - Z_w|^2 &= 4||x(v) - x(w)||^2 + \sum_{ij}(v_i - w_i)^2 B_{ij}^-(v_j^2 - w_j^2) \\ &\geq \mathbb{E}|\langle v, Xv\rangle - \langle w, Xw\rangle|^2.\end{aligned}$$

Zatem z lematu Slepiana [1, twierdzenie 13.3]

$$\mathbb{E}\sup_{v\in B_p}\langle v, Xv\rangle \leq \mathbb{E}\sup_{v\in B_p} Z_v. \tag{2.4}$$

Zastosowanie (2.1) daje tezę. □

2.2. Wnioski z twierdzenia 1

Wyprowadzimy teraz kilka oszacowań $\mathbb{E}\|X\|$ w terminach zależnych od $(a_{ij})_{ij}$ będących konsekwencjami udowodnionego w poprzedniej sekcji twierdzenia.

Wniosek 1. *Przy założeniach jak w twierdzeniu 1,*

$$\mathbb{E}||X||_{\ell_p\to\ell_{p^\star}} \leq K\sqrt{p^\star}\Big(\max_i\big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_i\big(\sum_j |a_{ij}|^4\big)^{\frac{1}{4}}\sqrt{\log(i+1)}\Big).$$

Jeżeli nie zakładamy, że X jest macierzą symetryczną, to

$$\begin{aligned}&\mathbb{E}||X||_{\ell_p\to\ell_{p^\star}} \\ &\leq K\sqrt{p^\star}\Big(\max_i\big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_j\big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_i\big(\sum_j |a_{ij}|^4\big)^{\frac{1}{4}}\sqrt{\log(i+1)} \\ &\qquad + \max_j\big(\sum_i |a_{ij}|^4\big)^{\frac{1}{4}}\sqrt{\log(j+1)}\Big).\end{aligned}$$

Dowód: Zauważmy, że $B^2 = (B^+)^2 + (B^-)^2$, a więc

$$\begin{aligned}\sum_j a_{ij}^4 = \sum_j B_{ij}^2 = \langle Be_i, Be_i\rangle = \langle B^2 e_i, e_i\rangle = \langle ((B^+)^2 + (B^-)^2)e_i, e_i\rangle \\ = \langle B^+ e_i, B^+ e_i\rangle + \langle B^- e_i, B^- e_i\rangle = \sum_j (B_{ij}^+)^2 + \sum_j (B_{ij}^-)^2.\end{aligned}$$

Co za tym idzie

$$\operatorname{Var} Y_i = B_{ii}^- \leq \sqrt{\sum_j a_{ij}^4} \tag{2.5}$$

dla dowolnego i. Skorzystamy teraz ze znanego oszacowania maksimum subgaussowskich zmiennych losowych (por. [8, Prop. 2.4.16]).

Lemat 1. *Niech $X_1, \ldots X_N$ będą zmiennymi losowymi spełniającymi dla $x \geq 0$*

$$\mathbb{P}(X_i > x) \leq \exp(-\frac{x^2}{L^2\sigma_i^2}).$$

Wtedy

$$\mathbb{E} \max_{i \leq n} X_i \leq KL \max_{i \leq n} \sigma_i^\star \sqrt{\log(i+1)},$$

gdzie $\sigma_i^\star$ jest nierosnącym uporządkowaniem σ_i.

Korzystając z tego lematu i (2.5) otrzymujemy

$$\mathbb{E} \max_i Y_i \leq K \max_i \Big(\sum_j a_{ij}^4\Big)^{\frac{1}{4}} \sqrt{\log(i+1)}.$$

Wystarczy więc oszacować wielkość $\mathbb{E}\max_i \Big(\sum_j |a_{ij} g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}}$. Oczywiście:

$$\begin{aligned} &\mathbb{E}\max_i \Big(\sum_j a_{ij}^{p^\star} |g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} \\ &\leq \max_i \mathbb{E}\Big(\sum_j |a_{ij} g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} + \mathbb{E}\max_i \Bigg(\Big(\sum_j |a_{ij} g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} - \mathbb{E}\Big(\sum_j |a_{ij} g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}}\Bigg). \end{aligned} \tag{2.6}$$

Jako, że funkcja $x \mapsto x^{\frac{1}{\alpha}}$ jest wklęsła dla $x \geq 0$, to z nierówności Jensena

$$\mathbb{E}\Big(\sum_j |a_{ij} g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} \leq \Big(\mathbb{E}\sum_j |a_{ij} g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} \leq K\sqrt{p^\star}\Big(\sum_j a_{ij}^{p^\star}\Big)^{\frac{1}{p^\star}} \tag{2.7}$$

gdyż $\mathbb{E}(|g_{ij}|^{p^\star})^{\frac{1}{p^\star}} \leq K\sqrt{p^\star}$. Jeżeli chodzi o drugi składnik oszacowania to rozważmy funkcję

$$f(x) = \Big(\sum_j |a_{ij} x_j|^{p^\star}\Big)^{\frac{1}{p^\star}}, \quad x \in \mathbb{R}^n.$$

Jest ona lipschitzowska ze stałą $\max_j |a_{ij}|$, gdyż $|f(x+y) - f(x)| \leq |f(y)|$ z nierówności trójkąta w $\ell_{p^\star}$ oraz

$$\sup_{y \in B_2} |f(y)| = \sup_{y \in B_2} (\sum_j |a_{ij} y_j|^{p^\star})^{\frac{1}{p^\star}} \leq \sup_{y \in B_{p^\star}} (\sum_j |a_{ij} y_j|^{p^\star})^{\frac{1}{p^\star}} \leq \max_j |a_{ij}|.$$

Z nierówności koncentracyjnej dla funkcji lipschitzowskich [6, twierdzenia 5.1 i 5.3]:

$$\mathbb{P}\Big(\big(\sum_j |a_{ij}g_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} - \mathbb{E}\big(\sum_j |a_{ij}g_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} > u\Big) \le \exp(-u^2/2\max_j a_{ij}^2).$$

Korzystając ponownie z lematu 1, oraz nierówności (2.6) i (2.7) otrzymujemy

$$\mathbb{E}\max_i \Big(\sum_j |a_{ij}g_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} \le K\sqrt{p^\star}\Big(\max_i \Big(\sum_j |a_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} + \max_i \sqrt{\log(i+1)}\max_j |a_{ij}|\Big) \quad (2.8)$$

Łącząc oszacowania dla składników procesu Z_v otrzymujemy

$$\mathbb{E}||X|| \le K\sqrt{p^\star}\Big(\max_i \big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_i \sqrt{\log(i+1)}\Big(\sum_j a_{ij}^4\Big)^{\frac{1}{4}}\Big),$$

gdyż

$$\max_i \sqrt{\log(i+1)}\max_j |a_{ij}| \le \max_i \sqrt{\log(i+1)}\Big(\sum_j a_{ij}^4\Big)^{\frac{1}{4}}.$$

Przypadek niesymetryczny jest wnioskiem z uwagi 1. $\square$

Uwaga 4. Rozumując jak w dowodzie równoważności hipotez 1.1 i 1.2 można dowieść przeciwnego oszacowania ze stałą $\frac{1}{K}$ zamiast $K\sqrt{p^\star}$.

Wniosek 2. *Niech* $\frac{4}{3} \le p \le 2$ *i* $X = (a_{ij}g_{ij})_{i,j\le n}$*, gdzie* $g_{ij} \sim \mathcal{N}(0,1)$ *są niezależne. Wtedy*

$$\mathbb{E}||X||_{\ell_p\to\ell_{p^\star}} \le K\sqrt{\log(n+1)}\Big(\max_i \big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_j \big(\sum_i |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_{ij}|a_{ij}|\Big)$$

dla pewnej stałej uniwersalnej K*.*

Dowód: Założenie, że $\frac{4}{3} \le p \le 2$ jest równoważne $2 \le p^\star \le 4$. Dla $p^\star \le 4$ mamy

$$\begin{aligned}\big(\sum_j |a_{ij}|^4\big)^{\frac{1}{4}} &= \big(\sum_j |a_{ij}|^{p^\star}|a_{ij}|^{4-p^\star}\big)^{\frac{1}{4}} \le \max_i |a_{ij}|^{\frac{4-p^\star}{4}}\big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{4}} \\ &\le \frac{4-p^\star}{4}\max_i |a_{ij}| + \frac{p^\star}{4}\big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}}\end{aligned}$$

Podstawiając powyższe oszacowanie do tezy wniosku 1 otrzymujemy tezę. $\square$

Jeżeli macierz X traktowana jest jako odwzorowanie z ℓ_1 w $\ell_{p^\star}$, gdzie $p \le 2$, to prawdziwy jest następujący wynik.

Twierdzenie 2. *Niech* $1 \le p \le 2$ *i niech* $X = (a_{ij}g_{ij})_{i,j\le n}$*, gdzie* $g_{ij} \sim \mathcal{N}(0,1)$ *są niezależne. Wówczas:*

$$\mathbb{E}||X||_{\ell_1\to\ell_{p^\star}} \le K\sqrt{p^\star}\Big(\max_i \Big(\sum_j |a_{ij}|^{p^\star}\Big)^{\frac{1}{p^\star}} + \max_i \sqrt{\log(i+1)}\max_j |a_{ij}|\Big).$$

Dowód: Wystarczy zauważyć, że

$$||X||_{\ell_1 \to \ell_{p^\star}} = \max_i \Big(\sum_j |a_{ij} g_{ij}|^{p^\star} \Big)^{\frac{1}{p^\star}}$$

i zastosować oszacowanie (2.8) z dowodu wniosku 1. □

Niech $\mathcal{M}_{n\times m}$ będzie zbiorem rzeczywistych macierzy wymiaru $n \times m$

Twierdzenie 3. *Niech A będzie macierzą blokowo-diagonalną tzn. istnieją macierze $A_k \in \mathcal{M}(n_k \times n_k)$, $k = 1, 2, \ldots m$ oraz $a_{(s_k+i)(s_k+j)} = (A_k)_{ij}$, gdzie $s_k = n_1 + n_2 + \ldots n_{k-1}$ i $a_{ij} = 0$ w przeciwnym przypadku. Wtedy dla $\frac{4}{3} \leq p \leq 2$ i macierzy X z $X_{ij} = a_{ij} g_{ij}$*

$$\mathbb{E}||X||_{\ell_p \to \ell_{p^\star}} \leq K \sqrt{\log(\max_i n_i + 1)} \Big(\max_i \Big(\sum_j |a_{ij}|^{p^\star} \Big)^{\frac{1}{p^\star}} + \max_j \Big(\sum_i |a_{ij}|^{p^\star} \Big)^{\frac{1}{p^\star}} + \mathbb{E} \max_{ij} |a_{ij} g_{ij}| \Big).$$

W dowodzie skorzystamy z następującego lematu.

Lemat 2. *[6, twierdzenie 7.1] Niech $(X(t))_{t\in T}$ będzie scentrowanym, ośrodkowym procesem gaussowskim spełniającym $\mathbb{P}(\sup_t |X(t)| < \infty) > 0$ i niech $\sigma^2 = \sup_t \mathbb{E}X^2(t)$. Wtedy dla dowolnego $u > 0$*

$$\mathbb{P}\big(\sup_t |X(t)| \geq \mathbb{E} \sup_t |X(t)| + u \big) \leq e^{-\frac{u^2}{2\sigma^2}}.$$

Dowód twierdzenia 3: Celem poprawy czytelności pisząc symbol normy będziemy mieli na myśli normę macierzy jako odwzorowania z ℓ_p w $\ell_{p^\star}$. Określimy $a_{ij}^{(k)} := a_{(s_k+i)(s_k+j)}$ oraz $g_{ij}^{(k)} := g_{(s_k+i)(s_k+j)}$. Niech $X_k = (a_{ij}^{(k)} g_{ij}^{(k)})_{i,j\leq n_k}$. Zauważmy, że skoro $p \leq 2$, to $\|X\| = \max_k \|X_k\|$. Istotnie

$$\|X\| = \sup_{u,v\in B_p} \sum_{ij} v_i a_{ij} u_j = \sup_{\lambda,\mu \in B_p} \sum_k \lambda_k \mu_k \sup_{u,v\in B_p^{n_k}} \sum_{ij} v_i a_{ij}^{(k)} g_{ij}^{(k)} u_j = \max_k \|X_k\|,$$

gdyż $(\lambda_k \mu_k)_k \in \ell_1$. Możemy też założyć, że wartość $\max_{ij} |a_{ij}^{(k)}|$ maleje wraz z k. Niech $(Y_k(u,v))_{u,v\in B_p^{n_k}}$ będzie procesem określonym wzorem

$$Y_k(u,v) = \sum_{i,j} v_i a_{ij}^{(k)} g_{ij}^{(k)} u_j.$$

Wtedy Y_k jest scentrowanym, ośrodkowym procesem Gaussowskim oraz

$$\|X_k\|_{\ell_p \to \ell_{p^\star}} = \sup_{u,v\in B_p^{n_k}} Y_k(u,v).$$

Ponadto z poprzednich rozważań wynika, że $\mathbb{E}\|X_k\| < \infty$ a zatem $\sup_{u,v} |Y_k(u,v)| < \infty$ p.n. Ponadto

$$\mathbb{E}Y_k(u,v)^2 = \sum_{i,j} v_i^2 (a_{ij}^{(k)})^2 u_j^2 = \sum_i v_i^2 \sum_j (a_{ij}^{(k)})^2 u_j^2.$$

Zauważmy, że $(u_1^2, \ldots, u_{n_k}^2) \in B_1^{n_k}$, oraz $(v_1^2, \ldots, v_{n_k}^2) \in B_1^{n_k}$. Zatem

$$\begin{aligned} \sup_{u\in B_p^{n_k}} \sup_{v\in B_p^{n_k}} \sum_i v_i^2 \sum_j (a_{ij}^{(k)})^2 u_j^2 &\leq \sup_{u\in B_1^{n_k}} \sup_{v\in B_1^{n_k}} \sum_i v_i^2 \sum_j (a_{ij}^{(k)})^2 u_j^2 \\ &= \sup_{u\in B_1^{n_k}} \max_i \sum_j (a_{ij}^{(k)})^2 u_j^2 = \max_{ij} (a_{ij}^{(k)})^2, \end{aligned}$$

gdzie w ostatnich dwóch równościach korzystaliśmy odpowiednio z tego, że funkcja liniowa i maksimum funkcji liniowych są funkcjami wypukłymi a te swoje suprema osiągają w punktach ekstremalnych.

Niech $M := \max_k \mathbb{E}\|X_k\|$. Wtedy na mocy lematu 2

$$\forall_k \quad \mathbb{P}(\|X_k\| > M + u) \leq e^{-\frac{u^2}{2\sigma_k^2}}, \tag{2.9}$$

gdzie $\sigma_k^2 = \sup_{u,v} \mathbb{E}Y_k(u,v)^2$. Zauważmy teraz, że

$$\mathbb{E}\|X\| = \mathbb{E}\max_k \|X_k\| = M + \mathbb{E}(\|X\| - M) = M + \mathbb{E}\max_k(\|X_k\| - M).$$

Na mocy (2.9) i lematu 1 zachodzi więc

$$\mathbb{E}\|X\| \leq \max_k \mathbb{E}\|X_k\| + K \max_k \sqrt{\log(k+1)} \max_{i,j} |a_{ij}^{(k)}|.$$

Aby zakończyć dowód wystarczy pokazać, że dla pewnej stałej K

$$\max_k \sqrt{\log(k+1)} \max_{i,j} |a_{ij}^{(k)}| \leq K\mathbb{E}\max_{ij} |a_{ij}g_{ij}|.$$

Teza wyniknie wówczas z wniosku 2. Potrzebny fakt wynika z nierówności

$$\mathbb{E}\max_{ij} |a_{ij}g_{ij}| \geq \mathbb{E}\max_k \left|(\max_{ij} |a_{ij}^{(k)}|)g_k\right|$$

oraz poniższego lematu 3 □

Lemat 3. *[7, lemat 2.4] Niech $X_i \sim \mathcal{N}(0, \sigma_i^2)$, $i = 1, 2, \ldots n$ będą niezależnymi zmiennymi losowymi i niech σ_k^2 są uporządkowane nierosnąco. Wówczas*

$$\mathbb{E}\max_i |X_i| \geq K \max_i \sigma_i \sqrt{\log(i+1)}.$$

Wniosek 3. *Niech A będzie macierzą kwadratową spełniającą $a_{ij} = 0$ dla $|i - j| > k$ i niech $X_{ij} = a_{ij}g_{ij}$, gdzie g_{ij} są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(0, 1)$. Wówczas dla $\frac{4}{3} \leq p \leq 2$,*

$$\mathbb{E}||X||_{\ell_p \to \ell_{p^\star}} \leq K\sqrt{\log(k+1)}\Big(\max_i \big(\sum_j |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \max_j \big(\sum_i |a_{ij}|^{p^\star}\big)^{\frac{1}{p^\star}} + \mathbb{E}\max_{ij} |a_{ij}g_{ij}|\Big).$$

Dowód: Bez utraty ogólności można założyć, że $k < n$. Niech macierz $\tilde{A}$ będzie rozszerzeniem macierzy A przez zera, tak aby jej wymiar był podzielny przez k. Wystarczy wykazać tezę dla $\tilde{A}$. Określmy A_1 i A_2 następującymi wzorami:

$$(A_1)_{ij} = \tilde{A}_{ij}\mathbf{1}(2ck < i, j \leq 2(c+1)k \text{ dla pewnego } c \in \mathbb{N}),$$

$$(A_2)_{ij} = \tilde{A}_{ij}\mathbf{1}((2c+1)k < i, j \leq (2c+3)k \text{ dla pewnego } c \in \mathbb{N} \text{ i } (A_1)_{ij} = 0).$$

Wówczas $\tilde{A} = A_1 + A_2$ a zatem z nierówności trójkąta $\|A\| = \|\tilde{A}\| \le \|A_1\| + \|A_2\|$. Twierdzenie 3 daje:

$$\begin{aligned}
&\mathbb{E}\|\big((A_1)_{ij} g_{ij}\big)\|_{\ell_p \to \ell_{p^\star}} \\
&\quad \le K\sqrt{\log(k+1)}\Big(\max_i \big(\sum_j |a_{ij}^{p^\star}|\mathbf{1}((A_1)_{ij} \neq 0)\big)^{\frac{1}{p^\star}} + \max_j \big(\sum_i |a_{ij}^{p^\star}|\mathbf{1}((A_1)_{ij} \neq 0)\big)^{\frac{1}{p^\star}} \\
&\qquad\qquad + \mathbb{E}\max_{ij} |a_{ij}\mathbf{1}((A_1)_{ij} \neq 0) g_{ij}|\Big) \\
&\quad \le K\sqrt{\log(k+1)}\Big(\max_i (\sum_j |a_{ij}^{p^\star}|)^{\frac{1}{p^\star}} + \max_j (\sum_i |a_{ij}^{p^\star}|)^{\frac{1}{p^\star}} + \mathbb{E}\max_{ij} |a_{ij} g_{ij}|\Big)
\end{aligned}$$

oraz

$$\begin{aligned}
&\mathbb{E}\|\big((A_2)_{ij} g_{ij}\big)\|_{\ell_p \to \ell_{p^\star}} \\
&\quad \le K\sqrt{\log(k+1)}\Big(\max_i \big(\sum_j |a_{ij}^{p^\star}|\mathbf{1}((A_2)_{ij} \neq 0)\big)^{\frac{1}{p^\star}} + \max_j \big(\sum_i |a_{ij}^{p^\star}|\mathbf{1}((A_2)_{ij} \neq 0)\big)^{\frac{1}{p^\star}} \\
&\qquad\qquad + \mathbb{E}\max_{ij} |a_{ij}\mathbf{1}((A_2)_{ij} \neq 0) g_{ij}|\Big) \\
&\quad \le K\sqrt{\log(k+1)}\Big(\max_i (\sum_j |a_{ij}^{p^\star}|)^{\frac{1}{p^\star}} + \max_j (\sum_i |a_{ij}^{p^\star}|)^{\frac{1}{p^\star}} + \mathbb{E}\max_{ij} |a_{ij} g_{ij}|\Big).
\end{aligned}$$

Łącząc powyższe nierówności otrzymujemy tezę. $\square$

Rozdział 3

Zastosowanie metod z [5]

Metody zastosowane w [5] pozwalają na wyprowadzenie oszacowań dla macierzy X traktowanej jako odwzorowanie $\ell_p \to \ell_{q^\star}$, gdzie q może być różne od p. Jako, że dowód jest nieco techniczny, rozbijemy go na kilka pomniejszych faktów.

Lemat 4. *Niech $X = (X_{ij})_{i,j\leq n}$ będzie jak w sekcji 1.1, wówczas dla dowolnych $u, v \in \mathbb{R}^n$,*

$$\mathbb{P}(\langle v, Xu\rangle > t) \leq \exp(-\frac{t^2}{2\sigma(u,v)^2}),$$

gdzie $\sigma(u,v) = \sqrt{\sum_{i,j} v_i^2 a_{ij}^2 u_j^2}$.

Dowód: Wynika to z koncentracji zmiennych gaussowskich oraz tego, że zmienna $\langle v, Xu\rangle$ ma rozkład $\mathcal{N}\big(0, \sigma^2(u,v)\big)$. $\square$

Lemat 5 (Przeformułowanie lematu 2 z [5])**.** *Niech $(Y_x)_{x\in U}$ - zmienne losowe, oraz nieujemne liczby $(b_x)_{x\in U}$ spełniają dla pewnego $\alpha \geq 1$ nierówność:*

$$\mathbb{P}(Y_x \geq tb_x) \leq e^{-t^\alpha}.$$

Wtedy dla dowolnego $u \geq 0$,

$$\mathbb{E}\max_{x\in U}[Y_x - b_x(u + (\log \#U)^{\frac{1}{\alpha}})] \leq K_\alpha e^{-u^\alpha} \max_{x\in U} b_x,$$

gdzie $K_\alpha = \Gamma(\frac{\alpha+1}{\alpha}) \leq 1$. W szczególności

$$\mathbb{E}\max_{x\in U}[Y_x - b_x(\log \#U)^{\frac{1}{\alpha}}] \leq \max_{x\in U} b_x.$$

Dowód: Niech $Z_x = Y_x - b_x(u + (\log \#U)^{\frac{1}{\alpha}})$. Wtedy

$$\begin{aligned}
\mathbb{E}\max_{x\in U} Z_x &\leq \int_0^\infty \mathbb{P}\big(\max_{x\in U} Z_x \geq s\big)ds \leq \sum_{x\in U}\int_0^\infty \mathbb{P}\big(Y_x \geq b_x(u + (\log \#U)^{\frac{1}{\alpha}} + \frac{s}{b_x})\big) \\
&\leq \sum_{x\in U}\int_0^\infty \exp\big(-(u + (\log \#U)^{\frac{1}{\alpha}} + \frac{s}{b_x})^\alpha\big)ds \\
&\leq \sum_{x\in U}\int_0^\infty \exp\big(-u^\alpha - \log \#U - (\frac{s}{b_x})^\alpha\big)ds \\
&\leq e^{-u^\alpha}\frac{1}{\#U}\sum_{x\in U}\int_0^\infty e^{-(\frac{s}{b_x})^\alpha} ds \leq e^{-u^\alpha} K_\alpha \max_{x\in U} b_x.
\end{aligned}$$

Pozostaje zauważyć, że

$$K_\alpha = \int_0^\infty e^{-t^\alpha} dt = \frac{1}{\alpha}\int_0^\infty s^{\frac{1}{\alpha}-1} e^{-s} ds = \frac{1}{\alpha}\Gamma(\frac{1}{\alpha}) = \Gamma(1+\frac{1}{\alpha}).$$

Oczywiście $\Gamma(1) = \Gamma(2) = 1$ a $\ln\Gamma$ jest funkcją wypukłą (zob. [3, ustęp 533]), więc $K_\alpha \leq 1$. □

Lemat 6. *Niech U i V będą skończonymi zbiorami zaś $I(u,v)$ podzbiorem $\mathbb{N}\times\mathbb{N}$. Wówczas:*

$$\mathbb{E}\max_{u\in U, v\in V}\Big[\sum_{i,j\in I(u,v)} v_i a_{ij} g_{ij} u_j - \sqrt{2}\sqrt{\sum_{ij} v_i^2 a_{ij}^2 u_j^2}\sqrt{\log(\#U\#V)}\Big] \leq \sqrt{2}\max_{u\in U, v\in V}\sqrt{\sum_{ij} v_i^2 a_{ij}^2 u_j^2}$$

Dowód: Dowód polega na zastosowaniu lematów 4 i 5 z $\alpha = 2$ i $b_{u,v} = \sqrt{2}\sigma(u,v)$. □

Niech dla $1 < r \leq 2$:

$$D_r^n\colon = \{u \in B_r^{2^n} \quad | \quad \forall i \quad u_i = 0 \text{ lub } \exists k \in \mathbb{Z}\colon |u^i|^r = 2^{-k}\}.$$

Dla $u \in D_r^n$ i $k \in \mathbb{N}$ połóżmy

$$A_{k,n}^r(u)\colon = \{i\colon u \in D_r^n\colon |u_i|^r \geq 2^{-k}\}.$$

Określmy także zbiór

$$\Pi_{k,n}^r = \{(u_i \mathbf{1}(i \in A_{k,n}^r(u))\colon u \in D_r^n\}.$$

Niech w końcu

$$D_r = \{u \in B_r \quad | \quad \forall i \quad u_i = 0 \text{ lub } \exists k \in \mathbb{Z}\colon |u^i|^r = 2^{-k}\}.$$

Zauważmy, że zbiór D_r^n może być wówczas traktowany jako podzbiór zbioru D_r. Jeżeli wartość r będzie można wywnioskować z kontekstu, bądź w danym momencie nie będzie ona istotna, to indeks r będziemy pomijać i zamiast $A_{k,n}^r$ będziemy pisać A_k^n zaś zamiast $\Pi_{k,n}^r$ - Π_k^n.

Zbiory $\Pi_{k,n}^p$ i $\Pi_{k,n}^q$ są równoliczne dla dowolnych p i q. Zatem w wypadku, gdy interesowała nas będzie jedynie moc zbioru $\Pi_{k,n}^p$, również będziemy pomijać indeks p i będziemy pisać Π_k^n.

Lemat 7. *Zachodzi nierówność*

$$\log(\#\Pi_k^n) \leq \begin{cases} K\cdot 2^k((n-k)+\log(k+1)) & k \leq n, \\ K\cdot 2^n\log(k+1) & k > n. \end{cases}$$

Dowód: Proste kombinatoryczne rozumowanie pokazuje że dla $k \leq n$:

$$\#\Pi_k^n \leq \binom{2^n}{2^k}(2k+3)^{2^k} \leq e^{2^k} 2^{(n-k)2^k}(2k+3)^{2^k},$$

zaś dla $k > n$:

$$\#\Pi_k^n \leq (2k+3)^{2^n}.$$

W pierwszym oszacowaniu skorzystaliśmy z nierówności $\binom{n}{k} \leq \frac{n^k}{k!} \leq (\frac{en}{k})^k$. □

Lemat 8. *Dla dowolnego $u \in B_r$ istnieje $\bar{u} \in D_r$ spełniający*

$$\|u - \bar{u}\|_r \leq 1 - 2^{-\frac{1}{r}}.$$

W szczególności dla dowolnego układu liczb $(a_{ij})_{ij}$ dla którego spełniony jest warunek $\sup_{\substack{u \in B_p \\ v \in B_q}} |\sum_{i,j} v_i a_{ij} u_j| < \infty$ zachodzi nierówność

$$\|A\| \colon = \sup_{\substack{u \in B_p \\ v \in B_q}} \Big|\sum_{i,j} v_i a_{ij} u_j\Big| \leq 4 \sup_{\substack{u \in B_p \\ v \in B_q}} \Big|\sum_{i,j} v_i a_{ij} u_j\Big| =: 4\|A\|_D.$$

Dowód: Wystarczy udowodnić, że dla dowolnego $a \in [-1, 1]$ istnieje $\bar{a} \in [-1, 1]$ spełniające:

$$|\bar{a}|^r \in \{0\} \cup \{2^{-k} \colon k \in \mathbb{Z}\},$$
$$|\bar{a}| \leq |a| \text{ oraz, } |a - \bar{a}| \leq (1 - 2^{-\frac{1}{r}})|a|.$$

Istotnie, niech $u \in B_r$. Określmy $\bar{u}$ jako $(\bar{u}_i)_i$. Wówczas:

$$\|u - \bar{u}\|_r = (\sum_i |u_i - \bar{u}_i|^r)^{\frac{1}{r}} \leq (1 - 2^{-\frac{1}{r}})(\sum_i |u_i|^r)^{\frac{1}{r}} \leq (1 - 2^{-\frac{1}{r}}).$$

Aby dowieść istnienie odpowiedniego $\bar{a}$ dla a zauważmy, że jeśli $a = 0$, to można wziąć $\bar{a} = 0$. W przeciwnym przypadku, $2^{-k} \leq |a|^r < 2^{-k+1}$ dla pewnego $k \in \mathbb{Z}$. Wystarczy wówczas przyjąć $\bar{a} = 2^{-\frac{k}{r}} \operatorname{sgn}(a)$.
W celu wykazania drugiej części lematu zauważmy, że jeśli liczby $(b_i)_i$ spełniają $\sup_{u \in D_r} \sum_i b_i u_i < \infty$ to

$$\sup_{u \in B_r} \sum_i b_i u_i \leq \sup_{u \in B_r} \sum_i b_i (u_i - \bar{u}_i) + \sup_{\bar{u} \in D_r} \sum_i b_i \bar{u}_i \leq (1 - 2^{-\frac{1}{r}}) \sup_{u \in B_r} \sum_i b_i u + \sup_{\bar{u} \in D_r} \sum_i b_i \bar{u}_i,$$

a co za tym idzie

$$\sup_{u \in B_r} \sum_i b_i u_i \leq 2^{\frac{1}{r}} \sup_{\bar{u} \in D_r} \sum_i b_i \bar{u}_i \leq 2 \sup_{\bar{u} \in D_r} \sum_i b_i \bar{u}_i.$$

Zachodzi więc

$$\begin{aligned}\|A\| \colon &= \sup_{u \in B_p} \sup_{v \in B_q} \Big|\sum_{i,j} v_i a_{ij} u_j\Big| \leq 2 \sup_{u \in B_p} \sup_{v \in D_q} \Big|\sum_{i,j} v_i a_{ij} u_j\Big| = 2 \sup_{v \in D_q} \sup_{u \in B_p} \Big|\sum_{i,j} v_i a_{ij} u_j\Big| \\ &\leq 4 \sup_{v \in D_q} \sup_{u \in D_p} \Big|\sum_{i,j} v_i a_{ij} u_j\Big| = 4\|A\|_D,\end{aligned}$$

czego należało dowieść. □

Lemat 9. *Niech dla $u \in D_r^n$, $A_{-1,n}^r = \emptyset$ oraz*

$$B_{k,n}^r(u) = A_{k,n}^r(u) \setminus A_{k-1,n}^r(u) = \{i \colon |u_i|^r = 2^{-k}\}.$$

Wprowadźmy dodatkowo oznaczenia

$$b_k(u,v) = 2\sqrt{\log \#\Pi_k^n}\Big(\sum_{\substack{i\in B_k^n(v)\\ j\in A_{k-1}^m(u)}} v_i^2 a_{ij}^2 u_j^2\Big)^{\frac{1}{2}},$$

$$c_k(u,v) = 2\sqrt{\log \#\Pi_k^n}\Big(\sum_{\substack{i\in A_k^n(v)\\ j\in B_k^n(u)}} v_i^2 a_{ij}^2 u_j^2\Big)^{\frac{1}{2}},$$

$$B_k = \sqrt{2}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}\Big(\sum_{\substack{i\in B_k^n(v)\\ j\in A_{k-1}^n(u)}} v_i^2 a_{ij}^2 u_j^2\Big)^{\frac{1}{2}},$$

$$C_k = \sqrt{2}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}\Big(\sum_{i\in A_k^n(v), j\in B_k^n(u)} v_i^2 a_{ij}^2 u_j^2\Big)^{\frac{1}{2}}.$$

Zachodzi nierówność:

$$\mathbb{E}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}\Big[\sum_{i,j} v_i a_{ij} g_{ij} u_j - \sum_{k=0}^{\infty} b_k(u,v) - \sum_{k=0}^{\infty} c_k(u,v)\Big] \le \sum_{k=0}^{\infty}(B_k + C_k).$$

Dowód: Niech

$$S_{k,l}(u,v) = \sum_{\substack{i\in A_k^n(v)\\ j\in A_l^n(u)}} v_i a_{ij} g_{ij} u_j$$

Udowodnimy przez indukcję, że

$$\mathbb{E}\sup_{\substack{u\in D_p^n\\ v\in D_p^n}}\Big[S_{k,k}(u,v) - \sum_{l=0}^{k} b_l(u,v) - \sum_{l=0}^{k} c_l(u,v)\Big] \le \sum_{l=0}^{k}(B_l + C_l)$$

Istotnie, jeżeli zauważymy, że w wyrażeniu powyżej $\sup_{u\in D_p^n, v\in D_q^n}$ można zastąpić, przez $\sup_{u\in \Pi_k^n, v\in \Pi_k^n}$ oraz, że $\log(\#\Pi_{k,n}^p \#\Pi_{k,n}^q) = 2\log\#\Pi_k^n$, to teza dla $k=0$ wynika z lematu 6 zastosowanego do $U = \Pi_{0,n}^p$, $V = \Pi_{0,n}^q$ i $I(u,v) = A_0^n(u)\times A_0^n(v)$, zaś przejście z k do $k+1$

wynika z nierówności:

$$
\begin{aligned}
&\mathbb{E}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}[S_{k+1,k+1}(u,v)-\sum_{l=0}^{k+1}b_l(u,v)-\sum_{l=0}^{k+1}c_l(u,v)]\\
&=\mathbb{E}\max_{\substack{u\in\Pi^p_{k+1,n}\\ v\in\Pi^q_{k+1,n}}}[S_{k+1,k+1}(u,v)-\sum_{l=0}^{k+1}b_l(u,v)-\sum_{l=0}^{k+1}c_l(u,v)]\\
&\le\mathbb{E}\max_{\substack{u\in\Pi^p_{k+1,n}\\ v\in\Pi^q_{k+1,n}}}[S_{k+1,k+1}(u,v)-S_{k+1,k}(u,v)-c_{k+1}(u,v)]\\
&\quad+\mathbb{E}\max_{\substack{u\in\Pi^p_{k+1,n}\\ v\in\Pi^q_{k+1,n}}}[S_{k+1,k}(u,v)-S_{k,k}(u,v)-b_{k+1}(u,v)]\\
&\quad+\mathbb{E}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}[S_{k,k}(u,v)-\sum_{l=0}^{k}b_l(u,v)-\sum_{l=0}^{k}c_l(u,v)]\\
&\le C_{k+1}+B_{k+1}+\sum_{l=0}^{k}(B_l+C_l).
\end{aligned}
$$

Ostatnia z nierówności wynika z lematu 6 zastosowanego do $U=\Pi^p_{k+1,n}$, $V=\Pi^q_{k+1,n}$, $I(u,v)=B^n_{k+1}(u)\times A^n_{k+1}(v)$; lematu 6 zastosowanego do $U=\Pi^p_{k+1,n}$, $V=\Pi^q_{k+1,n}$, $I(u,v)=A^n_k(u)\times B^n_{k+1}(v)$ oraz założenia indukcyjnego. Teza wynika z udowodnionej przez indukcję nierówności, faktu, że:

$$
\begin{aligned}
&\sup_{\substack{u\in D_p^n\\ v\in D_p^n}}\Big[\sum_{i,j}v_ia_{ij}g_{ij}u_j-\sum_{l=0}^{\infty}b_l(u,v)-\sum_{l=0}^{\infty}c_l(u,v)\Big]\\
&=\lim_{k\to\infty}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}\Big[S_{k,k}(u,v)-\sum_{l=0}^{k}b_l(u,v)-\sum_{l=0}^{k}c_l(u,v)\Big]
\end{aligned}
$$

oraz lematu Fatou. □

Przejdziemy teraz do udowodnienia głównego wyniku w tym rozdziale. Niech $X=(X_{ij})_{i\le n,j\le m}=(a_{ij}g_{ij})_{i\le n,j\le m}$. Bez utraty ogólności (uzupełniając macierz zerami, możemy założyć, że $m=n$.

Twierdzenie 4. *Dla dowolnych $p,q\in(1,2]$:*

$$\mathbb{E}\|X\|_{\ell_p\to\ell_{q^\star}}\le K\log n\big[\max_i(\sum_j a_{ij}^{p^\star})^{\frac{1}{p^\star}}+\max_j(\sum_i a_{ij}^{q^\star})^{\frac{1}{q^\star}}\big].$$

Dowód: Na mocy lematu 8 wystarczy wykazać, że

$$\mathbb{E}\sup_{\substack{u\in D_p^n\\ v\in D_q^n}}\sum_{i,j}v_ia_{ij}g_{ij}u_j\le K\log n\big[\max_i(\sum_j a_{ij}^{p^\star})^{\frac{1}{p^\star}}+\max_j(\sum_i a_{ij}^{q^\star})^{\frac{1}{q^\star}}\big].$$

Niech $N = \lceil \log(n+1) \rceil$. Z lematu 9,

$$\mathbb{E} \sup_{\substack{u \in D_p^n \\ v \in D_p^n}} \Big[\sum_{i,j} v_i a_{ij} g_{ij} u_j - \sum_{l=0}^{\infty} b_l(u,v) - \sum_{l=0}^{\infty} c_l(u,v) \Big] \leq \sum_{l=0}^{\infty} (B_l + C_l). \tag{3.1}$$

Zauważmy, że na mocy lematu 7, nierówności Schwarza oraz faktu, że

$$\sum_{k=0}^{N} \log(k+1) = \log(N+1)! \leq KN \log N \leq KN^2$$

zachodzi nierówność:

$$\begin{aligned}
\sum_{k=0}^{N} b_k(u,v) &\leq K \sum_{k=0}^{N} \big(\sqrt{N-k} + \sqrt{\log(k+1)} \big) \Big(2^k \sum_{\substack{i \in B_k^N(v) \\ j \in A_{k-1}^N(u)}} v_i^2 a_{ij}^2 u_j^2 \Big)^{\frac{1}{2}} \\
&\leq KN \Big(\sum_{k=0}^{N} 2^k \sum_{\substack{i \in B_k^N(v) \\ j \in A_{k-1}^N(u)}} v_i^2 a_{ij}^2 u_j^2 \Big)^{\frac{1}{2}}.
\end{aligned}$$

Z definicji zbiorów B_k^N, jeśli $i \in B_k^N(v)$, to $|v_i|^q = 2^{-k}$, więc wyrażenie powyżej jest nie większe niż

$$KN \sqrt{\sum_{j \in A_{N-1}^N(u)} u_j^2 \sum_{i \in A_N^N(v)} v_i^{2-q} a_{ij}^2}.$$

Podobnie

$$\begin{aligned}
\sum_{k=N+1}^{\infty} b_k(u,v) &\leq K \sum_{k=N+1}^{\infty} 2^{\frac{N-k}{2}} \sqrt{\log(k+1)} \Big(2^k \sum_{\substack{i \in B_{k,N}^q(v) \\ j \in A_{k-1,N}^p(u)}} v_i^2 a_{ij}^2 u_j^2 \Big)^{\frac{1}{2}} \\
&\leq K \sqrt{\log(N+1)} \sqrt{\sum_j u_j^2 \sum_i v_i^{2-q} a_{ij}^2}.
\end{aligned}$$

Łącząc powyższe dwie nierówności otrzymujemy i korzystając z nierówności Höldera oraz faktu, że $(|u_j|^2)_j \in L^1$,

$$\begin{aligned}
\sum_{k=0}^{\infty} b_k(u,v) &\leq KN \sqrt{\sum_j u_j^2 \sum_i v_i^{2-q} a_{ij}^2} \leq KN \max_j \sqrt{\big(\sum_i |v_i|^q \big)^{\frac{2-q}{q}} \big(\sum_i |a_{ij}|^{q^\star} \big)^{\frac{2}{q^\star}}} \\
&\leq KN \max_j \big(\sum_i |a_{ij}|^{q^\star} \big)^{\frac{1}{q^\star}}.
\end{aligned} \tag{3.2}$$

Podobnie wykazujemy, że:

$$\sum_{k=0}^{\infty} c_k(u,v) \leq KN \max_i \big(\sum_j |a_{ij}|^{p^\star} \big)^{\frac{1}{p^\star}}. \tag{3.3}$$

W analogiczny sposób, korzystając z nierówności Höldera dowodzimy, że

$$B_k \leq 2^{\frac{-k+2}{2}} \max_j \big(\sum_i |a_{ij}|^{q^\star} \big)^{\frac{1}{q^\star}},$$

więc

$$\sum_{k=0}^{\infty} B_k \leq K \max_j \big(\sum_i |a_{ij}|^{q^\star} \big)^{\frac{1}{q^\star}}. \tag{3.4}$$

Podobnie

$$\sum_{k=0}^{\infty} C_k \leq K \max_i \big(\sum_j |a_{ij}|^{p^\star} \big)^{\frac{1}{p^\star}}. \tag{3.5}$$

Łącząc (3.1), (3.2), (3.3), (3.4) oraz (3.5) otrzymujemy tezę. $\square$

Bibliografia

[1] Stéphane Boucheron, Gábor Lugosi, Pascal Massart *Concentration Inequalities. A Nonasymptotic Theory of Independence*, Oxford University Press, 2013

[2] Alfonso S. Bandeira, Ramon van Handel, *Sharp nonasymptotic bounds on the norm of random matrices with independent entries*, Ann. Probab. 44 (2016) 2479–2506.

[3] G. M. Fichtenholz, *Rachunek różniczkowy i całkowy*, tom 2, Państwowe Wydawnictwo Naukowe, 1980.

[4] Olivier Guédon, Aicke Hinrichs, Alexander E. Litvak, Joscha Prochno, *On the expectation of operator norms of random matrices*, Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics 2169, Springer (2017) 151–162.

[5] Rafał Latała *Some estimates of norms of random matrices*, Proc. Amer. Math. Soc. 133 (2005) 1273–1282.

[6] Michel Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs 89, American Mathematical Society, 2001.

[7] Ramon van Handel, *On the spectral norm of Gaussian random matrices*, arxiv:1502.05003.

[8] Michel Talagrand *Upper and lower bounds for stochastic processes*, Springer, 2014.