Metody Losowe w Geometrii Wypukłej

Rafał Latała

21 stycznia 2013

Poniższe notatki powstają na podstawie (wybranych) wykładów z Metod Losowych w Geometrii Wypukłej prowadzonych w semestrze zimowym 2012/13.

Przepraszam za wszystkie nieścisłości i omyłki mogące pojawić się w tekście i jednocześnie zwracam się z prośbą do czytelników, którzy zauważyli błędy lub mają jakieś inne uwagi na temat notatek o kontakt mailowy na adres `rlatala@mimuw.edu.pl` z podaniem wersji notatek (daty) do której chcą się ustosunkować.

Dziękuje panu Robertowi Boguckiemu za uważną lekturą i wychwycenie licznych literówek.

Rafał Latała

Spis treści

1 Wstęp

Celem wykładu jest omówienie pewnych zagadnień związanych z zastosowaniami rachunku prawdopodobieństwa w geometrii wysokowymiarowych zbiorów wypukłych. Metody losowe, o których mowa w tytule, będą polegać na

- wykazywaniu pewnych ciekawych i zaskakujących własności deterministycznych ciał wypukłych (np istnieniu prawie kulistych przekrojów),
- konstruowaniu losowych ciał (wielościanów) wypukłych o nieoczywistych własnościach, których deterministyczna konstrukcja jest często nieznana,
- badaniu własności obiektów probabilistycznych w naturalny sposób pojawiających się w geometrii wypukłej (np. rozkładów jednostajnych na ciałach wypukłych).

Zanim jednak przejdziemy do struktur probabilistycznych w pierwszych wykładach przedstawimy szereg deterministycznych wyników, które wprowadzą nas w tematykę geometrii wypukłej.

2 Nierówność Brunna-Minkowskiego

Zacznijmy od następującej definicji.

Definicja 2.1. *Sumą Minkowskiego* dwóch niepustych zbiorów $A, B \subset \mathbb{R}^n$ nazywamy zbiór

$$A + B = \{x + y\colon \ x \in A, y \in B\}.$$

Dla $\emptyset \neq A \subset \mathbb{R}^n$ i $\lambda \in \mathbb{R}$ definiujemy

$$\lambda A := \{\lambda x\colon \ x \in A\}.$$

Będziemy też pisać $-A$ zamiast $(-1)A$, $A - B$ zamiast $A + (-B)$ oraz $A \pm x$ zamiast $A \pm \{x\}$.

Uwaga 2.2. Jeśli A i B są zwarte, to zbiór $A + B$ jest zwarty. Jeśli jeden ze zbiorów A lub B jest otwarty, to $A + B$ jest otwarty. Ogólnie zbiór $A + B$ nie musi być borelowski, jeśli A i B są borelowskie. Korzystając z teorii zbiorów analitycznych można jednak udowodnić, że suma Minkowskiego zbiorów borelowskich jest mierzalna w sensie Lebesgue'a.

Przykłady.
1. Suma Minkowskiego zbiorów wypukłych jest zbiorem wypukłym. Suma Minkowskiego wielościanów jest wielościanem. Dla dowolnego zbioru wypukłego K i $a, b > 0$, $aK + bK = (a + b)K$.
2. Niech $B(0, r) = \{x \in \mathbb{R}^n \colon |x| < r\}$ będzie euklidesową kulą otwartą o środku w zerze i promieniu $r > 0$. Dla dowolnego niepustego zbioru $A \subset \mathbb{R}^n$,

$$A_r := A + B(0, r) = A + rB(0, 1) = \{x \in \mathbb{R}^n \colon \ \mathrm{dist}(x, A) < r\}.$$

Zbiór A_r nazywamy *r-otoczką* zbioru A.

Głównym celem pierwszego wykładu jest udowodnienie następującego twierdzenia

Twierdzenie 2.3 (Brunn-Minkowski)**.** *Dla dowolnych niepustych zwartych zbiorów $A, B \subset \mathbb{R}^n$,*

$$\mathrm{vol}_n^{1/n}(A + B) \geqslant \mathrm{vol}_n^{1/n}(A) + \mathrm{vol}_n^{1/n}(B). \qquad (1)$$

Uwaga 2.4. Założenie zwartości można zastąpić borelowskością.

Przykład. Załóżmy, że A jest zbiorem wypukłym oraz $B = aA$ dla pewnego $a > 0$. Wówczas $A + B = (1 + a)A$ i jak łatwo widać w nierówności Brunna-Minkowskiego (1) zachodzi wtedy równość.

Notacja. Dla uproszczenia oznaczeń będziemy często pisać $|A|$ zamiast $\mathrm{vol}_n(A)$ dla oznaczenia miary Lebesgue'a zbioru mierzalnego $A \subset \mathbb{R}^n$.

Lemat 2.5. *Twierdzenie 1 zachodzi dla $n = 1$.*

Dowód. Z uwagi na niezmienniczość obu stron nierówności ze względu na przesunięcia zbiorów A i B możemy zakładać, że $\inf A = \sup B = 0$. Zauważmy, że wtedy $A \cap B = \{0\}$, więc $A + B \supset A \cup B$ oraz $|A \cup B| = |A| + |B \setminus A| = |A| + |B|$. $\square$

Dla $n > 1$ twierdzenie Brunna-Minkowskiego już nie jest aż tak oczywiste, choć obecnie znanych jest kilka niezbyt trudnych dowodów. My oprzemy się na pewnej pokrewnej nierówności funkcyjnej, odkrytej przez Prékopę i Leindlera.

2.1 Nierówność Prékopy-Leindlera

Twierdzenie 2.6 (Prékopa-Leindler)**.** *Załóżmy, że $\lambda \in (0,1)$ oraz funkcje mierzalne $f, g, h \colon \mathbb{R}^n \to [0,\infty)$ spełniają warunek*

$$h(\lambda x + (1-\lambda)y) \geqslant f(x)^{\lambda} g(y)^{1-\lambda} \qquad \textit{dla wszystkich } x, y \in \mathbb{R}^n. \tag{2}$$

Wówczas

$$\int_{\mathbb{R}^n} h(x)dx \geqslant \left(\int_{\mathbb{R}^n} f(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}^n} g(x)dx\right)^{1-\lambda}. \tag{3}$$

Dowód. Dowód przeprowadzimy przez indukcję względem n. Zacznijmy od przypadku $n = 1$. Możemy zakładać, że funkcje f i g są ograniczone, wykorzystując zaś jednorodność wystarczy rozpatrzyć przypadek $\sup f = \sup g = 1$. Weźmy $t \in [0,1]$, wówczas

$$\begin{aligned}
|\{x\colon\ h(x) \geqslant t\}| &\geqslant |\lambda\{x\colon\ f(x) \geqslant t\} + (1-\lambda)\{x\colon\ g(x) \geqslant t\}| \\
&\geqslant |\lambda\{x\colon\ f(x) \geqslant t\}| + |(1-\lambda)\{x\colon\ g(x) \geqslant t\}|.
\end{aligned}$$

Stąd, całkując przez części, dostajemy

$$\begin{aligned}
\int_{\mathbb{R}} h(x)dx &\geqslant \int_0^1 \{x\colon\ h(x) \geqslant t\}|dt \\
&\geqslant \lambda \int_0^1 |\{x\colon\ f(x) \geqslant t\}|dt + (1-\lambda) \int_0^1 |\{x\colon\ g(x) \geqslant t\}|dt \\
&= \lambda \int_{\mathbb{R}} f(x)dx + (1-\lambda) \int_{\mathbb{R}} g(x)dx \geqslant \left(\int_{\mathbb{R}} f(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}} g(x)dx\right)^{1-\lambda},
\end{aligned}$$

gdzie ostatnia nierównośc wynika z porównania średniej arytmetycznej i geometrycznej.

Dowód kroku indukcyjnego. Załóżmy, że $n > 1$ i nierówność Prékopy-Leindlera zachodzi w wymiarze $n-1$. Weźmy funkcje f, g i h spełniające (2) i określmy $F, G, H \colon \mathbb{R}^{n-1} \to \mathbb{R}$ wzorami

$$F(x) := \int_{\mathbb{R}} f(x,y)dy, \quad G(x) := \int_{\mathbb{R}} g(x,y)dy, \quad \text{i} \quad H(x) := \int_{\mathbb{R}} h(x,y)dy.$$

Wówczas z udowodnionego już jednowymiarowego przypadku wynika, że

$$H(\lambda x + (1-\lambda)y) \geqslant F(x)^{\lambda} G(y)^{1-\lambda} \qquad \text{dla wszystkich } x, y \in \mathbb{R}^{n-1},$$

zatem z założenia indukcyjnego otrzymujemy

$$\int_{\mathbb{R}^n} h(x)dx = \int_{\mathbb{R}^{n-1}} H(x)dx \geqslant \left(\int_{\mathbb{R}^{n-1}} F(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}^{n-1}} G(x)dx\right)^{1-\lambda}$$
$$= \left(\int_{\mathbb{R}^n} f(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}^n} g(x)dx\right)^{1-\lambda}.$$

□

2.2 Dowód twierdzenia Brunna-Minkowskiego

Z nierówności Prékopy-Leindlera natychmiast można wyprowadzić nierówność Brunna-Minkowskiego w wersji logarytmicznej.

Wniosek 2.7. *Dla dowolnych niepustych zwartych zbiorów* $A, B \subset \mathbb{R}^n$ *oraz* $\lambda \in [0,1]$,

$$\mathrm{vol}_n(\lambda A + (1-\lambda)B) \geqslant \mathrm{vol}_n(A)^{\lambda}\mathrm{vol}_n(B)^{1-\lambda}. \tag{4}$$

Dowód. Dla $\lambda \in \{0,1\}$ nierówność jest oczywista. Dla $\lambda \in (0,1)$ stosujemy nierówność Prékopy-Leindlera z $f = I_A$, $g = I_B$ oraz $h = I_{\lambda A+(1-\lambda)B}$. □

Dowód twierdzenia 2.3. Jeśli $\mathrm{vol}_n(B) = 0$, to wybierając dowolny punkt $b \in B$ mamy

$$\mathrm{vol}_n^{1/n}(A+B) \geqslant \mathrm{vol}_n^{1/n}(A+b) = \mathrm{vol}_n^{1/n}(A) = \mathrm{vol}_n^{1/n}(A) + \mathrm{vol}_n^{1/n}(B),$$

podobny argument działa dla $\mathrm{vol}_n(A) = 0$. Możemy więc zakładać, że $0 < \mathrm{vol}_n(A), \mathrm{vol}_n(B) < \infty$. Określmy

$$\tilde{A} := \mathrm{vol}_n(A)^{-1/n}A, \quad \tilde{B} := \mathrm{vol}_n(A)^{-1/n}B, \quad \lambda := \frac{\mathrm{vol}_n(A)^{1/n}}{\mathrm{vol}_n(A)^{1/n} + \mathrm{vol}_n(B)^{1/n}}.$$

Wówczas $\mathrm{vol}_n(\tilde{A}) = \mathrm{vol}_n(\tilde{B}) = 1$ oraz $\lambda\tilde{A}+(1-\lambda)\tilde{B} = (\mathrm{vol}_n(A)^{1/n}+\mathrm{vol}_n(B)^{1/n})^{-1}(A+B)$, zatem z Wniosku 2.7,

$$\mathrm{vol}_n^{1/n}(A+B) = (\mathrm{vol}_n(A)^{1/n} + \mathrm{vol}_n(B)^{1/n})\mathrm{vol}_n(\lambda\tilde{A} + (1-\lambda)\tilde{B})$$
$$\geqslant \mathrm{vol}_n(A)^{1/n} + \mathrm{vol}_n(B)^{1/n}.$$

□

2.3 Nierówność izoperymetryczna

Zacznijmy od sformułowania otoczeniowej wersji nierówności izoperymetrycznej dla miary Lebesgue'a. Przypomnijmy, że $A_t = A + B(0,t)$.

Twierdzenie 2.8. *Jeśli* A *jest podzbiorem borelowskim* $\mathbb{R}^n$ *takim, że* $\mathrm{vol}_n(A) = \mathrm{vol}_n(B(x_0,r))$, *to dla dowolnego* $t > 0$,

$$\mathrm{vol}_n(A_t) \geqslant \mathrm{vol}_n(B(x_0,r)_t) = \mathrm{vol}_n(B(x_0,r+t)).$$

Dowód. Niech $c_n = \mathrm{vol}_n(B(0,1))$, wówczas $\mathrm{vol}_n(A) = c_n r^n$ i na podstawie nierówności Brunna-Minkowskiego (1),

$$\begin{aligned}\mathrm{vol}_n(A_t) &= \mathrm{vol}_n(A + B(0,t)) \geqslant (\mathrm{vol}_n(A)^{1/n} + \mathrm{vol}_n(B(0,t))^{1/n})^n \\ &= c_n(r+t)^n = \mathrm{vol}_n(B(x_0, r+t)).\end{aligned}$$

□

Definicja 2.9. Dla miary μ na przestrzeni metrycznej (X,d) określamy *zewnętrzną miarę brzegową* μ^+ wzorem

$$\mu^+(A) := \liminf_{t\to 0+} \frac{\mu(A_t) - \mu(A)}{t}.$$

Uwaga 2.10. Jeśli miara μ na $\mathbb{R}^n$ ma ciągłą gęstość $g(x)$ oraz zbiór A ma gładki brzeg, to

$$\mu^+(A) = \int_{\partial A} g(x) dH_{n-1}(x),$$

gdzie H_{n-1} oznacza $n-1$ wymiarową miarę Haussdorffa. W szczególności $\mathrm{vol}_n^+(A) = \mathrm{vol}_{n-1}(\partial A)$ dla zbiorów borelowskich A w $\mathbb{R}^n$ z gładkim brzegiem.

Równoważna różniczkowa forma klasycznej nierówności izoperymetrycznej mówi, że spośród zbiorów o ustalonej objętości najmniejszą powierzchnię brzegu ma kula. Dokładniej:

Twierdzenie 2.11. *Jeśli A jest podzbiorem borelowskim $\mathbb{R}^n$ takim, że $\mathrm{vol}_n(A) = \mathrm{vol}_n(B(x_0,r))$, to*

$$\mathrm{vol}_n^+(A) \geqslant \mathrm{vol}_n^+(B(x_0,r)) = n c_n^{1/n} (\mathrm{vol}_n(A))^{(n-1)/n},$$

gdzie

$$c_n = \mathrm{vol}_n(B(0,1)) = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

Pierwsza historycznie forma nierówności izoperymetrycznej mówiła, że z podzbiorów płaszczyzny o ustalonym obwodzie największe pole ma koło. W tej formie równoważna postać Twierdzeń 2.8 i 2.11 jest następująca:

Twierdzenie 2.12. *Jeśli A jest podzbiorem borelowskim $\mathbb{R}^n$, $n \geqslant 2$, takim, że $\mathrm{vol}_n^+(A) = \mathrm{vol}_n^+(B(x_0,r))$, to $\mathrm{vol}_n(A) \leqslant \mathrm{vol}_n(B(x_0,r))$.*

Uwaga 2.13. Ogólnie dla miary μ na przestrzeni metrycznej (X,d) można postawić problem izoperymetryczny w wersji klasycznej (różniczkowej) polegający na wyznaczeniu *funkcji izoperymetrycznej*

$$I_\mu(c) = \inf\{\mu^+(A) \colon \mu(A) = c\}$$

oraz w wersji otoczeniowej pytający o *funkcję koncentracji*

$$R_\mu(c,t) = \inf\{\mu(A_t) \colon \mu(A) = c\}.$$

Dokładna postać funkcji I_μ i R_μ jest znana tylko w kilku przypadkach (obejmujących między innymi miarę powierzchniową na S^{n-1} i rozkład gaussowski). Dla nieco szerszej klasy miar znane są dość dobre oszacowania funkcji izoperymetrycznej i koncentracyjnej. Jedno z takich oszacowań dla miary gaussowskiej zaprezentujemy w czasie kolejnych wykładów.

3 Koncentracja Gaussowska

Przez γ_n będziemy oznaczać kanoniczny rozkład gaussowski na $\mathbb{R}^n$, tzn. rozkład z gęstością $(2\pi)^{-n/2}\exp(-|x|^2/2)$.

Nasze pierwsze twierdzenie mówi o koncentracji funkcji Lipschitzowskich względem γ_n.

Twierdzenie 3.1. *Załóżmy, że $F\colon \mathbb{R}^n \to \mathbb{R}$ jest funkcją L-Lipschitzowską, wówczas dla $t \geqslant 0$,*

$$\gamma_n\Big\{x\colon\; F(x) - \int_{\mathbb{R}^n} F d\gamma_n > Lt\Big\} \leqslant \exp\Big(-\frac{2}{\pi^2}t^2\Big) \tag{5}$$

oraz

$$\gamma_n\Big\{x\colon\; \Big|F(x) - \int_{\mathbb{R}^n} F d\gamma_n\Big| > Lt\Big\} \leqslant 2\exp\Big(-\frac{2}{\pi^2}t^2\Big) \tag{6}$$

Dowód. Ponieważ funkcja $-F$ jest również Lipschitzowska, więc szacowanie (6) wynika z (5). By udowodnić (5) możemy, stosując argument aproksymacyjny, założyć, że funkcja F jest gładka. Niech zmienne X i Y mają kanoniczny rozkład gaussowski, określmy dla $\theta \in [0, \pi/2]$ zmienne

$$X(\theta) = X\sin(\theta) + Y\cos(\theta) \text{ oraz } X'(\theta) := \frac{d}{d\theta}X(\theta) = X\cos(\theta) - Y\sin(\theta).$$

Wówczas

$$F(X) - F(Y) = F(X(\frac{\pi}{2})) - F(X(0)) = \int_0^{\pi/2}\frac{d}{d\theta}F(X(\theta))d\theta = \int_0^{\pi/2}\langle\nabla F(X(\theta)), X'(\theta)\rangle d\theta.$$

Ustalmy $\lambda \in \mathbb{R}$, z wypukłości funkcji $t \to \exp(\lambda t)$ i nierówności Jensena otrzymujemy

$$\exp(\lambda(F(X) - F(Y))) \leqslant \frac{2}{\pi}\int_0^{\pi/2}\exp(\lambda\frac{\pi}{2}\langle\nabla F(X(\theta)), X'(\theta)\rangle)d\theta.$$

Jeszcze raz stosując nierówność Jensena dostajemy

$$\mathbb{E}\exp(\lambda(F(X) - \mathbb{E}F(Y))) \leqslant \mathbb{E}\exp(\lambda(F(X) - F(Y))) \leqslant \mathbb{E}\frac{2}{\pi}\int_0^{\pi/2}\exp(\lambda\frac{\pi}{2}\langle\nabla F(X(\theta)), X'(\theta)\rangle)d\theta.$$

Zauważmy, że para zmiennych $(X(\theta), X'(\theta))$ ma dla każdego θ ten sam rozkład, co (X, Y), zatem na mocy twierdzenia Fubiniego,

$$\mathbb{E}\exp(\lambda(F(X) - \mathbb{E}F(Y))) \leqslant \mathbb{E}\exp(\lambda\frac{\pi}{2}\langle\nabla F(X), Y\rangle) = \mathbb{E}\exp(\lambda^2\frac{\pi^2}{8}|\nabla F(X)|^2),$$

gdzie ostatnia równość wynika z faktu, że $\mathbb{E}\exp(\langle t, Y\rangle) = \exp(|t|^2/2)$ dla $t \in \mathbb{R}^n$. Z Lipschitzowskości F wynika, że $|\nabla F(x)| \leqslant L$ dla wszystkich x, zatem

$$\mathbb{E}\exp(\lambda(F(X) - \mathbb{E}F(Y))) \leqslant \mathbb{E}\exp(\lambda^2 \frac{\pi^2}{8} L^2)$$

i z nierówności Czebyszewa dostajemy, że dla dowolnego $\lambda, t \geqslant 0$,

$$\mathbb{P}(F(X) - \mathbb{E}F(Y) > Lt) \leqslant \exp(\lambda^2 \frac{\pi^2}{8} L^2 - L\lambda t).$$

Optymalizując powyższą nierówność po $\lambda \geqslant 0$, czyli wybierając $\lambda = 4t/(\pi^2 L)$ dostajemy

$$\mathbb{P}(F(X) - \mathbb{E}F(Y) > Lt) \leqslant \exp(-\frac{2}{\pi^2} t^2)$$

czyli nierówność (5). □

Uwaga 3.2. Można udowodnić (zob. np. rodział 5.1 w [2]), że stałą $\frac{2}{\pi^2}$ w (5) i (6) można poprawić do optymalnej stałej $\frac{1}{2}$.

Przypomnijmy, że X jest wektorem gaussowskim w $\mathbb{R}^n$, jeśli ma taki sam rozkład jak afiniczny obraz kanonicznego wektora gaussowskiego.

Wniosek 3.3. *Załóżmy, że X jest wektorem gaussowskim w $\mathbb{R}^n$, a $\| \ \|$ normą na $\mathbb{R}^n$. Wówczas dla $t \geqslant 0$,*

$$\mathbb{P}(|\|X\| - \mathbb{E}\|X\|| > \sigma t) \leqslant 2\exp\Big(-\frac{2}{\pi^2} t^2\Big),$$

gdzie

$$\sigma = \sigma(X) = \sup_{\|u\|_* \leqslant 1} (\mathrm{Var}(\langle u, X\rangle))^{1/2},$$

a $\| \ \|_*$ oznacza normę dualną do $\| \ \|$.

Dowód. Wiemy, że X ma ten sam rozkład co $x_0 + AG$, gdzie G jest kanonicznym wektorem gaussowskim, $x_0 \in \mathbb{R}^n$, zaś A pewnym przekształceniem liniowym w $\mathbb{R}^n$. Zauważmy, że

$$\|x_0 + AG\| = \sup_{\|u\|_* \leqslant 1} \langle u, x_0 + AG\rangle = \sup_{\|u\|_* \leqslant 1} (\langle u, x_0\rangle + \langle A^* u, G\rangle).$$

Funkcja $F(x) = \sup_{\|u\|_* \leqslant 1}(\langle u, x_0\rangle + \langle A^* u, x\rangle)$ jak łatwo sprawdzić jest Lipschitzowska ze stałą

$$\sup_{\|u\|_* \leqslant 1} |A^* u| = \sup_{\|u\|_* \leqslant 1} \mathrm{Var}^{1/2}(\langle A^* u, G\rangle) = \sup_{\|u\|_* \leqslant 1} \mathrm{Var}^{1/2}(\langle u, X\rangle).$$

□

Wniosek 3.4. *Przy założeniach i notacji Wniosku 3.3 mamy dla $p \geqslant 1$,*

$$(\mathbb{E}\|X\|^p)^{1/p} \leqslant \mathbb{E}\|X\| + C_1 \sqrt{p}\sigma(X) \leqslant C_2 \sqrt{p}]\mathbb{E}\|X\|,$$

gdzie C_1 i C_2 są pewnymi stałymi absolutnymi.

Dowód. Pierwsza nierówność wynika łatwo z Wniosku 3.3 i z całkowania przez części, by udowodnić drugą wystarczy zauważyć, że dla dowolnego u takiego, że $\|u\|_* \leqslant 1$ mamy

$$(\mathrm{Var}(\langle u, X\rangle))^{1/2} = \sqrt{\frac{\pi}{2}}\mathbb{E}|\langle u, X - \mathbb{E}X\rangle| \leqslant \sqrt{\frac{\pi}{2}}\mathbb{E}\|X - \mathbb{E}X\| \leqslant \sqrt{2\pi}\mathbb{E}\|X\|.$$

□

Uwaga 3.5. Wnioski 3.3 i 3.4 są prawdziwe dla wektorów gaussowskich w ośrodkowych przestrzeniach Banacha. Jest wiele równoważnych definicji wektora gaussowskiego w takich przestrzeniach – np. X jest gaussowski, jeśli $\varphi(X)$ jest zmienną gaussowską dla każdego funkcjonału φ.

4 Ciała wypukłe. Podstawowe fakty i definicje

Definicja 4.1. *Ciałem wypukłym* w $\mathbb{R}^n$ nazywamy zwarty wypukły podzbiór $\mathbb{R}^n$ o niepustym wnętrzu.

Ciało wypukłe K nazywamy *środkowosymetrycznym* lub krócej *symetrycznym* jeśli jest symetryczne względem środka układu współrzędnych, tzn. jeśli $x \in K$, to $-x \in K$.

Przykłady ciał wypukłych.
1.Kula w l_p

$$B_p^n = \{x \in \mathbb{R}^n\colon\ \|x\|_p \leqslant 1\}, \quad 1 \leqslant p \leqslant \infty,$$

gdzie dla $x \in \mathbb{R}^n$

$$\|x\|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}, \quad \|x\|_\infty = \max_{i \leqslant n} |x_i|.$$

Dla uproszczenia notacji będziemy pisać $|x|$ zamiast $\|x\|_2$.
2.Kula Orlicza. Załóżmy, że $\varphi\colon [0,\infty) \to [0,\infty)$ jest niezdegenerowaną funkcją wypukłą taką, że $\varphi(0) = 0$, *kulą Orlicza* wyznaczoną przez φ nazywamy zbiór

$$B_\varphi^n = \{x \in \mathbb{R}^n\colon\ \sum_{i=1}^{n} \varphi(|x_i|) \leqslant 1\}.$$

3.Załóżmy, że A jest zwartym podzbiorem $\mathbb{R}^n$, który nie jest zawarty we właściwej podprzestrzeni afinicznej $\mathbb{R}^n$, wówczas $\mathrm{conv}(A)$ jest ciałem wypukłym. Jeśli A jest zbiorem skończonym, to $\mathrm{conv}(A)$ jest wielościanem. Jeśli A jest zbiorem złożonym z $n+1$ punktów w położeniu ogólnym, to A jest sympleksem.

Uwaga 4.2. Każde symetryczne ciało wypukłe K wyznacza normę $\|\ \|_K$ zadaną wzorem

$$\|x\|_K := \inf\{t > 0\colon x/t \in K\}.$$

Ciało K jest domkniętą kulą jednostkową w normie $\|\ \|_K$.

Na odwrót, dla dowolnej normy $\|\ \|$ na $\mathbb{R}^n$ istnieje symetryczne ciało wypukłe $K = \{x\colon\ \|x\| \leqslant 1\}$ takie, że $\|\ \| = \|\ \|_K$.

4.1 Metryka Hausdorffa

Określimy kilka metryk na zbiorze ciał wypukłych. Zaczniemy od przypomnienia odległości Hausdorffa.

Definicja 4.3. Załóżmy, że A i B są niepustymi zwartymi podzbiorami $\mathbb{R}^n$. Definiujemy wówczas odległość Hausdorffa A i B wzorem

$$d_H(A,B) := \inf\{t > 0\colon\ A \subset B_t \text{ i } B \subset A_t\}.$$

Łatwo sprawdzić, że d_H spełnia wszystkie warunki metryki. Ponadto granica zbiorów wypukłych (wypukłych symetrycznych) w metryce Hausdorffa jest zbiorem wypukłym (wypukłym symetrycznym). Granica ciał wypukłych nie musi być ciałem, gdyż może mieć puste wnętrze.

Uwaga 4.4. Równoważnie możemy określić

$$d_H(A,B) = \max\Big\{\sup_{x\in A}\inf_{y\in B}|x-y|, \sup_{y\in B}\inf_{x\in A}|x-y|\Big\}.$$

Twierdzenie 4.5. *Załózmy, że K jest zwartym podzbiorem $\mathbb{R}^n$. Wówczas przestrzeń $\mathcal{C}(K)$ wszystkich niepustych zwartych podzbiorów K z metryką Hausdorffa jest przestrzenią zwartą.*

By udowodnić to twierdzenie skorzystamy z następującej charakteryzacji zwartych przestrzeni metrycznych.

Twierdzenie 4.6. *Załóżmy, że (E,ρ) jest przestrzenią metryczną. Wówczas E jest zwarta wtedy i tylko wtedy, gdy jest zupełna i całkowicie ograniczona, tzn. dla każdego $\varepsilon > 0$, E da się pokryć skończoną liczbą kul o promieniu ε.*

Dowód Twierdzenia 4.5. Najpierw udowodnimy całkowitą ograniczoność $\mathcal{C}(K)$. Ustalmy $\varepsilon > 0$, wówczas istnieją $x_1,\ldots,x_N \in K$ takie, że $K = \bigcup_{i\leqslant N} B(x_i,\varepsilon)$. Niech $C = \{x_1,\ldots,x_N\}$, wystarczy pokazać, że jeśli $A \in \mathcal{C}(K)$ to istnieje $B \subset C$ taki, że $d_H(B,A) \leqslant \varepsilon$. Przyjmijmy $B := \overline{A_\varepsilon} \cap C$, oczywiście $B \subset \overline{A_\varepsilon}$, ponadto dla $x \in A$ istnieje x_i takie, że $d(x_i,x) \leqslant \varepsilon$, stąd $x_i \in B$, a więc $A \subset \overline{B_\varepsilon}$, czyli istotnie $d_H(B,A) \leqslant \varepsilon$.

By wykazać zupełność, weźmy ciąg Cauchy'ego $(A_k)_{k\geqslant 1}$ w $\mathcal{C}(K)$ i połóżmy $A := \bigcap_{k\geqslant 1}\overline{\bigcup_{l\geqslant k} A_l}$. Wówczas A jest niepusty i zwarty jako przecięcie zstępującego ciągu niepustych zbiorów zwartych, pokażemy, że A jest granicą zbiorów A_k. Niech $\varepsilon > 0$, wtedy dla $k,l \geqslant N$, $A_k \subset (A_l)_\varepsilon$, ustalmy $k \geqslant N$, wystarczy udowodnić, że $d_H(A_k,A) \leqslant \varepsilon$. Niech $x \in A_k$, wówczas istnieją $x_l \in A_l$ takie, że $|x_l - x| \leqslant \varepsilon$ dla $l \geqslant k$, biorąc granicę podciągu zbieżnego x_{l_j} dostajemy $y \in A$ takie, że $|y-x| \leqslant \varepsilon$, zatem $A_k \subset \overline{A_\varepsilon}$. Z drugiej strony jeśli, $x \in A$, to $x \in \overline{\bigcup_{l\geqslant k} A_l} \subset \overline{(A_k)_\varepsilon}$, zatem istotnie $d_H(A_k,A) \leqslant \varepsilon$. $\square$

4.2 Odległość Banacha-Mazura. Twierdzenie Johna

Odległość Hausdorffa ma wiele zalet, ale chcielibyśmy mieć metrykę na zbiorach wypukłych, która mierzy jak bardzo różnią się kształty zbiorów.

Definicja 4.7. Niech K i L będą ciałami wypukłymi w $\mathbb{R}^n$.
Odległością geometryczną K i L nazywamy liczbę

$$d_G(K,L) := \inf\{\frac{b}{a}\colon\ \exists_{x,y}\ a(K-x) \subset L-y \subset b(K-x)\}.$$

Odległością Banacha-Mazura K i L nazywamy liczbę

$$d_{\mathrm{BM}}(K,L) := \inf\{\frac{b}{a}\colon\ \exists_{x,y}\exists_{T\in GL(n)}\ a(K-x) \subset TL-y \subset b(K-x)\}.$$

Zacznijmy od prostej obserwacji.

Fakt 4.8. *W przypadku gdy K i L są symetryczne można w defnicjach odległości geometrycznej i Banacha-Mazura pozbyć się przesunięć x i y.*

Dowód. Załóżmy, że K i L są ciałami symetrycznymi oraz $L-y \subset b(K-x)$. Wówczas $L-(y-bx) \subset bK$ czyli $L+(y-bx) = -(L+(y-bx)) \subset -bK = bK$, a więc

$$L = \frac{1}{2}(L-(y-bx)) + \frac{1}{2}(L+(y-bx)) \subset \frac{1}{2}bK + \frac{1}{2}bK = bK.$$

Analogicznie pokazujemy, że warunek $a(K-x) \subset L-y$ implikuje $aK \subset L$, stąd teza dla odległości geometrycznej (i dla Banacha-Mazura, bo ciało TL też jest symetryczne). □

Tak naprawdę ani odległość geometryczna ani Banacha-Mazura nie są odległościami w ścisłym znaczeniu tego słowa. Po pierwsze spełniają multyplikatywną a nie addytywną nierówność trójkąta (czyli odległość K od L mierzy raczej $\ln d(K,L)$ a nie $d(K,L)$), a po drugie ciała w minimalnej odległości jeden nie muszą się pokrywać.

Fakt 4.9. *i) $d_G(K,L) \geqslant 1$ oraz $d_G(K,L) = 1$ wtedy i tylko wtedy, gdy $K = aL + x$ dla pewnego $a \in \mathbb{R}$ i x.*
ii) $d_{\mathrm{BM}}(K,L) \geqslant 1$ oraz $d_{\mathrm{BM}}(K,L) = 1$ wtedy i tylko wtedy, gdy $K = TL + x$ dla pewnego $T \in GL(n)$ i x.
iii) $d_G(K,M) \leqslant d_G(K,L)d_G(L,M)$ i $d_{\mathrm{BM}}(K,M) \leqslant d_{\mathrm{BM}}(K,L)d_{\mathrm{BM}}(L,M)$.

Prosty dowód pozostawiamy jako ćwiczenie.

Metryka Banacha-Mazura została pierwotnie wprowadzona jako odległość między przestrzeniami Banacha.

Definicja 4.10. Załóżmy, że X i Y są dwiema przestrzeniami unormowanymi. Określamy wówczas odległość Banacha-Mazura X i Y wzorem

$$d_{\mathrm{BM}}(X,Y) := \inf\{\|T\|_{X\to Y}\|T^{-1}\|_{Y\to X}\colon\ T\colon X \to Y \text{ izomorfizm}\}.$$

Fakt 4.11. *Niech K i L będą symetrycznymi ciałami wypukłymi w $\mathbb{R}^n$, wówczas $d_{\mathrm{BM}}(K,L) = d_{\mathrm{BM}}((\mathbb{R}^n, \|\ \|_K), (\mathbb{R}^n, \|\ \|_L))$. W szczególności $d_{\mathrm{BM}}(K,L) = 1$ wtedy i tylko wtedy, gdy przestrzenie $(\mathbb{R}^n, \|\ \|_K)$ i $(\mathbb{R}^n, \|\ \|_L)$ są izometryczne.*

Przypomnijmy, że elipsoidą nazywamy niezdegenerowany afiniczny obraz kuli jednostkowej czyli ciało postaci

$$\mathcal{E} = AB_2^n + y = \big\{x \in \mathbb{R}^n \colon |A^{-1}(x-y)| \leqslant 1\big\} = \big\{x \in \mathbb{R}^n \colon \langle C(x-y), (x-y)\rangle \leqslant 1\big\},$$

gdzie $A \in GL(n)$, $y \in \mathbb{R}^n$, $C = (A^{-1})^* A^{-1}$. Zauważmy, że C jest macierzą symetryczną dodatnią, diagonalizuje się więc w bazie ortogonalnej i ma wartości własne dodatnie. Zatem

$$\mathcal{E} = \Big\{x \in \mathbb{R}^n \colon \sum_{i=1}^{n} \frac{\langle x-y, u_i\rangle^2}{d_i^2} \leqslant 1\Big\}, \tag{7}$$

gdzie $(u_i)_{i\leqslant n}$ to baza ortonormalna $\mathbb{R}^n$, a $(d_i)_{i\leqslant n}$ liczby dodatnie (długości półosi elipsoidy). Ponadto elipsoida $\mathcal{E}$ postaci (7) jest symetryczna względem y i ma objętość

$$\mathrm{vol}_n(\mathcal{E}) = \mathrm{vol}_n(B_2^n)\prod_{i=1}^{n} d_i.$$

Twierdzenie 4.12 (John)**.** *Niech K będzie symetrycznym ciałem wypukłym w $\mathbb{R}^n$, wówczas istnieje dokładnie jedna elipsoida symetryczna maksymalnej objętości $\mathcal{E}_{\max}$ wpisana w K. Ponadto $\mathcal{E}_{\max} \subset K \subset \sqrt{n}\mathcal{E}_{\max}$.*

Wniosek 4.13. *Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K, $d_{\mathrm{BM}}(K, B_2^n) \leqslant \sqrt{n}$.*

Dowód twierdzenia Johna zacznijmy od następującego lematu.

Lemat 4.14. *Załóżmy, że $d > 1$, $x_0 = (d, 0, \ldots, 0)$, $K := \mathrm{conv}(B_2^n, x_0, -x_0)$ oraz*

$$\mathcal{E} = \Big\{x \in \mathbb{R}^n \colon \frac{x_1^2}{a^2} + \sum_{i\geqslant 2} \frac{x_i^2}{b^2} \leqslant 1\Big\}.$$

Wówczas $\mathcal{E} \subset K$, jeśli $|b| \leqslant 1$ i $a^2 + (d^2-1)b^2 \leqslant d^2$.

Dowód. Z uwagi na rotacyjną niezmienniczość K i $\mathcal{E}$ na obroty względem osi x_1, wystarczy rozpatrzeć przypadek $n = 2$. Przypadek $|a| \leqslant 1$ jest oczywisty, więc będziemy zakładać, że $|a| > 1 \geqslant |b|$.

Zauważmy, że prosta przechodząca przez punkt $x_0 = (d, 0)$ oraz punkt $(0, d/(d^2-1)^{1/2})$ jest styczna do kuli B_2^2 w punkcie $(1/d, (d^2-1)^{1/2})$ i ma równanie $x_1 + (d^2-1)^{1/2} x_2 = d$. Weźmy teraz punkt $x = (x_1, x_2) \in \mathcal{E}$, pokażemy, że $x \in K$. Z uwagi na symetrię możemy zakładać, że $x_1, x_2 \geqslant 0$. Rozpatrzymy dwa przypadki.

i) $x_1 \leqslant 1/d$. Wówczas

$$\begin{aligned} x_1^2 + x_2^2 &\leqslant x_1^2 + b^2 - \frac{b^2}{a^2}x_1^2 = \Big(1 - \frac{b^2}{a^2}\Big)x_1^2 + b^2 \leqslant \Big(1 - \frac{b^2}{a^2}\Big)\frac{1}{d^2} + b^2 \\ &= \frac{1}{a^2}\Big(\frac{a^2}{d^2} + \Big(1 - \frac{1}{d^2}\Big)b^2\Big) + b^2\Big(1 - \frac{1}{a^2}\Big) \leqslant \frac{1}{a^2} + 1 - \frac{1}{a^2} = 1. \end{aligned}$$

czyli $x \in B_2^n \subset K$.

ii) $x_2 \geqslant 1/d$, wtedy na mocy nierówności Schwarza

$$x_1 + (d^2-1)^{1/2}x_2 \leqslant \Big(\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}\Big)^{1/2}(a^2 + (d^2-1)b^2)^{1/2} \leqslant d,$$

skąd łatwo wynika, że $x \in K$. □

Dowód Twierdzenia 4.12. Istnienie elipsoidy maksymalnej objętości wynika stąd, że bazy ortonormalne tworzą zwarty podzbiór $(\mathbb{R}^n)^n$ oraz długości półoś elipsoid zawartych w K są ograniczone (ze zwartości K). Przekształcając liniowo K możemy założyć $\mathcal{E}_{\max} = B_2^n$.

Załóżmy, że K nie jest zawarte w $\sqrt{n}B_2^n$, wówczas w K istnieje punkt x_0 długości $d > \sqrt{n}$. Bez straty ogólności możemy założyć $x_0 = (d, 0, \dots, 0)$. Ponieważ oczywiście $K \supset \mathrm{conv}(B_2^n, x_0, -x_0)$, więc z Lematu 4.14 elipsoida

$$\mathcal{E} := \Big\{x \in \mathbb{R}^n\colon\ \frac{nx_1^2}{d^2} + \sum_{i\geqslant 2}\frac{n(d^2-1)x_i^2}{(n-1)d^2} \leqslant 1\Big\} \subset K.$$

Mamy jednak

$$\mathrm{vol}_n(\mathcal{E}) = \mathrm{vol}_n(B_2^n)\frac{d}{\sqrt{n}}\bigg(\sqrt{\frac{(n-1)d^2}{n(d^2-1)}}\bigg)^{n-1} = \mathrm{vol}_n(B_2^n)\sqrt{\frac{f(\frac{1}{n})}{f(\frac{1}{d^2})}},$$

gdzie $f(t) = t(1-t)^{n-1}$. Łatwo sprawdzić, że f jest rosnąca na $(0, n^{-1/2})$, więc $\mathcal{E}$ ma większą objętość niż $\mathcal{E}_{\max} = B_2^n$, co jest niemożliwe.

By wykazać jednoznaczność, załóżmy, że $\mathcal{E}$ jest elipsoidą zawartą w K o tej samej objętości co $\mathcal{E}_{\max} = B_2^n$. Ewentualnie obracając K możemy założyć, że

$$\mathcal{E} = \Big\{x \in \mathbb{R}^n\colon\ \sum_{i=1}^n \frac{x_i^2}{a_i^2} \leqslant 1\Big\},$$

gdzie $a_1, \dots, a_n > 0$, wówczas $1 = \mathrm{vol}(\mathcal{E}')/\mathrm{vol}(B_2^n) = a_1a_2\cdots a_n$. Niech

$$\tilde{\mathcal{E}} := \Big\{x \in \mathbb{R}^n\colon\ \sum_{i=1}^n \frac{4x_i^2}{(1+a_i)^2} \leqslant 1\Big\}.$$

Zauważmy, że jeśli $x \in \mathcal{E}$ oraz zdefiniujemy y i z wzorami $y_i = 2a_ix_i/(1+a_i)$ oraz $z_i = 2/(1+x_i)$, to $x \in \mathcal{E}$, $y \in B_2^n$ oraz $x = (y+z)/2$. Zatem $\tilde{\mathcal{E}} \subset \frac{1}{2}(\mathcal{E} + B_2^n) \subset K$. Ponadto

$$\mathrm{vol}_n(\tilde{\mathcal{E}}) = \mathrm{vol}_n(B_2^n)\prod_{i=1}^n \frac{1+a_i}{2} \geqslant \mathrm{vol}_n(B_2^n)\prod_{i=1}^n \sqrt{a_i} = \mathrm{vol}_n(B_2^n)$$

oraz równość zachodzi wtedy i tylko wtedy, gdy $a_i = 1$ dla wszystkich i, czyli gdy $\mathcal{E} = B_2^n$. □

Przykład. Mamy $d_{\mathrm{BM}}(B_1^n, B_2^n) = \sqrt{n}$, czyli stała $\sqrt{n}$ w twierdzeniu Johna jest optymalna.

Istotnie, niech $d := d_{\mathrm{BM}}(B_1^n, B_2^n)$, wtedy istnieje elipsoida $\mathcal{E}$ taka, że $\mathcal{E} \subset B_\infty^n \subset d\mathcal{E}$. Mamy wtedy $\|x\|_1 \leqslant \|x\|_{\mathcal{E}} \leqslant d\|x\|_1$ dla wszystkich x, stąd na mocy tożsamości równoległoboku,

$$\begin{aligned} d^2 n = d^2 \sum_{i=1}^n \|e_i\|_1^2 \geqslant \sum_{i=1}^n \|e_i\|_{\mathcal{E}}^2 = \frac{1}{2^n} \sum_{\eta_1,\eta_2,\ldots,\eta_n=\pm 1} \Big\| \sum_{i=1}^n \eta_i e_i \Big\|_{\mathcal{E}}^2 \\ \geqslant \frac{1}{2^n} \sum_{\eta_1,\eta_2,\ldots,\eta_n=\pm 1} \Big\| \sum_{i=1}^n \eta_i e_i \Big\|_1^2 = n^2, \end{aligned}$$

zatem $d \geqslant \sqrt{n}$. Z drugiej strony $n^{-1/2} B_2^n \subset B_1^n \subset B_2^n$, więc $d = \sqrt{n}$.

Wniosek 4.15. *Zbiór n-wymiariowych symetrycznych ciał wypukłych z metryką Banacha-Mazura jest zwarty.*

Dowód. Niech (K_m) będzie ciągiem ciał wypukłych symetrycznych. Ponieważ metryka Banacha-Mazura jest niezmiennicza na przekształcenia liniowe, więc możemy zakładać, że $B_2^n \subset K_m \subset \sqrt{n} B_2^n$. Zbiór $\sqrt{n} B_2^n$ jest zwarty więc pewien podciąg K_{m_j} zbiega do zbioru zwartego K w metryce Hausdorffa. Zbiór K jest wypukły, symetryczny oraz $B_2^n \subset K$, więc K jest ciałem. Zauważmy też, że jeśli L jest ciałem wypukłym takim, że $B_2^n \subset L$, to $L_\varepsilon \subset L + \varepsilon L = (1+\varepsilon)L$, więc $d_{\mathrm{BM}}(K_{m_j}, K) \leqslant (1 + d_H(K_{m_j}, K))^2$, czyli K_{m_j} zbiega do K w metryce Banacha-Mazura. $\square$

Uwaga 4.16. Zachodzi niesymetryczna wersja twierdzenia Johna: jeśli K jest ciałem wypukłym, to istnieje dokładnie jedna elipsoida maksymalnej objętości $\mathcal{E}_{\max}$ wpisana w K oraz $\mathcal{E}_{\max} - x \subset K - x \subset n(\mathcal{E}_{\max} - x)$, gdzie x jest środkiem $\mathcal{E}$. W szczególności $d_{\mathrm{BM}}(K, B_2^n) \leqslant n$. Zbiór symetrycznych ciał wypukłych w $\mathbb{R}^n$ z metryką Banacha-Mazura jest zwarty.

4.3 Polarność

Definicja 4.17. Dla niepustego zbioru $A \subset \mathbb{R}^n$ definiujemy

$$A^0 := \{y \in \mathbb{R}^n \colon \ \langle x, y\rangle \leqslant 1 \text{ dla wszystkich } x \in A\}.$$

Fakt 4.18. *i) A^0 jest domkniętym, wypukłym podzbiorem $\mathbb{R}^n$ zawierającym* 0,
ii) $A^0 = \mathrm{conv}(A)^0$,
iii) A^0 jest zwarty, jeśli $0 \in \mathrm{int}(\mathrm{conv}(A))$,
iv) $0 \in \mathrm{int}(A^0)$, *jeśli A jest ograniczony,*
v) A^0 jest symetryczny, jeśli A jest symetryczny,
vi) $(T^*)^{-1}(A^0) = (T(A))^0$ *dla* $T \in \mathrm{GL}(n)$, *w szczególności* $(tA)^0 = \frac{1}{t} A^0$ *dla* $t > 0$,
vii) $A^0 \supset B^0$, *jeśli* $A \subset B$.

Prosty dowód pozostawiamy jako ćwiczenie.

Wniosek 4.19. *Zbiór A^0 jest ciałem wypukłym, jeśli A jest ciałem wypukłym zawierającym 0 w swoim wnętrzu oraz A^0 jest symetrycznym ciałem wypukłym, jeśli A jest symetrycznym ciałem wypukłym.*

Uwaga 4.20. Jeśli K jest kulą jednostkową w n-wymiarowej przestrzeni Banacha X, to K^0 jest kulą jednostkową w przestrzeni dualnej X^*.

Ważną własnością polarności jest jej inwolucyjność.

Fakt 4.21. *Dla dowolnego niepustego A, $A \subset (A^0)^0$. Jeśli A jest ciałem wypukłym zawierającym 0, to $(A^0)^0 = A$.*

Dowód. Pierwsza część jest oczywista. By udowodnić drugą wystarczy pokazać zawieranie $(A^0)^0 \subset A$. Załóżmy, że $x \in \mathbb{R}^n \setminus A$, wówczas A i x można oddzielić pewną hiperpłaszczyzną, tzn. istnieje $y \in \mathbb{R}^n$ takie, że $A \subset \{z\colon \langle z, y\rangle < a\}$ oraz $\langle x, y\rangle > a$. Ponieważ $0 \in A$, to $a > 0$, stąd $y/a \in A^0$, co implikuje, że $x \notin (A^0)^0$. □

Wniosek 4.22. *Dla dowolnych symetrycznych ciał wypukłych mamy $d_{\mathrm{BM}}(K, L) = d_{\mathrm{BM}}(K^0, L^0)$.*

Dowód. Jeśli $aK \subset TL \subset bK$, to $\frac{1}{a}K^0 = (aK)^0 \supset (TL)^0 = (T^*)^{-1}(L^0) \supset (bK)^0 = \frac{1}{b}K^0$, stąd $d_{\mathrm{BM}}(K, L) \geqslant d_{\mathrm{BM}}(K^0, L^0)$. Przeciwna nierówność wynika stąd, że $(K^0)^0 = K$ i $(L^0)^0 = L$. □

Uwaga 4.23. W przypadku niesymetrycznych ciał wypukłych K wygodnie jest często zdefiniować polarność względem innego punktu niż 0, kładąc dla $z \in \mathbb{R}^n$

$$K^z = \{y \in \mathbb{R}^n\colon\ \langle x - z, y\rangle \leqslant 1 \text{ dla wszystkich } x \in K\}.$$

4.4 Liczby pokryciowe

Definicja 4.24. Załóżmy, że K i L są ciałami wypukłymi w $\mathbb{R}^n$. Definiujemy wówczas *liczbę pokryciową K względem L* wzorem

$$N(K, L) := \inf\Big\{N\colon\ K \subset \bigcup_{i=1}^{N}(x_i + L) \text{ dla pewnych } x_1, \ldots, x_N \in K\Big\}.$$

Będziemy bardzo często wykorzystywali następujące oszacowania liczb pokryciowych oparte na porównywaniu objętości.

Fakt 4.25. *i) Dla dowolnych ciał wypukłych K, L w $\mathbb{R}^n$,*

$$N(K, L) \geqslant \frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(L)}.$$

ii) Dla dowolnych n-wymiarowych symetrycznych ciał wypukłych

$$N(K, L) \leqslant \frac{\mathrm{vol}_n(2K + L)}{\mathrm{vol}_n(L)}.$$

W szczególności, jeśli $L \subset K$, to $N(K, L) \leqslant 3^n \mathrm{vol}_n(K)/\mathrm{vol}_n(L)$.

Dowód. i) Mamy $\left|\bigcup_{i=1}^N (x_i + L)\right| \leqslant \sum_{i=1}^N |x_i + L| = N|L|$.

ii) Wybierzmy maksymalny zbiór $A = \{x_1, \ldots, x_N\} \subset K$ taki, że $\|x_i - x_j\|_L \geqslant 1$ dla $i \neq j$. Wówczas ciała $x_i + \frac{1}{2}L$, $1 \leqslant i \leqslant N$, mają rozłączne wnętrza oraz są zawarte w $K + \frac{1}{2}L$, stąd

$$N\Big|\frac{1}{2}L\Big| = \sum_{i=1}^N \Big|x_i + \frac{1}{2}L\Big| = \left|\bigcup_{i=1}^N x_i + \frac{1}{2}L\right| \leqslant \Big|K + \frac{1}{2}L\Big|.$$

Ponadto z maksymalności A, $K \subset \bigcup_{i=1}^N (x_i + L)$, zatem $N(K, L) \leqslant N \leqslant |K + \frac{1}{2}L|/|\frac{1}{2}L| = |2K + L|/|L|$. □

Wniosek 4.26. *Dla dowolnego symetrycznego ciała K i $\varepsilon > 0$,*

$$\frac{1}{\varepsilon^n} \leqslant N(K, \varepsilon K) \leqslant \Big(1 + \frac{2}{\varepsilon}\Big)^n.$$

Uwaga 4.27. i) Dla przestrzeni metrycznej (T, d) definiujemy liczbę pokryciową $N(T, d, \varepsilon)$ jako najmniejszą liczbę kul domkniętych o środkach w T i promieniu ε, która pokrywa T. Wtedy dla ciała symetrycznego L, $N(K, \varepsilon L) = N(K, \|\ \|_L, \varepsilon)$.
ii) Dla przestrzeni metrycznej (T, d) możemy zdefiniować

$$S(T, d, \varepsilon) := \sup\Big\{N\colon\ \ \text{istnieją } x_1, \ldots, x_N \in T,\ d(x_i, x_j) > \varepsilon \text{ dla } i \neq j\Big\},$$

wtedy $N(T, d, \varepsilon) \leqslant S(T, d, \varepsilon) \leqslant N(T, d, \varepsilon/2)$.
iii) Możemy zmodyfikować definicję $N(K, L)$ nie żądając by środki przesunięć należały do K:

$$\tilde{N}(K, L) := \inf\Big\{N\colon\ K \subset \bigcup_{i=1}^N (x_i + L) \text{ dla pewnych } x_1, \ldots, x_N \in \mathbb{R}^n\Big\}.$$

Wówczas oczywiście $\tilde{N}(K, L) \leqslant N(K, L)$, ponadto dla symetrycznych ciał L, $N(K, 2L) \leqslant \tilde{N}(K, L)$.

Definicja 4.28. Zbiór $A \subset K$ taki, że $K \subset A + \varepsilon K$ będziemy nazywać ε-siecią w K.

Fakt 4.29. *Załóżmy, że $\varepsilon \in (0, 1)$ oraz A jest ε-siecią w K, wówczas dla dowolnego operatora liniowego T z $(\mathbb{R}^n, \|\ \|_K)$ w przestrzeń unormowaną $(X, \|\ \|)$ zachodzi*

$$\|T\| = \sup_{\|x\|_K \leqslant 1} \|Tx\| \leqslant \frac{1}{1 - \varepsilon} \sup_{x \in A} \|Tx\|.$$

Dowód. Jeśli $x \in K$ to istnieje $y \in A$ taki, że $\|x - y\|_K \leqslant \varepsilon$ i wtedy $\|Tx\| \leqslant \|Ty\| + \|T(x - y)\| \leqslant \|Ty\| + \varepsilon\|T\|$, biorąc supremum po wszystkich $x \in K$ dostajemy tezę. □

5 Twierdzenie Gluskina

Z twierdzenia Johna i podmultyplikatywności odległości Banacha-Mazura natychmiast wynika

Wniosek 5.1. *Dla dowolnych n-wymiarowych ciał wypukłych K i L, $d_{\mathrm{BM}}(K,L) \leqslant n$.*

Przez długi czas problem dolnego oszacowania średnicy kompaktu Banacha-Mazura był otwarty. Rozwiązał go z początkiem lat 80-tych Efim Głuskin.

Twierdzenie 5.2 (Głuskin)**.** *Dla dowolnego n istnieją symetryczne ciała wypukłe K i L takie, że $d_{\mathrm{BM}}(K,L) \geqslant cn$, gdzie $c > 0$ jest stałą uniwersalną.*

Konstrukcja Głuskina da się prosto opisać. Niech $X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_m$ będą niezależnymi wektorami losowymi o rozkładzie jednostajnym na S^{n-1}. Kładziemy

$$K_m := \mathrm{conv}(\pm e_1, \ldots, \pm e_n, \pm X_1, \ldots, \pm X_m), \quad L_m := \mathrm{conv}(\pm e_1, \ldots, \pm e_n, \pm Y_1, \ldots, \pm Y_m).$$

Oczywiście K_m i L_m są (losowymi) symetrycznymi ciałami wypukłymi. Pokażemy, że (dla odpowiednio dobranego m) odległość Banacha-Mazura tych ciał jest z dużym prawdopodobieństwem większa niż cn.

W dowodzie wykorzystamy szereg użytecznych faktów, które sformułujemy osobno. Będziemy też potrzebowali pewnych ustaleń notacyjnych.

Przez $M_{n\times n}$ będziemy oznaczać przestrzeń macierzy (rzeczywistych) $n \times n$, którą możemy utożsamiać zarówno z przestrzenią przekształceń liniowych $\mathbb{R}^n$ w $\mathbb{R}^n$ jak i z $\mathbb{R}^{n^2}$. Z tymi utożsamieniami są ściśle związane dwie normy macierzy $A = (a_{ij}) \in M_{n\times n}$ – operatorowa i Hilberta-Schmidta:

$$\|A\|_{\mathrm{op}} := \|A \colon l_2^n \mapsto l_2^n\| = \sup\{|Ax| \colon \ |x| \leqslant 1\},$$

$$\|A\|_{\mathrm{HS}} := \Big(\sum_{i=1}^n |Ae_i|^2\Big)^{1/2} = \Big(\sum_{i=1}^n |a_{ij}|^2\Big)^{1/2}.$$

Przez U_{op} i U_{HS} będziemy oznaczać kule jednostkowe w $M_{n\times n}$ w odpowiednich normach, tzn.

$$U_{\mathrm{op}} := \{A \in M_{n\times n} \colon \ \|A\|_{\mathrm{op}} \leqslant 1\}, \quad U_{\mathrm{HS}} := \{A \in M_{n\times n} \colon \ \|A\|_{\mathrm{HS}} \leqslant 1\}.$$

Dla uproszczenia notacji będziemy też pisać dla operatora liniowego T na $\mathbb{R}^n$ i dwóch symetrycznych n-wymiarowych ciał wypukłych K i L,

$$\|T \colon K \mapsto L\| := \|T \colon (\mathbb{R}^n, \| \ \|_K) \mapsto (\mathbb{R}^n, \| \ \|_L)\|.$$

Fakt 5.3 (Chevet)**.** *Niech $G = (g_{ij})_{i,j\leqslant n}$ będzie losową macierzą $n \times n$, której współrzędne g_{ij} są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(0,1)$. Wówczas $\mathbb{E}\|G\|_{\mathrm{op}}^2 \leqslant Cn$.*

Dowód. Niech $A \subset B_2^n$ będzie 1/2-siecią w B_2^N taka, że $\#A \leqslant N(B_2^n, \frac{1}{2}B_2^n) \leqslant 5^n$. Wówczas

$$\|G\|_{\mathrm{op}} = \sup\{|\langle Gx, y\rangle|:\ x, y \in B_2^n\} \leqslant 4\sup\{|\langle Gx, y\rangle|:\ x, y \in A\}.$$

Zauważmy, że dla dowolnego $x, y \in B_2^n$, zmienna $\langle Gx, y\rangle$ ma rozkład normalny o średniej zero i wariancji $\sigma^2(x,y) = |x|^2|y|^2 \leqslant 1$. Stąd

$$\mathbb{P}(\|G\|_{\mathrm{op}} \geqslant t) \leqslant \sum_{x,y\in A} \mathbb{P}(|\langle Gx, y\rangle| \geqslant t/4) \leqslant \sum_{x,y\in A} \exp(-\frac{t^2}{32\sigma^2(x,y)}) \leqslant 25^n \exp(-\frac{t^2}{32}).$$

Zatem otrzymujemy dla $t_0 > 0$,

$$\begin{aligned}\mathbb{E}\|G\|_{\mathrm{op}}^2 &= 2\int_0^\infty t\mathbb{P}(\|G\|_{\mathrm{op}} \geqslant t)dt \leqslant 2\int_0^{t_0} tdt + 2\int_{t_0}^\infty t25^n \exp(-\frac{t^2}{32})dt \\ &= t_0^2 + 32\cdot 25^n \exp(-\frac{t_0^2}{32}).\end{aligned}$$

Wybierając np $t_0 = 16\sqrt{n}$ otrzymujemy $\mathbb{E}\|G\|_{\mathrm{op}}^2 \leqslant 16^2 n + 1 \leqslant 17^2 n$. □

Fakt 5.4. *Dla dowolnego symetrycznego ciała wypukłego K w $\mathbb{R}^n$ zachodzi*

$$\mathrm{vol}_n(K) = \mathrm{vol}_n(B_2^n)\int_{S^{n-1}} \|\theta\|_K^{-n} d\sigma_{n-1}(\theta).$$

Dowód. Zauważmy, że dla $\theta \in S^{n-1}$ i $r > 0$, $r\theta \in K$ wtedy i tylko wtedy gdy $r \leqslant \|\theta\|_K^{-1}$, stąd całkując we współrzędnych biegunowych dostajemy

$$\begin{aligned}\mathrm{vol}_n(K) &= \int_{\mathbb{R}^n} I_K(x)dx = \mathrm{vol}_{n-1}(S^{n-1})\int_{S^{n-1}}\int_0^\infty I_K(r\theta)r^{n-1}drd\sigma_{n-1}(\theta) \\ &= \mathrm{vol}(B_2^n)\int_{S^{n-1}}\int_0^{\|\theta\|_K^{-1}} nr^{n-1}drd\sigma_{n-1}(\theta) = \mathrm{vol}_n(B_2^n)\int_{S^{n-1}}\|\theta\|_K^{-n}d\sigma_{n-1}(\theta).\end{aligned}$$

□

Zauważmy, że $\|A\|_{\mathrm{op}} \leqslant \|A\|_{\mathrm{HS}} \leqslant \sqrt{n}\|A\|_{\mathrm{op}}$, stąd $U_{\mathrm{HS}} \subset U_{\mathrm{op}} \subset \sqrt{n}U_{\mathrm{HS}}$, czyli

$$\mathrm{vol}_{n^2}(U_{\mathrm{HS}}) \leqslant \mathrm{vol}_{n^2}(U_{\mathrm{op}}) \leqslant n^{n^2/2}\mathrm{vol}_{n^2}(U_{\mathrm{HS}}).$$

Następny lemat pokazuje, że górne szacowanie jest bliskie optymalnemu.

Lemat 5.5. *Istnieje stała c niezależna od wymiaru taka, że dla dowolnego n,*

$$\frac{\mathrm{vol}_{n^2}(U_{\mathrm{op}})}{\mathrm{vol}_{n^2}(U_{\mathrm{HS}})} \geqslant (cn)^{n^2/2}.$$

Dowód. Oczywiście możemy zakładać, że $n \geqslant 2$ (bo dla $n = 1$, $U_{\mathrm{op}} = U_{\mathrm{HS}} = [-1,1]$). Dla uproszczenia notacji będziemy pisać S zamiast S^{n^2-1} i σ zamiast σ_{n^2-1}. Zauważmy, że kulę U_{HS} możemy utożsamiać z $B_2^{n^2}$, więc na mocy Faktu 5.4,

$$\frac{\mathrm{vol}_{n^2}(U_{\mathrm{op}})}{\mathrm{vol}_{n^2}(U_{HS})} = \int_S \|\theta\|_{\mathrm{op}}^{-n^2} d\sigma(\theta).$$

Z nierówności Höldera otrzymujemy

$$\begin{aligned}1 = \int_S 1 d\sigma(\theta) &\leqslant \Big(\int_S \|\theta\|_{\mathrm{op}}^2 d\sigma(\theta)\Big)^{1/2}\Big(\int_S \|\theta\|_{\mathrm{op}}^{-2} d\sigma(\theta)\Big)^{1/2} \\ &\leqslant \Big(\int_S \|\theta\|_{\mathrm{op}}^2 d\sigma(\theta)\Big)^{1/2}\Big(\int_S \|\theta\|_{\mathrm{op}}^{-n^2} d\sigma(\theta)\Big)^{1/n^2}.\end{aligned}$$

Zatem

$$\frac{\mathrm{vol}_{n^2}(U_{\mathrm{op}})}{\mathrm{vol}_{n^2}(U_{HS})} \geqslant \Big(\int_S \|\theta\|_{\mathrm{op}}^2 d\sigma(\theta)\Big)^{-n^2/2}.$$

Niech $G = (g_{ij})$ będzie macierzą gaussowską, której współczynniki są niezależne o rozkładzie $\mathcal{N}(0,1)$. Wówczas G można utożsamiać z kanonicznym wektorem gaussowskim w $\mathbb{R}^{n^2}$, stąd $G/|G|$ ma jednostajny rozkład na S oraz jest niezależne od $|G|$. Zatem

$$\mathbb{E}\|G\|_{\mathrm{op}}^2 = \mathbb{E}|G|^2 \mathbb{E}\Big\|\frac{G}{|G|}\Big\|_{\mathrm{op}}^2 = n^2 \int_S \|\theta\|_{\mathrm{op}}^2 d\sigma(\theta)$$

i z Faktu 5.3 dostajemy $\int_S \|\theta\|_{\mathrm{op}}^2 d\sigma(\theta) \leqslant C/n$. □

Lemat 5.6. *Dla dowolnego ciała wypukłego K w $\mathbb{R}^n$ takiego, że $0 \in \mathrm{int}(K)$ zachodzi*

$$\sigma_{n-1}(K \cap S^{n-1}) \leqslant \frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(B_2^n)}.$$

Dowód. Określmy

$$\tilde{K} := \{x \in \mathbb{R}^n : \frac{x}{|x|} \in K\}.$$

Zauważmy, że jeśli $x \in \tilde{K}$, to $x = |x|\frac{x}{|x|} + (1-|x|)0 \in K$, zatem $\tilde{K} \subset K$. Stąd

$$\sigma_{n-1}(K \cap S^{n-1}) = \frac{\mathrm{vol}_n(\tilde{K})}{\mathrm{vol}_n(B_2^n)} \leqslant \frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(B_2^n)}.$$

□

Lemat 5.7. *Niech $L = \mathrm{conv}(\pm y_1, \ldots, \pm y_k)$ dla pewnych $y_1, y_2, \ldots, y_k \in B_2^n$, $k \geqslant n$. Wówczas*

$$\mathrm{vol}_n(L) \leqslant \binom{2k}{n}\frac{1}{n!} \leqslant \alpha^n \Big(\frac{k}{n^{3/2}}\Big)^n \mathrm{vol}_n(B_2^n),$$

gdzie $\alpha = (6e^3/\pi)^{1/2}$.

Dowód. Zauważmy, że $L = \bigcup \mathrm{conv}(0, x_1, \ldots, x_n)$ gdzie sumowanie przebiega po wszystkich wyborach $x_1, \ldots, x_n$ spośród wektorów $\pm y_1, \ldots, \pm y_k$, a takich wyborów jest $\binom{2k}{n}$. Ponadto na podstawie nierówności Hadamarda,

$$\mathrm{vol}_n(\mathrm{conv}(0, x_1, \ldots, x_n)) = \frac{1}{n!}\det[(x_1, \ldots, x_n)] \leqslant \frac{1}{n!}\prod_{i=1}^{n}|x_i| = \frac{1}{n!}.$$

Nierówność $n! \geqslant (\frac{n}{e})^n$ (która np. wynika ze wzoru Stirlinga) implikuje dla $m \geqslant n$, $\binom{m}{n} \leqslant \frac{m^n}{n!} \leqslant (\frac{em}{n})^n$. Stąd

$$\mathrm{vol}_n(L) \leqslant \sum \mathrm{vol}_n(\mathrm{conv}(0, x_1, \ldots, x_n)) \leqslant \binom{2k}{n}\frac{1}{n!} \leqslant \Big(\frac{2e^2k}{n^2}\Big)^n.$$

Zauważmy, że $\mathrm{vol}(B_2^n) = \pi^{n/2}/\Gamma(n/2+1)$. By zakończyć dowód wystarczy więc wykazać, że

$$\Gamma(\frac{n}{2}+1) \leqslant \Big(\frac{3n}{2e}\Big)^{n/2}.$$

Oszacowanie to sprawdzamy indukcyjnie (zwiększając w kroku indukcyjnym n o 2). Dla $n = 1, 2$ nierówność zachodzi i zauważamy, że

$$\Gamma(\frac{n+2}{2}+1) = (\frac{n}{2}+1)\Gamma(\frac{n}{2}+1) \leqslant \frac{3(n+2)}{2e}\Gamma(\frac{n}{2}+1).$$

□

Lemat 5.8. *Niech L i α będą jak w Lemacie 5.7, wówczas dla $T \in \mathrm{GL}(n)$ takiego, że $|\det(T)| = 1$*

$$\mathbb{P}\Big(\|T\colon K_m \mapsto L\| \leqslant t\frac{n^{3/2}}{\alpha k}\Big) \leqslant t^{nm} \text{ dla } 0 < t < 1.$$

Dowód. Na mocy definicji K_m i niezależności występujących w niej zmiennych X_j,

$$\mathbb{P}(\|T\colon K_m \mapsto L\| \leqslant r) \leqslant \mathbb{P}(TX_j \in rL, 1 \leqslant j \leqslant n) = \mathbb{P}(X_1 \in rT^{-1}L)^n.$$

Ale na podstawie Lematu 5.6,

$$\mathbb{P}(X_1 \in rT^{-1}L) = \sigma_{n-1}(rT^{-1}L \cap S^{n-1}) \leqslant \frac{\mathrm{vol}_n(rT^{-1}L)}{\mathrm{vol}_n(B_2^n)} = r^n\frac{\mathrm{vol}_n(L)}{\mathrm{vol}_n(B_2^n)},$$

gdzie ostatnia równość wynika stąd, że $|\mathrm{det}T| = 1$. Podstawiając $r = tn^{3/2}/(\alpha k)$ i wykorzystując oszacowanie z Lematu 5.7 dostajemy tezę. □

Lemat 5.9. *Niech $L = \mathrm{conv}(\pm e_1, \ldots, \pm e_n, \pm x_1, \ldots, \pm x_l)$ dla pewnych $x_1, \ldots, x_l \in B_2^n$ oraz*

$$\mathcal{M}_L = \{T \in M_{n\times n}\colon |\det(T)| = 1, \|Te_i\|_L \leqslant \sqrt{n}, 1 \leqslant i \leqslant n\}$$

Istnieje stała uniwersalna a taka, że dla $0 < \varepsilon < 1$,

$$N(\mathcal{M}_L, \|\ \|_{\mathrm{op}}, \varepsilon) \leqslant \Big(a\frac{n+l}{\varepsilon n}\Big)^{n^2}$$

Dowód. Zauważmy, że przy utożsamieniu $M_{n\times n}$ z $(\mathbb{R}^n)^n$ danym jako $T \to (Te_1, \dots, Te_n)$ otrzymujemy, że $\mathcal{M}_L \subset (\sqrt{n}L)^n$. Stąd na mocy Uwagi 4.27 iii) dostajemy

$$N(\mathcal{M}_L, \|\ \|_{\mathrm{op}}, 2\varepsilon) \leqslant N((\sqrt{n}L)^n, \|\ \|_{\mathrm{op}}, \varepsilon) = N((\sqrt{n}L)^n, \varepsilon U_{\mathrm{op}}).$$

By oszacować ostatnią wielkość zauważmy, że $B_1^n \subset L$, więc dla $x \in \mathbb{R}^n$

$$\|Tx\|_L \leqslant \|Tx\|_1 \leqslant \sqrt{n}|Tx| \leqslant \sqrt{n}\|T\|_{\mathrm{op}}|x|,$$

zatem $U_{\mathrm{op}} \subset (\sqrt{n}L)^n$. Stąd $2(\sqrt{n}L)^n + \varepsilon U_{\mathrm{op}} \subset (2+\varepsilon)(\sqrt{n}L)^n)$ i na mocy Faktu 4.25

$$\begin{aligned} N\big((\sqrt{n}L)^n, \varepsilon U_{\mathrm{op}}\big) &\leqslant \frac{\mathrm{vol}_{n^2}((2+\varepsilon)(\sqrt{n}L)^n)}{\mathrm{vol}_{n^2}(\varepsilon U_{\mathrm{op}})} \leqslant \Big(1 + \frac{2}{\varepsilon}\Big)^{n^2} \frac{(\mathrm{vol}_n(L))^n}{\mathrm{vol}_{n^2}(U_{\mathrm{op}})} \\ &\leqslant \Big(\frac{3}{\varepsilon}\Big)^{n^2} \frac{\alpha^{n^2}(\frac{n+l}{n^{3/2}})^{n^2}(\mathrm{vol}_n(B_2^n))^n}{(cn)^{n^2/2}\mathrm{vol}_{n^2}(B_2^{n^2})} \leqslant \Big(C\frac{n+l}{\varepsilon n}\Big)^{n^2}, \end{aligned}$$

dla pewnej stałej uniwersalnej C, gdzie druga nierówność wynika z Lematów 5.5 i 5.7, a ostatnie oszacowanie z tego, że $(\mathrm{vol}_n(B_2^n))^{1/n} \sim n^{-1/2}$.

□

Wniosek 5.10. *Niech L i a będą jak w poprzednim lemacie, α jak w Lemacie 5.7. Wówczas dla $0 < \varepsilon < \rho < 1$,*

$$\mathbb{P}\Big(\exists_{T\in M_{n\times n}, |\det T|=1} \|T\colon K_m \mapsto L\| \leqslant (\rho-\varepsilon)\sqrt{n}\Big) \leqslant \Big(a\frac{(n+l)}{\varepsilon n}\Big)^{n^2}\Big(\alpha\rho\frac{n+l}{n}\Big)^{nm}.$$

W szczególności,

$$\mathbb{P}\Big(\exists_{T\in M_{n\times n}, |\det T|=1} \|T\colon K_m \mapsto L_m\| \leqslant (\rho-\varepsilon)\sqrt{n}\Big) \leqslant \Big(a\frac{(n+m)}{\varepsilon n}\Big)^{n^2}\Big(\alpha\rho\frac{n+m}{n}\Big)^{nm}.$$

Dowód. Zauważmy najpierw, że ponieważ $K_m \subset B_2^n$, to $\|\mathrm{Id}\colon K_m \mapsto B_2^n\| \leqslant 1$, ponadto $B_2^n \subset \sqrt{n}B_1^n \subset \sqrt{n}L$, zatem $\|\mathrm{Id}\colon B_2^n \mapsto L\| \leqslant \sqrt{n}$, stąd dla $S \in M_{n\times n}$,

$$\|S\colon K_m \mapsto L\| \leqslant \|\mathrm{Id}\colon K_m \mapsto B_2^n\|\|S\colon B_2^n \mapsto B_2^n\|\|\mathrm{Id}\colon B_2^n \mapsto L\| \leqslant \sqrt{n}\|S\|_{\mathrm{op}}. \tag{8}$$

Zauważmy, że jeśli $\|T\colon K_m \mapsto L\| \leqslant \sqrt{n}$, to $Te_i \in \sqrt{n}L$, więc, jeśli dodatkowo założymy, że $|\det T| = 1$, to $T \in \mathcal{M}_L$. Niech $\mathcal{N}$ będzie ε-siecią w normie operatorowej w $\mathcal{M}_L$, wówczas na podstawie definicji ε-sieci i szacowania (8) dostajemy

$$\begin{aligned} \mathbb{P}(\|T\colon K_m \mapsto L\| &\leqslant (\rho+\varepsilon)\sqrt{n} \text{ dla pewnego } T \in M_{n\times n}, |\det T| = 1) \\ &\leqslant \mathbb{P}(\|T\colon K_m \mapsto L\| \leqslant \rho\sqrt{n} \text{ dla pewnego } T \in \mathcal{N}) \\ &\leqslant \sum_{T\in\mathcal{N}} \mathbb{P}(\|T\colon K_m \mapsto L\| \leqslant 2\rho\sqrt{n}) \leqslant \#\mathcal{N}\Big(\alpha\rho\frac{n+l}{n}\Big)^{nm}, \end{aligned}$$

gdzie ostatnia nierówność wynika z Lematu 5.8. Na podstawie Lematu 5.9 można wybrać taką ε-sieć, by $\#\mathcal{N} \leqslant (a(n+l)/(\varepsilon n))^{n^2}$.

Drugie szacowanie z lematu wynika z pierwszego i niezależności L_m od K_m.

□

Twierdzenie 5.11. *Istnieje stała uniwersalna $c > 0$ taka, że $d_{\mathrm{BM}}(K_{2n}, L_{2n}) \geqslant cn$ z prawdopodobieństwem większym niż $1 - 2\exp(-n^2) > 0$.*

Dowód. Zauważmy, że dla dowolnych ciał wypukłych K i L,

$$\begin{aligned} d_{\mathrm{BM}}(K, L) &= \inf\{\|T\colon K \mapsto L\| \|T^{-1}\colon K \mapsto L\|\colon\ T \in \mathrm{GL}(n)\} \\ &= \inf\{\|T\colon K \mapsto L\| \|T^{-1}\colon K \mapsto L\|\colon\ T \in \mathrm{GL}(n), |\det T| = 1\}, \end{aligned}$$

stąd

$$\begin{aligned} &\mathbb{P}(d_{\mathrm{BM}}(K_{2n}, L_{2n}) < \frac{1}{4}\rho^2 n) \\ &\quad \leqslant \mathbb{P}(\|T\colon K_{2n} \mapsto L_{2n}\| \leqslant \frac{1}{2}\rho\sqrt{n} \text{ dla pewnego } T \in M_{n\times n}, |\det T| = 1) \\ &\qquad + \mathbb{P}(\|T^{-1}\colon L_{2n} \mapsto K_{2n}\| \leqslant \frac{1}{2}\rho\sqrt{n} \text{ dla pewnego } T \in M_{n\times n}, |\det T| = 1). \end{aligned}$$

Ponieważ $|\det T^{-1}| = |\det T|^{-1}$ oraz zbiory losowe K_{2n} i L_{2n} są niezależne i mają taki sam rozkład, więc ostatnie dwa prawdopodobieństwa są równe. Stąd przyjmując we Wniosku 5.10 $n = 2m$ oraz $\varepsilon = \rho/2$ dostajemy dla $\rho \in (0, 1)$

$$\mathbb{P}(d_{\mathrm{BM}}(K_{2n}, L_{2n}) < \frac{1}{4}\rho^2 n) \leqslant 2\Big(\frac{6a}{\rho}\Big)^{n^2} (3\alpha\rho)^{2n^2} = 2(54\rho a\alpha^2)^{n^2}$$

i wystarczy przyjąć $\rho := \min(1, (54ea\alpha^2)^{-1})$. □

Uwaga 5.12. Konstrukcja Głuskina jest losowa. Nie jest znana deterministyczna konstrukcja symetrycznych ciał wypukłych w $\mathbb{R}^n$, których odległość Banacha-Mazura jest większego rzędu niż $n^{1/2}$.

Uwaga 5.13. i) Przyjęcie $m = 2n$ w Twierdzeniu 5.11 było arbitralne. Modyfikując dowód można przyjąć, że $m = \delta n$ z $\delta > 0$, oczywiście stała c wówczas zależy od δ.
ii) W konstrukcji Głuskina wierzchołki losowych wielościanów wypukłych wybiera się korzystając z rozkładu jednostajnego na sferze. Można go zastąpić (odpowiednio znormalizowanym) innym „porządnym" rozkładem, np. gaussowskim, czy jednostajnym na kostce dyskretnej $\{-1, 1\}^n$ - to ostatnie wymaga jednak istotnych zmian w dowodzie.

6 Twierdzenie Dvoretzky'ego

W tej sekcji $g_1, g_2, \ldots$ oznaczają niezależne zmienne $\mathcal{N}(0, 1)$. Przypomnijmy, że jeśli $X = \sum_{i=1}^n v_i g_i$ jest scentrowanym wektorem gaussowskim w n-wymiarowej przestrzeni unormowanej $(\mathbb{R}^n, \| \ \|)$, to definiujemy

$$\begin{aligned} \sigma(X) = \sigma(X, \| \ \|) &= \sup_{\|u\|_* \leqslant 1} (\mathbb{E}|\langle u, X\rangle|^2)^{1/2} = \sup_{\|u\|_* \leqslant 1} (\sum_{i=1}^n |\langle u, v_i\rangle|^2)^{1/2} \\ &= \sup_{\|u\|_* \leqslant 1} \sup_{|t| \leqslant 1} \sum_{i=1}^n \langle u, v_i\rangle t_i = \sup_{|t| \leqslant 1} \sup_{\|u\|_* \leqslant 1} \langle u, \sum_{i=1}^n v_i t_i\rangle = \sup_{|t| \leqslant 1} \|\sum_{i=1}^n t_i v_i\|. \end{aligned}$$

Definicja 6.1. Dla wektora gaussowskiego X w przestrzeni unormowanej $(\mathbb{R}^n, \| \|)$ określamy

$$d(X) = d(X, \| \|) := \Big(\frac{\mathbb{E}\|X\|}{\sigma(X, \| \|)}\Big)^2.$$

Dla symetrycznego ciała wypukłego K w $\mathbb{R}^n$ definiujemy

$$d(K) := \sup\Big\{d(X, \| \|_K)\colon\ X \text{ scentrowany wektor gaussowski w } \mathbb{R}^n\Big\}.$$

Przykład. Biorąc $X = \sum_{i=1}^n e_i g_i$ widzimy, że $\mathbb{E}\|X\|_1 = n\mathbb{E}|g_1| = n(2/\pi)^{1/2}$ oraz $\sigma(X, \| \|_1) = \sup_{|t|=1} \|t\|_1 = \sqrt{n}$, zatem $d(B_1^n) \geqslant d(X) = \frac{2}{\pi}n$. Ogólniej, rozważając taki sam X, ale w przestrzeni l_p^n widzimy, że dla $1 \leqslant p \leqslant 2$, $\sigma(X, \| \|_p) = n^{1/p-1/2}$ oraz, na mocy Wniosku 3.4, $\mathbb{E}\|X\|_p \geqslant c(\mathbb{E}\|X\|_p^p)^{1/p} \geqslant c'n^{1/p}$. Wykazaliśmy zatem, że

$$d(B_p^n) \geqslant d(X) \geqslant cn \text{ dla } 1 \leqslant p \leqslant 2.$$

Dla $2 < p < \infty$ mamy $\sigma(X, \| \|_p) = 1$. By oszacować $\mathbb{E}\|X\|_p$ zauważmy, ponownie korzystając z Wniosku 3.4, że

$$\mathbb{E}\|X\| + C\sqrt{p} \geqslant (\mathbb{E}\|X\|_p^p)^{1/p} = n^{1/p}(\mathbb{E}|g_1|^p)^{1/p} \geqslant c_1\sqrt{p}n^{1/p},$$

czyli dla $\log n \geqslant C'p$ mamy $d(X) = (\mathbb{E}\|X\|_p)^2 \geqslant \frac{1}{2}c_1^2 pn^{2/p}$. Ponadto dla dowolnego $p \geqslant 2$, $\mathbb{E}\|X\|_p \geqslant \mathbb{E}\|X\|_\infty \geqslant c\log^{1/2}(n+1)$. W ten sposób pokazaliśmy, że

$$d(B_p^n) \geqslant d(X) \geqslant \begin{cases} cpn^{2/p} & \text{dla } 2 \leqslant p \leqslant \log(n+1) \\ c\log(n+1) & \text{dla } p \geqslant \log(n+1). \end{cases}$$

Możemy teraz sformułować twierdzenie Dvoretzky'ego w wersji pochodzącej od Vitaliego Milmana.

Twierdzenie 6.2. *Dla dowolnego $\varepsilon > 0$ istnieje stała $c(\varepsilon) > 0$ o następującej własności. Dla każdego symetrycznego ciała wypukłego K i $k \leqslant c(\varepsilon)d(K)$ istnieje podprzestrzeń liniowa V wymiaru k taka, że $d_{\mathrm{BM}}(K \cap V, B_2^k) \leqslant 1 + \varepsilon$.*

Lemat 6.3. *Niech A będzie ε-siecią w S^{n-1} dla pewnego $\varepsilon \in [0, 1]$. Wówczas dla dowolnych wektorów $v_1, \ldots, v_n$ w $\mathbb{R}^n$ i dowolnej normy $\| \|$,*

$$\sup_{|t|=1}\Big\|\sum_{i=1}^n v_i t_i\Big\| \leqslant \frac{1}{1-\varepsilon}\sup_{s\in A}\Big\|\sum_{i=1}^n v_i s_i\Big\|$$

oraz

$$\inf_{|t|=1}\Big\|\sum_{i=1}^n v_i t_i\Big\| \geqslant \inf_{s\in A}\Big\|\sum_{i=1}^n v_i s_i\Big\| - \frac{\varepsilon}{1-\varepsilon}\sup_{s\in A}\Big\|\sum_{i=1}^n v_i s_i\Big\|.$$

Dowód. Z definicji ε-sieci łatwo wynika, że dowolny wektor $t \in S^{n-1}$ można zapisać w postaci $t = s + au$, $s \in A$, $u \in S^{n-1}$, $a \leqslant \varepsilon$, stąd

$$\Big|\Big\|\sum_{i=1}^n v_i t_i\Big\| - \Big\|\sum_{i=1}^n v_i s_i\Big\|\Big| \leqslant a\Big\|\sum_{i=1}^n v_i u_i\Big\| \leqslant \varepsilon \sup_{u\in S^{n-1}}\Big\|\sum_{i=1}^n v_i u_i\Big\|.$$

$\square$

Dowód Twierdzenia 6.2. Bez straty ogólności możemy przyjąć, że $\varepsilon < 1$, dobierzmy $\delta > 0$ takie, że $(1+\delta)/(1-\delta) \leqslant \sqrt{1+\varepsilon}$ oraz $1-\delta-\delta(1+\delta)/(1-\delta) \geqslant 1/\sqrt{1+\varepsilon}$ (np. można przyjąć $\delta = \varepsilon/20$). Niech A będzie δ-siecią w S^{k-1} mocy nie większej niż $(\frac{3}{\delta})^k$. Zauważmy, iż wystarczy wykazać, że istnieją wektory $u_1, \ldots, u_k$ w $\mathbb{R}^n$ takie, że

$$1-\delta \leqslant \Big\| \sum_{i=1}^{k} t_i u_i \Big\|_K \leqslant 1+\delta \text{ dla } t \in A. \tag{9}$$

Istotnie wobec Lematu 6.3 nierówność (9) implikuje

$$\frac{1}{\sqrt{1+\varepsilon}} \leqslant \Big\| \sum_{i=1}^{k} x_i u_i \Big\|_K \leqslant \sqrt{1+\varepsilon} \text{ dla } x \in S^{k-1}.$$

Zatem, jeśli zdefiniujemy $T\colon \mathbb{R}^k \to V := \mathrm{Lin}(u_1, u_2, \ldots, u_k)$ wzorem $Tx = \sum_{i=1}^k x_i u_i$, to $\|T\colon B_2^k \mapsto K \cap V\|, \|T^{-1}\colon K\cap V \mapsto B_2^k\| \leqslant \sqrt{1+\varepsilon}$, czyli $d_{\mathrm{BM}}(B_2^k, K \cap V) \leqslant 1+\varepsilon$.

By wykazać istnienie $u_1, \ldots, u_k$ spełniających (9) wybierzmy $X = \sum_{i=1}^n v_i g_i$ spełniający $d(X) \geqslant d(K)/2$. Dla uproszczenia notacji przyjmijmy $\|\ \| = \|\ \|_K$, $d = d(K)$, $\sigma = \sigma(X)$ oraz $M = \mathbb{E}\|X\|$. Na podstawie Wniosku 3.3 dostajemy

$$\mathbb{P}\Big(\Big|\frac{\|X\|}{M} - 1\Big| \geqslant \delta\Big) = \mathbb{P}(|\|X\| - \mathbb{E}\|X\|| \geqslant \delta M) = 2\exp\Big(-\frac{2\delta^2 M^2}{\pi^2\sigma^2}\Big).$$

Niech $X_1, \ldots, X_k$ będą niezależnymi kopiami zmiennej X. Wówczas, dla dowolnego $t \in S^{n-1}$, zmienna $\sum_{i=1}^k t_i X_i$ ma ten sam rozkład co X. Zatem

$$\mathbb{P}\Big(\Big|\Big\|\sum_{i=1}^{k} t_i \frac{X_i}{M}\Big\| - 1\Big| \geqslant \delta\Big) = \mathbb{P}\Big(\Big|\frac{\|X\|}{M} - 1\Big| \geqslant \delta\Big) \leqslant 2e^{-2\pi^{-2}\delta^2 d(X)} \leqslant 2e^{-\delta^2 d/10},$$

czyli

$$\mathbb{P}\Big(\exists_{t\in A}\Big|\Big\|\sum_{i=1}^{k} t_i \frac{X_i}{M}\Big\| - 1\Big| \geqslant \delta\Big) \leqslant 2\#A e^{-\delta^2 d/10} \leqslant 2\Big(\frac{3}{\delta}\Big)^k e^{-\delta^2 d/10} < 1,$$

jeśli tylko $k \leqslant c(\delta)d$, czyli można za u_i przyjąć $X_i(\omega)/M$ dla pewnego ω. □

Uwaga 6.4. Powyższy dowód pokazuje, że można przyjąć $c(\varepsilon) = c\varepsilon^2/\log(3/\varepsilon)$ dla $\varepsilon \leqslant 1$.

By móc efektywnie korzystać z Twierdzenia 6.2 musimy umieć szacować z dołu $d(K)$. Zacznijmy od geometrycznego lematu.

Lemat 6.5 (Dvoretzky-Rogers). *Niech K będzie n-wymiarowym symetrycznym ciałem wypukłym. Wówczas istnieją wektory $v_1, \ldots, v_n$ takie, że $\|\sum_{i=1}^n t_i v_i\|_K \leqslant |t|$ dla $t \in \mathbb{R}^n$ oraz $\|v_i\|_K \geqslant \frac{n-i+1}{n}$ dla $1 \leqslant i \leqslant n$.*

Dowód. Zbiór przekształceń liniowych $T\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ spełniających $\|T\| := \|T\colon B_2^n \mapsto K\| \leqslant 1$ jest zwarty (w topologii $\mathbb{R}^{n^2}$), w szczególności istnieje operator T o normie jeden taki, że

$$\det(T) = \max\{\det(S)\colon \ \|S\colon B_2^n \to K\| \leqslant 1\}.$$

Wówczas dla dowolnego $\varepsilon > 0$,

$$\det(T + \varepsilon S) \leqslant \det(T)\|T + \varepsilon S\|^n.$$

Mamy jednak

$$\det(T + \varepsilon S) = \det(T)\det(I + \varepsilon T^{-1}S) = \det(T)(1 + \varepsilon \mathrm{tr}(T^{-1}S) + o(\varepsilon)),$$

więc

$$1 + \varepsilon \mathrm{tr}(T^{-1}S) \leqslant \|T + \varepsilon S\|^n + o(\varepsilon) \leqslant (\|T\| + \varepsilon\|S\|)^n + o(\varepsilon) \leqslant 1 + \varepsilon n\|S\| + o(\varepsilon).$$

Wykazaliśmy w ten sposób, że $\mathrm{tr}(T^{-1}S) \leqslant n\|S\|$ dla dowolnego przekształcenia liniowego S. Niech $S = TP$, gdzie P oznacza rzut ortogonalny na podrzestrzeń V. Wówczas

$$n\|TP\| \geqslant \mathrm{tr}(T^{-1}TP) = \mathrm{tr}(P) = \dim V.$$

Niech $y_1 \in \mathbb{R}^n$ będzie takie, że $|y_1| = 1$ i $\|Ty_1\|_K = \|T\|$. Jeśli wybraliśmy już $y_1, \ldots, y_k$, to kładziemy $V_k = \{y_1, \ldots, y_k\}^\perp$ i wybieramy $y_{k+1} \in V_k$ takie, że $|y_{k+1}| = 1$ oraz $\|Ty_{k+1}\|_K = \|TP_k\| \geqslant (n-k)/n$, gdzie P_k oznacza rzut ortogonalny na V_k. Wystarczy, że położymy $v_k = Ty_k$. □

Lemat 6.6. *Niech ε_i będą niezależnymi symetrycznymi zmiennymi losowymi przyjmującymi wartości ± 1. Wówczas dla dowolnych wektorów $u_i \in \mathbb{R}^n$ i dowolnej normy $\|\ \|$, $\mathbb{E}\|\sum_{i=1}^n u_i\varepsilon_i\| \geqslant \max_{1\leqslant i\leqslant n}\|u_i\|$.*

Dowód. Mamy dla dowolnego $1 \leqslant i \leqslant n$,

$$\begin{aligned}\mathbb{E}\Big\|\sum_{k=1}^n u_k\varepsilon_k\Big\| &= \mathbb{E}\frac{1}{2}\Big(\Big\|u_i\varepsilon_i + \sum_{k\neq i} u_k\varepsilon_k\Big\| + \Big\|u_i\varepsilon_i - \sum_{k\neq i} u_k\varepsilon_k\Big\|\Big)\\ &\geqslant \mathbb{E}\|u_i\varepsilon_i\| = \|u_i\|.\end{aligned}$$

□

Wniosek 6.7. *Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K, $d(K) \geqslant \frac{1}{C}\log(n+1)$, gdzie C jest stałą uniwersalną.*

Dowód. Niech v_i będą takie jak w Lemacie 6.5 oraz $X = \sum_{i=1}^n v_i g_i$. Wówczas $\sigma(X, \|\ \|_K) = 1$. Niech (ε_i) będzie ciągiem niezależnych symetrycznymi zmiennych losowych przyjmujących wartości ± 1, niezależnym od (g_i). Wówczas na mocy symetrii g_i,

$$\mathbb{E}\|X\|_K = \mathbb{E}\Big\|\sum_{i=1}^n \varepsilon_i v_i g_i\Big\|_K \geqslant \mathbb{E}\max_{i\leqslant n}\|v_i g_i\|_K \geqslant \frac{1}{2}\mathbb{E}\max_{i\leqslant\lceil n/2\rceil}|g_i| \geqslant c\sqrt{\log(n+1)}.$$

□

Wniosek 6.8. *Dla każdego symetrycznego ciała wypukłego K i $\varepsilon > 0$ istnieje podprzestrzeń liniowa V wymiaru $k \geqslant c(\varepsilon)\log(n+1)$, gdzie $c(\varepsilon)$ jest stałą zależną tylko od ε taka, że $d_{\mathrm{BM}}(K \cap V, B_2^k) \leqslant 1 + \varepsilon$.*

Definicja 6.9. Dla symetrycznego ciała wypukłego K i $\varepsilon > 0$ określmy

$$d_\varepsilon(K) := \sup\{k\colon\ d_{\mathrm{BM}}(V \cap K, B_2^k) \leqslant 1+\varepsilon \text{ dla pewnej podprzestrzeni } V \text{ wymiaru } k\}.$$

Uwaga 6.10. Twierdzenie Dvoretzky'ego mówi, że $d_\varepsilon(K) \geqslant c(\varepsilon)d(K)$, nietrudno udowodnić, że $d(K) \geqslant c'(\varepsilon)d_\varepsilon(K)$.

Poniższy fakt pokazuje w szczególności, że logarytmicznej zależności od wymiaru we Wniosku 6.8 nie można poprawić.

Fakt 6.11. *Mamy $d_\varepsilon(B_p^n) \leqslant C(1+\varepsilon)^2 pn^{2/p}$ dla $2 \leqslant p \leqslant \log(n+1)$ oraz $d_\varepsilon(B_p^n) \leqslant C(1+\varepsilon)^2\log(n+1)$ dla $p \geqslant \log(n+1)$.*

Dowód. Załóżmy napierw, że $p < \infty$ oraz wektory $u_i = (u_i(j))_{j\leqslant n} \in \mathbb{R}^n$, $1 \leqslant i \leqslant k$ spełniają

$$|t| \leqslant \Big\|\sum_{i=1}^k t_i u_i\Big\|_p \leqslant (1+\varepsilon)|t| \text{ dla } t \in \mathbb{R}^k.$$

Niech $X = \sum_{i=1}^k u_i g_i$, wówczas

$$k = \mathbb{E}|(g_1,\ldots,g_k)|^2 \leqslant \mathbb{E}\|X\|_p^2 \leqslant (\mathbb{E}\|X\|_p^p)^{2/p} = \Big(\sum_{j=1}^n \mathbb{E}\Big|\sum_{i=1}^k u_i(j)g_i\Big|^p\Big)^{2/p}$$
$$= \Big(\sum_{j=1}^n \mathbb{E}|g_1|^p\Big(\sum_{i=1}^k |u_i(j)|^2\Big)^{p/2}\Big)^{2/p}.$$

Zauważmy jednak, że dla $1 \leqslant j \leqslant n$,

$$\Big(\sum_{i=1}^k |u_i(j)|^2\Big)^{1/2} = \sup_{|t|=1}\Big|\sum_{i=1}^k t_i u_i(j)\Big| \leqslant \sup_{|t|=1}\Big\|\sum_{i=1}^k t_i u_i\Big\|_p \leqslant (1+\varepsilon).$$

Stąd otrzymujemy $d_\varepsilon(B_p^n) \leqslant C(1+\varepsilon)^2 pn^{2/p}$ dla $2 \leqslant p < \infty$. By otrzymać postulowane oszacowanie dla $\log(n+1) \leqslant p \leqslant \infty$ wystarczy zauważyć, że dla $q := \max(2, \log(n+1)) \leqslant p$ zachodzi $\|x\|_p \leqslant \|x\|_q \leqslant e\|x\|_p$. □

Fakt 6.11 oraz Przykład po Definicji 6.1 pokazują, że

$$d_\varepsilon(B_p^n) \sim_\varepsilon d(B_p^n) \sim \begin{cases} n & \text{dla } 1 \leqslant p \leqslant 2 \\ pn^{2/p} & \text{dla } 2 \leqslant p \leqslant \log(n+1) \\ \log(n+1) & \text{dla } p \geqslant \log(n+1). \end{cases}$$

7 Twierdzenie Kaszina-Szarka

Z twierdzenia Dvoretzky'ego wynika, że kula B_1^n ma prawie-kuliste przekroje wymiaru cn, gdzie c jest stałą uniwersalną. Jednak ani to twierdzenie ani metoda dowodu nie pozwalają nic wywnioskować o przekrojach wymiaru λn, gdy λ jest liczbą mniejszą od 1, ale nie bardzo bliską 0, np $\lambda = 1/2$. Kaszin znalazł metodę pokazującą, że takie przekroje też są bliskie euklidesowym (ze stałą zależną oczywiście od λ) i wykazał wiele ciekawych konsekwencji tego fenomenu. Nieco później Szarek zauważył, że w dowodzie Kaszina wykorzystuje się zasadniczo to, że elipsoida Johna wypełnia dużą część B_1^n i wprowadził pojęcie stosunku objętości.

Definicja 7.1. Dla n-wymiarowego ciała wypukłego K definiujemy *stosunek objetości* (ang. volume ratio) wzorem

$$\mathrm{vr}(K) := \left(\frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(\mathcal{E}_{\max})}\right)^{1/n},$$

gdzie $\mathcal{E}_{\max}$ oznacza elipsoidę maksymalnej objętości zawartą w K.

Uwaga 7.2. Oczywiście

$$\mathrm{vr}(K) = \inf\left\{\left(\frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(\mathcal{E})}\right)^{1/n} : \ \mathcal{E} \text{ elipsoida zawarta w } K\right\},$$

ponadto $\mathrm{vr}(K) = \mathrm{vr}(TK)$ dla $T \in \mathrm{GL}(n)$.

Przykład. Dla kuli B_1^n mamy $n^{-1/2}B_2^n \subset B_1^n$ stąd

$$\mathrm{vr}(B_1^n) \leqslant \left(\frac{\mathrm{vol}_n(B_1^n)}{\mathrm{vol}_n(n^{-1/2}B_2^n)}\right)^{1/n} = \left(\frac{2^n n^{n/2}\Gamma(\frac{n}{2}+1)}{n!\pi^{n/2}}\right)^{1/n} \leqslant \left(\frac{2e}{\pi}\right)^{1/2}.$$

Dla $1 < p < 2$ mamy $n^{-1/2}B_2^n \subset n^{1/p-1}B_p^n \subset B_1^n$, więc

$$\mathrm{vr}(B_p^n) = \mathrm{vr}(n^{1/p-1}B_p^n) \leqslant \mathrm{vr}(B_1^n) \leqslant \left(\frac{2e}{\pi}\right)^{1/2}.$$

Zanim sformułujemy główne twierdzenie tego wykładu, wprowadźmy pewne oznaczenia. Przez $G_{n,k}$ będziemy oznaczać rozmaitość Grassmana wszystkich k-wymiarowych podprzestrzeni $\mathbb{R}^n$ z odległością np. zadaną wzorem $\rho(V,V') = d_H(V \cap S^{n-1}, V' \cap S^{n-1})$. Można wykazać, że $G_{n,k}$ jest zwartą rozmaitością bez brzegu, ale nam wystarczy obserwacja, że jest to przestrzeń metryczna na której zwarta grupa metryczna $O(n)$ działa izometrycznie i tranzytywnie w naturalny sposób, tzn. $V \mapsto TV$ dla $T \in O(n)$. W związku z tym, jeśli m oznacza miarę Haara na $O(n)$, to wzór

$$\mu_{n,k}(\Gamma) = m(\{T \in O(n) \colon \ TV_0 \in \Gamma\})$$

nie zależy od wyboru $V_0 \in G_{n,k}$ i określa miarę probabilistyczną na $G_{n,k}$, niezmienniczą ze względu na działanie grupy $O(n)$.

Twierdzenie 7.3. *Załóżmy, że K jest symetrycznym ciałem wypukłym w $\mathbb{R}^n$ takim, że $B_2^n \subset K$ oraz $\mathrm{vol}_n(K) \leqslant A^n \mathrm{vol}_n(B_2^n)$ dla pewnej liczby $A \geqslant 1$. Wówczas dla $1 \leqslant k \leqslant n$ istnieje przestrzeń V wymiaru k taka, że*

$$B_2^n \cap V \subset K \cap V \subset (4\pi A)^{\frac{n}{n-k}} B_2^n \cap V.$$

Co więcej, jeśli Ω_0 jest zbiorem tych przestrzeni $V \in G_{n,k}$, które spełniają powyższą nierówność, to $\mu_{n,k}(\Omega_0) > 1 - 2^{-n}$.

Zanim przystąpimy do dowodu twierdzenia sformułujemy dwa proste lematy.

Lemat 7.4. *Dla dowolnej ograniczonej funkcji mierzalnej $f\colon S^{n-1} \to \mathbb{R}$,*

$$\int_{S^{n-1}} f(x) d\sigma_{n-1}(x) = \int_{G_{n,k}} \int_{S_V} f(x) d\sigma_V(x) d\mu_{n,k}(V),$$

gdzie $S_V = S^{n-1} \cap V$, a σ_V oznacza unormowaną miarę powierzchniową na S_V.

Dowód. Wzór

$$\mu(A) := \int_{G_{n,k}} \int_{S_V} I_A(x) d\sigma_V(x) d\mu_{n,k}(V)$$

określa miarę probabilistyczną na S^{n-1} niezmienniczą na działanie $O(n)$, czyli $\mu = \sigma_{n-1}$. □

Lemat 7.5. *Dla dowolnego $x \in S^{k-1}$, $k \geqslant 2$, oraz $0 < \delta < \sqrt{2}$ zachodzi*

$$\sigma_{k-1}(\{y \in S^{k-1}\colon\ |x-y| \leqslant \delta\}) > \left(\frac{\delta}{\pi}\right)^k.$$

Dowód. Zauważmy, że szukana miara to miara kuli w metryce geodezyjnej na S^{k-1} o środku w x i promieniu α spełniającym $\sin(\alpha/2) = \delta/2$, zatem wynosi ona $F(\alpha)/F(\pi)$, gdzie $F(s) = \int_0^s \sin^{k-2}(t) dt$. Oczywiście $F(\pi) \leqslant \pi$ oraz

$$\begin{aligned}(k-1)F(\alpha) &\geqslant (k-1)\int_0^s \sin^{k-2}(t)\cos(t) dt = \sin^{k-1}(\alpha) = \Big(2\sin\Big(\frac{\alpha}{2}\Big)\cos\Big(\frac{\alpha}{2}\Big)\Big)^{k-1} \\ &= \Big(\delta\Big(1 - \frac{\delta^2}{4}\Big)^{1/2}\Big)^{k-1} \geqslant \Big(\frac{\delta}{\sqrt{2}}\Big)^{k-1}.\end{aligned}$$

Zatem

$$\frac{F(\alpha)}{F(\pi)} \geqslant \frac{1}{\pi(k-1)}\Big(\frac{\delta}{\sqrt{2}}\Big)^{k-1} = \Big(\frac{\delta}{\pi}\Big)^k \Big(\frac{\pi}{\sqrt{2}}\Big)^{k-1} \frac{1}{\delta(k-1)} \geqslant \Big(\frac{\delta}{\pi}\Big)^k \frac{2^{k-1}}{\sqrt{2}(k-1)} > \Big(\frac{\delta}{\pi}\Big)^k.$$

□

Dowód Twierdzenia 7.3. Niech $\|x\| = \|x\|_K$, z założenia wiemy, że $\|x\| \leqslant |x|$. Przypomnijmy, że na mocy Faktu 5.4 zachodzi

$$A^n \geqslant \frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(B_2^n)} = \int_{S^{n-1}} \|x\|^{-n} d\sigma_{n-1}(x).$$

Określmy

$$\Omega_0 := \Big\{V \in G_{n,k} \colon \int_{S_V} \|x\|^{-n} d\sigma_V(x) \leqslant (2A)^n\Big\},$$

wówczas Lemat 7.4 oraz nierówność Czebyszewa implikują, że $\mu_{n,k}(G_{n,k} \setminus \Omega_0) < 2^{-n}$. Wystarczy, że pokażemy, że jeśli $V \in \Omega_0$, to $|x| \leqslant (4\pi A)^{n/(n-k)}\|x\|$, czyli z jednorodności norm, że

$$\|x\| \geqslant (4\pi A)^{-n/(n-k)} \text{ dla } x \in S_V,\ V \in \Omega_0. \tag{10}$$

Ustalmy $V \in \Omega_0$, wówczas z definicji Ω_0 i nierówności Czebyszewa wynika, że

$$\sigma_V(\{y \in S_V \colon \|y\| \leqslant r\}) \leqslant (2Ar)^n \text{ dla } r > 0.$$

Wybierzmy $r > 0$ tak, by $(2Ar)^n = (\frac{r}{2\pi})^k$, wówczas, na mocy Lematu 7.5 dla $\delta = r/2$, zbiór $L_r := \{y \in S_V \colon \|y\| \leqslant r\}$ nie zawiera kuli w S_V o promieniu $r/2$ w metryce euklidesowej. Wybierzmy $x \in S_V$, wówczas istnieje $y \in S_V \setminus L_r$ takie, że $|x - y| \leqslant r/2$, stąd

$$\|x\| \geqslant \|y\| - \|x - y\| \geqslant \|y\| - |x - y| \geqslant r - \frac{r}{2} = \frac{r}{2}.$$

Zauważmy, że

$$\Big(\frac{r}{2}\Big)^{n-k} = \frac{1}{4^n \pi^k A^n} \geqslant (4\pi A)^{-n}$$

i nierówność (10), a co za tym idzie całe twierdzenie, została udowodniona. □

Wniosek 7.6. *Dla $1 \leqslant k \leqslant n$ i dowolnego symetrycznego n-wymiarowego ciała wypukłego K istnieje podprzestrzeń V wymiaru k taka, że $d_{\mathrm{BM}}(K \cap V, B_2^k) \leqslant (4\pi \mathrm{vr}(K))^{n/(n-k)}$.*

W szczególności, korzystając z tego, że $\mathrm{vr}(B_1^n) \leqslant C$ dostajemy

Wniosek 7.7. *Dla $\lambda \in (0,1)$ istnieje stała $C(\lambda)$ taka, że dla dowolnego n można znaleźć podprzestrzeń V wymiaru $k \geqslant \lambda n$ dla której $d_{\mathrm{BM}}(B_1^n \cap V, B_2^k) \leqslant C(\lambda)$.*

Wniosek 7.8. *Załóżmy, że $n = 2k$, K jest symetrycznym, n-wymiarowym ciałem wypukłym takim, że $B_2^n \subset K$ oraz $\mathrm{vol}_n(K) \leqslant A^n \mathrm{vol}_n(B_2^n)$ dla pewnej liczby $A \geqslant 1$. Wówczas istnieją dwie ortogonalne podprzestrzenie E_1 i E_2 wymiaru k takie, że $\mathbb{R}^n = E_1 \oplus E_2$ oraz*

$$\frac{1}{(4\pi A)^2}|x| \leqslant \|x\|_K \leqslant |x| \ \text{dla } x \in E_1 \cup E_2.$$

Dowód. Zauważmy, że przekształcenie $V \mapsto V^\perp$ jest niezmiennicze względem miary $\mu_{n,k}$ dla $n = 2k$. Niech Ω_0 będzie jak w Twierdzeniu 7.3, wówczas

$$\begin{aligned}\mu_{n,k}(V \colon\ V, V^\perp \in \Omega_0) &\geqslant 1 - \mu_{n,k}(V \notin \Omega_0) - \mu_{n,k}(V^\perp \notin \Omega_0)\\ &= 1 - 2\mu_{n,k}(V \notin \Omega_0) > 1 - 2^{1-n} > 0.\end{aligned}$$

Wystarczy przyjąć $E_1 = V$, $E_2 = V^\perp$ dla $V \in G_{n,k}$ takiego, że $V, V^\perp \in \Omega_0$. □

Wniosek 7.9. *Załóżmy, że $n = 2k$, wówczas w przestrzeni l_2^n można znaleźć dwie ortogonalne k-wymiarowe podprzestrzenie E_1 i E_2 takie, że $l_2^n = E_1 \oplus E_2$ oraz*

$$\|x\|_1 \leqslant \sqrt{n}\|x\|_2 \leqslant C\|x\|_1 \ \text{dla } x \in E_1 \cup E_2,$$

gdzie $C = 32e\pi$.

Dowód. Stosujemy poprzedni wniosek do $K = n^{-1/2}B_2^n$ i $A = (2e/\pi)^{1/2}$ (zob. Przykład po Definicji 7.1). □

Wniosek 7.10. *Istnieje stała C i dwie ortogonalne podprzestrzenie $E_1, E_2 \subset L_2[0,1]$ takie, że $E_1 \oplus E_2 = L_2[0,1]$ oraz*

$$\|f\|_1 \leqslant \|f\|_2 \leqslant C\|f\|_1 \ \text{dla } f \in E_1 \cup E_2.$$

Dowód. Niech $(h_n)_{n\geqslant 1}$ będzie układem Haara na $[0,1]$, tzn. $h_0 = I_{[0,1]}$, $h_1 = I_{[0,1/2)} - I_{[1/2,1)}$ oraz $h_{2^m+k}(t) = 2^{m/2}h_2(2^m t - k) = 2^{m/2}(I_{[k2^{-m},(2k+1)2^{-m-1})}(t) - I_{[(2k+1)2^{-m-1},(k+1)2^{-m})}(t))$ dla $m = 1, 2, \ldots$ i $0 \leqslant k \leqslant 2^m - 1$. Wówczas h_n jest bazą ortonormalną $L_2[0,1]$. Określmy

$$X_0 := \mathrm{Lin}(h_0) \text{ i } X_m := \mathrm{Lin}(h_n\colon\ 2^m \leqslant n < 2^{m+1}) \text{ dla } m = 1, 2, \ldots.$$

Z rozłączności nośników odpowiednich funkcji Haara łatwo wynika, że dla $f = \sum_{k=0}^{2^m-1} a_k h_{2^m+k} \in X_m$, $\|f\|_2 = \|a\|_2$ i $\|f\|_1 = 2^{-m/2}\|a\|_1$. Stąd Wniosek 7.9 daje rozkład ortogonalny $X_m = X_{m,1} \oplus X_{m,2}$ taki, że

$$\|f\|_1 \leqslant \|f\|_2 \leqslant C_0\|f\|_1 \text{ dla } f \in X_{m,1} \cup X_{m,2} \tag{11}$$

dla pewnej stałej uniwersalnej C_0. Możemy dodatkowo położyć (z zachowaniem oszacowania (11)) $X_{0,1} := X_0$, $X_{1,1} := 0$. Zdefiniujmy $E_1 := \oplus_{m\geqslant 0} X_{m,1}$, $E_2 := \oplus_{m\geqslant 0} X_{m,2}$. Oczywiście przestrzenie E_1 i E_2 są ortogonalne. By pokazać porównywalność norm, oznaczmy przez P_m rzut ortogonalny z $L_2[0,1]$ na X_m i ustalmy $1 < p < 2$, np. $p = 3/2$. Ponieważ układ Haara jest bazą bezwarunkową L_p oraz L_p ma kotyp 2, więc istnieje stała $C_p < \infty$ taka, że

$$\Big(\sum_{m\geqslant 0}\|P_m f\|_p^2\Big)^{1/2} \leqslant C_p\|f\|_p \text{ dla } f \in L_2[0,1] \subset L_p[0,1].$$

Stąd dla $f \in E_1 \cup E_2$, wobec (11)

$$\|f\|_2 = (\sum_{m\geqslant 0}\|P_m f\|_2^2)^{1/2} \leqslant C_0(\sum_{m\geqslant 0}\|P_m f\|_1^2)^{1/2} \leqslant C_0(\sum_{m\geqslant 0}\|P_m f\|_p^2)^{1/2} \leqslant C_0 C_p\|f\|_p.$$

Ale, jeśli $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2}$, to z nierówności Höldera $\|f\|_p \leqslant \|f\|_1^{1-\theta}\|f\|_2^{\theta}$ i

$$\|f\|_2 \leqslant (C_0 C_p)^{1/(1-\theta)}\|f\|_1 \text{ dla } f \in E_1 \cup E_2.$$

□

Uwaga 7.11. Ostatni wniosek można przeformułować w następujący sposób - istnieje baza o.n. $(f_n)_{n\geqslant 1}$ przestrzeni $L_2[0,1]$ taka, że na przestrzeniach $\overline{\mathrm{Lin}(f_{2k} : k \geqslant 1)}$ i $\overline{\mathrm{Lin}(f_{2k-1} : k \geqslant 1)}$ normy L_1 i L_2 są porównywalne. Nie jest znana żadna deterministyczna konstrukcja takiej bazy.

8 Elipsoidy Milmana

Definicja 8.1. Dla symetrycznych ciał wypukłych K, L zdefiniujmy liczbę

$$M(K,L) := \left(\frac{\mathrm{vol}_n(K+L)}{\mathrm{vol}_n(K \cap L)} \cdot \frac{\mathrm{vol}_n(K^0 + L^0)}{\mathrm{vol}_n(K^0 \cap L^0)} \right)^{1/n}.$$

Oczywiście $M(K,L) \geqslant 1$, niewielka liczba $M(K,L)$ mówi, że ciała K i L oraz ich polary są w „podobnym” położeniu.

Uwaga 8.2. i) Być może nieco naturalniej jest zdefiniować

$$\tilde{M}(K,L) := \left(\frac{\mathrm{vol}_n(\mathrm{conv}(K,L))}{\mathrm{vol}_n(K \cap L)} \cdot \frac{\mathrm{vol}_n(\mathrm{conv}(K^0, L^0))}{\mathrm{vol}_n(K^0 \cap L^0)} \right)^{1/n}.$$

Ponieważ $\mathrm{conv}(K,L) \subset K + L \subset 2\mathrm{conv}(K,L)$, więc $\tilde{M}(K,L) \leqslant M(K,L) \leqslant 4\tilde{M}(K,L)$. Ponadto $\tilde{M}(K,L) \geqslant 1$ i równość zachodzi wtedy i tylko wtedy, gdy $K = L$.
ii) Zauważmy, że jeśli $L \subset K \subset dL$, to $M(K,L) \leqslant (1+d)^2$. W szczególności dla dowolnego ciała wypukłego K istnieje elipsoida $\mathcal{E}$ taka, że $M(K,\mathcal{E}) \leqslant (1 + d_{\mathrm{BM}}(K, B_2^n))^2$. Nierówność tę można jednak znacząco wzmocnić.

Twierdzenie 8.3 (Milman)**.** *Dla dowolnego symetrycznego n-wymiarowego ciała wypukłego K istnieje elipsoida $\mathcal{E}$ taka, że $M(K,\mathcal{E}) \leqslant C$, gdzie C jest pewną stałą uniwersalną.*

Elipsoidę spełniającą warunek $M(K,\mathcal{E}) \leqslant C$ będziemy nazywać *elipsoidą Milmana* dla ciała K. Oczywiście taka elipsoida (w przeciwieństwie np. do elipsoidy Johna) nie jest wyznaczona jednoznacznie.

Podczas tego wykładu, jak zawsze, $g_1, g_2, \ldots$ jest ciągiem niezależnych zmiennych losowych $\mathcal{N}(0,1)$. Przez c, C, c_1, C_1 itp. będziemy oznaczać stałe uniwersalne, przy czym stałe C i c mogą miec inną wartość przy każdym wystąpieniu, a wartość stałych „numerowanych” będziemy się starali trzymać (dla wygody Czytelnika) ustaloną.

8.1 Minoryzacja Sudakowa

Zacznijmy od sformułowania ważnego twierdzenia, zwanego zasadą minoryzacyjną Sudakowa, które pozwala szacować pewne liczby pokryciowe przez średnią supremum odpowiedniego procesu gaussowskiego.

Twierdzenie 8.4. *Istnieje stała $C_1 \leqslant 64$ taka, że dla dowolnego wypukłego, symetrycznego n-wymiarowego ciała wypukłego K*

$$\sup_{\varepsilon > 0} \varepsilon \sqrt{\log N(K, \varepsilon B_2^n)} \leqslant C_1 \mathbb{E} \sup_{t \in K} \sum_{i=1}^n t_i g_i.$$

Udowodnimy najpierw dualną wersję tej nierówności:

Twierdzenie 8.5. *Istnieje stała $C_2 \leqslant 8$ taka, że dla dowolnego wypukłego, symetrycznego n-wymiarowego ciała wypukłego K,*

$$\sup_{\varepsilon>0} \varepsilon\sqrt{\log N(B_2^n, \varepsilon K^0)} \leqslant C_2 \mathbb{E}\sup_{t\in K}\sum_{i=1}^n t_i g_i.$$

Dowód. Dla uproszczenia notacji niech $a := \mathbb{E}\sup_{t\in K}\sum_{i=1}^n t_i g_i$. Zauważmy, że jeśli $x \notin uK^0$, to $\sup_{t\in K}\sum_{i=1}^n t_i x_i \geqslant u$, zatem, na mocy nierówności Czebyszewa,

$$\gamma_n(4aK^0) = 1 - \gamma_n(\mathbb{R}^n \setminus 4aK^0) \geqslant \frac{3}{4}.$$

Załóżmy, że $N = N(B_2^n, \varepsilon K^0) = N(\frac{8a}{\varepsilon}B_2^n, 8aK^0)$, wtedy istnieją $x_1, \ldots, x_N \in \frac{8a}{\varepsilon}B_2^N$ takie, że $x_i - x_j \notin 8aK^0$ dla $i \neq j$, czyli zbiory $x_i + 4aK^0$, $1 \leqslant i \leqslant N$ są rozłączne. Mamy

$$\begin{aligned}1 &\geqslant \gamma_n\Big(\bigcup_{i=1}^N (x_i + 4aK^0)\Big) = \sum_{i=1}^N \gamma_n(x_i + 4aK^0) \geqslant \sum_{i=1}^N e^{-|x_i|^2/2}\gamma_n(4aK^0)\\ &\geqslant \frac{3}{4}Ne^{-32a^2\varepsilon^{-2}},\end{aligned}$$

gdzie druga nierówność wynika z lematu sformułowanego poniżej. Stąd $\frac{3}{4}N \leqslant \exp(32a^2\varepsilon^{-2})$ i $\varepsilon\sqrt{\log N} \leqslant 8a$ (bo albo $N = 1$ albo $\frac{3}{4}N \geqslant \sqrt{N}$). □

Lemat 8.6. *Dla dowolnego $z \in \mathbb{R}^n$ i symetrycznego zbioru borelowskiego A, $\gamma_n(z + A) \geqslant e^{-|z|^2/2}\gamma_n(A)$.*

Dowód. Mamy

$$\begin{aligned}\gamma_n(A+z) &= (2\pi)^{-n/2}\int_{z+A} e^{-|x|^2/2}dx = (2\pi)^{-n/2}\int_A e^{-|x-z|^2/2}dx\\ &= e^{-|z|^2/2}\int_A e^{\langle x,z\rangle}d\gamma_n(x) = e^{-|z|^2/2}\int_A \frac{1}{2}(e^{\langle x,z\rangle} + e^{-\langle x,z\rangle})d\gamma_n(x)\\ &\geqslant e^{-|z|^2/2}\int_A d\gamma_n(x) = e^{-|z|^2/2}\gamma_n(A),\end{aligned}$$

gdzie w przedostatniej równości wykorzystaliśmy symetrię miary γ_n i zbioru A. □

Twierdzenie 8.4 jest konsekwencją Twierdzenia 8.5 i następującego lematu:

Lemat 8.7. *Dla dowolnego symetrycznego ciała wypukłego K w $\mathbb{R}^n$,*

$$\sup_{\varepsilon>0}\varepsilon\sqrt{\log N(K, \varepsilon B_2^n)} \leqslant 8\sup_{\varepsilon>0}\varepsilon\sqrt{\log N(B_2^n, \varepsilon K^0)}.$$

Dowód. Najpierw wykażemy, że dla symetrycznego ciała K,

$$N(K,L) = N(K, 2K \cap L). \tag{12}$$

Nierówność " $\leqslant$ " wynika stąd, że $2K \cap L \subset L$. By udowodnić " $\geqslant$ " wystarczy zauważyć, że dla $x \in K$, $(x+L) \cap K \subset x + (2K \cap L)$.

Jeśli $x \in 2K \cap \frac{\varepsilon^2}{2} K^0$, to $|x|^2 = \langle x, x\rangle \leqslant 2\frac{\varepsilon^2}{2} = \varepsilon^2$, co pokazuje, że $2K \cap \frac{\varepsilon^2}{2} K^0 \subset \varepsilon B_2^n$, stąd

$$N(K, \varepsilon B_2^n) \leqslant N\Big(K, 2K \cap \frac{\varepsilon^2}{2} K^0\Big) \leqslant N\Big(K, \frac{\varepsilon^2}{2} K^0\Big) \leqslant N(K, 2\varepsilon B_2^n) N\Big(B_2^n, \frac{\varepsilon}{4} K^0\Big).$$

Zatem

$$\varepsilon\sqrt{\log N(K, \varepsilon B_2^n)} \leqslant \varepsilon\sqrt{\log N(K, 2\varepsilon B_2^n)} + \varepsilon\sqrt{\log N\Big(B_2^n, \frac{\varepsilon}{4} K^0\Big)}$$

Przyjmując $f(\varepsilon) := \varepsilon\sqrt{\log N(K, \varepsilon B_2^n)}$ i $M := \sup_{\varepsilon>0} \varepsilon \sqrt{\log N(B_2^n, \varepsilon K^0)}$ dostajemy $f(\varepsilon) \leqslant \frac{1}{2} f(2\varepsilon) + 4M$ i przez łatwą iterację (wobec $f(\varepsilon) = 0$ dla dużych ε) otrzymujemy $f(\varepsilon) \leqslant 8M$. $\square$

Uwaga 8.8. Otwartym problemem jest czy istnieją stałe absolutne c, C takie, że dla dowolnych symetrycznych ciał wypukłych $\log N(K,L) \leqslant C \log N(L^0, cK^0)$. Wiadomo, że tak jest w wielu szczególnych przypadkach, m.in. gdy K lub L jest elipsoidą (co jest oczywiście znacznym wzmocnieniem Lematu 8.7).

Wprowadźmy teraz następujące definicje, które uproszczą sformułowania wielu wyników.

Definicja 8.9. Dla operatora liniowego $T\colon \mathbb{R}^n \to X$, gdzie $(X, \|\ \|)$ jest przestrzenią Banacha określmy

$$l(T) := \Big(\mathbb{E}\Big\|\sum_{i=1}^n g_i T(e_i)\Big\|^2\Big)^{1/2} = \Big(\int_{\mathbb{R}^n} \|Tx\|^2 d\gamma_n(x)\Big)^{1/2}.$$

Definicja 8.10. Dla przekształcenia liniowego $T\colon X \to Y$, gdzie X i Y są przestrzeniami Banacha z kulami jednostkowymi B_X i B_Y definiujemy *liczby entropijne* $e_k(T)$, $k \geqslant 1$ wzorem

$$e_k(T) := \inf\{\varepsilon > 0\colon\ N(T(B_X), \varepsilon B_Y) \leqslant 2^{k-1}\}.$$

Uwaga 8.11. Zauważmy, że $\|T\| = e_1(T) \geqslant e_2(T) \geqslant \ldots$. Ponadto, jeśli $i\colon Y \to Y_1$ jest izometrycznym włożeniem, to $\frac{1}{2} e_k(T) \leqslant e_k(iT) \leqslant e_k(T)$, a jeśli q jest rzutem z X' na $X = X'/X_0$ to $e_k(Tq) = e_k(T)$.

Twierdzenia 8.4 i 8.5 implikują:

Wniosek 8.12. *Dla dowolnego przekształcenia liniowego $T\colon l_2^n \to X$ zachodzi*

$$\sqrt{k} \max\{e_k(T), e_k(T^*)\} \leqslant C_3 l(T).$$

Dowód. Stosując Twierdzenie 8.4 do $K = T^*B_{X^*}$ dostajemy

$$\varepsilon\sqrt{\log N(T^*(B_{X^*}), \varepsilon B_2^n)} \leqslant C_1 \int \sup_{y \in T^* B_X^*} \langle y, x\rangle d\gamma_n(x) = C_1 \int \|Tx\| d\gamma_n(x).$$

Ponieważ dla $\varepsilon < e_k(T^*)$, $N(T^*(B_{X^*}), \varepsilon B_2^n) \geqslant 2^{k-1} + 1 \geqslant 2^{k/2}$, więc powyższe szacowanie implikuje $\sqrt{k} e_k(T^*) \leqslant \sqrt{2/\log 2} C_1 l(T)$.

By udowodnić drugie załóżmy najpierw, że T jest izomorfizmem. Wówczas zastosujmy Twierdzenie 8.5 do K jak wyżej zauważając, że $(T^*B_{X^*})^0 = T^{-1}(B_X)$ i otrzymamy

$$\varepsilon\sqrt{\log N(T(B_2^n), \varepsilon B_X)} = \varepsilon\sqrt{\log N(B_2^n, \varepsilon(T^*B_{X^*})^0)} \leqslant C_2 \int \sup_{y \in T^* B_X^*} \langle y, x\rangle d\gamma_n(x)$$
$$= C_2 \int \|Tx\| d\gamma_n(x).$$

i jak powyżej dostajemy $\sqrt{k} e_k(T) \leqslant \sqrt{2/\log 2} C_2 l(T)$. Jeśli T nie jest izomorfizmem, to $T = iT_1 q$, gdzie q jest rzutem na $l_2^n/\mathrm{Ker}(T) \simeq l_2^m$, T_1 izomorfizmem między l_2^m a $T(l_2^n)$, a i włożeniem $T(l_2^n)$ w X. Wówczas $l(T) = l(T_1)$ oraz $e_n(T) \leqslant 2e_n(T_1)$. □

Zakończmy tę część faktem pokazującym związek oszacowań liczb entropijnych i $M(K, \mathcal{E})$.

Fakt 8.13. *Dla dowolnego odwracalnego przekształcenia $T\colon l_2^n \to (\mathbb{R}^n, \| \ \|_K)$ mamy*

$$M(K, T(B_2^n)) \leqslant 16(1 + e_n(T^{-1}))(2 + e_n(T))(1 + e_n(T^*))(2 + e_n((T^*)^{-1})).$$

Dowód. Niech $\mathcal{E} = T(B_2^n)$, zauważmy, że wtedy $(T^*)^{-1}(B_2^n) = \mathcal{E}^0$, stąd

i) $N(\mathcal{E}, e_n(T)K) = N(T(B_2^n), e_n(T)K) \leqslant 2^{n-1}$,

ii) $N(K, e_n(T^{-1})\mathcal{E}) = N(T^{-1}(K), e_n(T^{-1})B_2^n) \leqslant 2^{n-1}$,

iii) $N(\mathcal{E}^0, e_n((T^*)^{-1})K^0) = N((T^*)^{-1}(B_2^n), e_n((T^*)^{-1})K^0) \leqslant 2^{n-1}$,

iv) $N(K^0, e_n(T^*)\mathcal{E}^0) = N(T^*(K^0), e_n(T^*)B_2^n) \leqslant 2^{n-1}$.

Na mocy ii) mamy

$$\mathrm{vol}(\mathcal{E} + K) \leqslant N(K, e_n(T^{-1})\mathcal{E})\mathrm{vol}(\mathcal{E} + e_n(T^{-1})\mathcal{E}) \leqslant 2^{n-1}(1 + e_n(T^{-1}))^n \mathrm{vol}(\mathcal{E}).$$

Ponieważ dla $a > 0$, $N(\mathcal{E}, aK) = N(\mathcal{E}, aK \cap 2\mathcal{E})$ i $aK \cap 2\mathcal{E} \subset (a+2)(K \cap \mathcal{E})$, więc i) implikuje

$$\mathrm{vol}(\mathcal{E}) \leqslant N(\mathcal{E}, e_n(T)K \cap 2\mathcal{E})\mathrm{vol}(e_n(T)K \cap 2\mathcal{E}) \leqslant 2^{n-1}(e_n(T) + 2)^n \mathrm{vol}(K \cap \mathcal{E}),$$

zatem

$$\left(\frac{\mathrm{vol}(K + \mathcal{E})}{\mathrm{vol}(K \cap \mathcal{E})}\right)^{1/n} \leqslant 4(1 + e_n(T^{-1}))(2 + e_n(T)).$$

Analogicznie z iii) i iv) wynika, że

$$\left(\frac{\mathrm{vol}(K^0+\mathcal{E}^0)}{\mathrm{vol}(K^0\cap\mathcal{E}^0)}\right)^{1/n} \leqslant 4(1+e_n(T^*))(2+e_n((T^*)^{-1})).$$

$\square$

8.2 Twierdzenie Lewisa

By wykazać twierdzenie Milmana spróbujemy znaleźć operator T między l_2^n a przestrzenią $(\mathbb{R}^n, \|\ \|_K)$ taki, że zarówno $l(T)$, jak i $l((T^*)^{-1})$ są małe. Pierwszym krokiem w tym kierunku będzie twierdzenie Lewisa.

Załóżmy, że α jest normą na przestrzeni macierzy losowych $M_{n\times n}$ (którą utożsamiamy z przestrzenią przekształceń liniowych na $\mathbb{R}^n$). Określmy normę α^* na $M_{n\times n}$ śladowo-dualną do α wzorem

$$\alpha^*(T) = \sup\{\mathrm{tr}(TS)\colon \alpha(S)\leqslant 1\}.$$

Zauważmy, że z definicji wynika, że $|\mathrm{tr}(TS)| \leqslant \alpha(T)\alpha^*(S)$, w szczególności $\alpha(T)\alpha^*(T^{-1}) \geqslant n$. Poniższe twierdzenie, pochodzące od Lewisa, pokazuje, że istnieje operator dla którego zachodzi w tym oszacowaniu równość.

Twierdzenie 8.14. *Dla dowolnej normy α na $M_{n\times n}$ istnieje przekształcenie odwracalne T takie, że $\alpha(T)=1$ i $\alpha^*(T^{-1})=n$.*

Dowód. Zbiór $K=\{S\in M_{n\times n}\colon \alpha(S)\leqslant 1\}$ jest zwarty, wyznacznik jest ciągłą funkcją na K, więc istnieje $T\in K$ takie, że

$$|\det(T)| = \sup\{|\det(S)|\colon\ \alpha(S)\leqslant 1\}.$$

Stąd dla dowolnego $S\in M_{n\times n}$,

$$|\det(T+S)| = \left|\det\left(\frac{T+S}{\alpha(T+S)}\right)\right|\alpha^n(T+S) \leqslant |\det(T)|\alpha^n(T+S).$$

Oczywiście T ma niezerowy wyznacznik, więc dzieląc powyższą nierówność stronami przez $\det(T)$ i zastępując S przez εS dostajemy dla $S\in M_{n\times n}$,

$$|\det(I+\varepsilon T^{-1}S)| \leqslant \alpha^n(T+\varepsilon S) \leqslant (\alpha(T)+\varepsilon\alpha(S))^n \text{ dla wszystkich } \varepsilon>0.$$

Ponieważ $\det(I+\varepsilon T^{-1}S) = 1+\varepsilon\mathrm{tr}(T^{-1}S)+o(\varepsilon)$ przy $\varepsilon\to 0$, więc $\mathrm{tr}(T^{-1}S)\leqslant n\alpha(S)$ dla wszystkich macierzy S, skąd $\alpha^*(T^{-1})\leqslant n$ na mocy definicji α^*. Ponieważ $\alpha^*(T^{-1})\alpha(T)\geqslant n$, więc musi zachodzić $\alpha(T)=1$ i $\alpha^*(T^{-1})=n$. $\square$

8.3 Porównywanie $l(T^*)$ z $l^*(T)$

Twierdzenie Lewisa pozwala nam znaleźć operator T taki, że $l(T)=l^*(T^{-1})=\sqrt{n}$. By przełożyć ostatnią równość na oszacowanie $l((T^{-1})^*)$ chcielibyśmy wiedzieć, że dla operatorów $T\colon X\to l_2^n$, $l(T^*)\leqslant C_X l^*(T)$, gdzie C_X jest stałą zależną tylko od przestrzeni X.

Spróbujmy zatem udowodnić takie oszacowanie, w tym celu ustalmy $T\colon X \to l_2^n$, gdzie X jest n wymiarową przestrzenią Banacha, wówczas $T^*\colon l_2^n \to X^*$ i jeśli utożsamimy X i X^* z $\mathbb{R}^n$ a działanie funkcjonału na wektorze z braniem iloczynu skalarnego, to

$$\begin{aligned} l(T^*) &= \Big(\int_{\mathbb{R}^n} \|T^*x\|^2 d\gamma_n\Big)^{1/2} \\ &= \sup\Big\{\int_{\mathbb{R}^n} \langle T^*x, F(x)\rangle d\gamma_n\colon\ F\colon \mathbb{R}^n \to X, \int_{\mathbb{R}^n} \|F(x)\|^2 d\gamma_n \leqslant 1\Big\}. \end{aligned}$$

Przekształcenie F, które wybija supremum nie musi być liniowym odwzorowaniem. Wydzielmy więc „liniową część” F kładąc

$$PF(x) = \sum_{i=1}^n x_i \int_{\mathbb{R}^n} F(y) y_i d\gamma_n(y). \tag{13}$$

Wówczas $PF\colon \mathbb{R}^n \to X$ liniowe oraz dla dowolnego j

$$\int_{\mathbb{R}^n} PF(x) x_j d\gamma_n(x) = \sum_{i=1}^n \int_{\mathbb{R}^n} F(y) y_i d\gamma_n(y) \int_{\mathbb{R}_n} x_i x_j d\gamma_n(x) = \int_{\mathbb{R}^n} F(x) x_j d\gamma_n(x),$$

czyli

$$\begin{aligned} \int_{\mathbb{R}^n} \langle T^*x, F(x)\rangle d\gamma_n = \int_{\mathbb{R}^n} \langle T^*x, PF(x)\rangle d\gamma_n &= \int_{\mathbb{R}^n} \langle x, TPF(x)\rangle d\gamma_n \\ = \operatorname{tr}(T(PF)) &\leqslant l^*(T) l(PF). \end{aligned}$$

Wystarczy więc umieć szacować $l(PF)$ poprzez $(\int_{\mathbb{R}^n} \|F(X)\|^2 d\gamma_n)^{1/2}$, czyli normę operatora P na przestrzeni

$$L_2(X) = L_2(\gamma_n, X) = \Big\{f\colon \mathbb{R}^n \to X \text{ mierzalne}, \|f\|_{L_2(X)} := \int_{\mathbb{R}^n} \|f\|_X^2 d\gamma_n < \infty\Big\}.$$

W ten sposób udowodniliśmy następujący fakt.

Fakt 8.15. *Jeśli T jest przekształceniem liniowym między n-wymiarową przestrzenią unormowaną X a l_2^n, to*

$$l(T^*) \leqslant \|P\colon L_2(X) \to L_2(X)\| l^*(T),$$

gdzie P jest przekształceniem zadanym wzorem (13).

Zanim przejdziemy do oszacowań normy P na $L_2(X)$ wprowadzimy trochę notacji.

Wielomiany Hermite'a

$$h_k(x) = \frac{(-1)^n}{\sqrt{n!}} e^{x^2/2} \frac{d^n}{dx^n}(e^{-x^2/2}),\ k = 0, 1, 2, \ldots$$

stanowią bazę ortogonalną przestrzeni $L_2(\mathbb{R},\gamma_1)$. Zauważmy, że w szczególności $h_1(x)=x$. Bazą ortogonalną $L_2(\mathbb{R}^n,\gamma_n)$ są n-wymiarowe wielomiany Hermite'a, tzn. układ

$$h_\alpha(x)=\prod_{i=1}^n h_{\alpha_i}(x_i),\ \alpha=(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}^n.$$

Przez Q_k oznacza rzut $L_2(\mathbb{R}^n,\gamma_n)$ na $\mathrm{Lin}(h_\alpha\colon\ |\alpha|=k)$. W szczególności Q_1 jest rzutem na $\mathrm{Lin}(x_1,\ldots,x_n)$.

Dla przestzeni Banacha X definiujemy

$$L_2\otimes X:=\Big\{F\in L_2(X)\colon\ F=\sum_{i=1}^k y_if_i,\ f_i\in L_2(\gamma_n),y_i\in X\Big\}.$$

Jeśli X jest skończenie wymiarowe, to $L_2\otimes X=L_2(X)$, a jeśli X jest ośrodkowe, to $L_2\otimes X$ jest gęste w $L_2(X)$. Dla operatora T na $L_2(\gamma_n)$ określamy operator $T\otimes\mathrm{Id}$ wzorem

$$T\otimes\mathrm{Id}\Big(\sum_{i=1}^k y_if_i\Big)=\sum_{i=1}^k y_iTf_i.$$

Mamy

$$\varphi(T\otimes\mathrm{Id}f)=T\varphi(f)\text{ dla }\varphi\in X^*,\tag{14}$$

co implikuje, że operator $T\otimes\mathrm{Id}$ jest dobrze określony. Nietrudno sprawdzić, że operator P zadany przez (13) to $Q_1\otimes\mathrm{Id}$.

Mamy następujące łatwe oszacowanie

Lemat 8.16. *Jeśli X jest n-wymiarową przestrzenią Banacha, to*

$$\|Q_k\otimes\mathrm{Id}\colon L_2(X)\to L_2(X)\|\leqslant d_{BM}(X,l_2^n),\quad k=0,1,2,\ldots.$$

Dowód. Wystarczy wybrać bazę e_i przestrzeni X taką, że $|t|\leqslant\|\sum_{i=1}^n t_ie_i\|\leqslant d|t|$, gdzie $d=d_{\mathrm{BM}}(x,l_2^n)$ i zauważyć, że $Q_k\otimes\mathrm{Id}$ jest rzutem na $\mathrm{Lin}(e_ih_\alpha\colon\ |\alpha|=k)$. □

Uwaga 8.17. Jeśli X jest przestrzenią Banacha, to supremum po $n=1,2,\ldots$ norm $\|Q_1\otimes\mathrm{Id}\|$ na $L_2(\gamma_n,X)$ nazywamy stałą K-wypukłości X. Mówimy, że X jest K-wypukłe, jeśli ta stała jest skończona. Twierdzenie Pisiera mówi, że przestrzeń X jest K-wypukła wtedy i tylko wtedy, gdy nie zawiera jednostajnie l_1^n.

Głównym wynikiem tej części jest następujące twierdzenie.

Twierdzenie 8.18. *Jeśli X jest n wymiarową przestrzenią Banacha, to*

$$\|Q_1\otimes\mathrm{Id}\colon L_2(X)\to L_2(X)\|\leqslant C_4(1+\log d_{BM}(X,l_2^n)).$$

Co wobec Faktu 8.15 implikuje

Wniosek 8.19. *Jeśli T jest przekształceniem liniowym między $(\mathbb{R}^n,\|\ \|_K)$ a l_2^n, to*

$$l(T^*)\leqslant C_4(1+\log d_{BM}(K,B_2^n))l^*(T).$$

W dowodzie Twierdzenia 8.18 wykorzystamy *półgrupę Ornsteina-Uhlebecka* określoną wzorem

$$P_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + (1-e^{-2t})^{1/2}y)d\gamma_n(y) \quad t \geqslant 0, f \in L_2(\gamma_n).$$

Łatwo sprawdzić, że istotnie jest to półgrupa, tzn. $P_0 f = f$, $P_{t+s} = P_t P_s$, ponadto P_t są operatorami samosprzężonymi.

Zauważmy, że dla $n = 1$, $P_t h_k$ jest wielomianem stopnia k, ponadto dla $l < k$, $\langle P_t h_k, h_l \rangle = \langle h_k, P_t h_l \rangle = 0$, gdyż h_k jest prostopadły do wszystkich wielomianów stopnia mniejszego niż k. Stąd $P_t h_k = c(t,k)h_k$ i z porównania współczynników przy x^k dostajemy $P_t h_k = e^{-t|k|}h_k$. Stąd łatwo dostajemy, że w przypadku dowolnego n, $P_t h_\alpha = e^{-t|\alpha|}h_\alpha$, czyli innymi słowy

$$P_t f = \sum_{k=0}^{\infty} e^{-tk} Q_k f.$$

Dla uproszczenia notacji określmy też dla $\varepsilon \in [-1,1]$ operator $Q(\varepsilon)$ wzorem

$$Q(\varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k Q_k.$$

Fakt 8.20. *Mamy dla $f \in L_2(\gamma_n)$,*

$$Q(\varepsilon)f(x) = \int_{\mathbb{R}^n} f(\varepsilon x + (1-\varepsilon^2)^{1/2}y)d\gamma_n(y), \quad \varepsilon \in [-1,1].$$

Ogólniej, dla dowolnej przestrzeni Banacha X i $F \in L_2 \otimes X$,

$$Q(\varepsilon) \otimes \mathrm{Id}F(x) = \int_{\mathbb{R}^n} F(\varepsilon x + (1-\varepsilon^2)^{1/2}y)d\gamma_n(y), \quad \varepsilon \in [-1,1].$$

Stąd $\|Q(\varepsilon) \otimes \mathrm{Id}F\|_{L_2(X)} \leqslant \|F\|_{L_2(X)}$

Dowód. Dla $\varepsilon \in [0,1]$, $Q(\varepsilon) = P_{-\ln(\varepsilon)}$, a dla $\varepsilon \in [-1,1]$, $Q(\varepsilon) = SP_{-\ln(\varepsilon)}$, gdzie S jest izometrią $L_2(X)$ daną wzorem $F(x) \mapsto F(-x)$. Stąd wynika pierwsza tożsamość. Druga wynika z pierwszej i (14). Ostatnie szacowanie wynika z nierówności Jensena oraz tego, że $\varepsilon G_1 + (1-\varepsilon^2)^{1/2}G_2 \sim G$, jeśli G, G_1, G_2 są wektorami w $\mathbb{R}^n$ o rozkładzie $\gamma(n)$ oraz G_1 i G_2 są niezależne. $\square$

Lemat 8.21. *Załóżmy, że $(x_k)_{k\geqslant 0}$ jest ciągiem wektorów z przestrzeni Banacha X takim, że $A := \max_{k\geqslant 0}\|x_k\| < \infty$. Jeśli $\sup_{-1/2\leqslant t\leqslant 1/2}\|\sum_{k=0}^{\infty} t^k x_k\| \leqslant 1$, to*

$$\|x_1\| \leqslant C(1 + \log_+(A)),$$

gdzie C jest stałą uniwersalną.

W dowodzie wykorzystamy słabszą wersję nierówności Bernsteina

Lemat 8.22. *Niech $Q(s) = \sum_{k=-n}^{n} a_k e^{iks}$ będzie wielomianem trygonometrycznym stopnia n o współczynnikach zespolonych. Wówczas $\|Q'\|_\infty \leqslant 2n\|Q\|_\infty$.*

Uwaga 8.23. Prawdziwa jest nieco silniejsza nierówność $\|Q'\|_\infty \leqslant n\|Q\|_\infty$.

Dowód. Niech $F_n(t) = \sum_{|j|\leqslant n}(1 - \frac{|j|}{n+1})e^{ijt}$ będzie jądrem Fejera oraz $\psi_n(t) = 2nF_{n-1}(t)\sin(nt)$, wówczas $\|F_n\|_1 = 1$, zatem $\|\psi_n\|_1 \leqslant 2n$. Łatwo sprawdzić (patrząc na współczynniki Fouriera), że $Q' = -Q*\psi_n$, czyli $\|Q'\|_\infty \leqslant \|\psi_n\|_1\|Q\|_\infty \leqslant 2n\|Q\|_\infty$. □

Dowód Lematu 8.21. Ustalmy funkcjonał $\varphi \in X^*$ taki, że $\|\varphi\| \leqslant 1$ i określmy

$$P_n(t) = \sum_{j=0}^{n} t^j \varphi(x_j) \text{ oraz } Q_n(s) := P_n\Big(\frac{1}{2}\sin s\Big).$$

Wówczas Q_n jest wielomianem trygonometrycznym stopnia n i $Q_n'(0) = \frac{1}{2}P_n'(0) = \frac{1}{2}\varphi(x_1)$. Stąd na mocy Lematu 8.22,

$$\begin{aligned} |\varphi(x_1)| &\leqslant 2\|Q_n'\|_\infty \leqslant 4n\|Q_n\|_\infty \leqslant 4n \sup_{|t|\leqslant 1/2}\Big|\sum_{k=0}^{n} t^k\varphi(x_k)\Big| \\ &\leqslant 4n\Big(\sup_{|t|\leqslant 1/2}\Big|\sum_{k=0}^{\infty} t^k\varphi(x_k)\Big| + \sum_{k=n+1}^{\infty} 2^{-k}|\varphi(x_k)|\Big) \leqslant 4n(1 + A2^{-n}). \end{aligned}$$

Biorąc supremum po wszystkich funkcjonałach otrzymujemy $\|x_1\| \leqslant 4n(1 + A2^{-n})$ i wystarczy przyjąć $n = \lceil \ln_+(A)\rceil$. □

Dowód Twierdzenia 8.18. Teza natychmiast wynika z Faktu 8.20 oraz Lematów 8.16 i 8.21. □

8.4 Liczby aproksymacyjne

Definicja 8.24. Niech $T\colon X \to Y$ będzie operatorem liniowym między przestrzeniami Banacha. Definiujemy *liczby aproksymacyjne* $a_k(T)$, *liczby Gelfanda* $c_k(T)$ oraz *liczby Kołmogorowa* $d_k(T)$ dla $k = 1, 2, \ldots$ wzorami

$$\begin{aligned} a_k(T) &:= \inf\{\|T - S\|\colon\ S\colon X \to Y,\ \mathrm{rk}(S) < k\} \\ c_k(T) &:= \inf\{\|T_{|Z}\|\colon\ Z \subset X,\ \mathrm{codim}(Z) < k\} \\ d_k(T) &:= \inf\{\|\pi_Z T\|\colon\ Z \subset Y,\ \dim(Z) < k\}, \end{aligned}$$

gdzie π_Z oznacza kanoniczny rzut Y na Y/Z.

Oczywiście $a_1(T) = c_1(T) = d_1(T) = \|T\|$ oraz ciągi $(a_k(T))$, $(c_k(T))$ i $(d_k(T))$ są nierosnące.

Fakt 8.25. *Załóżmy, że T jest przekształceniem liniowym między przestrzeniami Banacha X i Y.*
i) Jeśli $X = l_1(\Gamma)/Z$ oraz $q\colon l_1(\Gamma) \to l_1(\Gamma)/Z$ jest kanonicznym rzutem, to $d_k(T) = a_k(Tq)$.
ii) Jeśli $j\colon Y \to l_\infty(\Gamma)$ jest izometrycznym włożeniem, to $c_k(T) = a_k(jT)$.
iii) $c_k(T) = d_k(T^*)$.

Dowód pozostawiamy jako proste ćwiczenie.

Twierdzenie 8.26. *Dla dowolnego $\alpha > 0$ istnieje liczba $C(\alpha)$ zależna tylko od α taka, że dla dowolnego operatora liniowego T między przestrzeniami Banacha X i Y zachodzi*

$$\sup_{k\leqslant n} k^\alpha e_k(T) \leqslant C(\alpha) \sup_{k\leqslant n} k^\alpha s_k(T) \text{ dla wszystkich } n \geqslant 1,$$

gdzie s_k oznacza a_k, c_k lub d_k.

Dowód. Z uwagi na poprzedni fakt oraz to, że każda przestrzeń Banacha jest izometryczna z ilorazem pewnego $l_1(\Gamma)$ i wkłada się izometrycznie w pewne $l_\infty(\Gamma)$, wystarczy udowodnić twierdzenie dla $s_k = a_k$. Wykorzystując jednorodność oraz monotoniczność liczb e_k i a_k wystarczy pokazać, że jeśli $\sup_{k\leqslant n} k^\alpha a_k(T) < 1$ i $n = 2^N$ dla pewnego $N \geqslant 0$, to $n^\alpha e_n(T) \leqslant C'(\alpha)$.

Załóżmy więc, że zachodzi powyżej wzmiankowany warunek. Z definicji liczb a_k znajdziemy operatory $S_m\colon X \to Y$, $m = 0, 1, \ldots, N$ takie, że $\mathrm{rk}(S_m) < 2^m$ oraz $\|T - S_m\| < 2^{-m\alpha}$. Określmy $\Delta_m = S_m - S_{m-1}$ dla $1 \leqslant m \leqslant N$ i $\Delta_0 = S_0$, wówczas $T = \sum_{m=0}^N \Delta_m + T - S_N$, $\|\Delta_m\| \leqslant 2^{\alpha+1}2^{-m\alpha}$ oraz $\mathrm{rk}(\Delta_m) < 2^{m+1}$. Niech $K := T(B_X)$ oraz $K_m := \Delta_m(B_X)$. Zauważmy, że $K_m \subset \|\Delta_m\| B_Y$, zatem

$$N(K_m, \varepsilon\|\Delta_m\|B_Y) \leqslant \Big(1 + \frac{2}{\varepsilon}\Big)^{\mathrm{rk}(\Delta_m)} \leqslant \Big(1 + \frac{2}{\varepsilon}\Big)^{2^{m+1}}.$$

Mamy

$$K \subset \sum_{m=0}^N K_m + (T - S_m)B_X \subset \sum_{m=0}^N K_m + 2^{-N\alpha}B_Y$$

stąd dla $\varepsilon_m > 0$ otrzymujemy

$$N\Big(K, \Big(\sum_{m=0}^N \varepsilon_m\|\Delta_m\| + 2^{-N\alpha}\Big)B_Y\Big) \leqslant \prod_{m=0}^N (K_m, \varepsilon_m\|\Delta_m\|B_Y) \leqslant \prod_{m=0}^N \Big(1 + \frac{2}{\varepsilon_m}\Big)^{2^{m+1}}.$$

Wybierzmy $\varepsilon_m = t2^{2\alpha(m-N)}$, $t > 0$, wówczas

$$\sum_{m=0}^N \varepsilon_m\|\Delta_m\| + 2^{-N\alpha} \leqslant \sum_{m=0}^N t2^{2\alpha(m-N)}2^{\alpha+1}2^{-m\alpha} + 2^{-N\alpha} \leqslant 2^{-N\alpha}(C_1(\alpha)t + 1).$$

Niech $r = \min(1, \frac{1}{4\alpha})$, zauważmy, że $1 + s \leqslant (1 + s^r)^{1/r} \leqslant \exp(s^r/r)$ dla $s > 0$, zatem

$$\prod_{m=0}^N \Big(1 + \frac{2}{\varepsilon_m}\Big)^{2^{m+1}} \leqslant \exp\Big(\sum_{m=0}^N \Big(\frac{2}{\varepsilon_m}\Big)^r \frac{2^{m+1}}{r}\Big) \leqslant \exp(C_2(\alpha)t^{-r}2^N),$$

czyli wykazaliśmy, że dla $t > 0$,

$$N(K, 2^{-N\alpha}(C_1(\alpha)t + 1)) \leqslant \exp(C_2(\alpha)t^{-r}2^N).$$

Zauważmy teraz, że jeśli $t(\alpha)$ jest odpowiednio duże to $2\exp(C_2(\alpha)t(\alpha)^{-r}2^N) \leqslant 2^{2^N}$, czyli

$$e_n(T) \leqslant 2^{-N\alpha}(C_1(\alpha)t(\alpha) + 1) = n^{-\alpha}C'(\alpha),$$

co mieliśmy pokazać. □

Twierdzenie 8.27. *Istnieje stała uniwersalna C taka, że dla dowolnego operatora liniowego T między przestrzenią Banacha X a l_2^n,*

$$c_k(T) \leqslant C_5\sqrt{\frac{n}{k}}e_k(T) \text{ dla } k \leqslant n.$$

Zanim przejdziemy do dowodu powyższego twierdzenia sformułujemy jeden prosty lemat dotyczący macierzy gaussowskich.

Lemat 8.28. *Niech $G_{n,k} = (g_{ij})_{i\leqslant k, j\leqslant n}$ będzie macierzą gaussowską, której współczynniki są niezależnymi zmiennymi $\mathcal{N}(0,1)$. Wówczas*
i) $\mathbb{E}\|G_{n,k}\|_{l_2^n\to l_2^k} \leqslant C_6\max(n,k)$,
ii) jeśli $A \subset l_2^n$ jest zbiorem mocy co najwyżej 2^k, to

$$\mathbb{P}(\forall_{x\in A}\ |G_{n,k}x| \geqslant c_7\sqrt{k}|x|) \geqslant \frac{1}{2}.$$

Dowód. i) Niech $m = \max(\sqrt{n}, \sqrt{k})$ wówczas macierz $G_{n,k}$ możemy traktować jako obcięcie macierzy $G_{m,m}$ i teza wynika z Faktu 5.3.

ii) Zauważmy, że

$$|G_{n,k}x|^2 = \sum_{i=1}^k \Big|\sum_{j=1}^n g_{ij}x_j\Big|^2 \sim |x|^2\sum_{i=1}^k g_i^2.$$

Stąd dla $t > 0$

$$\begin{aligned}\mathbb{P}(|G_{n,k}x| < t\sqrt{k}|x|) = \mathbb{P}\Big(\sum_{i=1}^k g_i^2 < kt^2\Big) &= \gamma_k(t\sqrt{k}B_2^k) \\ &\leqslant (2\pi)^{-k/2}\mathrm{vol}(t\sqrt{k}B_2^k) \leqslant (C_0t)^k\end{aligned}$$

dla pewnej stałej uniwersalnej C_0. Zatem

$$\mathbb{P}(\exists_{x\in A}\ |G_{n,k}x| < t\sqrt{k}|x|) \leqslant |A|(C_0t)^k \leqslant (2C_0t)^k \leqslant \frac{1}{2}.$$

dla $t = 1/(4C_0)$. □

Dowód Twierdzenia 8.27. Ustalmy $1 \leqslant k \leqslant n-1$, niech $\varepsilon = e_{k+1}(T)$ wówczas możemy wybrać $A \subset l_2^n$ będący ε-siecią w odległości euklidesowej dla $T(B_X)$ oraz taki, że $|A| \leqslant 2^k$.

Niech $G_{n,k}$ będzie macierzą gaussowską jak w poprzednim Lemacie, wtedy zbiór

$$\Omega_1 = \{\omega\colon \; \|G_{n,k}(\omega)\|_{l_2^k \to l_2^n} \leqslant 3C_6\sqrt{n}, \; \forall_{y\in A} \; |G_{n,k}(\omega)y| \geqslant c_7\sqrt{k}|y|\}$$

ma dodatnie prawdopodobieństwo, czyli jest niepusty. Niech $G = G_{n,k}(\omega)$ dla pewnego $\omega \in \Omega_1$ oraz $F = \mathrm{Ker}(GT)$. Wówczas $GT\colon X \to l_2^k$, czyli F jest podprzestrzenią X kowymiaru conajwyżej k, stąd $c_{k+1}(T) \leqslant \|GT_{|F}\|$. By oszacować tę ostatnią liczbę ustalmy $x \in F$, $\|x\| \leqslant 1$, z definicji A istnieje $y \in A$ taki, że $|y - Tx| \leqslant \varepsilon$. Wówczas

$$|Gy| = |Gy - GTx| \leqslant \|G\||y - Tx| \leqslant 3C_6\sqrt{n}\varepsilon$$

oraz

$$|y| \leqslant c_7^{-1}k^{-1/2}|Gy| \leqslant 3C_6c_7^{-1}\sqrt{\frac{n}{k}}\varepsilon.$$

Ponieważ $|Ty| \leqslant |y| + \varepsilon$ i y było dowolnym wektorem z kuli jednostkowej w F, to

$$\begin{aligned} c_{k+1}(T) &\leqslant \|GT_{|F}\| \leqslant \Big(3C_6c_7^{-1}\sqrt{\frac{n}{k}} + 1\Big)\varepsilon = \Big(3C_6c_7^{-1}\sqrt{\frac{n}{k}} + 1\Big)e_{k+1}(T) \\ &\leqslant C'\sqrt{\frac{n}{k+1}}e_{k+1}(T). \end{aligned}$$

Mamy też $e_1(T) = c_1(T) = \|T\|$, więc teza twierdzenia została udowodniona. □

Możemy teraz przejść do dowodu kluczowego faktu, który jest wnioskiem z Twierdzenia Kaszina-Szarka oraz wcześniej udowodnionych oszacowań liczb aproksymacyjnych i entropijnych.

Wniosek 8.29. *Istnieje stała uniwersalna C_8 taka, że dla dowolnego przekształcenia T z przestrzeni $X = (\mathbb{R}^n, \|\;\|_K)$ w l_2^n zachodzi*

$$\sqrt{n}\max(e_n(T), e_n(T^*)) \leqslant C_8(1 + \log \mathrm{vr}(K))l^*(T).$$

Dowód. W poniższym dowodzie wielokrotnie wykorzystujemy konwencje, że C oznacza stałą uniwersalną, która może zmieniać wartość przy każdym wystąpieniu.

Dla uproszczenia notacji niech $A := \mathrm{vr}(K) \geqslant 1$. Ustalmy k takie, że $2k \leqslant n$, wówczas z Twierdzenia 7.3 istnieje podprzestrzeń F w $\mathbb{R}^n$ kowymiaru k taka, ze $d_{\mathrm{BM}}(K \cap F, B_2^{n-k}) \leqslant (4\pi A)^{n/k}$. Ponieważ $l^*(T_{|F}) \leqslant l^*(T)$, więc na mocy Wniosku 8.19

$$l((T_{|F})^*) \leqslant C_4(1 + \log d_{\mathrm{BM}}(K \cap F, B_2^{n-k}))l^*(T_{|F}) \leqslant C\frac{n}{k}(1 + \log A)l^*(T).$$

Stąd z Twierdzenia 8.27 i Wniosku 8.12,

$$c_k(T_{|F}) \leqslant C_5\sqrt{\frac{n-k}{k}}e_k(T_{|F}) \leqslant C_3C_5\frac{\sqrt{n-k}}{k}l((T_{|F})^*) \leqslant C\frac{n^{3/2}}{k^2}(1+\log A)l^*(T).$$

Otrzymujemy zatem

$$c_{2k}(T) \leqslant c_k(T_{|F}) \leqslant C\frac{n^{3/2}}{k^2}(1+\log A)l^*(T),$$

zatem z monotoniczności c_k i Twierdzenia 8.26 z $\alpha = 2$

$$n^2 e_n(T) \leqslant C(2)\max_{k\leqslant n} k^2 c_k(T) \leqslant C\max_{2k\leqslant n}(2k)^2 c_{2k}(T) \leqslant Cn^{3/2}(1+\log A)l^*(T).$$

Ponieważ $c_k(T) = d_k(T^*)$, więc analogicznie jak wyżej Twierdzenie 8.26 implikuje, że $n^{1/2}e_n(T^*) \leqslant C(1+\log A)l^*(T)$. □

8.5 Konkluzja dowodu

Wniosek 8.30. *Dla dowolnego symetrycznego ciała wypukłego K istnieje elipsoida $\mathcal{E}$ taka, że*

$$M(K,\mathcal{E}) \leqslant C(1+\log \mathrm{vr}(K))^2.$$

Dowód. Stosując twierdzenie Lewisa znajdujemy $T\colon l_2^n \to X$ takie, że $l(T) = \sqrt{n}$ i $l^*(T^{-1}) \leqslant \sqrt{n}$. Wówczas Wniosek 8.12 implikuje $e_n(T), e_n(T^*) \leqslant C$, zaś z Faktu 8.29 wynika $e_n(T^{-1}), e_n(T{-}1^*) \leqslant C(1+\log \mathrm{vr}(K))$. Przyjmujemy zatem $\mathcal{E} = T(B_2^n)$ i korzystamy z Faktu 8.13. □

Jesteśmy gotowi do dowodu głównego twierdzenia.

Dowód Twierdzenia 8.3. Ustalmy $n \geqslant 1$ i określmy

$$m_n := \sup_K \inf\{M(K,\mathcal{E})\colon\ \mathcal{E} \text{ elipsoida w } \mathbb{R}^n\},$$

gdzie supremum przebiega po wszystkich symetrycznych ciałach wypukłych K. Naszym celem jest pokazanie, że m_n jest ograniczone przez stałą uniwersalną. Zauważmy, że $m_n \leqslant C(1+\log n)^2 < \infty$.

Ustalmy ciało wypukłe K i wybierzmy elipsoidę $\mathcal{E}$ taką, że $M(K,\mathcal{E}) \leqslant 4m_n$.Pokażemy, że wówczas istnieje elipsoida $\mathcal{E}_1$ taka, że

$$M(K,\mathcal{E}_1) \leqslant C_9 m_n^{1/2}(1+\log m_n)^2 \tag{15}$$

dla pewnej stałej uniwersalnej C_9.

Zauważmy, że $M(K,\mathcal{E}) = M_1(K,\mathcal{E})M_2(K,\mathcal{E})$, gdzie

$$M_1(K,\mathcal{E}) := \left(\frac{\mathrm{vol}(K+\mathcal{E})}{\mathrm{vol}(K)}\cdot\frac{\mathrm{vol}(K^0)}{\mathrm{vol}(K^0\cap\mathcal{E}^0)}\right)^{1/n},$$

$$M_2(K,\mathcal{E}) := M_1(K^0,\mathcal{E}^0) = \left(\frac{\mathrm{vol}(K^0+\mathcal{E}^0)}{\mathrm{vol}(K^0)}\cdot\frac{\mathrm{vol}(K)}{\mathrm{vol}(K\cap\mathcal{E})}\right)^{1/n}.$$

Mamy albo $M_1(K,\mathcal{E}) \leqslant 2m_n^{1/2}$ lub $M_1(K^0,\mathcal{E}^0) \leqslant 2m_n^{1/2}$. Ponieważ $M(K,\mathcal{E}_1) = M(K^0,\mathcal{E}_1^0)$, więc (ewentualnie zamieniając rolami K z K^0 i $\mathcal{E}$ z $\mathcal{E}^0$) możemy zakładać, że zachodzi ta pierwsza możliwość.

Połóżmy

$$\alpha := \left(\frac{\mathrm{vol}(K+\mathcal{E})}{\mathrm{vol}(K)}\right)^{1/n}, \quad \beta := \left(\frac{\mathrm{vol}(K^0)}{\mathrm{vol}(K^0\cap\mathcal{E}^0)}\right)^{1/n}$$

wówczas $\alpha\beta = M_1(K,\mathcal{E}) \leqslant 2m_n^{1/2}$. Na mocy Faktu 4.25b) mamy

$$N(K+\mathcal{E},\mathcal{E}) \leqslant (3\alpha)^n, \quad N(K^0,K^0\cap\mathcal{E}^0) \leqslant (3\beta)^n.$$

Mamy też

$$\mathrm{vr}(K+\mathcal{E}) \leqslant \left(\frac{\mathrm{vol}(K+\mathcal{E})}{\mathrm{vol}(\mathcal{E})}\right)^{1/n} \leqslant M(K,\mathcal{E}) \leqslant 4m_n.$$

Fakt 8.29 implikuje zatem istnienie elipsoidy $\mathcal{E}_1$ takiej, że

$$M(K+\mathcal{E},\mathcal{E}_1) \leqslant C_1(1+\log(4m_n))^2.$$

Pokażemy, że elipsoida $\mathcal{E}_1$ spełnia (15).

Mamy

$$N((K+\mathcal{E})\cap\mathcal{E}_1, K\cap 2\mathcal{E}_1) \leqslant N((K+\mathcal{E})\cap\mathcal{E}_1, K) \leqslant N(K+\mathcal{E},K) \leqslant (3\alpha)^n,$$

stąd

$$\mathrm{vol}(((K+\mathcal{E})\cap\mathcal{E}_1) \leqslant (3\alpha)^n\mathrm{vol}(K\cap 2\mathcal{E}_1) \leqslant (6\alpha)^n\mathrm{vol}(K\cap\mathcal{E}_1)$$

i, wobec oczywistej nierówności $\mathrm{vol}(K+\mathcal{E}_1) \leqslant \mathrm{vol}(K+\mathcal{E}+\mathcal{E}_1)$, dostajemy

$$\frac{\mathrm{vol}(K+\mathcal{E}_1)}{\mathrm{vol}(K\cap\mathcal{E}_1)} \leqslant (6\alpha)^n\frac{\mathrm{vol}(K+\mathcal{E}+\mathcal{E}_1)}{\mathrm{vol}((K+\mathcal{E})\cap\mathcal{E}_1)}. \tag{16}$$

Z drugiej strony,

$$\begin{aligned}\mathrm{vol}(K^0+\mathcal{E}_1^0) &\leqslant N(K^0,K^0\cap\mathcal{E}^0)\mathrm{vol}(K^0\cap\mathcal{E}^0+\mathcal{E}_1^0)\\ &\leqslant (3\beta)^n\mathrm{vol}(\mathrm{conv}(K,\mathcal{E})^0+\mathcal{E}_1^0) \leqslant (6\beta)^n\mathrm{vol}((K+\mathcal{E})^0+\mathcal{E}_1^0).\end{aligned}$$

Ponadto, oczywiście $\mathrm{vol}((K+\mathcal{E})^0\cap\mathcal{E}_1^0) \leqslant \mathrm{vol}(K^0\cap\mathcal{E}_1^0)$, zatem

$$\frac{\mathrm{vol}(K^0+\mathcal{E}_1^0)}{\mathrm{vol}(K^0\cap\mathcal{E}_1^0)} \leqslant (6\beta)^n\frac{\mathrm{vol}((K+\mathcal{E})^0+\mathcal{E}_1^0)}{\mathrm{vol}((K+\mathcal{E})^0\cap\mathcal{E}_1^0)}. \tag{17}$$

Szacowania (16) i (17) po wymnożeniu stronami dają

$$M(K,\mathcal{E}_1) \leqslant 36\alpha\beta M(K+\mathcal{E},\mathcal{E}_1) \leqslant 72m_n^{1/2}C_1(1+\log(4m_n))^2,$$

czyli dostaliśmy (15). Stąd (wobec dowolności wyboru ciała K) mamy

$$m_n \leqslant 72C_1m_n^{1/2}(1+\log(4m_n))^2,$$

skąd już łatwo wynika, że $m_n \leqslant C$ dla pewnej stałej uniwersalnej C. $\square$

9 Wnioski z Twierdzenia o Elipsoidzie Milmana

Definicja 9.1. Mówimy, że n-wymiarowe ciało symetryczne K jest *w pozycji regularnej ze stałą* C, jeśli istnieje $a > 0$ takie, że $M(K, aB_2^n) \leqslant C$.

Uwaga 9.2. i) Ponieważ $M(TK, TL) = M(K, L)$, więc, na mocy twierdzenia Milmana, istnieje stała uniwersalna C_0 taka, że dla dowolnego ciała symetrycznego K istnieje takie T, że TK jest w pozycji regularnej ze stałą C_0. Ponadto możemy wybrać T, takie, że $\det(T) = 1$.
ii) Ponieważ $M(K, aB_2^n) = M(K^0, a^{-1}B_2^n)$, więc ciało K jest w pozycji regularnej wtedy i tylko wtedy, gdy ciało K^0 jest w pozycji regularnej.

9.1 Odwrotna nierówność Brunna-Minkowskiego

Przypomnijmy, że nierówność Brunna-Minkowskiego mówi, że dla dowolnych zbiorów zwartych w $\mathbb{R}^n$

$$\mathrm{vol}_n(A + B)^{1/n} \geqslant \mathrm{vol}(A)^{1/n} + \mathrm{vol}(B)^{1/n}.$$

Biorąc na przykład $n = 2$, $K = [-1, 1] \times [-m, m]$, $L = [-m, m] \times [-1, 1]$ widzimy, że niemożliwe jest zachodzenie nierówności odwrotnej $\mathrm{vol}(K+L)^{1/n} \leqslant C(\mathrm{vol}(K)^{1/n} + \mathrm{vol}(L)^{1/n})$ dla dowolnych symetrycznych ciał wypukłych K i L. Okazuje się, że jednak taka nierówność jest prawdziwa, jeśli K i L mają jednokładne elipsoidy Milmana.

Twierdzenie 9.3. *Załóżmy, że $K_1, \ldots, K_m$ są symetrycznymi ciałami wypukłymi w $\mathbb{R}^n$ w pozycji regularnej ze stałą C. Wówczas dla dowolnych $t_1, t_2, \ldots, t_m > 0$ mamy*

$$\mathrm{vol}(t_1K_1 + \ldots + t_mK_m)^{1/n} \leqslant C(3C)^m\Big(t_1\mathrm{vol}(K_1)^{1/n} + \ldots + t_m\mathrm{vol}(K_m)^{1/n}\Big)$$

oraz

$$\mathrm{vol}(t_1K_1^0 + \ldots + t_mK_m^0)^{1/n} \leqslant C(3C)^m\Big(t_1\mathrm{vol}(K_1^0)^{1/n} + \ldots + t_m\mathrm{vol}(K_m^0)^{1/n}\Big).$$

Dowód. W dowodzie wielokrotnie będziemy wykorzystywać Fakt 4.25 pokazujący związek między liczbami pokryciowymi a ilorazami objętości.

Ponieważ ciała t_iK_i oraz $t_iK_i^0$ też są w pozycji regularnej, wystarczy udowodnić pierwsze z oszacowań dla $t_1 = \ldots = t_m = 1$. Załóżmy zatem, że dla $B_i = a_iB_2^n$ mamy $M(K_i, B_i) \leqslant C$. Wówczas

$$N(K_i, B_i) \leqslant N(K_i + B_i, K_i \cap B_i) \leqslant 3^n \frac{\mathrm{vol}(K_i + B_i)}{\mathrm{vol}(K_i \cap B_i)} \leqslant (3C)^n.$$

Stąd dla dowolnego ciała L,

$$\begin{aligned}\mathrm{vol}(K_i + L) &\leqslant N(K_i + L, B_i + L)\mathrm{vol}(B_i + L) \leqslant N(K_i, B_i)\mathrm{vol}(B_i + L)\\ &\leqslant (3C)^n\mathrm{vol}(B_i + L).\end{aligned}$$

W szczególności

$$\mathrm{vol}(K_1 + K_2 + \ldots + K_m)^{1/n} \leqslant 3C\mathrm{vol}(B_1 + K_2 + \ldots + K_m)^{1/n}$$

i powtarzając ten argument dostajemy

$$\mathrm{vol}(K_1 + \ldots + K_m)^{1/n} \leqslant (3C)^m \mathrm{vol}(B_1 + \ldots + B_m)^{1/n}$$

Zauważmy jednak, że kule B_i są jednokładne, zatem

$$\mathrm{vol}(B_1 + \ldots + B_m)^{1/n} = \mathrm{vol}(B_1)^{1/n} + \ldots + \mathrm{vol}(B_m)^{1/n}.$$

Ponadto

$$\frac{\mathrm{vol}(B_i)}{\mathrm{vol}(K_i)} \leqslant \frac{\mathrm{vol}(K_i + B_i)}{\mathrm{vol}(K_i \cap B_i)} \leqslant M(K_i, B_i)^n \leqslant C^n$$

zatem wykazaliśmy żądaną nierówność

$$\mathrm{vol}(K_1 + \ldots + K_m)^{1/n} \leqslant C(3C)^m \Big(\mathrm{vol}(K_1)^{1/n} + \ldots + \mathrm{vol}(K_m)^{1/n}\Big).$$

□

9.2 Twierdzenie o rzucie przekroju

Dla podprzestrzeni $E \subset \mathbb{R}^n$ przez P_E będziemy oznaczać zarówno rzut ortogonalny $\mathbb{R}^n$ na E. Będziemy często wykorzystywać fakt, że $(K \cap E)^0 = P_E(K^0)$.

Twierdzenie 9.4. *Dla $\delta \in (0, 1)$ istnieje stała $C(\delta)$ zależna tylko od δ taka, że dla dowolnego wypukłego ciała symetrycznego K w $\mathbb{R}^n$ istnieją podprzestrzenie $F \subset E \subset \mathbb{R}^n$ takie, że $l = \dim(F) \geqslant \delta n$ oraz*

$$d(P_F(E \cap K), B_2^l) \leqslant C(\delta).$$

Zanim przejdziemy do dowodu udowodnimy następujący fakt.

Fakt 9.5. *Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K i k-wymiarowej podprzestrzeni $E \subset \mathbb{R}^n$ mamy*

$$\frac{1}{\binom{n}{k}} \mathrm{vol}_{n-k}(E^\perp \cap K)\mathrm{vol}_k(P_E K) \leqslant \mathrm{vol}_n(K) \leqslant \mathrm{vol}_{n-k}(E^\perp \cap K)\mathrm{vol}_k(P_E K).$$

Dowód. Na mocy nierówności Brunna-Minkowskiego funkcja

$$x \mapsto \mathrm{vol}_{n-k}(K \cap (E^\perp + x))^{1/(n-k)}$$

jest wklęsła na $P_E K$, ponadto jest symetryczna, zatem osiąga maksimum w zerze. Stąd i z twierdzenia Fubiniego mamy

$$\mathrm{vol}_n(K) = \int_{P_E K} \mathrm{vol}_{n-k}(K \cap (E^\perp + x)) \leqslant \mathrm{vol}_k(P_E K)\mathrm{vol}_{n-k}(E^\perp \cap K).$$

By udowodnić odwrotne szacowanie zauważmy, że jeśli $x \in E$ i $\|x\|_{P_E K} = t$ to $x/t \in P_E K$, czyli zbiór $K \cap (E^\perp + \frac{x}{t})$ jest niepusty. Ponadto, jeśli dodatkowo $t \leqslant 1$, to

$$K \cap (E^\perp + x) \supset (1-t)(K \cap E^\perp) + t\Big(K \cap \Big(E^\perp + \frac{x}{t}\Big)\Big),$$

więc

$$\mathrm{vol}_{n-k}(K \cap (E^\perp + x)) \geqslant (1-t)^{n-k} \mathrm{vol}_{n-k}(K \cap E^\perp).$$

Zatem

$$\mathrm{vol}_n(K) = \int_{P_E K} \mathrm{vol}_{n-k}(K \cap (E^\perp + x)) \geqslant \mathrm{vol}_{n-k}(K \cap E^\perp) \int_{P_E K} (1 - \|x\|_{P_E K})^{n-k}.$$

Mamy jednak dla ciała wypukłego $L := P_E K \subset E$

$$\begin{aligned}\int_L (1 - \|x\|_L)^{n-k} dx &= \int_L \int_{\|x\|_L}^1 (n-k)(1-t)^{n-k-1} dt dx \\ &= (n-k) \int_0^1 (1-t)^{n-k-1} \int_E I_{\{x \in L:\ \|x\|_L \leqslant t\}} dx dt \\ &= (n-k) \int_0^1 (1-t)^{n-k-1} t^k \mathrm{vol}_k(L) dt = \frac{1}{\binom{n}{k}} \mathrm{vol}_k(L).\end{aligned}$$

□

Dowód Twierdzenia 9.4. Możemy bez straty ogólności zakładać, że $n \geqslant \frac{2}{1-\delta}$, w przeciwnym przypadku wystarczy wziąć $F = E = \mathbb{R}^n$ i skorzystać z twierdzenia Johna.

Niech $\mathcal{E}$ będzie elipsoidą Milmana dla K, czyli $M(K, \mathcal{E}) \leqslant C_0$. Wówczas

$$\left(\frac{\mathrm{vol}(K + \mathcal{E})}{\mathrm{vol}(\mathcal{E})}\right)^{1/n} \leqslant M(K, \mathcal{E}) \leqslant C_0.$$

Na mocy Twierdzenia 7.3 dla $1 \leqslant k \leqslant n-1$ istnieje podprzestrzeń E wymiaru k taka, że

$$K \cap E \subset (K + \mathcal{E}) \cap E \subset (4\pi C_0)^{\frac{n}{n-k}} \mathcal{E} \cap E.$$

Dla uproszczenia notacji niech $\lambda := (4\pi C_0)^{n/(n-k)}$, wówczas dualizując powyższą nierówność dostajemy

$$P_E(K^0) \supset \frac{1}{\lambda} P_E(\mathcal{E}^0). \tag{18}$$

Mamy

$$\begin{aligned}\mathrm{vol}_k(P_E(K^0)) &\leqslant 2^n \frac{\mathrm{vol}_n(K^0)}{\mathrm{vol}_{n-k}(E^\perp \cap K^0)} \leqslant (2C_0)^n \frac{\mathrm{vol}_n(K^0 \cap \mathcal{E}^0)}{\mathrm{vol}_{n-k}(E^\perp \cap K^0)} \\ &\leqslant (2C_0)^n \frac{\mathrm{vol}_k(P_E(K^0 \cap \mathcal{E}^0)) \mathrm{vol}_{n-k}(E^\perp \cap (K^0 \cap \mathcal{E}^0))}{\mathrm{vol}_{n-k}(E^\perp \cap K^0)} \\ &\leqslant (2C_0)^n \mathrm{vol}_k(P_E(K^0 \cap \mathcal{E}^0)) \leqslant (2C_0)^n \mathrm{vol}_k(P_E(\mathcal{E}^0)),\end{aligned}$$

gdzie pierwsza i trzecia nierówność wynikają z Faktu 9.5 (i tego, że $\binom{n}{k} \leqslant 2^n$), a druga z oszacowania $M(K, \mathcal{E}) \leqslant C_0$. Zatem dla $k \geqslant n/2$,

$$\left(\frac{\mathrm{vol}_k(P_E(K^0))}{\mathrm{vol}_k(P_E(\mathcal{E}^0))}\right)^{1/k} \leqslant (2C_0)^2$$

co razem z (18) implikuje, że $\mathrm{vr}(P_E(K^0)) \leqslant (2C_0)^2\lambda$. Stąd, ponownie stosując Twierdzenie 7.3, dla $1 \leqslant l < k$ znajdujemy podprzestrzeń $F \subset E$ wymiaru l taką, że

$$\frac{1}{\lambda}P_E(\mathcal{E}^0) \cap F \subset P_E(K^0) \cap F \subset (4\pi(2C_0)^2\lambda)^{\frac{k}{k-l}} \frac{1}{\lambda}P_E(\mathcal{E}^0) \cap F.$$

Dualizując otrzymujemy

$$P_F(\lambda(\mathcal{E} \cap E)) \supset P_F(K \cap E) \supset (4\pi(2C_0)^2\lambda)^{-\frac{k}{k-l}} P_F(\lambda(\mathcal{E} \cap E)),$$

czyli

$$d_{\mathrm{BM}}(P_F(K \cap E), B_2^l) \leqslant (4\pi(2C_0)^2\lambda)^{\frac{k}{k-l}}.$$

Pozostaje teraz wybrać k, l w ten sposób by $n - k = k - l = \lfloor \frac{1-\delta}{2} n \rfloor \geqslant \frac{1-\delta}{4} n$ i wtedy $k \geqslant n/2$, $l \geqslant \delta n$ oraz

$$(4\pi(2C_0)^2\lambda)^{\frac{k}{k-l}} \leqslant C(\delta) = \left(4\pi(2C_0)^2(4\pi C_0)^{4/(1-\delta)}\right)^{4/(1-\delta)}.$$

□

9.3 Odwrotna nierówność Santaló

Definicja 9.6. *Objętością Mahlera* symetrycznego ciała wypukłego K nazywamy liczbę

$$s(K) = \mathrm{vol}(K)^{1/n}\mathrm{vol}(K^0)^{1/n}.$$

Zauważmy, że

$$s(TK) = s(K) \quad \text{dla } T \in \mathrm{GL}(n).$$

W szczególności dla dowolnej elipsoidy $\mathcal{E}$ w $\mathbb{R}^n$

$$s(\mathcal{E}) = s(B_2^n) = \mathrm{vol}(B_2^n)^{2/n} = \frac{\pi}{\Gamma(n/2+1)^{2/n}}.$$

Nierówność Santaló, którą udowodnimy nieco później mówi, że $s(K) \leqslant s(B_2^n)$ dla dowolnego symetrycznego ciała wypukłego K. Twierdzenie Milmana pozwala w jednej linijce wykazać jej odwrotną wersję – ten trudny i nieoczywisty wynik pierwotnie udowodnili Bourgain i Milmana.

Twierdzenie 9.7 (Odwrotna nierówność Santaló)**.** *Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K*

$$\mathrm{vol}(K)^{1/n}\mathrm{vol}(K^0)^{1/n} \geqslant \frac{1}{C_0}\mathrm{vol}(B_2^n)^{2/n},$$

gdzie C_0 jest stałą z twierdzenia o elipsoidzie Milmana.

Dowód. Wystarczy zauważyć, że dla elipsoidy Milmana $\mathcal{E}$

$$C_0 \geqslant M(K,\mathcal{E}) \geqslant \frac{s(\mathcal{E})}{s(K)}.$$

□

Uwaga 9.8. Ponieważ $M(K,\mathcal{E}) \geqslant s(K)/s(\mathcal{E})$, więc taki sam dowód jak wyżej pokazuje też, że $s(K) \leqslant C_0 s(\mathcal{E})$. Stąd dla dowolnych ciał wypukłych symetrycznych K i L, $s(K) \leqslant C_0^2 s(L)$.

9.4 Dualność liczb entropijnych modulo czynnik wykładniczy

Twierdzenie 9.9 (König-Milman)**.** *Istnieje stała uniwersalna C_1 taka, że dla dowolnych symetrycznych ciał wypukłych K i L w $\mathbb{R}^n$,*

$$\frac{1}{C_1^n} N(L^0,K^0) \leqslant N(K,L) \leqslant C_1^n N(L^0,K^0).$$

Lemat 9.10. *Dla dowolnych n-wymiarowych symetrycznych ciał wypukłych K i L,*

$$\mathrm{vol}(K) \leqslant N(K,L)\mathrm{vol}(K\cap L)$$

Dowód. Niech $N = N(K,L)$ wówczas istnieją wektory $x_1,\dots,x_N$ takie, że

$$K \subset \bigcup_{i=1}^N (x_i+L), \text{ czyli } K = \bigcup_{i=1}^N (x_i+L)\cap K.$$

Zauważmy, że funkcja $x \mapsto \mathrm{vol}_n((x+L)\cap K))^{1/n}$ jest symetryczna i wklęsła na $K+L$, stąd przyjmuje maksimum w zerze i

$$\mathrm{vol}(K) \leqslant \sum_{i=1}^N \mathrm{vol}((x_i+L)\cap K) \leqslant N\mathrm{vol}(K\cap L).$$

□

Dowód Twierdzenia 9.9. Lemat 9.10 oraz odwrotna nierówność Santaló wraz z uwagą po niej implikują

$$N(K,L) \geqslant \frac{\mathrm{vol}(K)}{\mathrm{vol}(K\cap L)} \geqslant \frac{1}{C_0^{2n}} \frac{\mathrm{vol}((K\cap L)^0)}{\mathrm{vol}(K^0)}.$$

Mamy jednak

$$(K\cap L)^0 = \mathrm{conv}(K^0,L^0) \supset \frac{K^0+L^0}{2},$$

stąd

$$N(K,L) \geqslant \frac{1}{(4C_0^2)^n}\frac{\mathrm{vol}(2K^0+2L^0)}{\mathrm{vol}(K^0)} \geqslant \frac{1}{(4C_0^2)^n}\frac{\mathrm{vol}(K^0+2L^0)}{\mathrm{vol}(K^0)} \geqslant \frac{1}{(4C_0^2)^n} N(L^0,K^0),$$

gdzie ostatnia nierówność wynika z Faktu 4.25ii). Wykazaliśmy, że $N(K,L) \geqslant C_1^{-n} N(L^0, K^0)$ zamieniając K z L^0 i L z K^0 dostajemy drugą nierówność z tezy. $\square$

10 Nierówność Santaló. Hipoteza Mahlera

Zacznijmy od sformułowania nierówności Santaló.

Twierdzenie 10.1. *Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K,*

$$\mathrm{vol}(K)\mathrm{vol}(K^0) \leqslant \mathrm{vol}(B_2^n)^2.$$

Dowód. Dowód przeprowadzimy używając symetryzacji Steinera. Ustalmy ciało wypukłe K oraz $n-1$ wymiarową podprzestrzeń E, niech $a \in S^{n-1} \cap E^{\perp}$, czyli $E = a^{\perp}$. Zdefiniujemy zbiór $S_E(K)$ w następujący sposób. Dla $x \in P_E(K)$ kładziemy

$$S_E(K) \cap (x + E^{\perp}) = x + \Big[-\frac{1}{2}\mathrm{vol}_1(K \cap (x + \mathrm{Lin}(a))), \frac{1}{2}\mathrm{vol}_1(K \cap (x + \mathrm{Lin}(a)))\Big]a,$$

a dla $x \notin P_E(K)$ definiujemy $s_E(K) \cap (x + E^{\perp}) = \emptyset$. Innymi słowy, zbiór $S_E(K)$ to zbiór którego cięcia prostopadłe do E są odcinkami symetrycznymi o tej samej długości co odpowiednie cięcia K. Z jednowymiarowej nierówności Brunna-Minkowskiego łatwo wynika, że zbiór $S_E(K)$ jest wypukły, nietrudno też wykazać, że jest on zwarty, zatem jest ciałem wypukłym. Z twierdzenia Fubiniego natychmiast dostajemy, że $\mathrm{vol}_n(S_E K) = \mathrm{vol}_n(K)$. Wykażemy, że

$$\mathrm{vol}(S_E K)\mathrm{vol}((S_E K)^0) \geqslant \mathrm{vol}(K)\mathrm{vol}(K^0). \tag{19}$$

W tym celu wystarczy udowodnić, że $\mathrm{vol}_n((S_E K)^0) \geqslant \mathrm{vol}_n(K^0)$. Dla $s \in \mathbb{R}$ połóżmy

$$K^0(s) := \{y \in E \colon y + sa \in K^0\}, \quad (S_E K)^0(s) := \{y \in E \colon y + sa \in (S_E K)^0\}.$$

Z twierdzenia Fubiniego,

$$\mathrm{vol}_n(K^0) = \int \mathrm{vol}_{n-1}(K^0(s))ds, \quad \mathrm{vol}_n((S_E K)^0) = \int \mathrm{vol}_{n-1}((S_E K)^0(s))ds,$$

zatem by udowodnić (19) wystarczy pokazać, że

$$\mathrm{vol}_{n-1}((S_E K)^0(s)) \geqslant \mathrm{vol}_{n-1}(K^0(s)) \quad \text{dla } s \in \mathbb{R}. \tag{20}$$

Zauważmy, że

$$K^0(s) = \big\{y \in E \colon\ \forall_{x \in E} \forall_{t \in \mathbb{R}}\ \langle x, y\rangle + ts \leqslant 1,\ \text{jeśli } x + ta \in K\big\},$$

$$S_E K = \Big\{x + \frac{t_1 - t_2}{2}a \colon\ x \in E,\ x + t_1 a, x + t_2 a \in K\Big\},$$

$$(S_E K)^0(s) = \Big\{ y \in E\colon\ \forall_{x\in E}\ \langle x, y\rangle + \frac{t_1 - t_2}{2} s \leqslant 1, \text{ jeśli } x + t_i a \in K, i = 1, 2 \Big\}.$$

Jeśli $y_1 \in K^0(s)$, $y_2 \in K^0(-s)$ oraz $x + t_1 a \in K$, $x + t_2 a \in K$, to

$$\langle x, y_1\rangle + t_1 s \leqslant 1, \quad \langle x, y_2\rangle - t_2 s \leqslant 1,$$

czyli

$$\langle x, \frac{y_1 + y_2}{2}\rangle + \frac{t_1 - t_2}{2} s \leqslant 1,$$

co pokazuje, że $\frac{1}{2}(y_1 + y_2) \in (S_E K)^0(s)$. Zatem

$$\frac{1}{2}(K^0(s) + K^0(-s)) \subset (S_E K)^0(s)$$

i z nierówności Brunna Minkowskiego,

$$\mathrm{vol}_{n-1}((S_E K)^0(s)) \geqslant \big(\mathrm{vol}_{n-1}(K^0(s))\mathrm{vol}_{n-1}(K^0(-s))\big)^{1/2}$$

i (20) wynika stąd, że $K^0(-s) = -K^0(s)$ na mocy symetrii K^0.

Udowodniliśmy zatem (19) dla dowolnej przestrzeni E wymiaru $n - 1$. Dla dowolnego symetrycznego ciała wypukłego K o tej samej objętości co cB_2^n istnieje ciąg symetryzacji Steinera taki, że $K_i = S_{E_i} \cdots S_{E_2} S_{E_1} K$ zbiega do cB_2^n w odległości Hausdorffa, a zatem również geometrycznej. Stąd K_i^0 zbiega do $c^{-1}B_2^n$ w odległości geometrycznej. Stąd z (19)

$$\begin{aligned} \mathrm{vol}_n(K)\mathrm{vol}_n(K^0) &\leqslant \mathrm{vol}_n(K_i)\mathrm{vol}_n(K_i^0) \to \mathrm{vol}_n(cB_2^n)\mathrm{vol}_n(c^{-1}B_2^n) \\ &= \mathrm{vol}_n(B_2^n)\mathrm{vol}_n(B_2^n). \end{aligned}$$

□

Hipoteza 10.2 (Mahler)**.** *Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K,*

$$s(K)^n = \mathrm{vol}(K)\mathrm{vol}(K^0) \geqslant s(B_1^n)^n = \mathrm{vol}(B_1^n)\mathrm{vol}(B_\infty^n) = \frac{4^n}{n!}.$$

Uwaga 10.3. Zauważmy, że $s(B_2^n) = \pi\Gamma(\frac{n}{2} + 1)^{-2/n} = \frac{1}{n}(2\pi e + o(1))$ oraz $s(B_1^n) = 4(n!)^{-1/n} = \frac{1}{n}(4e + o(1))$. Twierdzenie 9.7 pokazuje zatem, że $s(K) \geqslant cs(B_1^n)$ dla pewnej stałej uniwersalnej $c > 0$. Greg Kuperberg wykazał, że $s(K) \geqslant \frac{\pi}{4} s(B_1^n)$, co jest najlepszym znanym do tej pory wynikiem.

Uwaga 10.4. Można wykazać, że hipoteza Mahlera jest spełniona w $\mathbb{R}^2$. Przypadek trzy i więcej wymiarowy jest otwarty.

Uwaga 10.5. Nietrudno wykazać, że $K \mapsto s(K)$ jest ciągłe względem odległości Banacha-Mazura, w szczególności przyjmuje minimum dla jakiegoś ciała wypukłego symetrycznego K. Nazarov, Ryabogin i Zvavitch wykazali, że B_1^n jest lokalnym minimum dla objętości Mahlera na kompakcie Banacha-Mazura. Wiadomo też, że każde lokalne minimum nie może być ciałem gładkim.

Definicja 10.6. Ciało wypukłe w $\mathbb{R}^n$ nazywamy *bezwarunkowym*, jeśli dla dowolnego $x = (x_1, \ldots, x_n) \in K$ oraz dowolnych znaków $\eta_1, \ldots, \eta_n \in \{\pm 1\}$, $(\eta_1 x_1, \ldots, \eta_n x_n) \in K$.

Twierdzenie 10.7 (Saint-Raymond)**.** *Jeśli K jest n-wymiarowym bezwarunkowym ciałem wypukłym, to*

$$s(K)^n = \mathrm{vol}(K)\mathrm{vol}(K^0) \geqslant s(B_1^n)^n = \mathrm{vol}(B_1^n)\mathrm{vol}(B_\infty^n) = \frac{4^n}{n!}.$$

Dowód. Udowodnimy twierdzenie przez indukcję po n. Dla $n = 1$ teza jest oczywista, bo $s([-a, a]) = 4$ dla dowolnego $a > 0$. Załóżmy zatem, że $n > 1$ i twierdzenie zachodzi dla $n-1$-wymiarowych bezwarunkowych ciał wypukłych. Niech K będzie bezwarunkowym ciałem wypukłym w $\mathbb{R}^n$. Określmy

$$K^+ := K \cap \mathbb{R}^n_+ = \{x \in K\colon\ x_1, \ldots, x_n \geqslant 0\}.$$

Dla $1 \leqslant i \leqslant n$ i $x \in K^+$ zdefiniujmy ostrosłup o wierzchołku w x i podstawie $K^+ \cap e_i^\perp$,

$$K_i^+(x) := \mathrm{conv}\{x, K^+ \cap e_i^\perp\} \subset K^+.$$

Zauważmy, że przecięcie $K_i^+(x)$ i $K_j^+(x)$ dla $i \neq j$ ma wymiar co najwyżej $n-1$, czyli zerową objetość. Zatem

$$\begin{aligned}\mathrm{vol}(K) = 2^n \mathrm{vol}(K^+) &\geqslant 2^n \sum_{i=1}^n \mathrm{vol}(K_i^+(x)) = 2^n \sum_{i=1}^n \frac{1}{n} x_i \mathrm{vol}_{n-1}(K^+ \cap e_i^\perp) \\ &= \sum_{i=1}^n x_i \frac{2}{n} \mathrm{vol}_{n-1}(K \cap e_i^\perp).\end{aligned}$$

Stąd otrzymujemy, że

$$\sum_{i=1}^n x_i \frac{2\mathrm{vol}_{n-1}(K \cap e_i^\perp)}{n\mathrm{vol}_n(K)} \leqslant \sum_{i=1}^n |x_i| \frac{2\mathrm{vol}_{n-1}(K \cap e_i^\perp)}{n\mathrm{vol}_n(K)} \leqslant 1 \text{ dla } x \in K,$$

czyli

$$\left(\frac{2\mathrm{vol}_{n-1}(K \cap e_i^\perp)}{n\mathrm{vol}_n(K)} \right)_{i=1}^n \in K^0.$$

Zamieniając rolami K z K^0 (które też oczywiście jest ciałem bezwarunkowym) dostajemy

$$\left(\frac{2\mathrm{vol}_{n-1}(K^0 \cap e_i^\perp)}{n\mathrm{vol}_n(K^0)} \right)_{i=1}^n \in K.$$

Ponieważ $\langle x, y \rangle \leqslant 1$ dla $x \in K$, $y \in K^0$, więc

$$\sum_{i=1}^n \frac{2\mathrm{vol}_{n-1}(K \cap e_i^\perp)}{n\mathrm{vol}_n(K)} \cdot \frac{2\mathrm{vol}_{n-1}(K^0 \cap e_i^\perp)}{n\mathrm{vol}_n(K^0)} \leqslant 1. \tag{21}$$

Zauważmy, że $(K \cap e_i^\perp)^0 = P_{e_i^\perp}(K^0) = K^0 \cap e_i^\perp$, więc z założenia indukcyjnego i (21) dostajemy

$$\mathrm{vol}_n(K)\mathrm{vol}_n(K^0) \geqslant \frac{4}{n^2}\sum_{i=1}^{n} \mathrm{vol}_{n-1}(K\cap e_i^\perp)\mathrm{vol}_{n-1}(K^0\cap e_i^\perp) \geqslant \frac{4}{n}\cdot\frac{4^{n-1}}{(n-1)!} = \frac{4^n}{n!}.$$

□

11 Hyperplane Conjecture

Zacznijmy od sformułowania ważnego otwartego problemu, któremu poświęcimy tę część wykładu.

Problem 11.1 (Hyperplane Conjecture, Slicing Problem)**.** *Czy istnieje stała uniwersalna $c > 0$ taka, że dla dowolnego symetrycznego n–wymiarowego ciała wypukłego o objętości jeden istnieje podprzestrzeń H wymiaru $n-1$ dla której $\mathrm{vol}_{n-1}(K \cap H) \geqslant c$?*

Uwaga 11.2. Można rozpatrywać niesymetryczne ciała wypukłe, ale wtedy za H trzeba brać $n-1$–wymiarowe podprzestrzenie afiniczne.

Definicja 11.3. Mówimy, że ciało wypukłe K znajduje się w *pozycji izotropowej*, jeśli
i) $\mathrm{vol}_n(K) = 1$;
ii) $\int_K x_i dx = 0$ dla $1 \leqslant i \leqslant n$;
iii) $\int_K x_i x_j dx = \alpha\delta_{ij}$ dla $1 \leqslant i, j \leqslant n$ i pewnego $\alpha > 0$.

Uwaga 11.4. Warunek ii) oznacza, że zero jest środkiem ciężkości K, oczywiście jest on spełniony dla ciał symetrycznych. Warunek iii) mówi, że $\mathrm{Cov}(X_K) = \alpha\mathrm{Id}$, gdzie X_K oznacza wektor losowy rozłożony jednostajnie na K.

Fakt 11.5. *i) Dla każdego ciała wypukłego istnieje takie przekształcenie afiniczne $Tx = Ax + b$, $\det(A) \neq 0$, że $TK = AK + b$ jest w pozycji izotropowej.*
ii) Jeśli $K_1 = T_1K$ i $K_2 = T_2K$ dwa przekształcenia afiniczne n-wymiarowego ciała K będące w pozycji izotropowej, to $K_2 = AK_1$ dla pewnego $A \in O(n)$.

Dowód. i) Najpierw przesuwamy ciało K tak, by środek ciężkości stał się zerem, potem je przeskalowujemy, by objętość była równa 1, a na końcu przekształcamy liniowo z zachowaniem objętości tak, by macierz kowariancji była wielokrotnością identyczności (korzystamy tu z prostej obserwacji, że $\mathrm{Cov}(X_{AK}) = A\mathrm{Cov}(X_K)A^T$).

ii) Załóżmy, że $K_2 = AK_1 + b$, ponieważ K_1 i K_2 mają równą objetość, więc $\det(A) = 1$, zero jest środkiem ciężkości K_1 i K_2, zatem $b = 0$. Wreszcie

$$\alpha_2\mathrm{Id} = \mathrm{Cov}(X_{K_2}) = \mathrm{Cov}(X_{AK_1}) = A\mathrm{Cov}(X_{K_1})A^T = \alpha_1 AA^T.$$

Z porównania wyznaczników $\alpha_1 = \alpha_2$, stąd $AA^T = \mathrm{Id}$, czyli $A \in O(n)$. □

Definicja 11.6. *Stałą izotropową n-wymiarowego ciała wypukłego K nazywamy liczbę* $L_K = \mathrm{vol}_n(K)^{-1/n}(\det \mathrm{Cov}(X_K))^{1/2n}$.

Uwaga 11.7. Dla dowolnego nieosobliwego przekształcenia afinicznego T, $L_{TK} = L_K$. Jeśli ciało K jest w pozycji izotropowej, to $\mathrm{Cov}(X_K) = L_K^2 \mathrm{Id}$.

Problem 11.8. *Czy istnieje taka stała C, że dla dowolnego symetrycznego ciała wypukłego K, $L_K \leqslant C$?*

By powiązać Problemy 11.1 i 11.8 potrzebujemy kilku prostych lematów.

Lemat 11.9. *Niech g będzie gęstością miary probabilistycznej na $\mathbb{R}^n$, wówczas*

$$\|g\|_\infty^{2/n} \int_{\mathbb{R}^n} |x|^2 g(x)dx \geqslant \int_{c_n B_2^n} |x|^2 dx, \tag{22}$$

gdzie $c_n > 0$ takie, że $\mathrm{vol}_n(c_n B_2^n) = 1$.

Dowód. Niech μ oznacza miarę z gęstością g. Zauważmy, że jeśli przeskalujemy liniowo miarę μ, tzn. g zmienimy na $g_a(x) = a^n g(ax)$ dla $a > 0$, to lewa strona (22) się nie zmieni. Możemy zatem zakładać, że $\|g\|_\infty = 1$. Mamy wówczas dla $t > 0$

$$\mu\{x\colon |x| \geqslant t\} = 1 - \int_{tB_2^n} g(x)dx \geqslant 1 - \mathrm{vol}_n(tB_2^n).$$

Całkując przez części otrzymujemy:

$$\begin{aligned}\int_{\mathbb{R}^n} |x|^2 g(x)dx &= 2\int_0^\infty t\mu\{x\colon |x| \geqslant t\}dt \geqslant 2\int_0^{c_n} t(1 - \mathrm{vol}_n(tB_2^n))dt \\ &= 2\int_0^{c_n} t\mathrm{vol}_n\{x \in B_2^n\colon |x| \geqslant t\} = \int_{c_n B_2^n} |x|^2 dx.\end{aligned}$$

□

Wniosek 11.10. *Dla dowolnego ciała wypukłego K w $\mathbb{R}^n$, $L_K \geqslant L_{B_2^n} \geqslant \frac{1}{\sqrt{2\pi e}}$.*

Dowód. Pierwsza nierówność wynika natychmiast z Lematu 11.9 zastosowanego do gęstości rozkładu jednostajnego na K. By obliczyć $L_{B_2^n}$ zauważmy, że

$$\frac{1}{\mathrm{vol}_n(B_2^n)} \int_{B_2^n} |x|^2 dx = 2\int_0^1 t\frac{\mathrm{vol}_n\{x \in B_2^n\colon |x| \geqslant t\}}{\mathrm{vol}_n(B_2^n)}dt = 2\int_0^1 t(1-t^n) = \frac{n}{n+2},$$

czyli $\mathrm{Cov}(X_{B_2^n}) = \frac{1}{n+2}\mathrm{Id}$. Zatem

$$L_{B_2^n}^{2n} = \frac{1}{\mathrm{vol}_n^2(B_2^n)(n+2)^n} = \frac{\Gamma^2(n/2+1)}{\pi^n(n+2)^n} \geqslant \frac{1}{(2\pi e)^n},$$

gdzie ostatnią nierówność dla $n = 1, 2$ sprawdzamy na palcach, a dla $n \geqslant 3$ korzystamy z wzoru Stirlinga i oszacowania dla $x \geqslant 3/2$

$$\Gamma(x+1) \geqslant \sqrt{2\pi x}x^x e^{-x} \geqslant \sqrt{3\pi}(x+1)^x e^{-x-1} \geqslant (x+1)^x e^{-x}.$$

□

Lemat 11.11. *Niech g będzie funkcją logarytmicznie wklęsłą na $\mathbb{R}$ (tzn. $g = e^{-f}$, $f\colon \mathbb{R} \to [0,\infty]$ wypukła), parzystą oraz $\int_{\mathbb{R}} g dx = 1$. Wówczas*

$$\frac{1}{12} \leqslant g(0)^2 \int_{\mathbb{R}} x^2 g(x) dx \leqslant \frac{1}{2}.$$

Dowód. Niech μ będzie miarą z gęstością g, wówczas na mocy nierówności Prékopy-Leindlera, μ jest logarytmicznie wklęsła, tzn. $\mu(tA + (1-t)B) \geqslant \mu(A)^t \mu(B)^{1-t}$. Zauważmy, że $g(0) = \|g\|_\infty$ (bo parzysta funkcja wypukła ma minimum w 0), więc dolne oszacowanie wynika z (22). By udowodnić górne oszacowanie, możemy przeskalowując liniowo μ założyć, że $g(0) = 1$. Zdefiniujmy $h\colon [0,\infty) \to [0,\infty]$ wzorem

$$h(t) = -\ln(2\mu([t,\infty))).$$

Z logarytmicznej wklęsłości μ wynika wypukłość h, ponadto $h(0) = 0$, $h'(0) = 2g(0) = 2$, stąd $h(t) \geqslant 2t$, czyli

$$g(0) \int_{\mathbb{R}} x^2 g(x) dx = 2 \int_0^\infty t e^{-h(t)} dt \leqslant 2 \int_0^\infty t e^{-2t} dt = \frac{1}{2}.$$

□

Twierdzenie 11.12. *i) Dla dowolnego n-wymiarowego symetrycznego ciała wypukłego K o objetości jeden istnieje hiperpłaszczyzna H taka, że*

$$\mathrm{vol}_{n-1}(K \cap H) \geqslant \frac{1}{2\sqrt{3} L_K}$$

ii) Załóżmy, że ciało K jest w pozycji izotropowej. Wówczas dla dowolnej hiperpłaszczyzny H,

$$\frac{1}{2\sqrt{3} L_K} \leqslant \mathrm{vol}_{n-1}(K \cap H) \leqslant \frac{1}{\sqrt{2} L_K}.$$

Dowód. i) Z definicji L_K wynika, że $\det \mathrm{Cov}(X_K) = L_K^{2n}$. Macierz $\mathrm{Cov}(X_K)$ jest symetryczna, nieujemna, więc jej najmniejsza wartość własna $\lambda_1 \leqslant L_K^2$. Niech $u \in S^{n-1}$ takie, że $\mathrm{Cov}(X_K)u = \lambda_1 u$ oraz

$$g_u(t) = \mathrm{vol}_{n-1}(K \cap (tu + u^\perp)), \quad t \geqslant 0.$$

Funkcja g_u jest symetryczna, logarytmicznie wklęsła, czyli z Lematu 11.11

$$\begin{aligned} \frac{1}{12} &\leqslant g_u^2(0) \int_{\mathbb{R}} t^2 g_u(t) dt = g_u^2(0) \int_K \langle x, u\rangle dx = g_u(0)^2 \langle \mathrm{Cov}(X_K)u, u\rangle \\ &= \mathrm{vol}_{n-1}^2(K \cap u^\perp)\lambda_1 \leqslant \mathrm{vol}_{n-1}^2(K \cap u^\perp) L_K^2. \end{aligned}$$

Zatem $\mathrm{vol}_{n-1}(K \cap u^\perp) \geqslant \frac{1}{2\sqrt{3}L_K}$.

ii) Oszacowanie dolne już wykazaliśmy w i), by wykazać oszacowanie górne załóżmy, że $H = u^\perp$ i postępujemy jak wcześniej wykorzystując jednak tym razem górne oszacowanie z Lematu 11.11. □

Wniosek 11.13. *Problemy 11.1 i 11.8 są równoważne. Dokładniej, jeśli*

$$C_n = \sup\left\{L_k\colon\ K \textit{sym. ciało wypukłe w } \mathbb{R}^n\right\},$$

$$c_n := \inf\left\{\sup_{H\in G_{n-1,n}} \mathrm{vol}_{n-1}(K\cap H)\colon K \textit{sym. ciało wypukłe w } \mathbb{R}^n, \mathrm{vol}_n(K)=1\right\},$$

to $\frac{1}{2\sqrt{3}} \leqslant c_nC_n \leqslant \frac{1}{\sqrt{2}}$.

Lemat 11.14. *Załóżmy, że rzeczywista zmienna losowa X ma symetryczny rozkład z logarytmicznie wklęsłą gęstością, wówczas*

$$(\mathbb{E}|X|^p)^{1/p} \leqslant \frac{\Gamma(p+1)^{1/p}}{\Gamma(q+1)^{1/q}}(\mathbb{E}|X|^q)^{1/q} \textit{ dla } p \geqslant q > 0.$$

Równość zachodzi dla zmiennych o symetrycznym rozkładzie wykładniczym.

Dowód. Odpowiednio przeskalowując X możemy założyć, że $\mathbb{E}|X|^q = \mathbb{E}|Y|^q$, gdzie Y symetryczny rozkład wykładniczy z parametrem 1, tzn. rozkład z gęstością $\frac{1}{2}e^{-x}$. Musimy pokazać, że $\mathbb{E}|Y|^p \geqslant \mathbb{E}|X|^p$. Zauważmy, że $\mathbb{P}(Y \geqslant t) = \frac{1}{2}e^{-t}$. Zdefiniujmy

$$h(t) = -\ln \mathbb{P}(X \geqslant t) \text{ dla } t \geqslant 0,$$

wówczas h jest funkcją wypukłą, $h(0) = \ln 2 = -\ln \mathbb{P}(Y \geqslant 0)$. Ponieważ $\mathbb{E}|X|^q = \mathbb{E}|Y|^q$, więc istnieje $t_0 > 0$ takie, że $h(t_0) = -\ln \mathbb{P}(Y \geqslant t_0) = \ln 2 + t_0$. Z wypukłości h wnioskujemy, że $h(t) \leqslant \ln 2 + t$ dla $t \in [0, t_0]$ oraz $h(t) \geqslant \ln 2 + t$ dla $t \geqslant t_0$. Zatem

$$\left(\left(\frac{t}{t_0}\right)^{p-1} - \left(\frac{t}{t_0}\right)^{q-1}\right)\big(\mathbb{P}(|Y| \geqslant t) - \mathbb{P}(|X| \geqslant t)\big) \geqslant 0$$

dla wszystkich $t > 0$. Stąd

$$\begin{aligned}
\mathbb{E}|Y|^p - \mathbb{E}|X|^p &= pt_0^{p-1}\int\left(\frac{t}{t_0}\right)^{p-1}(\mathbb{P}(|Y| \geqslant t) - \mathbb{P}(|X| \geqslant t))dt \\
&\geqslant pt_0^{p-1}\int\left(\frac{t}{t_0}\right)^{q-1}(\mathbb{P}(|Y| \geqslant t) - \mathbb{P}(|X| \geqslant t))dt \\
&= \frac{p}{q}t_0^{p-q}(\mathbb{E}|Y|^q - \mathbb{E}|X|^q) = 0.
\end{aligned}$$

□

Twierdzenie 11.15. *Dla dowolnego bezwarunkowego ciała wypukłego K, $L_K \leqslant \frac{1}{\sqrt{2}}$.*

Dowód. Zamieniając ciało K na DK gdzie D jest dodatnią macierzą diagonalną, możemy zakładać, że K jest w pozycji izotropowej. Określmy $K_+ := 2K \cap \mathbb{R}^n_+$, wówczas $\mathrm{vol}_n(K_+) = 1$. Zauważmy, że jeśli $x \in K_+$, to $[0, x_1] \times [0, x_2] \times \cdots \times [0, x_n] \subset K_+$, czyli $x_1x_2\cdots x_n \leqslant \mathrm{vol}_n(K_+) = 1$. Niech

$$V := \{x \in \mathbb{R}^n_+\colon x_1x_2\cdots x_n > 1\},$$

wówczas V jest zbiorem wypukłym, rozłącznym ze zbiorem K_+. Z twierdzenia o oddzielaniu wynika, że istnieją liczby $\lambda_1, \ldots, \lambda_n$ takie, że

$$V \subset \{x\colon \sum_{i=1}^n \lambda_i x_i > n\} \text{ oraz } K_+ \subset \{x\colon \sum_{i=1}^n \lambda_i x_i \leqslant n\}. \tag{23}$$

Ponieważ $(a^n + 1, 1/a, 1/a, \ldots, 1/a) \in V$ dla dowolnego $a > 0$, to biorąc $a \to \infty$ widzimy, że $\lambda_1 > 0$, analogicznie $\lambda_i > 0$ dla $i = 2, \ldots, n$. Gdyby $\prod_{i=1}^n \lambda_i < 1$, to $(1/\lambda_1, \ldots, 1/\lambda_n) \in V$, co przeczyłoby (23). Zatem $\prod_{i=1}^n \lambda_i \geqslant 1$ i z nierówności między średnią arytmetyczną i geometryczną wynika, że

$$\begin{aligned} 1 = \mathrm{vol}(K_+) &\geqslant \int_{K_+} \frac{1}{n}\sum_{i=1}^n \lambda_i x_i dx = \frac{1}{n}\sum_{i=1}^n \lambda_i \int_{K_+} x_i dx \\ &\geqslant \Big(\prod_{i=1}^n \lambda_i \prod_{i=1}^n \int_{K_+} x_i dx\Big)^{1/n} \geqslant \Big(\prod_{i=1}^n \int_{K_+} x_i dx\Big)^{1/n}. \end{aligned}$$

Ale wykorzystując nierówność z Lematu 11.14,

$$\int_{K_+} x_i dx = 2\int_K |x_i| dx \geqslant \sqrt{2}\Big(\int_K x_i^2 dx\Big)^{1/2} = \sqrt{2} L_K$$

i otrzymujemy tezę. □

L_K można łatwo oszacować przez odległość od kuli euklidesowej

Fakt 11.16. *Dla dowolnego wypukłego symetrycznego ciała K w $\mathbb{R}^n$, $L_K \leqslant d_{\mathrm{BM}}(K, B_2^n)$.*

Dowód. Ponieważ $L_K = L_{TK}$ możemy założyć, że $\mathrm{vol}(K) = 1$ oraz $aB_2^n \subset K \subset adB_2^n$, gdzie $d = d_{\mathrm{BM}}(K, B_2^n)$. Z porównania objętości wynika, że $a \leqslant \sqrt{n}$, zatem $K \subset d\sqrt{n}B_2^n$ i oczywiście

$$\mathrm{tr}(\mathrm{Cov}(X_K)) = \int_K |x|^2 dx \leqslant \sup_{x\in K} |x|^2 \leqslant d^2 n.$$

Zatem z nierówności między średnią arytmetyczną i geometryczną,

$$L_K^2 = (\det \mathrm{Cov}(X_K))^{1/n} \leqslant \frac{1}{n}\mathrm{tr}(\mathrm{Cov}(X_K)) \leqslant d^2.$$

□

Uwaga 11.17. Poprzedni Fakt i twierdzenie Johna implikują, że $L_K \leqslant \sqrt{n}$. Najlepsze znane oszacowanie stałej L_K pochodzi od Klartaga i mówi, że $L_K \leqslant Cn^{1/4}$.

12 Rozkłady logarytmicznie wklęsłe

Definicja 12.1. Miarę μ na $\mathbb{R}^n$ nazywamy *logarytmicznie wklęsłą* (lub krócej *log-wklęsłą*), jeśli dla dowolnych niepustych zbiorów zwartych A i B zachodzi

$$\mu(\lambda A + (1-\lambda)B) \geqslant \mu(A)^\lambda \mu(B)^{1-\lambda} \text{ dla } 0 < \lambda < 1.$$

Z nierówności Brunna-Minkowskiego łatwo wynika, że miara Lebesgue'a oraz rozkład jednostajny na ciele wypukłym są logarytmicznie wklęsłe. By wprowadzić nieco szerszą klasę miar będziemy potrzebowali definicji funkcji logarytmicznie wklęsłej.

Definicja 12.2. Funkcja nieujemna g na $\mathbb{R}^n$ (lub ogólniej na wypukłym podzbiorze $\mathbb{R}^n$) jest *logarytmicznie wklęsła*, jeśli $g = e^{-h}$, gdzie h jest funkcją wypukłą o wartościach w $(-\infty, \infty]$ (równoważnie $f(\lambda x + (1-\lambda)y) \geqslant f(x)^\lambda f(y)^{1-\lambda}$).

Fakt 12.3. *Załóżmy, że g jest funkcją logarytmicznie wklęsłą na k-wymiarowej afinicznej podprzestrzeni $E \subset \mathbb{R}^n$. Wówczas miara μ na $\mathbb{R}^n$ skoncentrowana na E i mająca gęstość g, tzn. miara zadana wzorem*

$$\mu(A) = \int_{E\cap A} g(x)dx$$

jest logarytmicznie wklęsła.

Dowód. Łatwo widać, że wystarczy rozpatrzeć przypadek $E = \mathbb{R}^n$. Ustalmy zbiory zwarte A, B w $\mathbb{R}^n$ i $\lambda \in (0,1)$, wówczas funkcje $F := gI_A$, $G := gI_B$ i $H = gI_{\lambda A + (1-\lambda)B}$ spełniają założenia nierówności Prékopy-Leindlera (Twierdzenie 2.6). Zatem

$$\mu(\lambda A + (1-\lambda)B) = \int H dx \geqslant \Big(\int F dx\Big)^\lambda \Big(\int G dx\Big)^{1-\lambda} = \mu(A)^\lambda \mu(B)^{1-\lambda}.$$

□

Okazuje się, że nie ma innych miar logarytmicznie wklęsłych.

Twierdzenie 12.4 (Borel)**.** *Miara μ na $\mathbb{R}^n$, która jest skończona na zbiorach zwartych, jest logarytmicznie wklęsła wtedy i tylko wtedy, gdy istnieje podprzestrzeń afiniczna E wymiaru $0 \leqslant k \leqslant n$ w której leży nośnik μ oraz logarytmicznie wklęsła funkcja nieujemna g na E będąca gęstością μ.*

Wniosek 12.5. *Jeśli nośnik miary σ-skończonej μ na $\mathbb{R}^n$ jest pełnowymiarowy (tzn. μ nie jest skupiona na afinicznej podprzestrzeni mniejszego wymiaru), to miara μ jest logarytmicznie wklęsła wtedy i tylko wtedy, gdy ma logarytmicznie wklęsłą gęstość.*

W czasie tego wykładu będziemy rozważali miary (rozkłady) probabilistyczne.

Przykłady rozkładów logarytmicznie wklęsłych:
1) Rozkład jednostajny na ciele wypukłym K. Wówczas $g = I_K$;
2) Rozkład gaussowski. Wówczas $g = \exp(\langle Ax, x\rangle + \langle b, x\rangle)$ dla $x \in E$;
3) Produktowy symetryczny rozkład wykładniczy, czyli rozkład z gęstością

$$g_\lambda(x) = (2\lambda)^{-n} \exp(-\lambda \|x\|_1) \quad \lambda > 0;$$

4) Rozkład z gęstością $c_{p,K} \exp(-\|x\|_K^p)$, $p \geqslant 1$, K ciało wypukłe w $\mathbb{R}^n$.

Następny fakt pokazuje podstawowe, użyteczne własności rozkładów log-wklęsłych.

Fakt 12.6. *Następujące rozkłady są logarytmicznie wklęsłe:*
i) afiniczny obraz rozkładu logarytmicznie wklęsłego (tzn. rozkład postaci $\mu \circ T^{-1}$, gdzie μ jest logarytmicznie wklęsły na $\mathbb{R}^n$, a T jest afinicznym przekształceniem $\mathbb{R}^n$ w $\mathbb{R}^m$);
ii) produkt rozkładów logarytmicznie wklęsłych;
iii) splot rozkładów logarytmicznie wklęsłych;
iv) słaba granica rozkładów logarytmicznie wklęsłych.

Dowód. i) jest oczywista z definicji, ii) wynika z charakteryzacji Borela, a iii) wynika z ii) i i), bo splot to obraz produktu przy przekształceniu $(x, y) \mapsto x + y$.

By udowodnić iv) załóżmy, że rozkłady logwklęsłe μ_n zbiegają słabo do rozkładu μ. Ustalmy $\lambda \in (0, 1)$ i zbiory zwarte A i B, niech $C = \lambda A + (1 - \lambda)B$. Oczywiście $\mu(\partial A)$ nie musi być zerem, więc nie wiemy, że $\mu_n(A) \to \mu(A)$. Ale da się znaleźć ciąg t_k zbiezny do zera taki, że $\mu(\partial A_{t_k}) = \mu(\partial B_{t_k}) = \mu(\partial C_{t_k}) = 0$ i wtedy jak łatwo sprawdzić $C_{t_k} \supset \lambda A_{t_k} + (1 - \lambda) B_{t_k}$, zatem

$$\begin{aligned} \mu(C_{t_k}) &= \lim_{n\to\infty} \mu_n(C_{t_k}) \geqslant \lim_{n\to\infty} \mu_n(A_{t_k})^\lambda \mu_n(B_{t_k})^{1-\lambda} \\ &= \mu(A_{t_k})^\lambda \mu(B_{t_k})^{1-\lambda} \geqslant \mu(A)^\lambda \mu(B)^{1-\lambda}. \end{aligned}$$

Ze zwartości C wynika, że $\bigcap_k C_{t_k} = C$, czyli $\mu(C_{t_k}) \to \mu(C)$ przy $k \to \infty$. □

Lemat 12.7. *Załóżmy, że miara probabilistyczna μ ma gęstość g taką, że funkcja g^s jest wklęsła na zbiorze zwartym wypukłym K, który zawiera nośnik g. Niech $\frac{1}{m} \leqslant s$ i*

$$K_m = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \colon x \in K,\ |y| \leqslant (g(x)/c_m)^{1/m}\},$$

gdzie $c_m = \mathrm{vol}_m(B_2^m)$. Wówczas K_m jest ciałem wypukłym o objętości 1, ponadto jeśli P_m jest kanonicznym rzutem $\mathbb{R}^n \times \mathbb{R}^m$ na $\mathbb{R}^n$, to $P_m X_{K_m}$ ma rozkład μ.

Dowód. Z twierdzenia Fubiniego $\mathrm{vol}_{n+m}(K_m) = \int g = 1$, to, że wektor $P_m X_{K_m}$ ma gęstość g jest oczywiste. Funkcja $g^{1/m}$ jest wklęsła na K jako złożenie funkcji wklęsłej g^s i funkcji wklęsłej niemalejącej $t \mapsto t^{1/(ms)}$ na $[0, \infty)$, a stąd łatwo wynika, że K_m jest ciałem wypukłym. □

Wniosek 12.8. *Załóżmy, że g jest logarytmicznie wklęsłą gęstością na $\mathbb{R}^n$, wówczas istnieją ciała wypukłe K_m w $\mathbb{R}^n \times \mathbb{R}^m$ o objętości 1 takie, że jeśli P_m jest kanonicznym rzutem $\mathbb{R}^n \times \mathbb{R}^m$ na $\mathbb{R}^n$, to gęstości zmiennych $P_m X_{K_m}$ zbiegają punktowo do g.*

Dowód. Dla $s > 0$ określmy

$$\tilde{g}_s(x) = \Big(1 + s \ln g(x)\Big)_+^{1/s}$$

wówczas oczywiście $\tilde{g}_s^s$ jest funkcją wklęsłą na swoim nośniku. Ponadto $0 \leqslant g_s \leqslant g$, czyli na mocy twierdzenia Lebesgue'a o zbieżności zmajoryzowanej,

$$c_s := \int \tilde{g}_s(x)dx \to \int g dx = 1.$$

Określmy (dla na tyle małych s by $c_s > 0$) $g_s := c_s^{-1}\tilde{g}_s$. Wtedy funkcja $g_{1/m}$ spełnia założenia Lematu 12.7 i możemy skonstruować ciało K_m by $g_{1/m}$ była gęstością $P_m X_{K_m}$. □

Z charakteryzacji Borela i poprzedniego wniosku (wykorzystując też fakt, że zbieżność gęstości implikuje słabą zbieżność rozkładów) natychmiast otrzymujemy:

Wniosek 12.9. *Każdy rozkład log-wklęsły jest słabą granicą rzutów rozkładów jednostajnych na ciałach wypukłych.*

Mamy zatem następujący wniosek pokazujący dlaczego klasa rozkładów logwklęsłych jest naturalnym obiektem.

Wniosek 12.10. *Klasa rozkładów logarytmicznie wklęsłych to najmniejsza klasa rozkładów probabilistycznych zawierających rozkłady jednostajne, która jest zamknięta ze względu na przekształcenia afiniczne (niekoniecznie zachowujące wymiar) i słabe granice.*

Lemat 12.11 (Borel)**.** *Załóżmy, że μ jest rozkładem logarytmicznie wklęsłym na $\mathbb{R}^n$, a K wypukłym ciałem symetrycznym w $\mathbb{R}^n$ oraz $\mu(K) = \theta > 0$. Wówczas*

$$1 - \mu(tK) \leqslant \theta\Big(\frac{1-\theta}{\theta}\Big)^{\frac{1+t}{2}} \qquad \text{dla } t \geqslant 1.$$

Dowód. Zauważmy, że jeśli $\|x\|_K > t$ oraz $\|y\|_K \leqslant 1$, to

$$\Big\|\frac{2}{t+1}x + \frac{t-1}{t+1}y\Big\|_K \geqslant \frac{2}{t+1}\|x\|_K - \frac{t-1}{t+1}\|y\|_K > 1,$$

co pokazuje, że

$$\frac{2}{t+1}\big(\mathbb{R}^n \setminus tK\big) + \frac{t-1}{t+1}K \subset \mathbb{R}^n \setminus K,$$

czyli

$$1 - \theta = \mu\big(\mathbb{R}^n \setminus K\big) \geqslant \mu\big(\mathbb{R}^n \setminus tK\big)^{\frac{2}{t+1}}\mu(K)^{\frac{t-1}{t+1}} = (1-\mu(tK))^{\frac{2}{t+1}}\theta^{\frac{t-1}{t+1}}$$

i po prostych przekształceniach dostajemy dowodzone oszacowanie. □

Wniosek 12.12. *Istnieje stała uniwersalna C taka, że dla dowolnej normy $\| \ \|$ na $\mathbb{R}^n$ i dowolnej miary logarytmicznie wklęsłej na $\mathbb{R}^n$, jeśli $\mu\{\|x\| \leqslant m\} \geqslant 2/3$, to*

$$\Big(\int \|x\|^p d\mu(x)\Big)^{1/p} \leqslant Cpm \ \textit{dla}\ p \geqslant 1.$$

Ponadto

$$\Big(\int \|x\|^p d\mu(x)\Big)^{1/p} \leqslant C\frac{p}{q}\Big(\int \|x\|^q d\mu(x)\Big)^{1/q} \ \textit{dla}\ p \geqslant q \geqslant 1.$$

Dowód. Niech X ma rozkład μ, wówczas na mocy Lematu 12.11 z $\theta = 2/3$ otrzymujemy

$$\mathbb{P}(\|X\| \geqslant tm) \leqslant \Big(\frac{1}{2}\Big)^{\frac{1+t}{2}} \qquad \text{dla } t \geqslant 1$$

Stąd całkując przez części

$$\begin{aligned}\mathbb{E}\|X\|^p &= p\int_0^\infty t^{p-1}\mathbb{P}(\|X\| \geqslant t)dt \leqslant m^p\Big(1 + p\int_1^\infty t^{p-1}\mathbb{P}(\|X\| \geqslant tm)dt\Big)\\ &\leqslant m^p\Big(1 + p\int_1^\infty t^{p-1}2^{-\frac{1+t}{2}}dt\Big) \leqslant (Cm)^p.\end{aligned}$$

Oznaczmy $\|X\|_q := (\mathbb{E}\|X\|^q)^{1/q}$, wówczas na mocy nierówności Czebyszewa,

$$\mathbb{P}(\|X\| \geqslant 2e^2\|X\|_q) \leqslant (2e^2)^{-q} \leqslant \frac{1}{2}e^{-2q},$$

skąd z Lematu 12.11,

$$\mathbb{P}(\|X\| \geqslant 2e^2 t\|X\|_q) \leqslant e^{-q(t+1)} \qquad \text{dla } t \geqslant 1$$

i całkując przez części jak wcześniej dostajemy

$$\mathbb{E}\|X\|^p \leqslant (2e^2\|X\|_q)^p\Big(1 + p\int_1^\infty t^{p-1}e^{-(t+1)q}dt\Big) \leqslant \Big(C\frac{p}{q}\|X\|_q\Big)^p.$$

□

Uwaga 12.13. Oszacowania z Wniosku 12.12 są optymalnego rzędu, bo dla zmiennej X o rozkładzie symetrycznym wykładniczym na prostej $\|X\|_p = \Gamma(p+1)^{1/p} \sim p$ dla $p \geqslant 1$.

Uwaga 12.14. Liczbę $\frac{2}{3}$ w definicji m można zastąpić przez dowolną liczbę p z przedziału $(0,1)$, gdyż można pokazać, że jeśli $\mu\{x\colon \|x\| \leqslant m_p\} \leqslant 1 - p$ i $p > 0$, to istnieje stała C_p zależna tylko od p taka, że $\mu\{x\colon \|x\| \leqslant tm_p\} \leqslant C_p t$ dla $t \in (0,1)$.

13 Oszacowania stałej izotropowej

Twierdzenie 13.1 (Bourgain)**.** *Dla dowolnego wypukłego symetrycznego ciała K, $L_K \leqslant Cn^{1/4}\log n$.*

Definicja 13.2. Dla wypukłego ciała symetrycznego K określamy *średnią szerokość* K jako

$$w(K) := \int_{S^{n-1}} \sup_{x\in K} |\langle x,\theta\rangle| d\sigma_{n-1}(\theta) = \int_{S^{n-1}} \sup_{x\in K} \|\theta\|_{K^0} d\sigma_{n-1}(\theta).$$

Uwaga 13.3. Alternatywnie się często określna

$$M(K) = \Big(\int_{S^{n-1}} \|\theta\|_K^2 d\sigma_{n-1}(\theta)\Big)^{1/2}.$$

Wówczas $\frac{1}{C}M(K^0) \leqslant w(K) \leqslant M(K^0)$. Ponadto, jeśli G ma rozkład γ_n, to

$$w(K) = \frac{1}{\mathbb{E}|G|}\mathbb{E}\|G\|_{K^0} \leqslant \frac{C}{\sqrt{n}}\mathbb{E}\|G\|_{K^0}.$$

Wniosek 8.19 razem z Twierdzeniem Lewisa 8.14 implikują, że istnieje takie $T \in \mathrm{GL}(n)$, że

$$w(TK)w((TK)^0) \leqslant M(TK)M((TK)^0) \leqslant C(1+\log d_{\mathrm{BM}}(K, B_2^n)) \leqslant C\log(n+1).$$

Ponieważ $w(cK) = cw(K)$ możemy oczywiście zakładać, że $T \in \mathrm{SL}(n)$. Oszacowanie powyższe się nazywa MM^*-estimate.

Wniosek 13.4. *Dla dowolnego wypukłego ciała symetrycznegi K istnieje nieujemny symetryczny operator $T \in \mathrm{SL}(n)$ taki, że $w(TK) \leqslant C\sqrt{n}\log(n+1)\mathrm{vol}(K)^{1/n}$.*

Dowód. Niech $T \in \mathrm{SL}(n)$ spełnia $w(TK)w((TK)^0) \leqslant C\log(n+1)$. Z Faktu 5.4,

$$\begin{aligned}\mathrm{vol}_n(K)^{1/n} = \mathrm{vol}_n(TK)^{1/n} &= \Big(\mathrm{vol}_n(B_2^n)\int_{S^{n-1}} \|\theta\|_{TK}^{-n} d\sigma_{n-1}(\theta)\Big)^{1/n} \\ &\geqslant \frac{1}{C\sqrt{n}}\Big(\int_{S^{n-1}} \|\theta\|_{TK}^{-n} d\sigma_{n-1}(\theta)\Big)^{1/n}.\end{aligned}$$

Zauważmy, że

$$\Big(\int_{S^{n-1}} \|\theta\|_{TK}^{-n} d\sigma_{n-1}(\theta)\Big)^{-1/n} \leqslant \int_{S^{n-1}} \|\theta\|_{TK} d\sigma_{n-1}(\theta) = w((TK)^0),$$

skąd wynika, że $w(TK) \leqslant C\sqrt{n}\log(n+1)\mathrm{vol}(K)^{1/n}$. By zakończyć dowód zauważmy, że $w(TK) = w(UTK)$ dla dowolnego $U \in O(n)$, więc dobierając U tak, by $UT = |T|$ dostajemy symetrię i nieujemność. □

Lemat 13.5. *Załóżmy, że X jest symetrycznym wektorem logwklęsłym w $\mathbb{R}^n$ takim, że $\mathrm{Cov}(X) = a^2\mathrm{Id}$ oraz S jest skończonym podzbiorem $\mathbb{R}^n$. Wtedy*

$$\mathbb{E}\max_{s\in S}|\langle s, X\rangle| \leqslant Ca\log(\#S+1)\sup_{t\in S}|t|.$$

Dowód. Mamy $\|\langle s, X\rangle| = a|s|$, zatem dla $u \geqslant 0$,

$$\mathbb{P}(|\langle s, X\rangle| \geqslant ua|s|) \leqslant 2e^{-u/C}$$

i

$$\mathbb{P}(\max_{s\in S} |\langle s, X\rangle| \geqslant ua \max_{s\in S} |s|) \leqslant 2\#S e^{-u/C}.$$

Teza wynika przez całkowanie przez części. □

Dowód Twierdzenia 13.1. Bez straty ogólności możemy założyć, że ciało K jest w pozycji izotropowej czyli $\mathrm{vol}(K) = 1$ oraz $\mathrm{Cov}(X_K) = L_K^2 \mathrm{Id}$. Korzystając z Wniosku 13.4 znajdujemy takie $T \in \mathrm{SL(n)}$ nieujemne i symetryczne, że $w(TK) \leqslant C\sqrt{n}\log(n+1)$. Ponadto $\mathrm{tr}(T)/n \geqslant \det(T)^{1/n} = 1$, więc

$$nL_K^2 =\leqslant \frac{1}{n}\mathrm{tr}(T)L_K^2 = \int_K \langle Tx, x\rangle dx \leqslant \int_K \sup_{y\in TK} |\langle y, x\rangle| dx.$$

Niech $R > 0$ będzie najmniejszą liczbą taką, że $TK \subset RB_2^n$. Stąd

$$nL_K^2 \leqslant \int_K \sup_{y\in TK} |\langle y, x\rangle| dx \leqslant R\int_K |x|dx \leqslant R\sqrt{n}L_K,$$

więc możemy zakładać, że $R \geqslant n^{3/4}\log(n+1)$.

Zauważmy, że

$$\mathbb{E}\|G\|_{(TK)^0} = \mathbb{E}|G|w(TK) \leqslant Cn\log(n+1),$$

więc korzystając z minoryzacji Sudakowa dla $j = 1, 2, \ldots$ (Twierdzenie 8.4),

$$\log N(TK, R2^{-j}B_2^n) \leqslant C\frac{2^{2j}}{R^2}(\mathbb{E}\|G\|_{(TK)^0})^2 \leqslant C\frac{2^{2j}}{R^2}n^2\log^2(n+1).$$

Możemy znaleźć zatem zbiory $A_j \subset TK$ oraz przekształcenia $\pi_j\colon TK \to A_j$ takie, że $\log\#A_j \leqslant C2^{2j}R^{-2}n^2\log^2(n+1)$ i $|y - \pi_j(y)| \leqslant R2^{-j}$ dla $y \in TK$. Połóżmy dodatkowo $A_0 = \{0\}$ i $\pi_0(y) = 0$. Wtedy dla liczby całkowitej m, którą dobierzemy później i $y \in TK$,

$$y = y - 0 = y - \pi_m(y) + \sum_{i=i}^{m}(\pi_i(y) - \pi_{i-1}(y)),$$

stąd dla $x \in \mathbb{R}^n$

$$\begin{aligned}\sup_{y\in TK} |\langle y, x\rangle| &\leqslant \sup_{y\in TK} |\langle y - \pi_m(y), x\rangle| + \sup_{y\in TK}\sum_{i=i}^{m} |\langle \pi_i(y) - \pi_{i-1}(y), x\rangle| \\ &\leqslant R2^{-m}|x| + \sum_{i=i}^{m}\sup_{y\in TK} |\langle \pi_i(y) - \pi_{i-1}(y), x\rangle|.\end{aligned}$$

Mamy $\pi_i(y) - \pi_{i-1}(y) \in A_i - A_{i-1}$,

$$\log \#(A_i - A_{i-1}) \leqslant \log \#A_i + \log \#A_{i-1} \leqslant C2^{2i}R^{-2}n^2\log^2(n+1) \quad i = 1, 2, \ldots$$

oraz $|\pi_i(y) - \pi_{i-1}(y)| \leqslant |\pi_i(y) - y| + |y - \pi_{i-1}(y)| \leqslant 3R2^{-i}$. Zatem z Lematu 13.5 dla $X = X_K$,

$$\int_K \sup_{y \in TK} |\langle \pi_i(y) - \pi_{i-1}(y), x\rangle| \leqslant C_1 L_K \frac{2^i}{R} n^2 \log^2(n+1).$$

Ponieważ $\int_K |x| dx \leqslant \sqrt{n} L_K$, więc

$$\begin{aligned} nL_K^2 &\leqslant \int_K \sup_{y \in TK} |\langle y, x\rangle| dx \leqslant R2^{-m}\sqrt{n}L_K + C_1 \sum_{i=1}^m \frac{2^i}{R} n^2 \log^2(n+1) L_K \\ &\leqslant R2^{-m}\sqrt{n}L_K + C_1 \frac{2^{m+1}}{R} n^2 \log^2(n+1) L_K. \end{aligned}$$

Wybierzmy największe $m \geqslant 0$ takie, że $R2^{-m}\sqrt{n}L_K \geqslant 2^m R^{-1} n^2 \log^2(n+1) L_K$ dostajemy

$$nL_K^2 \leqslant C\Big(R\sqrt{n}L_K R^{-1} n^2 \log^2(n+1) L_K\Big)^{1/2} = Cn^{5/4}\log(n+1)L_K.$$

□

Uwaga 13.6. Klartag poprawił nieco oszacowanie Bourgaina i najlepsze obecnie znane oszacowanie stałej izotropowej to $L_K \leqslant Cn^{1/4}$.

Uwaga 13.7. Mówimy, że wypukłe ciało symetryczne K jest ciałem Ψ_α ze stałą A jeśli dla $u > 0$ i $t \in \mathbb{R}^n$,

$$\mathbb{P}(|\langle X, t\rangle| > Au\|\langle X, t\rangle\|_2) \leqslant 2e^{-t^\alpha}.$$

Każde ciało wypukłe jest Ψ_1 ze stałą uniwersalną. Podobny argument jak powyżej pokazuje, że dla ciała Ψ_α, $1 \leqslant \alpha < 2$ w pozycji izotropowej,

$$\mathbb{E}\max_{t \in S} |\langle t, X\rangle| \leqslant CA\log(\#S + 1)^{1/\alpha} \sup_{t \in S} |t|$$

oraz

$$L_K \leqslant C_\alpha A^{\alpha/2} n^{1/2 - \alpha/4} \log(n+1).$$

14 Otwarte pytania

Najważniejszym otwartym problemem w przedstawianej dziedzinie wydaje się pytanie o ograniczoność (niezależną od wymiaru) stałej izotropowej dla ciał wypukłych. W ostatnim rozdziale tego skryptu przedstawiamy szereg innych, ważnych problemów, często ściśle związanych z „hyperplane conjecture".

14.1 Hipoteza KLS

Definicja 14.1. Dla miary probabilistycznej μ na $\mathbb{R}^n$ określamy

$$\mu_+(A) := \liminf_{t\to 0+} \frac{\mu(A_t)-\mu(A)}{t}.$$

Zauważmy, że gdy ν oznacza symetryczny rozkład wykładniczy na prostej z gęstością $\frac{1}{2}e^{-|x|}$ oraz $\nu(-\infty,x]=p$, to $\nu_+(-\infty,x]=\frac{1}{2}e^{-|x|}=\min\{p,1-p\}$

Definicja 14.2. Powiemy, że miara probabilistyczna μ na $\mathbb{R}^n$ spełnia *nierówność Cheegera* ze stałą D, jeśli

$$\mu_+(A) \geqslant \frac{1}{D}\min\{\mu(A),1-\mu(A)\} \quad \text{dla } A\in\mathcal{B}(\mathbb{R}^n). \tag{24}$$

Najmniejszą stałą D dla której zachodzi powyższa nierówność będziemy oznaczać $D_{\mathrm{Che}}(\mu)$.

Uwaga 14.3. Nietrudno wykazać, że nierówność (24) jest równoważna

$$\forall_{A\in\mathcal{B}(\mathbb{R}^n)}\ \mu(A)=\nu(-\infty,x] \ \Rightarrow\ \mu(A_{Dt})=\nu(-\infty,x+t] \text{ dla } t\geqslant 0.$$

Nietrywialny wynik Bobkowa i Houdré mówi, że jeśli $\mu=\otimes_{k\leqslant m}\mu_k$, to $D_{\mathrm{Che}}(\mu)\leqslant\sqrt{6}\max_{k\leqslant m}D_{\mathrm{Che}}(\mu_k)$.

Hipoteza 14.4 (Kannan-Lóvasz-Simonovits)**.** *Istnieje stała uniwersalna C taka, że $D_{\mathrm{Che}}(\mu)\leqslant C$ dla dowolnej izotropowej miary logarytmicznie wklęsłej μ.*

W dalszej części wskażemy wielkości równoważne $D_{\mathrm{Che}}(\mu)$ dla miary log-wklęsłej μ.

Definicja 14.5. Powiemy, że miara probabilistyczna μ na $\mathbb{R}^n$ spełnia *nierówność Poincaré* ze stałą D, jeśli dla dowolnej funkcji gładkiej ograniczonej f

$$\mathrm{Var}_\mu(f)=\int\Big(f-\int f d\mu\Big)^2 d\mu \leqslant D^2\int|\nabla f|^2 d\mu$$

Najmniejszą stałą $D\geqslant 0$ dla której zachodzi powyższa nierówność będziemy oznaczać $D_{\mathrm{Poin}}(\mu)$.

Zaletą nierówności Poincaré jest jej tensoryzowalność – jeśli $\mu=\otimes_{k\leqslant m}\mu_k$ oraz μ_k spełniają nierówność Poincaré, to μ też spełnia nierówność Poincaré oraz $D_{\mathrm{Poin}}(\mu)\leqslant\max_{k\leqslant m}D_{\mathrm{Poin}}(\mu_k)$.

Mazya i Cheeger (przy pomocy „co-area formula”) pokazali, że nierówność Cheegera implikuje nierówność Poincaré i stała $D_{\mathrm{Poin}}(\mu)\leqslant 2D_{\mathrm{Che}}(\mu)$. Odwrotna implikacja nie jest prawdziwa – miara μ_α na $[-1,1]$ z gęstością $\frac{1+\alpha}{2}|x|^\alpha$ dla $\alpha\in(0,1)$ spełnia nierówność Poincaré, a nie spełnia nierówności Cheegera, bo $\mu([-1,0])=\frac{1}{2}$ i $\mu_+([-1,0])=0$. Buser i Ledoux wykazali, że dla miar logarytmicznie wklęsłych μ nierówność Poincaré implikuje nierówność Cheegera i $D_{\mathrm{Che}}(\mu)\leqslant CD_{\mathrm{Poin}}(\mu)$.

Definicja 14.6. Powiemy, że miara probabilistyczna μ na $\mathbb{R}^n$ spełnia *wykładniczą koncentrację* ze stałą D, jeśli dla dowolnego zbioru borelowskiego A takiego, że $\mu(A) \geqslant 1/2$,

$$\mu(A_t) = \mu(A + tB_2^n) \geqslant 1 - 2e^{-t/D} \quad \text{dla } t \geqslant 0.$$

Najmniejszą stałą $D \geqslant 0$ dla której zachodzi powyższa nierówność będziemy oznaczać $D_{\exp}(\mu)$.

Nietrudno zauważyć, że nierówność Cheegera implikuje wykładniczą koncentrację oraz $D_{\exp}(\mu) \leqslant D_{\mathrm{Che}}(\mu)$. Gromov i Milman jako pierwsi zauważyli, że koncentracja wykładnicza wynika również z nierówności Poincaré. Stosując np. tzw. argument Herbsta, spopularyzowany przez Ledoux, można pokazać, że $D_{\exp}(\mu) \leqslant 2D_{\mathrm{Poin}}(\mu)$. Biorąc np. miarę z gęstością $\frac{1}{2}(I_{[0,1]} + I_{[2,3]})$ widzimy, że odwrotne implikacje są fałszywe.

Można też rozważyć słabszą formę nierówności Poincaré mianowicie, że dla gładkich funkcji ograniczonych

$$\int \Big| f - \int f d\mu \Big| d\mu \leqslant D \sup_x |\nabla f(x)|. \tag{25}$$

Zauważmy, że $\sup_x |\nabla f(x)|$ to stała Lipschitza funkcji f. Najmniejszą stałą D w powyższej nierówności będziemy oznaczać przez $D_{\mathrm{Lip}}(\mu)$.

Twierdzenie 14.7 (E.Milman)**.** *Dla miary logarytmicznie wklęsłej μ na $\mathbb{R}^n$ nierówności Cheegera, Poincaré, słabsza nierówność Poincaré 25 oraz wykładnicza koncentracja są równoważne. Co więcej stałe $D_{\mathrm{Che}}(\mu)$, $D_{\mathrm{Poin}}(\mu)$, $D_{\mathrm{Lip}}(\mu)$ i $D_{\exp}(\mu)$ porównują się z dokładnością do stałych uniwersalnych.*

Najlepsze znane obecnie oszacowania, to $D_{\mathrm{Che}}(\mu) \leqslant Cn^{1/3}\log^{1/2} n$ dla miar izotropowych log-wklęsłych (wynik Eldana w oparciu o szacowania z pracy Guédona i Milmana) i $D_{\mathrm{Che}}(\mu) \leqslant C\log n$ dla bezwarunkowych log-wklęsłych miar izotropowych (Klartag przy użyciu wyników Milmana).

14.2 Szacowania „thin-shell”

Kluczowym elementem dowodu Klartaga centralnego twierdzenia granicznego dla wektorów logwklęsłych było pokazanie, że jeśli X jest izotropowy log-wklęsły, to $|X|/\sqrt{n}$ ma małą wariancję, czyli się silnie koncentruje wokół 1.

Hipoteza 14.8. *Dla dowolnego izotropowego wektora log-wklęsłego X, $\mathbb{E}(|X| - \sqrt{n})^2 \leqslant C$.*

Zauważmy, że funkcja $|x|$ jest 1-Lipschitzowska, więc jeśli X ma rozkład μ, to $\mathrm{Var}(|X|)^{1/2} \leqslant D_{\mathrm{Poin}}(\mu)$. Zauważmy też, że $\mathbb{E}|X|^2 = n$ oraz $|(\mathbb{E}|X|^2)^{1/2} - \mathbb{E}|X|| \leqslant (\mathbb{E}||X| - \mathbb{E}|X||^2)^{1/2}$, stąd

$$((\mathbb{E}(|X| - \sqrt{n})^2)^{1/2} \leqslant \mathrm{Var}(|X|) + |\sqrt{n} - \mathbb{E}|X|| \leqslant 2\mathrm{Var}(|X|) \leqslant 2D_{\mathrm{Poin}}(\mu),$$

czyli hipoteza KLS implikuje Hipotezę 14.8. Ciekawe jest, że (z dokładnością do logarytmu) zachodzi też odwrotna implikacja. By ją dokładnie sformułować zdefiniujmy

$$D_n := \max\left\{D_{\mathrm{Che}}(\mu)\colon\ \mu \text{ log-wklęsła izotropowa w } \mathbb{R}^n\right\}$$

oraz

$$\sigma_n := \max\left\{\left(\mathbb{E}(|X|-\sqrt{n})^2\right)^{1/2}\colon\ X \text{ log-wklęsły izotropowy w } \mathbb{R}^n\right\}$$

Twierdzenie 14.9 (Eldan)**.** *Dla dowolnego $n \geqslant 2$,*

$$D_n \leqslant C\sqrt{\log n \sum_{k=1}^{n} \frac{\sigma_k^2}{k}}.$$

Uwaga 14.10. Powyższe szacowanie w przeciwieństwie do poprzednio omawianych szacowań jest „globalne”. Nie wiadomo czy jeśli X jest izotropowy o rozkładzie μ oraz umiemy oszacować $\mathbb{E}(|X|-\sqrt{n})^2$ (oraz być może podobne wielkości dla rzutów X), to otrzymujemy ograniczenie dla $D_{\mathrm{Poin}}(\mu)$.

Wielkość σ_n ogranicza też stałe izotropowe. Mianowicie niech

$$L_n := \sup\left\{L_K\colon\ K \text{ ciało wypukłe w } \mathbb{R}^n\right\}.$$

Twierdzenie 14.11 (Eldan-Klartag)**.** *Dla dowolnego n, $L_n \leqslant C\sigma_n$.*

Wynik ten pokazuje w szczególności, że Hipoteza 14.8 implikuje hyperplane conjecture, zatem również prawdziwość hipotezy KLS pociągałaby za sobą ograniczoność stałych izotropowych – ten ostatni wynik został wcześniej zaanonsowany przez Balla (choć nie był opublikowany). Twierdzenie Eldana-Klartaga też jest „globalne”. Ball i Nguyen niedawno wykazali, że $L_\mu \leqslant \exp(16 D_{\mathrm{Poin}}(\mu))$ dla izotropowej miary log-wklęsłej μ.

14.3 Nierówności splotu infimum i pokrewne nierówności koncentracyjne

14.4 Porównywanie słabych i silnych momentów

Literatura

[1] E. Gluskin, *The diameter of Minkowski compactum roughly equals to n*, Funct. Anal. Appl. 15 (1981), 57–58.

[2] M. Ledoux, *The concentration of measure phenomenon*, AMS 2001.

[3] M. Ledoux, M. Talagrand, *Probability in Banach spaces. Isoperimetry and processes*, Springer-Verlag, Berlin, 1991.

[4] V.D. Milman, G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Math. 1200, Springer 1986.

[5] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Univ. Press 1989.