Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Marcin Socha
Nr albumu: 418253

Nierówności koncentracyjne dla miar produktowych

Praca licencjacka
na kierunku MATEMATYKA

Praca wykonana pod kierunkiem
dr hab. Radosława Adamczaka

Warszawa, wrzesień 2022

Streszczenie

W pracy omówiono pewne problemy oszacowań probabilistycznych za pomocą nierówności. W szczególności przedyskutowano wyprowadzenie nierówności Azumy, Chinczyna i Talagranda oraz zagadnienie zjawiska koncentracji miary. Praca ma charakter kompilacyjny, zawiera zmodyfikowane dowody podstawowych twierdzeń uzupełnione o analityczne przykłady zastosowań: pakowania do koszy, problemu komiwojażera, drzewa Steinera. Ponadto omówiono związek zjawiska koncentracji miary z funkcjami lipschitzowskimi oraz przeprowadzono porównanie nierówności Talagranda z nierównościami wynikającymi z nierówności Azumy.

Słowa kluczowe

Nierówność Azumy, zjawisko koncentracji miary, nierówność Talagranda.

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.1 Matematyka

Klasyfikacja tematyczna

60E15 Inequalities; stochastic orderings

Tytuł pracy w języku angielskim

Concentration inequalities for product measures

Pragnę złożyć najszczersze podziękowania Promotorowi niniejszej pracy Panu Profesorowi Radosławowi Adamczakowi za poświęcony czas oraz cenne rady, a także za wsparcie i wyrozumiałość.

Spis treści

1. Wprowadzenie 7

2. Wstępne definicje i podstawowe nierówności 9

2.1. Podstawowe definicje 9

2.1.1. Funkcje koncentracyjne i lipschitzowskie 9

2.8.1. Podstawowe fakty teorii martyngałów 11

2.12.1. Klasyczne definicje oraz nierówności 11

2.16. Nierówność Azumy 12

2.25. Wnioski z nierówności Azumy 19

2.25.1. Nierówność Hoeffdinga 19

2.26.1. Nierówność Chinczyna 19

2.27.1. Pakowanie do koszy [7] 20

3. Nierówność *convex distance* Talagranda 21

3.2.1. Nierówność convex distance 23

3.4.1. Przykłady 27

3.7.1. Funkcje konfiguracyjne [7] 32

3.11.1. Problem komiwojażera [7] 36

3.12.1. Drzewo Steinera [7] 37

4. Podsumowanie 39

Bibliografia 41

Rozdział 1

Wprowadzenie

Problemy nierówności izoperymetrycznych były rozważane już w starożytnej Grecji [8, 2]. Szukano wtedy figur geometrycznych o maksymalnym polu przy ustalonym obwodzie. Na początku próbowano rozwiązywać je doświadczalnie, ale w końcu rozwiązania zaczynały przypominać dzisiejsze dowody. Jedną z pierwszych osób, która zajmowała się tymi zagadnieniami był Zenodor (III i II w.p.n.e). Pierwsze współczesne dowody nierówności izoperymetrycznych przedstawił w XIX w. Steiner, zaś w pełni poprawne rozwiązania problemu izoperymetrycznego w przestrzeni $\mathbf{R}^n$ pojawiły się dopiero w XX w. dzięki Schwartzowi. W drugiej połowie XX w. zauważono, że nierówności izoperymetryczne dla miar innych niż Lebesgue'a, na przykład dla miary powierzchniowej na sferze, mają zastosowania w rachunku prawdopodobieństwa i prowadzą do daleko idących wzmocnień klasycznych nierówności Czebyszewa czy Markowa. Związek ten został rozszerzony na przypadek miar probabilistycznych na przestrzeniach produktowych, w których w zależności od rozpatrywanej odległości otrzymuje się różne nierówności dla odchyleń zmiennych losowych od swojej średniej lub mediany.

Celem pracy jest przybliżenie Czytelnikowi zagadnień związanych z nierównościami Azumy, Hoeffdinga oraz Talagranda oraz ich zastosowań. Pokażemy na przykładach, że czasami niewielka zmiana założeń może przynosić nieintuicyjnie mocniejsze rezultaty. Praca składa się z trzech rozdziałów. W rozdziale 1 wprowadzimy kilka podstawowych pojęć potrzebnych w dalszej części pracy oraz pokażemy w jaki sposób łączy się ze sobą zjawisko koncentracji miary z funkcjami lipschitzowskimi. W rozdziale 2 pokażemy dowód nierówności Azumy, dzięki której można oszacować bardzo dobrze prawdopodobieństwo odchyleń niektórych funkcji od ich wartości oczekiwanych. W kolejnych podrozdziałach udowodnimy korzystając z nierówności Azumy nierówność Hoeffdinga oraz Chinczyna. Pokażemy również przykład zastosowania nierówności różnic skończonych. W rozdziale 3 wprowadzimy odległość convex distance oraz pokażemy, że udowodniona przez Michela Talagranda [13] nierówność wykorzystująca zjawisko koncentracji miary daje znacznie lepsze wyniki.

Rozdział 2

Wstępne definicje i podstawowe nierówności

2.1. Podstawowe definicje

2.1.1. Funkcje koncentracyjne i lipschitzowskie

W tym podrozdziale zajmiemy się funkcjami opisującymi zjawisko koncentracji miary oraz ich znaczeniu w kontekście funkcji lipschitzowskich.

Definicja 2.2. [5] Oznaczmy przez (X, d) przestrzeń metryczną z metryką d. Funkcję $F : (X, d) \to \mathbf{R}$ nazywamy *lipschitzowską* jesli

$$\|F\|_{Lip} := \sup_{x \neq y} \frac{|F(x) - F(y)|}{d(x, y)} < \infty. \tag{2.1}$$

Mówimy, że funkcja jest *L-lipschitzowską* jeśli $\|F\|_{Lip} \leq L$, tzn. $|F(x) - F(y)| \leq Ld(x, y)$ dla wszystkich $x, y \in X$.

Definicja 2.3. [5] Niech (X, d) będzie przestrzenią metryczną z metryką d oraz zbiór A będzie dowolnym podzbiorem X. Niech $d(x, A) := \inf_{a \in A} d(x, a)$. Wówczas jeśli $t > 0$, to *t-otoczeniem zbioru A* nazywamy zbiór

$$A_t := \{x \in X : d(x, A) < t\} = \bigcup_{y \in A} B(y, t), \tag{2.2}$$

gdzie $B(y, t)$ oznacza kulę otwartą w X o środku w y oraz promieniu t.

Okazuje się, że w wielu ważnych przykładach miar probabilistycznych ψ na (X, d) jeśli $\psi(A) \geq 1/2$, to miara $\psi(A_t)$ szybko zbiega do 1 gdy $t \to \infty$. Badaniem tego zjawiska, zwanego *zjawiskiem koncentracji miary* zajmiemy się w kolejnych rozdziałach.

Ważnym aspektem w pracy jest zrozumienie powiązania między koncentracją zapisaną w formie odległości od zbioru, a zapisaną w języku funkcji lipschitzowskich. Z uwagi na to, że jest to istotne, może nie tyle dla dowodu, co filozoficznego zrozumienia nierówności Talagranda, warto się temu przyjrzeć. Wprowadźmy więc definicje.

Definicja 2.4. [5] Niech ψ będzie miarą probabilistyczną na (X, d). *Funkcją koncentracji miary* ψ nazywamy funkcję α_ψ zdefiniowaną jako

$$\alpha_\psi(t) = \alpha_{(X,d,\psi)}(t) := \sup\left\{1 - \psi(A_t) : \psi(A) \geq \frac{1}{2}, A \in B(X)\right\}, \tag{2.3}$$

gdzie $B(X)$ *to* σ*-ciało zbiorów borelowskich.*

Definicja 2.5. Niech ψ będzie miara probabilistyczną na (X, d) oraz $f : X \to \mathbf{R}$. *Medianą funkcji* f *względem miary* ψ nazywamy taką liczbę $m = Med_\psi(f)$, że

$$\psi\left(\{x : f(x) \geq m\}\right) \geq \frac{1}{2} \tag{2.4}$$

oraz

$$\psi(\{x : f(x) \leq m\}) \geq \frac{1}{2}. \tag{2.5}$$

Modułem ciągłości funkcji f nazywamy funkcję $\omega_f : \mathbf{R}_+ \to \mathbf{R}_+ \cup \{+\infty\}$

$$w_f(t) := \sup_{x,y \in X}\left(|f(x) - f(y)| : d(x, y) \leq t\right). \tag{2.6}$$

Warto zauważyć, że mediana zawsze istnieje, lecz nie musi być jednoznacznie wyznaczona. W dalszej części pracy będziemy pisali zazwyczaj, że $Med_\psi(f) = m$.

Lemat 2.6. [5] *Jeśli* $f : X \to \mathbf{R}$ *jest dowolną funkcją, to*

$$\psi(\{x : f(x) > m + w_f(t)\}) \leq \alpha_\psi(t). \tag{2.7}$$

Dowód. [5] Oznaczmy przez A zbiór $\{x : f(x) \leq m\}$. Zatem oczywiście $\psi(A) \geq \frac{1}{2}$. Z (2.3) wynika, że $\alpha_\psi(t) \geq 1 - \psi(A_t)$, zatem $\psi(A_t) \geq 1 - \alpha_\psi(t)$. Jeśli dodatkowo $x \in A_t$, to istnieje $y \in A$, taki że $d(x, y) < t$, więc $f(x) - f(y) \leq w_f(t)$, zatem $f(x) \leq f(y) + w_f(t)$. Dalej $f(x) \leq f(y) + w_f(t) \leq m + w_f(t)$ (bo $f(y) \leq m$ z definicji zbioru A), skąd wynika teza. □

Lemat 2.7. [5] *Niech* f *będzie funkcją lipschitzowską ze stałą* L*. Wówczas dla każdego* $t > 0$ *zachodzi*

$$\psi(\{x : f(x) > m + t\}) \leq \alpha_\psi(\frac{t}{L}). \tag{2.8}$$

Dowód. [5] Zauważmy, że skoro f jest funkcją lipschitzowską, to $w_f(s) \leq sL$, więc podstawiając w lemacie 2.6 $t := \frac{s}{L}$ dostajemy tezę. □

Lemat 2.8. [5] *Jeśli dla każdej funkcji 1-lipschitzowskiej* f *i ustalonego* $t \geq 0$ *istnieje pewna* $\alpha \in \mathbf{R}$*, taka że zachodzi*

$$\psi(\{x : f(x) \geq m + t\}) \leq \alpha, \tag{2.9}$$

to $\alpha_\psi(t) \leq \alpha$*.*

Dowód. [5] Niech A będzie ustalonym zbiorem w przestrzeni X, takim że $\psi(A) \geq \frac{1}{2}$ oraz $f(x) := d(x, A)$. Wówczas oczywiście f jest lipschitzowska ze stałą 1 oraz jeśli $x \in A$, to $f(x) = 0$, więc mediana f jest równa 0. W takim razie

$$\alpha \geq \psi(\{x : f(x) \geq t\}) = \psi(\{x : d(x, A) \geq t\}) = 1 - \psi(A_t), \tag{2.10}$$

skąd wynika teza. □

2.8.1. Podstawowe fakty teorii martyngałów

Przypomnimy teraz podstawowe definicje i fakty dotyczące teorii martyngałów potrzebnych nam w zrozumieniu późniejszych twierdzeń. Ich dowody można znaleźć np. w [4]. Zakładamy, że mamy ustalona przestrzeń probabilistyczna $(\Omega, \mathcal{F}, \mathbf{P})$. Niech $T \subseteq \{0, 1, 2, ..\}$ będzie dowolnym podzbiorem liczb naturalnych.

Definicja 2.9. [4] *Filtracją* nazywamy niemalejącą rodzinę σ-ciał $(\mathcal{F}_t)_{t\in T}$, czyli dla $t_1, t_2 \in T$ zachodzi $t_1 < t_2 \Leftrightarrow \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$, gdzie $\mathcal{F}_t \subseteq \mathcal{F}$ dla $t \in T$.

Definicja 2.10. [4] Rodzina zmiennych losowych $(X_t)_{t\in T}$ jest *adaptowalna do filtracji* $(\mathcal{F}_t)_{t\in T}$, jeśli zmienne losowe X_t są $\mathcal{F}_t$-mierzalne dla wszystkich $t \in T$

W szczególności rodzina zmiennych losowych $(X_t)_{t\in T}$ jest adaptowalna do *naturalnej filtracji* $(\mathcal{F}_t)_{t\in T}$, gdzie $\mathcal{F}_t = \sigma(X_s : s \leq t)$, $t \in T$.

Definicja 2.11. [4] Niech $\mathcal{F}$ będzie σ-ciałem, a X zmienną losową całkowalną. *Warunkową wartością oczekiwaną* X nazywamy zmienną losową $\mathbf{E}[X|\mathcal{F}]$ spełniającą warunki
i) $\mathbf{E}[X|\mathcal{F}]$ jest $\mathcal{F}$-mierzalna,
ii) dla każdego $A \in \mathcal{F}$

$$\int_A X d\mathbf{P} = \int_A \mathbf{E}[X|\mathcal{F}]\, d\mathbf{P}. \tag{2.11}$$

Zakładamy, że mamy ustalone filtracje $(\mathcal{F}_t)_{t\in T}$ i adaptowany do niej ciąg zmiennych losowych $(X_t)_{t\in T}$.

Definicja 2.12. [4] Rodzina $(X_t, \mathcal{F}_t)_{t\in T}$, gdzie zmienne losowe X_t są całkowalne dla $t \in T$ jest *martyngałem*, jeśli dla $s \leq t, s, t \in T$ zachodzi

$$\mathbf{E}[X_t|\mathcal{F}_s] = X_s. \tag{2.12}$$

2.12.1. Klasyczne definicje oraz nierówności

W tym podrozdziale wprowadzimy kilka podstawowych definicji oraz nierówności potrzebnych w dalszej części pracy.

Przypomnijmy klasyczną nierówność Markowa.

Twierdzenie 2.13. [4] *Niech X będzie nieujemną zmienną losową o wartości oczekiwanej μ. Wówczas dla każdego $\epsilon > 0$ zachodzi*

$$\mathbf{P}(X \geq \epsilon) \leq \frac{\mu}{\epsilon}. \tag{2.13}$$

Wprowadzimy teraz definicję transformaty Laplace'a i udowodnimy proste twierdzenie pomocnicze, które przyda nam się do wzmocnienia nierówności Markowa przy dodatkowych założeniach dotyczących całkowalności.

Definicja 2.14. [5] Dla $\lambda \in \mathbf{R}$ *transformatą Laplace'a zmiennej losowej Z* nazywamy

$$L_Z(\lambda) := \mathbf{E}\left[e^{\lambda Z}\right]. \tag{2.14}$$

Podobnie dla miary probabilistycznej ψ określonej na pewnej przestrzeni X oraz dla funkcji $F : X \to \mathbf{R}$ *transformatą Laplace'a* określamy

$$L_{F,\psi}(\lambda) := \int_X e^{\lambda F(x)} d\psi(x). \tag{2.15}$$

Twierdzenie 2.15. [5] *Niech Z będzie dowolną zmienną losową. Wówczas dla dowolnego $t \geq 0$*

$$\mathbf{P}(Z \geq t) \leq \inf_{\lambda \geq 0} e^{-\lambda t} L_Z(\lambda). \tag{2.16}$$

Dowód. [5] Zauważmy, że dla dowolnych $\lambda \in \mathbf{R}_+$ zachodzi nierówność $e^{Z\lambda} \geq 0$, więc korzystając z nierówności Markowa otrzymujemy

$$\begin{aligned}\mathbf{P}(Z \geq t) &= \mathbf{P}(e^Z \geq e^t) \\ &= \inf_{\lambda \geq 0} \mathbf{P}(e^{\lambda Z} \geq e^{\lambda t}) \\ &\leq \inf_{\lambda \geq 0} \frac{\mathbf{E}\left[e^{\lambda Z}\right]}{e^{\lambda t}}. \end{aligned} \tag{2.17}$$

□

2.16. Nierówność Azumy

Efektywnym sposobem szacowania odchyleń zmiennej losowej jest przedstawienie jej w postaci sumy różnic martyngałowych i oszacowanie warunkowych transformat Laplace'a. W przypadku gdy różnice martyngałowe są ograniczone, odpowiednie oszacowanie zostało udowodnione przez Azumę.

Twierdzenie 2.17. [1] *Niech $(M_k, \mathcal{F}_k)_{k=0}^n$ będzie martyngałem o ograniczonych przyrostach, takim że $a_k \leq M_k - M_{k-1} \leq b_k$, gdzie a_k, b_k są pewnymi stałymi skończonymi. Wówczas*

$$\mathbf{P}\left(|M_n - M_0| \geq t\right) \leq 2\exp\left(-\frac{2t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right). \tag{2.18}$$

Dowód nierówności Azumy. Oznaczmy przez $d_k := M_k - M_{k-1}$ dla $k \in \{1, 2, ..., n\}$, wówczas $\mathbf{E}\left[d_k | \mathcal{F}_{k-1}\right] = 0$. Skoro funkcja $f(x) = \exp(\lambda x)$ jest funkcją wypukłą dla $\lambda > 0$, więc

$$\begin{aligned}\mathbf{E}\left[e^{\lambda d_k} \middle| \mathcal{F}_{k-1}\right] &\leq \mathbf{E}\left[\left(\frac{d_k - a_k}{b_k - a_k} e^{\lambda b_k} + \frac{b_k - d_k}{b_k - a_k} e^{\lambda a_k}\right) \middle| \mathcal{F}_{k-1}\right] \\ &= \frac{e^{\lambda b_k}}{b_k - a_k} \mathbf{E}\left[(d_k - a_k)\,|\mathcal{F}_{k-1}\right] + \frac{e^{\lambda a_k}}{b_k - a_k} \mathbf{E}\left[(b_k - d_k)\,|\mathcal{F}_{k-1}\right] \\ &= \frac{e^{\lambda a_k} b_k - e^{\lambda b_k} a_k}{b_k - a_k} \\ &= (1-p)e^{-py} + pe^{(1-p)y} \\ &= e^{-py}(1 - p + pe^y) = e^{f(y)}, \end{aligned} \tag{2.19}$$

gdzie

$$p = \frac{a_k}{a_k - b_k}, \; y = (b_k - a_k)\,\lambda, \; f(y) = -py + \ln(1 - p + pe^y). \tag{2.20}$$

Zauważmy, że

$$\begin{aligned} f'(y) &= -p + \frac{pe^y}{(1-p) + pe^y} \\ &= -p + \frac{p}{p + (1-p)e^{-y}} \end{aligned} \tag{2.21}$$

oraz

$$f''(y) = \frac{p(1-p)e^{-y}}{(p+(1-p)e^{-y})^2} \le \frac{1}{4}. \tag{2.22}$$

Nierówność (2.22) wynika z nierówności między średnią arytmetyczną i geometryczną dla dwóch dodatnich liczb

$$\frac{p}{p+(1-p)e^{-y}}, \quad \frac{(1-p)e^{-y}}{p+(1-p)e^{-y}}. \tag{2.23}$$

Powyższe liczby są dodatnie, ponieważ $a_k \le 0 \le b_k$, więc $p \in [0,1)$. Zauważmy, że $f(0) = f'(0) = 0$, więc ze wzoru Taylora dla funkcji $f(y)$ otrzymujemy

$$f(y) \le \frac{1}{8}y^2 = \frac{1}{8}(b_k - a_k)^2\lambda^2. \tag{2.24}$$

Z nierówności (2.19) oraz (2.24) dostajemy, że

$$\begin{aligned}
\mathbf{E}\left[e^{\lambda(M_n - M_0)}\right] &= \mathbf{E}\left[e^{\lambda(M_{n-1} - M_0 + d_n)}\right] \\
&= \mathbf{E}\left[e^{\lambda(M_{n-1} - M_0)}\mathbf{E}\left[e^{\lambda d_n}\middle| \mathcal{F}_{n-1}\right]\right] \\
&\le \exp\left(\frac{1}{8}(b_n - a_n)^2\lambda\right)\mathbf{E}\left[e^{\lambda(M_{n-1} - M_0)}\right].
\end{aligned} \tag{2.25}$$

Iterując nierówność (2.25), tzn. zmieniając $(M_n - M_0)$ w wykładniku $\mathbf{E}[e^{(M_n - M_0)}]$ na $\mathbf{E}[e^{(M_{n-1} - M_0)}]$ dostajemy

$$\begin{aligned}
L_{(M_n - M_0)}(\lambda) &= \mathbf{E}\left[e^{\lambda(M_n - M_0)}\right] \\
&\le \prod_{k=1}^{n} \exp\left(\frac{1}{8}(b_k - a_k)^2\lambda^2\right) \\
&= \exp\left(\frac{1}{8}\sum_{k=1}^{n}(b_k - a_k)^2\lambda^2\right).
\end{aligned} \tag{2.26}$$

Stąd oraz korzystając z twierdzenia 2.15 dostajemy, że

$$\begin{aligned}
\mathbf{P}\left((M_n - M_0) \ge t\right) &\le \inf_{\lambda \ge 0}\left(e^{-\lambda t}\mathbf{E}\left[e^{\lambda(M_n - M_0)}\right]\right) \\
&\le \inf_{\lambda \ge 0} e^{-\lambda t}\left(\exp\left(\frac{1}{8}\sum_{k=1}^{n}(b_k - a_k)^2\lambda^2\right)\right) \\
&= \inf_{\lambda \ge 0}\left(\exp\left(-\lambda t + \frac{1}{8}\sum_{k=1}^{n}(b_k - a_k)^2\lambda^2\right)\right).
\end{aligned} \tag{2.27}$$

Zauważmy, że z nierówności $\sum_{k=1}^{n}(b_k - a_k)^2 \ge 0$ oraz z faktu, że funkcja ze względu na λ postaci $-\lambda t + \frac{1}{8}\sum_{k=1}^{n}(b_k - a_k)^2\lambda^2$ jest kwadratowa i osiąga minimum równe $-(2t^2)/\left(\sum_{k=1}^{n}(b_k - a_k)^2\right)$ wynika, że nierówność (2.27) przyjmuje postać

$$\mathbf{P}(M_n - M_0 \ge t) \le \exp\left(-\frac{2t^2}{\sum_{k=1}^{n}(b_k - a_k)^2}\right). \tag{2.28}$$

W takim razie

$$\mathbf{P}(|M_n - M_0| \geq t) = \mathbf{P}(M_n - M_0 \geq t) + \mathbf{P}(-M_n + M_0 \geq t) \\ \leq 2\exp\left(-\frac{2t^2}{\sum_{k=1}^{n}(b_k - a_k)^2}\right), \tag{2.29}$$

ponieważ $(-M_k)$ również jest martyngałem oraz spełnia założenia twierdzenia postaci $-b_k \leq -M_k - (-M_{k-1}) \leq -a_k$, więc możemy również zastosować (2.28). $\square$

Wnioskiem z nierówności Azumy jest na przykład nierówność różnic skończonych.

Twierdzenie 2.18. *Niech $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ będzie rodziną niezależnych zmiennych losowych, gdzie dla każdego k, X_k będzie przyjmowała wartości z pewnego zbioru A_k. Załóżmy, że pewna funkcja rzeczywista f zdefiniowana na $\prod_{k=1}^{n} A_k$ spełnia nierówność*

$$|f(\boldsymbol{x}) - f(\boldsymbol{x}')| \leq c_k \tag{2.30}$$

dla dowolnych wektorów $\boldsymbol{x}$ oraz $\boldsymbol{x}'$ różniących się jedynie na k-tej współrzędnej. Wówczas dla każdego $t \geq 0$ zachodzi

$$\mathbf{P}\left((f(\boldsymbol{X}) - \mathbf{E}[f(\mathbf{X})]) \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{k=1}^{n} c_k^2}\right). \tag{2.31}$$

Dowód. Niech $\mathcal{F}_k$ oznacza σ-ciało generowane przez zmienne losowe $X_1, X_2, ..., X_k$. Wówczas zachodzi $\mathcal{F}_k \subseteq \mathcal{F}_{k+1}$. Niech $M_k = \mathbf{E}\left[f(X_1, X_2, ..., X_n)|\mathcal{F}_k\right]$. Łatwo sprawdzić, że M_k jest martyngałem, bowiem

$$\begin{aligned} \mathbf{E}\left[(M_{k+1}|\mathcal{F}_k)\right] &= \mathbf{E}\left[(\mathbf{E}[(f(X_1, X_2, ..., X_n)|\mathcal{F}_{k+1})]|\mathcal{F}_k)\right] \\ &= \mathbf{E}\left[f(X_1, X_2, ..., X_n)|\mathcal{F}_k\right] \\ &= M_k. \end{aligned} \tag{2.32}$$

Oznaczmy przez $\mathbf{E}_{X_{t_1},...,X_{t_k}}\left[f(X_1, ..., ..., X_n)\right]$ wartość oczekiwaną funkcji f względem zmiennych $X_{t_1}, ..., X_{t_k}$. Stąd widzimy, że $M_n = f(X_1, X_2, ..., X_n)$ oraz $M_0 = \mathbf{E}\left[f(X_1, X_2, ..., X_n)\right]$ i zdefiniujmy $\mu = M_0$. W takim razie

$$\begin{aligned} |M_k - M_{k-1}| &= \left|\mathbf{E}_{X_{k+1},...,X_n}\left[f(X_1, X_2, ..., X_n)\right] - \mathbf{E}_{X_k,...,X_n}\left[f(X_1, X_2, ..., X_n)\right]\right| \\ &= \left|\mathbf{E}_{\hat{X}_k, X_{k+1},...,X_n}\left[(f(X_1, ..., X_k, ..., X_n) - f(X_1, ..., \hat{X}_k, ..., X_n))\right]\right|, \end{aligned} \tag{2.33}$$

gdzie $\hat{X}_k$ jest niezależną kopią X_k. Niech $\mathbf{x} = (X_1, ..., X_k..., X_n)$ oraz $\mathbf{x}' = (X_1, ..., \hat{X}_k..., X_n)$. Zauważmy, że z nierówności $|f(\mathbf{x}) - f(\mathbf{x}')| \leq c_k$ wynika

$$|M_k - M_{k-1}| \leq c_k, \tag{2.34}$$

a tym samym spełnione są nierówności występujące w założeniach twierdzenia 2.17 (nierówność Azumy) $a_k \leq M_k - M_{k-1} \leq b_k$, gdzie a_k, b_k są pewnymi stałymi skończonymi. Oznaczmy przez

$$A_k = \inf_x \left(\mathbf{E}_{X_1,...,X_{k-1},x}[f(\mathbf{X})] - \mathbf{E}_{X_1,...,X_{k-1}}[f(\mathbf{X})]\right), \tag{2.35}$$

$$B_k = \sup_x \left(\mathbf{E}_{X_1,...,X_{k-1},x}[f(\mathbf{X})] - \mathbf{E}_{X_1,...,X_{k-1}}[f(\mathbf{X})]\right), \tag{2.36}$$

dla $k = 1, ..., n$. Wówczas widzimy, że

$$A_k \leq M_k - M_{k-1} \leq B_k. \tag{2.37}$$

Zauważmy, że podobnie jak w równaniach (2.33) oraz nierówności (2.34)

$$\begin{aligned} 0 \leq B_k - A_k &= \sup_x \left(\mathbf{E}_{X_1,...,X_{k-1},x}[f(\mathbf{X})] - \mathbf{E}_{X_1,...,X_{k-1}}[f(\mathbf{X})]\right) \\ &\quad - \inf_x \left(\mathbf{E}_{X_1,...,X_{k-1},x}[f(\mathbf{X})] - \mathbf{E}_{X_1,...,X_{k-1}}[f(\mathbf{X})]\right) \\ &= \sup_{x,y} \mathbf{E}_{X_{k+1},...,X_n}[f(X_1,...,X_{k-1},x,X_{k+1},...,X_n) - f(X_1,...,X_{k-1},y,X_{k+1},...,X_n)] \\ &\leq c_k. \end{aligned} \tag{2.38}$$

Stąd z nierówności (2.38) widzimy, że możemy skorzystać z nierówności Azumy (twierdzenie 2.17)

$$\mathbf{P}(f(\mathbf{X}) - \mu \geq t) = \mathbf{P}(M_n - M_0) \geq t \leq \exp\left(-\frac{2t^2}{\sum_{k=1}^n (B_k - A_k)^2}\right) \leq \exp\left(-\frac{2t^2}{\sum_{k=1}^n c_k^2}\right). \tag{2.39}$$

□

Zinterpretujmy tę nierówność w terminach koncentracji dla ważonych odległości Hamminga. Wprowadźmy więc na początku definicje.

Definicja 2.19. Niech $\Omega_1, ..., \Omega_n$ będą pewnymi przestrzeniami oraz niech Ω oznacza przestrzeń produktową $\prod_{k=1}^n \Omega_k$. Niech $\alpha = (\alpha_1, ..., \alpha_n)$ będzie n-wymiarowym wektorem o nieujemnych współrzędnych. Jeśli $\mathbf{x} = (x_1, ..., x_n), \mathbf{y} = (y_1, ..., y_n) \in \Omega$, to wówczas *odległością α-Hamminga* $d_\alpha(\mathbf{x}, \mathbf{y})$ nazywamy sumę wartości α_i, jeśli $x_i \neq y_i$, zaś dla $A \subseteq \Omega$ definiujemy jako $d_\alpha(\mathbf{x}, A) = \inf\{d_\alpha(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in A\}$.

Definicja 2.20. [7] *Odległość Hamminga* $d_H(\mathbf{x}, \mathbf{y})$ definiujemy jako $d_\alpha(\mathbf{x}, \mathbf{y})$ dla $\alpha = (1, ..., 1)$.

Definicja 2.21. Przestrzenią l_2^n nazwiemy zbiór wektorów o nieujemnych współrzędnych $\alpha = (\alpha_1, ..., \alpha_n)$ z normą zdefiniowaną wzorem $\|\alpha\|_2 := \sqrt{\sum_{k=1}^n \alpha_k^2}$. Wektor α będziemy nazywali *wektorem jednostkowym,* jeśli $\|\alpha\|_2 = 1$.

Zacytujemy twierdzenie, które pokazuje związek między nierównością różnic skończonych oraz zjawiskiem koncentracji miary.

Twierdzenie 2.22. [7] *Niech $\mathbf{X} = (X_1, ..., X_n)$ będzie rodziną niezależnych zmiennych losowych oraz niech $\alpha = (\alpha_1, ..., \alpha_n)$ będzie n-wymiarowym wektorem jednostkowym o współrzędnych nieujemnych z przestrzeni l_2, zaś niech A będzie podzbiorem przestrzeni produktowej. Wówczas dla każdego $t \geq 0$ zachodzi*

$$\mathbf{P}(\mathbf{X} \in A)\mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t) \leq \exp\left(-\frac{t^2}{2}\right). \tag{2.40}$$

Dowód. [7] Niech $\rho = \mathbf{P}(\mathbf{X} \in A)$ oraz $\mu = \mathbf{E}[d_\alpha(\mathbf{X}, A)]$. Będziemy zakładali, że $\rho > 0$. Zauważmy, że jeśli wektory $\mathbf{X}$ oraz $\mathbf{X}'$ różnią się na k-tej współrzędnej, to z nierówności trójkąta zachodzi

$$d_\alpha(\mathbf{X}, A) \leq d_\alpha(\mathbf{X}', A) + d_\alpha(\mathbf{X}, \mathbf{X}'), \tag{2.41}$$

co jest równoważne stwierdzeniu, że

$$d_\alpha(\mathbf{X}, A) - d_\alpha(\mathbf{X}', A) \leq d_\alpha(\mathbf{X}, \mathbf{X}') = \alpha_k. \tag{2.42}$$

Stąd możemy skorzystać z nierówności różnic skończonych (twierdzenie 2.18), z której dostajemy dla $t \geq 0$

$$\mathbf{P}\left(d_\alpha(\mathbf{X}, A) - \mu \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{k=1}^n \alpha_k^2}\right) = \exp(-2t^2) \tag{2.43}$$

oraz podobnie

$$\mathbf{P}(d_\alpha(\mathbf{X}, A) - \mu \leq -t) \leq \exp(-2t^2). \tag{2.44}$$

Zauważmy, że jeśli $\mathbf{x} \in A$, to $d_\alpha(\mathbf{x}, A) = 0$, więc podstawiając w nierówności (2.44) w miejsce $t = \mu$ dostajemy

$$\begin{aligned} \rho = \mathbf{P}(\mathbf{X} \in A) &= \mathbf{P}(d_\alpha(\mathbf{X}, A) \leq 0) \\ &= \mathbf{P}(d_\alpha(\mathbf{X}, A) - \mu \leq -\mu) \leq \exp(-2\mu^2), \end{aligned} \tag{2.45}$$

w takim razie

$$\mu \leq \left(\frac{1}{2}\ln(\frac{1}{\rho})\right)^{\frac{1}{2}}, \tag{2.46}$$

w związku z tym $\mu \leq t_0$, gdzie $t_0 = \left(\frac{1}{2}\ln(\frac{1}{\rho})\right)^{\frac{1}{2}}$. Wykorzystując ograniczenie t_0 w nierówności (2.43) w miejsce μ dostajemy

$$\begin{aligned} \mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t + t_0) &\leq \mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t + \mu) \\ &\leq \exp(-2t^2). \end{aligned} \tag{2.47}$$

Zatem w nierówności (2.47) dla $t \geq t_0$ podstawiając w miejsce t różnicę $(t - t_0)$ dostajemy

$$\mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t) \leq \exp(-2(t - t_0)^2). \tag{2.48}$$

Zauważmy, że dla $t \geq 2a$ zachodzi $(t-a)^2 \geq \frac{t^2}{4}$, więc jeśli weźmiemy $t \geq 2t_0$, to zachodzi $(t - t_0)^2 \geq \frac{t^2}{4}$, więc z nierówności (2.48) dostajemy

$$\mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t) \leq \exp(-2(t - t_0)^2) \leq \exp\left(-\frac{t^2}{2}\right). \tag{2.49}$$

Zauważmy, że z definicji t_0 wynika, że $\exp(-2t_0^2) = \rho$. Natomiast dla $0 \leq t \leq 2t_0$ prawa strona nierówności (2.44) jest równa co najmniej $\exp(-2t_0^2) = \rho = \mathbf{P}(\mathbf{X} \in A)$. Zatem dla dowolnego $t \geq 0$

$$\min\left\{\mathbf{P}(\mathbf{X} \in A), \mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t)\right\} \leq \exp\left(-\frac{t^2}{2}\right), \tag{2.50}$$

skąd wynika teza. □

Twierdzenie 2.22 można uogólnić przy słabszych założeniach (dla dowolnego wektora $\alpha = (\alpha_1, ..., \alpha_n)$) i przeprowadzając analogiczny dowód otrzymujemy następujące twierdzenie.

Twierdzenie 2.23.

$$\mathbf{P}(\mathbf{X} \in A)\mathbf{P}(d_\alpha(\mathbf{X}, A) \geq t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n \alpha_k^2}\right). \tag{2.51}$$

Twierdzenie 2.22 da nam wyniki podobne do nierówności różnic skończonych (twierdzenie 2.18) jednak zamiast wartości oczekiwanej, jak we wzorze (2.31) otrzymamy wzór zawierający medianę. Przyjrzyjmy się bliżej analizie różnicy między wartością oczekiwaną i medianą. W tym celu zacytujemy następujący lemat.

Lemat 2.24. [7] *Niech X będzie zmienną losową z wartością oczekiwaną równą μ, medianą m. Niech ponadto a, b będą stałymi nieujemnymi.*

a) Jeśli $\mathbf{P}(X - m \geq t) \leq a \exp\left(-\frac{t^2}{b}\right)$ dla każdego $t \geq 0$, to

$$\mu - m \leq a\sqrt{b}(\sqrt{\pi}/2). \tag{2.52}$$

Jeśli ponadto $\mathbf{P}(X - m \leq -t) \leq a \exp\left(-\frac{t^2}{b}\right)$ dla każdego $t \geq 0$, to

$$|\mu - m| \leq a\sqrt{b}(\sqrt{\pi}/2). \tag{2.53}$$

b) Jeśli $\mathbf{P}(X - m \geq t) \leq a \exp\left(-\frac{t^2}{b(m+t)}\right)$ dla każdego $t > 0$, to

$$\mu - m \leq a\sqrt{(\pi/2)bm} + 2ab \exp\left(-\frac{m}{2b}\right). \tag{2.54}$$

Dowód. Zauważmy, że

$$\mu - m = \mathbf{E}[X - m] \leq \mathbf{E}\left[(X - m)^+\right] = \int_0^\infty \mathbf{P}(X - m > t)dt. \tag{2.55}$$

Skoro $\mathbf{P}(X - m \geq t) \leq a \exp\left(-\frac{t^2}{b}\right)$, to

$$\int_0^\infty \mathbf{P}(X - m > t)dt \leq \int_0^\infty a \exp\left(-\frac{t^2}{b}\right) dt = a\sqrt{b}(\sqrt{\pi}/2), \tag{2.56}$$

co kończy dowód nierówności (2.52). Przejdźmy do dowodu nierówności (2.53). Zauważmy, że medianą $(-X)$ jest $(-m)$ oraz $\mathbf{P}((-X) - (-m) \geq t) = \mathbf{P}(X - m \leq (-t))$, więc z nierówności (2.52), którą już udowodniliśmy wynika, że

$$m - \mu = \mathbf{E}\left[(-X) - (-m)\right] \leq a\sqrt{b}(\sqrt{\pi}/2). \tag{2.57}$$

Stąd wynika nierówność (2.53). Przeprowadźmy teraz dowód nierówności (2.54). Korzystając ponownie z równości (2.55) dostajemy

$$\begin{aligned}\int_0^\infty \mathbf{P}(X - m > t)dt &\leq \int_0^\infty a \exp\left(-\frac{t^2}{b(m+t)}\right) dt \\ &\leq a \int_0^m \exp\left(-\frac{t^2}{2bm}\right) dt + a \int_m^\infty \exp\left(-\frac{t}{2b}\right) dt \\ &\leq a\sqrt{(\pi/2)bm} + 2ab \exp\left(-\frac{m}{2b}\right). \qquad (2.58)\end{aligned}$$

□

Rozważmy funkcję f zdefiniowaną na $\prod_{k=1}^{n}\Omega_k$ oraz niech $\mathbf{c}=(c_1,...,c_n)$. Wówczas, podobnie jak w założeniach nierówności różnic skończonych (twierdzenie 2.18), zakładamy, że zachodzi nierówność $|f(\mathbf{x})-f(\mathbf{x}')|\leq c_k$ o ile wektory $\mathbf{x}$ i $\mathbf{x}'$ różnią się tylko k-tą współrzędną co jest równoznaczne ze stwierdzeniem, że $|f(\mathbf{x})-f(\mathbf{x}')|\leq d_{\mathbf{c}}(\mathbf{x},\mathbf{x}')$. Niech

$$A_a=\{\mathbf{y}\in\prod_{k=1}^{n}\Omega_k : f(\mathbf{y})\leq a\}. \tag{2.59}$$

Rozważmy $\mathbf{x}\in\prod_{k=1}^{n}\Omega_k$. W takim razie dla każdego $\mathbf{y}\in A_a$ zachodzi

$$f(\mathbf{x})\leq f(\mathbf{y})+d_{\mathbf{c}}(\mathbf{x},\mathbf{y})\leq a+d_{\mathbf{c}}(\mathbf{x},\mathbf{y}). \tag{2.60}$$

Pierwsza nierówność w (2.60) zachodzi, ponieważ jeśli $\mathbf{x}$ różni się od $\mathbf{y}$ na m współrzędnych $k_1,...,k_m$, to biorąc wektory z przestrzeni produktowej $a_1,...,a_{m-1}$, które różnią się między sobą jedną współrzędną w taki sposób, że $\mathbf{x}$ oraz a_1 różnią się jedną współrzędną, zaś $\mathbf{y}$ oraz a_{m-1} też tylko jedną dostaniemy zachodzenie zależności rekurencyjnych

$$\begin{aligned}
&f(\mathbf{x})-f(a_1)\leq c_{k_1}\\
&f(\mathbf{x})-f(a_2)=f(\mathbf{x})-f(a_1)+f(a_1)-f(a_2)\leq c_{k_1}+c_{k_2}\\
&\qquad\qquad\cdots
\end{aligned}$$

$$f(\mathbf{x})-f(a_{m-1})=f(\mathbf{x})-f(a_1)+f(a_1)-f(a_2)+\cdots-f(a_{m-1})\leq\sum_{i=1}^{m}c_{k_i}. \tag{2.61}$$

Stąd właśnie zachodzi $f(\mathbf{x})\leq f(\mathbf{y})+d_{\mathbf{c}}(\mathbf{x},\mathbf{y})$. W takim razie minimalizując $f(\mathbf{y})+d_{\mathbf{c}}(\mathbf{x},\mathbf{y})$ po $\mathbf{y}\in A_a$ dostajemy, że

$$f(\mathbf{x})\leq a+d_{\mathbf{c}}(\mathbf{x},A_a). \tag{2.62}$$

Niech $c'=||\mathbf{c}||_2$ oraz niech α będzie wektorem jednostkowym postaci $\frac{\mathbf{c}}{c'}$. Wówczas jeśli $f(\mathbf{x})\geq a+t$, to

$$d_\alpha(\mathbf{x},A_a)=d_{\mathbf{c}}(\mathbf{x},A_a)/c'\geq\frac{f(\mathbf{x})-a}{c'}\geq\frac{t}{c'}. \tag{2.63}$$

Stąd z nierówności (2.63) oraz (2.51), dla każdego $t\geq 0$ dostajemy

$$\mathbf{P}(f(\mathbf{X})\leq a)\mathbf{P}(f(\mathbf{X})\geq a+t)\leq\mathbf{P}(\mathbf{X}\in A_a)\mathbf{P}(d_\alpha(\mathbf{X},A_a)\geq\frac{t}{c'})\leq\exp\left(-\frac{t^2}{2c'^2}\right), \tag{2.64}$$

gdzie pierwsza nierówność wynika z (2.63), a druga z (2.51). Niech m będzie medianą $f(\mathbf{X})$, czyli zgodnie z definicją 2.5 $\mathbf{P}(f(\mathbf{X})\leq m)\geq\frac{1}{2}$ oraz $\mathbf{P}(f(\mathbf{X})\geq m)\geq\frac{1}{2}$. Podstawiając w miejsce a medianę m w nierówności (2.64) dostajemy, że

$$\mathbf{P}(f(\mathbf{X})\geq m+t)\leq\frac{1}{\mathbf{P}(f(\mathbf{X})\leq m)}\exp\left(-\frac{t^2}{2c^2}\right)\leq 2\exp\left(-\frac{t^2}{2c^2}\right). \tag{2.65}$$

Zaś podstawiając w miejsce a różnicę $m-t$ dostajemy

$$\mathbf{P}(f(\mathbf{X})\leq m-t)\leq\frac{1}{\mathbf{P}(f(\mathbf{X})\geq m)}\exp\left(-\frac{t^2}{2c^2}\right)\leq 2\exp\left(-\frac{t^2}{2c^2}\right). \tag{2.66}$$

Korzystając z lematu 2.24 dla $a=2$ i $b=2c^2$ oraz założeń (2.65), (2.66) wynika, że

$$|\mathbf{E}\left[f(\mathbf{X})\right]-m|\leq\sqrt{2\pi}c, \tag{2.67}$$

co pokazuje, że w praktyce nie ma wielkiej różnicy czy będziemy rozpatrywali medianę czy średnią.

2.25. Wnioski z nierówności Azumy

Jednym z podstawowych wniosków jest nierówność Hoeffdinga.

2.25.1. Nierówność Hoeffdinga

Twierdzenie 2.26. *Niech rzeczywiste zmienne losowe $X_1, X_2, ..., X_n$ będą niezależne oraz ograniczone, tak że dla każdego $k \in 1, ..., n$ istnieją takie stałe a_k, b_k, że $a_k \leq X_k \leq b_k$. Oznaczmy przez $S_n = \sum_{k=1}^n X_k$ oraz niech $\mu = \mathbf{E}[S_n]$. Wówczas dla dowolnego $t \geq 0$,*

$$\mathbf{P}(|S_n - \mu| \geq t) \leq 2\exp\left(-\frac{2t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right) \tag{2.68}$$

Dowód. Rozważmy funkcję f daną wzorem $f(x_1, ..., x_n) = x_1 + ... + x_n$. Widzimy zatem, że $f(X_1, ..., X_n) = S_n$. Spełnia ona oczywiście założenia (2.30) dla $c_k = b_k - a_k$, więc możemy skorzystać z nierówności różnic skończonych (twierdzenie 2.18) i dostajemy

$$\begin{aligned} \mathbf{P}\left(|S_n - \mu| \geq t\right) &= \mathbf{P}(|M_n - M_0| \geq t) \\ &\leq 2\exp\left(-\frac{2t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right). \end{aligned} \tag{2.69}$$

□

Kolejnym wnioskiem jest nierówność Chinczyna.

2.26.1. Nierówność Chinczyna

Twierdzenie 2.27. *Niech $\{\epsilon_n\}_{k=1}^n$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie, gdzie $\mathbf{P}(\epsilon_k = \pm 1) = \frac{1}{2}$ dla $k \in \{1, 2, ..., n\}$. Niech $2 < p < \infty$ oraz niech $x_1, x_2, ..., x_n \in \mathbf{R}$. Wówczas*

$$\left(\mathbf{E}\left[\left|\sum_{k=1}^n \epsilon_k x_k\right|^p\right]\right)^{\frac{1}{p}} \leq C_p \left(\sum_{k=1}^n |x_k|^2\right)^{\frac{1}{2}}, \tag{2.70}$$

gdzie C_p jest pewną stałą zależną jedynie od p.

Dowód. Zauważmy, że skoro $\mathbf{E}\left[\sum_{k=1}^n \epsilon_k x_k\right] = 0$, więc korzystając z nierówności Hoeffdinga dostajemy

$$\mathbf{P}\left(\left|\sum_{k=1}^n \epsilon_k x_k\right| \geq t\right) \leq 2\exp\left(-\frac{2t^2}{\sum_{k=1}^n x_k^2}\right). \tag{2.71}$$

Stąd oraz ze wzoru na p-ty moment zmiennej losowej X

$$\mathbf{E}\left[X^p\right] = \int_0^\infty pt^{p-1}\mathbf{P}(X > t)dt \tag{2.72}$$

dostajemy

$$\begin{aligned} \mathbf{E}\left[\left|\sum_{k=1}^n \epsilon_k x_k\right|^p\right] &= \int_0^\infty pt^{p-1}\mathbf{P}\left(\left|\sum_{k=1}^n \epsilon_k x_k\right| \geq t\right) dt \\ &\leq 2\int_0^\infty pt^{p-1}\exp\left(-\frac{2t^2}{\sum_{k=1}^n x_k^2}\right) dt. \end{aligned} \tag{2.73}$$

Niech $C = \sum_{k=1}^{n}(x_k^2)$. Stosując całkowanie przez podstawienie dla $x = \frac{2t^2}{C}, dx = \frac{4t}{C}dt$ oraz warunków $C > 0$ i $t = \sqrt{\frac{Cx}{2}}$, otrzymamy

$$2\int_0^\infty pt^{p-1}\exp\left(-x\right)\frac{C}{4t}dx = \frac{pC}{2}\int_0^\infty \left(\frac{Cx}{2}\right)^{\frac{p-2}{2}}\exp(-x)dx \tag{2.74}$$

$$= C^{(p/2)}2^{(-p/2)}\int_0^\infty x^{\frac{p-2}{2}}\exp(-x)dx \tag{2.75}$$

Widzimy, że dla $p > 2$, zachodzi

$$\left(2^{(-p/2)}\int_0^\infty x^{\frac{p-2}{2}}\exp(-x)dx\right)^{\frac{1}{p}} = 2^{(-\frac{1}{2})}\Gamma\left(\frac{p}{2}\right)^{\frac{1}{p}}, \tag{2.76}$$

gdzie $2^{(-\frac{1}{2})}\Gamma\left(\frac{p}{2}\right)^{\frac{1}{p}}$ jest pewną stałą zależną od p i oznaczmy ją jako B_p, wówczas dostajemy

$$\left(\mathbf{E}\left[\left|\sum_{k=1}^{n}\epsilon_k x_k\right|^p\right]\right)^{\frac{1}{p}} \leq B_p C^{1/2} = B_p\left(\sum_{k=1}^{n}(x_k^2)\right)^{\frac{1}{2}} \tag{2.77}$$

□

Nierówność Azumy, a konkretnie jej wniosek, czyli nierówność różnic skończonych ma zastosowania również w przykładach kombinatorycznych. Omówimy teraz jeden z nich.

2.27.1. Pakowanie do koszy [7]

Dany jest n-wymiarowy wektor $\mathbf{x} = (x_1, ..., x_n)$, gdzie składowe wektora $\mathbf{x}$ są objętościami przedmiotów spełniającymi nierówności $0 \leq x_k \leq 1$ dla każdego $k \in \{1, ..., n\}$. Niech $B(\mathbf{x})$ będzie zdefiniowana jako najmniejsza liczba koszy o objętości jednostkowej (równej 1) potrzebnych do zmieszczenia przedmiotów o objętościach $x_1, ..., x_n$. Niech $\mathbf{X} = (X_1, ..., X_n)$ będzie rodziną niezależnych zmiennych losowych ograniczonych z przedziału $[0, 1]$. Widzimy więc, że jeśli zamienimy tylko jedną z wielkości przedmiotów pakowanych do koszy, to wartość $B(\mathbf{x})$ zmieni się maksymalnie o 1, więc jeśli $\mathbf{X}'$ będzie wektorem różniącym się od $\mathbf{X}$ jedynie na k-tej współrzędnej, to $|B(\mathbf{X}) - B(\mathbf{X}')| \leq c_k = 1$, więc jeśli $\mu = \mathbf{E}[B(\mathbf{X})]$, to z nierówności różnic skończonych (twierdzenie 2.18) dostajemy, że dla $t \geq 0$

$$\mathbf{P}(|B(\mathbf{X}) - \mu| \geq t) \leq 2e^{-2t^2/n}. \tag{2.78}$$

Widzimy, że jeśli $n \to \infty$ i $\omega(n)$ będzie dowolnym ciągiem rozbieżnym, czyli $\omega(n) \to \infty$, wówczas prawdopodobieństwo, że $B(\mathbf{X})$ odbiega od swojej średniej o więcej niż $t = \omega(n)\sqrt{n}$ zbiega do zera. Wynika to z podstawienia $t = \omega(n)\sqrt{n}$ do wzoru (2.78). Otrzymamy wtedy

$$\mathbf{P}(|B(\mathbf{X}) - \mu| \geq \omega(n)\sqrt{n}) \leq 2e^{-2\omega_n^2}. \tag{2.79}$$

Rozdział 3

Nierówność *convex distance* Talagranda

Oznaczmy przez $\Omega_1, ..., \Omega_n$ przestrzenie probabilistyczne oraz niech Ω będzie oznaczała przestrzeń produktową $\Omega = \prod_{k=1}^{n} \Omega_k$. Niech $\mathbf{X} = (X_1, ..., X_n)$ będzie rodziną niezależnych zmiennych losowych, gdzie X_k posiada wartości z przestrzeni Ω_k.

Definicja 3.1. *Odległością convex distance* $f_c(\mathbf{x}, A)$ nazywamy

$$f_c(\mathbf{x}, A) = \sup_{\|\alpha\|_2 = 1} (d_\alpha(\mathbf{x}, A)), \tag{3.1}$$

gdzie $\mathbf{x} \in X$, $A \subseteq X$ oraz supremum wybieramy po n-wymiarowych wektorach o normie euklidesowej równej 1 (to znaczy $||\alpha||_2 = 1$).

Wówczas rozważając n-wymiarowy wektor α o każdej współrzędnej równej $\frac{1}{\sqrt{n}}$ widzimy, że $f_c(\mathbf{x}, A) \geq d_\alpha(\mathbf{x}, A) = \frac{1}{\sqrt{n}} d_H(\mathbf{x}, A)$. Stąd wynika, że odległość convex distance jest ograniczeniem górnym odległości Hamminga (definicja 2.20).

W dalszych rozważaniach będziemy korzystać z dwóch oznaczeń. Zdefiniujemy pewien zbiór wektorów $\mathbf{s}$ zero-jedynkowych o n współrzędnych takich, że dla każdego $\mathbf{s}$ istnieje taki $\mathbf{y}$ należący do A, że jeśli i-ta współrzędna wektora $\mathbf{s}$ jest równa zero, to $x_i = y_i$

$$U_A(\mathbf{x}) = \{(s_i)_{i \leq n} \in \{0, 1\}^n : \exists \mathbf{y} \in A,\ \forall_i\ s_i = 0 \Rightarrow x_i = y_i\} \subseteq \mathbf{R}^n, \tag{3.2}$$

Drugim zdefiniowanym zbiorem jest $V_A(\mathbf{x})$ będący otoczką wypukłą zbioru $U_A(\mathbf{x})$.

Lemat 3.2. [7]

$$f_c(\mathbf{x}, A) = \min \{\|\mathbf{v}\| : \mathbf{v} \in V_A(\mathbf{x})\}, \tag{3.3}$$

gdzie $\|\cdot\|$ *oznacza normę euklidesową.*

Dowód. [7] Jeśli $\mathbf{x} \in A$, to obie strony równości są równe 0, więc możemy zakładać, że $\mathbf{x} \notin A$ i wtedy obie strony równości są dodatnie. Oznaczmy prawą stronę równości (3.3) jako ρ. Niech $\alpha = (\alpha_1, ..., \alpha_n)$ będzie wektorem jednostkowym o nieujemnych współrzędnych. Iloczyn skalarny wektorów α oraz $\mathbf{u} = (u_1, ..., u_n)$ zapiszemy w postaci $\langle \alpha, \mathbf{u} \rangle = \sum_{k=1}^{n} \alpha_k u_k$. Stąd

$$d_\alpha(\mathbf{x}, A) = \min_{\mathbf{y} \in A} d_\alpha(\mathbf{y}, A) = \min_{\mathbf{u} \in U_A(\mathbf{x})} \langle \alpha, u \rangle = \min_{\mathbf{v} \in V_A(\mathbf{x})} \langle \alpha, v \rangle, \tag{3.4}$$

ostatnia równość zachodzi, ponieważ minimum funkcji liniowej określonej na otoczce wypukłej $V_A(\mathbf{x})$ skończonego zbioru $U_A(\mathbf{x})$ musi być osiągalne w punkcie należącym do $U_A(\mathbf{x})$. Z nierówności Cauchy'ego-Schwarza dostajemy

$$\langle \alpha, \mathbf{v} \rangle \leq ||\alpha|| \, ||\mathbf{v}|| = ||\mathbf{v}||. \tag{3.5}$$

W takim razie z ostatniej równości (3.4) oraz nierówności (3.5) wynika, że

$$d_\alpha(\mathbf{x}, A) \leq \rho. \tag{3.6}$$

Z uwagi na to, że (3.6) zachodzi dla każdego α, to

$$f_c(\mathbf{x}, A) \leq \rho. \tag{3.7}$$

Wykażemy teraz nierówność przeciwną do nierówności (3.7). Zauważmy, że minimum w zbiorze występującym po prawej równości (3.3), tzn. $\min\{\|\mathbf{v}\| : \mathbf{v} \in V_A(\mathbf{x})\}$ jest osiągane w punkcie $\hat{\mathbf{v}} \in V$ z normą równą $\|\hat{\mathbf{v}}\| = \rho$, ponieważ $V_A(\mathbf{x})$ jest zwarty (domknięty i ograniczony w $\mathbf{R}^n$). Stąd $\alpha = \frac{\hat{\mathbf{v}}}{\rho}$ jest wektorem jednostkowym. Rozważmy dowolny punkt $\mathbf{v} \in V_A(\mathbf{x})$. Z uwagi na to, że $V_A(\mathbf{x})$ jest wypukła wynika, że $\hat{\mathbf{v}} + \lambda(\mathbf{v} - \hat{\mathbf{v}}) \in V_A(\mathbf{x})$ dla każdej $0 \leq \lambda \leq 1$, więc

$$\langle (\hat{\mathbf{v}} + \lambda(\mathbf{v} - \hat{\mathbf{v}})), (\hat{\mathbf{v}} + \lambda(\mathbf{v} - \hat{\mathbf{v}})) \rangle \geq \langle \hat{\mathbf{v}}, \hat{\mathbf{v}} \rangle . \tag{3.8}$$

Skąd dostajemy, że

$$2\lambda \langle \hat{\mathbf{v}}, (\mathbf{v} - \hat{\mathbf{v}}) \rangle + \lambda^2 \langle (\mathbf{v} - \hat{\mathbf{v}}), (\mathbf{v} - \hat{\mathbf{v}}) \rangle \geq 0. \tag{3.9}$$

Widzimy, że biorąc małą λ dostajemy, że

$$\langle \hat{\mathbf{v}}, (\mathbf{v} - \hat{\mathbf{v}}) \rangle \geq 0. \tag{3.10}$$

Podstawiając w nierówności (3.10) $\hat{\mathbf{v}} = \rho\alpha$ w miejsce pierwszego wektora w iloczynie skalarnym otrzymujemy

$$\rho \langle \alpha, \mathbf{v} \rangle \geq \rho \langle \alpha, \hat{\mathbf{v}} \rangle . \tag{3.11}$$

Zatem

$$\langle \alpha, \mathbf{v} \rangle \geq \langle \alpha, \hat{\mathbf{v}} \rangle = \frac{\langle \hat{\mathbf{v}}, \hat{\mathbf{v}} \rangle}{\rho} = \frac{\|\hat{\mathbf{v}}\|^2}{\rho} = \rho \tag{3.12}$$

dla wszystkich $\mathbf{v} \in V_A(\mathbf{x})$. Z definicji $f_c(\mathbf{x}, A)$, z równości (3.4) oraz z nierówności (3.12) wynika

$$f_c(\mathbf{x}, A) \geq d_\alpha(\mathbf{x}, A) = \min_{\mathbf{v} \in V_A(\mathbf{x})} \langle \alpha, \mathbf{v} \rangle = \rho. \tag{3.13}$$

Z nierówności (3.6) oraz (3.13)

$$\rho \leq f_c(\mathbf{x}, A) \leq \rho, \tag{3.14}$$

czyli zachodzi równość (3.3), tzn. teza lematu. □

Zdefiniujmy też zbiór, który przyda nam się w dalszej części pracy.

$$A_t^c = \{\mathbf{x} \in \Omega : f_c(A, \mathbf{x}) \leq t\}. \tag{3.15}$$

3.2.1. Nierówność convex distance

Przejdźmy teraz do nierówności Talagranda. Zaczniemy od dowodu następującego lematu.

Lemat 3.3. [13] *Dla każdego* $0 \leq r \leq 1$ *zachodzi*

$$\inf_{0\leq\lambda\leq 1} r^{-\lambda} \exp\left(\frac{(1-\lambda)^2}{4}\right) \leq 2 - r. \tag{3.16}$$

Dowód. [13] Zdefiniujmy funkcję $f(r)$ następująco

$$f(r) = \ln(2-r) + \ln r + (\ln r)^2. \tag{3.17}$$

Skoro $0 \leq r \leq 1$, więc $(2-r)^{-2} \leq 1 \leq r^{-1}$, w takim razie zachodzi

$$(2-r)^{-2} - r^{-1} \leq 0. \tag{3.18}$$

Obliczmy następujące pochodne ze względu na r

$$f'(r) = \frac{1}{r-2} + \frac{1}{r} + 2\frac{1}{r}\ln r, \tag{3.19}$$

$$\begin{aligned} (rf'(r))' &= \frac{r-2-r}{(r-2)^2} + 0 + 2\frac{1}{r}, \\ &= -2((r-2)^{-2} - r^{-1}). \end{aligned} \tag{3.20}$$

Z nierówności (3.18) wynika $(rf'(r))' \geq 0$. W takim razie $rf'(r)$ jest rosnąca na $[0,1]$. Z uwagi na to, że $1 \cdot f'(1) = 0$, to funkcja $rf'(r) \leq 0$, a co za tym idzie $f'(r) \leq 0$. Stąd wynika, że $f(r)$ jest nierosnąca na $[0,1]$. Skoro więc $f(1) = 0$, to $f(r) \geq 0$. Stąd

$$\inf_{0\leq r\leq 1} f(r) = \inf_{0\leq r\leq 1} \left[\ln(2-r) + \ln r + (\ln r)^2\right] = 0, \tag{3.21}$$

więc dla $0 \leq r \leq 1$ zachodzi

$$-\ln r - (\ln r)^2 \leq \ln(2-r). \tag{3.22}$$

Stąd

$$(2-r)r^{1+\ln r} = \exp\left(f(r)\right) \geq 1. \tag{3.23}$$

Załóżmy, że λ jest postaci

$$\lambda = \begin{cases} 1 + 2\ln r, \text{ dla } 1 \geq r \geq \exp\left(-\frac{1}{2}\right) \\ 0, \text{ dla } 0 \leq r < \exp\left(-\frac{1}{2}\right). \end{cases} \tag{3.24}$$

Zauważmy, że dla $r = 0$ spełniona jest teza lematu, ponieważ

$$\inf_{0\leq\lambda\leq 1} r^{-\lambda} \exp\left(\frac{(1-\lambda)^2}{4}\right) = \exp\left(\frac{1}{4}\right) \leq 2 - r. \tag{3.25}$$

Rozważmy więc pozostałe przypadki, gdy $r > 0$. Zauważmy, że jeśli $0 < r < \exp\left(-\frac{1}{2}\right)$ oraz $\lambda = 0$, to wówczas

$$\begin{aligned} r^{-\lambda} \exp\left(\frac{(1-\lambda)^2}{4}\right) &= \exp\left(\frac{1}{4}\right) \\ &\leq 2 - \exp\left(-\frac{1}{2}\right) \\ &< 2 - r. \end{aligned} \tag{3.26}$$

Analogicznie jeśli $r \geq \exp\left(\frac{-1}{2}\right)$ oraz $\lambda = 1 + 2\ln r$, to z nierówności (3.22) zachodzi

$$\begin{aligned} \ln\left(r^{-\lambda}\exp\left(\frac{(1-\lambda)^2}{4}\right)\right) &= (-1-2\ln r)\ln r + (\ln r)^2 \\ &= -\ln r - (\ln r)^2 \\ &\leq \ln(2-r), \end{aligned} \tag{3.27}$$

co jest równoważne tezie, ponieważ funkcja $\ln(\cdot)$ jest rosnąca oraz jest zdefiniowana dla $r \geq \exp\left(-\frac{1}{2}\right)$. $\square$

Głównym twierdzeniem tego rozdziału jest następująca nierówność, udowodniona przez M. Talagranda [13].

Twierdzenie 3.4. [13] *Niech* $\mathbf{P}$ *będzie miarą produktową na* $\Omega = \prod_{k=1}^{n}\Omega_k$. *Wówczas dla każdego zbioru* $A \subseteq \Omega$ *zachodzi nierówność*

$$\int_\Omega \exp\left(\frac{1}{4}f_c^2(A,\boldsymbol{x})\right)d\mathbf{P}(\boldsymbol{x}) \leq \frac{1}{\mathbf{P}(A)}, \tag{3.28}$$

w szczególności

$$\mathbf{P}(A)\mathbf{P}(f_c(A,\boldsymbol{x}) \geq t) \leq \exp\left(-\frac{t^2}{4}\right). \tag{3.29}$$

Dowód. [13] Oznaczmy miarę prawdopodobieństwa na Ω_k przez ψ_k, zaś miarę produktową $\psi_1 \times ... \times \psi_n$ przez $\mathbf{P}$. Dowód przeprowadzimy indukcyjnie ze względu na n. Na początku pokażemy, że nierówność (3.28) zachodzi dla $n = 1$. Zauważmy, że wówczas jeśli $\beta \in A$, to $U_A(\beta) = \{0,1\}$, zaś gdy $\gamma \notin A$, to $U_A(\gamma) = \{1\}$. Wówczas $V_A(\beta) = [0,1]$ oraz $V_A(\gamma) = \{1\}$, więc $f_c^2(A,\beta) = 0$ oraz $f_c^2(A,\gamma) = 1$. W takim razie, jeśli oznaczmy przez A' dopełnienia zbioru A w przestrzeni Ω, to możemy zapisać

$$\begin{aligned} \int_{\Omega_1}\exp\left(\frac{1}{4}f_c^2(A,\mathbf{x})\right)d\mathbf{P}(\mathbf{x}) &= \int_A \exp\left(\frac{1}{4}f_c^2(A,\mathbf{x})\right)d\mathbf{P}(\mathbf{x}) + \int_{A'}\exp\left(\frac{1}{4}f_c^2(A,\mathbf{x})\right)d\mathbf{P}(\mathbf{x}) \\ &= \int_A \exp\left(0\cdot\frac{1}{4}\right)d\mathbf{P}(\mathbf{x}) + \int_{A'}\exp\left(\frac{1}{4}\right)d\mathbf{P}(\mathbf{x}) \\ &= \mathbf{P}(A) + \exp\left(\frac{1}{4}\right)(1-\mathbf{P}(A)) \\ &\leq \frac{1}{\mathbf{P}(A)}. \end{aligned} \tag{3.30}$$

Ostatnia nierówność zachodzącej we wzorze (3.30) zachodzi, ponieważ przyjmując, że $\mathbf{P}(A) = a$ otrzymamy nierówność

$$a\left(1-\exp\left(\frac{1}{4}\right)\right) + \exp\left(\frac{1}{4}\right) \leq \frac{1}{a}, \tag{3.31}$$

która jest spełniona dla wszystkich $a \in (0,1]$.

Załóżmy więc teraz, że nierówność (3.28) zachodzi dla wszystkich wymiarów nie większych od n. Niech A będzie podzbiorem $\Omega \times \Omega_{n+1}$, zaś B będzie jego rzutem na Ω. Dla $\omega \in \Omega_{n+1}$ oznaczmy

$$A(\omega) = \{\mathbf{x} \in \Omega; (\mathbf{x},\omega) \in A\}. \tag{3.32}$$

Zauważmy, że jeśli $\mathbf{x} \in \Omega$, $\omega \in \Omega_{n+1}$, $\mathbf{z} = (\mathbf{x}, \omega)$, to zachodzą dwie następujące implikacje

$$\mathbf{s} \in U_{A(\omega)}(\mathbf{x}) \Rightarrow (\mathbf{s}, 0) \in U_A(\mathbf{z}), \tag{3.33}$$

$$\mathbf{t} \in U_B(\mathbf{x}) \Rightarrow (\mathbf{t}, 1) \in U_A(\mathbf{z}). \tag{3.34}$$

Uzasadniając poprawność zachodzenia implikacji (3.33) zauważmy, że $A(\omega)$ jest zbiorem rzutów punktów należących do A, o ostatniej współrzędnej $\omega \in \Omega_n$ na Ω. Następnie jeśli $\mathbf{s} \in U_{A(\omega)}(\mathbf{x})$, to znaczy, że istnieje $\mathbf{y}_0 \in A(\omega)$, że jeśli $s_i = 0$, to $x_i = y_i$. Skoro $\mathbf{z} = (\mathbf{x}, \omega)$, to ten sam $(\mathbf{y}_0, \omega) \in A(\omega) \times \{\omega\} \subseteq A$ zapewnia nam przynależność $(\mathbf{s}, 0)$ do $U_A(\mathbf{z})$. Analogicznie można uzasadnić poprawność implikacji (3.34). Stąd dla $s \in V_{A(\omega)}(\mathbf{x})$, $t \in V_B(\mathbf{x})$, $\lambda \in [0,1]$ zachodzi $(\lambda\mathbf{s} + (1-\lambda)\mathbf{t}, 1-\lambda) \in V_A(\mathbf{z})$. Zauważmy, że odległość zbioru $V_A(\mathbf{z})$ od zera jest nie mniejsza od odległości dowolnego z tego punktu zbioru od 0, czyli możemy zapisać

$$f_c(A, \mathbf{z}) \leq \|(\lambda\mathbf{s} + (1-\lambda)\mathbf{t}, 1-\lambda)\|_2 . \tag{3.35}$$

Stąd oraz z wypukłości funkcji $u \mapsto u^2$ wynika

$$\begin{aligned} f_c^2(A, \mathbf{z}) &\leq \|(\lambda\mathbf{s} + (1-\lambda)\mathbf{t}, 1-\lambda)\|_2^2 \\ &= \left(\sum_{k=1}^{N} (\lambda s_k + (1-\lambda)t_k)^2\right) + (1-\lambda)^2 \\ &\leq \left(\sum_{k=1}^{N} \left(\lambda (s_k)^2 + (1-\lambda)t_k^2\right)\right) + (1-\lambda)^2 \\ &= \lambda \sum_{k=1}^{n} s_k^2 + (1-\lambda)\sum_{k=1}^{n} t_k^2 + (1-\lambda)^2 \end{aligned} \tag{3.36}$$

Zatem biorąc infimum po $\mathbf{s}$ oraz po $\mathbf{t}$ dostajemy nierówność

$$f_c^2(A, \mathbf{z}) \leq (1-\lambda)^2 + \lambda f_c^2(A(\omega), \mathbf{x}) + (1-\lambda) f_c^2(B, x). \tag{3.37}$$

Przekształcając nierówność (3.37) otrzymujemy

$$\begin{aligned} &\int_\Omega \exp\left(\frac{1}{4} f_c^2(A, (\mathbf{x}, \omega))\right) d\mathbf{P}(\mathbf{x}) \\ &\leq \exp\left(\frac{1}{4}(1-\lambda)^2\right)\left(\int_\Omega e^{\left(\frac{1}{4}f_c^2(A(\omega),\mathbf{x})\right)\lambda} \cdot e^{\left(\frac{1}{4}f_c^2(B,\mathbf{x})\right)(1-\lambda)} d\mathbf{P}(\mathbf{x})\right). \end{aligned} \tag{3.38}$$

Stąd korzystając z nierówności (3.38) oraz z nierówności Höldera dla ($\frac{1}{p} = \lambda$, $\frac{1}{q} = 1-\lambda$) dostajemy

$$\begin{aligned} \int_\Omega \exp\left(\frac{1}{4} f_c^2(A, \mathbf{z})\right) d\mathbf{P}(\mathbf{x}) \leq \exp\left(\frac{1}{4}(1-\lambda)^2\right) &\left(\int_\Omega \exp\left(\frac{1}{4} f_c^2(A(\omega), \mathbf{x})\right) d\mathbf{P}(\mathbf{x})\right)^\lambda \\ &\times \left(\int_\Omega \exp\left(\frac{1}{4} f_c^2(B, \mathbf{x})\right) d\mathbf{P}(\mathbf{x})\right)^{1-\lambda}. \end{aligned} \tag{3.39}$$

Z nierówności (3.39) oraz z założeń indukcyjnych wynika

$$\begin{aligned} \int_\Omega \exp\left(\frac{1}{4} f_c^2(A, \mathbf{z})\right) d\mathbf{P}(\mathbf{x}) &\leq \exp\left(\frac{1}{4}(1-\lambda)^2\right)\left(\frac{1}{\mathbf{P}(A(\omega))}\right)^\lambda \left(\frac{1}{\mathbf{P}(B)}\right)^{1-\lambda} \\ &= \frac{1}{\mathbf{P}(B)} \exp\left(\frac{1}{4}(1-\lambda)^2\right)\left(\frac{\mathbf{P}(A(\omega))}{\mathbf{P}(B)}\right)^{-\lambda}. \end{aligned} \tag{3.40}$$

Nierówność (3.40) zachodzi dla dowolnego $\lambda \in [0,1]$, więc możemy zapisać, że

$$\int_{\Omega} \exp\left(\frac{1}{4} f_c^2(A, \mathbf{z})\right) d\mathbf{P}(\mathbf{x}) \leq \frac{1}{\mathbf{P}(B)} \inf_{0 \leq \lambda \leq 1} \left(\exp\left(\frac{1}{4}(1-\lambda)^2\right) \left(\frac{\mathbf{P}(A(\omega))}{\mathbf{P}(B)}\right)^{-\lambda} \right). \tag{3.41}$$

Korzystając z nierówności (3.41) oraz z lematu 3.3 dla $0 \leq r = \frac{\mathbf{P}(A(\omega))}{\mathbf{P}(B)} \leq 1$ dostajemy

$$\int_{\Omega} \exp\left(\frac{1}{4} f_c^2(A, (\mathbf{x}, \omega)\right) d\mathbf{P}(\mathbf{x}) \leq \frac{1}{\mathbf{P}(B)} \left(2 - \frac{\mathbf{P}(A(\omega))}{\mathbf{P}(B)}\right). \tag{3.42}$$

Całkując obie strony nierówności (3.42) ze względu na ω oraz korzystając z twierdzenia Fubiniego i z faktu, że $\delta(2-\delta) \leq 1$ dla wszystkich δ dostajemy

$$\begin{aligned} \int_{\Omega^{n+1}} \exp\left(\frac{1}{4} f_c^2(A, \cdot)\right) d(\mathbf{P} \otimes \psi_{n+1}) &\leq \frac{2 - \frac{(\mathbf{P} \otimes \psi_{n+1})(A)}{\mathbf{P}(B)}}{\mathbf{P}(B)} \\ &\leq \frac{1}{(\mathbf{P} \otimes \psi_{n+1})(A)}, \end{aligned} \tag{3.43}$$

dla $\delta = \frac{(\mathbf{P} \otimes \psi_{n+1})(A)}{\mathbf{P}(B)}$. Z nierówności (3.43) wynika, że nierówność (3.28) jest spełniona dla każdego $A \in \Omega \times \Omega_{n+1}$, co kończy dowód indukcyjny nierówności (3.28). Następnie, udowodnimy, że z nierówności (3.28) wynika nierówność (3.29). Zauważmy, że całkę

$$\int_{\Omega} \exp\left(\frac{1}{4} f_c^2(A, \mathbf{x})\right) d\mathbf{P}(\mathbf{x}) \tag{3.44}$$

można zapisać w postaci sumy

$$\begin{aligned} &\int_{\Omega} \exp\left(\frac{1}{4} f_c^2(A, \mathbf{x})\right) d\mathbf{P}(\mathbf{x}) \\ &= \int_{A_t^c} \exp\left(\frac{1}{4} f_c^2(A, \mathbf{x})\right) d\mathbf{P}(\mathbf{x}) + \int_{(A_t^c)'} \exp\left(\frac{1}{4} f_c^2(A, \mathbf{x})\right) d\mathbf{P}(\mathbf{x}), \end{aligned} \tag{3.45}$$

gdzie $(A_t^c)'$ oznacza dopełnienie $A_t^c = \{\mathbf{x} \in \Omega : f_c(A, \mathbf{x}) \leq t\}$ zawartego w Ω. Zauważmy, że z definicji 3.15 zbioru A_t^c wynika

$$\begin{aligned} \int_{(A_t^c)'} \exp\left(\frac{1}{4} f_c^2(A, x)\right) d\mathbf{P}(x) &\geq \int_{(A_t^c)'} \exp\left(\frac{1}{4} t^2\right) d\mathbf{P}(x) \\ &= \exp\left(\frac{1}{4} t^2\right) \mathbf{P}\left((A_t^c)'\right) \\ &= \exp\left(\frac{1}{4} t^2\right) (1 - \mathbf{P}\left(A_t^c\right)) \end{aligned} \tag{3.46}$$

oraz

$$\int_{A_t^c} \exp\left(\frac{1}{4} f_c^2(A, x)\right) d\mathbf{P}(x) \geq 0. \tag{3.47}$$

Z nierówności (3.28), (3.46), (3.47) wynika, że

$$\begin{aligned} \frac{1}{\mathbf{P}(A)} &\geq \int_{\Omega} \exp\left(\frac{1}{4} f_c^2(A, x)\right) d\mathbf{P}(x) \\ &\geq \exp\left(\frac{1}{4} t^2\right) (1 - \mathbf{P}\left(A_t^c\right)), \end{aligned} \tag{3.48}$$

co jest równoważne nierówności (3.29) i to kończy dowód.

□

Porównując nierówność convex distance (3.29)

$$\mathbf{P}(A)\mathbf{P}(f_c(A,\mathbf{x}) \geq t) \leq \exp\left(-\frac{t^2}{4}\right) \tag{3.49}$$

z twierdzeniem (2.22), czyli nierównością (2.40)

$$\mathbf{P}(\mathbf{X} \in A)\mathbf{P}(d_\alpha(\mathbf{X},A) \geq t) \leq \exp\left(-\frac{t^2}{2}\right) \tag{3.50}$$

widzimy, że prawe strony tych nierówności są postaci $\exp\left(-t^2 c\right)$, gdzie c oznacza pewną stałą niezależną od wymiaru α. Wiemy, że

$$f_c(\mathbf{x},A) \geq d_\alpha(\mathbf{x},A), \tag{3.51}$$

więc

$$\{d_\alpha(\mathbf{x},A) \geq t\} \subseteq \{f_c(\mathbf{x},A) \geq t\}. \tag{3.52}$$

Stąd widzimy, że występujący po lewej stronie nierówności convex distance (3.29) zbiór postaci $\{f_c(A,\mathbf{x}) \geq t\}$ jest większy według miary probabilistycznej niż zbiór występujący po lewej stronie (2.40) z twierdzenia (2.22) postaci $\{d_\alpha(\mathbf{x},A) \geq t\}$. W takim razie nierówność convex distance (3.29) więc jest filozoficznie "mocniejsza", siłą nierówności (3.49) jest fakt, że rozpatrujemy ją po wszystkich wektorach jednostkowych α, a nie po konkretnym, co zobaczymy w kolejnych przykładach.

3.4.1. Przykłady

W tym podrozdziale omówimy kilka przykładów zastosowania nierówności convex distance (twierdzenie 3.4). Pierwszym przykładem jest następujące twierdzenie.

Twierdzenie 3.5. [14] *Niech $f : [-1,1]^n \to \mathbf{R}$ będzie wypukłą funkcją lipschitzowską ze stałą Lipschitza L. Niech $\boldsymbol{X} = (X_1, ..., X_n)$, gdzie X_i są niezależnymi zmiennymi losowymi o wartościach w $[-1,1]$. Wówczas jeśli przez m oznaczymy medianę $f(\boldsymbol{X})$, to dla każdego $t \geq 0$ zachodzi*

$$\mathbf{P}(|f(\boldsymbol{X}) - m| \geq t) \leq 4\exp(-t^2/(16L^2)). \tag{3.53}$$

Dowód. [14] Definiujemy funkcję $h : \Omega \times \Omega \to \mathbf{R}^n$, *taką że dla dowolnych $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ zachodzi*

$$h(\mathbf{x},\mathbf{y})_i = \begin{cases} 1, & dla\ x_i \neq y_i \\ 0, & dla\ x_i = y_i, \end{cases} \tag{3.54}$$

gdzie x_i oznacza i-tą współrzędną wektora $\mathbf{x}$ i analogiczne pozostałe oznaczenia. Oznaczmy przez A_a zbiór punktów z kostki $[-1,1]^n$ zdefiniowany wzorem

$$A_a = \{\mathbf{x} \in [-1,1]^n : f(\mathbf{x}) \leq a\}, \text{ dla dowolnego } a \in \mathbf{R}. \tag{3.55}$$

Rozważmy dowolny punkt $\mathbf{y} \in A_a$. Dla dowolnego $i \in \{1, ..., n\}$ zachodzi

$$|x_i - y_i| \leq 2h(\mathbf{x},\mathbf{y})_i, \tag{3.56}$$

ponieważ lewa strona jest równa co najwyżej 2 oraz równa 0 tylko wtedy, gdy prawa jest różna od 2. Rozważmy dowolny skończony ciąg M-elementowy wektorów z A_a jako $(\mathbf{y}^k)_{1\leq k\leq M}$. Wówczas dla kombinacji wypukłej i-tej współrzędnej wektorów $(\mathbf{y}^k)_{1\leq k\leq M}$ tzn. $\sum_{k=1}^M \alpha_k y_i^k$, gdzie $\sum_{k=1}^M \alpha_k = 1$ oraz $\alpha_k \geq 0$ dla wszystkich $1 \leq k \leq M$ zachodzi następująca nierówność

$$\left| x_i - \sum_{k=1}^M \alpha_k y_i^k \right| \leq 2 \sum_{k=1}^M \alpha_k h(\mathbf{x}, \mathbf{y}^k)_i \ . \tag{3.57}$$

W takim razie skoro f jest funkcją wypukłą, to zbiór A_a jest wypukły, więc

$$d\left(\mathbf{x}, A_a\right) \leq \left\| \mathbf{x} - \sum_{k=1}^M \alpha_k \mathbf{y}^k \right\|_2 \leq 2 \left(\sum_{i=1}^n \left(\sum_k^M \alpha_k h\left(\mathbf{x}, \mathbf{y}^k\right)_i \right)^2 \right)^{1/2}, \tag{3.58}$$

gdzie $||\cdot||_2$ oznacza odległość euklidesową. Z definicji zbioru $U_A(x)$ 3.2 dla A_a, tzn.

$$U_{A_a}(\mathbf{x}) = \{(s_i)_{i\leq n} \in \{0,1\}^n : \exists \mathbf{y} \in A_a, \ \forall_i \ s_i = 0 \Rightarrow x_i = y_i\} \subseteq \mathbf{R}^n, \tag{3.59}$$

wynika, że $h(\mathbf{x}, \mathbf{y}^k) \in U_{A_a}(\mathbf{x})$. Skoro $V_{A_a}(\mathbf{x})$ jest otoczką wypukłą $U_{A_a}(\mathbf{x})$, więc $\sum_{k=1}^M \alpha_k h(\mathbf{x}, \mathbf{y}^k) \in V_{A_a}(\mathbf{x})$. Z dowolności ciągów $(\mathbf{y}^k)_{1\leq k\leq M}$, $(\alpha_k)_{1\leq k\leq M}$, takich że $\sum_{k=1}^M \alpha_k = 1$ oraz $\alpha_k \geq 0$ dla wszystkich $1 \leq k \leq M$ oraz faktu, że $V_{A_a}(\mathbf{x})$ jest zbiorem zwartym istnieją takie $(\mathbf{y}^l)_{1\leq l\leq M}$, $(\alpha_l)_{1\leq l\leq M}$, że $\hat{\mathbf{v}} = \sum_{l=1}^M \alpha_l h(\mathbf{x}, \mathbf{y}^l)$ realizuje minimum odległości $V_{A_a}(\mathbf{x})$ od zera. Stąd z definicji odległości Talagranda dostajemy, że

$$d(\mathbf{x}, A_a) \leq 2 \sum_{l=1}^M \alpha_l h(\mathbf{x}, \mathbf{y}^l) = 2f_c(A_a, \mathbf{x}), \tag{3.60}$$

gdzie $d(\mathbf{x}, A_a)$ oznacza odległość euklidesową wektora $\mathbf{x}$ od zbioru A_a. Ustalmy pewne $\mathbf{x} \in [-1,1]^n$, więc skoro funkcja $f(\mathbf{x})$ jest lipschitzowska ze stałą L, to dla każdego wektora $\mathbf{y} \in A_a$ zachodzi

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq Ld(\mathbf{x}, \mathbf{y}). \tag{3.61}$$

Z nierówności (3.61) wynika

$$f(\mathbf{x}) - f(\mathbf{y}) \leq |f(\mathbf{x}) - f(\mathbf{y})| \leq Ld(\mathbf{x}, \mathbf{y}). \tag{3.62}$$

Z dowolności $\mathbf{y} \in A_a$ oraz z nierówności (3.62) dostajemy, że

$$f(\mathbf{x}) \leq \inf_{\mathbf{y}\in A_a} (f(\mathbf{y}) + Ld(\mathbf{x}, \mathbf{y})) \leq a + Ld(\mathbf{x}, A_a). \tag{3.63}$$

Z nierówności (3.63) oraz (3.60)

$$f(\mathbf{x}) \leq a + Ld(\mathbf{x}, A_a) \leq a + 2Lf_c(A_a, \mathbf{x}). \tag{3.64}$$

Niech $\mathbf{P}$ oznacza rozkład prawdopodobieństwa zmiennej losowej $\mathbf{X} = (X_1, ..., X_n)$ na $\Omega = [-1,1]^n$, więc z nierówności Talagranda (3.29) wynika

$$\mathbf{P}(f(\mathbf{X}) \geq a + t) \leq \frac{1}{\mathbf{P}(A_a)} \exp\left(-\frac{t^2}{16L^2}\right). \tag{3.65}$$

Podstawiając $a = m$ w zależności (3.65) otrzymamy

$$\begin{aligned} \mathbf{P}(f(\mathbf{X}) - m \geq t) &\leq \frac{1}{\mathbf{P}(A_m)} \exp\left(-\frac{t^2}{16L^2}\right) \\ &\leq 2\exp\left(-\frac{t^2}{16L^2}\right), \end{aligned} \tag{3.66}$$

bowiem $\mathbf{P}(A_m) \geq \frac{1}{2}$ z definicji mediany. Natomiast dla $a = (m - t)$ z nierówności (3.65) dostajemy

$$\begin{aligned} \frac{1}{2} &\leq \mathbf{P}(f(\mathbf{X}) \geq m - t + t) \\ &\leq \frac{1}{\mathbf{P}(A_{(m-t)})} \exp\left(-\frac{t^2}{16L^2}\right) \\ &= \frac{1}{\mathbf{P}(f(\mathbf{X}) \leq m - t)} \exp\left(-\frac{t^2}{16L^2}\right), \end{aligned} \tag{3.67}$$

ponieważ $\frac{1}{2} \leq \mathbf{P}(f(\mathbf{X}) \geq m)$. Z nierówności (3.67) wynika

$$\mathbf{P}(f(\mathbf{X}) \leq m - t) \leq 2\exp\left(-\frac{t^2}{16L^2}\right). \tag{3.68}$$

Korzystając z nierówności (3.65) oraz (3.67) dostajemy

$$\mathbf{P}(|f(\mathbf{X}) - m| \geq t) \leq 4\exp(-t^2/(16L^2)), \tag{3.69}$$

czyli tezę. □

Korzystając z metod użytych w dowodzie twierdzenia 3.5 pokażemy na przykładzie dlaczego nierówność convex dinstance (twierdzenie 3.4) daje lepsze szacowanie od nierówności różnic skończonych (twierdzenie 2.18). Rozważmy przykład funkcji $f : \{0,1\}^n \to \mathbf{R}$ danej wzorem $f(\mathbf{x}) = \|\mathbf{x}\|_2$, co oznacza normę euklidesową $\mathbf{x}$, μ oznacza wartość oczekiwaną $f(\mathbf{X})$, a m oznacza jej medianę. Niech $\mathbf{x}$ oraz $\mathbf{x}'$ oznaczają pewne wektory z przestrzeni $\{0,1\}^n$ różniące się jedynie na k-tej współrzędnej. Wówczas zdefiniujmy wektor $\mathbf{s}$ wzorem

$$\mathbf{s} = \mathbf{x} - \mathbf{x}'. \tag{3.70}$$

Z równości (3.70) wynika, że $\mathbf{s}$ zawiera k-tą współrzędną równą 1 lub (-1) oraz 0 na pozostałych miejscach. W takim razie z nierówności trójkąta dostajemy nierówność

$$\|\mathbf{x}\|_2 \leq \left\|\mathbf{x}'\right\|_2 + \|\mathbf{s}\|_2 = \left\|\mathbf{x}'\right\|_2 + 1. \tag{3.71}$$

Nierówność (3.71) gwarantuje zachodzenie nierówności (2.30)

$$\left|f(\mathbf{x}) - f(\mathbf{x}')\right| \leq 1. \tag{3.72}$$

Możemy więc skorzystać z nierówności różnic skończonych (twierdzenie 2.18) i otrzymamy

$$\mathbf{P}(f(\mathbf{X}) \geq \mu + t) \leq \exp\left(-\frac{t^2}{2n}\right). \tag{3.73}$$

Znajdźmy teraz jednak ograniczenie górne lewej strony nierówności (3.73). Korzystając z twierdzenia 3.5 dostajemy

$$\mathbf{P}\left((f(\mathbf{X}) - m) \geq t\right) \leq \mathbf{P}\left(f_c(D, \mathbf{X}) \geq t\right) \leq 2\exp\left(-\frac{t^2}{4}\right). \tag{3.74}$$

W takim razie z lematu 2.24 przy stałych $a = 2$ oraz $b = 4$ dla nierówności (3.74) dostajemy, że

$$|\mu - m| \leq 2\sqrt{\pi}. \tag{3.75}$$

Z nierówności (3.74), (3.75) wynika

$$\begin{aligned} 2\exp\left(-\frac{t^2}{4}\right) &\geq \mathbf{P}\left((f(\mathbf{X}) - m) \geq t\right) \\ &\geq \mathbf{P}\left((f(\mathbf{X}) - \mu) \geq t + m - \mu\right) \\ &\geq \mathbf{P}\left((f(\mathbf{X}) - \mu) \geq t + |m - \mu|\right) \\ &\geq \mathbf{P}\left((f(\mathbf{X}) - \mu) \geq t + 2\sqrt{\pi}\right). \end{aligned} \tag{3.76}$$

Podstawiając w nierównościach (3.76) zamiast t różnicę $(t - 2\sqrt{\pi})$, dla $t \geq 2\sqrt{\pi}$ dostajemy

$$\mathbf{P}\left(f(\mathbf{X}) \geq (t + \mu)\right) \leq 2\exp\left(-\frac{(t - 2\sqrt{\pi})^2}{4}\right). \tag{3.77}$$

Widzimy, że prawa strona nierówności (3.77) (wynikająca z nierówności convex distance (twierdzenie 3.4)) równa wyrażeniu

$$2\exp\left(-\frac{(t - 2\sqrt{\pi})^2}{4}\right) \tag{3.78}$$

jest niezależna od wymiaru dziedziny funkcji f, czyli $\{0, 1\}^n$ w przeciwieństwie do prawej strony nierówności (3.73) (wynikającej z nierówności różnic skończonych (twierdzenie 2.18)) równej wyrażeniu

$$\exp\left(-\frac{t^2}{2n}\right), \tag{3.79}$$

które bezpośrednio zależy od n. W ten sposób widzimy, że jeśli $n \to \infty$, to nierówność (3.73) przyjmuje postać

$$\mathbf{P}(f(\mathbf{x}) \geq (\mu + t)) \leq \lim_{n\to\infty} \exp\left(-\frac{t^2}{2n}\right) = \exp(0) = 1, \tag{3.80}$$

zaś nierówność (3.77) pozostaje bez zmian. Widzimy więc, że dla dużych wymiarów nierówność convex distance (twierdzenie 3.4) daje znacznie lepsze ograniczenia niż nierówność różnic skończonych (twierdzenie 2.18).

Drugim przykładem zastosowania nierówności convex distance (twierdzenie 3.4) jest następujące twierdzenie.

Twierdzenie 3.6. [12] *Rozważmy przestrzeń Banacha B i przestrzeń do niej sprzężoną B^* oraz ciąg $(x_i)_{i\leq n}$ elementów z B. Normy w przestrzeniach B oraz B^* będziemy oznaczali tak samo. Niech*

$$L^2 = \sup\left\{\sum_{i=1}^{n} (x^*(x_i))^2 : x^* \in B^*, ||x^*|| \leq 1\right\}. \tag{3.81}$$

Wówczas dla $(\epsilon_i)_{i\leq n}$ ciągu n-elementowego niezależnych zmiennych losowych Radamachera, zachodzi dla wszystkich $t \geq 0$

$$\mathbf{P}\left(\left|\left\|\sum_{i=1}^{n}\epsilon_i x_i\right\| - m\right| \geq t\right) \leq 4\exp\left(-\frac{t^2}{16L^2}\right), \tag{3.82}$$

gdzie m oznacza medianę funkcji $\|\sum_{i=1}^{n}\epsilon_i x_i\|$.

Dowód. [12] Niech $\mathbf{u} = (u_1, ..., u_n), \mathbf{v} = (v_1, ..., v_n)$ będą n-wymiarowymi wektorami o współrzędnych będącymi niezależnymi zmiennymi Radamachera oraz niech $f : \mathbf{R}^n \to \mathbf{R}$ będzie zdefiniowana wzorem

$$f(\mathbf{u}) = \left\|\sum_{i=1}^{n} u_i x_i\right\|, \tag{3.83}$$

która jest oczywiście wypukła. Pokażemy teraz, że jest ona lipschitzowska. Weźmy dowolny funkcjonał x^* o normie nie większej niż 1. Wówczas z liniowości funkcjonałów wynika

$$\left|x^*\left(\sum_{i=1}^{n} u_i x_i\right) - x^*\left(\sum_{i=1}^{n} v_i x_i\right)\right| \leq \left|\left(\sum_{i=1}^{n}(u_i - v_i)x^*(x_i)\right)\right|. \tag{3.84}$$

Z nierówności (3.84) oraz nierówności Cauchy'ego-Schwarza dostajemy

$$\left|x^*\left(\sum_{i=1}^{n} u_i x_i\right) - x^*\left(\sum_{i=1}^{n} v_i x_i\right)\right| \leq \left(\sum_{i=1}^{n}(x^*(x_i))^2\right)^{(1/2)} \|\mathbf{u} - \mathbf{v}\|_2. \tag{3.85}$$

Stąd z nierówności (3.85) oraz z równości (3.81) wynika

$$\left|x^*\left(\sum_{i=1}^{n} u_i x_i\right) - x^*\left(\sum_{i=1}^{n} v_i x_i\right)\right| \leq L\|\mathbf{u} - \mathbf{v}\|_2. \tag{3.86}$$

Z twierdzenia Hahna-Banacha, dla dowolnego $x \in B$, $\|x\| = \max(x^*(x) \colon x^* \in B^*, \|x^*\| \leq 1)$. Istnieje zatem x_0^* o normie nie większej niż 1, taki że $x_0^*\left(\sum_{i=1}^{n} u_i x_i\right) = \|\sum_{i=1}^{n} u_i x_i\|$, więc

$$\left\|\left(\sum_{i=1}^{n} u_i x_i\right)\right\| - \left\|\left(\sum_{i=1}^{n} v_i x_i\right)\right\| = x_0^*\left(\sum_{i=1}^{n} u_i x_i\right) - x_0^*\left(\sum_{i=1}^{n} v_i x_i\right). \tag{3.87}$$

Zamieniając rolami $\mathbf{u}$ i $\mathbf{v}$ oraz korzystając z nierówności (3.86) dostajemy, że

$$\left|\left\|\left(\sum_{i=1}^{n} u_i x_i\right)\right\| - \left\|\left(\sum_{i=1}^{n} v_i x_i\right)\right\|\right| \leq L\|\mathbf{u} - \mathbf{v}\|_2. \tag{3.88}$$

Zauważmy, że lewa strona nierówności (3.88) jest równa $|f(\mathbf{u}) - f(\mathbf{v})|$, więc możemy zapisać, że

$$|f(\mathbf{u}) - f(\mathbf{v})| \leq L\|\mathbf{u} - \mathbf{u}\|_2. \tag{3.89}$$

Stąd z nierówności (3.89) funkcja f jest lipschitzowska ze stałą L. Korzystając więc z twierdzenia (3.5) dostajemy tezę twierdzenia (3.6). □

Jako wniosek z poprzedniego przykładu rozpatrzmy wzmocnioną wersję twierdzenia Chinczyna. Uzyskujemy je całkując przez części podobnie jak w dowodzie nierówności Chinczyna (twierdzenia 2.27).

Twierdzenie 3.7. [12] *Niech B oznacza pewną przestrzeń Banacha oraz niech $x_1, ..., x_n \in B$ oraz $\epsilon = (\epsilon_i)_{i \leq n}$ jest ciągiem n-elementowym o współrzędnych będących niezależnymi zmiennymi losowymi Radamachera. Wówczas istnieje uniwersalna stała D (niezależna od ϵ, $\boldsymbol{x}$, n oraz B), taka że dla każdego $p \geq 1$ zachodzi*

$$\left\| \sum_{i=1}^{n} \epsilon_i x_i \right\|_p \leq \left\| \sum_{i=1}^{n} \epsilon_i x_i \right\|_1 + DLp^{\frac{1}{2}}, \tag{3.90}$$

gdzie L jest zdefiniowane analogicznie jak w równości (3.81). *Zauważmy, że $L \leq \| \sum_{i=1}^{n} x_i \|_2$, więc jest to wzmocnienie nierówności Chinczyna do przestrzeni nieskończenie wymiarowych.*

Na zakończenie przedstawimy kilka zastosowań kombinatorycznych nierówności Talagranda.

3.7.1. Funkcje konfiguracyjne [7]

Dany jest ciąg liczb rzeczywistych $\mathbf{x} = (x_1, ..., x_n)$. Oznaczmy przez $inc(\mathbf{x})$ długość najdłuższego podciągu $\mathbf{x}$ rosnącego. Można zatem zapisać, że $inc(\mathbf{x})$ jest równe maksimum $|K|$ po wszystkich zbiorach K będących podzbiorami zbioru $\{1, ..., n\}$, takich że jeśli $i, j \in K$ oraz $i < j$, to $x_i < x_j$. Niech $\mathbf{X} = (X_1, ..., X_n)$ będzie rodziną niezależnych zmiennych losowych bezatomowych o wartościach rzeczywistych. W dalszych rozważaniach zajmiemy się koncentracją zmiennej losowej $inc(\mathbf{X})$. Oznaczmy przez μ średnią $inc(\mathbf{X})$. Stąd z nierówności różnic skończonych dla każdego $t \geq 0$ dostajemy

$$\mathbf{P}(|inc(\mathbf{X}) - \mu| \geq t) \leq \exp\left(-\frac{2t^2}{n}\right). \tag{3.91}$$

Pokazuje to, że dla dużego n z dużym prawdopodobieństwem $inc(\mathbf{X})$ jest ograniczony na odcinku długości $O(\sqrt{n})$. Z nierówności Talagranda otrzymamy jednak niedługo znacznie lepszy rezultat, którego szczególnym przypadkiem będzie następujące twierdzenie

Twierdzenie 3.8. *Dla każdego $t \geq 0$ zachodzi*

$$\mathbf{P}(inc(\boldsymbol{X}) \geq m + t) \leq 2\exp\left(-\frac{t^2}{4(m+t)}\right) \tag{3.92}$$

oraz

$$\mathbf{P}(inc(\boldsymbol{X}) \leq m - t) \leq 2\exp\left(-\frac{t^2}{4m}\right). \tag{3.93}$$

A.M. Vershik i S.V. Kerov [15] oraz B.F. Logan i L.A. Shepp [6] wykazali, że m jest rzędu $\sqrt{n}$, dzięki czemu widać dlaczego ta nierówność jest polepszona.

Definicja 3.9. [7] Niech f będzie funkcją zdefiniowaną na zbiorze n-wymiarowym zbiorze Ω. Wówczas dla pewnego $c \in \mathbf{R}$ funkcję f będziemy nazywali *funkcją c-konfiguracyjną* jeśli dla każdego wektora $\mathbf{x} \in \Omega$ będzie istniał jednostkowy wektor α o nieujemnych współrzędnych, taki że dla każdego $\mathbf{y} \in \Omega$ będzie zachodziła nierówność

$$f(\mathbf{y}) \geq f(\mathbf{x}) - d_\alpha(\mathbf{x}, \mathbf{y})\sqrt{cf(\mathbf{x})}. \tag{3.94}$$

Zauważmy, że funkcja $f(\mathbf{x}) = inc(\mathbf{x})$ ma następujące właściwości. Dla każdego $\mathbf{x} \in \Omega$ istnieje podzbiór $K = K(\mathbf{x})$ (zależny od $\mathbf{x}$) zbioru $\{1, ..., n\}$, taki że $f(\mathbf{x}) = |K|$ oraz dla każdego $\mathbf{y} \in \Omega$ zachodzi

$$f(\mathbf{y}) \geq |\{i \in K(\mathbf{x}) : y_i = x_i\}| = f(\mathbf{x}) - |\{i \in K(\mathbf{x}) : y_i \neq x_i\}| . \tag{3.95}$$

Zatem z nierówności (3.95) istnieje n-wymiarowy wektor jednostkowy $\alpha = K(\mathbf{x})/\sqrt{f(\mathbf{x})}$ spełniający nierówność

$$f(\mathbf{y}) \geq f(\mathbf{x}) - d_\alpha(\mathbf{x}, \mathbf{y})\sqrt{cf(\mathbf{x})}. \tag{3.96}$$

W takim razie z definicji 3.9 funkcja $inc(\mathbf{x})$ jest 1-konfiguracyjną.

Twierdzenie 3.10. [7] *Niech f będzie funkcją c-konfiguracyjną oraz oznaczmy przez m jej medianę. Wówczas dla każdego $t \geq 0$ zachodzi*

$$\mathbf{P}(f(\boldsymbol{X}) \geq m + t) \leq 2\exp\left(-\frac{t^2}{4(m+t)}\right) \tag{3.97}$$

oraz

$$\mathbf{P}(f(\boldsymbol{X}) \leq m - t) \leq 2\exp\left(-\frac{t^2}{4m}\right). \tag{3.98}$$

Dowód. [7] Ustalmy pewien wektor $\mathbf{x} \in \Omega$. Skoro f jest c-konfiguracyjna, to istnieje wektor jednostkowy α o nieujemnych współrzędnych, taki, że dla każdego $\mathbf{y} \in \Omega$ zachodzi

$$f(\mathbf{x}) \leq f(\mathbf{y}) + d_\alpha(\mathbf{x}, \mathbf{y})\sqrt{cf(\mathbf{x})}. \tag{3.99}$$

Oznaczmy przez A_a zbiór zdefiniowany wzorem

$$A_a = \{\mathbf{y} \in \Omega : f(\mathbf{y}) \leq a\}. \tag{3.100}$$

Stąd oraz z nierówności (3.99) dostajemy, że dla każdego $\mathbf{y} \in A_a$ zachodzi

$$f(\mathbf{x}) \leq a + d_\alpha(\mathbf{x}, \mathbf{y})\sqrt{cf(\mathbf{x})}. \tag{3.101}$$

Minimalizując więc prawą stronę nierówności (3.101) po $\mathbf{y} \in A_a$ oraz z definicji 3.1 otrzymujemy

$$f(\mathbf{x}) \leq a + d_\alpha(\mathbf{x}, A_a)\sqrt{cf(\mathbf{x})} \leq a + f_c(\mathbf{x}, A)\sqrt{cf(\mathbf{x})}. \tag{3.102}$$

Zatem jeżeli $f(\mathbf{x}) \geq a + t$, to skoro funkcja $g(t) = (t-a)/\sqrt{t}$ jest rosnąca dla $t > a$, to stąd oraz z nierówności (3.102) wynika

$$f_c(\mathbf{x}, A_a) \geq \frac{f(\mathbf{x}) - a}{\sqrt{c}\sqrt{f(\mathbf{x})}} \geq \frac{t}{\sqrt{c}\sqrt{a+t}}. \tag{3.103}$$

W takim razie możemy zapisać, że

$$\mathbf{P}(f(\mathbf{X}) \geq a + t) \leq \mathbf{P}\left(f_c(\mathbf{X}, A_a) \geq \frac{t}{\sqrt{(c(a+t))}}\right). \tag{3.104}$$

Korzystając z nierówności Talagranda (3.29) dostajemy, że

$$\begin{aligned}&\mathbf{P}(f(\mathbf{X}) \leq a)\mathbf{P}(f(\mathbf{X}) \geq a+t) \\ &\leq \mathbf{P}(\mathbf{X} \in A_a)\mathbf{P}\left(f_c(\mathbf{X}, A_a) \geq \frac{t}{\sqrt{c(a+t)}}\right) \\ &\leq \exp\left(-\frac{t^2}{4c(a+t)}\right).\end{aligned} \tag{3.105}$$

Podstawiając pod a medianę m funkcji f w nierówności (3.105) dostajemy pierwszą nierówność z tezy, czyli nierówności (3.10), zaś podstawiając pod a różnicę $a = (m - t)$ dostajemy drugą nierówność z tezy, czyli nierówność (3.98). $\square$

Stąd z twierdzenia 3.10 oraz skoro funkcja $inc(\mathbf{x})$ jest 1-konfiguracyjna, to zachodzi twierdzenie 3.8. Rozważmy teraz zagadnienie wspólnych podciągów dwóch ciągów. Dane są dwa ciągi $\mathbf{x} = (x_1, ..., x_{n_1})$ oraz $\mathbf{y} = (y_1, ..., y_{n_2})$. Oznaczmy przez $com(\mathbf{x}, \mathbf{y})$ funkcję zwracającą maksymalną długość wspólnego podciągu $\mathbf{x}$ oraz $\mathbf{y}$. Niech $\mathbf{X} = (X_1, ..., X_{n_1})$ oraz $\mathbf{Y} = (Y_1, ..., Y_{n_2})$ będą dwoma niezależnymi ciągami niezależnych zmiennych losowych. Będziemy rozpatrywali koncentrację zmiennej losowej $com(\mathbf{X}, \mathbf{Y})$. Oznaczmy przez μ średnią wartość $com(\mathbf{X}, \mathbf{Y})$. Analogicznie jak w przypadku funkcji $inc(\mathbf{x})$ możemy skorzystać z nierówności różnic skończonych (twierdzenie 2.18), z której dostajemy, że dla każdego $t \geq 0$ zachodzi

$$\mathbf{P}(|com(\mathbf{X}, \mathbf{Y}) - \mu| \geq t) \leq 2\exp\left(-\frac{2t^2}{n_1 + n_2}\right). \tag{3.106}$$

Stąd widzimy, że dla dużych $n_1 = n_2 = n$ z dużym prawdopodobieństwem $com(\mathbf{X}, \mathbf{Y})$ jest ograniczone w przedziale długości $O(\sqrt{n})$. Zauważmy jednak, że rozważając funkcję $com(\mathbf{x}, \mathbf{y})$ jako funkcję $(n_1 + n_2)$ zmiennych to jest ona funkcją 2-konfiguracyjną. W takim razie, jeśli m oznaczymy jako medianę funkcji $com(\mathbf{X}, \mathbf{Y})$, to z twierdzenia 3.8 dostajemy

$$\mathbf{P}(com(\mathbf{X}, \mathbf{Y}) \geq m + t) \leq 2\exp\left(-\frac{t^2}{8(m+t)}\right) \tag{3.107}$$

oraz

$$\mathbf{P}(com(\mathbf{X}, \mathbf{Y}) \geq m - t) \leq 2\exp\left(-\frac{t^2}{8m}\right). \tag{3.108}$$

Rozważymy teraz zastosowania nierówności Talagranda (twierdzenie 3.4) do długości drogi w problemie komiwojażera oraz drzewa Steinera w kwadracie jednostkowym [7]. W tym celu wprowadźmy na początku lemat.

Lemat 3.11. [7] *Niech $\boldsymbol{X} = (X_1, ..., X_n)$ będzie rodziną niezależnych zmiennych losowych dla X_i przyjmującej wartości w zbiorze Ω_i oraz niech $\Omega = \prod_{i=1}^{n} \Omega_i$. Niech f będzie funkcją o wartościach rzeczywistych zdefiniowaną na Ω spełniającą warunek, że dla każdego $x \in \Omega$ istnieje nieujemny jednostkowy wektor n-wymiarowy wektor α, taki że*

$$f(\boldsymbol{x}) \leq f(\boldsymbol{y}) + cd_\alpha(\boldsymbol{x}, \boldsymbol{y}) \text{ dla każdego } \boldsymbol{y} \in \Omega. \tag{3.109}$$

Jeśli m jest medianą $f(\boldsymbol{X})$, to

$$\mathbf{P}(|f(\boldsymbol{X}) - m| \geq t) \leq 4\exp\left(-\frac{t^2}{4c^2}\right). \tag{3.110}$$

Taki sam wynik dostaniemy jeśli zamienimy założenie (3.109) *na to, że*

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) + cd_\alpha(\boldsymbol{x}, \boldsymbol{y}) \text{ dla każdego } \boldsymbol{y} \in \Omega. \tag{3.111}$$

Dowód. [7] Wprowadźmy oznaczenie. Niech

$$A_a = \{\mathbf{y} \in \Omega : f(\mathbf{y}) \leq a\}. \tag{3.112}$$

Rozważmy dowolny wektor $\mathbf{x} \in \Omega$. Wówczas istnieje nieujemny wektor jednostkowy n-wymiarowy α oraz stała $c \in \mathbf{R}$, taka że dla każdego $y \in \Omega$ zachodzi

$$f(\mathbf{x}) \leq f(\mathbf{y}) + cd_\alpha(\mathbf{x}, \mathbf{y}) \tag{3.113}$$

i stąd

$$f(\mathbf{x}) \leq a + cd_\alpha(\mathbf{x}, \mathbf{y}) \tag{3.114}$$

dla każdego $\mathbf{y} \in A_a$. Minimalizując prawą stronę nierówności (3.114) po $\mathbf{y}$ zauważamy, że

$$f(\mathbf{x}) \leq a + cd_\alpha(\mathbf{x}, A_a) \leq a + cd_T(\mathbf{x}, A_a). \tag{3.115}$$

Zatem jeśli $f(\mathbf{x}) \geq a = t$, to $f_c(\mathbf{x}, A_a) \geq \frac{t}{c}$. Stąd

$$\begin{aligned}\mathbf{P}(f(\mathbf{X}) \leq a)\mathbf{P}(f(\mathbf{X}) \geq a + t) &\leq \mathbf{P}(\mathbf{X} \in A_a)\mathbf{P}\left(f_c(\mathbf{X}, A_a) \geq \frac{t}{c}\right) \\ &\leq \exp\left(-\frac{t^2}{4c^2}\right),\end{aligned} \tag{3.116}$$

przy czym druga nierówność w (3.116) wynika z nierówności Talagranda (twierdzenie 3.4). Jeśli w nierówności (3.116) podstawimy $a = m$ otrzymamy

$$\mathbf{P}(f(\mathbf{X}) \geq m + t) \leq 2\exp\left(-\frac{t^2}{4c^2}\right), \tag{3.117}$$

podobnie jeśli podstawimy pod $a = m - t$ dostajemy, że

$$\mathbf{P}\left(f(\mathbf{X}) \leq m - t\right) \leq 2\exp\left(-\frac{t^2}{4c^2}\right), \tag{3.118}$$

co kończy dowód gdy zachodzi założenie (3.109).

Rozpatrzmy zachodzenie drugiego warunku w lemacie, tzn. (3.111). Niech $g(\mathbf{x}) = -f(\mathbf{x})$. Wówczas $g(\mathbf{x})$ spełnia pierwsze założenie (3.109) oraz $(-m)$ jest medianą $g(\mathbf{x})$ i podobnie jak wyżej dostajemy, że

$$\mathbf{P}(|f(\mathbf{X}) - m| \geq t) = \mathbf{P}(|g(\mathbf{X}) - (-m)| \geq t) \leq 4\exp\left(-\frac{t^2}{4c^2}\right), \tag{3.119}$$

co kończy dowód lematu. □

Rozważmy rodzinę $\mathbf{X} = (X_1, ..., X_n)$ niezależnych zmiennych losowych, gdzie każda zmienna X_i posiada wartości z kwadratu jednostkowego $[0,1]^2$, tzn. $\Omega = \left([0,1]^2\right)^n$, co rozumiemy jako zbiór n punktów w kwadracie $[0,1]^2$.

3.11.1. Problem komiwojażera [7]

Ścieżką komiwojażera nazywamy najkrótszą ścieżkę zamknięta w grafie przechodzącą przez wszystkie jego wierzchołki. Dany jest punkt $\mathbf{x} \in \Omega$. Oznaczmy przez $tsp(\mathbf{x})$ minimum długości ścieżki komiwojażera przechodzącej przez wszystkie punkty w $\mathbf{x}$. Skorzystamy z faktu bez pokazywania dowodu, który można znaleźć w [10, 11], że istnieje stała $c \geq 0$, taka że spełnione są następujące warunki. Dla każdego n oraz każdego $\mathbf{x} \in \Omega$ istnieje ścieżka $T(\mathbf{x})$ przechodząca przez punkty z $\mathbf{x}$, taka że suma kwadratów długości krawędzi w tej ścieżce jest równa co najwyżej c. Użyjemy tej funkcji do zdefiniowania odpowiedniego wektora α, gdzie k-ta współrzędna α_k odpowiada 'awkwardness' punktu $\mathbf{x}_k$. Dany jest $\mathbf{x} \in \Omega$ oraz niech β_k będzie sumą długości dwóch krawędzi wychodzących z $\mathbf{x}_k$ w ścieżce $T(\mathbf{x})$. Zatem

$$\sum_{k=1}^{n} \beta_k^2 \leq 4c, \tag{3.120}$$

ponieważ $(a+b)^2 \leq 2a^2 + 2b^2$ dla każdego $a, b \in \mathbf{R}$. Wprowadźmy następujący lemat wykorzystujący powyższe zmienne.

Lemat 3.12. [7] *Dla dowolnego* $\boldsymbol{y} \in \Omega$ *zachodzi*

$$tsp(\boldsymbol{x}) \leq tsp(\boldsymbol{y}) + d_\beta(\boldsymbol{x}, \boldsymbol{y}) \leq tsp(\boldsymbol{y}) + (2\sqrt{c}) d_\alpha(\boldsymbol{x}, \boldsymbol{y}), \tag{3.121}$$

gdzie $\alpha = \frac{\beta}{\|\beta\|_2}$ *jest wektorem jednostkowym.*

Dowód. [7] Oznaczmy przez x, y zbiory punktów, których elementy są równe współrzędnym kolejno $\mathbf{x}, \mathbf{y}$, czyli $x = \{\mathbf{x}_1, ...\mathbf{x}_n\}$ oraz $y = \{\mathbf{y}_1, ...\mathbf{y}_n\}$. Rozpatrzmy dwa przypadki.

(i) Jeśli $x \cap y = \emptyset$, to

$$d_\beta(\mathbf{x}, \mathbf{y}) = 2T(\mathbf{x}) \tag{3.122}$$

oraz

$$d_\alpha(\mathbf{x}, \mathbf{y}) = \alpha_1 + ... + \alpha_n = \frac{\beta_1}{\|\beta\|_2} + ... + \frac{\beta_n}{\|\beta\|_2} = \frac{2T(\mathbf{x})}{\|\beta\|_2}. \tag{3.123}$$

Z nierówności (3.120) oraz (3.123) wynika, że

$$2T(\mathbf{x}) \leq \frac{2\sqrt{c}}{\|\beta\|_2} 2T(\mathbf{x}) = d_\alpha(\mathbf{x}, \mathbf{y}) 2\sqrt{c}. \tag{3.124}$$

Stąd z równości (3.122) oraz nierówności (3.124) wynika

$$d_\beta(\mathbf{x}, \mathbf{y}) \leq d_\alpha(\mathbf{x}, \mathbf{y}) 2\sqrt{c}, \tag{3.125}$$

skąd dostajemy tezę.

(ii) Załóżmy, że $x \cap y \neq \emptyset$. Niech F będzie multizbiorem (zbiór, w którym niektóre elementy mogą się powtarzać wiele razy) krawędzi łączących punkty z x następująco. Dla każdego segmentu ścieżki $T(\mathbf{x})$ postaci $a, v_1, ..., v_j, b$, gdzie $a, b \in x \cap y$ oraz $v_1, ..., v_j \notin x \setminus y$ (zauważmy, że $a = b$ jeśli $|x \cap y| = 1$), wstawiamy po dwa razy do zbioru F każdą z krawędzi $v_i v_{i+1}$, dla $i = 1, ..., j-1$ oraz krótszą z krawędzi av_1, $v_j b$, również podwójnie. W ten sposób odpowiednio dla każdego segmentu otrzymujemy cykl zawierający dokładnie jeden punkt z y z sumą krawędzi równą co najwyżej sumie współrzędnych wektora β odpowiadających v_i. Te cykle pomiędzy nimi zakrywają wszystkie punkty z $x \setminus y$, zaś suma długości krawędzi w F jest nie większa niż $d_\beta(\mathbf{x}, \mathbf{y})$. Niech $T^*(\mathbf{y})$ będzie pewną optymalna ścieżką dla $\mathbf{y}$. Rozważmy multigraf G ze zbiorem wierzchołków $x \cup y$ oraz zbiorem krawędzi składającym się z krawędzi z $T^*(\mathbf{y})$ oraz krawędzi z F. Graf G jest spójny oraz każdy stopień wierzchołka jest parzysty, więc G posiada ścieżkę Eulera (przechodzącą przez każdą krawędź dokładnie raz). Ta ścieżka może być skrócona tak, aby dała nam ścieżkę komiwojażera dla x, ponieważ jest ona podzbiorem ścieżki Eulera, co z nierówności trójkąta ma długość nie większą od sumy długości krawędzi w G, która jest równa co najwyżej $tsp(\mathbf{y}) + d_\beta(\mathbf{x}, \mathbf{y})$. Stąd dostajemy zależność (3.121) co kończy dowód (ii), a tym samym lematu 3.12. □

Korzystając z zależności (3.121) oraz lematu 3.11 dostajemy

$$\mathbf{P}(|tsp(\mathbf{X}) - m| \geq t) \leq 4 \exp\left(-\frac{t^2}{16c}\right), \tag{3.126}$$

gdzie m jest medianą $tsp(\mathbf{X})$.

3.12.1. Drzewo Steinera [7]

Drzewo Steinera dla zbioru x punktów z kwadratu jednostkowego jest drzewem, w którym zbiór wierzchołków jest podzbiorem punktów na płaszczyźnie zawierającym x. Dany jest $\mathbf{x} \in \Omega$, oznaczmy przez $st(\mathbf{x})$ minimalną długość drzewa Steinera dla danego zbioru x, którego elementy odpowiadają współrzędnym $\mathbf{x}$. Użyjemy ścieżki $T(\mathbf{x})$ analogicznie jak w problemie komiwojażera, aby zdefiniować odpowiedni wektor β oraz wektor jednostkowy $\alpha = \beta / \|\beta\|_2$. Będziemy stosowali te same oznaczenie jak w lemacie 3.12.

Niech $\mathbf{y} \in \Omega$ oraz niech $S^*(\mathbf{y})$ oznacza optymalne drzewo Steinera dla zbioru y odpowiadającemu $\mathbf{y}$.

W przypadku gdy $x \cap y = \emptyset$ dostaniemy nierówność

$$st(\mathbf{x}) \leq st(\mathbf{y}) + d_\beta(\mathbf{x}, \mathbf{y}) \leq st(\mathbf{y}) + (2\sqrt{c}) d_\alpha(\mathbf{x}, \mathbf{y}), \tag{3.127}$$

którą dowodzi się analogicznie jak przypadek (i) w lemacie 3.12. Będziemy chcieli jednak udowodnić tę nierówność dla dowolnych $\mathbf{x}, \mathbf{y} \in \Omega$. Zajmiemy się więc analizą drugiego przypadku, gdy $x \cap y \neq \emptyset$.

Rozważmy zbiór E krawędzi zawierający krawędzie z $S^*(\mathbf{y})$ oraz te krawędzie z $T(\mathbf{x})$, które mają co najmniej jeden koniec w zbiorze $x \setminus y$. Oznaczmy przez $q(E)$ całkowitą sumę krawędzi w zbiorze E. W takim razie możemy zapisać, że

$$q(E) \leq st(\mathbf{y}) + d_\beta(\mathbf{x}, \mathbf{y}). \tag{3.128}$$

Dodatkowo wiemy z poprzedniego przykładu o problemie komiwojażera, że $\sum_{k=1}^n \beta_k^2 < 4c$. Kluczową obserwacją jest to, że graf G na $x \cup y$ ze zbiorem krawędzi E jest grafem spójnym. Wynika to z faktu, że $T(\mathbf{x})$ łączy wszystkie punkty ze zbioru x, gdzie niektóre z nich są równocześnie ze zbioru y. Zauważmy, że również $S^*(\mathbf{y})$ łączy wszystkie punkty ze zbioru y, a z założenia $x \cap y \neq \emptyset$, więc istotnie graf G jest spójny. Stąd dostajemy, że $st(\mathbf{x})$ jest równe co najwyżej sumie długości krawędzi w E, ponieważ graf G jest spójnym grafem łączącym punkty zbioru na płaszczyźnie zawierającym zbiór x, więc istnieje pewne drzewo Steinera D będące podgrafem grafu G. W takim razie suma długości krawędzi grafu G jest nie mniejsza od sumy krawędzi drzewa D. W takim razie z definicji $st(\mathbf{x})$ suma długości krawędzi grafu D jest większa bądź równa sumie długości krawędzi w optymalnym drzewie Steinera dla zbioru odpowiadającego $\mathbf{x}$. Stąd skoro suma długości krawędzi ze zbioru E jest nie większa niż $st(\mathbf{y}) + d_\beta(\mathbf{x}, \mathbf{y})$ (nierówność (3.128)), to możemy zapisać, że

$$st(\mathbf{x}) \leq q(E) \leq st(\mathbf{y}) + d_\beta(\mathbf{x}, \mathbf{y}) \leq st(\mathbf{y}) + \left(2\sqrt{c}\right) d_\alpha(\mathbf{x}, \mathbf{y}). \tag{3.129}$$

Ostatnią nierówność zależności (3.129) dowodzi się analogicznie jak przypadek (i) w lemacie 3.12. W takim razie korzystając z lematu 3.11 dostajemy, że dla m będącą medianą $st(\mathbf{X})$

$$\mathbf{P}(|st(\mathbf{X} - m)| \geq t) \leq 4 \exp\left(-\frac{t^2}{16c}\right). \tag{3.130}$$

Rozdział 4

Podsumowanie

W pracy zaprezentowano zjawisko koncentracji miary oraz jego powiązania z funkcjami lipchitzowskimi. Dodatkowo omówiono podstawową nierówność martyngałową Azumy oraz wynikające z niej nierówność różnic skończonych oraz jej zastosowanie, nierówność Hoeffdinga i Chinczyna. Niektóre z nich omówiono w terminach koncentracji miary. Zauważono fakt, że dla nierówności omawianych w pracy różnica między medianą, a wartością oczekiwaną rozpatrywanych funkcji jest względnie mała, dzięki czemu można było porównywać w dalszej części koncentrację wartości funkcji wokół obu z tych wartości. Następnie wprowadzono odległość convex distance za pomocą dwóch, różnych równoważnych definicji. Pokazano oraz udowodniono nierówność convex distance Talagranda i wywnioskowano z niej nierówność korzystającą z przestrzeni Banacha oraz wzmocnioną wersję nierówności Chinczyna. Ponadto porównano nierówność Talagranda oraz wnioski z niej płynące z nierównością różnic skończonych na przykładach funkcji konfiguracyjnych i funkcji normy euklidesowej. Udowodniono nierówność szacującą prawdopodobieństwo koncentracji wokół mediany funkcji wypukłych lipschitzowskich. Rozważono dodatkowo znane zagadnienia algorytmiczne dla losowych wierzchołków grafów na kwadracie jednostkowym na płaszczyźnie, tzn. problem komiwojażera i drzewo Steinera, zinterpretowano w języku koncentracji miary funkcje charakteryzujące te zagadnienia i zastosowano do oszacowania ich wartości nierówność Talagranda.

Bibliografia

[1] K. Azuma *Weighted Sums of Certain Dependent Random Variables*, Tôhoku Mathematical Journal, vol. 19, 1967 s, 357-367.

[2] J. Górnicki, *Własności ekstremalne figur izoperymetrycznych na płaszczyźnie*, Matematyka-Społeczeństwo-Nauczanie, vol. 5, 43-48.

[3] J.A. Gubner, *The Gamma Function and Stirling's Formula*, https://gubner.ece.wisc.edu/notes/GammaFunctionStirling.pdf

[4] J. Jakubowski, R. Sztencel. *Wstęp do teorii prawdopodobieństwa*, wyd. III, Warszawa, 2004.

[5] R. Latała, Notatki do wykładu Koncentracja Miary, https://www.mimuw.edu.pl/~ rlatala/testy/mon/wykkoncw09.pdf

[6] B.F. Logan i L.A. Shepp, *A variational problem for random Young tableaux*, Advances in Math., vol.26, 1977 s, 206-222.

[7] C. McDiarmid, *Concentration*, Probabilistic Methods for Algorithmic Discrete Mathematics, Part of the book series: Algorithms and Combinatorics (AC, vol. 16), 1998 s, 195–248.

[8] D. S. Mitrinović, *Elementarne nierówności*, PWN, Warszawa, 1972.

[9] O. Mordhorst, *The optimal constants in Khintchine's inequality for the case 2<p<3*, Colloq. Math., vol. 147(2), 2017 s, 203–216.

[10] L.K. Platzman i J.J. Bartholdi, *Spacefilling curves and the planar traveling salesman problem*, J. Assoc. Comput. Mach., vol.36, 1989 s, 719-737.

[11] J.M. Steele, *Probability Theory and Combinatorial*, SIAM CBMS, vol.69, 1997 s.

[12] M. Talagrand, *An isoperimetric theorem on the cube and the Kintchine-Kahane inequalities*, Proceedings of the American Mathematical Society, vol. 104, n. 3, 1988 s, 905-905

[13] M. Talagrand, *Concentration of measure and isometric inequalities in product spaces*, Publ. Math. Institut des Hautes Etudes Scientifiques, vol. 81, 1995 s, 73-205.

[14] M. Talagrand, *A new look at independence*, The Annals of Probability, vol.24, n. 1, 1996 s, 1-34.

[15] A.M. Vershik i S.V. Kerov, *Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group*, Functional Anal. Appl., vol.19, 1985 s, 21-31.