

Nierówności wyrocennie dla problemów odwrotnych

Grzegorz Mika

3 marca 2018

1 Streszczenie

W pracy rozważany będzie problem estymacji nieznanego elementu f na podstawie niebezpośrednich i zaburzonych obserwacji. Niech Λ będzie skończonym zbiorem estymatorów liniowych. Celem będzie konstrukcja metody wyboru estymatora z rodziny Λ naśladującego estymator o minimalnym ryzyku w tej klasie. Okaże się, że można to osiągnąć poprzez minimalizację nieobciążonego estymatora ryzyka. W pierwszej części pracy zostaną przedstawione wyniki dotyczące operatorów zwartych. W drugiej części wyniki zostaną uogólnione na przypadek operatorów, które niekoniecznie są zwarte. Głównym wynikiem pracy jest zaprezentowanie odpowiednich nieasymptotycznych nierówności wyroczeni w obu przypadkach, które okazują się być asymptotycznie dokładne.

Spis treści

1	Streszczenie	1
2	Wstęp	1
3	Wyroczenie	10
4	Główne rezultaty I	11
5	Główne rezultaty II	25
6	Lematy pomocnicze	36

2 Wstęp

Na początek wprowadzimy potrzebną notację i oznaczenia oraz opiszemy model, w którym będziemy pracować w dalszej części. Niech H oraz G będą dwoma ośrodkowymi przestrzeniami Hilberta z iloczynem

skalarnym oznaczanym odpowiednio $\langle \cdot, \cdot \rangle_H$ oraz $\langle \cdot, \cdot \rangle_G$ (lub gdy nie prowadzi to do nieporozumień krótko $\langle \cdot, \cdot \rangle$), natomiast A niech będzie liniowym i ograniczonym operatorem między tymi przestrzeniami. Naszym celem jest znalezienie takiego $f \in H$, by mając dany $g \in G$, zachodziło

$$Af = g.$$

Problem nazwiemy dobrze postawionym (*well-posed*) wg Hadamarda, gdy:

- dla dowolnego $g \in G$ istnieje $f \in H$ spełniający zadane równanie,
- rozwiązanie jest jedyne,
- rozwiązanie jest stabilne, czyli zależy w sposób ciągły od prawej strony równania.

Jeżeli choć jeden z warunków nie jest spełniony problem nazywamy źle postawionym (*ill-posed*). W przypadku braku stabilności, operator odwrotny A^{-1} jest nieograniczony, co może prowadzić do eksplozji rozwiązania nawet w przypadku niewielkiego zaburzenia wartości g .

W dalszej części obserwacje będą zaburzone przez pewien losowy szum, zatem przypuścimy, że dysponujemy następującym modelem

$$Y = Af + \epsilon \xi, \quad (1)$$

w którym celem jest odzyskanie informacji na temat elementu f na bazie zakłóconych obserwacji Y . Przez ξ rozumiemy odpowiednio zdefiniowany poniżej stochastyczny szum, natomiast przez $\epsilon > 0$ jego poziom.

Definicja 1. *Stochastycznym błędem ξ nazwiemy proces na przestrzeni Hilberta, czyli ograniczony liniowy operator $\xi: G \rightarrow \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ taki, że dla dowolnych elementów $g_1, g_2 \in G$ mamy zdefiniowane zmienne losowe $\langle \xi, g_i \rangle$ takie, że $\mathbb{E}\langle \xi, g_i \rangle = 0$ oraz możemy zdefiniować kowariancję Cov_ξ jako ograniczony liniowy operator ($\|\text{Cov}_\xi\| \leq 1$) z przestrzeni G w przestrzeń G taki, że $\langle \text{Cov}_\xi g_1, g_2 \rangle = \text{Cov}(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle)$ oraz indukowane zmienne $\langle \xi, g \rangle$ są stochastycznie niezależne. Przestrzeń $(\Omega, \mathcal{F}, \mathbb{P})$ jest podstawową przestrzenią probabilistyczną, natomiast $\mathcal{L}^2(\cdot)$ jest przestrzenią wszystkich funkcji całkownych z kwadratem na zadanej przestrzeni z miarą.*

W pracy ograniczymy się do rozważań w modelu białego szumu.

Definicja 2. *Powiemy, że losowy błąd ξ jest białym szumem (white noise), jeśli $\text{Cov}_\xi = I$ oraz indukowane zmienne losowe są gaussowskie, czyli dla dowolnych elementów $g_1, g_2, \dots, g_k \in G$ i dla dowolnego $k \in \mathbb{N}$ mamy, że $\langle \xi, g_i \rangle \sim \mathcal{N}(0, \|g_i\|^2)$, $(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle, \dots, \langle \xi, g_k \rangle) \sim \mathcal{N}_k(\mathbf{0}, \mathbf{\Sigma})$. oraz $\text{Cov}(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle) = \langle g_1, g_2 \rangle$.*

Lemat 1. *Niech ξ będzie białym szumem w przestrzeni G oraz niech $\{u_i\}_{i \in I}$ będzie ortonormalną bazą tej przestrzeni. Oznaczając $\xi_k = \langle \xi, u_k \rangle$ dostajemy, że $\{\xi_i\}_{i \in I}$ są niezależnymi zmiennymi losowymi o tym samym standardowym rozkładzie gaussowskim.*

Dowód. Z definicji $\xi_k = \langle \xi, u_k \rangle \sim \mathcal{N}(0, \|u_k\|^2) = \mathcal{N}(0, 1)$ oraz $Cov(\langle \xi, u_n \rangle, \langle \xi, u_k \rangle) = \langle u_n, u_k \rangle = \delta_{nk}$, gdzie δ_{nk} oznacza symbol Kroneckera, co wraz z założeniem o łącznym rozkładzie gaussowskim kończy dowód. \square

Zauważmy, że gdy ξ jest białym szumem, Y nie jest elementem przestrzeni G a staje się operatorem działającym na przestrzeni G w następujący sposób

$$\forall_{g \in G} \langle Y, g \rangle = \langle Af, g \rangle + \epsilon \langle \xi, g \rangle$$

gdzie $\langle \xi, g \rangle \sim \mathcal{N}(0, \|g\|^2)$.

Wprowadzimy teraz kilka faktów dotyczących operatorów liniowych na przestrzeniach Hilberta.

Rozważmy element $A \in L(H, G)$ przestrzeni operatorów liniowych między dwoma przestrzeniami Hilberta H, G . Założymy, że $D(A) = \{f \in H : \exists_{g \in G} Af = g\} = H$.

Operatorem sprzężonym do operatora A nazywamy operator A^* taki, że $\forall_{f \in H} \forall_{g \in G} \langle Af, g \rangle = \langle f, A^*g \rangle$, natomiast operator, który jest swoim własnym sprzężeniem nazwiemy samosprzężonym.

Operator $A: H \rightarrow H$ jest nieujemny, gdy $\forall_{f \in H} \langle Af, f \rangle \geq 0$ oraz dodatni, gdy $\forall_{f \in H} \langle Af, f \rangle > 0$.

Poniższe twierdzenie pokazuje bardzo użyteczną możliwość rozkładu odpowiednich przestrzeni na pewne składowe wzajemnie ortogonalne.

Twierdzenie 1. *Niech $A \in L(H, G)$. Wtedy*

- $Ker A = (Range A^*)^\perp$ oraz $\overline{Range A} = (Ker A^*)^\perp$,
- jeśli A jest iniektywny, to A^*A też,
- $A^*A \in L(H)$ oraz A^*A jest dodatni i samosprzężony.

Dowód. Zauważmy, że $Range A^\perp = \{g \in G : \langle Ag, g \rangle = 0 \forall f \in H\}$.

Wtedy dla dowolnych $f \in Ker A$ i $g \in G$ mamy, że $0 = \langle Af, g \rangle = \langle f, A^*g \rangle$ a stąd $Ker A = (Range A^*)^\perp$. Zamieniając A z A^* otrzymujemy, że $(Ker A^*)^\perp = Range A^\perp$, czyli $(Ker A^*)^\perp = (Range A^\perp)^\perp = \overline{Range A}$, jako że rozpatrujemy przestrzeń Hilberta.

Korzystając z równości $\langle A^*Af, f \rangle = \langle Af, Af \rangle = \|Af\|^2$, widzimy, że $Ker A = Ker A^*A$.

Analogicznie otrzymujemy, że $\langle A^*Af, f \rangle = \langle Af, Af \rangle = \langle f, A^*Af \rangle$ oraz $\langle A^*Af, f \rangle = \|Af\|^2 \geq 0$, zatem operator A^*A jest samosprzężony i nieujemny. \square

Wniosek 1. • $H = Ker A \oplus Ker A^\perp = Ker A \oplus \overline{Range A^*}$,

- $G = \overline{Range A} \oplus Range A^\perp = \overline{Range A} \oplus Ker A^*$.

W pierwszej części pracy ograniczymy się do rozważania tylko zwartych operatorów liniowych, jednak dzięki temu uda się uzyskać dającą szerokie możliwości reprezentację według wartości osobliwych. Założenie o zwartości badanego

operatora jest naturalnym i często pojawiającym się założeniem w kontekście badania problemów odwrotnych w statystyce z uwagi na częste występowanie operatorów z tej klasy w praktycznych problemach, a także z uwagi na właśnie bardzo wygodną reprezentację.

Definicja 3. Operator $A: H \rightarrow G$ nazywamy *zwartym (compact)*, jeżeli dla każdego ograniczonego zbioru w H , jego obraz przez operator A jest względnie zwarty w G , czyli jego domknięcie jest zwarte w G . Przez $K(H, G)$ będziemy oznaczać zbiór operatorów zwartych między przestrzeniami H i G .

Przykładem operatorów zwartych są operatory całkowe postaci $(Ku)(x) = \int_a^b K(x, y)u(y)dy$, $x \in [a, b]$, $u \in C([a, b])$, gdzie jądro $K(x, y)$ jest takie, że $\int_a^b \int_a^b |K(x, y)|^2 dx dy < +\infty$ lub słabo osobliwe, czyli postaci $\frac{\mathcal{H}(x, y)}{|x-y|^\alpha}$, gdzie $\alpha \in (0, 1)$ a funkcja \mathcal{H} jest funkcją mierzalną i ograniczoną na odcinku $[a, b]$.

Konsekwencją braku ograniczoności kuli jednostkowej w przestrzeniach nieskończenie wymiarowych jest bardzo istotna z punktu widzenia stabilności rozwiązania następująca uwaga

Uwaga 1. Jeżeli $A \in K(H, G)$ oraz $\dim H = \infty$ to operator A^{-1} jest nieograniczony.

Fakt ten powoduje, że dla dowolnego zwartego operatora w przestrzeni nieskończenie wymiarowej, każdy związany z nim problem odwrotny jest źle postawiony.

Poniżej bez dowodu przytaczamy znane twierdzenie dotyczące reprezentacji spektralnej dla operatorów zwartych i samosprężonych potrzebne do wykazania istnienia reprezentacji według wartości osobliwych dla operatorów zwartych, ale już niekoniecznie samosprężonych.

Twierdzenie 2 (Reprezentacja spektralna). Niech A będzie samosprężonym operatorem zwartym na przestrzeni Hilberta H . Wtedy istnieje zupełny układ funkcji własnych $E = \{f_j, j \in I\} \subset H$. Niech $J = \{j \in I: \lambda_j \neq 0\}$ oznacza zbiór tych indeksów dla których odpowiednie wartości własne są niezerowe, wtedy zbiór J jest przeliczalny oraz

$$\forall_{f \in H} Af = \sum_{j \in J} \lambda_j \langle f, f_j \rangle f_j.$$

Ponadto dla każdego $\delta > 0$ zbiór $J_\delta = \{j \in I: |\lambda_j| \geq \delta\}$ jest skończony a jedynym możliwym punktem skupienia zbioru wartości własnych jest zero.

Dowód. Dowód w [1]. □

Możemy teraz wprowadzić reprezentację według wartości singularnych dla operatora zwartego.

Twierdzenie 3 (Reprezentacja według wartości osobliwych). Niech $A: H \rightarrow G$ będzie operatorem zwartym na przestrzeniach Hilberta H, G . Wtedy istnieją skończony lub zbieżny do zera ciąg liczb dodatnich $\{b_n\}_{n \in I}$ oraz układy ortonormalne $\{v_n\}_{n \in I} \subset H$, $\{u_n\}_{n \in I} \subset G$ takie, że

- $\text{Ker} A^\perp = \overline{\text{span}\{v_n, n \in I\}},$
- $\overline{\text{Range} A} = \overline{\text{span}\{u_n, n \in I\}},$
- $Af = \sum_n b_n \langle f, v_n \rangle u_n$ oraz $A^*g = \sum_n b_n \langle g, u_n \rangle v_n.$

Ponadto $g \in \text{Range} A$ wtedy i tylko wtedy, gdy spełniony jest tzw. warunek Picarda

$$\sum_n b_n^{-2} |\langle g, u_n \rangle|^2 < \infty \text{ oraz } g = \sum_n \langle g, u_n \rangle u_n$$

Wtedy rozwiązanie równania $Af = g$ mają postać

$$f = f_0 + \sum_n b_n^{-1} \langle g, u_n \rangle v_n$$

przy czym $f_0 \in \text{Ker} A$ jest dowolne.

Układ (u_n, v_n, b_n) nazywamy układem singularnym operatora A a jego reprezentację w postaci $Af = \sum_n \lambda_n \langle f, v_n \rangle u_n$ nazywamy dekompozycją według wartości osobliwych (*singular value decomposition– SVD*) operatora A .

Dowód. Dowód twierdzenia opiera się na wykorzystaniu twierdzenia spektralnego do operatora A^*A .

Operator A^*A jest samosprężony, zwarty i dodatni, a zatem istnieją liczby $b_1^2 \geq b_2^2 \geq \dots \geq 0$ oraz funkcje ortonormalne v_n takie, że $A^*Av_n = b_n^2 v_n$. Niech $I = \{n: b_n > 0\}$ oraz przez u_n oznaczmy znormalizowane obrazy wektorów v_n , czyli $u_n = b_n^{-1} Av_n$ dla $n \in I$. Zauważmy, że $\langle u_k, u_l \rangle = b_k^{-1} b_l^{-1} \langle Av_k, Av_l \rangle = b_k^{-1} b_l^{-1} \langle v_k, A^*Av_l \rangle = b_k^{-1} b_l^{-1} \langle v_k, b_l^2 v_l \rangle = \delta_{kl}$.

Korzystając w wykazanego wcześniej twierdzenia dostajemy, że $\text{Ker} A^\perp = (\text{Ker} A^*A)^\perp = \overline{\text{Range} A^*A} = \overline{\text{span}\{v_n, n \in I\}}.$

Analogicznie rozpatrując operator AA^* z rozkładem spektralnym $AA^*u_n = b_n^2 u_n$ dostajemy, że $\overline{\text{Range} A} = \overline{\text{span}\{u_n, n \in I\}}.$

Tożsamości $Af = \sum_n b_n \langle f, v_n \rangle u_n$ oraz $A^*g = \sum_n b_n \langle g, u_n \rangle v_n$ otrzymujemy, zauważając, że $Af = \sum_n \langle Af, u_n \rangle u_n = \sum_n \langle Af, b_n^{-1} Av_n \rangle u_n = \sum_n \langle f, b_n^{-1} A^*Av_n \rangle u_n = \sum_n \langle f, b_n^{-1} b_n^2 v_n \rangle u_n = \sum_n b_n \langle f, v_n \rangle u_n$ oraz drugą analogicznie.

Z nierówności Bessela dostajemy, że $\sum_n |\langle f, v_n \rangle|^2 < \infty$, bo $f \in H$ a stąd $\sum_n |\langle f, v_n \rangle|^2 = \sum_n b_n^{-4} |\langle f, b_n^2 v_n \rangle|^2 = \sum_n b_n^{-4} |\langle f, A^*Av_n \rangle|^2 = \sum_n b_n^{-4} |\langle Af, Av_n \rangle|^2 = \sum_n b_n^{-2} |\langle g, b_n^{-1} Av_n \rangle|^2 = \sum_n b_n^{-2} |\langle g, u_n \rangle|^2 < \infty$. W drugą stronę wnioskujemy, że jeśli spełniony jest warunek Picarda to możemy wypisać jawny wzór na rozwiązanie, gdyż odpowiedni szereg norm współczynników jest zbieżny i g jest sumą swojego szeregu Fouriera..

Ostatecznie możemy wnioskować, że $f = f_0 + \sum_n b_n^{-1} \langle g, u_n \rangle v_n$, gdzie $f_0 \in \text{Ker} A$, gdyż na mocy powyższych faktów mamy, że $Af = A(f_0 + \sum_n b_n^{-1} \langle g, u_n \rangle v_n) = \sum_n \langle g, u_n \rangle b_n^{-1} Av_n = g$. \square

Udało nam się zaprezentować działanie zwartego operatora w postaci jego rozwinięcia według wartości osobliwych w postaci $Af = \sum_n b_n \langle f, v_n \rangle u_n$ oraz

uzyskać postać szukanych rozwiązań w postaci $f = f_0 + \sum_n b_n^{-1} \langle g, u_n \rangle v_n$. Jednak takie rozwiązanie sytuacji stawia przed nami nowe problemy. Po pierwsze zauważmy, że jeżeli tylko g posiada niezerowe składowe w przestrzeni ortogonalnej do domknięcia obrazu operatora A równanie $Af = g$ nie może być spełnione dokładnie. Niech $P: G \rightarrow \overline{Range A}$ będzie rzutem ortogonalnym, czyli $\forall_{g \in G} Pg = \sum_n \langle g, u_n \rangle u_n$. Wtedy dla dowolnego elementu $f \in H$ mamy, że $\|Af - g\|^2 = \|Af - Pg\|^2 + \|(1 - P)g\|^2 \geq \|(1 - P)g\|^2$.

Drugi problem związany jest ze zbieżnością szeregu w warunku Picarda. Z twierdzenia o reprezentacji spektralnej operatora zwartego samosprężonego wiemy, że liczby $b_n \rightarrow 0$ gdy $n \rightarrow \infty$, a zatem liczby $b_n^{-2} \rightarrow \infty$ gdy $n \rightarrow \infty$, a nie mamy żadnej gwarancji, że liczby $\langle g, u_n \rangle$ zbiegają do zera odpowiednio szybko by zrównoważyć ten przyrost szczególnie w przypadku zaburzonej wartości y .

W kolejnym kroku ograniczając się do badania operatorów zwartych w modelu białego szumu wprowadzimy równoważną formę wyjściowego zagadnienia

$$Y = Af + \epsilon \xi$$

w postaci modelu przestrzeni ciągów (*sequence space model*).

Rozważmy układ singularny (u_n, v_n, b_n) operatora zwartego A oraz niech ξ będzie białym szumem. Możemy wtedy zapisać, rozpatrując projekcję Y na układ $\{u_n\}$, że

$$\langle Y, u_n \rangle = \langle Af, u_n \rangle + \epsilon \langle \xi, u_n \rangle = \langle Af, b_n^{-1} A v_n \rangle + \epsilon \xi_n = b_n^{-1} \langle A^* A f, v_n \rangle + \epsilon \xi_n =$$

$$b_n^{-1} \left\langle \sum_k b_k^2 \langle f, v_k \rangle v_k, v_n \right\rangle + \epsilon \xi_n = b_n \theta_n + \epsilon \xi_n$$

gdzie $\theta_n = \langle f, v_n \rangle$ są współczynnikami w rozwinięciu Fouriera funkcji f w bazie $\{v_n\}$.

Oznaczając przez $y_n = \langle Y, u_n \rangle$ możemy wyjściowy problem $Y = Af + \epsilon \xi$ zapisać w równoważnej postaci *sequence space model* jako

$$y_n = b_n \theta_n + \epsilon \xi_n, \quad n = 1, 2, \dots$$

W tej postaci widać dokładnie trudności związane ze stochastycznymi problemami odwrotnymi. Jako że b_n są wartościami osobliwymi operatora zwartego mamy, że $b_n \rightarrow 0$ gdy $n \rightarrow \infty$, czyli widać, że wraz ze wzrostem n sygnał $b_n \theta_n$ staje się coraz słabszy i coraz trudniej estymować θ_n . Dodatkową trudnością jest fakt, że naszym celem jest estymacja współczynników θ_n a nie współczynników $b_n \theta_n$, dlatego możemy zapisać równoważną postać problemu

$$x_n = \theta_n + \epsilon \sigma_n \xi_n, \quad n = 1, 2, \dots \quad (2)$$

gdzie $x_n = y_n / b_n$ oraz $\sigma_n = b_n^{-1}$, czyli $\sigma_n \rightarrow \infty$ gdy $n \rightarrow \infty$.

Korzystając z powyższego modelu wprowadzimy notację i pojęcia związane z konstrukcją rozważanych estymatorów i ich własnościami.

Mając pełną i niezaburzoną informację o współczynnikach θ_n , $n = 1, 2, \dots$ można by uzyskać pełną informację o poszukiwanym elemencie f z dokładnością do składowej znajdującej się w dopełnieniu ortogonalnym domknięcia obrazu operatora A kładąc $f = \sum_n \theta_n v_n$. W naturalny sposób można by zatem estymować współczynniki θ_n przez odpowiednie zaobserwowane wartości x_n , gdyż $\mathbb{E}_f(x_n) = \theta_n$. Uzasadnione może jednak być estymowanie współczynników rozwinięcia nie bezpośrednio przez zaobserwowane wartości, a przez przeskalowane w pewien sposób wartości.

Definicja 4. Niech $\lambda = (\lambda_1, \lambda_2, \dots)$ będzie nielosowym ciągiem liczbowym. Estymatorem liniowym współczynników θ_n w modelu (2) nazwiemy estymator $\theta(\hat{\lambda}) = (\hat{\theta}_1, \hat{\theta}_2, \dots)$, gdzie

$$\hat{\theta}_i = \lambda_i x_i, \quad i = 1, 2, \dots$$

Ciąg λ nazywać będziemy filtrem lub wagami.

Przykładowo estymatory rzutowe estymujące poszukiwany element f przez początkowe N składników w rozwinięciu w szereg Fouriera za współczynniki przyjmując zaobserwowane wartości odpowiadają filtrom $\lambda = (\lambda_1, \lambda_2, \dots)$, w którym $\lambda_i = \mathbf{1}_{\{i \leq N\}}$, gdzie $\mathbf{1}_A$ oznacza indyktor zbioru A . Innymi często stosowanymi wagami są wagi Tichonowa–Phillipsa postaci $\lambda_i = \frac{1}{1+(i/w)^a}$, $w > 0$, $a > 0$ lub Pinskera postaci $\lambda_i = \max\{0, 1 - (i/w)^a\}$, $w > 0$, $a > 0$.

Jakość estymatora \hat{f} elementu f mierzona będzie przy pomocy scałkowanego ryzyka średniokwadratowego (*mean integrated squared risk*).

Definicja 5. Scałkowanym ryzykiem średniokwadratowym estymatora \hat{f} elementu f nazywamy wyrażenie

$$\mathcal{R}(\hat{f}, f) = \mathbb{E}_f \|f - \hat{f}\|^2.$$

W przypadku modelu (2) estymator \hat{f} możemy zapisać w postaci $\hat{f} = \sum_n \hat{\theta}_n v_n$. Wtedy dostajemy, że

$$\mathcal{R}(\hat{f}, f) = \mathbb{E}_f \|f - \hat{f}\|^2 = \mathbb{E}_\theta \sum_n \left(\theta_n - \hat{\theta}_n \right)^2 = \mathbb{E}_\theta \|\theta - \hat{\theta}\|^2,$$

gdzie ostatnia równość wynika z tożsamości Parsevala, a norma $\|\cdot\|$ rozumiana jest odpowiednio w przestrzeni \mathcal{L}^2 lub l^2 , natomiast wartość oczekiwana jest liczona odpowiednio względem Y lub $X = (x_1, x_2, \dots)$.

Zatem w przypadku modelu (2) analiza ryzyka $\mathcal{R}(\hat{f}, f)$ jest równoważna analizie ryzyka $\mathcal{R}(\hat{\theta}, \theta) = \mathbb{E}_\theta \|\theta - \hat{\theta}\|^2$. W przypadku estymatorów liniowych wyrażenie na ryzyko estymatora przyjmuje postać

$$\mathbb{E}_\theta \|\theta - \hat{\theta}\|^2 = \mathbb{E}_\theta \sum_{n=1}^{\infty} \left(\theta_n - \hat{\theta}_n(\lambda) \right)^2 = \mathbb{E}_\theta \sum_{n=1}^{\infty} (\theta_n - \lambda_n x_n)^2 =$$

$$= \sum_{n=1}^{\infty} (1 - \lambda_n)^2 \theta_n^2 + \epsilon^2 \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2. \quad (3)$$

Pierwszy składnik odpowiada za obciążenie estymatora, natomiast drugi za jego wariancję.

Definicja 6. *Ryzykiem minimaxowym w klasie funkcji \mathcal{F} nazywamy wyrażenie*

$$r(\mathcal{F}) = \inf_{\hat{f}} \sup_{f \in \mathcal{F}} \mathcal{R}(\hat{f}, f),$$

gdzie $\inf_{\hat{f}}$ wzięte jest po wszystkich możliwych estymatorach elementu f .

W przypadku nieparametrycznego podejścia do estymacji wyznaczenie estymatora realizującego ryzyko minimaxowe jest zwykle niemożliwe, dlatego poszukiwać będziemy estymatora asymptotycznie minimaxowego.

Definicja 7. *Przypuśćmy, że istnieje estymator \tilde{f} , taki, że istnieją stałe $0 < C_1 \leq C_2 < \infty$ takie, że gdy $\epsilon \rightarrow 0$ zachodzi*

$$\sup_{f \in \mathcal{F}} \mathcal{R}(\tilde{f}, f) \leq C_2 a_\epsilon$$

$$\inf_{\hat{f}} \sup_{f \in \mathcal{F}} \mathcal{R}(\hat{f}, f) \geq C_1 a_\epsilon,$$

gdzie ciąg nieujemnych wartości a_ϵ jest taki, że $a_\epsilon \rightarrow 0$, gdy $\epsilon \rightarrow 0$. Mówimy wtedy, że estymator \tilde{f} jest optymalny lub osiąga optymalne tempo zbieżności. W przypadku gdy $C_1 = C_2$, mówimy, że estymator \tilde{f} jest asymptotycznie minimaxowy.

Na koniec wprowadzimy jeszcze pewne pojęcia związane z założeniami o gładkości estymowanego elementu f w zależności od własności wygładzających danego operatora.

Niech zatem $f \in H$ i niech A będzie operatorem zwartym.

Definicja 8. *Powiemy, że dla elementu f istnieją warunki źródłowe (source condition), jeżeli istnieje $w \in H$, $L > 0$ oraz $\mu \geq 0$ takie, że*

$$f = (A^* A)^\mu w \text{ oraz } \|w\|^2 \leq L.$$

Przez $H_{\mu,L}$ oznaczać będziemy klasę funkcji takich, że

$$H_{\mu,L} = \{f \in H : f = (A^* A)^\mu w, w \in H, \|w\|^2 \leq L\}.$$

Mając do dyspozycji reprezentację spektralną operatora $A^* A$ oraz rachunek funkcyjny ([9]) dostajemy, że warunki źródłowe dla funkcji f są równoważne wyrażeniu

$$f = (A^* A)^\mu w = \sum_{k=1}^{\infty} b_k^{2\mu} w_k v_k,$$

gdzie $w_k = \langle w, v_k \rangle$. Oznaczając przez $\theta_k = \langle f, v_k \rangle = b_k^{2\mu} w_k$ dostajemy, że

$$\|w\|^2 \leq L \iff \sum_{k=1}^{\infty} w_k^2 = \sum_{k=1}^{\infty} b_k^{-4\mu} \theta_k^2 \leq L.$$

Zatem założenie o warunkach źródłowych są równoważne założeniu, że współczynniki rozwinięcia Fouriera funkcji f w odpowiedniej bazie należą do pewnej elipsoidy w przestrzeni l^2 . Oznaczmy taką elipsoidę przez

$$\Theta(a, L) = \left\{ \theta \in {}^2: \sum_{k=1}^{\infty} a_k^2 \theta_k^2 \leq L \right\},$$

gdzie $L > 0$ oraz ciąg $a = \{a_k\}$ jest ciągiem nieujemnych liczb rozbieżnym do nieskończoności.

Wprowadzimy teraz pojęcie klasy Sobolewa funkcji.

Definicja 9. Niech H będzie pewną przestrzenią Hilberta a $\{\phi_i\}$ układem ortonormalnym w tej przestrzeni. Klasę Sobolewa nazywamy klasę postaci

$$\mathcal{W}(\alpha, L) = \left\{ f \in H: f = \sum_{i=1}^{\infty} \theta_i \phi_i, \theta \in \Theta^\alpha(a, L) \right\},$$

gdzie ciąg a jest taki, że

$$a_i = \begin{cases} (i-1)^\alpha, & \text{dla } i \text{ nieparzystego,} \\ i^\alpha, & \text{dla } i \text{ parzystego.} \end{cases}$$

W przypadku, gdy α jest liczbą całkowitą, $H = \mathcal{L}^2[0, 1]$ oraz $\{\phi_i\}$ jest układem trygonometrycznym, klasa Sobolewa ma równoważne przedstawienie postaci

$$\mathcal{W}(\alpha, L) = \left\{ f \in H: \int_0^1 \left(f^{(\alpha)}(t) \right)^2 dt \leq L, f^{(j)}(0) = f^{(j)}(1) = 0, j = 0, 1, \dots, \alpha - 1 \right\}.$$

W przypadku problemów z wielomianowym tempem wzrostu współczynników σ_k , czyli takich, że $\sigma_k^{-1} = b_k = k^{-\beta}$ warunki źródłowe są równoważne warunkowi

$$\sum_{k=1}^{\infty} b_k^{-4\mu} \theta_k^2 = \sum_{k=1}^{\infty} k^{4\mu\beta} \theta_k^2 \leq L,$$

czyli założeniu, że funkcja f jest z klasy Sobolewa $\mathcal{W}(2\mu\beta, L)$.

W przypadku problemów z wykładniczym tempem wzrostu współczynników σ , czyli takich, że $\sigma_k^{-1} = b_k = \exp(-\beta)$, warunki źródłowe prowadzą do warunku

$$\sum_{k=1}^{\infty} b_k^{-4\mu} \theta_k^2 = \sum_{k=1}^{\infty} e^{4\mu\beta k} \theta_k^2 \leq L,$$

czyli założenia, że funkcja f należy do klasy funkcji analitycznych postaci

$$\mathcal{A}(\alpha, L) = \left\{ f \in H : f = \sum_{k=1}^{\infty} \theta_k \phi_k, \theta \in \Theta_{\mathcal{A}}^{\alpha}(a, L) \right\},$$

gdzie elipsoida $\Theta_{\mathcal{A}}^{\alpha}(a, L)$ jest taka, że ciąg a zdefiniowany jest jako $a_k = \exp(\alpha k)$ z $\alpha = 2\mu\beta$.

3 Wyrocznie

Przypuśćmy przez chwilę, że zajmujemy się estymacją współczynników θ_i w następującym modelu

$$x_i = \theta_i + \sigma \epsilon_i, \quad i = 1, 2, \dots, n,$$

gdzie σ jest stała dla wszystkich obserwacji, a ϵ_i są niezależnymi zmiennymi losowymi o standardowym rozkładzie normalnym. Będziemy je estymować przy pomocy estymatorów liniowych ze stałymi wagami, czyli postaci $\lambda X = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$. W takim przypadku ryzyko takiego estymatora wynosi

$$\mathcal{R}(\hat{\theta}, \theta) = (1 - \lambda)^2 \|\theta\|_n^2 + n\lambda^2 \sigma^2,$$

gdzie $\|\theta\|_n^2 = \sum_{i=1}^n \theta_i^2$.

Ryzyko to jest minimalizowane przez ustalenie wag jako

$$\tilde{\lambda} = \frac{\|\theta\|_n^2}{n\sigma^2 + \|\theta\|_n^2}.$$

Wtedy estymator $\tilde{\lambda}X$ osiąga minimalne ryzyko w rozważanej klasie estymatorów równe

$$\mathcal{R}(\tilde{\lambda}X, \theta) = \frac{\|\theta\|_n^2}{n\sigma^2 + \|\theta\|_n^2} = \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

gdzie $\Lambda = \{\lambda X : \lambda \in [0, 1]\}$ jest rozważaną klasą estymatorów, w której wystarczy ograniczyć się do rozważania wag z przedziału $[0, 1]$, gdyż wagi spoza tego przedziału prowadzą do estymatorów niedopuszczalnych. Zauważmy jednak, że nie możemy zastosować tak wyznaczonego estymatora, gdyż korzysta on z niedostępnej dla nas informacji o estymowanym elemencie poprzez $\|\theta\|_n^2$. Minimalne ryzyko może zostać osiągnięte jedynie przez wyrocznie, która zna estymowany element. Naszym celem byłaby konstrukcja takiej metody wyznaczania wagi λ , która korzystając jedynie z informacji zawartej w obserwowanej próbie starałaby się naśladować zachowanie wyroczy w kontekście osiąganego ryzyka. Tak postawione zagadnienie rozwiązywane jest przez znalezienie estymatora θ^* , którego ryzyko daje się kontrolować przez nieasymptotyczne nierówności wyrocznie postaci

$$\mathcal{R}(\theta^*, \theta) \leq C_1 \gamma \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + C_2 \gamma^{-1}, \quad (4)$$

gdzie stałe C_1, C_2, γ nie zależą od estymowanego elementu.

Dodatkowo pożądaną własnością takiego estymatora, która uzasadniałaby dodatkowo jego optymalne własności oraz wskazywałaby tempo zbieżności do estymatora asymptotycznie minimaksowego, jest, by powyższa nierówność wyrocznia prowadziła do dokładnych nierówności postaci

$$\mathcal{R}(\theta^*, \theta) \leq (1 + o(1)) \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) \quad (5)$$

gdzie $\sigma \rightarrow 0$.

W przypadku rozważanego na początku rozdziału problemu poszukiwanym estymatorem naśladowującym wyrocznię jest estymator Jamesa–Steina z wagą

$$\lambda^* = 1 - \frac{(n-2)\sigma^2}{\sum_{i=1}^n x_i^2}$$

spełniający nierówność wyrocznię postaci

$$\mathcal{R}(\lambda^* X, \theta) \leq 2\sigma^2 + \mathcal{R}(\tilde{\lambda} X, \theta).$$

W badanym w pracy zagadnieniu zaprezentowane zostaną analogiczne nierówności wyrocznie postaci (4) i (5) dla wszystkich badanych klas operatorów. Mając je do dyspozycji można uzasadnić optymalny wybór parametrów wygładzających w przypadku na przykład estymatorów typu rzutowego, Tichonowa–Phillipsa czy Pinsker.

4 Główne rezultaty I

Przypuśćmy, że dysponujemy skończonym zbiorem filtrów $\Lambda = (\lambda^1, \dots, \lambda^N)$ i związanych z nimi estymatorów liniowych, gdzie $\lambda^i = (\lambda_1^i, \lambda_2^i, \dots)$, $i = 1, 2, \dots, N$. Naszym celem będzie konstrukcja filtra opartego na obserwacjach $\lambda^*(X) = (\lambda_1^*, \lambda_2^*, \dots)$ o wartościach w zbiorze Λ o asymptotycznie minimalnym ryzyku przy prawdziwej wartości θ , który równocześnie naśladowuje ryzyko najlepszego estymatora w tej klasie w każdej skończonej próbie. Okaze się, że filtr ten może zostać zdefiniowany jako element minimalizujący względem $\lambda \in \Lambda$ nieobciążony estymator ryzyka (*unbiased risk estimator*).

Zgodnie z (3) ryzyko estymatora liniowego wyraża się wzorem

$$\mathcal{R}(\hat{\theta}, \theta) = \sum_{n=1}^{\infty} (1 - \lambda_n)^2 \theta_n^2 + \epsilon^2 \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2.$$

By analiza ryzyka jakiegokolwiek estymatora liniowego miała sens, musi założyć, że ryzyko to jest skończone, a zatem należy dobrać filtry tak, by drugi człon był skończony.

Założenie 1.

$$\forall \lambda \in \Lambda \quad 0 < \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2 < \infty$$

Założenie od dodatniości powyższej sumy implikuje, że żaden z rozważanych filtrów nie może być tożsamościowo równy zeru.

Dodatkowo z postaci wyrażenia ma ryzyko widać, że wystarczy ograniczyć się do rozpatrywania wag takich, że $\forall_{\lambda \in \Lambda} \forall_i \lambda_i \in [0, 1]$, gdyż w przeciwnym wypadku uzyskane estymatory stają się niedopuszczalne. Mimo tego narzucimy na rozważane wagi nieco słabsze wymagania.

Założenie 2.

$$\max_{\lambda \in \Lambda} \sup_i |\lambda_i| \leq 1$$

Dopuszczenie ujemnych wartości dla wag pozwala rozpatrywać potencjalnie także wagi związane na przykład z estymatorami jądrowymi przyjmującymi czasem ujemne wartości co może być uzasadnione ich lepszym zachowaniem w badanym problemie (głębsza dyskusja w [10]).

Kolejnym krokiem jest wyznaczenie nieobciążonego estymatora wyrażenia (3). W dalszym ciągu rozważań założymy, że poziom szumu ϵ jest znany, natomiast czynniki σ_n związane są z rozważanym operatorem, którego pełną znajomość także jest zakładana. Pozostaje znalezienie nieobciążonego estymatora dla składników θ_n^2 . Zgodnie z modelem (2) estymatorem takim jest $x_n^2 - \sigma_n^2 \epsilon^2$, gdyż $x_n \sim \mathcal{N}(\theta_n, \sigma_n^2 \epsilon^2)$. Wstawiając to wyrażenie do wzoru opisującego ryzyko estymatora liniowego dostajemy

$$\begin{aligned} \mathcal{E} &= \sum_{n=1}^{\infty} (1 - \lambda_n)^2 (x_n^2 - \sigma_n^2 \epsilon^2) + \epsilon^2 \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2 = \\ &= \sum_{n=1}^{\infty} (x_n^2 - \sigma_n^2 \epsilon^2) + \sum_{n=1}^{\infty} (\lambda_n^2 - 2\lambda_n) (x_n^2 - \sigma_n^2 \epsilon^2) + \epsilon^2 \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2 = \\ &= \sum_{n=1}^{\infty} (x_n^2 - \sigma_n^2 \epsilon^2) + \sum_{n=1}^{\infty} (\lambda_n^2 - 2\lambda_n) x_n^2 + 2\epsilon^2 \sum_{n=1}^{\infty} \lambda_n \sigma_n^2. \end{aligned}$$

\mathcal{E} jest nieobciążonym estymatorem ryzyka. Naszym celem jest jednak minimalizacja tego estymatora ze względu na wagi λ_i , stąd wystarczy ograniczyć się do rozpatrywania wyrażenia

$$U(\lambda, X) = \sum_{n=1}^{\infty} (\lambda_n^2 - 2\lambda_n) x_n^2 + 2\epsilon^2 \sum_{n=1}^{\infty} \lambda_n \sigma_n^2 \quad (6)$$

będącego nieobciążonym estymatorem wyrażenia $\mathcal{R}(\hat{\theta}, \theta) - \sum_{n=1}^{\infty} \theta_n^2$.

Uwaga 2. *Wariancja funkcjonału $U(\lambda, X)$ wyraża się wzorem*

$$\begin{aligned} \text{Var} U(\lambda, X) &= 2\epsilon^4 \sum_{n=1}^{\infty} \lambda_n^4 \sigma_n^4 - \sum_{n=1}^{\infty} \lambda_n^4 \theta_n^4 - 2\epsilon^2 \sum_{n=1}^{\infty} \theta_n^2 \sigma_n^2 \lambda_n^4 + \\ &+ 8\epsilon^4 \sum_{n=1}^{\infty} \sigma_n^4 \lambda_n^2 - 4 \sum_{n=1}^{\infty} \lambda_n^2 \theta_n^4 - 8\epsilon^2 \sum_{n=1}^{\infty} \lambda_n^2 \sigma_n^2 \theta_n^2 - 8\epsilon^4 \sum_{n=1}^{\infty} \lambda_n^3 \sigma_n^4 + \end{aligned}$$

$$\begin{aligned}
& +4 \sum_{n=1}^{\infty} \lambda_n^3 \theta_n^4 + 8\epsilon^2 \sum_{n=1}^{\infty} \theta_n^2 \sigma_n^2 \lambda_n^3 \leq \\
& \leq 2\epsilon^4 \sum_{n=1}^{\infty} \lambda_n^4 \sigma_n^4 + 8\epsilon^4 \sum_{n=1}^{\infty} \sigma_n^4 \lambda_n^2.
\end{aligned}$$

Zatem dla skończoności wariancji przy założeniach 1 i 2 potrzebne jest dodatkowo założenie o skończoności następującej sumy dla dowolnego $\lambda \in \Lambda$

Założenie 3. (nie było w artykule)

$$\forall \lambda \in \Lambda \quad \sum_{n=1}^{\infty} \lambda_n^2 \sigma_n^4 < \infty.$$

Mając do dyspozycji odpowiednie wyrażenie związane z estymatorem ryzyka, możemy zdefiniować poszukiwany filtr.

Definicja 10. Niech funkcjonal $U(\lambda, X)$ będzie zdefiniowany jak w (6). Poszukiwanym filtrem jest element minimalizujący względem $\lambda \in \Lambda$ funkcjonal $U(\lambda, X)$, czyli

$$\lambda^* = \arg \min_{\lambda \in \Lambda} U(\lambda, X). \quad (7)$$

Dla tak zdefiniowanej metody wyboru wag pokazane i udowodnione zostaną nierówności wyrocznie.

Zdefiniujemy teraz kilka wartości użytych później w konstrukcji nierówności wyroczni.

Definicja 11. Wprowadźmy następujące oznaczenia:

$$\rho(\lambda) = \sup_n \sigma_n^2 |\lambda_n| \left[\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4 \right]^{-1/2},$$

$$\rho = \max_{\lambda \in \Lambda} \rho(\lambda),$$

$$S = \frac{\max_{\lambda \in \Lambda} \sup_n \sigma_n^2 \lambda_n^2}{\min_{\lambda \in \Lambda} \sup_n \sigma_n^2 \lambda_n^2},$$

$$M = \sum_{\lambda \in \Lambda} \exp \left(\frac{-1}{\rho(\lambda)} \right),$$

$$L_\lambda = \ln(NS) + \rho^2 \ln^2(MS).$$

Wartość $\rho(\lambda)$ jest pewnym sposobem mierzenia wielkości poszczególnych filtrów biorącym pod uwagę zarówno tempo znikania dalekich wyrazów poprzez $[\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4]^{-1/2}$ jak i rozrzut wokół zera poprzez $\sup_n \sigma_n^2 |\lambda_n|$. Parametry S i M mierzą natomiast zachowanie rodziny wag Λ . Liczbę S można interpretować jako rozrzut bądź zmienność w rodzinie Λ natomiast M jest czynnikiem kontrolującym masowość tej rodziny (dalszy komentarz na temat znaczenia

parametru M można znaleźć w pracy [2]). Zauważmy ponadto, że na mocy założeń 1 i 3 wszystkie te wartości są liczbami skończonymi.

Z założenia 2 wynika następująca nierówność

$$\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4 \leq \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^2.$$

Dodatkowo wymagać będziemy istnienia następującej stałej

Założenie 4.

$$\exists_{C_1 > 0} \forall_{\lambda \in \Lambda} \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^2 \leq C_1 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4.$$

Oznacza to, że będziemy wymagać by obie sumy były tego samego rzędu. Jak widzieliśmy z postaci wariancji funkcjonału $U(\lambda, X)$ sumy $\epsilon^4 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4$ i $\epsilon^4 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^2$ są głównymi jej składnikami. Z drugiej strony ze wzoru 3 mamy, że $\mathcal{R}(\hat{\theta}, \theta) \geq \epsilon^2 \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2$ oraz

$$\frac{(\epsilon^4 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4)^{1/2}}{\epsilon^2 \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leq \rho, \quad (8)$$

ponieważ z uwagi na założenie 2

$$\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4 \leq \sum_{k=1}^{\infty} \sigma_k^4 |\lambda_k|^3 = \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2 \cdot \sigma_k^2 |\lambda_k| \leq \sup_k \sigma_k^2 |\lambda_k| \leq \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2$$

a stąd i z definicji $\rho(\lambda)$ dostajemy, że dla dowolnego $\lambda \in \Lambda$

$$\frac{(\epsilon^4 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4)^{1/2}}{\epsilon^2 \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leq \rho(\lambda) \leq \rho.$$

Zatem parametr ρ pozwala kontrolować wielkość stosunku odchylenia standardowego do wartości oczekiwanej funkcjonału $U(\lambda, X)$, czyli $\mathbb{V}ar^{1/2} U(\lambda, X) / \mathcal{R}(\hat{\theta}, \theta)$ jednostajnie względem λ i θ .

Dodatkowo z uwagi na fakt, że $\sup_k \sigma_k^2 |\lambda_k| \leq \sqrt{\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^2}$ i założenie 4 mamy, że

$$\forall_{\lambda \in \Lambda} \rho(\lambda) \leq \sqrt{C_1}.$$

Przed wypowiedzeniem głównego twierdzenia zauważmy jeszcze, że zawsze zachodzi związek

$$M \leq N,$$

gdzie N oznacza licznosc rodziny Λ .

W poniższych dwóch twierdzeniach zebrane są główne wyniki otrzymane dla rozważanych problemów odwrotnych z operatorami zwartymi.

Twierdzenie 4. Niech założenia 1–4 będą spełnione. Wtedy dla dowolnego $\theta \in l^2$, dla dowolnego $B > B_0$ i dla estymatora liniowego θ^* z filtrem wybranym zgodnie z (7) zachodzi

$$\mathcal{R}(\theta^*, \theta) \leq (1 + \gamma_1 B^{-1}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + \gamma_2 B \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda),$$

gdzie stałe $B_0 > 0, \gamma_1 > 0, \gamma_2 > 0$ zależą tylko od stałej C_1 , wyrażenie $\min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta)$ rozumiane jest jako minimum wzięte po wszystkich estymatorach $\hat{\theta}$ postaci λX , $\lambda \in \Lambda$, a funkcja $\omega(x)$ jest postaci

$$\omega(x) = \max_{\lambda \in \Lambda} \sup_k \left[\sigma_k^2 \lambda_k^2 \mathbf{1} \left(\sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2 \leq x \sup_k \sigma_k^2 \lambda_k^2 \right) \right], \quad x > 0.$$

Twierdzenie 5. Niech założenia 1–4 będą spełnione. Wtedy istnieją stałe $\gamma_3 > 0, \gamma_4 > 0$ zależące tylko od C_1 , takie że dla dowolnego $\theta \in l^2$ i dla estymatora liniowego θ^* z filtrem wybranym zgodnie z (7) zachodzi

$$\mathcal{R}(\theta^*, \theta) \leq (1 + \gamma_3 \rho \sqrt{L_\Lambda}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

o ile $\rho \sqrt{L_\Lambda} < \gamma_4$, a minimum rozumiane jest jak w poprzednim twierdzeniu.

Zanim przejdziemy do dowodu powyższych twierdzeń podamy wniosek z twierdzenia 5, który pozwoli wnioskować o asymptotycznej dokładności podanych nierówności wyroczni postaci (5).

Wniosek 2. Niech założenia 1–4 będą spełnione, ponadto niech zachodzi $\lim_{\epsilon \rightarrow 0} \rho^2 \ln(NS) = 0$. Wtedy istnieją stałe $C_2 > 0, C_3 > 0$ zależące tylko od stałej C_1 , takie że dla $\rho^2 \ln(NS) < C_2$ i dla dowolnego $\theta \in l^2$ zachodzi

$$\mathcal{R}(\theta^*, \theta) \leq \left(1 + C_3 \rho \sqrt{\ln(NS)} \right) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

gdzie estymator $\hat{\theta}$ i minimum rozumiane są jak poprzednio.

Dowód. Skoro zachodzi $\lim_{\epsilon \rightarrow 0} \rho^2 \ln(NS) = 0$, zatem ciąg ten jest ograniczony przez pewną stałą zależną tylko od C_1 . Z twierdzenia 5 mamy, że stała γ_3 zależy tylko od stałej C_1 . Wystarczy zatem pokazać, że $L_\Lambda < C \ln(NS)$.

$$L_\Lambda = \ln(NS) + \rho^2 \ln^2(MS) \leq \ln(NS) + \rho^2 \ln^2(NS) \leq (1 + C_2) \ln(NS),$$

gdyż $M \leq N$. □

Zatem warunek $\lim_{\epsilon \rightarrow 0} \rho^2 \ln(NS) = 0$ jest warunkiem wystarczającym do otrzymania dokładnych asymptotycznie nierówności wyroczni postaci (5). Zanim przejdziemy do dowodu twierdzeń 4 i 5 podamy trzy lematy z których będziemy korzystać, a których dowody znajdują się w rozdziale 6.

Lemat 2. Niech $\{\xi_i\}$ będzie ciągiem niezależnych zmiennych losowych o tym samym standardowym rozkładzie normalnym i niech $v = \{v_i\}_{i=1}^{\infty} \in l^2$ będzie losowym ciągiem tak samo mierzalnym jak $\{\xi_i\}$ i takim, że przyjmuje on wartości w skończonym zbiorze $V \subset l^2$ o liczności $N > 1$. Wtedy dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} \left| \sum_{k=1}^{\infty} v_k \xi_k \right| \leq \sqrt{2 \ln(NK)} \left(\mathbb{E} \|v\| + \sqrt{2 \mathbb{E} \|v\|^2 / K} \right).$$

Lemat 3. Niech $\{\xi_i\}$ będzie ciągiem niezależnych zmiennych losowych o tym samym standardowym rozkładzie normalnym i niech $v = \{v_i\}_{i=1}^{\infty} \in l^2$ będzie losowym ciągiem tak samo mierzalnym jak $\{\xi_i\}$ i takim, że przyjmuje on wartości w skończonym zbiorze $V \subset l^2$ o liczności $N > 1$. Niech ponadto $v \neq 0$ dla dowolnego $v \in V$. Oznaczmy przez $m(v) = \sup_i |v_i| / \|v\|$, $m_V = \max_{v \in V} m(v)$ oraz

$$M(q) = \sum_{v \in V} \exp(-q/m_V), \quad q > 0.$$

Wtedy istnieje stała D zależna tylko od q , taka, że dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} \left| \sum_{k=1}^{\infty} v_k (\xi_k^2 - 1) \right| \leq D \left(\sqrt{\ln(NK)} + m_V \ln(M(q)K) \right) \left(\mathbb{E} \|v\| + \sqrt{\mathbb{E} \|v\|^2 / K} \right).$$

Lemat 4. Niech $\hat{\theta}_i = \hat{\lambda}_i(X) X_i$ będzie liniowym estymatorem z wagami z przedziału $[-1, 1]$ przyjmującym wartości w zbiorze Λ . Oznaczmy przez

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2, \quad \lambda \in \Lambda.$$

Wtedy istnieje absolutna stała $C > 0$ taka, że dla dowolnego $B > 0$ zachodzi

$$\mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|^2 \leq (1 + 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + CB \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].$$

Przejdziemy teraz do dowodów twierdzenia 4, a później do dowodu twierdzenia 5.

Dowód twierdzenia 4. Niech $\lambda^*, \lambda^* \in \Lambda$. Wtedy korzystając ze wzoru na różnicę kwadratów i szacowania $(a + b)^2 \leq 2(a^2 + b^2)$, można pokazać, że

$$[(1 - \lambda_i^*)^2 - (1 - \lambda_i^*)^2] \leq 4[(1 - \lambda_i^*)^2 + (1 - \lambda_i^*)^2][\lambda_i^{*2} + \lambda_i^{*2}].$$

(W artykule stała była 2, ale $(0.55, -1)$).

Niech teraz $\tilde{\lambda} \in \Lambda$ będzie takim filtrem, że związany z nim estymator jest wyrocznią, czyli $\tilde{\theta} = \arg \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta)$, natomiast przez λ^* oznaczmy filtr definiowany przez (7) i konsekwentnie związany z nim estymator przez θ^* .

W rozpatrywanym modelu (2) $x_i = \theta_i + \epsilon \sigma_i \xi_i$, zatem $x_i^2 = \theta_i^2 + \epsilon^2 \sigma_i^2 \xi_i^2 + 2\epsilon \sigma_i \theta_i \xi_i$. Wstawiając to wyrażenie do wzoru na nieobciążony estymator ryzyka (6) mamy, że

$$U[\lambda^*, X] = \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) (x_i^2 - \epsilon^2 \sigma_i^2) + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} =$$

$$\begin{aligned}
&= \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) (\theta_i^2 + \epsilon^2 \sigma_i^2 \xi_i^2 + 2\epsilon \sigma_i \theta_i \xi_i - \epsilon^2 \sigma_i^2) + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} = \\
&= 2\epsilon \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i \theta_i \xi_i + \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) (\theta_i^2 + \epsilon^2 \sigma_i^2 \xi_i^2 - \epsilon^2 \sigma_i^2) + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} = \\
&= 2\epsilon \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i \theta_i \xi_i + \epsilon^2 \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) + \\
&\quad + \left[\sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \theta_i^2 + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} + \sum_{i=1}^{\infty} \theta_i^2 \right] - \sum_{i=1}^{\infty} \theta_i^2 = \\
&= 2\epsilon \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i - 2\epsilon \sum_{i=1}^{\infty} \sigma_i \theta_i \xi_i - \sum_{i=1}^{\infty} \theta_i^2 + \epsilon^2 \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) + \mathcal{R}(\theta^*, \theta).
\end{aligned}$$

Obliczając wartość oczekiwaną powyższego wyrażenia dostajemy, że

$$\mathbb{E}_{\theta} U[\lambda^*, X] = \mathbb{E}_{\theta} \mathcal{R}(\theta^*, \theta) - \sum_{i=1}^{\infty} \theta_i^2 + \quad (9)$$

$$+ 2\epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i + \epsilon^2 \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1).$$

Znajdziemy teraz dolne oszacowania na dwa ostatnie składniki powyższego wyrażenia.

Zauważmy, że

$$\begin{aligned}
&\epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i = \epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i - 0 = \\
&= \epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i - \epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (1 - \tilde{\lambda}_i)^2 \sigma_i \theta_i \xi_i = \\
&= \epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2] \sigma_i \theta_i \xi_i \geq \\
&\geq -\epsilon \mathbb{E}_{\theta} \left| \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2] \sigma_i \theta_i \xi_i \right|.
\end{aligned}$$

Korzystając z lematu 2 z $K = S$ i $v_i = [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2] \sigma_i \theta_i$ dostajemy

$$\begin{aligned}
&-\epsilon \mathbb{E}_{\theta} \left| \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2] \sigma_i \theta_i \xi_i \right| \geq -\epsilon \sqrt{2 \ln(NS)} \mathbb{E}_{\theta} \left(\sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2] \sigma_i^2 \theta_i^2 \right)^{1/2} - \\
&- 2\epsilon \sqrt{\ln(NS)/S} \left(\mathbb{E}_{\theta} \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2] \sigma_i^2 \theta_i^2 \right)^{1/2}.
\end{aligned}$$

Następnie korzystając z nierówności wskazanej na początku dowodu dostajemy dolne oszacowanie postaci

$$\begin{aligned} & -2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2](\lambda_i^{*2}+\tilde{\lambda}_i^2)\theta_i^2\sigma_i^2\right)^{1/2} - \\ & -4\epsilon\sqrt{\ln(NS)/S}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2](\lambda_i^{*2}+\tilde{\lambda}_i^2)\theta_i^2\sigma_i^2\right)^{1/2} \end{aligned}$$

W kolejnym kroku będziemy korzystać z nierówności $2ab \leq B^{-1}a^2 + Bb^2$ zachodzącej dla dowolnego $B > 0$ i najpierw oszacujemy pierwszy składnik powyższego wyrażenia.

$$\begin{aligned} & 2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2](\lambda_i^{*2}+\tilde{\lambda}_i^2)\theta_i^2\sigma_i^2\right)^{1/2} \leq \\ & \leq 2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left\{\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda}_i^2)\sigma_i^2\right\}\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2\right\}^{1/2} = \\ & = \mathbb{E}_\theta 2\epsilon\sqrt{2\ln(NS)\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda}_i^2)\sigma_i^2\right\}}\left(\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2\right)^{1/2} \leq \\ & \leq 2B\epsilon^2\ln(NS)\mathbb{E}_\theta\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda}_i^2)\sigma_i^2\right\}+B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2 \leq \\ & \leq 2B\epsilon^2\ln(NS)\mathbb{E}_\theta\left(\sup_i\lambda_i^{*2}\sigma_i^2+\sup_i\tilde{\lambda}_i^2\sigma_i^2\right)+B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2. \end{aligned}$$

Analogicznie postępując z drugim wyrażeniem dostajemy

$$\begin{aligned} & 4\epsilon\sqrt{\ln(NS)/S}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2](\lambda_i^{*2}+\tilde{\lambda}_i^2)\theta_i^2\sigma_i^2\right)^{1/2} \leq \\ & \leq 4\epsilon\sqrt{\ln(NS)/S}\left(\mathbb{E}_\theta\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda}_i^2)\sigma_i^2\right\}\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2\right)^{1/2} \leq \\ & \leq 2\cdot 2\epsilon\sqrt{\ln(NS)/S}\left(\max_{\lambda\in\Lambda}\sup_i\left\{(\lambda_i^2+\tilde{\lambda}_i^2)\sigma_i^2\right\}\right)^{1/2}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2\right)^{1/2} \leq \\ & \leq 4B\epsilon^2\ln(NS)/S\max_{\lambda\in\Lambda}\sup_i\left\{(\lambda_i^2+\tilde{\lambda}_i^2)\sigma_i^2\right\}+B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda}_i)^2]\theta_i^2. \end{aligned}$$

Zauważmy, że skoro

$$S = \frac{\max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}{\min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}$$

to wyrażenie $\max_{\lambda \in \Lambda} \sup_i \{(\lambda_i^2 + \tilde{\lambda}_i^2) \sigma_i^2\} / S$ można oszacować przez

$$\begin{aligned} \max_{\lambda \in \Lambda} \sup_i \{(\lambda_i^2 + \tilde{\lambda}_i^2) \sigma_i^2\} / S &= \frac{\min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}{\max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2} \max_{\lambda \in \Lambda} \sup_i \{(\lambda_i^2 + \tilde{\lambda}_i^2) \sigma_i^2\} \leq \\ &= \frac{\min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2 \cdot \max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}{\max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2} + \frac{\min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2 \cdot \sup_i \tilde{\lambda}_i^2 \sigma_i^2}{\max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2} = \\ &= \min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2 + \frac{1}{S} \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \leq \sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2. \end{aligned}$$

Zatem dostajemy oszacowanie postaci

$$\begin{aligned} &4\epsilon \sqrt{\ln(NS)/S} \left(\mathbb{E}_\theta \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 + (1 - \tilde{\lambda}_i)^2] (\lambda_i^{*2} + \tilde{\lambda}_i^2) \theta_i^2 \sigma_i^2 \right)^{1/2} \leq \\ &\leq 4B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) + B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 + (1 - \tilde{\lambda}_i)^2] \theta_i^2, \end{aligned}$$

czyli takie samo z dokładnością do stałych jak dla czynnika pierwszego. Zatem możemy napisać, że

$$\begin{aligned} &\epsilon \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i \geq \\ &\geq -6B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) - 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 + (1 - \tilde{\lambda}_i)^2] \theta_i^2 \end{aligned}$$

Zachodzi również, że

$$\sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 \leq \mathcal{R}(\theta^*, \theta) \text{ oraz } \ln(NS) \leq L_\Lambda.$$

Przy oznaczeniu $\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2$ jak w lemacie 4 prowadzi to do oszacowania postaci

$$\begin{aligned} &-6B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) - 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 + (1 - \tilde{\lambda}_i)^2] \theta_i^2 \geq \\ &\geq -6B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) - 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 - 2B^{-1} \sum_{i=1}^{\infty} (1 - \tilde{\lambda}_i)^2 \theta_i^2 \geq \\ &\geq -2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 - 2B^{-1} \mathcal{R}(\tilde{\theta}, \theta) - 6B \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 6B \mathbb{E}_\theta \Delta^\epsilon[\tilde{\lambda}]. \quad (10) \end{aligned}$$

Znajdziemy teraz oszacowanie dla składnika $\epsilon^2 \mathbb{E}_\theta \sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1)$. Zauważmy na początek, że z uwagi na to, że $|\lambda_i| \leq 1$ dla dowolnego $\lambda \in \Lambda$ z założenia 2 zachodzi $\lambda_i^2 \leq (\lambda_i^2 - 2\lambda_i)^2 \leq 9\lambda_i^2$. Do oszacowania analizowanego wyrażenia posłużymy się lematem 3 z $K = S$, $q = 3$ i $v_i = (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2$. Zatem

$$\begin{aligned} \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) &\leq \epsilon^2 \mathbb{E}_\theta \left| \sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) \right| \leq \\ &\leq \epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta \|v\| + \sqrt{\mathbb{E}_\theta \|v\|^2 / S} \right). \end{aligned}$$

Oszacujemy teraz niektóre z elementów pojawiających się powyżej:

$$\begin{aligned} m(v) &= \sup_i |v_i| / \|v\| = \frac{\sup_i |\lambda_i^{*2} - 2\lambda_i^*| \sigma_i^2}{\sqrt{\sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*)^2 \sigma_i^4}} \leq \frac{3 \sup_i |\lambda_i| \sigma_i^2}{\sqrt{\sum_{i=1}^\infty \lambda_i^{*4} \sigma_i^4}} \leq 3\rho(\lambda), \\ M(3) &= \sum_v \exp(-3/m(v)) = \sum_{\lambda \in \Lambda} \exp(-3/m(v)) \leq \sum_{\lambda \in \Lambda} \exp(-1/\rho(\lambda)) = M, \\ m_V &= \max_v m(v) \leq 3\rho = 3 \max_{\lambda \in \Lambda} \rho(\lambda). \end{aligned}$$

Stąd dostajemy, że

$$\begin{aligned} \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right)^2 &\leq \left(\sqrt{\ln(NS)} + 3\rho \ln(MS) \right)^2 \leq \\ &\leq 2 \ln(NS) + 6\rho^2 \ln^2(MS) \leq 6(\ln(NS) + \rho^2 \ln^2(MS)), \end{aligned}$$

a stąd

$$\sqrt{\ln(NS)} + m_V \ln(SM(3)) \leq C \sqrt{L_\Lambda},$$

gdzie C jest pewną stałą. W dalszym ciągu rozważań wszystkie stałe generyczne zależne tylko od C_1 oznaczać będziemy przez C . Pozwala nam to oszacować badane wyrażenie od dołu przez

$$\begin{aligned} \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) &\geq -\epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta \|v\| + \sqrt{\mathbb{E}_\theta \|v\|^2 / S} \right) \geq \\ &\geq -\epsilon^2 C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*)^2 \sigma_i^4 \right)^{1/2} - \epsilon^2 C \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*)^2 \sigma_i^4 \right)^{1/2} \geq \\ &\geq -\epsilon^2 3C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\sum_{i=1}^\infty \lambda_i^{*4} \sigma_i^4 \right)^{1/2} - \epsilon^2 3C \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \sum_{i=1}^\infty \lambda_i^{*4} \sigma_i^4 \right)^{1/2} \geq \\ &\geq -\epsilon^2 C \sqrt{C_1} \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\sum_{i=1}^\infty \lambda_i^{*4} \sigma_i^4 \right)^{1/2} - \epsilon^2 C \sqrt{C_1} \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \sum_{i=1}^\infty \lambda_i^{*4} \sigma_i^4 \right)^{1/2}, \end{aligned}$$

z uwagi na to, że $(\lambda_i^{*2} - 2\lambda_i^*)^2 \leq 9\lambda_i^{*2}$ oraz założenie 4.

W dalszej części skorzystamy z faktu, że $\forall \omega \in \Omega \quad \min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2 \leq \sup_i \lambda_i^{*2} \sigma_i^2 =$

$\sup_i \lambda_i^{*2}(\omega) \sigma_i^2$, a stąd $\min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2 \leq \mathbb{E}_\theta \sup_i \lambda_i^{*2} \sigma_i^2$.
Analogicznie jak poprzednio

$$\begin{aligned} S^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^4 \lambda_i^{*4} &\leq \frac{\min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2}{\max_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2} \mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^{*2} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \\ &\leq \frac{\min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2}{\max_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2} \max_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \leq \\ &\leq \mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^{*2} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2}. \end{aligned}$$

Podobnie dla drugiego wyrażenia

$$\mathbb{E}_\theta \left(\sum_{i=1}^{\infty} \sigma_i^4 \lambda_i^{*4} \right)^{1/2} \leq \mathbb{E}_\theta \left(\sup_i \sigma_i^2 \lambda_i^{*2} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right)^{1/2}.$$

Ponownie korzystając z nierówności $2ab \leq B^{-1}a^2 + Bb^2$ dostajemy

$$\begin{aligned} C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\sup_i \sigma_i^2 \lambda_i^{*2} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right)^{1/2} &= \mathbb{E}_\theta \left(2C \sqrt{L_\Lambda \sup_i \sigma_i^2 \lambda_i^{*2}} \left(2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right)^{1/2} \right) \leq \\ &\leq B \mathbb{E}_\theta C^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^{*2} + 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2}. \end{aligned}$$

Oraz dla drugiego wyrażenia

$$\begin{aligned} C \sqrt{L_\Lambda} \left(\mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^{*2} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right)^{1/2} &= 2 \left(\mathbb{E}_\theta C L_\Lambda \sup_i \sigma_i^2 \lambda_i^{*2} \right)^{1/2} \left(2 \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right)^{1/2} \leq \\ &\leq B \mathbb{E}_\theta C L_\Lambda \sup_i \sigma_i^2 \lambda_i^{*2} + 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2}. \end{aligned}$$

Łącząc oba oszacowania oraz wykorzystując oznaczenie $\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2$ dostajemy

$$\begin{aligned} &\epsilon^1 \mathbb{E}_\theta \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) \geq \\ &\geq C \epsilon^2 B \mathbb{E}_\theta L_\Lambda \sup_i \sigma_i^2 \lambda_i^{*2} - 4 \epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} = \\ &= -CB \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 4 \epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2}. \end{aligned} \tag{11}$$

Łącząc wyrażenia (10) oraz (11) dostajemy, że

$$\begin{aligned}
& 2\epsilon\mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i + \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) \geq \\
& \geq -4B^{-1}\mathbb{E}_\theta \left[\sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right] - 4B^{-1}\mathcal{R}(\tilde{\theta}, \theta) - \\
& \quad - 12B\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 12B\Delta^\epsilon[\tilde{\lambda}] = \\
& = -4B^{-1}\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - 4B^{-1}\mathcal{R}(\tilde{\theta}, \theta) - CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - CB\Delta^\epsilon[\tilde{\lambda}]. \quad (12)
\end{aligned}$$

Możemy teraz przejść do ostatecznego szacowania ryzyka badanego estymatora. Zgodnie z (9) mamy wykorzystując oszacowania (10) i (11) (bądź od razu (12)), że

$$\begin{aligned}
& \mathbb{E}_\theta U[\lambda^*, X] = \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - \sum_{i=1}^{\infty} \theta_i^2 + \\
& + 2\epsilon\mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i + \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) \geq \\
& \geq \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - \sum_{i=1}^{\infty} \theta_i^2 - 4B^{-1}\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - 4B^{-1}\mathcal{R}(\tilde{\theta}, \theta) - CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - CB\Delta^\epsilon[\tilde{\lambda}].
\end{aligned}$$

Zatem zachodzi, że

$$(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq \mathbb{E}_\theta U[\lambda^*, X] + \sum_{i=1}^{\infty} \theta_i^2 + 4B^{-1}\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}].$$

Jednak, jako że filtr λ^* był zdefiniowany w (7) jako argument minimalizujący nieobciążony estymator ryzyka (6), musi zachodzić, że

$$\mathbb{E}_\theta U[\lambda^*, X] \leq \mathbb{E}_\theta U[\tilde{\lambda}, X] = \mathcal{R}(\tilde{\theta}, \theta) + \sum_{i=1}^{\infty} \theta_i^2.$$

Dochodzimy zatem do oszacowania postaci

$$(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}]. \quad (13)$$

Zauważmy, że dla dowolnego $x > 0$ zachodzi następujące oszacowanie

$$\begin{aligned}
& \sup_i \sigma_i^2 \lambda_i^2 = \sup_i \sigma_i^2 \lambda_i^2 \mathbf{1} \left\{ x \sup_i \sigma_i^2 \lambda_i^2 < \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \right\} + \sup_i \sigma_i^2 \lambda_i^2 \mathbf{1} \left\{ x \sup_i \sigma_i^2 \lambda_i^2 \geq \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \right\} \leq \\
& \leq \frac{1}{x} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 + \max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2 \mathbf{1} \left\{ x \sup_i \sigma_i^2 \lambda_i^2 \geq \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \right\} \leq
\end{aligned}$$

$$\leq \frac{1}{x} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 + \omega(x) \leq \frac{1}{x\epsilon^2} \mathcal{R}(\hat{\theta}, \theta) + \omega(x) \quad \forall \lambda \in \Lambda.$$

Stąd dostajemy oszacowanie na składnik

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2 \leq \frac{L_\Lambda}{x} \mathcal{R}(\hat{\theta}, \theta) + \epsilon^2 L_\Lambda \omega(x).$$

Wykorzystamy teraz powyższe nierówności dla nierówności $(1-4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1+4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}]$.

$$\begin{aligned} (1-4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) &\leq (1+4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}] \leq \\ &\leq CB \frac{L_\Lambda}{x} \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(x) + CB\epsilon^2 L_\Lambda \omega(x) + CB \frac{L_\Lambda}{x} \mathcal{R}(\tilde{\theta}, \theta) + (1+4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) = \\ &= CB \frac{L_\Lambda}{x} \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(x) + \left(CB \frac{L_\Lambda}{x} + 1 + 4B^{-1} \right) \mathcal{R}(\tilde{\theta}, \theta). \end{aligned}$$

Mamy stąd, że

$$\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq \frac{CB\epsilon^2 L_\Lambda \omega(x)}{1-4B^{-1}-CB \frac{L_\Lambda}{x}} + \frac{CB \frac{L_\Lambda}{x} + 1 + 4B^{-1}}{1-4B^{-1}-CB \frac{L_\Lambda}{x}} \mathcal{R}(\tilde{\theta}, \theta).$$

Korzystając z lematu 4 mamy ponadto, że

$$\begin{aligned} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1+4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] \leq \\ &\leq (1+4B^{-1} + CB \frac{L_\Lambda}{x})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(x). \end{aligned}$$

Niech teraz $x = B^2 L_\Lambda$ oraz $\gamma = 4 + C$. Wtedy z powyższych nierówności dostajemy, że

$$\begin{aligned} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1+\gamma B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) \leq \\ &\leq \frac{1+\gamma B^{-1}}{1-\gamma B^{-1}} CB\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) + CB\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) + \frac{(1+\gamma B^{-1})^2}{1-\gamma B^{-1}}. \end{aligned}$$

Zauważmy następnie, że z uwagi na fakt, że $\lim_{x \rightarrow \infty} \frac{x+\gamma}{x-\gamma} = 1$, mamy, że $B \frac{1+\gamma B^{-1}}{1-\gamma B^{-1}} = C'B$ dla pewnej stałej C' zależnej tylko od C_1 (poprzez γ) o ile tylko $B > B_0$ dla pewnego $B_0 > \gamma$. Analogicznie widzimy, że $\frac{(1+\gamma B^{-1})^2}{1-\gamma B^{-1}} = (1+\gamma B^{-1}) \cdot C''$ o ile znowu $B > B_0$. Zwiększając stałą γ i oznaczając ją przez γ_1 dostajemy, że $\frac{(1+\gamma B^{-1})^2}{1-\gamma B^{-1}} = 1 + \gamma_1 B^{-1}$. Ostatecznie prowadzi nas do nierówności

$$\mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq (1+\gamma_1 B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + \gamma_2 B\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda).$$

Nierówność ta kończy dowód twierdzenia 4. \square

Dowód twierdzenia 5. Będziemy stosować te same oznaczenia na estymator i wyrocznie jak w dowodzie twierdzenia 4.

Zauważmy na początek, że z uwagi (8) mamy, że

$$\frac{(\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4)^{1/2}}{\sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leq \rho \implies \frac{1}{\sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leq \frac{\rho}{(\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4)^{1/2}}.$$

Zatem możemy oszacować wyrażenie $\frac{\sup_i \sigma_i^2 \lambda_i^2}{\sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2}$ w następujący sposób

$$\frac{\sup_i \sigma_i^2 \lambda_i^2}{\sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leq \rho \frac{\sup_i \sigma_i^2 \lambda_i^2}{(\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4)^{1/2}} \leq \rho^2$$

korzystając z definicji ρ . Zauważmy ponadto, że dla dowolnego $\lambda \in \Lambda$ zachodzi, że

$$\mathcal{R}(\hat{\theta}, \theta) = \sum_{i=1}^{\infty} (1 - \lambda_i) \theta_i^2 + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \geq \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2.$$

Łącząc te dwie nierówności dostajemy, że

$$\epsilon^2 \sup_i \sigma_i^2 \lambda_i^2 \leq \epsilon^2 \rho^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \leq \rho^2 \mathcal{R}(\hat{\theta}, \theta).$$

Pozwala nam to oszacować wyrażenie $\Delta^\epsilon[\lambda]$ dla dowolnego $\lambda \in \Lambda$ i dowolnego $\theta \in l^2$ przez

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2 \leq \rho^2 L_\Lambda \mathcal{R}(\hat{\theta}, \theta).$$

Następnie analogicznie jak w dowodzie twierdzenia 4 możemy dostać nierówność postaci (oszacowanie (13))

$$(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}].$$

Co prowadzi do nierówności

$$(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\rho^2 L_\Lambda \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\rho^2 L_\Lambda \mathcal{R}(\tilde{\theta}, \theta),$$

a stąd dostajemy

$$\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq \frac{1 + 4B^{-1} + CB\rho^2 L_\Lambda}{1 - 4B^{-1} - CB\rho^2 L_\Lambda} \mathcal{R}(\tilde{\theta}, \theta).$$

Ponownie korzystając z lematu 4 dostajemy, że

$$\begin{aligned} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1 + 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] \leq \\ &\leq (1 + 4B^{-1} + CB\rho^2 L_\Lambda)\mathcal{R}(\tilde{\theta}, \theta). \end{aligned}$$

Łącząc te dwie nierówności mamy

$$\mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq \frac{(1 + 4B^{-1} + CB\rho^2 L_\Lambda)^2}{1 - 4B^{-1} - CB\rho^2 L_\Lambda} \mathcal{R}(\tilde{\theta}, \theta).$$

Zauważmy, że istnieje taka stała $\gamma_4 > 0$, że jeśli tylko zachodzi $\rho^2 L_\Lambda \leq \gamma_4$, to wybór B jako $(\rho^2 L_\Lambda)^{-1/2}$ prowadzi do nierówności $4B^{-1} + CB\rho^2 L_\Lambda < 1/2$, a stąd $1 - 4B^{-1} - CB\rho^2 L_\Lambda \geq 1/2$, czyli jest odcięte od zera. Wtedy wybór $B = (\rho^2 L_\Lambda)^{-1/2}$ prowadzi do nierówności

$$\mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq \frac{(1 + (4 + C)\rho^2 \sqrt{L_\Lambda})^2}{1 - (4 + C)\rho^2 \sqrt{L_\Lambda}} \mathcal{R}(\tilde{\theta}, \theta),$$

co rozumując analogicznie do dowodu twierdzenia 5 prowadzi do nierówności

$$\mathcal{R}(\tilde{\theta}, \theta) \leq (1 + \gamma_3 \rho \sqrt{L_\Lambda}) \mathcal{R}(\tilde{\theta}, \theta),$$

która kończy dowód. \square

5 Główne rezultaty II

W tej części pracy podejmiemy próbę zbudowania analogicznych jak w rozdziale poprzednim nierówności wyrocni w przypadku, gdy w rozważanym problemie

$$Y = Af + \epsilon \xi$$

operator A jest operatorem samosprzężonym, ale już niekoniecznie zwartym.

Będziemy potrzebowali pojęcia operatora unitarnego.

Definicja 12. Operator $U: H \rightarrow H$ na ośrodkowej przestrzeni Hilberta nazywamy operatorem unitarnym, jeżeli jest ciągły, liniowy, bijektywny i zachodzi $U^* = U^{-1}$.

Uwaga 3. Zauważmy, że jeżeli operator U jest unitarny, to zachowuje on iloczyn skalarny, ponieważ dla dowolnych $f, g \in H$ zachodzi

$$\langle f, g \rangle = \langle U^{-1}Uf, g \rangle = \langle U^*Uf, g \rangle = \langle Uf, Ug \rangle.$$

Okazuje się, że dla operatorów normalnych można uogólnić twierdzenie o reprezentacji według wartości singularnych do następującej postaci

Twierdzenie 6. Niech A będzie operatorem samosprzężonym na ośrodkowej przestrzeni Hilberta H . Wtedy istnieje σ -skończona przestrzeń (S, \mathcal{S}, μ) , rzeczywista funkcja $b \in L_\infty(S, \mathcal{S}, \mu)$, operator unitarny $U: H \rightarrow L_2(S, \mathcal{S}, \mu)$, takie, że

$$UAU^{-1} = M_b,$$

gdzie M_b jest operatorem mnożenia przez funkcję b zdefiniowanym jako $(M_b g)(x) = b(x)g(x)$.

Dowód. Dowód w [9]. \square

Zauważmy, że operatory zwarte i ich reprezentacja według wartości singu-
larnych jest specjalnym przypadkiem powyższego twierdzenia, gdzie $S = \mathbb{N}$,
 $L_2(S, \mathcal{S}, \mu) = l^2(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ z μ jako miarą liczącą, funkcją b taką, że $b(k) = b_k$
oraz operatorem U przeprowadzającym funkcję f w ciąg jej współczynników
Fouriera.

Jeżeli ξ jest szumem na przestrzeni H , natomiast operator T jest określony na
tej samej przestrzeni, to możemy zdefiniować działanie operatora T na ξ w
następujący sposób ([1])

$$\langle T\xi, f \rangle = \langle \xi, T^* f \rangle, \quad \forall f \in H.$$

Korzystając z powyższego twierdzenia spektralnego możemy rozważany model

$$Y = Af + \epsilon\xi = U^{-1}M_b Uf + \epsilon\xi$$

przekształcić do następującej postaci stosując operator U

$$UY = U(U^{-1}M_b U)f + \epsilon U\xi,$$

co można zapisać jako

$$Z = b\theta + \epsilon\eta,$$

gdzie $Z = UY$, $\theta = Uf$ oraz $\eta = U\xi$.

Analogicznie jak w przypadku operatorów zwartych możemy zapisać to wyrażenie
w postaci analogicznej do (2)

$$X = \theta + \epsilon\sigma\eta, \tag{14}$$

gdzie $\sigma = b^{-1}$.

Twierdzenie 6 mówi tylko, że funkcja b należy do $L_\infty(S, \mathcal{S}, \mu)$ nie mówiąc nic
więcej o jej zachowaniu. Zatem o źle postawionym problemie będziemy mogli
mówić tylko wtedy, gdy $b \rightarrow 0$ w jakimś sensie.

Przed dalszym rozumowaniem, udowodnimy, czym jest η . Okazuje się, że jest
ona gaussowskim białym szumem na przestrzeni $L_2(S, \mathcal{S}, \mu)$.

Twierdzenie 7. η zdefiniowana jako $\eta = U\xi$ jest gaussowskim białym szumem
na przestrzeni $L_2(S, \mathcal{S}, \mu)$.

Dowód. Jako, że operator U jest unitarny mamy, że dla dowolnego elementu
 $F \in L_2(S, \mathcal{S}, \mu)$ $F = Uf$ dla pewnego $f \in H$ oraz

$$\langle \eta, F \rangle = \langle U\xi, Uf \rangle = \langle \xi, U^* Uf \rangle = \langle \xi, f \rangle \sim \mathcal{N}(0, \|f\|_H^2).$$

Ponadto zachodzi, że dla $F = Uf$, $G = Ug$

$$\text{Cov}(\langle \eta, F \rangle, \langle \eta, G \rangle) = \text{Cov}(\langle \xi, f \rangle, \langle \xi, g \rangle) = \langle f, g \rangle.$$

Analogicznie dostajemy także, że dla $F_i = Uf_i$, $i = 1, 2, \dots, k$ mamy

$$(\langle \eta, F_1 \rangle, \langle \eta, F_2 \rangle, \dots, \langle \eta, F_k \rangle) = (\langle \xi, f_1 \rangle, \langle \xi, f_2 \rangle, \dots, \langle \xi, f_k \rangle) \sim \mathcal{N}_k(\mathbf{0}, \mathbf{\Sigma}).$$

□

Niech teraz $\hat{\theta}$ oznacza estymator elementu θ w modelu (14) na podstawie obserwacji X . Wtedy estymatorem poszukiwanego elementu f jest $U^{-1}\hat{\theta} = \hat{f}$ i możemy wyrazić ryzyko tego estymatora w następujący sposób

$$\mathcal{R}(\hat{f}, f) = \mathbb{E}_f \|\hat{f} - f\|_H^2 = \mathbb{E}_\theta \|\hat{\theta} - \theta\|_S^2 = \mathbb{E}_\theta \int_S |\hat{\theta} - \theta|^2 d\mu,$$

gdzie w drugiej równości skorzystaliśmy z faktu, że U jest liniowym operatorem unitarnym, czyli zachowuje normę elementu.

Możemy teraz uogólnić pojęcie estymatora liniowego na rozważany przypadek.

Definicja 13. Niech λ będzie nielosową funkcją zespoloną lub rzeczywistą taką, że $\lambda \in L_2(S, \mathcal{S}, \mu)$. Wtedy estymatorem liniowym $\hat{\theta}$ elementu θ w modelu (14) nazwiemy wyrażenie

$$\hat{\theta} = \lambda X.$$

Funkcję λ będziemy konsekwentnie nazywać wagą lub filtrem.

W przypadku estymatorów liniowych wyrażenie na ich ryzyko sprowadza się do postaci

$$\mathcal{R}(\hat{\theta}, \theta) = \mathbb{E}_\theta \|\hat{\theta} - \theta\|_2^2 = \|(1 - \lambda)\theta\|_2^2 + \epsilon^2 \|\sigma\lambda\|_2^2 = \int_S |1 - \lambda|^2 |\theta|^2 d\mu + \epsilon^2 \int_S \sigma^2 |\lambda|^2 d\mu,$$

gdź $\sigma = 1/b$ jest funkcją rzeczywistą.

W dalszej części ograniczymy się tylko do rozważania rzeczywistych filtrów, w przypadku których wyrażenie na ryzyko przyjmuje postać

$$\mathcal{R}(\hat{\theta}, \theta) = \int_S (1 - \lambda)^2 |\theta|^2 d\mu + \epsilon^2 \int_S \sigma^2 \lambda^2 d\mu.$$

Podobnie jak w przypadku rozważanych wcześniej modeli z operatorem zwartym wyrażenie $|X|^2 - \epsilon^2 \sigma^2$ jest nieobciążonym estymatorem dla $|\theta|^2$, co prowadzi nas do analogicznej do (6) definicji nieobciążonego estymatora ryzyka.

Definicja 14. Nieobciążonym estymatorem ryzyka w modelu (14) nazywamy wyrażenie

$$U(\lambda, X) = \int_S (\lambda^2 - 2\lambda) |X|^2 d\mu - 2\epsilon \int_S \sigma^2 \lambda^2 d\mu, \quad (15)$$

będące nieobciążonym estymatorem wyrażenia $\mathcal{R}(\hat{\theta}, \theta) - \int_S \theta^2 d\mu$ dla estymatora postaci $\hat{\theta} = \lambda X$.

Niech teraz badane filtry należą do pewnej skończonej rodziny $\Lambda = \{\lambda^1, \dots, \lambda^N\}$. Naszym celem będzie wybór na podstawie tylko i wyłącznie obserwacji takiego filtra z tej rodziny, by związany z nim estymator naśladował ryzyko najlepszego estymatora w Λ .

Analogicznie do rozważanego wcześniej przypadku wymagać będziemy, by odpowiednie człony występujące w wyrażeniu na ryzyko estymatora oraz wariancję $U(\lambda, X)$ były skończone oraz by estymatory nie były zbyt duże. Założymy także skończoność wyrażeń postaci $\|\sigma\lambda^2\|_\infty$ oraz $\|\sigma^2\lambda^2\|_\infty$ występujących później w wyrażeniach występujących w nierównościach wyroczniach dla badanych estymatorów.

Założenie 5. *Założmy, że*

$$\begin{aligned}
& \forall \lambda \in \Lambda \quad 0 < \int_S \sigma^2 \lambda^2 d\mu < \infty, \\
& \max_{\lambda \in \Lambda} \|\lambda\|_\infty \leq 1, \\
& \forall \lambda \in \Lambda \quad \int_S \lambda^2 \sigma^4 d\mu < \infty, \\
& \exists C_1 > 0 \forall \lambda \in \Lambda \quad \int_S \sigma^4 \lambda^2 d\mu \leq C_1 \int_S \sigma^4 \lambda^4 d\mu, \\
& \forall \lambda \in \Lambda \quad \|\sigma \lambda^2\|_\infty < \infty, \\
& \forall \lambda \in \Lambda \quad \|\sigma^2 \lambda^2\|_\infty < \infty.
\end{aligned}$$

W kolejnym kroku wprowadzimy potrzebne oznaczenia.

Definicja 15. *Wprowadźmy następujące oznaczenia:*

$$\begin{aligned}
\rho(\lambda) &= \|\sigma_n^2 \lambda\|_\infty \left[\int_S \sigma_k^4 \lambda_k^4 d\mu \right]^{-1/2}, \\
\rho &= \max_{\lambda \in \Lambda} \rho(\lambda), \\
S &= \frac{\max_{\lambda \in \Lambda} \|\sigma^2 \lambda^2\|_\infty}{\min_{\lambda \in \Lambda} \|\sigma^2 \lambda^2\|_\infty}, \\
M &= \sum_{\lambda \in \Lambda} \exp \left(\frac{-1}{\rho(\lambda)} \right), \\
L_\lambda &= \ln(NS) + \rho^2 \ln^2(MS).
\end{aligned}$$

Interpretacje powyższych wyrażeń przenoszą się z rozważanego wcześniej dyskretnego przypadku.

Mając już odpowiednie założenia możemy zdefiniować poszukiwany estymator naśladowujący ryzyko najlepszego estymatora w klasie Λ . Ponownie wykorzystamy w tym celu zdefiniowany wcześniej nieobciążony estymator ryzyka, który w pewien sposób przybliża prawdziwe ryzyko, które chcielibyśmy zminimalizować.

Definicja 16. *Niech funkcjonal $U(\lambda, X)$ będzie zdefiniowany jak w (15). Poszukiwanym filtrem jest element minimalizujący względem $\lambda \in \Lambda$ funkcjonal $U(\lambda, X)$, czyli*

$$\lambda^* = \arg \min_{\lambda \in \Lambda} U(\lambda, X). \quad (16)$$

Przedstawimy teraz uogólnione wersje lematów 2, 3 i 4 na rozważany obecnie przypadek. Dowody znajdują się w części 6.

Lemat 5. Niech η będzie gaussowskim białym szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ i niech $v \in L_2(S, \mathcal{S}, \mu)$ będzie losowym elementem tej przestrzeni ze skończonego zbioru $V \subset L_2(S, \mathcal{S}, \mu)$ o liczności $N > 1$. Wtedy dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} |\langle \eta, v \rangle| \leq \sqrt{2 \ln(NK)} \left(\mathbb{E} \|v\|_2 + \sqrt{2 \mathbb{E} \|v\|_2^2 / K} \right).$$

Lemat 6. Niech η będzie gaussowskim białym szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ i niech $v \in L_2(S, \mathcal{S}, \mu)$ będzie losowym elementem tej przestrzeni ze skończonego zbioru $V \subset L_2(S, \mathcal{S}, \mu)$ o liczności $N > 1$. Niech ponadto $v \neq 0$ dla dowolnego $v \in V$. Oznaczmy przez $m(v) = \|v\|_\infty / \|v\|_2$, $m_V = \max_{v \in V} m(v)$ oraz

$$M(q) = \sum_{v \in V} \exp(-q/m_V), \quad q > 0.$$

Wtedy istnieje stała D zależna tylko od q , taka, że dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} |\langle \eta^2 - 1, v \rangle| \leq D \left(\sqrt{\ln(NK)} + m_V \ln(M(q)K) \right) \left(\mathbb{E} \|v\| + \sqrt{\mathbb{E} \|v\|^2 / K} \right).$$

Uwaga 4. W powyższym lemacie pojawiło się wyrażenie postaci $\eta^2 - 1$. W ogólności możemy zdefiniować wielomian, którego argumentem jest biały szum w następujący sposób. Niech $P = P(t_1, t_1, \dots, t_n)$ będzie zwykłym wielomianem zmiennych rzeczywistych t_1, t_2, \dots, t_n oraz niech f_1, f_2, \dots, f_n będą elementami przestrzeni H . Wtedy wyrażenie $P(\eta)$ zdefiniowane jako złożenie P i funkcji $\langle \eta, f_1 \rangle, \dots, \langle \eta, f_n \rangle$, czyli $P(\langle \eta, f_1 \rangle, \dots, \langle \eta, f_n \rangle)$ jest wielomianem względem białego szumu η . Jeżeli elementy te są wzajemnie ortogonalne, to rząd wielomianu od szumu jest taki sam jak wielomianu wyjściowego. Można pokazać, że dowolny taki wielomian jest całkowalny z kwadratem ([7]).

Lemat 7. Niech $\hat{\theta} = \hat{\lambda}(X)X$ będzie liniowym estymatorem z wagą z wartościami z przedziału $[-1, 1]$ przyjmującym wartości w zbiorze Λ . Oznaczmy przez

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \|\sigma^2 \lambda^2\|_\infty, \quad \lambda \in \Lambda.$$

Wtedy istnieje absolutna stała $C > 0$ taka, że dla dowolnego $B > 0$ zachodzi

$$\mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|_2^2 \leq (1 + 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + CB \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].$$

Możemy teraz przejść do wypowiedzenia dwóch twierdzeń będących zarazem głównym wynikiem tego rozdziału i uogólnieniem twierdzeń 4 i 5 na przypadek, gdy badany problem jest modelowany przy pomocy operatora, który nie musi być zwarty, jednak jest samosprężony.

Twierdzenie 8. Niech założenie 5 będzie spełnione. Wtedy dla dowolnego $\theta \in L_2(S, \mathcal{S}, \mu)$, dla dowolnego $B > B_0$ i dla estymatora liniowego θ^* z filtrem wybranym zgodnie z (16) zachodzi

$$\mathcal{R}(\theta^*, \theta) \leq (1 + \gamma_1 B^{-1}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + \gamma_2 B \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda),$$

gdzie stałe $B_0 > 0, \gamma_1 > 0, \gamma_2 > 0$ zależą tylko od stałej C_1 , wyrażenie $\min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta)$ rozumiane jest jako minimum wzięte po wszystkich estymatorach $\hat{\theta}$ postaci λX , $\lambda \in \Lambda$, a funkcja $\omega(x)$ jest postaci

$$\omega(x) = \max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2 \mathbf{1} \left(\int_S \sigma^2 \lambda^2 d\mu \leq x \|\sigma^2 \lambda^2\|_\infty \right) \right\|_\infty, \quad x > 0.$$

Twierdzenie 9. Niech założenie 5 będzie spełnione.. Wtedy istnieją stałe $\gamma_3 > 0, \gamma_4 > 0$ zależące tylko od C_1 , takie że dla dowolnego $\theta \in L_2(S, \mathcal{S}, \mu)$ i dla estymatora liniowego θ^* z filtrem wybranym zgodnie z (16) zachodzi

$$\mathcal{R}(\theta^*, \theta) \leq (1 + \gamma_3 \rho \sqrt{L_\Lambda}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

o ile $\rho \sqrt{L_\Lambda} < \gamma_4$, a minimum rozumiane jest jak w poprzednim twierdzeniu.

Dowód twierdzenia 8. Dowód tego twierdzenia jest naturalnym uogólnieniem dowodu twierdzenia 4. Niech $\tilde{\lambda} \in \Lambda$ będzie takim filtrem, że związany z nim estymator jest wyrocznią, czyli $\tilde{\theta} = \arg \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta)$, natomiast przez λ^* oznaczymy filtr definiowany przez (16) i związany z nim estymator przez θ^* .

W rozpatrywanym modelu (14) $X = \theta + \epsilon \sigma \eta$. Dostajemy stąd, że

$$U[\lambda^*, X] = 2\epsilon \int_S (1 - \lambda^*)^2 \sigma |\theta| \eta d\mu - 2\epsilon \int_S \sigma |\theta| \eta d\mu - \int_S |\theta|^2 d\mu + \epsilon^2 \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) + \mathcal{R}(\theta^*, \theta).$$

A stąd mamy

$$\begin{aligned} \mathbb{E}_\theta U[\lambda^*, X] &= \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - \int_S |\theta|^2 d\mu + \\ &+ 2\epsilon \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 \sigma |\theta| \eta d\mu + \epsilon^2 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu. \end{aligned}$$

Korzystając z wprowadzonych wcześniej lematów, oszacujemy dwa ostatnie składniki tego wyrażenia.

Zauważmy, że zachodzi

$$\begin{aligned} \epsilon \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 \sigma |\theta| \eta d\mu &= \epsilon \mathbb{E}_\theta \int_S [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma |\theta| \eta d\mu \geq \\ &\geq -\epsilon \mathbb{E}_\theta \left| \int_S [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma |\theta| \eta d\mu \right|. \end{aligned}$$

Korzystając z lematu 5 z $K = S$ i $v = [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma |\theta|$ dostajemy

$$\begin{aligned} -\epsilon \mathbb{E}_\theta \left| \int_S [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma |\theta| \eta d\mu \right| &\geq -\epsilon \sqrt{2 \ln(NS)} \mathbb{E}_\theta \left(\int_S [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma^2 |\theta|^2 d\mu \right)^{1/2} - \\ &- 2\epsilon \sqrt{\ln(NS)/S} \left(\mathbb{E}_\theta \int_S [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma^2 |\theta|^2 d\mu \right)^{1/2} \geq \end{aligned}$$

$$\begin{aligned} &\geq -2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta \left(\int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda}^2)|\theta|^2\sigma^2 d\mu \right)^{1/2} - \\ &-4\epsilon\sqrt{\ln(NS)/S} \left(\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda}^2)|\theta|^2\sigma^2 d\mu \right)^{1/2}. \end{aligned}$$

Korzystając z nierówności $2ab \leq B^{-1}a^2 + Bb^2$ zachodzącej dla dowolnego $B > 0$ dla pierwszego składnika dostajemy

$$\begin{aligned} &2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta \left(\int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda}^2)|\theta|^2\sigma^2 d\mu \right)^{1/2} \leq \\ &\leq \mathbb{E}_\theta 2\epsilon\sqrt{2\ln(NS)} \left\| (\lambda^{*2} + \tilde{\lambda}^2)\sigma^2 \right\|_\infty \left(\int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]|\theta|^2 d\mu \right)^{1/2} \leq \\ &\leq 2B\epsilon^2 \ln(NS)\mathbb{E}_\theta \left\| (\lambda^{*2} + \tilde{\lambda}^2)\sigma^2 \right\|_\infty + B^{-1}\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]|\theta|^2 d\mu \leq \\ &\leq 2B\epsilon^2 \ln(NS)\mathbb{E}_\theta \left(\left\| \lambda^{*2}\sigma^2 \right\|_\infty + \left\| \tilde{\lambda}^2\sigma^2 \right\|_\infty \right) + B^{-1}\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]|\theta|^2 d\mu. \end{aligned}$$

Postępując podobnie z drugim wyrażeniem otrzymujemy

$$\begin{aligned} &4\epsilon\sqrt{\ln(NS)/S} \left(\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda}^2)|\theta|^2\sigma^2 d\mu \right)^{1/2} \leq \\ &\leq 4B\epsilon^2 \ln(NS)/S \max_{\lambda \in \Lambda} \left\| (\lambda^2 + \tilde{\lambda}^2)\sigma^2 \right\|_\infty + B^{-1}\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]|\theta|^2 d\mu. \end{aligned}$$

Korzystając z definicji wielkości S mamy, że

$$\frac{\max_{\lambda \in \Lambda} \left\| (\lambda^2 + \tilde{\lambda}^2)\sigma^2 \right\|_\infty}{S} \leq \left\| \lambda^{*2}\sigma^2 \right\|_\infty + \left\| \tilde{\lambda}^2\sigma^2 \right\|_\infty.$$

A stąd mamy oszacowanie postaci

$$\begin{aligned} &4\epsilon\sqrt{\ln(NS)/S} \left(\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda}^2)|\theta|^2\sigma^2 d\mu \right)^{1/2} \leq \\ &\leq 4B\epsilon^2 \ln(NS)\mathbb{E}_\theta \left(\left\| \lambda^{*2}\sigma^2 \right\|_\infty + \left\| \tilde{\lambda}^2\sigma^2 \right\|_\infty \right) + B^{-1}\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]|\theta|^2 d\mu. \end{aligned}$$

Jako, że zachodzi

$$\int_S (1-\lambda^*)^2|\theta|^2 d\mu \leq \mathcal{R}(\theta^*, \theta) \text{ oraz } \ln(NS) \leq L_\Lambda,$$

to łącząc powyższe dwa oszacowania dostajemy, że

$$\epsilon\mathbb{E}_\theta \int_S (1-\lambda^*)^2\sigma|\theta|\eta d\mu \geq$$

$$\geq -2B^{-1}\mathbb{E}_\theta \int_S (1-\lambda^*)^2 |\theta|^2 d\mu - 2B^{-1}\mathcal{R}(\tilde{\theta}, \theta) - 6B\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 6B\mathbb{E}_\theta \Delta^\epsilon[\tilde{\lambda}].$$

Znajdziemy teraz oszacowanie dla składnika $\epsilon^2 \mathbb{E}_\theta \left| \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \right|$. Zauważmy, że dla dowolnego $\lambda \in \Lambda$ zachodzi $\lambda^2 \leq (\lambda^2 - 2\lambda)^2 \leq 9\lambda^2$ rozumiane jako nierówność funkcyjna zachodząca dla wszystkich argumentów funkcji λ . Posłużymy się teraz lematem 6 z $K = S$, $q = 3$ i $v = (\lambda^{*2} - 2\lambda^*)\sigma^2$.

$$\begin{aligned} & \epsilon^2 \mathbb{E}_\theta \left| \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \right| \leq \\ & \leq \epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta \|v\| + \sqrt{\mathbb{E}_\theta \|v\|^2 / S} \right). \end{aligned}$$

Analogicznie jak w rozważanym wcześniej przypadku dekompozycji według wartości singularnych mamy, że

$$m(v) \leq 3\rho(\lambda), \quad M(3) \leq M, \quad m_V \leq 3 \max_{\lambda \in \Lambda} \rho(\lambda) = 3\rho.$$

Stąd dostajemy, że

$$\sqrt{\ln(NS)} + m_V \ln(SM(3)) \leq C\sqrt{L_\Lambda},$$

gdzie C jest pewną stałą zależną tylko od C_1 . Możemy teraz kontynuować szacowanie analizowanego wyrażenia.

$$\begin{aligned} & \epsilon^2 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \geq -\epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta \|v\| + \sqrt{\mathbb{E}_\theta \|v\|^2 / S} \right) \geq \\ & \geq -\epsilon^2 C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\int_S (\lambda^{*2} - 2\lambda^*)^2 \sigma^4 d\mu \right)^{1/2} - \epsilon^2 C \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*)^2 \sigma^4 d\mu \right)^{1/2} \geq \\ & \geq -\epsilon^2 C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^{*4} \sigma^4 d\mu \right)^{1/2} - \epsilon^2 C \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \int_S \lambda^{*4} \sigma^4 d\mu \right)^{1/2}. \end{aligned}$$

Korzystając z faktu, że $\min_{\lambda \in \Lambda} \|\lambda^2 \sigma^2\|_\infty \leq \mathbb{E}_\theta \|\lambda^{*2} \sigma^2\|_\infty$ mamy, że

$$S^{-1} \mathbb{E}_\theta \int_S \sigma^4 \lambda^{*4} d\mu \leq \mathbb{E}_\theta \|\sigma^2 \lambda^{*2}\|_\infty \mathbb{E}_\theta \int_S \sigma^2 \lambda^{*2} d\mu.$$

Podobnie dla drugiego wyrażenia

$$\mathbb{E}_\theta \left(\int_S \sigma^4 \lambda^{*4} d\mu \right)^{1/2} \leq \mathbb{E}_\theta \left(\|\sigma^2 \lambda^{*2}\|_\infty \int_S \sigma^2 \lambda^{*2} d\mu \right)^{1/2}.$$

Ponownie korzystając z nierówności $2ab \leq B^{-1}a^2 + Bb^2$ dostajemy

$$C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\|\sigma^2 \lambda^{*2}\|_\infty \int_S \sigma^2 \lambda^{*2} d\mu \right)^{1/2} \leq B \mathbb{E}_\theta C^2 L_\Lambda \|\sigma^2 \lambda^{*2}\|_\infty + 2B^{-1} \mathbb{E}_\theta \int_S \sigma^2 \lambda^{*2} d\mu.$$

Możemy także zapisać

$$C\sqrt{L_\Lambda} \left(\mathbb{E}_\theta \|\sigma^2 \lambda^{*2}\|_\infty \mathbb{E}_\theta \int_S \sigma^2 \lambda^{*2} d\mu \right)^{1/2} \leq B\mathbb{E}_\theta CL_\Lambda \|\sigma^2 \lambda^{*2}\|_\infty + 2B^{-1}\mathbb{E}_\theta \int_S \sigma^2 \lambda^{*2} d\mu.$$

Łącząc oba te oszacowania dostajemy

$$\epsilon^1 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \geq -CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 4\epsilon^2 B^{-1} \mathbb{E}_\theta \int_S \sigma^2 \lambda^{*2} d\mu.$$

Powyższe rozważania pozwalają zapisać nam następujący wniosek

$$\begin{aligned} & 2\epsilon \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 \sigma |\theta| \eta d\mu + \epsilon^2 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \geq \\ & \geq -4B^{-1} \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - 4B^{-1} \mathcal{R}(\tilde{\theta}, \theta) - CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - CB\Delta^\epsilon[\tilde{\lambda}]. \end{aligned}$$

Możemy teraz przejść do szacowania ryzyka badanego estymatora.

$$\begin{aligned} \mathbb{E}_\theta U[\lambda^*, X] &= \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - \int_S |\theta|^2 d\mu + \\ &+ 2\epsilon \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 \sigma |\theta| \eta d\mu + \epsilon^2 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \geq \\ &\geq \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - \int_S |\theta|^2 d\mu - 4B^{-1} \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - 4B^{-1} \mathcal{R}(\tilde{\theta}, \theta) - CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - CB\Delta^\epsilon[\tilde{\lambda}]. \end{aligned}$$

Zatem zachodzi, że

$$(1 - 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq \mathbb{E}_\theta U[\lambda^*, X] + \int_S |\theta|^2 d\mu + 4B^{-1} \mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}].$$

Jednak, jako że filtr λ^* był zdefiniowany jako argument minimalizujący nieobciążony estymator ryzyka, musi zachodzić, że

$$\mathbb{E}_\theta U[\lambda^*, X] \leq \mathbb{E}_\theta U[\tilde{\lambda}, X] = \mathcal{R}(\tilde{\theta}, \theta) + \int_S |\theta|^2 d\mu.$$

Otrzymujemy zatem, że

$$(1 - 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1}) \mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}].$$

Zauważmy, że dla dowolnego $x > 0$ zachodzi następujące oszacowanie

$$\begin{aligned} \|\sigma^2 \lambda^2\|_\infty &= \left\| \sigma^2 \lambda^2 \mathbf{1} \left\{ x \|\sigma^2 \lambda^2\|_\infty < \int_S \sigma^2 \lambda^2 d\mu \right\} \right\|_\infty + \left\| \sigma^2 \lambda^2 \mathbf{1} \left\{ x \|\sigma^2 \lambda^2\|_\infty \geq \int_S \sigma^2 \lambda^2 d\mu \right\} \right\|_\infty \leq \\ &\leq \frac{1}{x} \int_S \sigma^2 \lambda^2 d\mu + \max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2 \mathbf{1} \left\{ x \|\sigma^2 \lambda^2\|_\infty \geq \int_S \sigma^2 \lambda^2 d\mu \right\} \right\|_\infty \leq \\ &\leq \frac{1}{x} \int_S \sigma^2 \lambda^2 d\mu + \omega(x) \leq \frac{1}{x\epsilon^2} \mathcal{R}(\hat{\theta}, \theta) + \omega(x) \quad \forall \lambda \in \Lambda. \end{aligned}$$

Stąd dostajemy

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \|\sigma^2 \lambda^2\|_\infty \leq \frac{L_\Lambda}{x} \mathcal{R}(\hat{\theta}, \theta) + \epsilon^2 L_\Lambda \omega(x).$$

Wykorzystamy teraz powyższe nierówności do wyrażenia $(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}]$.

$$\begin{aligned} (1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) &\leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}] \leq \\ &\leq CB \frac{L_\Lambda}{x} \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(x) + \left(CB \frac{L_\Lambda}{x} + 1 + 4B^{-1} \right) \mathcal{R}(\tilde{\theta}, \theta). \end{aligned}$$

Mamy stąd, że

$$\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq \frac{CB\epsilon^2 L_\Lambda \omega(x)}{1 - 4B^{-1} - CB \frac{L_\Lambda}{x}} + \frac{CB \frac{L_\Lambda}{x} + 1 + 4B^{-1}}{1 - 4B^{-1} - CB \frac{L_\Lambda}{x}} \mathcal{R}(\tilde{\theta}, \theta).$$

Korzystając z lematu 7 mamy ponadto, że

$$\begin{aligned} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1 + 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] \leq \\ &\leq (1 + 4B^{-1} + CB \frac{L_\Lambda}{x})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(x). \end{aligned}$$

Niech teraz $x = B^2 L_\Lambda$ oraz γ będzie stałą zależną tylko od C (zależące tylko od stałej C_1). Wtedy

$$\begin{aligned} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1 + \gamma B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) \leq \\ &\leq \frac{1 + \gamma B^{-1}}{1 - \gamma B^{-1}} CB\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) + CB\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) + \frac{(1 + \gamma B^{-1})^2}{1 - \gamma B^{-1}}. \end{aligned}$$

Korzystając z faktu, że $\lim_{x \rightarrow \infty} \frac{x + \gamma}{x - \gamma} = 1$, powyższe rozważania prowadzą nas do nierówności

$$\mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq (1 + \gamma_1 B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + \gamma_2 B\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda),$$

która kończy dowód twierdzenia 8. \square

Dowód twierdzenia 9. Będziemy stosować te same oznaczenia na estymator i wyrocznię jak w dowodzie twierdzenia 5. Dowód ten ponownie stanowi naturalne uogólnienie dowodu twierdzenia 5.

Wyrażenie $\frac{\|\sigma^2 \lambda^2\|_\infty}{\int_S \sigma^2 \lambda^2 d\mu}$ możemy oszacować w następujący sposób

$$\frac{\|\sigma^2 \lambda^2\|_\infty}{\int_S \sigma^2 \lambda^2 d\mu} \leq \rho \frac{\|\sigma^2 \lambda^2\|_\infty}{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}} \leq \rho^2.$$

Zauważmy ponadto, że dla dowolnego $\lambda \in \Lambda$ zachodzi, że

$$\mathcal{R}(\hat{\theta}, \theta) = \int_S (1 - \lambda) |\theta|^2 d\mu + \epsilon^2 \int_S \sigma^2 \lambda^2 d\mu \geq \epsilon^2 \int_S \sigma^2 \lambda^2 d\mu.$$

Skąd dostajemy, że

$$\epsilon^2 \|\sigma^2 \lambda^2\|_\infty \rho^2 \mathcal{R}(\hat{\theta}, \theta).$$

Co z kolei prowadzi do oszacowania

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \|\sigma^2 \lambda^2\|_\infty \leq \rho^2 L_\Lambda \mathcal{R}(\hat{\theta}, \theta).$$

W dowodzie twierdzenia 8 uzyskaliśmy nierówność następującej postaci

$$(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB\Delta^\epsilon[\tilde{\lambda}].$$

Wykorzystując wyprowadzone oszacowanie mamy stąd, że

$$(1 - 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1})\mathcal{R}(\tilde{\theta}, \theta) + CB\rho^2 L_\Lambda \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\rho^2 L_\Lambda \mathcal{R}(\tilde{\theta}, \theta),$$

co prowadzi do

$$\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq \frac{1 + 4B^{-1} + CB\rho^2 L_\Lambda}{1 - 4B^{-1} - CB\rho^2 L_\Lambda} \mathcal{R}(\tilde{\theta}, \theta).$$

Ponownie korzystając z lematu 7 dostajemy, że

$$\begin{aligned} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1 + 4B^{-1})\mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) + CB\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] \leq \\ &\leq (1 + 4B^{-1} + CB\rho^2 L_\Lambda)\mathcal{R}(\tilde{\theta}, \theta). \end{aligned}$$

Łącząc te dwie nierówności mamy

$$\mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq \frac{(1 + 4B^{-1} + CB\rho^2 L_\Lambda)^2}{1 - 4B^{-1} - CB\rho^2 L_\Lambda} \mathcal{R}(\tilde{\theta}, \theta).$$

Zauważmy teraz, że istnieje taka stała $\gamma_4 > 0$, że jeśli tylko zachodzi $\rho^2 L_\Lambda \leq \gamma_4$, to wybór B jako $(\rho^2 L_\Lambda)^{-1/2}$ prowadzi do nierówności $4B^{-1} + CB\rho^2 L_\Lambda < 1/2$, a stąd $1 - 4B^{-1} - CB\rho^2 L_\Lambda \geq 1/2$, czyli jest odcięte od zera. Wtedy wybór $B = (\rho^2 L_\Lambda)^{-1/2}$ prowadzi do nierówności

$$\mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq \frac{(1 + (4 + C)\rho^2 \sqrt{L_\Lambda})^2}{1 - (4 + C)\rho^2 \sqrt{L_\Lambda}} \mathcal{R}(\tilde{\theta}, \theta),$$

która z kolei prowadzi nas do postulowanej na początku nierówności

$$\mathcal{R}(\tilde{\theta}, \theta) \leq (1 + \gamma_3 \rho \sqrt{L_\Lambda}) \mathcal{R}(\tilde{\theta}, \theta),$$

która kończy dowód twierdzenia 9. □

Wniosek 3. Niech założenie 5 będzie spełnione, ponadto niech zachodzi $\lim_{\epsilon \rightarrow 0} \rho^2 \ln(NS) = 0$. Wtedy istnieją stałe $C_2 > 0, C_3 > 0$ zależące tylko od stałej C_1 , takie że dla $\rho^2 \ln(NS) < C_2$ i dla dowolnego $\theta \in l^2$ zachodzi

$$\mathcal{R}(\theta^*, \theta) \leq \left(1 + C_3 \rho \sqrt{\ln(NS)}\right) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

gdzie estymator $\tilde{\theta}$ i minimum rozumiane są jak poprzednio.

6 Lematy pomocnicze

W rozdziale tym podamy dowody lematów użytych do dowodzenia głównych wyników pracy.

Lemat. Niech $\{\xi_i\}$ będzie ciągiem niezależnych zmiennych losowych o tym samym standardowym rozkładzie normalnym i niech $v = \{v_i\}_{i=1}^\infty \in l^2$ będzie losowym ciągiem tak samo mierzalnym jak $\{\xi_i\}$ i takim, że przyjmuje on wartości w skończonym zbiorze $V \subset l^2$ o liczności N . Wtedy dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} \left| \sum_{k=1}^\infty v_k \xi_k \right| \leq \sqrt{2 \ln(NK)} \left(\mathbb{E} \|v\| + \sqrt{2 \mathbb{E} \|v\|^2 / K} \right).$$

Dowód. Na początek zauważmy pewien użyteczny fakt.

Niech $X \sim \mathcal{N}(0, 1)$ będzie zmienną losową. Wtedy zachodzi oszacowanie

$$\mathbb{E} X^2 \mathbf{1}_{\{|X| > a\}} \leq \frac{2}{\sqrt{2\pi}} (a + 2a^{-1}) e^{-a^2/2} \quad \forall a > 0.$$

Istotnie możemy napisać, że

$$\begin{aligned} \mathbb{E} X^2 \mathbf{1}_{\{|X| > a\}} &= \int_{-\infty}^{-a} \frac{1}{\sqrt{2\pi}} x^2 e^{-x^2/2} dx + \int_a^{\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-x^2/2} dx = \frac{2}{\sqrt{2\pi}} \int_a^{\infty} x^2 e^{-x^2/2} dx \leq \\ &\leq \frac{2}{\sqrt{2\pi}} \int_a^{\infty} \frac{x^3}{a} e^{-x^2/2} dx = \frac{4}{a\sqrt{2\pi}} \int_{a^2/2}^{\infty} t e^{-t} dt = \frac{2}{\sqrt{2\pi}} (a + 2a^{-1}) e^{-a^2/2}. \end{aligned}$$

Przejdziemy teraz do dowodu właściwej części lematu.

Oznaczmy przez $\zeta_v = \frac{1}{\|v\|} \sum_{k=1}^\infty v_k \xi_k$. Zachodzi wtedy oczywiście, że $|\zeta_v| \|v\| = |\sum_{k=1}^\infty v_k \xi_k|$. Możemy zatem zapisać, że

$$\mathbb{E} \left| \sum_{k=1}^\infty v_k \xi_k \right| = \mathbb{E} |\zeta_v| \|v\| \leq \mathbb{E} \|v\| \max_{v \in V} |\zeta_v| =$$

$$\begin{aligned} \mathbb{E} \|v\| \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| \leq \sqrt{2 \ln(NK)}\}} + \mathbb{E} \|v\| \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} &\leq \\ &\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \mathbb{E} \|v\| \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}}. \end{aligned}$$

Skorzystamy następnie dla drugiego członu z nierówności Cauchy'ego–Schwartz.

$$\begin{aligned} &\sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \mathbb{E} \|v\| \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \leq \\ &\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \left(\mathbb{E} \|v\|^2 \right)^{1/2} \left(\mathbb{E} \max_{v \in V} |\zeta_v|^2 \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2}. \end{aligned}$$

Rozważmy teraz funkcję $F(t) = t^2 \mathbf{1}_{\{t > \sqrt{2 \ln(NK)}\}}$. Pokażemy, że zachodzi

$$F(\max_{v \in V} |\zeta_v|) \leq \sum_{v \in V} F(|\zeta_v|).$$

Z uwagi na monotoniczność funkcji kwadratowej dla dodatnich argumentów zachodzi, że

$$F(\max_{v \in V} |\zeta_v|) = \max_{v \in V} F(|\zeta_v|),$$

gdym

$$\begin{aligned} P &= \max_{v \in V} \left(|\zeta_v| \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^2 = \max_{v \in V} \left(|\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right) = \\ &= \max_{v \in V} \left(|\zeta_v|^2 \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \right) = \max_{v \in V} (|\zeta_v|)^2 \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} = L. \end{aligned}$$

A stąd dostajemy juŹ wprost, że

$$F(\max_{v \in V} |\zeta_v|) = \max_{v \in V} F(|\zeta_v|) \leq \sum_{v \in V} F(|\zeta_v|),$$

poniewaŹ $\forall_t F(t) \geq 0$.

ZauwaŹmy ponadto, że skoro zmienne $\{\xi_i\}$ s niezaleŹnymi zmiennymi losowymi o rozkadzie $\mathcal{N}(0, 1)$, to zmienne ζ_v teŹ maj takie same rozkady normalne $\mathcal{N}(0, 1)$ dla kaŹdego $v \in V$. Zatem moŹemy napisać, że

$$\begin{aligned} \mathbb{E} \left| \sum_{k=1}^{\infty} v_k \xi_k \right| &\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \left(\mathbb{E} \|v\|^2 \right)^{1/2} \left(\mathbb{E} \max_{v \in V} |\zeta_v|^2 \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} \leq \\ &\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \left(\mathbb{E} \|v\|^2 \right)^{1/2} \left(\sum_{v \in V} \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} = \\ &= \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \left(\mathbb{E} \|v\|^2 \right)^{1/2} \left(N \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} \end{aligned}$$

Następnie korzystajc z oszacowania pokazanego na pocztku dowodu z $a = \sqrt{2 \ln(NK)}$ dostajemy, że

$$\begin{aligned} N \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} &\leq \frac{2N}{\sqrt{2\pi}} \left(\sqrt{2 \ln(NK)} + \frac{2}{\sqrt{2 \ln(NK)}} \right) \frac{1}{NK} \leq \\ &\leq \frac{4}{\sqrt{2\pi}} \sqrt{2 \ln(NK)} / K \leq \frac{2}{K} \sqrt{2 \ln(NK)} \leq \frac{2}{K} \cdot 2 \ln(NK), \end{aligned}$$

o ile tylko zachodzi, że $NK \geq 2$. Przy załoŹeniu na K jest to spełnione dla kaŹdego nietrywialnego problemu z licznosci $V > 1$. Łczajc te nierowności dostajemy, że

$$\begin{aligned} \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \left(\mathbb{E} \|v\|^2 \right)^{1/2} \left(N \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} &\leq \\ &\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \sqrt{2 \ln(NK)} \left(\mathbb{E} \|v\|^2 \cdot \frac{2}{K} \right)^{1/2} \end{aligned}$$

co kończy dowód. □

Lemat. Niech $\{\xi_i\}$ będzie ciągiem niezależnych zmiennych losowych o tym samym standardowym rozkładzie normalnym i niech $v = \{v_i\}_{i=1}^\infty \in l^2$ będzie losowym ciągiem tak samo mierzalnym jak $\{\xi_i\}$ i takim, że przyjmuje on wartości w skończonym zbiorze $V \subset l^2$ o liczności N . Niech ponadto $v \neq 0$ dla dowolnego $v \in V$. Oznaczmy przez $m(v) = \sup_i |v_i|/\|v\|$, $m_V = \max_{v \in V} m(v)$ oraz

$$M(q) = \sum_{v \in V} \exp(-q/m_V), \quad q > 0.$$

Wtedy istnieje stała D zależna tylko od q , taka, że dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} \left| \sum_{k=1}^{\infty} v_k(\xi_k^2 - 1) \right| \leq D \left(\sqrt{\ln(NK)} + m_V \ln(M(q)K) \right) \left(\mathbb{E}\|v\| + \sqrt{\mathbb{E}\|v\|^2/K} \right).$$

Dowód. Zauważmy na początek, że z nierówności Markowa dostajemy, że dla dowolnego $t > 0$ zachodzi, że $P(X > \epsilon) \leq e^{-t\epsilon} \mathbb{E}e^{tX}$. Policzmy pomocniczo następującą wartość oczekiwaną

$$\mathbb{E} \exp(a\xi_i^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ax^2} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}(1-2a)} dx = \sqrt{\frac{1}{1-2a}}.$$

Niech teraz $\eta_v = \frac{1}{\sqrt{2}\|v\|} \sum_{i=1}^{\infty} v_i(\xi_i^2 - 1)$. Korzystając z powyższych faktów możemy napisać dla dowolnego $t > 0$

$$\begin{aligned} P(\eta_v > x) &\leq \exp(-tx) \mathbb{E} \exp(t\eta_v) = \exp(-tx) \mathbb{E} \exp \left(t \frac{1}{\sqrt{2}\|v\|} \sum_{i=1}^{\infty} v_i(\xi_i^2 - 1) \right) = \\ &= \exp(-tx) \prod_{i=1}^{\infty} \mathbb{E} \exp \left(t \frac{1}{\sqrt{2}\|v\|} v_i(\xi_i^2 - 1) \right) = \\ &= \exp(-tx) \prod_{i=1}^{\infty} \exp \left(-\frac{tv_i}{\sqrt{2}\|v\|} \right) \mathbb{E} \exp \left(\frac{t}{\sqrt{2}\|v\|} v_i \xi_i^2 \right) = \\ &= \exp(-tx) \prod_{i=1}^{\infty} \exp \left(-\frac{tv_i}{\sqrt{2}\|v\|} \right) \sqrt{\frac{\|v\|}{\|v\| - \sqrt{2}tv_i}} = \\ &= \exp(-tx) \prod_{i=1}^{\infty} \exp \left(-\frac{tv_i}{\sqrt{2}\|v\|} - \frac{1}{2} \ln \left(1 - \frac{\sqrt{2}tv_i}{\|v\|} \right) \right). \end{aligned}$$

W powyższym wyrażeniu skorzystać będziemy chcieli z następującego rozwinięcia $-\ln(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k}$ dla $|x| < 1$. Zatem musimy założyć dodatkowo, że $t < \frac{1}{\sqrt{2}m(v)}$. Stąd

$$\exp(-tx) \prod_{i=1}^{\infty} \exp \left(-\frac{tv_i}{\sqrt{2}\|v\|} - \frac{1}{2} \ln \left(1 - \frac{\sqrt{2}tv_i}{\|v\|} \right) \right) =$$

$$\begin{aligned}
&= \exp(-tx) \prod_{i=1}^{\infty} \exp \left(-\frac{tv_i}{\sqrt{2}\|v\|} + \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{k} \left(\frac{\sqrt{2}tv_i}{\|v\|} \right)^k \right) = \\
&= \exp(-tx) \exp \left(\sum_{i=1}^{\infty} \sum_{k=2}^{\infty} \frac{1}{2k} \left(\frac{\sqrt{2}tv_i}{\|v\|} \right)^k \right) = \exp(-tx) \exp \left(\sum_{k=2}^{\infty} \sum_{i=1}^{\infty} \frac{1}{2k} \left(\frac{\sqrt{2}tv_i}{\|v\|} \right)^k \right) = \\
&= \exp(-tx) \exp \left(\sum_{k=2}^{\infty} \frac{(\sqrt{2}t)^k}{2k} \sum_{i=1}^{\infty} \left(\frac{v_i}{\|v\|} \right)^2 \left(\frac{v_i}{\|v\|} \right)^{k-2} \right) = \\
&\leq \exp(-tx) \exp \left(\sum_{k=2}^{\infty} \frac{(\sqrt{2}t)^k}{2k} (m(v))^{k-2} \frac{\sum_{i=1}^{\infty} v_i^2}{\|v\|^2} \right) = \\
&= \exp(-tx) \exp \left(\frac{1}{2m^2(v)} \sum_{k=2}^{\infty} \frac{1}{k} \left(\sqrt{2}tm(v) \right)^k \right) = \\
&= \exp(-tx) \exp \left(-\frac{1}{2m^2(v)} \ln \left(1 - \sqrt{2}tm(v) \right) - \frac{t}{\sqrt{2}m(v)} \right).
\end{aligned}$$

Minimalizując ostatnie wyrażenie względem t dostajemy, że w punkcie $t = \frac{1}{\sqrt{2}m(v)} - \frac{1}{2m^2(v) + \sqrt{2}m(v)}$ osiągane jest minimum o wartości $\exp \left(\frac{1}{2m^2(v)} \ln \left(1 + \sqrt{2}m(v)x \right) - \frac{x}{\sqrt{2}m(v)} \right)$. Możemy zatem zapisać, że

$$\begin{aligned}
P(\eta_v > x) &\leq \exp \left(\frac{1}{2m^2(v)} \ln \left(1 + \sqrt{2}m(v)x \right) - \frac{x}{\sqrt{2}m(v)} \right) = \\
&= \exp \left(\frac{1}{2m^2(v)} \left(\ln \left(1 + \sqrt{2}m(v)x \right) - \sqrt{2}m(v)x \right) \right).
\end{aligned}$$

Zauważmy następnie, że zachodzi następująca zależność

$$\ln(1+u) - u = u \int_0^1 \left(-\frac{tu}{1+tu} \right) dt \leq - \int_0^1 \frac{tu^2}{1+u} dt = -\frac{u^2}{2(1+u)}.$$

Zatem powyższe oszacowanie sprowadza się do postaci

$$P(\eta_v > x) \leq \exp \left(-\frac{x^2}{2(1 + \sqrt{2}m(v)x)} \right).$$

Skąd dostajemy oszacowanie

$$P(|\eta_v| > x) \leq 2 \exp \left(-\frac{x^2}{2(1 + \sqrt{2}m(v)x)} \right).$$

Wyrażenie $-\frac{x^2}{2(1 + \sqrt{2}m(v)x)}$ możemy ograniczyć w następujący sposób

$$-\frac{x^2}{2(1 + \sqrt{2}m(v)x)} \leq \begin{cases} -\frac{x^2}{4}, & \sqrt{2}m(v)x < 1 \\ -\frac{x}{\sqrt{32}m(v)}, & \sqrt{2}m(v)x > 1 \end{cases}.$$

Zauważmy następnie, że dla nieujemnej zmiennej losowej X zachodzi, że $\mathbb{E}X^2 = 2 \int_0^\infty tP(X > t)dt$. Zatem dla dowolnego $Q > 0$ mamy, że

$$\begin{aligned}\mathbb{E}\eta_v^2 \mathbf{1}\{|\eta_v| > Q\} &= 2 \int_Q^\infty tP(|\eta_v| > t)dt \leq 4 \int_Q^\infty t \exp\left(-\frac{t^2}{2(1+\sqrt{2}m(v)t)}\right) dt \leq \\ &\leq 4 \int_Q^{\frac{1}{\sqrt{2}m(v)}} t \exp\left(-\frac{t^2}{4}\right) dt + 4 \int_Q^\infty t \exp\left(-\frac{t}{\sqrt{32}m(v)}\right) dt \leq \\ &\leq C \exp\left(-\frac{Q^2}{4}\right) + CQ \exp\left(-\frac{Q}{\sqrt{32}m(v)}\right).\end{aligned}$$

Następnie możemy zapisać, że

$$\sum_{v \in V} \exp\left(-\frac{Q}{\sqrt{32}m(v)}\right) = \sum_{v \in V} \exp\left(-\frac{q}{m(v)}\right) \exp\left(-\frac{Q/\sqrt{32}-q}{m(v)}\right) \leq M(q) \exp\left(-\frac{Q/\sqrt{32}-q}{m_V}\right),$$

o ile $Q > q\sqrt{32}$.

Będziemy teraz postępować analogicznie jak w dowodzie poprzedniego lematu dla dowolnego $Q > q\sqrt{32}$.

$$\begin{aligned}\mathbb{E}\left|\sum_{i=1}^\infty v_i(\xi_i^2 - 1)\right| &= \mathbb{E}\left|\sqrt{2}\|v\| |\eta_v|\right| \leq \sqrt{2}\mathbb{E}\|v\| \max_{v \in V} |\eta_v| \leq \\ &\leq \sqrt{2}\mathbb{E}\|v\| \max_{v \in V} |\eta_v| \mathbf{1}\{\max_{v \in V} |\eta_v| \leq Q\} + \sqrt{2}\mathbb{E}\|v\| \max_{v \in V} |\eta_v| \mathbf{1}\{\max_{v \in V} |\eta_v| > Q\} \leq \\ &\leq \sqrt{2}Q\mathbb{E}\|v\| + \sqrt{2}\mathbb{E}\|v\| \max_{v \in V} |\eta_v| \mathbf{1}\{\max_{v \in V} |\eta_v| > Q\} \leq \\ &\leq \sqrt{2}Q\mathbb{E}\|v\| + \left(2\mathbb{E}\|v\|^2\right)^{1/2} \left(\mathbb{E} \max_{v \in V} |\eta_v|^2 \mathbf{1}\{\max_{v \in V} |\eta_v| > Q\}\right)^{1/2} \leq \\ &\leq \sqrt{2}Q\mathbb{E}\|v\| + \left(2\mathbb{E}\|v\|^2\right)^{1/2} \left(\sum_{v \in V} \mathbb{E}|\eta_v|^2 \mathbf{1}\{\max_{v \in V} |\eta_v| > Q\}\right)^{1/2} \leq \\ &\leq \sqrt{2}Q\mathbb{E}\|v\| + \left(2\mathbb{E}\|v\|^2\right)^{1/2} \left(\sum_{v \in V} \left(C \exp\left(-\frac{Q^2}{4}\right) + CQ \exp\left(-\frac{Q}{\sqrt{32}m(v)}\right)\right)\right)^{1/2} \leq \\ &\leq \sqrt{2}Q\mathbb{E}\|v\| + \left(2\mathbb{E}\|v\|^2\right)^{1/2} \left(NC \exp\left(-\frac{Q^2}{4}\right) + CQM(q) \exp\left(-\frac{Q/\sqrt{32}-q}{m_V}\right)\right)^{1/2}.\end{aligned}$$

Przyjmując teraz $Q = 2\sqrt{\ln(NK)} + \sqrt{32}m_V \ln(M(q)K) + q\sqrt{32}$ dostajemy, że sprowadza się ono do następującej postaci

$$\left(2\sqrt{2}\sqrt{\ln(NK)} + 8m_V \ln(M(q)K) + 8q\right) \mathbb{E}\|v\| + \left(2\mathbb{E}\|v\|^2\right)^{1/2}.$$

$$\begin{aligned}
& \left(NC \exp(-\ln(NK)) \exp\left(-8m_V^2 \ln^2(M(q)K) - 8q - 8\sqrt{2}q\sqrt{\ln(NK)} - 32qm_V \ln(M(q)K) - \right. \right. \\
& \quad \left. \left. - 8\sqrt{2}\sqrt{\ln(NK)}m_V \ln(M(q)K) \right) + CM(q) \left(2\sqrt{\ln(NK)} + \sqrt{32}m_V \ln(M(q)K) + q\sqrt{32} \right) \right. \\
& \quad \left. \exp(-\ln(M(q)K)) \exp\left(-\frac{\sqrt{\ln(NK)}}{2\sqrt{2}m_V}\right) \right)^{1/2} \leq \\
& \leq D \left(\sqrt{\ln(NK)} + m_V \ln(M(q)K) \right) \left(\mathbb{E}\|v\| + \left(\frac{1}{K} \mathbb{E}\|v\|^2 \right)^{1/2} \right),
\end{aligned}$$

dla pewnej stałej D zależnej tylko od q .

Nierówność ta kończy dowód lematu. \square

Lemat. Niech $\hat{\theta}_i = \lambda_i(X)X_i$ będzie liniowym estymatorem z wagami z przedziału $[-1, 1]$ przyjmującym wartości w zbiorze Λ . Oznaczmy przez

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2, \quad \lambda \in \Lambda.$$

Wtedy istnieje absolutna stała $C > 0$ taka, że dla dowolnego $B > 0$ zachodzi

$$\mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|^2 \leq (1 + 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + CB \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].$$

Dowód. Przypomnijmy, że w naszym modelu (2) mamy, że $x_i = \theta_i + \epsilon \sigma_i \xi_i$, a stąd $x_i^2 = \theta_i^2 + \epsilon^2 \sigma_i^2 \xi_i^2 + 2\epsilon \theta_i \sigma_i \xi_i$.

Zgodnie z definicją możemy zapisać, że

$$\begin{aligned}
\mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|^2 &= \mathbb{E}_\theta \left[\sum_{i=1}^{\infty} (\lambda_i(X)x_i - \theta_i)^2 \right] = \mathbb{E}_\theta \left[\sum_{i=1}^{\infty} \lambda_i^2(X)x_i^2 + \sum_{i=1}^{\infty} \theta_i^2 - 2 \sum_{i=1}^{\infty} \lambda_i(X)x_i \theta_i \right] = \\
&= \mathbb{E}_\theta \left[\sum_{i=1}^{\infty} \lambda_i^2(X) (\theta_i^2 + \epsilon^2 \sigma_i^2 \xi_i^2 + 2\epsilon \theta_i \sigma_i \xi_i) + \sum_{i=1}^{\infty} \theta_i^2 - 2 \sum_{i=1}^{\infty} \lambda_i(X) \theta_i (\theta_i + \epsilon \sigma_i \xi_i) \right] = \\
&= \mathbb{E}_\theta \left[\sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 - 2\epsilon \sum_{i=1}^{\infty} (1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i \xi_i + \epsilon^2 \sum_{i=1}^{\infty} \lambda_i^2(X) \sigma_i^2 (\xi_i^2 - 1) + \epsilon^2 \sum_{i=1}^{\infty} \lambda_i^2(X) \sigma_i^2 \right] = \\
&= \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) - 2\epsilon \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i \xi_i + \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} \lambda_i^2(X) \sigma_i^2 (\xi_i^2 - 1).
\end{aligned}$$

Następnie korzystając z pierwszego z lematów z $K = S$ oszacujemy wyrażenie $\epsilon \mathbb{E}_\theta \left| \sum_{i=1}^{\infty} (1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i \xi_i \right|$. Ciągami v_i jest tym razem ciąg $(1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i$.

$$\begin{aligned}
& \epsilon \mathbb{E}_\theta \left| \sum_{i=1}^{\infty} (1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i \xi_i \right| \leq \\
& \leq \epsilon \sqrt{2 \ln(NS)} \mathbb{E}_\theta \left(\sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 \lambda_i^2(X) \sigma_i^2 \right)^{1/2} +
\end{aligned}$$

$$\begin{aligned}
& +2\epsilon\sqrt{\ln(NS)}S^{-1/2}\left(\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2\lambda_i^2(X)\sigma_i^2\right)^{1/2} \leq \\
& \leq \epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\sup_i\sigma_i|\lambda_i(X)|\left(\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2\right)^{1/2} + \\
& +2\epsilon\sqrt{\ln(NS)}S^{-1/2}\max_{\lambda\in\Lambda}\sup_i\sigma_i|\lambda_i|\left(\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2\right)^{1/2}.
\end{aligned}$$

W następnym kroku skorzystamy z nierówności $2ab \leq Ba^2 + B^{-1}b^2$ spełnionej dla dowolnego $B > 0$.

$$\begin{aligned}
& \epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left(\sup_i\sigma_i|\lambda_i(X)|\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2\right)^{1/2} + \\
& +2\epsilon\sqrt{\ln(NS)}S^{-1/2}\max_{\lambda\in\Lambda}\sup_i\sigma_i|\lambda_i|\left(\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2\right)^{1/2} \leq \\
& \leq \epsilon^2B\ln(NS)\mathbb{E}_\theta\sup_i\sigma_i^2\lambda_i^2(X) + B^{-1}\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2 + \\
& +\epsilon^2B\ln(NS)\frac{\max_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}{S} + B^{-1}\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2.
\end{aligned}$$

Zauważmy, że skoro $S = \frac{\max_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}{\min_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}$ zatem $\frac{\max_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}{S} = \min_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2$, a stąd powyższe wyrażenie redukuje się do postaci

$$\begin{aligned}
& 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2 + \epsilon^2B\ln(NS)\mathbb{E}_\theta\sup_i\sigma_i^2\lambda_i^2(X) + \epsilon^2B\ln(NS)\min_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2 \leq \\
& \leq 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2 + B\mathbb{E}_\theta\left(\epsilon^2\ln(NS)\sup_i\sigma_i^2\lambda_i^2(X)\right) + B\mathbb{E}_\theta\left(\epsilon^2\ln(NS)\sup_i\sigma_i^2\lambda_i^2(X)\right) \leq \\
& \leq 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^\infty(1-\lambda_i(X))^2\theta_i^2 + 2B\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)].
\end{aligned}$$

Następnie będziemy szacować wyrażenie $\epsilon^2\mathbb{E}_\theta\left|\sum_{i=1}^\infty\lambda_i^2(X)\sigma_i^2(\xi_i^2-1)\right|$ korzystając z drugiego z lematów z $K = S$, $q = 1$ i ciągu $v_i = \lambda_i^2(X)\sigma_i^2$. Ponadto zachodzi, że

$$m(v) = \sup_i \frac{|v_i|}{\|v\|} = \sup_i \lambda_i^2(X)\sigma_i^2 \left(\sum_{k=1}^\infty \lambda_k^4(X)\sigma_k^4\right)^{-1/2} \leq \rho(\lambda),$$

a stąd $m_V \leq \rho$ oraz $M(1) \leq M$ i możemy szacować wyrażenie $\sqrt{\ln(NS)} + m_V \ln(M(1)S)$ przez $\sqrt{2L_\Lambda}$. Możemy zatem zapisać, że

$$\begin{aligned} & \epsilon^2 \mathbb{E}_\theta \left| \sum_{i=1}^{\infty} \lambda_i^2(X) \sigma_i^2 (\xi_i^2 - 1) \right| \leq \\ & \leq \epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(M(1)S) \right) \left(\mathbb{E}_\theta \left(\sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2} + \left(\frac{1}{S} \mathbb{E}_\theta \sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2} \right) \leq \\ & \leq \epsilon^2 D \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2} + \frac{\epsilon^2 D \sqrt{2L_\Lambda}}{\sqrt{S}} \left(\mathbb{E}_\theta \sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2}. \end{aligned}$$

Zauważmy, że zachodzi następujący związek

$$\begin{aligned} S^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^4 \lambda_i^4(X) & \leq \frac{\min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2}{\max_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2} \mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^2(X) \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) \\ & \leq \frac{\min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2}{\max_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2} \max_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) \leq \\ & \leq \mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^2(X) \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X). \end{aligned}$$

Korzystając z tego faktu dostajemy, że

$$\begin{aligned} & \epsilon^2 D \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2} + \frac{\epsilon^2 D \sqrt{2L_\Lambda}}{\sqrt{S}} \left(\mathbb{E}_\theta \sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2} \leq \\ & \leq \epsilon^2 D \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\sum_{i=1}^{\infty} \lambda_i^4(X) \sigma_i^4 \right)^{1/2} + \epsilon^2 D \sqrt{2L_\Lambda} \left(\mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^2(X) \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) \right)^{1/2} \leq \\ & \leq \mathbb{E}_\theta \frac{\epsilon^2 B D^2 L_\Lambda}{4} \sup_i \sigma_i^2 \lambda_i^2(X) + 2\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) + \\ & + \mathbb{E}_\theta \frac{\epsilon^2 B D^2 L_\Lambda}{4} \sup_i \sigma_i^2 \lambda_i^2(X) + 2\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) = \\ & = 4\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) + \frac{\epsilon^2 B D^2}{2} \mathbb{E}_\theta L_\Lambda \sup_i \sigma_i^2 \lambda_i^2(X) \leq \\ & \leq 4\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) + \frac{B D^2}{2} \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)]. \end{aligned}$$

Łącząc te oszacowania dostajemy, że

$$\begin{aligned}
\mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|^2 &\leq \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) - 2\epsilon \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i \xi_i + \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} \lambda_i^2(X) \sigma_i^2 (\xi_i^2 - 1) \leq \\
&\leq \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + 4B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 + 2B \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)] + 4\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) + \frac{BD^2}{2} \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)] = \\
&= (1 + 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + CB \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)],
\end{aligned}$$

co kończy dowód lematu. \square

Lemat. Niech η będzie gaussowskim białym szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ i niech $v \in L_2(S, \mathcal{S}, \mu)$ będzie losowym elementem tej przestrzeni ze skończonego zbioru $V \subset L_2(S, \mathcal{S}, \mu)$ o liczności $N > 1$. Wtedy dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} |\langle \eta, v \rangle| \leq \sqrt{2 \ln(NK)} \left(\mathbb{E} \|v\|_2 + \sqrt{2 \mathbb{E} \|v\|_2^2 / K} \right).$$

Dowód. W dowodzie lematu ponownie skorzystamy z oszacowania

$$\mathbb{E} X^2 \mathbf{1}_{\{|X| > a\}} \leq \frac{2}{\sqrt{2\pi}} (a + 2a^{-1}) e^{-a^2/2} \quad \forall a > 0$$

zachodzącego dla zmiennych losowych o standardowym rozkładzie normalnym oraz z własności

$$F(\max_{v \in V} |\zeta_v|) \leq \sum_{v \in V} F(|\zeta_v|)$$

dla funkcji postaci $F(t) = t^2 \mathbf{1}_{\{t > \sqrt{2 \ln(NK)}\}}$.

Oznaczmy przez $\zeta_v = \frac{\langle \eta, v \rangle}{\|v\|_2}$. Zauważmy, że z faktu, że η jest gaussowskim białym szumem wynika, że ζ_v ma standardowy rozkład normalny $\mathcal{N}(0, 1)$. Możemy teraz napisać, że

$$\begin{aligned}
\mathbb{E} |\langle \eta, v \rangle| &= \mathbb{E} |\zeta| \|v\|_2 \leq \mathbb{E} \|v\|_2 \max_{v \in V} |\zeta_v| = \\
&\mathbb{E} \|v\|_2 \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| \leq \sqrt{2 \ln(NK)}\}} + \mathbb{E} \|v\|_2 \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \leq \\
&\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\|_2 + \mathbb{E} \|v\|_2 \max_{v \in V} |\zeta_v| \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \leq \\
&\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\|_2 + \left(\mathbb{E} \|v\|_2^2 \right)^{1/2} \left(\mathbb{E} \max_{v \in V} |\zeta_v|^2 \mathbf{1}_{\{\max_{v \in V} |\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} \cdot \\
&\leq \sqrt{2 \ln(NK)} \mathbb{E} \|v\|_2 + \left(\mathbb{E} \|v\|_2^2 \right)^{1/2} \left(\sum_{v \in V} \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} =
\end{aligned}$$

$$= \sqrt{2 \ln(NK)} \mathbb{E} \|v\|_2 + \left(\mathbb{E} \|v\|_2^2 \right)^{1/2} \left(N \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2}.$$

Korzystając ze wspomnianego na początku oszacowania z $a = \sqrt{2 \ln(NK)}$ dostajemy, że

$$\begin{aligned} & \sqrt{2 \ln(NK)} \mathbb{E} \|v\|_2 + \left(\mathbb{E} \|v\|_2^2 \right)^{1/2} \left(N \mathbb{E} |\zeta_v|^2 \mathbf{1}_{\{|\zeta_v| > \sqrt{2 \ln(NK)}\}} \right)^{1/2} \leq \\ & \leq \sqrt{2 \ln(NK)} \left(\mathbb{E} \|v\|_2 + \sqrt{2 \mathbb{E} \|v\|_2^2 / K} \right) \end{aligned}$$

co kończy dowód. \square

Lemat. Niech η będzie gaussowskim białym szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ i niech $v \in L_2(S, \mathcal{S}, \mu)$ będzie losowym elementem tej przestrzeni ze skończonego zbioru $V \subset L_2(S, \mathcal{S}, \mu)$ o liczności $N > 1$. Niech ponadto $v \neq 0$ dla dowolnego $v \in V$. Oznaczmy przez $m(v) = \|v\|_\infty / \|v\|_2$, $m_V = \max_{v \in V} m(v)$ oraz

$$M(q) = \sum_{v \in V} \exp(-q/m_V), \quad q > 0.$$

Wtedy istnieje stała D zależna tylko od q , taka, że dla dowolnego $K \geq 1$ zachodzi

$$\mathbb{E} |\langle \eta^2 - 1, v \rangle| \leq D \left(\sqrt{\ln(NK)} + m_V \ln(M(q)K) \right) \left(\mathbb{E} \|v\|_2 + \sqrt{\mathbb{E} \|v\|_2^2 / K} \right).$$

Lemat. Niech $\hat{\theta} = \hat{\lambda}(X)X$ będzie liniowym estymatorem z wagą z wartościami z przedziału $[-1, 1]$ przyjmującym wartości w zbiorze Λ . Oznaczmy przez

$$\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \|\sigma^2 \lambda^2\|_\infty, \quad \lambda \in \Lambda.$$

Wtedy istnieje absolutna stała $C > 0$ taka, że dla dowolnego $B > 0$ zachodzi

$$\mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|_2^2 \leq (1 + 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + CB \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].$$

Dowód. Przypomnijmy, że w naszym modelu (14) mamy, że $X = \theta + \epsilon \sigma \eta$, a stąd $|X|^2 = |\theta|^2 + \epsilon^2 \sigma^2 \eta^2 + 2\epsilon |\theta| \sigma \eta$.

Zgodnie z definicją możemy zapisać, że

$$\begin{aligned} \mathbb{E}_\theta \left\| \hat{\theta} - \theta \right\|_2^2 &= \mathbb{E}_\theta \left[\int_S (\lambda(X)X - \theta)^2 d\mu \right] = \\ &= \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) - 2\epsilon \mathbb{E}_\theta \int_S (1 - \lambda(X)) |\theta| \lambda(X) \sigma \eta d\mu + \epsilon^2 \mathbb{E}_\theta \int_S \lambda^2(X) \sigma^2 (\eta^2 - 1) d\mu. \end{aligned}$$

Następnie korzystając z pierwszego z lematów z $K = S$ oszacujemy wyrażenie $\epsilon \mathbb{E}_\theta \left| \int_S (1 - \lambda(X)) |\theta| \lambda(X) \sigma \eta d\mu \right|$.

$$\epsilon \mathbb{E}_\theta \left| \int_S (1 - \lambda(X)) |\theta| \lambda(X) \sigma \eta d\mu \right| \leq$$

$$\begin{aligned}
&\leq \epsilon \sqrt{2 \ln(NS)} \mathbb{E}_\theta \left(\int_S (1 - \lambda(X))^2 |\theta|^2 \lambda^2(X) \sigma^2 d\mu \right)^{1/2} + \\
&+ 2\epsilon \sqrt{\ln(NS)} S^{-1/2} \left(\mathbb{E}_\theta \int_S (1 - \lambda(X))^2 |\theta|^2 \lambda^2(X) \sigma^2 d\mu \right)^{1/2} \leq \\
&\leq \epsilon \sqrt{2 \ln(NS)} \mathbb{E}_\theta \left\| \sigma^2 \lambda^2(X) \right\|_\infty \left(\int_S (1 - \lambda(X))^2 |\theta|^2 d\mu \right)^{1/2} + \\
&+ 2\epsilon \sqrt{\ln(NS)} S^{-1/2} \max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty \left(\mathbb{E}_\theta \int_S (1 - \lambda(X))^2 |\theta|^2 d\mu \right)^{1/2}. \\
&\leq \epsilon^2 B \ln(NS) \mathbb{E}_\theta \left\| \sigma^2 \lambda^2(X) \right\|_\infty + B^{-1} \mathbb{E}_\theta \int_S (1 - \lambda(X))^2 |\theta|^2 d\mu + \\
&+ \epsilon^2 B \ln(NS) \frac{\max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty}{S} + B^{-1} \mathbb{E}_\theta \int_S (1 - \lambda(X))^2 |\theta|^2 d\mu.
\end{aligned}$$

Ponownie możemy zauważyć, że skoro $S = \frac{\max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty}{\min_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty}$ zatem $\frac{\max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty}{S} = \min_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty$, a stąd powyższe wyrażenie redukuje się do postaci

$$\begin{aligned}
&2B^{-1} \mathbb{E}_\theta \int_S (1 - \lambda(X))^2 |\theta|^2 d\mu + \epsilon^2 B \ln(NS) \mathbb{E}_\theta \left\| \sigma^2 \lambda^2(X) \right\|_\infty + \epsilon^2 B \ln(NS) \min_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty \leq \\
&\leq 2B^{-1} \mathbb{E}_\theta \int_S (1 - \lambda_i(X))^2 |\theta|^2 d\mu + 2B \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].
\end{aligned}$$

Następnie będziemy szacować drugie z wyrażeń, czyli $\epsilon^2 \mathbb{E}_\theta \left| \int_S \lambda^2(X) \sigma^2 (\eta^2 - 1) d\mu \right|$ korzystając z lematu 6. Zauważmy na początek, że

$$m(v) = \frac{\|v\|_\infty}{\|v\|_2} \leq \rho(\lambda),$$

a stąd $m_V \leq \rho$ oraz $M(1) \leq M$ i możemy szacować wyrażenie $\sqrt{\ln(NS)} + m_V \ln(M(1)S)$ przez $\sqrt{2L_\Lambda}$. Zachodzi zatem

$$\begin{aligned}
&\epsilon^2 \mathbb{E}_\theta \left| \int_S \lambda^2(X) \sigma^2 (\eta^2 - 1) d\mu \right| \leq \\
&\leq \epsilon^2 D \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^4(X) \sigma^4 d\mu \right)^{1/2} + \frac{\epsilon^2 D \sqrt{2L_\Lambda}}{\sqrt{S}} \left(\mathbb{E}_\theta \int_S \lambda^4(X) \sigma^4 d\mu \right)^{1/2}.
\end{aligned}$$

Analogicznie do poprzednich rozważań zachodzi

$$\begin{aligned}
&S^{-1} \mathbb{E}_\theta \int_S \sigma^4 \lambda^4(X) d\mu \leq \\
&\leq \mathbb{E}_\theta \left\| \sigma^2 \lambda^2(X) \right\|_\infty \mathbb{E}_\theta \int_S \sigma^2 \lambda^2(X) d\mu.
\end{aligned}$$

Dostajemy stąd następujące oszacowanie

$$\begin{aligned}
& \epsilon^2 D \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^4(X) \sigma^4 d\mu \right)^{1/2} + \frac{\epsilon^2 D \sqrt{2L_\Lambda}}{\sqrt{S}} \left(\mathbb{E}_\theta \int_S \lambda^4(X) \sigma^4 d\mu \right)^{1/2} \leq \\
& \leq \epsilon^2 D \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^4(X) \sigma^4 d\mu \right)^{1/2} + \epsilon^2 D \sqrt{2L_\Lambda} \left(\mathbb{E}_\theta \|\sigma^2 \lambda^2(X)\|_\infty \mathbb{E}_\theta \int_S \sigma^2 \lambda^2(X) d\mu \right)^{1/2} \leq \\
& \leq \mathbb{E}_\theta \frac{\epsilon^2 B D^2 L_\Lambda}{4} \|\sigma^2 \lambda^2(X)\|_\infty + 2\epsilon^2 B^{-1} \mathbb{E}_\theta \int_S \sigma^2 \lambda^2(X) d\mu + \\
& + \mathbb{E}_\theta \frac{\epsilon^2 B D^2 L_\Lambda}{4} \|\sigma^2 \lambda^2(X)\|_\infty + 2\epsilon^2 B^{-1} \mathbb{E}_\theta \int_S \sigma^2 \lambda^2(X) d\mu = \\
& = 4\epsilon^2 B^{-1} \mathbb{E}_\theta \int_S \sigma^2 \lambda^2(X) d\mu + \frac{\epsilon^2 B D^2}{2} \mathbb{E}_\theta L_\Lambda \|\sigma^2 \lambda^2(X)\|_\infty \leq \\
& \leq 4\epsilon^2 B^{-1} \mathbb{E}_\theta \int_S \sigma^2 \lambda^2(X) d\mu + \frac{B D^2}{2} \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].
\end{aligned}$$

Łącząc te oszacowania dostajemy tezę lematu. \square

Literatura

- [1] P. Alquier, E. Gautier, G. Stoltz, *Inverse Problems and High-Dimensional Estimation* Springer-Verlag, 2011, wydanie zbiorowe,
- [2] L. Birge, *Model selection via testing: an alternative to (penalized) maximum likelihood estimators*, Ann. I. H. Poincaré, PR 42, 2006, pp. 273–325,
- [3] L. Birge, *Statistical estimation with model selection*, arXiv, 2006, The Brouwer Lecture, 2005,
- [4] L. Cavalier, *Inverse problems with non-compact operators*, Journal of Statistical Planning and Inference, 136, 2006, pp. 390–400,
- [5] L. Cavalier, G. K. Golubev, D. Picard, A.B. Tsybakov, *Oracle inequalities for inverse problems*, The Annals of Statistics, Vol. 30, No. 3, 2002, pp. 843–874,
- [6] P. R. Halmos, *What Does the Spectral Theorem Say?*, The American Mathematical Monthly, Vol. 70, No. 3, 1963, pp. 241–247,
- [7] T. Hida, *Brownian Motion*, Springer, 1980,
- [8] J. Kaipo, E. Somersalo, *Statistical and computational inverse problems*, Springer, 2004,
- [9] H. Lal Vasudeva, *Elements of Hilbert Spaces and Operator Theory*, Springer-Verlag, 2017,

- [10] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Springer-Science+Business Media, B.Y., 1986,
- [11] Z. Szkutnik, *Statystyczne problemy odwrotne*, notatki do wykładu,
- [12] L. Wasserman, *All of Nonparametric Statistics*, Springer Science+Business Media, Inc., 2006,