Grzegorz Mika

Nierówności wyrocznie dla problemów odwrotnych

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 odpowiedniego wyrażenia związanego z estymatorem 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 wyroczni w obu przypadkach.

Spis treści

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 $g \in G$, spełnić równanie

$$Af = g.$$

Definicja 1. *([21], str. 1)*
Problem nazwiemy dobrze postawionym 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 powyższej definicji nie jest spełniony, problem nazywamy źle postawionym. 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 przypuśćmy, że dysponujemy następującym modelem

$$Y = Af + \epsilon\xi, \tag{1}$$

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

Definicja 2. *([1], str. 7)*
Stochastycznym błędem ξ nazwiemy proces na przestrzeni Hilberta G, czyli ograniczony liniowy operator $\xi\colon G \to L^2(\Omega, \mathcal{F}, \mathbb{P})$ 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 operator kowariancji Cov_ξ określony jako ograniczony liniowy operator ($||Cov_\xi|| \leqslant 1$) z przestrzeni G w przestrzeń G taki, że $\langle Cov_\xi g_1, g_2 \rangle = Cov(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle)$. Przestrzeń $(\Omega, \mathcal{F}, \mathbb{P})$ jest podstawową przestrzenią probabilistyczną, natomiast $\mathcal{L}^2(\cdot)$ jest przestrzenią wszystkich funkcji całkowalnych z kwadratem na zadanej przestrzeni z miarą.

Definicja 3. *([1], str. 8)*
Powiemy, że losowy błąd ξ jest gaussowskim białym szumem, jeśli $Cov_\xi = I$ oraz indukowane zmienne losowe są gaussowskie, czyli dla dowolnych elementów $g_1, g_2, \ldots, g_k \in G$ i dla dowolnego $k \in \mathbb{N}$ mamy, że $(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle, \ldots, \langle \xi, g_k \rangle) \sim \mathcal{N}_k(\mathbf{0}, \mathbf{\Sigma})$. oraz $Cov(\langle \xi, g_i \rangle, \langle \xi, g_j \rangle) = \langle g_i, g_j \rangle$.

Lemat 1. *([1], str. 8)*
Niech ξ będzie białym szumem w przestrzeni G oraz niech $\{u_i\}_{i \in I}$ będzie ortonormalną bazą tej przestrzeni i niech $\xi_k = \langle \xi, u_k \rangle$. Wtedy $\{\xi_i\}_{i \in I}$ 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. □

Zauważmy, że gdy ξ jest białym szumem, to 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 ograniczonych operatorów liniowych między dwoma przestrzeniami Hilberta H, G. Założymy, że $D(A) = \{f \in H \colon \exists_{g \in G}\ Af = g\} = H$.

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 $A\colon H \to H$ jest nieujemny, gdy $\forall_{f \in H}\ \langle Af, f \rangle \geqslant 0$ oraz dodatni, gdy $\forall_{f \in H \setminus \{0\}}\ \langle Af, f \rangle > 0$.

Estymatorem liniowym elementu f w modelu (1) nazywamy estymator postaci $T(Y)$, gdzie T jest pewnym operatorem z przestrzeni $L(G, H)$, a działanie operatora T na proces Y zdefiniowane jest warunkiem $\langle T(Y), h \rangle = \langle Y, T^* h \rangle$ dla dowolnego elementu $h \in H$.

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

Twierdzenie 1. *([1], str. 9, [18], str. 10)*
Niech $A \in L(H, G)$. *Wtedy*
— $KerA = (RangeA^*)^\perp$ *oraz* $\overline{RangeA} = (KerA^*)^\perp$,
— *jeśli* A *jest iniektywny, to* A^*A *też,*
— $A^*A \in L(H)$ *oraz* A^*A *jest dodatni i samosprzężony.*

Dowód. Zauważmy, że $RangeA^\perp = \{g \in G \colon \langle Af, g\rangle = 0 \ \forall f \in H\}$. Wtedy dla dowolnych $f \in KerA$ i $g \in G$ mamy, że $0 = \langle Af, g\rangle = \langle f, A^*g\rangle$, a stąd $KerA = (RangeA^*)^\perp$. Zamieniając A z A^* otrzymujemy, że $KerA^* = RangeA^\perp$, czyli $(KerA^*)^\perp = (RangeA^\perp)^\perp = \overline{RangeA}$, gdyż dla dowolnej przestrzeni Hilberta H i dla dowolnej podprzestrzeni A tej przestrzeni zachodzi, że $(A^\perp)^\perp = \overline{A}$, ponieważ dopełnienie ortogonalne $A^\perp$ jest zbiorem domkniętym, co pokażemy korzystając z ciągłości iloczynu skalarnego. Niech $\{x_n\}$ będzie ciągiem Cauchy'ego elementów zbioru $A^\perp$ o granicy $x \in H$. Wtedy dla dowolnego elementu $y \in A$ zachodzi $\langle x, y\rangle = \langle x - x_n, y\rangle + \langle x_n, y\rangle = \langle x - x_n, y\rangle + 0 \to 0$, gdy $n \to \infty$ co dowodzi, że $x \in A^\perp$.

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

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

Wniosek 1. — $H = KerA \oplus KerA^\perp = KerA \oplus \overline{RangeA^*}$,
— $G = \overline{RangeA} \oplus RangeA^\perp = \overline{RangeA} \oplus KerA^*$.

W pierwszej części pracy ograniczymy się do rozważania tylko zwartych operatorów liniowych, jednak dzięki temu uda się uzyskać dającą szerokie możliwości reprezentację według wartości singularnych. 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 4. *([24], str. 264)*
Operator $A \colon H \to G$ *nazywamy zwartym, 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 $\int_a^b \int_a^b |K(x, y)|^2 dxdy < +\infty$ 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 zwartości 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* $dimH = \infty$ *to operator* A^{-1} *jest nieograniczony, o ile istnieje.*

Fakt ten powoduje, że dla dowolnego zwartego operatora na 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 samosprzężonych potrzebne do wykazania istnienia reprezentacji według wartości singularnych dla operatorów zwartych, ale już niekoniecznie samosprzężonych.

Twierdzenie 2 (Reprezentacja spektralna)**.** *Niech* A *będzie samosprzężonym operatorem zwartym na przestrzeni Hilberta* H*. Wtedy istnieje zupełny układ ortonormalnych wektorów własnych* $E = \{f_j, j \in I\} \subset H$*. Niech* $J = \{j \in$

$I\colon \lambda_j \neq 0\}$ 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 $J_\delta = \{j \in I\colon |\lambda_j| \geqslant \delta\}$ jest skończony, czyli jedynym możliwym punktem skupienia zbioru wartości własnych jest zero.

Dowód. Dowód w [1], str. 11. □

Możemy teraz wprowadzić reprezentację według wartości singularnych dla operatora zwartego.

Twierdzenie 3 (Reprezentacja według wartości singularnych)**.** *Niech $A\colon H \to G$ będzie operatorem zwartym między przestrzeniami Hilberta H i 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*
— $KerA^{\perp} = \overline{span\{v_n,\ n \in I\}}$,
— $\overline{RangeA} = \overline{span\{u_n,\ n \in I\}}$,
— $Af = \sum_n b_n\langle f, v_n\rangle u_n$ *oraz* $A^*g = \sum_n b_n\langle g, u_n\rangle v_n$.
Ponadto $g \in RangeA$ wtedy i tylko wtedy, gdy $g = \sum_n \langle g, u_n\rangle u_n$ oraz spełniony jest tzw. warunek Picarda

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

Wtedy rozwiązania równania $Af = g$ mają postać

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

przy czym $f_0 \in KerA$ 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 singularnym (*singular value decomposition– SVD*) operatora A.

Dowód. ([18], str. 11, [21], str. 10)
Dowód twierdzenia opiera się na wykorzystaniu twierdzenia spektralnego do operatora A^*A.

Operator A^*A jest samosprzężony, zwarty i nieujemny, a zatem istnieją liczby $b_1^2 \geqslant b_2^2 \geqslant \cdots \geqslant 0$ oraz wektory ortonormalne v_n takie, że $A^*Av_n = b_n^2 v_n$. Niech $I = \{n\colon b_n > 0\}$ 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 $KerA^{\perp} = (KerA^*A)^{\perp} = \overline{RangeA^*A} = \overline{span\{v_n,\ n \in I\}}$.

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

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 $f_0 \in KerA$ jest rozwiązaniem, gdyż na mocy powyższych faktów mamy, że $Af = A(f_0 + \sum_n b_n^{-1}\langle g, u_n\rangle v_n) = \sum_n\langle g, u_n\rangle b_n^{-1}Av_n = g$. Z drugiej strony, jeżeli $Af =$

g, to mamy, że $g = \sum_n \langle g, u_n \rangle u_n$ oraz $Af = \sum_n b_n \langle f, v_n \rangle u_n$ zatem musi zachodzić, że $b_n \langle f, v_n \rangle = \langle g, u_n \rangle$ dla dowolnego $n \in \mathbb{N}$, czyli $f = \sum_n b_n^{-1} \langle g, u_n \rangle + f_0$, gdzie $f_0 \in KerA$. □

Udało nam się zaprezentować działanie zwartego operatora w postaci jego rozwinięcia według wartości singularnym 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, gdy g jest znane tylko w przybliżeniu. 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 $P \colon G \to \overline{RangeA}$ 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 $||Af - g||^2 = ||Af - Pg||^2 + ||(1-P)g||^2 \geqslant ||(1-P)g||^2$.

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

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 ciągowego.

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

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

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

gdzie $\theta_n = \langle f, v_n \rangle$ są współczynnikami w rozwinięciu Fouriera funkcji f w bazie $\{v_n\}$,a $\xi_n = \langle \xi, u_n \rangle$ są zgodnie z lematem 1 niezależnymi zmiennymi losowymi o standardowym rozkładzie normalnym.

Oznaczając $y_n = \langle Y, u_n \rangle$ możemy wyjściowy problem $Y = Af + \epsilon\xi$ zapisać w równoważnej postaci modelu ciągowego jako

$$y_n = b_n \theta_n + \epsilon \xi_n,\ n = 1, 2, \ldots.$$

W tej postaci widać dokładnie trudności związane ze stochastycznymi problemami odwrotnymi. Jako że b_n są wartościami osobliwymi operatora zwartego mamy, że $b_n \to 0$ gdy $n \to \infty$, 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

$$x_n = \theta_n + \epsilon \sigma_n \xi_n,\ n = 1, 2, \ldots \tag{2}$$

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

Korzystając z powyższego modelu wprowadzimy notację i pojęcia związane z konstrukcją rozważanych estymatorów i ich własnościami.
Mając pełną i niezaburzoną informację o współczynnikach $\theta_n,\ n = 1, 2, \ldots$ można by uzyskać pełną informację o poszukiwanym elemencie f z dokładnością do składowej znajdującej się w dopełnieniu ortogonalnym jądra 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, aby uwzględnić różne poziomy szumu.

Definicja 5. *Niech* $\lambda = (\lambda_1, \lambda_2, \dots)$ *będzie nielosowym ciągiem liczbowym. Estymatorem liniowym współczynników* θ_n *w modelu (2) nazwiemy estymator* $\hat{\theta}(\lambda) = (\hat{\theta}_1, \hat{\theta}_2, \dots)$*, gdzie*

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

Ciąg λ *nazywać będziemy filtrem lub wagami.*

Powyższa definicja jest przeformułowaniem ogólnej definicji estymatora liniowego w modelu obserwacji w białym szumie. 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 $\lambda_i = \mathbf{1}_{\{i \leqslant N\}}$, gdzie $\mathbf{1}_A$ oznacza indykator 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.

Definicja 6. *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 druga 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 $X = (x_1, x_2, \dots)$, w zależności od przyjętego modelu obserwacji.
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

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

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

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

Definicja 7. *([1], str. 28)*
Ryzykiem minimaksowym 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 minimaksowe jest zwykle niemożliwe, dlatego poszukiwać będziemy estymatora asymptotycznie minimaksowego.

Definicja 8. *([1], str. 28)*
Przypuśćmy, że istnieje estymator $\tilde{f}$, taki, że istnieją stałe $0 < C_1 \leqslant C_2 < \infty$ takie, że gdy $\epsilon \to 0$, to zachodzi

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

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

gdzie ciąg nieujemnych wartości a_ϵ jest taki, że $a_\epsilon \to 0$, gdy $\epsilon \to 0$. 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 minimaksowy.

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 9. *Powiemy, że dla elementu f zachodzi warunek źródłowy, jeżeli istnieje $w \in H$, $L > 0$ oraz $\mu \geqslant 0$ takie, że*

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

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

$$H_{\mu,L} = \left\{ f \in H \colon f = (A^*A)^\mu w,\ w \in H,\ ||w||^2 \leqslant L \right\}.$$

Mając do dyspozycji reprezentację spektralną operatora A^*A oraz tzw. rachunek funkcyjny, pozwalający zdefiniować funkcję od operatora poprzez jego reprezentację spektralną ([24]) dostaniemy równoważną postać warunku źródłowego. Zapiszmy

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

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

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

Zatem założenie o warunku źródłowym jest równoważne założeniu, że współczynniki rozwinięcia Fouriera funkcji f w odpowiedniej bazie wektorów singularnych należą do pewnej elispoidy w przestrzeni l^2. Oznaczmy taką elipsoidę przez

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

gdzie $L > 0$ oraz ciąg $a = \{a_k\}$ jest ciągiem nieujemnych liczb rozbieżnym do nieskończoności.
Wprowadzimy teraz pojęcie klasy Sobolewa funkcji.

Definicja 10. *Niech H będzie pewną przestrzenią Hilberta a $\{\phi_i\}$ układem ortonormalnym w tej przestrzeni. Klasą Sobolewa nazywamy klasę postaci*

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

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

$$\mathcal{W}(\alpha, L) = \left\{ f \in H\colon \int_0^1 \left(f^{(\alpha)}(t)\right)^2 dt \leqslant L,\ f^{(j)}(0) = f^{(j)}(1) = 0,\ j = 0, 1, \ldots, \alpha - 1 \right\},$$

gdzie $f^{(\alpha)}$ jest rozumiane jako pochodna słaba funkcji f rzędu α.
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

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

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

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

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

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

gdzie $\Theta_{\mathcal{A}}(\alpha, L)$, to elipsoida $\Theta(a, L)$ z ciągiem a zdefiniowanym 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

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

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 $\lambda X = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$.
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 $\Lambda = \{\lambda X\colon \lambda \in [0, 1]\}$ 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 wyrocznię, 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 wyroczni w kontekście osiąganego ryzyka. Tak postawione zagadnienie rozwiązywane jest przez znalezienie estymatora θ^*, którego ryzyko daje się kontrolować przez nieasymptotyczne nierówności wyrocznie postaci

$$\mathcal{R}(\theta^*, \theta) \leqslant C_1 \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + C_2(\Lambda, n), \tag{4}$$

gdzie stałe $C_1, C_2(\Lambda, n)$ nie zależą od estymowanego elementu, jednak stała $C_2(\Lambda, n)$ może zależeć od liczności rodziny Λ bądź jej złożoności, gdy jest ona nieskończona oraz od wielkości próby n ([19], str. 177) .

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

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

gdy $\sigma \to 0$.

W przypadku rozważanego na początku rozdziału problemu poszukiwanym estymatorem naśladującym 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

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

Szczegółowe uzasadnienie można znaleźć w [25], str. 156.

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 Pinskera.

4. Główne rezultaty I

Rezultaty uzyskane w tym rozdziale są rezultatami uzyskanymi w pracy [8] z kilkoma modyfikacjami.

Przypuśćmy, że dysponujemy skończonym zbiorem filtrów $\Lambda = (\lambda^1, \ldots, \lambda^N)$ i związanych z nimi estymatorów liniowych, gdzie $\lambda^i = (\lambda_1^i, \lambda_2^i, \ldots)$, $i = 1, 2, \ldots, N$. Naszym celem będzie konstrukcja filtra opartego na obserwacjach $\lambda^*(X) = (\lambda_1^*, \lambda_2^*, \ldots)$ o wartościach w zbiorze Λ o asymptotycznie minimalnym ryzyku przy prawdziwej wartości θ, który równocześnie naśladuje ryzyko najlepszego estymatora w tej klasie w każdej skończonej próbie. Okaże się, że filtr ten może zostać zdefiniowany jako element minimalizujący względem $\lambda \in \Lambda$ nieobciążony estymator ryzyka. 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 minimalizacja ryzyka estymatora liniowego miała sens, trzeba założyć, że ryzyko jest skończone, a zatem należy dobierać filtry tak, by drugi człon był skończony. Ponadto założenie o dodatniości poniższej sumy implikuje, że żaden z rozważanych filtrów nie może być tożsamościowo równy zeru.

Założenie 1.

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

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.

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

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 [20]).

Ponadto będziemy zakładać skończoność poniższej sumy, by występujące później wyrażenia były skończone.

Założenie 3.

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

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 pełna 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

$$\sum_{n=1}^{\infty}(1-\lambda_n)^2(x_n^2 - \sigma_n^2\epsilon^2) + \epsilon^2\sum_{n=1}^{\infty}\sigma_n^2\lambda_n^2 =$$

$$= \sum_{n=1}^{\infty}(x_n^2 - \sigma_n^2\epsilon^2) + \sum_{n=1}^{\infty}(\lambda_n^2 - 2\lambda_n)(x_n^2 - \sigma_n^2\epsilon^2) + \epsilon^2\sum_{n=1}^{\infty}\sigma_n^2\lambda_n^2 =$$

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

Jest to nieobciążony estymator ryzyka. Naszym celem jest jednak minimalizacja tego estymatora ze względu na wagi λ_i, stąd wystarczy ograniczyć się do rozpatrywania wyrażenia

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

będącego nieobciążonym estymatorem $\mathcal{R}(\hat\theta, \theta) - \sum_{n=1}^{\infty}\theta_n^2$. Wyrażenie to jest skończone gdyż, dla dowolnego $\theta = \{\theta_i\}_{i=1}^{\infty}$ i λ spełniającego założenia 1–3 oraz ciągu zmiennych losowych określonych jako $\sqrt{\lambda_i^2 - 2\lambda_i}x_i$ o rozkładzie $\mathcal{N}(\sqrt{\lambda_i^2 - 2\lambda_i}\theta_i, (\lambda_i^2 - 2\lambda_i)\sigma_i^2\epsilon^2)$ mamy

$$\sum_{i=1}^{\infty}(\lambda_i^2 - 2\lambda_i)\theta_i^2 \leqslant 3\sum_{i=1}^{\infty}\theta_i^2 < \infty,$$

$$\sum_{i=1}^{\infty}(\lambda_i^2 - 2\lambda_i)\sigma_i^2\epsilon^2 = \epsilon^2\sum_{i=1}^{\infty}\lambda_i^2\sigma_i^2 - 2\epsilon^2\sum_{i=1}^{\infty}\lambda_i\sigma_i^2 < \infty,$$

co na mocy lematu 9 implikuje zbieżność $U(\lambda, X)$ w normie L_2.

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

$$\mathbb{V}ar U(\lambda, X) = 2\epsilon^4\sum_{n=1}^{\infty}\lambda_n^4\sigma_n^4 - \sum_{n=1}^{\infty}\lambda_n^4\theta_n^4 - 2\epsilon^2\sum_{n=1}^{\infty}\theta_n^2\sigma_n^2\lambda_n^4 +$$

$$+8\epsilon^4 \sum_{n=1}^{\infty} \sigma_n^4 \lambda_n^2 - 4 \sum_{n=1}^{\infty} \lambda_n^2 \theta_n^4 - 8\epsilon^2 \sum_{n=1}^{\infty} \lambda_n^2 \sigma_n^2 \theta_n^2 - 8\epsilon^4 \sum_{n=1}^{\infty} \lambda_n^3 \sigma_n^4 +$$

$$+4 \sum_{n=1}^{\infty} \lambda_n^3 \theta_n^4 + 8\epsilon^2 \sum_{n=1}^{\infty} \theta_n^2 \sigma_n^2 \lambda_n^3 \leqslant$$

$$\leqslant 2\epsilon^4 \sum_{n=1}^{\infty} \lambda_n^4 \sigma_n^4 + 8\epsilon^4 \sum_{n=1}^{\infty} \sigma_n^4 \lambda_n^2.$$

Wariancja ta jest skończona na mocy wcześniejszych założeń, gdyż zbieżność szeregu $\sum_{n=1}^{\infty} \lambda_n \sigma_n^2$ implikuje zbieżność szeregu $\sum_{n=1}^{\infty} \lambda_n^2 \sigma_n^4$.

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

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

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

Wprowadzimy następujące oznaczenia na występujące w późniejszych rozważaniach wielkości

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

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

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

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

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

Wielkość $\rho(\lambda)$ jest pewnym sposobem mierzenia wielkości poszczególnych filtrów biorącym pod uwagę zarówno tempo znikania dalekich wyrazów poprzez $\left[\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4\right]^{-1/2}$ 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 masywność tej rodziny (dalszy komentarz na temat znaczenia parametru M można znaleźć w pracy [4]). Zauważmy ponadto, że na mocy założeń 1 i 3 wszystkie te wartości są liczbami skończonymi.

Z założenia 2 wynika następująca nierówność

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

Dodatkowo wymagać będziemy istnienia następującej stałej

Założenie 4.

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

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 $\mathcal{R}(\hat{\theta}, \theta) \geqslant \epsilon^2 \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2$ oraz

$$\frac{\left(\epsilon^4 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4\right)^{1/2}}{\epsilon^2 \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leqslant \rho, \tag{8}$$

ponieważ z uwagi na założenie 2

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

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

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

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 $\sup_k \sigma_k^2 |\lambda_k| \leqslant \sqrt{\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^2}$ i założenie 4 mamy, że

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

Przed wypowiedzeniem głównego twierdzenia zauważmy jeszcze, że zawsze zachodzi związek

$$M \leqslant N,$$

gdzie N oznacza liczność rodziny Λ.
W poniższych dwóch twierdzeniach zebrane są główne wyniki otrzymane dla rozważanych problemów odwrotnych z operatorami zwartymi.

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

$$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant (1 + \gamma_1 B^{-1}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + \gamma_2 B \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda),$$

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*

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

Twierdzenie 5. *Niech będą spełnione założenia 1– 4. 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*

$$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant (1 + \gamma_3 \rho \sqrt{L_\Lambda}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

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 będą spełnione założenia 1– 4. Ponadto niech zachodzi* $\lim_{\epsilon \to 0} \rho^2 \ln(NS) = 0$*. Wtedy istnieją stałe* $\mathbb{C}_2 > 0, \mathbb{C}_3 > 0$ *zależące tylko od stałej* C_1*, takie że dla* $\rho^2 \ln(NS) < \mathbb{C}_2$ *i dla dowolnego* $\theta \in l^2$ *zachodzi*

$$\mathbb{E}_\theta \left\|\hat{\theta} - \theta\right\|^2 \leqslant \left(1 + \mathbb{C}_3 \rho \sqrt{ln(NS)}\right) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$$

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

Dowód. Skoro zachodzi $\lim_{\epsilon \to 0} \rho^2 \ln(NS) = 0$, 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 $L_\Lambda < C \ln(NS)$.

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

gdyż $M \leqslant N$. □

Zatem warunek $\lim_{\epsilon\to 0}\rho^2\ln(NS)=0$ 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 7.

Lemat 2. *Niech $\{\xi_i\}_{i=1}^{\infty}$ 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 mierzalnym względem tej samej przestrzeni co ciąg $\{\xi_i\}_{i=1}^{\infty}$ i takim, że przyjmuje on wartości w skończony zbiorze $V\subset l^2$ o liczności $N>1$. Wtedy dla dowolnego $K\geqslant 1$ zachodzi*

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

Lemat 3. *Niech $\{\xi_i\}_{i=1}^{\infty}$ 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 mierzalnym względem tej samej przestrzeni co ciąg $\{\xi_i\}_{i=1}^{\infty}$ i takim, że przyjmuje on wartości w skończonym zbiorze $V\subset l^2$ o liczności $N>1$. Niech ponadto $v\neq 0$ dla każdego $v\in V$. Oznaczmy przez $m(v)=\sup_i|v_i|/||v||$, $m_V=\max_{v\in V}m(v)$ oraz*

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

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

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

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

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

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

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

Różnica w stałej $1+4B^{-1}$ w porównaniu do pracy [8] jest wynikiem innego doboru stałych w dowodzie twierdzeń 4 i 5 z uwagi na niewłaściwe dobraną stałą w pracy [8]. Różnica ta nie ma wpływa na ostateczny wydźwięk tych twierdzeń i wnioski z nich.

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

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

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

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

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

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

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

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

$$+ \left[\sum_{i=1}^{\infty}(\lambda_i^{*2} - 2\lambda_i^*)\theta_i^2 + \epsilon^2\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2} + \sum_{i=1}^{\infty}\theta_i^2\right] - \sum_{i=1}^{\infty}\theta_i^2 =$$

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

Obliczając wartość oczekiwaną powyższego wyrażenia dostajemy

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

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

Znajdziemy teraz dolne oszacowania na dwa ostatnie składniki powyższego wyrażenia. Zauważmy, że

$$\epsilon\mathbb{E}_\theta\sum_{i=1}^{\infty}(1 - \lambda_i^*)^2\sigma_i\theta_i\xi_i = \epsilon\mathbb{E}_\theta\sum_{i=1}^{\infty}(1 - \lambda_i^*)^2\sigma_i\theta_i\xi_i - 0 =$$

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

$$= \epsilon\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2]\sigma_i\theta_i\xi_i \geqslant$$

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

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

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

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

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

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

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

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

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

$$2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda_i})^2](\lambda_i^{*2}+\tilde{\lambda_i^2})\theta_i^2\sigma_i^2\right)^{1/2} \leqslant$$

$$\leqslant 2\epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left\{\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda_i^2})\sigma_i^2\right\}\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda_i})^2]\theta_i^2\right)^{1/2}=$$

$$=\mathbb{E}_\theta 2\epsilon\sqrt{2\ln(NS)\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda_i^2})\sigma_i^2\right\}}\left(\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda_i})^2]\theta_i^2\right)^{1/2}\leqslant$$

$$\leqslant 2B\epsilon^2\ln(NS)\mathbb{E}_\theta\sup_i\left\{(\lambda_i^{*2}+\tilde{\lambda_i^2})\sigma_i^2\right\}+B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda_i})^2]\theta_i^2\leqslant$$

$$\leqslant 2B\epsilon^2\ln(NS)\mathbb{E}_\theta\left(\sup_i\lambda_i^{*2}\sigma_i^2+\sup_i\tilde{\lambda_i^2}\sigma_i^2\right)+B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda_i})^2]\theta_i^2.$$

Analogicznie postępując z drugim wyrażeniem dostajemy

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

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

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

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

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

$$\max_{\lambda\in\Lambda}\sup_i\left\{(\lambda_i^2+\tilde{\lambda_i^2})\sigma_i^2\right\}/S=\frac{\min_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}{\max_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}\max_{\lambda\in\Lambda}\sup_i\left\{(\lambda_i^2+\tilde{\lambda_i^2})\sigma_i^2\right\}\leqslant$$

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

$$=\min_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2+\frac{1}{S}\sup_i\tilde{\lambda_i^2}\sigma_i^2\leqslant\sup_i\lambda_i^{*2}\sigma_i^2+\sup_i\tilde{\lambda_i^2}\sigma_i^2,$$

a więc także przez wartość oczekiwaną ostatniego wyrażenia. Zatem dostajemy oszacowanie postaci

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

$$\leqslant 4B\epsilon^2\ln(NS)\mathbb{E}_\theta\left(\sup_i\lambda_i^{*2}\sigma_i^2+\sup_i\tilde{\lambda_i^2}\sigma_i^2\right)+B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}[(1-\lambda_i^*)^2+(1-\tilde{\lambda_i})^2]\theta_i^2,$$

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

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

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

Zachodzi również

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

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

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

$$\geqslant -6B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) - 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1-\lambda_i^*)^2 \theta_i^2 - 2B^{-1} \sum_{i=1}^{\infty} (1-\tilde{\lambda}_i)^2 \theta_i^2 \geqslant$$

$$\geqslant -2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1-\lambda_i^*)^2 \theta_i^2 - 2B^{-1} \mathcal{R}(\tilde{\theta}, \theta) - 6B \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 6B \mathbb{E}_\theta \Delta^\epsilon[\tilde{\lambda}]. \quad (10)$$

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 $|\lambda_i| \leqslant 1$ dla dowolnego $\lambda \in \Lambda$ z założenia 2, zachodzi $\lambda_i^2 \leqslant (\lambda_i^2 - 2\lambda_i)^2 \leqslant 9\lambda_i^2$. 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

$$\epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) \leqslant \epsilon^2 \mathbb{E}_\theta \left| \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) \right| \leqslant$$

$$\leqslant \epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta ||v|| + \sqrt{\mathbb{E}_\theta ||v||^2 / S} \right).$$

Oszacujemy teraz niektóre z elementów pojawiających się powyżej:

$$m(v) = \sup_i |v_i| / ||v|| = \frac{\sup_i |\lambda_i^{*2} - 2\lambda_i^*| \sigma_i^2}{\sqrt{\sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*)^2 \sigma_i^4}} \leqslant \frac{3 \sup_i |\lambda_i| \sigma_i^2}{\sqrt{\sum_{i=1}^{\infty} \lambda_i^{*4} \sigma_i^4}} \leqslant 3\rho(\lambda),$$

$$M(3) = \sum_v \exp(-3/m(v)) = \sum_{\lambda \in \Lambda} \exp(-3/m(v)) \leqslant \sum_{\lambda \in \Lambda} \exp(-1/\rho(\lambda)) = M,$$

$$m_V = \max_v m(v) \leqslant 3\rho = 3 \max_{\lambda \in \Lambda} \rho(\lambda).$$

Stąd dostajemy, że

$$\left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right)^2 \leqslant \left(\sqrt{\ln(NS)} + 3\rho \ln(MS) \right)^2 \leqslant$$

$$\leqslant 2 \ln(NS) + 6\rho^2 \ln^2(MS) \leqslant 6(\ln(NS) + \rho^2 \ln^2(MS)),$$

a stąd

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

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

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

$$\geqslant -\epsilon^2 C\sqrt{L_\Lambda}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}(\lambda_i^{*2}-2\lambda_i^*)^2\sigma_i^4\right)^{1/2} -\epsilon^2 C\sqrt{L_\Lambda}S^{-1/2}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}(\lambda_i^{*2}-2\lambda_i^*)^2\sigma_i^4\right)^{1/2} \geqslant$$

$$\geqslant -\epsilon^2 3C\sqrt{L_\Lambda}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}\lambda_i^{*2}\sigma_i^4\right)^{1/2} - \epsilon^2 3C\sqrt{L_\Lambda}S^{-1/2}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}\lambda_i^{*2}\sigma_i^4\right)^{1/2} \geqslant$$

$$\geqslant -\epsilon^2 C\sqrt{C_1}\sqrt{L_\Lambda}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}\lambda_i^{*4}\sigma_i^4\right)^{1/2} -\epsilon^2 C\sqrt{C_1}\sqrt{L_\Lambda}S^{-1/2}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}\lambda_i^{*4}\sigma_i^4\right)^{1/2},$$

z uwagi na to, że $(\lambda_i^{*2}-2\lambda_i^*)^2 \leqslant 9\lambda_i^{*2}$ oraz założenie 4.
W dalszej części skorzystamy z faktu, że $\forall_{\omega\in\Omega}\ \min_{\lambda\in\Lambda}\sup_i \lambda_i^2\sigma_i^2 \leqslant \sup_i \lambda_i^{*2}\sigma_i^2 = \sup_i \lambda_i^{*2}(\omega)\sigma_i^2$, a stąd $\min_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2 \leqslant \mathbb{E}_\theta \sup_i \lambda_i^{*2}\sigma_i^2$.
Analogicznie jak poprzednio

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

$$\leqslant \frac{\min_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2}{\max_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2}\max_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2} \leqslant$$

$$\leqslant \mathbb{E}_\theta\sup_i\sigma_i^2\lambda_i^{*2}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}.$$

Podobnie dla drugiego wyrażenia

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

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

$$C\sqrt{L_\Lambda}\mathbb{E}_\theta\left(\sup_i\sigma_i^2\lambda_i^{*2}\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}\right)^{1/2} = \mathbb{E}_\theta\left(2C\sqrt{L_\Lambda\sup_i\sigma_i^2\lambda_i^{*2}}\left(2\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}\right)^{1/2}\right) \leqslant$$

$$\leqslant B\mathbb{E}_\theta C^2 L_\Lambda\sup_i\sigma_i^2\lambda_i^{*2} + 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}.$$

Podobnie dla drugiego wyrażenia

$$C\sqrt{L_\Lambda}\left(\mathbb{E}_\theta\sup_i\sigma_i^2\lambda_i^{*2}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}\right)^{1/2} = 2\left(\mathbb{E}_\theta C L_\Lambda\sup_i\sigma_i^2\lambda_i^{*2}\right)^{1/2}\left(2\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}\right)^{1/2} \leqslant$$

$$\leqslant B\mathbb{E}_\theta C L_\Lambda\sup_i\sigma_i^2\lambda_i^{*2} + 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}.$$

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

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

$$\geqslant C\epsilon^2 B\mathbb{E}_\theta L_\Lambda\sup_i\sigma_i^2\lambda_i^{*2} - 4\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2} =$$

$$= -CB\mathbb{E}_\theta\Delta^\epsilon[\lambda^*] - 4\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^{*2}. \tag{11}$$

Łącząc wyrażenia (10) oraz (11) dostajemy

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

$$\geqslant -4B^{-1}\mathbb{E}_\theta\left[\sum_{i=1}^\infty (1-\lambda_i^*)^2\theta_i^2 + \epsilon^2\sum_{i=1}^\infty \sigma_i^2\lambda_i^{*2}\right] - 4B^{-1}\mathcal{R}(\tilde\theta,\theta)-$$

$$-12B\mathbb{E}_\theta\Delta^\epsilon[\lambda^*] - 12B\Delta^\epsilon[\tilde\lambda] =$$

$$= -4B^{-1}\mathbb{E}_\theta\mathcal{R}(\theta^*,\theta) - 4B^{-1}\mathcal{R}(\tilde\theta,\theta) - CB\mathbb{E}_\theta\Delta^\epsilon[\lambda^*] - CB\Delta^\epsilon[\tilde\lambda]. \qquad (12)$$

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

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

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

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

Zatem zachodzi

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

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

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

Dochodzimy zatem do oszacowania postaci

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

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

$$\sup_i \sigma_i^2\lambda_i^2 = \sup_i \sigma_i^2\lambda_i^2\mathbf{1}\left\{x\sup_i\sigma_i^2\lambda_i^2 < \sum_{i=1}^\infty\sigma_i^2\lambda_i^2\right\} + \sup_i\sigma_i^2\lambda_i^2\mathbf{1}\left\{x\sup_i\sigma_i^2\lambda_i^2 \geqslant \sum_{i=1}^\infty\sigma_i^2\lambda_i^2\right\} \leqslant$$

$$\leqslant \frac{1}{x}\sum_{i=1}^\infty\sigma_i^2\lambda_i^2 + \max_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2\mathbf{1}\left\{x\sup_i\sigma_i^2\lambda_i^2 \geqslant \sum_{i=1}^\infty\sigma_i^2\lambda_i^2\right\} \leqslant$$

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

Stąd dostajemy oszacowanie na składnik

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

Wykorzystamy teraz powyższe nierówności dla nierówności (13).

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

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

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

Mamy stąd

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

Korzystając z lematu 4 mamy ponadto, że

$$\mathbb{E}_\theta||\theta^* - \theta||^2 \leqslant (1 + 4B^{-1})\mathbb{E}_\theta\mathcal{R}(\theta^*, \theta) + CB\mathbb{E}_\theta\Delta^\epsilon[\lambda^*] \leqslant$$

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

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

$$\mathbb{E}_\theta||\theta^* - \theta||^2 \leqslant (1 + \gamma B^{-1})\mathbb{E}_\theta\mathcal{R}(\theta^*, \theta) + CB\epsilon^2 L_\Lambda\omega(B^2 L_\Lambda) \leqslant$$

$$\leqslant \frac{1 + \gamma B^{-1}}{1 - \gamma B^{-1}} CB\epsilon^2 L_\Lambda\omega(B^2 L_\Lambda) + CB\epsilon^2 L_\Lambda\omega(B^2 L_\Lambda) + \frac{(1 + \gamma B^{-1})^2}{1 - \gamma B^{-1}}\mathcal{R}(\tilde{\theta}, \theta).$$

Zauważmy następnie, że z uwagi na fakt, że $\lim_{x\to\infty} \frac{x+\gamma}{x-\gamma} = 1$, 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

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

Nierówność ta kończy dowód twierdzenia 4. □

Dowód twierdzenia 5. Będziemy stosować te same oznaczenia na estymator i wyrocznię jak w dowodzie twierdzenia 4.
Zauważmy na początek, że z uwagi (8) mamy

$$\frac{\left(\sum_{k=1}^\infty \sigma_k^4\lambda_k^4\right)^{1/2}}{\sum_{k=1}^\infty \sigma_k^2\lambda_k^2} \leqslant \rho \Longrightarrow \frac{1}{\sum_{k=1}^\infty \sigma_k^2\lambda_k^2} \leqslant \frac{\rho}{\left(\sum_{k=1}^\infty \sigma_k^4\lambda_k^4\right)^{1/2}}.$$

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

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

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

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

Łącząc te dwie nierówności dostajemy, że

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

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

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

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

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

Co prowadzi do nierówności

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

a stąd dostajemy

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

Ponownie korzystając z lematu 4 dostajemy

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

Łącząc te dwie nierówności mamy

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

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

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

co, rozumując analogicznie do dowodu twierdzenia 5, prowadzi do nierówności

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

która kończy dowód. □

5. Główne rezultaty II

Rezultaty zaprezentowane w tym rozdziale są modyfikacjami rezultatów uzyskanych w pracy [7].
W tej części pracy podejmiemy próbę zbudowania analogicznych jak w rozdziale poprzednim nierówności wyroczni w przypadku, gdy w rozważanym problemie

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

operator A jest operatorem niekoniecznie zwartym.

Będziemy potrzebowali pojęcia operatora unitarnego.

Definicja 11. *Niech H, G będą ośrodkowymi przestrzeniami Hilberta. Operator $U\colon H \to G$ nazwiemy 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*

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

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

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

Ponadto, jeżeli operator kowariancji ξ oznaczymy przez $\mathbf{Cov}_\xi$, to operator kowariancji $T\xi$ jest postaci $\mathbf{Cov}_{T\xi} = T\mathbf{Cov}_\xi T^*$ ([6], str. 2615). Wprowadźmy także następujące definicje

Definicja 12. *([23], str. 102)*
Niech $\xi\colon H \to {}_2(T, \mathbb{T}, \tau)$ będzie losowym szumem. Powiem, że ξ ma skończony słaby drugi moment, jeżeli dla dowolnego $f \in H$ zachodzi $\int_T |\langle \xi, f\rangle|^2 d\tau < \infty$ oraz, że ma skończony silny drugi moment, jeżeli $\int_T \|\xi\|^2 d\tau < \infty$. Oczywiście silny moment implikuje słaby.

Definicja 13. *([3], str. 13)*
Operator $T\colon H \to H$ *nazwiemy operatorem śladowym, jeżeli istnieją ciągi* $\{a_k\}_{k=1}^{\infty}, \{b_k\}_{k=1}^{\infty} \subset H$ *takie, że*

$$\sum_{k=1}^{\infty} \|a_k\| \, \|b_k\| < \infty$$

oraz

$$\forall h \in H \;\; Ah = \sum_{k=1}^{\infty} \langle h, a_k \rangle b_k.$$

Okazuje się, że dla operatorów samosprzężonych można uogólnić twierdzenie o reprezentacji według wartości singularnych do następującej postaci ([22], str. 97)

Twierdzenie 6. *Niech* A *będzie operatorem samosprzężonym na ośrodkowej przestrzeni Hilberta* H*. Wtedy istnieje przestrzeń mierzalna* $(S, \mathcal{S}, \mu)$*, rzeczywista funkcja mierzalna* b *określona na* S*, operator unitarny* $U\colon H \to L_2(S, \mathcal{S}, \mu)$*, takie, że*

$$A = U^{-1} M_b U,$$

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

Zauważmy, że samosprzężone operatory zwarte i ich reprezentacja według wartości singularnych 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 względem bazy wektorów własnych.

Pierwszym krokiem jaki wykonamy przed dalszą analizą, będzie prekondycjonowanie problemu $Y = Af + \epsilon\xi$, czyli przekształcenie go do postaci

$$A^*Y = A^*Af + \epsilon A^*\xi. \tag{14}$$

Podstawową korzyścią z takiego przekształcenia jest to, że w dalszej części możemy zajmować się już tylko i wyłącznie operatorami samosprzężonymi i dodatnimi. Fakt, że operator A^*A jest dodatni implikuje, że funkcja b z twierdzenia 6 jest dodatnia μ– prawie wszędzie, a zatem operator mnożenia M_b jest odwracalny i jego odwrotność ma postać $(M_b)^{-1} = M_{b^{-1}}$. Z drugiej strony wiadomo, że f jest najlepszym rozwiązaniem w sensie minimalizacji kwadratu normy różnicy wyrażenia $Af - g$ wtedy i tylko wtedy, gdy zachodzi $A^*Af = A^*g$ ([1], str. 13).

Korzystając z twierdzenia spektralnego 6 możemy rozważany model (14)

$$A^*Y = A^*Af + \epsilon A^*\xi = U^{-1}M_bUf + \epsilon A^*\xi$$

przekształcić do następującej postaci stosując operator U

$$UA^*Y = U(U^{-1}M_bU)f + \epsilon UA^*\xi,$$

co można zapisać jako

$$Z = b\theta + \epsilon\eta, \tag{15}$$

gdzie $Z = UA^*Y$, $\theta = Uf$ oraz $\eta = UA^*\xi$.
Podobnie jak w przypadku operatorów zwartych możemy zapisać to wyrażenie w postaci analogicznej do (2)

$$X = \theta + \epsilon\sigma\eta, \tag{16}$$

gdzie $\sigma\eta$ rozumiane jest jako $M_{b^{-1}}UA^*\xi$. Przejście z równości (15) do równości (16) rozumiane jest jako następujące przekształcenie równania operatorowego. Skoro (15) jest postaci $UA^*Y = M_bUf + \epsilon UA^*\xi$, zatem mnożąc lewostronnie to równanie przez $M_{b^{-1}}$ dostajemy równanie $M_{b^{-1}}UA^*Y = M_{b^{-1}}M_bUf + \epsilon M_{b^{-1}}UA^*\xi$, co po oznaczeniu przez $X = M_{b^{-1}}UA^*Y$ prowadzi do równania (16), gdyż $M_{b^{-1}}M_b$ jest operatorem identycznościowym. Natomiast równanie

(16) rozumiane jest tak, że dla dowolnego elementu $g \in L_2(S, \mathcal{S}, \mu)$ obserwowana jest zmienna losowa $\langle X, g\rangle = \langle \theta, g\rangle + \epsilon\langle \eta, b^{-1}g\rangle$, gdzie z kolei $\langle \eta, b^{-1}g\rangle$ jest zmienną losową z przestrzeni $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Twierdzenie 6 mówi tylko, że funkcja b jest mierzalną funkcją rzeczywistą nie mówiąc nic więcej o jej zachowaniu. Zatem o źle postawionym problemie będziemy mogli mówić tylko wtedy, gdy $b \to 0$ w jakimś sensie.
Przed dalszym rozumowaniem, udowodnimy, czym jest η. Okazuje się, że jest ona gaussowskim kolorowym szumem na przestrzeni $L_2(S, \mathcal{S}, \mu)$.

Twierdzenie 7. *Niech η będzie określona jako $\eta = UA^*\xi$. Wtedy $\sqrt{\sigma}\eta$ jest gaussowskim białym szumem na przestrzeni $L_2(S, \mathcal{S}, \mu)$.*

Dowód. Zauważmy na początek, że zachodzi $A = U^* M_{b^{1/2}} U$. A stąd mamy, że $\langle \sqrt{\sigma}\eta, F\rangle = \langle M_{b^{-1/2}} UA^*\xi, Uf\rangle = \langle \xi, AU^* M_{b^{-1/2}} Uf\rangle \sim \mathcal{N}(0, ||AU^* M_{b^{-1/2}} Uf||^2)$, gdzie $||AU^* M_{b^{-1/2}} Uf||^2 = ||U^* M_{b^{1/2}} UU^* M_{b^{-1/2}} Uf||^2 = ||f||^2$, gdyż $U^* = U^{-1} = U$. Natomiast operator kowariancji ma postać $\mathbf{Cov}_{\sqrt{\sigma}\eta} = M_{b^{-1/2}} UA^* \mathbf{Cov}_\xi AU^* M_{b^{-1/2}} = M_{b^{-1/2}} M_b M_{b^{-1/2}} = I$, gdyż $\mathbf{Cov}_\xi = I$. □

Niech teraz $\hat{\theta}$ oznacza estymator elementu θ w modelu (16) 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.
W pracy [7] pracowano przy założeniu, że obserwacje zaburzone są przez błąd pochodzący z białego szumu. Jednak z uwagi na to, że w przestrzeni nieskończenie wymiarowej operator identycznościowy, będący operatorem kowariancji dla białego szumu, nie jest operatorem zwartym, czyli nie może być operatorem śladowym, a więc nie może mieć skończonego silnego drugiego momentu. Z drugiej strony, jeżeli opeartor kowariancji pewnego szumu jest operatorem śladowym, to wtedy szum ten posiada silny drugi moment ([23], str. 175). Dlatego za [6] i korzystając ze wcześniejszego prekondycjonowania będziemy zakładać o operatorze A i szumie ξ, że zachodzą następujące warunki gwarantujące skończoność drugich momentów regularyzowanego rozwiązania

$$\forall f \in G \ \mathbb{E}\langle \xi, f\rangle = 0 \text{ oraz } \|\mathbf{Cov}_\xi\| \leqslant 1, \tag{17}$$

$$\mathbb{E} \|A^*\xi\|^2 < \infty.$$

Jeżeli spełnione są powyższe dwa warunki, możliwe jest takie dobranie operatora U w twierdzeniu 6, by zachodziło

$$\forall s \in S \ \boldsymbol{V}ar(UK^*\xi(s)) \leqslant b(s)$$

oraz $b \in L_1(S, \mathcal{S}, \mu)$ ([6], str. 2616). Przykłady modeli spełniających powyższe warunki znajdują sie w rozdziale 6 oraz w [6], str. 2620– 2624.
W ogólnym przypadku w związku ze złym uwarunkowaniem problemu $Z = Th + \epsilon\xi$ rozważane są estymatory elementu h postaci

$$\hat{h}_\alpha = \Phi_\alpha(T^*T)T^*Y,$$

gdzie funkcja Φ_α ma następujące własności ([17], str. 1426)

$$\sup_{t \geqslant 0} \Phi_\alpha(t) < \infty \ \forall \alpha > 0,$$

$$\sup_{\alpha > 0, t \geqslant 0} t\Phi_\alpha(t) < \infty,$$

$$\Phi_\alpha(t) \to t^{-1}, \ gdy \ \alpha \to 0, \ \forall t > 0.$$

W rozważanym przez nas problemie możemy rozważyć estymator liniowy elementu θ w postaci zaproponowanej powyżej, gdzie celem zachowania spójności ze wcześniejszymi oznaczeniami będziemy dalej pisać, że jest on postaci

$$\hat{\theta} = \lambda X, \tag{18}$$

gdzie $\lambda = \Phi_\alpha(A^*A)U^*M_b$ jest pewną nielosową funkcją, a funkcja Φ_α zależy od konkretnego wyboru estymatora. Przykładowo dla estymatorów projekcyjnych funkcja Φ_α przyjmuje postać

$$\Phi_\alpha(t) = \frac{1}{t}\mathbf{1}_{[\alpha,\infty)}.$$

Funkcję λ będziemy konsekwentnie nazywać wagą lub filtrem. W dalszej części ograniczymy się tylko do rozważania rzeczywistych filtrów i przestrzeni. Możemy teraz wyznaczyć górne oszacowanie ryzyka estymatora postaci (18). Warunki (17) narzucone na postać szumu i operator A gwarantują, że zarówno wyznaczone oszacowanie jak i samo ryzyko są skończone.

$$\begin{gathered}
\mathcal{R}(\hat{\theta},\theta) = \mathbb{E}_\theta||\hat{\theta}-\theta||_2^2 = \mathbb{E}_\theta\left\|\lambda\theta + \epsilon\lambda\sigma\eta - \theta\right\|_2^2 = \\
= \mathbb{E}_\theta\left\|\lambda\theta-\theta\right\|_2^2 + \mathbb{E}_\theta\left\|\epsilon\lambda\sigma\eta\right\|_2^2 = \left\|(1-\lambda)\theta\right\|_2^2 + \epsilon^2\mathbb{E}_\theta\left\|\lambda\sigma\eta\right\|_2^2 = \\
= \left\|(1-\lambda)\theta\right\|_2^2 + \epsilon^2\mathbb{E}_\theta\int_S\left(\lambda(s)\sigma(s)\eta(s)\right)^2 d\mu = \\
= \left\|(1-\lambda)\theta\right\|_2^2 + \epsilon^2\int_S\lambda^2(s)\sigma^2(s)\mathbb{E}_\theta\eta^2(s)d\mu \leqslant \\
\leqslant \left\|(1-\lambda)\theta\right\|_2^2 + \epsilon^2\int_S\lambda^2(s)\sigma(s)d\mu = \int_S(1-\lambda(s))^2\theta^2(s)d\mu + \epsilon^2\int_S\lambda^2(s)\sigma(s)d\mu
\end{gathered}$$

W dalszej części będziemy omijać argumenty w funkcjach podcałkowych. Warto zauważyć, że w powyższym wyrażeniu w wyrażeniu $\epsilon^2\int_S\lambda^2\sigma d\mu$ składnik σ związany ze złym uwarunkowaniem problemu występuje w potędze o jeden niższej niż w analogicznym wyrażeniu (3), jednak trzeba zauważyć, że oba modele różnią się charakterem szumu.

Definicja 14. *Wprowadzimy następujące oznaczenie*

$$\Psi(\lambda,\theta) = \int_S(1-\lambda)^2\theta^2 d\mu + \epsilon^2\int_S\lambda^2\sigma d\mu$$

na wyrażenie będące odpowiednikiem wyrażenia (3) w modelu rozważanym w rozdziale 4.

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

Definicja 15. *Nieobciążonym estymatorem wyrażenia $\Psi(\lambda,\theta)$ w modelu (16) nazywamy wyrażenie*

$$\psi(\lambda,X) = \int_S(\lambda^2-2\lambda)(X^2-\epsilon^2\sigma^2)d\mu + \epsilon^2\int_S\lambda^2\sigma d\mu, \tag{19}$$

będące nieobciążonym estymatorem wyrażenia $\Psi(\lambda,\theta) - \int_S\theta^2 d\mu$ dla estymatora postaci $\hat{\theta} = \lambda X$. Wyrażenie to jest skończone przy założeniu (5).

Niech teraz badane filtry należą do pewnej skończonej rodziny $\Lambda = \{\lambda^1,\ldots,\lambda^N\}$. Naszym celem będzie wybór na podstawie 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 $\Psi(\lambda,\theta)$ oraz jego estymator $\psi(\lambda,X)$ były skończone oraz by estymatory nie były zbyt duże. Założymy także skończoność wyrażeń występujących później w wyrażeniach występujących w nierównościach wyroczniach dla badanych estymatorów.

Założenie 5. *Załóżmy, że*

$$\forall\ \lambda \in \Lambda\ 0 < \int_S \sigma^2 \lambda^2 d\mu < \infty,$$

$$\max_{\lambda \in \Lambda} \|\lambda\|_\infty \leqslant 1,$$

$$\forall\ \lambda \in \Lambda\ \int_S \lambda \sigma^2 d\mu < \infty,$$

$$\forall\ \lambda \in \Lambda\ \int_S \sigma \lambda^2 d\mu < \infty$$

$$\exists\ C_2 > 0\ \forall\ \lambda \in \Lambda\ \int_S \sigma^4 \lambda^2 d\mu \leqslant C_2 \int_S \sigma^3 \lambda^4 d\mu,$$

Wprowadzimy także potrzebne oznaczenia

$$\rho(\lambda) = \left\|\sigma^2 \lambda\right\|_\infty \left[\int_S \sigma^4 \lambda^4 d\mu\right]^{-1/2},$$

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

$$S = \frac{\max_{\lambda \in \Lambda} \left\|\sigma^2 \lambda^2\right\|_\infty}{\min_{\lambda \in \Lambda} \left\|\sigma^2 \lambda^2\right\|_\infty},$$

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

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

Interpretacje powyższych wyrażeń przenoszą się z rozważanego wcześniej dyskretnego przypadku.
Mając już odpowiednie założenia możemy zdefiniować poszukiwany estymator naśladujący ryzyko najlepszego estymatora w klasie Λ. Ponownie wykorzystamy w tym celu zdefiniowane wcześniej wyrażenie $\psi(\lambda, X)$, który w pewien sposób przybliża górne ograniczenie na ryzyko, które chcielibyśmy zminimalizować.

Definicja 16. *Niech funkcjonał $\psi(\lambda, X)$ będzie zdefiniowany jak w (19). Poszukiwanym filtrem jest element minimalizujący względem $\lambda \in \Lambda$ funkcjonał $\psi(\lambda, X)$, czyli*

$$\lambda^* = \arg\min_{\lambda \in \Lambda} \psi(\lambda, X). \tag{20}$$

Przedstawimy teraz uogólnione wersje lematów 2, 3 i 4 na rozważany obecnie przypadek. Dowody znajdują się w części 7.

Lemat 5. *Niech η będzie gaussowskim szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ z operatorem kowariancji postaci TT^* dla pewnego ograniczonego operatora T i o skończonym silnym drugim momencie 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 $K \geqslant 1$ zachodzi*

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

Lemat 6. *Niech η będzie gaussowskim szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ z operatorem kowariancji postaci TT^* dla pewnego ograniczonego operatora T i o skończonym silnym drugim momencie 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*

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

Wtedy istnieje stała D zależna tylko od q i operatora T, taka, że dla dowolnego $K \geqslant 1$ zachodzi

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

Lemat 7. *Niech η będzie gaussowskim szumem na przestrzeni Hilberta $L_2(S, \mathcal{S}, \mu)$ z operatorem kowariancji postaci TT^* dla pewnego ograniczonego operatora T i o skończonym silnym drugim momencie. Niech ponadto $\hat{\theta} = \hat{\lambda}(X)X$ będzie liniowym estymatorem z wagą z wartościami z przedziału $[-1, 1]$ przyjmującym wartości w zbiorze Λ. Oznaczmy przez*

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

Wtedy istnieje stała $C > 0$ zależna tylko od operatora T taka, że dla dowolnego $B > 0$ zachodzi

$$\mathbb{E}_\theta \left\|\hat{\theta} - \theta\right\|_2^2 \leqslant \max\{1, C_2\}(1+4B^{-1}\max\{1, \left\|T\right\|^2\})\mathbb{E}_\theta\Psi(\hat{\lambda}, \theta)+CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)].$$

Możemy teraz przejść do wypowiedzenia dwóch twierdzeń będących zarazem głównym wynikiem tego rozdziału i uogólnieniem twierdzeń 4 i 5 na przypadek, gdy badany problem może być modelowany jako (14).

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 (20) zachodzi*

$$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant \max\{1, C_2\}\left((1+\gamma_1 B^{-1}\max\{1, \left\|A\right\|^2\})\min_{\lambda\in\Lambda}\Psi(\lambda, \theta) + \gamma_2 B\left\|A\right\|^2\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda)\right),$$

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

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

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 (20) zachodzi*

$$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant \max\{1, C_2\}(1+\gamma_3\left\|A\right\|^2\rho\sqrt{L_\Lambda})\min_{\lambda\in\Lambda}\Psi(\lambda, \theta),$$

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

Zanim udowodnimy powyższe twierdzenia, zauważmy, że nierówność występująca w tezie twierdzenia 8 możemy zapisać w postaci

$$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant C'(1+\gamma_1 B^{-1})\min_{\lambda\in\Lambda}\left(\mathbb{E}_\theta\left\|\lambda X - \theta\right\|^2 + pen_A(\lambda)\right)+\gamma_2 B\left\|A\right\|^2\epsilon^2 L_\Lambda\omega(B^2 L_\Lambda)$$

dla pewnej funkcji $pen_A\colon \Lambda \to \mathbb{R}_+$, która może być interpretowana jako funkcja kary w badanym problemie pojawiająca się w związku z brakiem zwartości ('skomplikowaniem') operatora A. Takie przedstawienie uzyskanego wyniku jest podobne do wyników uzyskanych w innych modelach (por. [2], str. 378, [14], str. 181, [14], str. 668, [10], str. 37).

Dowód twierdzenia 8. Niech TT^* będzie rozkładem Choleskiego operatora kowariancji UA^*AU^* szumu η. Niech $\tilde{\lambda} \in \Lambda$ będzie takim filtrem, że związany z nim estymator jest wyrocznią, czyli $\tilde{\theta} = \arg\min_{\lambda\in\Lambda}\Psi(\lambda, \theta)$, natomiast przez λ^* oznaczmy filtr definiowany przez (20) i związany z nim estymator przez θ^*. Oznaczmy ponadto $\max\{1, C_2\} = C'$ oraz $\max\{1, \left\|T\right\|^2\} = C''$. W rozpatrywanym modelu (16) $X = \theta + \epsilon\sigma\eta$. Dostajemy stąd, że

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

A stąd mamy

$$\mathbb{E}_\theta \psi[\lambda^*, X] = \mathbb{E}_\theta \Psi(\lambda^*, \theta) - \int_S \theta^2 d\mu +$$

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

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

$$\epsilon \mathbb{E}_\theta \int_S (1-\lambda^*)^2 \sigma\theta\eta d\mu = \epsilon \mathbb{E}_\theta \int_S [(1-\lambda^*)^2 - (1-\tilde{\lambda})^2]\sigma\theta\eta d\mu \geqslant$$

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

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

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

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

$$\geqslant -2\epsilon \|T\| \sqrt{2\ln(NS)} \mathbb{E}_\theta \left(\int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda^2})\theta^2\sigma^2 d\mu \right)^{1/2} -$$

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

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

$$2\epsilon \|T\| \sqrt{2\ln(NS)} \mathbb{E}_\theta \left(\int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2](\lambda^{*2} + \tilde{\lambda^2})\theta^2\sigma^2 d\mu \right)^{1/2} \leqslant$$

$$\leqslant \mathbb{E}_\theta 2\epsilon \|T\| \sqrt{2\ln(NS) \left\| (\lambda^{*2} + \tilde{\lambda^2})\sigma^2 \right\|_\infty} \left(\int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]\theta^2 d\mu \right)^{1/2} \leqslant$$

$$\leqslant 2B\epsilon^2 \|T\|^2 \ln(NS) \mathbb{E}_\theta \left\| (\lambda^{*2} + \tilde{\lambda^2})\sigma^2 \right\|_\infty + B^{-1} \mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]\theta^2 d\mu \leqslant$$

$$\leqslant 2B\epsilon^2 \|T\|^2 \ln(NS) \mathbb{E}_\theta \left(\left\| \lambda^{*2}\sigma^2 \right\|_\infty + \left\| \tilde{\lambda^2}\sigma^2 \right\|_\infty \right) + B^{-1} \mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]\theta^2 d\mu.$$

Postępując podobnie z drugim wyrażeniem otrzymujemy

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

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

Korzystając z definicji wielkości S mamy, że

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

A stąd mamy oszacowanie postaci

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

$$\leqslant 4B\epsilon^2 \|T\|^2 \ln(NS)\mathbb{E}_\theta \left(\left\|\lambda^{*2}\sigma^2\right\|_\infty + \left\|\tilde{\lambda^2}\sigma^2\right\|_\infty \right) + B^{-1}\mathbb{E}_\theta \int_S [(1-\lambda^*)^2 + (1-\tilde{\lambda})^2]\theta^2 d\mu.$$

Jako, że zachodzi $\ln(NS) \leqslant L_\Lambda$, to łącząc powyższe dwa oszacowania dostajemy, że

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

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

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 $\lambda^2 \leqslant (\lambda^2 - 2\lambda)^2 \leqslant 9\lambda^2$ rozumiane jako nierówność funkcyjna zachodząca dla wszystkich argumentów funkcji λ. Posłużymy się teraz lematem 6 z $K = S$, $q = 3$ i $v = (\lambda^{*2} - 2\lambda^*)\sigma^2$.

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

$$\leqslant \epsilon^2 \|T\|^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta ||v|| + \sqrt{\mathbb{E}_\theta ||v||^2/S} \right).$$

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

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

Stąd dostajemy, że

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

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

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

$$\geqslant -\epsilon^2 \|T\|^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3)) \right) \left(\mathbb{E}_\theta ||v|| + \sqrt{\mathbb{E}_\theta ||v||^2/S} \right) \geqslant$$

$$\geqslant -\epsilon^2 \|T\|^2 C\sqrt{L_\Lambda}\mathbb{E}_\theta \left(\int_S (\lambda^{*2} - 2\lambda^*)^2\sigma^4 d\mu \right)^{1/2} - \epsilon^2 \|T\|^2 C\sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*)^2\sigma^4 d\mu \right)^{1/2} \geqslant$$

$$\geqslant -\epsilon^2 \|T\|^2 C\sqrt{L_\Lambda}\mathbb{E}_\theta \left(\int_S \lambda^{*3}\sigma^4 d\mu \right)^{1/2} - \epsilon^2 \|T\|^2 C\sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \int_S \lambda^{*3}\sigma^4 d\mu \right)^{1/2}.$$

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

$$S^{-1}\mathbb{E}_\theta \int_S \sigma^3\lambda^{*4} d\mu \leqslant \mathbb{E}_\theta \left\|\sigma^2\lambda^{*2}\right\|_\infty \mathbb{E}_\theta \int_S \sigma\lambda^{*2} d\mu.$$

Podobnie dla drugiego wyrażenia

$$\mathbb{E}_\theta \left(\int_S \sigma^3\lambda^{*4} d\mu \right)^{1/2} \leqslant \mathbb{E}_\theta \left(\left\|\sigma^2\lambda^{*2}\right\|_\infty \int_S \sigma\lambda^{*2} d\mu \right)^{1/2}.$$

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

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

Możemy także zapisać

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

Łącząc oba te oszacowania dostajemy

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

Powyższe rozważania pozwalają zapisać nam następujący wniosek

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

$$\geqslant -4B^{-1}\left\|T\right\|^2\mathbb{E}_\theta\int_S\sigma\lambda^{*2}d\mu-4B^{-1}\left\|T\right\|^2\int_S\sigma\tilde{\lambda}^2d\mu-CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]-CB\left\|T\right\|^2\Delta^\epsilon[\tilde{\lambda}].$$

Możemy teraz przejść do dalszych oszacowań.

$$\mathbb{E}_\theta\psi[\lambda^*,X]=\mathbb{E}_\theta\Psi(\lambda^*,\theta)-\int_S\theta^2d\mu+$$

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

$$\geqslant\mathbb{E}_\theta\Psi(\lambda^*,\theta)-\int_S\theta^2d\mu-4B^{-1}C''\mathbb{E}_\theta\Psi(\lambda^*,\theta)-$$

$$-4B^{-1}C''\Psi(\tilde{\lambda},\theta)-CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]-CB\left\|T\right\|^2\Delta^\epsilon[\tilde{\lambda}].$$

Zatem zachodzi, że

$$(1-4B^{-1}C''\mathbb{E}_\theta\Psi(\lambda^*,\theta)\leqslant$$

$$\leqslant\mathbb{E}_\theta\psi[\lambda^*,X]+\int_S\theta^2d\mu+4B^{-1}C''\Psi(\tilde{\lambda},\theta)+CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]+CB\left\|T\right\|^2\Delta^\epsilon[\tilde{\lambda}].$$

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

$$\mathbb{E}_\theta\psi[\lambda^*,X]\leqslant\mathbb{E}_\theta\psi[\tilde{\lambda},X]=\Psi(\tilde{\lambda},\theta)+\int_S\theta^2d\mu.$$

Otrzymujemy zatem, że

$$(1-4B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta)\leqslant$$

$$\leqslant 1+4B^{-1}C'')\Psi(\tilde{\lambda},\theta)+CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]+CB\left\|T\right\|^2\Delta^\epsilon[\tilde{\lambda}].$$

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

$$\left\|\sigma^2\lambda^2\right\|_\infty=\left\|\sigma^2\lambda^2\mathbf{1}\left\{x\left\|\sigma^2\lambda^2\right\|_\infty<\int_S\sigma\lambda^2d\mu\right\}\right\|_\infty+\left\|\sigma^2\lambda^2\mathbf{1}\left\{x\left\|\sigma\lambda^2\right\|_\infty\geqslant\int_S\sigma\lambda^2d\mu\right\}\right\|_\infty\leqslant$$

$$\leqslant\frac{1}{x}\int_S\sigma\lambda^2d\mu+\max_{\lambda\in\Lambda}\left\|\sigma^2\lambda^2\mathbf{1}\left\{x\left\|\sigma^2\lambda^2\right\|_\infty\geqslant\int_S\sigma\lambda^2d\mu\right\}\right\|_\infty\leqslant$$

$$\leqslant\frac{1}{x}\int_S\sigma\lambda^2d\mu+\omega(x)\leqslant\frac{1}{x\epsilon^2}\Psi(\lambda,\theta)+\omega(x)\ \forall_{\lambda\in\Lambda}.$$

Stąd dostajemy

$$\Delta^\epsilon[\lambda]=\epsilon^2L_\Lambda\left\|\sigma^2\lambda^2\right\|_\infty\leqslant\frac{L_\Lambda}{x}\Psi(\lambda,\theta)+\epsilon^2L_\Lambda\omega(x).$$

Wykorzystamy teraz powyższe nierówności do wyrażenia $(1-4B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta)\leqslant(1+4B^{-1}C'')\Psi(\tilde{\lambda},\theta)+CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]+CB\left\|T\right\|^2\Delta^\epsilon[\tilde{\lambda}]$.

$$(1-4B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta)\leqslant(1+4B^{-1}C'')\Psi(\tilde{\lambda},\theta)+CB\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]+CB\left\|T\right\|^2\Delta^\epsilon[\tilde{\lambda}]\leqslant$$

$$\leqslant CB\left\|T\right\|^2\frac{L_\Lambda}{x}\mathbb{E}_\theta\Psi(\lambda^*,\theta)+CB\left\|T\right\|^2\epsilon^2L_\Lambda\omega(x)+\left(CB\left\|T\right\|^2\frac{L_\Lambda}{x}+1+4B^{-1}C''\right)\Psi(\tilde{\lambda},\theta).$$

Mamy stąd, że

$$\mathbb{E}_\theta \Psi(\lambda^*,\theta) \leqslant \frac{CB\,\|T\|^2\,\epsilon^2 L_\Lambda \omega(x)}{1-4B^{-1}C''-CB\,\|T\|^2\,\frac{L_\Lambda}{x}} + \frac{CB\,\|T\|^2\,\frac{L_\Lambda}{x}+1+4B^{-1}C''}{1-4B^{-1}C''-CB\,\|T\|^2\,\frac{L_\Lambda}{x}}\Psi(\tilde{\lambda},\theta).$$

Korzystając z lematu 7 mamy ponadto, że

$$\frac{1}{C'}\mathbb{E}_\theta ||\theta^*-\theta||^2 \leqslant (1+4B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta)+CB\,\|T\|^2\,\mathbb{E}_\theta\Delta^\epsilon[\lambda^*] \leqslant$$

$$\leqslant (1+4B^{-1}C''+CB\,\|T\|^2\,\frac{L_\Lambda}{x})\mathbb{E}_\theta\Psi(\lambda^*,\theta)+CB\,\|T\|^2\,\epsilon^2 L_\Lambda\omega(x).$$

Niech teraz $x = B^2 L_\Lambda$ oraz γ będzie stałą zależną tylko od C (zależące tylko od stałej C_2). Wtedy

$$\frac{1}{C'}\mathbb{E}_\theta ||\theta^*-\theta||^2 \leqslant (1+\gamma B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta)+CB\,\|T\|^2\,\epsilon^2 L_\Lambda\omega(B^2 L_\Lambda\,\|T\|^2) \leqslant$$

$$\leqslant \frac{1+\gamma B^{-1}C''}{1-\gamma B^{-1}C''}CB\,\|T\|^2\,\epsilon^2 L_\Lambda\omega(B^2L_\Lambda\,\|T\|^2)+CB\,\|T\|^2\,\epsilon^2 L_\Lambda\omega(B^2L_\Lambda\,\|T\|^2)+\frac{(1+\gamma B^{-1}C'')^2}{1-\gamma B^{-1}C''}\Psi(\tilde{\lambda},\theta).$$

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

$$\frac{1}{C'}\mathbb{E}_\theta ||\theta^*-\theta||^2 \leqslant (1+\gamma_1 B^{-1}C'')\Psi(\tilde{\lambda},\theta)+\gamma_2 B\,\|T\|^2\,\epsilon^2 L_\Lambda\omega(B^2L_\Lambda),$$

która kończy dowód twierdzenia 8. □

Dowód twierdzenia 9. Będziemy stosować te same oznaczenia na estymator, wyrocznię i stałe jak w dowodzie twierdzenia 8. Wyrażenie $\frac{\|\sigma^2\lambda^2\|_\infty}{\int_S \sigma^2\lambda^2 d\mu}$ możemy oszacować w następujący sposób

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

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

$$\Psi(\lambda,\theta)=\int_S(1-\lambda)^2\theta^2 d\mu+\epsilon^2\int_S\sigma\lambda^2 d\mu \geqslant \epsilon^2\int_S\sigma\lambda^2 d\mu.$$

Skąd dostajemy, że

$$\epsilon^2\,\|\sigma^2\lambda^2\|_\infty \leqslant C_2\rho^2\Psi(\lambda,\theta).$$

Co z kolei prowadzi do oszacowania

$$\Delta^\epsilon[\lambda]=\epsilon^2 L_\Lambda\,\|\sigma^2\lambda^2\|_\infty \leqslant C_2\rho^2 L_\Lambda\Psi(\lambda,\theta).$$

W dowodzie twierdzenia 8 uzyskaliśmy nierówność następującej postaci

$$(1-4B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta) \leqslant (1+4B^{-1}C'')\Psi(\tilde{\lambda},\theta)+CB\,\|T\|^2\,\mathbb{E}_\theta\Delta^\epsilon[\lambda^*]+CB\,\|T\|^2\,\Delta^\epsilon[\tilde{\lambda}].$$

Wykorzystując wyprowadzone oszacowanie mamy stąd, że

$$(1-4B^{-1})\mathbb{E}_\theta\Psi(\lambda^*,\theta) \leqslant (1+4B^{-1}C'')\Psi(\tilde{\lambda},\theta)+CBC_2\,\|T\|^2\,\rho^2 L_\Lambda\mathbb{E}_\theta\Psi(\lambda^*,\theta)+CBC_2\,\|T\|^2\,\rho^2 L_\Lambda\Psi(\tilde{\lambda},\theta),$$

co prowadzi do

$$\mathbb{E}_\theta\Psi(\lambda^*,\theta) \leqslant \frac{1+4B^{-1}C''+CB\,\|T\|^2\,\rho^2 L_\Lambda}{1-4B^{-1}C''-CB\,\|T\|^2\,\rho^2 L_\Lambda}\Psi(\tilde{\lambda},\theta).$$

Ponownie korzystając z lematu 7 dostajemy, że

$$\frac{1}{C'}\mathbb{E}_\theta ||\theta^*-\theta||^2 \leqslant (1+4B^{-1}C'')\mathbb{E}_\theta\Psi(\lambda^*,\theta)+CB\,\|T\|^2\,\mathbb{E}_\theta\Delta^\epsilon[\lambda^*] \leqslant$$

$$\leqslant (1 + 4B^{-1}C'' + CB \|T\|^2 \rho^2 L_\Lambda)\Psi(\tilde{\lambda}, \theta).$$

Łącząc te dwie nierówności mamy

$$\frac{1}{C'}\mathbb{E}_\theta||\theta^* - \theta||^2 \leqslant \frac{(1 + 4B^{-1}C'' + CB \|T\|^2 \rho^2 L_\Lambda)^2}{1 - 4B^{-1}C'' - CB \|T\|^2 \rho^2 L_\Lambda}\Psi(\tilde{\lambda}, \theta).$$

Zauważmy teraz, że istnieje taka stała $\gamma_4 > 0$, że jeśli tylko zachodzi $\|T\|^2 \rho^2 L_\Lambda \leqslant \gamma_4$, to wybór B jako $(\rho^2 L_\Lambda)^{-1/2}$ prowadzi do nierówności $4B^{-1}C''+CB \|T\|^2 \rho^2 L_\Lambda < \|T\|^2 1/2$, a stąd $1 - 4B^{-1} - CB \|T\|^2 \rho^2 L_\Lambda \geqslant 1/2$, czyli jest odcięte od zera. Wtedy wybór $B = (\rho^2 L_\Lambda)^{-1/2}$ prowadzi do nierówności

$$\frac{1}{C'}\mathbb{E}_\theta||\theta^* - \theta||^2 \leqslant \frac{(1 + C \|T\|^2 \rho\sqrt{L_\Lambda})^2}{1 - C \|T\|^2 \rho\sqrt{L_\Lambda}}\Psi(\tilde{\lambda}, \theta),$$

która z kolei prowadzi nas do postulowanej na początku nierówności

$$\frac{1}{C'}\mathbb{E}_\theta||\theta^* - \theta||^2 \leqslant (1 + \gamma_3 \|T\|^2 \rho\sqrt{L_\Lambda})\Psi(\tilde{\lambda}, \theta),$$

która kończy dowód twierdzenia 9. □

Wniosek 3. *Niech założenie 5 będzie spełnione, ponadto niech zachodzi* $\lim_{\epsilon\to 0} \rho^2 \ln(NS) = 0$. *Wtedy istnieją stałe* $\mathbb{C}_2 > 0, \mathbb{C}_3 > 0$ *zależące tylko od stałej* C_2 *i operatora* $T = UA^*$*, takie że dla* $\rho^2 \ln(NS) < \mathbb{C}_2$ *i dla dowolnego* $\theta \in L_2(S, \mathbb{S}, \mu)$ *zachodzi*

$$\mathbb{E}_\theta||\theta^* - \theta||^2 \leqslant \max\{1, C_2\} \left(1 + \mathbb{C}_3 \|T\|^2 \rho\sqrt{ln(NS)}\right) \min_{\lambda\in\Lambda} \Psi(\lambda, \theta),$$

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

6. Przykład

7. Lematy pomocnicze

W rozdziale tym podamy dowody pomocniczych lematów użytych w pracy.

Lemat 8. *Niech* $\{X_i\}_{i=1}^\infty$ *będzie ciągiem niezależnych zmiennych losowych o rozkładach odpowiednio* $X_i \sim \mathcal{N}(0, s_i^2)$ *i niech* $\sum_{i=1}^\infty s_i^2 = S < \infty$. *Wtedy istnieje zmienna losowa* $X \sim \mathcal{N}(0, S)$ *taka, że* $L_2 - \lim_{n\to\infty} \sum_{i=1}^n X_i = X$.

Dowód. Przestrzeń $L_2(\Omega, \mathcal{F}, \mathbb{P})$ jest przestrzenią zupełną, więc wystarczy pokazać, że ciąg $\{\sum_{i=1}^\infty X_i\}_{n=1}^\infty$ jest ciągiem Cauchy'ego i graniczna zmienna losowa ma odpowiedni rozkład. Niech $n > m$, wtedy

$$0 \leqslant \left\|\sum_{i=1}^n X_i - \sum_{i=1}^m X_i\right\|_2^2 = \left\|\sum_{i=m+1}^n X_i\right\|_2^2 \leqslant \sum_{m+1}^n \|X_i\|_2^2 = \sum_{m+1}^n s_i^2 \to 0,\ n, m \to \infty,$$

bo szereg $\sum_{i=1}^\infty s_i^2$ jest zbieżny. Zatem ciąg $\{X_i\}_{i=1}^\infty$ jest ciągiem Cauchy'ego, a zatem istnieje zmienna losowa X taka, że $L_2 - \lim_{n\to\infty} \sum_{i=1}^n X_i = X$ oraz z uwagi na niezależność zmiennych X_i, $\sum_{i=1}^n X_i \sim \mathcal{N}(0, \sum_{i=1}^n s_i^2)$, a jako, że zbieżność w normie L_2 implikuje słabą zbieżność dostajemy żądaną tezę o rozkładzie X. □

Lemat 9. *Niech* $\{X_i\}_{i=1}^\infty$ *będzie ciągiem niezależnych zmiennych losowych o rozkładach odpowiednio* $X_i \sim \mathcal{N}(\theta_i, \sigma_i^2)$ *oraz niech szeregi* $\sum_{i=1}^\infty \sigma_i^2$ *i* $\sum_{i=1}^\infty \theta_i^2$ *będą zbieżne. Wtedy istnieje zmienna losowa* $Y \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ *taka, że* $L_2 - \lim_{n\to\infty} \sum_{i=1}^n X_i^2 = Y$.

Dowód. Przestrzeń $L_2(\Omega,\mathcal{F},\mathbb{P})$ jest przestrzenią zupełną, więc wystarczy pokazać, że ciąg $\left\{\sum_{i=1}^{n} X_i^2\right\}_{n=1}^{\infty}$ jest ciągiem Cauchy'ego. Niech $n > m$, wtedy mamy

$$0 \leqslant \left\| \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{m} X_i^2 \right\|_2^2 = \left\| \sum_{i=m+1}^{n} X_i^2 \right\|_2^2 \leqslant \sum_{m+1}^{n} \left\| X_i^2 \right\|_2^2 .$$

Zauważmy teraz, że $\left\|X_i^2\right\|_2^2 = \mathbb{E}X_i^4 = \mathbb{E}\left((X_i - \mathbb{E}X_i) + \mathbb{E}X_i\right)^4 = 3\sigma_i^2 + 6\theta_i^2\sigma_i^2 + \theta_i^4$. Założenie o zbieżności szeregów $\sum_{i=1}^{\infty}\sigma_i^2$ i $\sum_{i=1}^{\infty}\theta_i^2$ implikuje zbieżność szeregów $\sum_{i=1}^{\infty}\sigma_i^4$, $\sum_{i=1}^{\infty}\theta_i^4$ oraz $\sum_{i=1}^{\infty}\sigma_i^2\theta_i^2$, co wraz z poprzednim oszacowaniem pokazuje, że ciąg $\left\{\sum_{i=1}^{\infty} X_i^2\right\}_{n=1}^{\infty}$ jest ciągiem Cauchy'ego, co na mocy zupełności $L_2(\Omega,\mathcal{F},\mathbb{P})$ implikuje istnienie żądanej zmiennej losowej Y. □

Lemat 10. *Niech $\eta\colon H \to L_2(T,\mathbb{T},\tau)$ będzie gaussowskim szumem o skończonym silnym drugim momencie oraz niech v będzie pewnym elementem z przestrzeni H takim, że $\|v\|_\infty < \infty$. Załóżmy, że zachodzi $\langle \eta, v\rangle \sim \mathcal{N}(0,1)$. Wtedy dla dowolnego $t > 0$ zachodzi*

$$\ln \mathbb{E}\exp\left(t\langle \eta^2 - 1, v\rangle\right) \leqslant \frac{t^2\left\|v\right\|_2^2}{1 - t\left\|v\right\|_\infty}.$$

Dowód. Zachodzenie powyższego lematu pokazano, gdy v jest funkcją prostą w [13], str. 1325. Jako że η ma skończone drugie momenty i v jest ograniczona, możemy zastosować metodę komplikacji by uzyskać żądaną tezę. □

Przejdziemy teraz do dowodów lematów użytych do dowodzenia głównych rezultatów pracy.

Dowód lematu 2. Na początek zauważmy pewien użyteczny fakt. Niech $X \sim \mathcal{N}(0,1)$ będzie zmienną losową. Wtedy zachodzi oszacowanie

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

Istotnie możemy napisać, że

$$\mathbb{E}X^2\mathbf{1}_{\{|X|>a\}} = 2\int_a^{\infty}\frac{1}{\sqrt{2\pi}}x^2e^{-x^2/2}dx = \frac{2}{\sqrt{2\pi}}\left[-xe^{-x^2/2}|_a^{\infty} + \int_a^{\infty}e^{-x^2/2}dx\right] \leqslant$$

$$\leqslant \frac{2}{\sqrt{2\pi}}\left[ae^{-a^2/2} + \frac{1}{a}e^{-a^2/2}\right],$$

gdzie skorzystaliśmy z nierówności $1 - \Phi(x) \leqslant x^{-1}\phi(x)$ zachodzącej dla dowolnego $x > 0$, gdzie Φ, ϕ oznaczają odpowiednio dystrybuantę i gęstość standardowego rozkładu normalnego ([9], str. 175). Przejdziemy teraz do dowodu właściwej części lematu.
Oznaczmy przez $\zeta_v = \frac{1}{\|v\|}\sum_{k=1}^{\infty} v_k\xi_k$. Wyrażenie to jest skończone na mocy lematu 8 i faktu, że $v \in l^2$. Zachodzi oczywiście, że $|\zeta_v|\,\|v\| = |\sum_{k=1}^{\infty} v_k\xi_k|$. Możemy zatem zapisać, że

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

$$\mathbb{E}\,\|v\|\max_{u\in V}|\zeta_u|\mathbf{1}_{\{\max_{u\in V}|\zeta_u|\leqslant\sqrt{2\ln(NK)}\}} +$$

$$+\mathbb{E}\,\|v\|\max_{u\in V}|\zeta_u|\mathbf{1}_{\{\max_{u\in V}|\zeta_u|>\sqrt{2\ln(NK)}\}} \leqslant$$

$$\leqslant \sqrt{2\ln(NK)}\mathbb{E}\,\|v\| + \mathbb{E}\,\|v\|\max_{u\in V}|\zeta_u|\mathbf{1}_{\{\max_{u\in V}|\zeta_u|>\sqrt{2\ln(NK)}\}}.$$

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

$$\sqrt{2\ln(NK)}\mathbb{E}\,\|v\| + \mathbb{E}\,\|v\|\max_{u\in V}|\zeta_u|\mathbf{1}_{\{\max_{u\in V}|\zeta_u|>\sqrt{2\ln(NK)}\}} \leqslant$$

$$\leqslant \sqrt{2\ln(NK)}\mathbb{E}\,\|v\| + \left(\mathbb{E}\,\|v\|^2\right)^{1/2}\left(\mathbb{E}\max_{u\in V}|\zeta_u|^2\mathbf{1}_{\{\max_{u\in V}|\zeta_u|>\sqrt{2\ln(NK)}\}}\right)^{1/2}.$$

Rozważmy teraz funkcję $F(t) = t^2\mathbf{1}_{\{t>\sqrt{2\ln(NK)}\}}$. Z uwagi na monotoniczność funkcji kwadratowej dla dodatnich argumentów zachodzi, że

$$F(\max_{uv\in V}|\zeta_u|) = \max_{u\in V}F(|\zeta_u|),$$

Ponownie na mocy lematu 8 i niezależności zmiennych losowych $\{\xi_i\}$ mamy, że zmienne losowe ζ_v mają takie same rozkłady normalne $\mathcal{N}(0,1)$ dla każdego $v\in V$. Zatem możemy napisać

$$\mathbb{E}\left|\sum_{k=1}^{\infty}v_k\xi_k\right| \leqslant$$

$$\leqslant \sqrt{2\ln(NK)}\mathbb{E}\,\|v\| + \left(\mathbb{E}\,\|v\|^2\right)^{1/2}\left(\mathbb{E}\max_{u\in V}|\zeta_u|^2\mathbf{1}_{\{\max_{u\in V}|\zeta_u|>\sqrt{2\ln(NK)}\}}\right)^{1/2} \leqslant$$

$$\leqslant \sqrt{2\ln(NK)}\mathbb{E}\,\|v\| + \left(\mathbb{E}\,\|v\|^2\right)^{1/2}\left(\sum_{u\in V}\mathbb{E}|\zeta_u|^2\mathbf{1}_{\{|\zeta_u|>\sqrt{2\ln(NK)}\}}\right)^{1/2} =$$

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

Następnie korzystając z oszacowania pokazanego na początku dowodu z $a = \sqrt{2\ln(NK)}$ dostajemy, że

$$N\mathbb{E}|\zeta_v|^2\mathbf{1}_{\{|\zeta_v|>\sqrt{2\ln(NK)}\}} \leqslant \frac{2N}{\sqrt{2\pi}}\left(\sqrt{2\ln(NK)} + \frac{1}{\sqrt{2\ln(NK)}}\right)\frac{1}{NK} \leqslant$$

$$\leqslant \frac{1}{K}\left(\sqrt{2\ln(NK)} + \frac{1}{\sqrt{2\ln(NK)}}\right) \leqslant \frac{1}{K}\cdot 2\ln(NK),$$

o ile tylko zachodzi, że $NK \geqslant 2$. Przy założeniu na K jest to spełnione dla każdego nietrywialnego problemu z licznością $V > 1$. Łącząc te nierówności dostajemy, że

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

$$\leqslant \sqrt{2\ln(NK)}\mathbb{E}\,\|v\| + \sqrt{2\ln(NK)}\left(\mathbb{E}\,\|v\|^2\cdot\frac{1}{K}\right)^{1/2}$$

co kończy dowód. □

Dowód lematu 3. Zauważmy na początek, że z nierówności Markowa dostajemy, że dla dowolnego $t>0$ zachodzi, że $P(X>\epsilon)\leqslant e^{-t\epsilon}\mathbb{E}e^{tX}$. Policzymy pomocniczo następującą wartość oczekiwaną, dla dowolnego $a\in(0,1/2)$

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

Niech teraz $\eta_v = \frac{1}{\sqrt{2}\|v\|}\sum_{i=1}^{\infty}v_i(\xi_i^2-1)$. Wyrażenie to jest skończone na mocy lematu 9 oraz faktu, że $v\in l^2$, a zatem szereg $\sum_{i=1}^{\infty}v_i$ jest skończony. Korzystając z powyższych faktów możemy napisać dla dowolnego $t>0$ i dla ustalonego $u\in V$

$$P(\eta_u > x) \leqslant \exp(-tx)\mathbb{E}\exp(t\eta_u) = \exp(-tx)\mathbb{E}\exp\left(t\frac{1}{\sqrt{2}\,\|U\|}\sum_{i=1}^{\infty}U_i(\xi_i^2-1)\right) =$$

$$= \exp(-tx)\prod_{i=1}^{\infty}\mathbb{E}\exp\left(t\frac{1}{\sqrt{2}\,\|u\|}u_i(\xi_i^2-1)\right) =$$

$$= \exp(-tx)\prod_{i=1}^{\infty}\exp\left(-\frac{tu_i}{\sqrt{2}\,\|u\|}\right)\mathbb{E}\exp\left(\frac{t}{\sqrt{2}\,\|u\|}u_i\xi_i^2\right) =$$

$$= \exp(-tx)\prod_{i=1}^{\infty}\exp\left(-\frac{tu_i}{\sqrt{2}\,\|u\|}\right)\sqrt{\frac{\|u\|}{\|u\|-\sqrt{2}tu_i}} =$$

$$= \exp(-tx)\prod_{i=1}^{\infty}\exp\left(-\frac{tu_i}{\sqrt{2}\,\|u\|}-\frac{1}{2}\ln\left(1-\frac{\sqrt{2}tu_i}{\|u\|}\right)\right).$$

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 $t < \frac{1}{\sqrt{2}m(u)}$. Stąd

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

$$= \exp(-tx)\prod_{i=1}^{\infty}\exp\left(-\frac{tu_i}{\sqrt{2}\,\|u\|}+\frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k}\left(\frac{\sqrt{2}tu_i}{\|u\|}\right)^k\right) =$$

$$= \exp(-tx)\exp\left(\sum_{i=1}^{\infty}\sum_{k=2}^{\infty}\frac{1}{2k}\left(\frac{\sqrt{2}tu_i}{\|u\|}\right)^k\right) =$$

$$= \exp(-tx)\exp\left(\sum_{k=2}^{\infty}\sum_{i=1}^{\infty}\frac{1}{2k}\left(\frac{\sqrt{2}tu_i}{\|u\|}\right)^k\right) =$$

$$= \exp(-tx)\exp\left(\sum_{k=2}^{\infty}\frac{(\sqrt{2}t)^k}{2k}\sum_{i=1}^{\infty}\left(\frac{u_i}{\|u\|}\right)^2\left(\frac{u_i}{\|u\|}\right)^{k-2}\right) =$$

$$\leqslant \exp(-tx)\exp\left(\sum_{k=2}^{\infty}\frac{(\sqrt{2}t)^k}{2k}(m(u))^{k-2}\frac{\sum_{i=1}^{\infty}u_i^2}{\|u\|^2}\right) =$$

$$= \exp(-tx)\exp\left(\frac{1}{2m^2(u)}\sum_{k=2}^{\infty}\frac{1}{k}\left(\sqrt{2}tm(u)\right)^k\right) =$$

$$= \exp(-tx)\exp\left(-\frac{1}{2m^2(u)}\ln\left(1-\sqrt{2}tm(u)\right)-\frac{t}{\sqrt{2}m(u)}\right).$$

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

$$P(\eta_u > x) \leqslant \exp\left(\frac{1}{2m^2(u)}\ln\left(1+\sqrt{2}m(u)x\right)-\frac{x}{\sqrt{2}m(u)}\right) =$$

$$= \exp\left(\frac{1}{2m^2(u)}\left(\ln\left(1+\sqrt{2}m(u)x\right)-\sqrt{2}m(u)x\right)\right).$$

Zauważmy następnie, że zachodzi następująca zależność

$$\ln(1+z)-z = z\int_0^1\left(-\frac{tz}{1+tz}\right)dt \leqslant -\int_0^1\frac{tz^2}{1+z}dt = -\frac{z^2}{2(1+z)}.$$

Zatem powyższe oszacowanie sprowadza się do postaci

$$P(\eta_u > x) \leqslant \exp\left(-\frac{x^2}{2(1+\sqrt{2}m(u)x)}\right).$$

Zauważmy, że zachodzi również $P(\eta_u < -x) = P(-\eta_u > x) = P(\eta_{-u} > x)$ i powyższe rozumowanie przenosi się na ten przypadek, a stąd dostajemy oszacowanie

$$P(|\eta_u| > x) \leqslant 2\exp\left(-\frac{x^2}{2(1+\sqrt{2}m(u)x)}\right).$$

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

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

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

$$\mathbb{E}\eta_u^2\mathbf{1}\{|\eta_u| > Q\} = 2\int_Q^\infty tP(|\eta_u| > t)dt \leqslant 4\int_Q^\infty t\exp\left(-\frac{t^2}{2(1+\sqrt{2}m(u)t)}\right)dt \leqslant$$

$$\leqslant 4\int_Q^{\frac{1}{\sqrt{2}m(u)}} t\exp\left(-\frac{t^2}{4}\right)dt + 4\int_Q^\infty t\exp\left(\frac{-t}{\sqrt{32}m(u)}\right)dt \leqslant$$

$$\leqslant C\exp\left(-\frac{Q^2}{4}\right) + CQ\exp\left(-\frac{Q}{\sqrt{32}m(u)}\right),$$

Gdy $Q \leqslant \frac{1}{\sqrt{2}m(u)}$. Natomiast gdy $Q > \frac{1}{\sqrt{2}m(u)}$ całkę $4\int_Q^\infty t\exp\left(-\frac{t^2}{2(1+\sqrt{2}m(u)t)}\right)dt$ można oszacować przez $CQ\exp\left(-\frac{Q}{\sqrt{32}m(u)}\right)$, co z uwagi na to, że $C\exp\left(-\frac{Q^2}{4}\right) \geqslant 0$ prowadzi do tego samego oszacowania. Następnie możemy zapisać, że

$$\sum_{u\in V}\exp\left(-\frac{Q}{\sqrt{32}m(u)}\right) = \sum_{u\in V}\exp\left(-\frac{q}{m(u)}\right)\exp\left(-\frac{Q/\sqrt{32}-q}{m(u)}\right) \leqslant$$

$$\leqslant M(q)\exp\left(-\frac{Q/\sqrt{32}-q}{m_V}\right),$$

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

$$\mathbb{E}\left|\sum_{i=1}^\infty v_i(\xi_i^2 - 1)\right| = \mathbb{E}\left|\sqrt{2}\,\|v\|\,|\eta_v|\right| \leqslant \sqrt{2}\mathbb{E}\,\|v\|\max_{v\in V}|\eta_v| \leqslant$$

$$\leqslant \sqrt{2}\mathbb{E}\,\|v\|\max_{u\in V}|\eta_u|\mathbf{1}\{\max_{u\in V}|\eta_u| \leqslant Q\} + \sqrt{2}\mathbb{E}\,\|v\|\max_{u\in V}|\eta_u|\mathbf{1}\{\max_{u\in V}|\eta_u| > Q\} \leqslant$$

$$\leqslant \sqrt{2}Q\mathbb{E}\,\|v\| + \sqrt{2}\mathbb{E}\,\|v\|\max_{u\in V}|\eta_u|\mathbf{1}\{\max_{u\in V}|\eta_u| > Q\} \leqslant$$

$$\leqslant \sqrt{2}Q\mathbb{E}\,\|v\| + \left(2\mathbb{E}\,\|v\|^2\right)^{1/2}\left(\mathbb{E}\max_{u\in V}|\eta_u|^2\mathbf{1}\{\max_{u\in V}|\eta_u| > Q\}\right)^{1/2} \leqslant$$

$$\leqslant \sqrt{2}Q\mathbb{E}\,\|v\| + \left(2\mathbb{E}\,\|v\|^2\right)^{1/2}\left(\sum_{u\in V}\mathbb{E}|\eta_u|^2\mathbf{1}\{\max_{u\in V}|\eta_u| > Q\}\right)^{1/2} \leqslant$$

$$\leqslant \sqrt{2}Q\mathbb{E}\,\|v\| + \left(2\mathbb{E}\,\|v\|^2\right)^{1/2}\left(\sum_{u\in V}\left(C\exp\left(-\frac{Q^2}{4}\right) + CQ\exp\left(-\frac{Q}{\sqrt{32}m(u)}\right)\right)\right)^{1/2} \leqslant$$

$$\leqslant \sqrt{2}Q\mathbb{E}\,\|v\| + \left(2\mathbb{E}\,\|v\|^2\right)^{1/2}\left(NC\exp\left(-\frac{Q^2}{4}\right) + CQM(q)\exp\left(-\frac{Q/\sqrt{32}-q}{m_V}\right)\right)^{1/2}.$$

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

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

$$\Big(NC \exp\left(-\ln(NK)\right) \exp\Big(-8m_V^2 \ln^2(M(q)K) - 8q - 8\sqrt{2}q\sqrt{\ln(NK)} - 32qm_V \ln(M(q)K) -$$

$$-8\sqrt{2}\sqrt{\ln(NK)}m_V \ln(M(q)K)\Big) + CM(q)\left(2\sqrt{\ln(NK)} + \sqrt{32}m_V \ln(M(q)K) + q\sqrt{32}\right)$$

$$\exp\left(-\ln(M(q)K)\right) \exp\left(-\frac{\sqrt{\ln(NK)}}{2\sqrt{2}m_V}\right)\Big)^{1/2} \leqslant$$

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

dla pewnej stałej D zależnej tylko od q i operatora T.
Nierówność ta kończy dowód lematu. □

Dowód lematu 4. Przypomnijmy, że w naszym modelu (2) mamy $x_i = \theta_i + \epsilon\sigma_i\xi_i$, 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

$$\mathbb{E}_\theta \left\|\hat{\theta} - \theta\right\|^2 = \mathbb{E}_\theta\left[\sum_{i=1}^{\infty} (\lambda_i(X)x_i - \theta_i)^2\right] =$$

$$= \mathbb{E}_\theta\left[\sum_{i=1}^{\infty} \lambda_i^2(X)x_i^2 + \sum_{i=1}^{\infty} \theta_i^2 - 2\sum_{i=1}^{\infty} \lambda_i(X)x_i\theta_i\right] =$$

$$= \mathbb{E}_\theta\left[\sum_{i=1}^{\infty} \lambda_i^2(X)\left(\theta_i^2 + \epsilon^2\sigma_i^2\xi_i^2 + 2\epsilon\theta_i\sigma_i\xi_i\right) + \sum_{i=1}^{\infty}\theta_i^2 - 2\sum_{i=1}^{\infty}\lambda_i(X)\theta_i\left(\theta_i + \epsilon\sigma_i\xi_i\right)\right] =$$

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

$$+ \epsilon^2\sum_{i=1}^{\infty}\lambda_i^2(X)\sigma_i^2(\xi_i^2 - 1) + \epsilon^2\sum_{i=1}^{\infty}\lambda_i^2(X)\sigma_i^2\Big] =$$

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

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

$$\epsilon\mathbb{E}_\theta\left|\sum_{i=1}^{\infty}(1 - \lambda_i(X))\theta_i\lambda_i(X)\sigma_i\xi_i\right| \leqslant$$

$$\leqslant \epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}(1 - \lambda_i(X))^2\theta_i^2\lambda_i^2(X)\sigma_i^2\right)^{1/2} +$$

$$+2\epsilon\sqrt{\ln(NS)}S^{-1/2}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}(1 - \lambda_i(X))^2\theta_i^2\lambda_i^2(X)\sigma_i^2\right)^{1/2} \leqslant$$

$$\leqslant \epsilon\sqrt{2\ln(NS)}\mathbb{E}_\theta\sup_i \sigma_i|\lambda_i(X)|\left(\sum_{i=1}^{\infty}(1 - \lambda_i(X))^2\theta_i^2\right)^{1/2} +$$

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

$$\leqslant \frac{1}{2}\epsilon^2 B\ln(NS)\mathbb{E}_\theta \sup_i \sigma_i^2\lambda_i^2(X) + B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))^2\theta_i^2 +$$

$$+\epsilon^2 B\ln(NS)\frac{\max_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2}{S} + B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))^2\theta_i^2,$$

gdzie skorzystaliśmy z nierówności $2ab \leqslant Ba^2 + B^{-1}b^2$ spełnionej dla dowolnego $B>0$. 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

$$2B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))^2\theta_i^2+\frac{1}{2}\epsilon^2 B\ln(NS)\mathbb{E}_\theta\sup_i\sigma_i^2\lambda_i^2(X)+\epsilon^2 B\ln(NS)\min_{\lambda\in\Lambda}\sup_i\sigma_i^2\lambda_i^2 \leqslant$$

$$\leqslant 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))^2\theta_i^2 + \frac{1}{2}B\mathbb{E}_\theta\left(\epsilon^2\ln(NS)\sup_i\sigma_i^2\lambda_i^2(X)\right)+$$

$$+B\mathbb{E}_\theta\left(\epsilon^2\ln(NS)\sup_i\sigma_i^2\lambda_i^2(X)\right)\leqslant$$

$$\leqslant 2B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))^2\theta_i^2 + \frac{3}{2}B\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)].$$

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

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

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

$$\epsilon^2\mathbb{E}_\theta\left|\sum_{i=1}^{\infty}\lambda_i^2(X)\sigma_i^2(\xi_i^2-1)\right|\leqslant$$

$$\leqslant\epsilon^2 D\left(\sqrt{\ln(NS)}+m_V\ln(M(1)S)\right)\left(\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2}+\left(\frac{1}{S}\mathbb{E}_\theta\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2}\right)\leqslant$$

$$\leqslant\epsilon^2 D\sqrt{2L_\Lambda}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2}+\frac{\epsilon^2 D\sqrt{2L_\Lambda}}{\sqrt{S}}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2}.$$

Zauważmy, że zachodzi następujący związek

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

$$\leqslant\frac{\min_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2}{\max_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2}\max_{\lambda\in\Lambda}\sup_i\lambda_i^2\sigma_i^2\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X)\leqslant$$

$$\leqslant\mathbb{E}_\theta\sup_i\sigma_i^2\lambda_i^2(X)\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X).$$

Korzystając z tego faktu i ponownie z nierówności $2ab\leqslant Ba^2+B^{-1}b^2$ z $B>0$ dostajemy

$$\epsilon^2 D\sqrt{2L_\Lambda}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2}+\frac{\epsilon^2 D\sqrt{2L_\Lambda}}{\sqrt{S}}\left(\mathbb{E}_\theta\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2}\leqslant$$

$$\leqslant \epsilon^2 D\sqrt{2L_\Lambda}\mathbb{E}_\theta\left(\sum_{i=1}^{\infty}\lambda_i^4(X)\sigma_i^4\right)^{1/2} + \epsilon^2 D\sqrt{2L_\Lambda}\left(\mathbb{E}_\theta \sup_i \sigma_i^2\lambda_i^2(X)\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X)\right)^{1/2} \leqslant$$

$$\leqslant \mathbb{E}_\theta\frac{\epsilon^2 BD^2 L_\Lambda}{4}\sup_i \sigma_i^2\lambda_i^2(X) + 2\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X)+$$

$$+\mathbb{E}_\theta\frac{\epsilon^2 BD^2 L_\Lambda}{4}\sup_i \sigma_i^2\lambda_i^2(X) + 2\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X) =$$

$$= 4\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X) + \frac{\epsilon^2 BD^2}{2}\mathbb{E}_\theta L_\Lambda \sup_i \sigma_i^2\lambda_i^2(X) \leqslant$$

$$\leqslant 4\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X) + \frac{BD^2}{2}\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)].$$

Łącząc te oszacowania dostajemy, że

$$\mathbb{E}_\theta\left\|\hat{\theta}-\theta\right\|^2 \leqslant \mathbb{E}_\theta\mathcal{R}(\hat{\theta},\theta) - 2\epsilon\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))\theta_i\lambda_i(X)\sigma_i\xi_i + \epsilon^2\mathbb{E}_\theta\sum_{i=1}^{\infty}\lambda_i^2(X)\sigma_i^2(\xi_i^2-1) \leqslant$$

$$\leqslant \mathbb{E}_\theta\mathcal{R}(\hat{\theta},\theta) + 4B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}(1-\lambda_i(X))^2\theta_i^2 + \frac{3}{2}B\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)]+$$

$$+4\epsilon^2 B^{-1}\mathbb{E}_\theta\sum_{i=1}^{\infty}\sigma_i^2\lambda_i^2(X) + \frac{BD^2}{2}\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)] =$$

$$= (1+4B^{-1})\mathbb{E}_\theta\mathcal{R}(\hat{\theta},\theta) + CB\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)],$$

co kończy dowód lematu. □

Dowód lematu 5. W dowodzie lematu ponownie skorzystamy z oszacowania

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

zachodzącego dla zmiennych losowych o standardowym rozkładzie normalnym oraz z własności

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

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

Oznaczmy przez $\zeta_v = \frac{\langle\eta,v\rangle}{\|T^*v\|_2}$. Zauważmy, że z faktu, że η jest gaussowskim szumem z operatorem kowariancji postaci TT^*, może być przedstawiona jako $T\xi$, gdzie ξ jest gaussowskim białym szumem. Wynika stąd, że ζ_v ma standardowy rozkład normalny $\mathcal{N}(0,1)$. Możemy teraz napisać, że

$$\mathbb{E}\left|\langle\eta,v\rangle\right| = \mathbb{E}|\zeta|\,\|T^*v\|_2 \leqslant \mathbb{E}|\zeta|\,\|T^*\|\,\|v\|_2 \leqslant \|T^*\|\,\mathbb{E}\,\|v\|_2\max_{v\in V}|\zeta_v| =$$

$$= \|T^*\|\,\mathbb{E}\,\|v\|_2\max_{v\in V}|\zeta_v|\mathbf{1}_{\{\max_{v\in V}|\zeta_v|\leqslant\|T^*\|\sqrt{2\ln(NK)}\}} + \|T^*\|\,\mathbb{E}\,\|v\|_2\max_{v\in V}|\zeta_v|\mathbf{1}_{\{\max_{v\in V}|\zeta_v|>\sqrt{2\ln(NK)}\}} \leqslant$$

$$\leqslant \|T^*\|\sqrt{2\ln(NK)}\mathbb{E}\,\|v\|_2 + \|T^*\|\,\mathbb{E}\,\|v\|_2\max_{v\in V}|\zeta_v|\mathbf{1}_{\{\max_{v\in V}|\zeta_v|>\sqrt{2\ln(NK)}\}} \leqslant$$

$$\leqslant \|T^*\|\sqrt{2\ln(NK)}\mathbb{E}\,\|v\|_2 + \|T^*\|\left(\mathbb{E}\,\|v\|_2^2\right)^{1/2}\left(\mathbb{E}\max_{v\in V}|\zeta_v|^2\mathbf{1}_{\{\max_{v\in V}|\zeta_v|>\sqrt{2\ln(NK)}\}}\right)^{1/2} \leqslant$$

$$\leqslant \|T^*\|\sqrt{2\ln(NK)}\mathbb{E}\,\|v\|_2 + \|T^*\|\left(\mathbb{E}\,\|v\|_2^2\right)^{1/2}\left(\sum_{v\in V}\mathbb{E}|\zeta_v|^2\mathbf{1}_{\{|\zeta_v|>\sqrt{2\ln(NK)}\}}\right)^{1/2} =$$

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

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

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

$$\leqslant \|T^*\| \sqrt{2\ln(NK)} \left(\mathbb{E}||v||_2 + \sqrt{2\mathbb{E}||v||_2^2/K}\right) = \|T\| \sqrt{2\ln(NK)} \left(\mathbb{E}||v||_2 + \sqrt{2\mathbb{E}||v||_2^2/K}\right),$$

gdyż $\|T^*\| = \|T\|$, co kończy dowód. □

Dowód lematu 6. Jeżeli η jest gaussowskim szumem z operatorem kowariancji TT^* to można go przedstawić w postaci $T\xi$, gdzie ξ jest gaussowskim białym szumem. Oznaczmy $\zeta_v = \frac{1}{\|T^*v\|_2^2}\langle \eta^2 - 1, v\rangle$. Zauważmy, że tak zdefiniowana zmienna losowa spełnia założenia lematu 10 z $v = \frac{v}{\|T^*v\|_2}$. Z nierówności Markowa dostajemy, że dla dowolnego $t > 0$ i $x > 0$ zachodzi i ustalonego $u \in V$

$$\ln P(\zeta_u > x) \leqslant -tx + \ln \mathbb{E} \exp\left(\frac{t}{\|T^*u\|_2^2}\langle \eta^2 - 1, u\rangle\right) \leqslant$$

$$\leqslant -tx + \frac{\left(\frac{t}{\|T^*u\|}\right)^2 \|u\|_2^2}{1 - \frac{t}{\|T^*u\|_2}\|u\|_\infty} \leqslant -tx + \frac{\left(\frac{t}{\|T^*u\|_2}\right)^2 \|u\|_2^2}{1 - \frac{t}{\|T^*u\|_2}\|u\|_2\, m(u)}.$$

Zauważmy następnie, że dla dowolnego $a > 0$ i $t \in [0, \frac{1}{\sqrt{2}a})$ zachodzi nierówność

$$\frac{t^2}{1 - at} \leqslant \frac{-1}{a^2}\ln(1 - \sqrt{2}at) - \frac{\sqrt{2}t}{a}.$$

Dostajemy zatem następującą nierówność

$$P(\zeta_u > x) \leqslant \exp(-tx)\exp\left(-\frac{1}{m^2(u)}\ln\left(1 - \frac{\sqrt{2}m(u)\,\|u\|_2}{\|T^*u\|_2}t\right) - \frac{\sqrt{2}\,\|u\|_2}{\|T^*u\|_2\, m(u)}t\right),$$

o ile $t < \frac{\|T^*u\|}{\sqrt{2}m(u)\|u\|_2}$. Zauważmy, że nierówność ta sprowadza się do analogicznej nierówności z dowodu lematu 3, gdy operator T jest operatorem identycznościowym, jak w przypadku białego szumu. Minimalizując prawą stronę powyższego wyrażenia względem t dostajemy, że w punkcie $t_{min} = \frac{\|T^*u\|}{\sqrt{2}m(u)\|u\|_2} - \frac{1}{m^2(u)x + \frac{\sqrt{2}m(u)\|u\|_2}{\|T^*u\|}}$ osiągane jest minimum o wartości

$$\exp\left(-\frac{\|T^*u\|_2\, x}{\sqrt{2}m(u)\,\|u\|_2} + \frac{1}{m^2(u)}\ln\left(1 + \frac{\|T^*u\|_2\, m(u)x}{\sqrt{2}\,\|u\|}\right)\right) =$$

$$= \exp\left(\frac{1}{m^2(u)}\left(\ln\left(1 + \frac{\|T^*u\|_2\, m(u)}{\sqrt{2}\,\|u\|_2}x\right) - \frac{\|T^*u\|_2\, m(u)}{\sqrt{2}\,\|u\|_2}\right)\right) \leqslant$$

$$\leqslant \exp\left(\frac{-\frac{\|T^*u\|_2^2}{\|u\|_2^2}x^2}{1 + \frac{\|T^*u\|m(u)}{\sqrt{2}\|u\|_2}x}\right), \tag{21}$$

gdzie ponownie skorzystaliśmy z nierówności $\ln(1+z) - z \leqslant -\frac{z^2}{2(1+z)}$ zachodzącej dla dodatnich z. Powyższe rozważania będą wykorzystane tylko w przypadku, gdy operator T^* jest operatorem ograniczonym i odwracalnym, co gwarantuje istnienie stałych $c, C > 0$ takich, że dla dowolnego v zachodzi ([16], str. 216)

$$c\,\|u\|_2 \leqslant \|Tu\|_2 \leqslant C\,\|u\|_2\,.$$

W dalszym ciągu rozważań największą ze stałych c dla których zachodzi powyższa nierówność będzie oznaczać przez $\|t\|$, natomiast najmniejszą ze stałych

C przez $\|T\|$. Oczywiście stała $\|T\|$ jest normą operatora T. Przy takich oznaczeniach oszacowanie (21) przyjmuje postać

$$\exp\left(\frac{-\|t\|^2 x^2}{1+2^{-1/2}\|T\|\, m(u)x}\right).$$

Zauważmy, że powyższe wyrażenie można oszacować następująco w zależności od wielkości x.

$$\exp\left(\frac{-\|t\|^2 x^2}{1+2^{-1/2}\|T\|\, m(u)x}\right) \leqslant \begin{cases} \frac{-\|t\|^2x^2}{2}, \ gdy\ 2^{-1/2}\|T\|\, m(u)x < 1, \\ \frac{-\|t\|^2 x}{\sqrt{2}\|T\| m(u)}, \ gdy\ 2^{-1/2}\|T\|\, m(u)x \geqslant 1. \end{cases}$$

Ostatecznie prowadzi nas to do następującego oszacowania

$$P(\zeta_u > x) \leqslant \begin{cases} \frac{-\|t\|^2x^2}{2}, \ gdy\ 2^{-1/2}\|T\|\, m(u)x < 1, \\ \frac{-\|t\|^2 x}{\sqrt{2}\|T\| m(u)}, \ gdy\ 2^{-1/2}\|T\|\, m(u)x \geqslant 1. \end{cases}$$

Zauważmy, że zachodzi również $P(\zeta_u < -x) = P(-\zeta_u > x) = P(\zeta_{-u} > x)$ i powyższe rozumowanie przenosi się na ten przypadek, a stąd dostajemy oszacowanie

$$P(|\zeta_u| > x) \leqslant 2 \begin{cases} \frac{-\|t\|^2x^2}{2}, \ gdy\ 2^{-1/2}\|T\|\, m(u)x < 1, \\ \frac{-\|t\|^2 x}{\sqrt{2}\|T\| m(u)}, \ gdy\ 2^{-1/2}\|T\|\, m(u)x \geqslant 1. \end{cases}$$

Następnie korzystając z uzyskanego oszacowania możemy zapisać dla dowolnego $\frac{\sqrt{2}}{\|T\| m(u)} \geqslant Q > 0$

$$\mathbb{E}\zeta_u^2 \mathbf{1}\{|\zeta_u| > Q\} = 2\int_Q^\infty zP\left(|\zeta_u| > z\right) dz \leqslant 4\int_Q^\infty z\exp\left(\frac{-\|t\|^2 z^2}{1+2^{-1/2}\|T\|\, m(u)z}\right) dt \leqslant$$

$$\leqslant 4\int_Q^{\frac{\sqrt{2}}{\|T\| m(u)}} z\exp\left(\frac{-\|t\|^2 z^2}{2}\right) dz + 4\int_Q^\infty z\exp\left(\frac{-\|t\|^2 z}{\sqrt{2}\|T\|\, m(u)}\right) dz \leqslant$$

$$\leqslant C\exp\left(\frac{-\|t\|^2 Q^2}{2}\right) + CQ\exp\left(\frac{-\|t\|^2 Q}{\sqrt{2}\|T\|\, m(u)}\right).$$

Podobnie jak w dowodzie lematu 3 nierówność tą możemy uogólnić dla dowolnego $Q > 0$. Zauważmy następnie, że zachodzi

$$\sum_{u\in V}\exp\left(\frac{-\|t\|^2 Q}{\sqrt{2}\|T\|\, m(u)}\right) = \sum_{u\in V}\exp\left(-\frac{q}{m(u)}\right)\exp\left(-\frac{\|t\|\, Q/(\sqrt{2}\|T\|) - q}{m(u)}\right) \leqslant$$

$$\leqslant M(q)\exp\left(-\frac{\|t\|\, Q/(\sqrt{2}\|T\|) - q}{m_V}\right),$$

o ile $Q > \frac{\sqrt{2}q\|T\|}{\|t\|}$. Niech zatem $Q > \frac{\sqrt{2}q\|T\|}{\|t\|}$, wtedy

$$\mathbb{E}\left|\langle \eta^2 - 1, u\rangle\right| = \mathbb{E}\|T^* u\|_2^2 |\zeta_u| \leqslant$$

$$\leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 |\zeta_u| \leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 \max_{u\in V}|\zeta_u| \leqslant$$

$$\leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 \max_{u\in V}|\zeta_u| \mathbf{1}\{\max_{u\in V}|\zeta_u| \leqslant Q\} + \|T\|^2 \mathbb{E}\|u\|_2^2 \max_{u\in V}|\zeta_u| \mathbf{1}\{\max_{u\in V}|\zeta_u| > Q\} \leqslant$$

$$\leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 \left(\mathbb{E}\|v\|_2^4\right)^{1/2} \left(\mathbb{E}\max_{u\in V}|\zeta_u|^2 \mathbf{1}\{\max_{u\in V}|\zeta_u| > Q\}\right)^{1/2} \leqslant$$

$$\leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 \left(\mathbb{E}\|v\|_2^4\right)^{1/2} \left(\sum_{u\in V}\mathbb{E}|\zeta_u|^2 \mathbf{1}\{\max_{u\in V}|\zeta_u| > Q\}\right)^{1/2} \leqslant$$

$$\leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 \left(\mathbb{E}\|v\|_2^4\right)^{1/2} \left(\sum_{u\in V} \left(C \exp\left(\frac{-\|t\|^2 Q^2}{2}\right) + CQ\exp\left(\frac{-\|t\|^2 Q}{\sqrt{2}\|T\| m(u)}\right)\right)\right)^{1/2} \leqslant$$

$$\leqslant \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 \left(\mathbb{E}\|v\|_2^4\right)^{1/2} \left(NC\exp\left(\frac{-\|t\|^2 Q^2}{2}\right) + CQM(q)\exp\left(-\frac{\|t\| Q/(\sqrt{2}\|T\|) - q}{m_V}\right)\right)^{1/2}$$

Przyjmując teraz $Q = \frac{1}{\|v\|}\left(\frac{\sqrt{2}}{\|t\|^2}\sqrt{\ln(NK)} + \frac{\sqrt{2}\|T\|}{\|t\|} m_V \ln(M(q)K)\right) + \frac{\sqrt{2}q\|T\|}{\|t\|}$ dostajemy tezę. □

Dowód lematu 7. Oznaczmy przez $C' = \max\{1, C_2\}$. Zauważmy, że z faktu istnienia stałej C_2 wynika

$$\int_S \lambda\sigma^2 d\mu \leqslant \int_S \frac{\lambda^2\sigma^4}{\lambda\sigma^2} d\mu \leqslant C_2 \int_S \frac{\lambda^4\sigma^3}{\lambda\sigma^2} d\mu = C_2 \int_S \lambda^3\sigma d\mu \leqslant C_2 \int_S \lambda^2\sigma d\mu.$$

Przypomnijmy, że w naszym modelu (16) mamy, że $X = \theta + \epsilon\sigma\eta$, a stąd $X^2 = \theta^2 + \epsilon^2\sigma^2\eta^2 + 2\epsilon\theta\sigma\eta$.

$$\mathbb{E}_\theta \left\|\hat{\theta} - \theta\right\|^2 = \mathbb{E}_\theta\left[\int_S (\lambda(X)X - \theta)^2 d\mu\right] =$$

$$= \mathbb{E}_\theta\left[\int_S \lambda^2(X)X^2 d\mu + \int_S \theta^2 d\mu - 2\int_S \lambda(X)X\theta d\mu\right] =$$

$$= \mathbb{E}_\theta\left[\int_S \lambda^2(X)\left(\theta^2 + \epsilon^2\sigma^2\eta^2 + 2\epsilon\theta\sigma\eta\right) d\mu + \int_S \theta^2 d\mu - 2\int_S \lambda(X)\theta\left(\theta + \epsilon\sigma\eta\right) d\mu\right] =$$

$$= \mathbb{E}_\theta\left[\int_S (1-\lambda(X))^2\theta^2 d\mu - 2\epsilon\int_S (1-\lambda(X))\theta\lambda(X)\sigma\eta d\mu + \right.$$

$$\left. + \epsilon^2\int_S \lambda^2(X)\sigma^2(\eta^2 - 1) d\mu + \epsilon^2\int_S \lambda^2(X)\sigma^2 d\mu\right] \leqslant$$

$$\leqslant \mathbb{E}_\theta C'\Psi(\lambda,\theta) - 2\epsilon\mathbb{E}_\theta\int_S (1-\lambda(X))\theta\lambda(X)\sigma\eta d\mu + \epsilon^2\mathbb{E}_\theta\int_S \lambda^2(X)\sigma^2(\eta^2-1) d\mu.$$

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

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

$$\leqslant \epsilon\|T\|\sqrt{2\ln(NS)}\mathbb{E}_\theta\left(\int_S (1-\lambda(X))^2\theta^2\lambda^2(X)\sigma^2 d\mu\right)^{1/2} +$$

$$+2\epsilon\|T\|\sqrt{\ln(NS)}S^{-1/2}\left(\mathbb{E}_\theta\int_S (1-\lambda(X))^2\theta^2\lambda^2(X)\sigma^2 d\mu\right)^{1/2} \leqslant$$

$$\leqslant \epsilon\|T\|\sqrt{2\ln(NS)}\mathbb{E}_\theta\left\|\sigma^2\lambda^2(X)\right\|_\infty\left(\int_S (1-\lambda(X))^2\theta^2 d\mu\right)^{1/2} +$$

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

$$\leqslant \frac{1}{2C'}\epsilon^2\|T\|^2 B\ln(NS)\mathbb{E}_\theta\left\|\sigma^2\lambda^2(X)\right\|_\infty + B^{-1}C'\mathbb{E}_\theta\int_S (1-\lambda(X))^2\theta^2 d\mu +$$

$$+\epsilon^2\frac{1}{C'}\|T\|^2 B\ln(NS)\frac{\max_{\lambda\in\Lambda}\left\|\sigma^2\lambda^2(X)\right\|_\infty}{S} + B^{-1}C'\mathbb{E}_\theta\int_S (1-\lambda(X))^2\theta^2 d\mu,$$

gdzie korzystaliśmy z nierówności $2ab \leqslant C'B^{-1}a^2 + BC'^{-1}b^2$ zachodzącej dla dowolnego $B > 0$.

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

$$2B^{-1}C'\mathbb{E}_\theta\int_S(1-\lambda(X))^2\theta^2d\mu+\frac{1}{2C'}\epsilon^2\left\|T\right\|^2 B\ln(NS)\mathbb{E}_\theta\left\|\sigma^2\lambda^2(X)\right\|_\infty+\epsilon^2B\frac{1}{C'}\left\|T\right\|^2\ln(NS)\min_{\lambda\in\Lambda}\left\|\sigma^2\lambda^2(X)\right\|_\infty$$

$$\leqslant 2B^{-1}C'\mathbb{E}_\theta\int_S(1-\lambda(X))^2\theta^2d\mu+\frac{3}{2C'}B\left\|T\right\|^2\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)].$$

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

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

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

$$\epsilon^2\mathbb{E}_\theta\left|\int_S\lambda^2(X)\sigma^2(\eta^2-1)d\mu\right|\leqslant$$

$$\leqslant\epsilon^2D\sqrt{C_2}\left\|T\right\|^2\sqrt{2L_\Lambda}\mathbb{E}_\theta\left(\int_S\lambda^4(X)\sigma^3d\mu\right)^{1/2}+\frac{\epsilon^2D\sqrt{C_2}\left\|T\right\|^2\sqrt{2L_\Lambda}}{\sqrt{S}}\left(\mathbb{E}_\theta\int_S\lambda^4(X)\sigma^3\mu\right)^{1/2}.$$

Analogicznie do poprzednich rozważań zachodzi

$$S^{-1}\mathbb{E}_\theta\int_S\sigma^3\lambda^4(X)d\mu\leqslant$$

$$\leqslant\mathbb{E}_\theta\left\|\sigma^2\lambda^2(X)\right\|_\infty\mathbb{E}_\theta\int_S\sigma\lambda^2(X)d\mu.$$

Dostajemy stąd następujące oszacowanie

$$\epsilon^2D\sqrt{C_2}\left\|T\right\|^2\sqrt{2L_\Lambda}\mathbb{E}_\theta\left(\int_S\lambda^4(X)\sigma^3d\mu\right)^{1/2}+\frac{\epsilon^2D\sqrt{C_2}\left\|T\right\|^2\sqrt{2L_\Lambda}}{\sqrt{S}}\left(\mathbb{E}_\theta\int_S\lambda^4(X)\sigma^3\mu\right)^{1/2}\leqslant$$

$$\leqslant\epsilon^2D\sqrt{C_2}\left\|T\right\|^2\sqrt{2L_\Lambda}\mathbb{E}_\theta\left(\int_S\lambda^4(X)\sigma^3d\mu\right)^{1/2}+\epsilon^2D\sqrt{C_2}\left\|T\right\|^2\sqrt{2L_\Lambda}\left(\mathbb{E}_\theta\left\|\sigma^2\lambda^2(X)\right\|_\infty\mathbb{E}_\theta\int_S\sigma\lambda^2(X)d\mu\right)^{1/2}$$

$$\leqslant\mathbb{E}_\theta\frac{\epsilon^2BD^2C_2\left\|T\right\|^4L_\Lambda}{4C'}\left\|\sigma^2\lambda^2(X)\right\|_\infty+2\epsilon^2B^{-1}C'\left\|T\right\|^2\mathbb{E}_\theta\int_S\sigma\lambda^2(X)d\mu+$$

$$+\mathbb{E}_\theta\frac{\epsilon^2BD^2C_2\left\|T\right\|^4L_\Lambda}{4C'}\left\|\sigma^2\lambda^2(X)\right\|_\infty+2\epsilon^2B^{-1}C'\left\|T\right\|^2\mathbb{E}_\theta\int_S\sigma\lambda^2(X)d\mu=$$

$$=4\epsilon^2B^{-1}C'\left\|T\right\|^2\mathbb{E}_\theta\int_S\sigma\lambda^2(X)d\mu+\frac{\epsilon^2BD^2C_2\left\|T\right\|^2}{2C'}\mathbb{E}_\theta L_\Lambda\left\|\sigma^2\lambda^2(X)\right\|_\infty\leqslant$$

$$\leqslant4\epsilon^2B^{-1}C'\left\|T\right\|^2\mathbb{E}_\theta\int_S\sigma\lambda^2(X)d\mu+\frac{BD^2C_2\left\|T\right\|^2}{2C'}\mathbb{E}_\theta\Delta^\epsilon[\lambda(X)].$$

Łącząc te oszacowania dostajemy tezę lematu. □

Literatura

[1] P. Alquier, E. Gautier, G. Stoltz, *Inverse Problems and High-Dimensional Estimation* Springer-Verlag, 2011, wydanie zbiorowe,

[2] A. Barron, L. Birge, P. Massart, *Risk bounds for model selection via penalization*, Probab. Theory Relat. Fields, 113, 1999, pp. 301– 413,

[3] M. Beśka, *Wykład monograficzny. Dodatek*,

[4] L. Birge, *Model selection via testing: an alternative to (penalized) maximum likelihood estimators*, Ann. I. H. Poincaré, PR 42, 2006, pp. 273–325,

[5] L. Birge, *Statistical estimation with model selection*, arXiv, 2006, The Brouwer Lecture, 2005,

[6] N. Bissantz, T. Hohange, A. Munk, F. Ruymgaart, *Convergence rates of general regularization methods for statistical inverse problems and applications*, SIAM J. Numer. Anal., Vol. 45, No. 6, 2007, pp. 2610-2636,

[7] L. Cavalier, *Inverse problems with non-compact operators*, Journal of Statistical Planning and Inference, 136, 2006, pp. 390– 400,

[8] 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,

[9] W. Feller, *An Introduction to Probability Theory and Its Applications. Volume 1*, John Wiley and Sons, 1968,

[10] C. Giraud, *Introduction to High– Dimensional Statistics*, CRC Press, 2015,

[11] P. R. Halmos, *What does the spectral theorem say?*, The American Mathematical Monthly, Vol. 70, No. 3, 1963, pp. 241-247,

[12] T. Hida, *Brownian Motion*, Springer, 1980,

[13] B.Laurent, P. Massart, *Adaptive estimation of a quadratic functional by model selection*, The Annals of Statistics, Vol. 28, No. 5, 2000, pp. 1302– 1338,

[14] J.– M. Loubes, C. Ludena, *Penalized estimators for non linear inverse problems*, ESAIM: PS Vol. 14, 2010, pp. 173– 191,

[15] J.– M. Loubes, C. Ludena, *Adaptive complexity regularization for linear inverse problems*, Electronic Journal of Statistics, Vol. 2, 2008, pp. 661– 677,

[16] L. A. Lusternik, V. J. Sobolew, *Elements of functional analysis*, John Wiley and Sons, 1974,

[17] B. A. Mair, F. H. Ruymgaart, *Statistical inverse estimation in Hilbert scales*, SIAM J. Appl. Math, Vol. 56, No. 5, 1996, pp. 1424– 1444,

[18] J. Kaipio, E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2004,

[19] C. Mitchell, S. van de Geer, *General oracle inequalities for model selection*, Electron. J. Statist, 3 (2009), pp. 176-204,

[20] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Springer-Science+Business Media, B.Y., 1986,

[21] Z. Szkutnik, *Statystyczne problemy odwrotne*, notatki do wykładu,

[22] M. E. Taylor, *Partial Differential Equations II. Qualitative Studies of Linear Equations*, Springer Science+Business Media, 2011,

[23] N. N. Vakhania, V. I. Tarieladze, *Probability Distributions on Banach Spaces*, D. Reidel Publishing Company, 1987,

[24] H. Lal Vasudeva, *Elements of Hilbert Spaces and Operator Theory*, Springer Verlag, 2017,

[25] L. Wasserman, *All of Nonparametric Statistics*, Springer Science+Business Media, Inc., 2006,