



Akademia Górniczo-Hutnicza
im. Stanisława Staszica
w Krakowie

Praca magisterska

Nierówności wyrocznie dla problemów odwrotnych

Grzegorz Mika

Kierunek: Matematyka
Specjalności: Matematyka finansowa,
Matematyka ubezpieczeniowa

Nr albumu: 267543	Promotor prof. dr hab. inż. Zbigniew Szkutnik
-------------------	--



Wydział Matematyki Stosowanej

Kraków 2018

Oświadczenie autora

Ja, niżej podpisany Grzegorz Mika oświadczam, że praca ta została napisana samodzielnie i wykorzystywała (poza zdobytą na studiach wiedzę) jedynie wyniki prac zamieszczonych w spisie literatury.

.....
(Podpis autora)

Oświadczenie promotora

Oświadczam, że praca spełnia wymogi stawiane pracom magisterskim.

.....
(Podpis promotora)

*Składam serdeczne podziękowania mojemu Promotorowi prof. dr hab. inż.
Zbigniewowi Szkutnikowi za wskazanie mi drogi, wszelką pomoc i cierpliwość.*

Spis treści

Streszczenie	3
Abstract	4
Wstęp	5
1. Pojęcia wstępne	7
2. Nierówności wyrocznie w modelu z operatorem zawartym	16
3. Nierówności wyrocznie w ogólnym modelu	29
4. Przykład	43
5. Lematy pomocnicze	47
Podsumowanie	61
Bibliografia	62

Streszczenie

W pracy rozważany jest problem estymacji nieznanego elementu f na podstawie niebezpośrednich i zaburzonych obserwacji. Niech Λ będzie skończonym zbiorem estymatorów liniowych. Celem jest konstrukcja metody wyboru estymatora z rodziny Λ naśladującego estymator o minimalnym ryzyku w tej klasie. Okazuje się, że można to osiągnąć poprzez minimalizację odpowiedniego wyrażenia związanego z ryzykiem empirycznym. W pierwszej części pracy przedstawione zostały wyniki dotyczące operatorów zwartych. W drugiej części zaprezentowano uogólnienie tych wyników na przypadek dowolnych operatorów liniowych i ograniczonych. Głównym wynikiem pracy jest zaprezentowanie odpowiednich nieasymptotycznych nierówności wyroczni w obu przypadkach.

Słowa kluczowe

statystyczne problemy odwrotne, nierówności wyrocznie, operatory zwarte, operatory niezwalne, minimalizacja ryzyka empirycznego

Klasyfikacja tematyczna AMS 2000: 62G05, 62G20, 62M15, 62M30, 62P99

Abstract

This thesis concerns the problem of estimation of an unknown element f on the basis of indirect and noisy observations. Let Λ be a finite set of possible linear estimators. The aim is to construct a method of selection of an estimator from a family Λ that mimics the estimator with the smallest risk in this set. It is demonstrated that this can be achieved by minimization of an adequate expression connected with the empirical risk. The results related to compact operators are presented in the first part of the work. In the second part the findings are generalized to any linear bounded operator. It is exhibited that in both cases non-asymptotic oracle inequalities can be constructed.

Keywords

statistical inverse problems, oracle inequalities, compact operators, non-compact operators, empirical risk minimization

AMS 2000 Subject Classification: 62G05, 62G20, 62M15, 62M30, 62P99

Wstęp

Statystyczne problemy odwrotne są zagadnieniem występującym w wielu dziedzinach takich, jak geofizyka, obrazowanie medyczne czy astrofizyka. Za każdym razem celem jest rekonstrukcja sygnału, który nie może być jednak obserwowany bezpośrednio. Dodatkowo tak przekształcony sygnał może być zniekształcony przez pewne losowe zaburzenie. Tak postawione zagadnienie może zostać ujęte w matematycznym opisie jako

$$Y = Af + \epsilon,$$

gdzie f jest poszukiwanym sygnałem, operator A modeluje transformację jakiej poddany został sygnał, natomiast ϵ jest pewnym zaburzeniem, którego źródłem może być przykładowo niedokładność aparatury pomiarowej. Najprostszym sposobem "odzyskania" elementu f mogłaby się wydawać konstrukcja estymatora Af i następnie poddanie go działaniu operatora odwrotnego A^{-1} . Jednak w większości przypadków okazuje się, że takie podejście prowadzi do nadmiernego zwiększenia błędu estymacji z uwagi na nieograniczoność operatora odwrotnego A^{-1} . Celem poradzenia sobie z tym problemem zaproponowano w literaturze szereg metod mających na celu niedopuszczenie do eksplozji błędu poprzez zastosowanie tak zwanej regularyzacji ([9]) pozwalającej na kontrolowanie zachowania odwrotności operatora. Metody te wymagają jednak najczęściej ustalenia pewnego parametru, kluczowego dla ich zachowania. Stąd pojawiła się konieczność konstrukcji odpowiednich procedur wyboru tych parametrów na podstawie danych dla skończonych prób. W tej pracy omawiana jest jedna z takich procedur, wybierająca estymator θ^* ze skończonej rodziny, naśladujący ryzyko najlepszego estymatora w tej rodzinie, polegająca na minimalizacji odpowiedniego wyrażenia związanego z ryzykiem empirycznym. W przypadku, gdy rozważany operator jest operatorem zwartym wyrażenie to jest nieobciążonym estymatorem ryzyka, natomiast w przypadku ogólnym dla operatorów liniowych i ograniczonych jest to estymator pewnego górnego oszacowania na ryzyko rozważanego estymatora elementu f . W obu przypadkach udaje się skonstruować nieasymptotyczne nierówności wyrocznie postaci

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

gdzie Λ jest rozważaną rodziną estymatorów. Pozwalają one w sposób systematyczny wybierać i oceniać ryzyko estymatora z pewnej skończonej rodziny. W literaturze

opisano także inne metody wyboru (przykładowo [8]), jednak zaprezentowana tutaj metoda charakteryzuje się dużą intuicyjnością.

Praca zorganizowana jest w następujący sposób. W rozdziale 1 zaprezentowane są podstawowe pojęcia wykorzystywane w pracy oraz wprowadzenie do pojęcia wyroczni, rozdziały 2 i 3 prezentują uzyskane rezultaty dla odpowiednio przypadku operatorów zwartych i dowolnych operatorów ograniczonych i liniowych, w rozdziale 4 zamieszczono przykład zastosowania uzyskanych wyników dla problemu dekonwolucji na prostej rzeczywistej, natomiast rozdział 5 zawiera dowody wszystkich lematów użytych do dowodzenia głównych rezultatów. Rozdział 2 jest prezentacją wyników uzyskanych w pracy [7] z pewnymi niewielkimi poprawkami, natomiast rozdział 3 został napisany w oparciu o pracę [6] z szerokimi zmianami i poprawkami wprowadzonymi przez autora.

Rozdział 1

Pojęcia wstępne

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 1.1 ([23], str. 1). *Problem nazwiemy dobrze postawionym w sensie 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. Będziemy pracować z następującym modelem

$$Y = Af + \epsilon \xi, \tag{1.1}$$

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

Definicja 1.2 ([1], str. 7). *Stochastycznym błędem ξ nazwiemy proces na przestrzeni Hilberta G , czyli ograniczony liniowy operator $\xi: G \rightarrow L_2(\Omega, \mathcal{F}, \mathbb{P})$ taki, że dla dowolnych elementów $g_1, g_2 \in G$ mamy zdefiniowane zmienne losowe $\langle \xi, g_i \rangle$ takie, że $\mathbb{E} \langle \xi, g_i \rangle = 0$. Operator kowariancji \mathbf{Cov}_ξ określony jest jako ograniczony liniowy operator ($\|\mathbf{Cov}_\xi\| \leq 1$) z przestrzeni G w przestrzeń G taki, że $\langle \mathbf{Cov}_\xi g_1, g_2 \rangle =$*

$\mathbf{Cov}(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle)$. Przestrzeń $(\Omega, \mathcal{F}, \mathbb{P})$ jest podstawową przestrzenią probabilistyczną, natomiast $L_2(\cdot)$ jest przestrzenią wszystkich funkcji całkowalnych z kwadratem na zadanej przestrzeni z miarą.

Definicja 1.3 ([1], str. 8). Powiemy, że losowy błąd ξ jest gaussowskim białym szumem, jeśli $\mathbf{Cov}_\xi = I$ oraz indukowane zmienne losowe są gaussowskie, czyli dla dowolnych elementów $g_1, g_2, \dots, g_k \in G$ i dla dowolnego $k \in \mathbb{N}$ mamy, że $(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle, \dots, \langle \xi, g_k \rangle) \sim \mathcal{N}_k(\mathbf{0}, \Sigma)$ oraz $\mathbf{Cov}(\langle \xi, g_i \rangle, \langle \xi, g_j \rangle) = \langle g_i, g_j \rangle$.

Lemat 1.1 ([1], str. 8). Niech ξ będzie białym szumem w przestrzeni G oraz niech $\{u_i\}_{i \in I}$ będzie bazą ortonormalną tej przestrzeni i niech $\xi_i = \langle \xi, u_i \rangle$. 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 $\mathbf{Cov}(\langle \xi, u_n \rangle, \langle \xi, u_k \rangle) = \langle u_n, u_k \rangle = \delta_{nk}$, gdzie δ_{nk} oznacza symbol Kroneckera, co wraz z założeniem o łącznym rozkładzie gaussowskim kończy dowód. \square

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

$$\forall g \in G \quad \langle Y, g \rangle = \langle Af, g \rangle + \epsilon \langle \xi, g \rangle$$

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 $D(A) = \{f \in H : \exists_{g \in G} Af = g\} = H$.

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

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

Operator $A: H \rightarrow G$ jest operatorem śladowym, jeśli istnieje baza $\{x_i\}_i$ przestrzeni H taka, że $\sum_i \langle (A^*A)^{1/2}x_i, x_i \rangle < \infty$ ([26], str. 285).

Estymatorem liniowym elementu f w modelu (1.1) nazywamy estymator postaci TY , gdzie T jest pewnym operatorem z przestrzeni $L(G, H)$, a działanie operatora T na proces Y zdefiniowane jest warunkiem $\langle TY, h \rangle = \langle Y, T^*h \rangle$ dla dowolnego elementu $h \in H$. Jeżeli błąd estymacji ma być mierzony w normie przestrzeni H , to trzeba zapewnić, że TY jest losowym elementem H , czyli procesem o trajektoriach w H , a nie tylko procesem na przestrzeni H w sensie definicji 1.2. Będzie tak, gdy operator kowariancji TY jest operatorem śladowym.

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

Twierdzenie 1.1 ([1], str. 9, [19], str. 10). Niech $A \in L(H, G)$. Wtedy

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

Dowód. Zauważmy, że $\text{Range} A^\perp = \{g \in G: \langle Af, g \rangle = 0 \ \forall f \in H\}$. Wtedy dla dowolnych $f \in \text{Ker} A$ i $g \in G$ mamy, że $0 = \langle Af, g \rangle = \langle f, A^*g \rangle$, a stąd $\text{Ker} A = (\text{Range} A^*)^\perp$. Zamieniając A z A^* otrzymujemy, że $\text{Ker} A^* = \text{Range} A^\perp$, czyli $(\text{Ker} A^*)^\perp = (\text{Range} A^\perp)^\perp = \overline{\text{Range} A}$, 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 $\langle x, y \rangle = \langle x - x_n, y \rangle + \langle x_n, y \rangle = \langle x - x_n, y \rangle + 0 \rightarrow 0$, gdy $n \rightarrow \infty$ co dowodzi, że $x \in A^\perp$.

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

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

Wniosek 1.1. Dla liniowego operatora $A: H \rightarrow G$ zachodzi

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

W pierwszej części pracy ograniczymy się do rozważania tylko zwartych operatorów liniowych, jednak dzięki temu uda się uzyskać dającą szerokie możliwości reprezentację według wartości 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 ich bardzo wygodną reprezentację.

Definicja 1.4 ([26], str. 264). Operator $A: H \rightarrow G$ 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 $\int_a^b \int_a^b |K(x, y)|^2 dx dy < +\infty$ lub słabo osobliwe, czyli postaci $\frac{\mathcal{H}(x, y)}{|x-y|^\alpha}$, gdzie $\alpha \in (0, 1)$ a funkcja \mathcal{H} jest funkcją mierzalną i ograniczoną na odcinku $[a, b]$.

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

Uwaga 1.1. Jeżeli $A \in K(H, G)$ oraz $\dim H = \infty$ 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 1.2 (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 $J = \{j \in I: \lambda_j \neq 0\}$ oznacza zbiór tych indeksów dla których odpowiednie wartości własne są niezerowe, wtedy zbiór J jest przeliczalny oraz*

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

Ponadto dla każdego $\delta > 0$ zbiór $J_\delta = \{j \in I: |\lambda_j| \geq \delta\}$ jest skończony, 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 1.3 (Reprezentacja według wartości singularnych). *Niech $A: H \rightarrow G$ 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 zachodzi*

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

Ponadto $g \in \text{Range} A$ wtedy i tylko wtedy, gdy $g = \sum_n \langle g, u_n \rangle u_n$ oraz spełniony jest tzw. warunek Picarda

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

Wtedy rozwiązania równania $Af = g$ mają postać

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

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

Układ (u_n, v_n, b_n) nazywamy układem singularnym operatora A a jego reprezentację w postaci $Af = \sum_n \lambda_n \langle f, v_n \rangle u_n$ nazywamy dekompozycją według wartości singularnych operatora A .

Dowód. ([19], str. 11, [23], 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 $b_1^2 \geq b_2^2 \geq \dots \geq 0$ oraz wektory ortonormalne v_n takie, że $A^*A v_n = b_n^2 v_n$. Niech $I = \{n: b_n > 0\}$ oraz przez u_n oznaczymy znormalizowane obrazy wektorów v_n , czyli $u_n = b_n^{-1} A v_n$ dla $n \in I$. Zauważmy, że $\langle u_k, u_l \rangle = b_k^{-1} b_l^{-1} \langle A v_k, A v_l \rangle = b_k^{-1} b_l^{-1} \langle v_k, A^* A v_l \rangle = b_k^{-1} b_l^{-1} \langle v_k, b_l^2 v_l \rangle = \delta_{kl}$.

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

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

Tożsamości $Af = \sum_n b_n \langle f, v_n \rangle u_n$ oraz $A^*g = \sum_n b_n \langle g, u_n \rangle v_n$ otrzymujemy, zauważając

$$\begin{aligned} Af &= \sum_n \langle Af, u_n \rangle u_n = \sum_n \langle Af, b_n^{-1} Av_n \rangle u_n = \sum_n \langle f, b_n^{-1} A^* Av_n \rangle u_n \\ &= \sum_n \langle f, b_n^{-1} b_n^2 v_n \rangle u_n = \sum_n b_n \langle f, v_n \rangle u_n \end{aligned}$$

oraz równość drugą analogicznie.

Z nierówności Bessela dostajemy, że $\sum_n |\langle f, v_n \rangle|^2 < \infty$, bo $f \in H$ a stąd

$$\begin{aligned} \sum_n |\langle f, v_n \rangle|^2 &= \sum_n b_n^{-4} |\langle f, b_n^2 v_n \rangle|^2 = \sum_n b_n^{-4} |\langle f, A^* Av_n \rangle|^2 = \sum_n b_n^{-4} |\langle Af, Av_n \rangle|^2 \\ &= \sum_n b_n^{-2} |\langle g, b_n^{-1} Av_n \rangle|^2 = \sum_n b_n^{-2} |\langle g, u_n \rangle|^2 < \infty. \end{aligned}$$

W drugą stronę wnioskujemy, że jeśli spełniony jest warunek Picarda to możemy wypisać jawny wzór na rozwiązanie, gdyż odpowiedni szereg norm współczynników jest zbieżny i g jest sumą swojego szeregu Fouriera.

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

$$Af = A(f_0 + \sum_n b_n^{-1} \langle g, u_n \rangle v_n) = \sum_n \langle g, u_n \rangle b_n^{-1} Av_n = g.$$

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 każdego $n \in \mathbb{N}$, czyli $f = \sum_n b_n^{-1} \langle g, u_n \rangle v_n + f_0$, gdzie $f_0 \in \text{Ker} A$. \square

Udało nam się zaprezentować działanie operatora zwartego w postaci jego rozwinięcia według wartości singularnych w postaci $Af = \sum_n b_n \langle f, v_n \rangle u_n$ oraz użyć postaci 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 $P: G \rightarrow \overline{\text{Range} A}$ będzie rzutem ortogonalnym, czyli $\forall_{g \in G} Pg = \sum_n \langle g, u_n \rangle u_n$. Wtedy dla dowolnego elementu $f \in H$ mamy, że $\|Af - g\|^2 = \|Af - Pg\|^2 + \|(1 - P)g\|^2 \geq \|(1 - P)g\|^2$.

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

Rozważmy układ singulary (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

$$\begin{aligned}\langle Y, u_n \rangle &= \langle Af, u_n \rangle + \epsilon \langle \xi, u_n \rangle = \langle Af, b_n^{-1} A v_n \rangle + \epsilon \xi_n = b_n^{-1} \langle A^* A f, v_n \rangle + \epsilon \xi_n \\ &= b_n^{-1} \left\langle \sum_k b_k^2 \langle f, v_k \rangle v_k, v_n \right\rangle + \epsilon \xi_n = b_n \theta_n + \epsilon \xi_n,\end{aligned}$$

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 1.1 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

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

W tej postaci widać dokładnie trudności związane ze stochastycznymi problemami odwrotnymi. Jako że b_n są wartościami singularnymi operatora zwartego mamy, że b_n → 0, gdy n → ∞, czyli widać, że wraz ze wzrostem n sygnał b_nθ_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_nθ_n, dlatego możemy zapisać równoważną postać problemu

$$x_n = \theta_n + \epsilon \sigma_n \xi_n, \quad n = 1, 2, \dots \quad (1.2)$$

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

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

Mając pełną i niezaburzoną informację o współczynnikach θ_n, n = 1, 2, ... można by uzyskać pełną informację o poszukiwanym elemencie f z dokładnością do składowej znajdującej się w jądrze 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 1.5. Niech $\lambda = (\lambda_1, \lambda_2, \dots)$ będzie nielosowym ciągiem liczbowym. Estymatorem liniowym współczynników θ_n w modelu (1.2) nazwiemy estymator $\hat{\theta}(\lambda) = (\hat{\theta}_1, \hat{\theta}_2, \dots)$, gdzie

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

Ciąg λ nazywać będziemy filtrem lub wagami.

Powyższa definicja jest przeformułowaniem ogólnej definicji estymatora liniowego w modelu (1.1) obserwacji z szumem. Przykładowo estymatory rzutowe estymujące poszukiwany element f przez początkowe N składników w rozwinięciu w szereg Fouriera ze współczynnikami równymi zaobserwowanym wartościom odpowiadają filtrom

$\lambda = (\lambda_1, \lambda_2, \dots)$, w których $\lambda_i = \mathbf{1}_{\{i \leq N\}}$, gdzie $\mathbf{1}_A$ oznacza indyktor zbioru A . Innymi często stosowanymi wagami są wagi Tichonowa–Phillipsa postaci $\lambda_i = \frac{1}{1+(i/w)^a}$, $w > 0$, $a > 0$ lub Pinskera postaci $\lambda_i = \max\{0, 1 - (i/w)^a\}$, $w > 0$, $a > 0$.

Jakość estymatora \hat{f} elementu f mierzona będzie przy pomocy scałkowanego ryzyka średniokwadratowego.

Definicja 1.6. *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 (1.2) estymator \hat{f} możemy zapisać w postaci $\hat{f} = \sum_n \hat{\theta}_n v_n$. Wtedy dostajemy, że

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

gdzie druga równość wynika z tożsamości Parsewala, a norma $\|\cdot\|$ rozumiana jest odpowiednio w przestrzeni 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 (1.2) 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

$$\begin{aligned} \mathbb{E}_\theta \|\theta - \hat{\theta}\|^2 &= \mathbb{E}_\theta \sum_{n=1}^{\infty} (\theta_n - \hat{\theta}_n(\lambda))^2 = \mathbb{E}_\theta \sum_{n=1}^{\infty} (\theta_n - \lambda_n x_n)^2 \\ &= \sum_{n=1}^{\infty} (1 - \lambda_n)^2 \theta_n^2 + \epsilon^2 \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2. \end{aligned} \quad (1.3)$$

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

Definicja 1.7 ([1], str. 28). *Ryzykiem minimaxowym w klasie funkcji \mathcal{F} nazywamy wyrażenie*

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

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

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

Definicja 1.8 ([1], str. 28). *Przypuśćmy, że istnieje estymator \tilde{f} , taki, że istnieją stałe $0 < C_1 \leq C_2 < \infty$ takie, że gdy $\epsilon \rightarrow 0$, to zachodzi*

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

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

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

W ostatniej części tego rozdziału wprowadzone zostanie pojęcie wyroczni. Przypuśćmy przez chwilę, że zajmujemy się estymacją współczynników θ_i w następującym modelu

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

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

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

gdzie $\|\theta\|_n^2 = \sum_{i=1}^n \theta_i^2$. Ryzyko to jest minimalizowane przez zadanie wag postaci

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

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

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

gdzie $\Lambda = \{\lambda X : \lambda \in [0, 1]\}$ jest rozważaną klasą estymatorów, w której wystarczy ograniczyć się do rozważania wag z przedziału $[0, 1]$, gdyż wagi spoza tego przedziału prowadzą do estymatorów niedopuszczalnych. Zauważmy jednak, że nie możemy zastosować tak wyznaczonego estymatora, gdyż korzysta on z niedostępnej dla nas informacji o estymowanym elemencie poprzez $\|\theta\|_n^2$. Minimalne ryzyko może zostać osiągnięte jedynie przez 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

$$\mathcal{R}(\theta^*, \theta) \leq C_1 \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + C_2(\Lambda, \sigma), \quad (1.4)$$

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 od jej odpowiednio zdefiniowanej złożoności, gdy jest ona nieskończona oraz od poziomu szumu σ ([21], str. 177). Dodatkowo pożądaną własnością takiego estymatora jest, by powyższa nierówność wyrocznia prowadziła do dokładnych nierówności postaci

$$\mathcal{R}(\theta^*, \theta) \leq (1 + o(1)) \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) \quad (1.5)$$

gdy $\sigma \rightarrow 0$. W przypadku rozważanego na początku rozdziału problemu poszukiwanym estymatorem naśladowującym wyrocznie jest estymator Jamesa–Steina z wagą

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

spełniający nierówność wyrocznie postaci

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

Szczegółowe uzasadnienie można znaleźć w [27], str. 156.

W badanym w pracy zagadnieniu zaprezentowane zostaną analogiczne nierówności wyrocznie postaci (1.4) i (1.5) 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.

Nierówności wyrocznie w modelu z operatorem zawartym

Rezultaty opisane w tym rozdziale są rezultatami uzyskanymi w pracy [7] z kilkoma modyfikacjami.

Przypuśćmy, że dysponujemy skończonym zbiorem filtrów $\Lambda = (\lambda^1, \dots, \lambda^N)$ i związanych z nimi estymatorów liniowych, gdzie $\lambda^i = (\lambda_1^i, \lambda_2^i, \dots)$, $i = 1, 2, \dots, N$. Naszym celem będzie konstrukcja filtra opartego na obserwacjach $\lambda^*(X) = (\lambda_1^*, \lambda_2^*, \dots)$ o wartościach w zbiorze Λ o asymptotycznie minimalnym ryzyku przy prawdziwej wartości θ , który równocześnie naśladuje ryzyko najlepszego estymatora w tej klasie w każdej skończonej próbie. Okaze się, że filtr ten może zostać zdefiniowany jako element minimalizujący względem $\lambda \in \Lambda$ nieobciążony estymator ryzyka. Zgodnie z (1.3) ryzyko estymatora liniowego wyraża się wzorem

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

Aby uzyskać skończone ryzyko, trzeba dobierać filtry tak, by drugi człon był skończony.

Założenie 2.1.

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

Ponadto założenie o dodatniości powyższej sumy implikuje, że żaden z rozważanych filtrów nie może być tożsamościowo równy zeru. Dodatkowo z postaci wyrażenia na ryzyko widać, że wystarczy ograniczyć się do rozpatrywania wag takich, że $\forall \lambda \in \Lambda \quad \forall i \quad \lambda_i \in [0, 1]$, gdyż w przeciwnym wypadku uzyskane estymatory stają się niedopuszczalne. Mimo tego narzucimy na rozważane wagi nieco słabsze wymagania, aby rozszerzyć zakres stosowalności wyników.

Założenie 2.2.

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

Ponadto będziemy zakładać skończoność poniższej sumy, by występujące później wyrażenia były skończone. Założenie to nie jest jawnie sformułowane w pracy [7].

Założenie 2.3.

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

Z założenia 2.2 wynika następująca nierówność

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

Dodatkowo wymagać będziemy istnienia następującej stałej, której istnienie implikuje, że obie te sumy są tego samego rzędu

Założenie 2.4.

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

Kolejnym krokiem jest wyznaczenie nieobciążonego estymatora wyrażenia (1.3). 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 (1.2) estymatorem takim jest $x_n^2 - \sigma_n^2 \epsilon^2$, gdyż $x_n \sim \mathcal{N}(\theta_n, \sigma_n^2 \epsilon^2)$. Wstawiając to wyrażenie do wzoru opisującego ryzyko estymatora liniowego dostajemy

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

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

$$U(\lambda, X) = \sum_{n=1}^{\infty} (\lambda_n^2 - 2\lambda_n) x_n^2 + 2\epsilon^2 \sum_{n=1}^{\infty} \lambda_n \sigma_n^2 \quad (2.1)$$

będącego nieobciążonym estymatorem $\mathcal{R}(\hat{\theta}, \theta) - \sum_{n=1}^{\infty} \theta_n^2$. Wyrażenie to jest skończone gdyż, dla dowolnego $\theta = \{\theta_i\}_{i=1}^{\infty} \in l^2$ i λ spełniającego założenia 2.1– 2.3 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

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

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

co na mocy lematu 5.2 implikuje zbieżność $U(\lambda, X)$ w normie L_2 .

Uwaga 2.1. Wariancja funkcjonału $U(\lambda, X)$ wyraża się wzorem

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

Wariancja ta jest skończona na mocy wcześniejszych założeń 2.1, 2.2 i 2.4.

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

$$\lambda^* = \arg \min_{\lambda \in \Lambda} U(\lambda, X). \quad (2.2)$$

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

Wprowadzimy następujące oznaczenia na występujące w późniejszych rozważaniach wielkości

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

Wielkość $\rho(\lambda)$ jest pewnym sposobem mierzenia wielkości poszczególnych filtrów biorącym pod uwagę zarówno tempo znikania dalekich wyrazów poprzez $[\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4]^{-1/2}$ jak i rozrzut wokół zera poprzez $\sup_n \sigma_n^2 |\lambda_n|$. Parametry S i M mierzą natomiast zachowanie rodziny wag Λ . Liczbę S można interpretować jako rozrzut bądź zmienność w rodzinie Λ natomiast M jest czynnikiem kontrolującym masywność tej rodziny (dalszy komentarz na temat znaczenia parametru M można znaleźć w pracy [3]). Zauważmy ponadto, że na mocy założeń 2.1 i 2.3 wszystkie te wartości są liczbami skończonymi.

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

$$\frac{(\epsilon^4 \sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4)^{1/2}}{\epsilon^2 \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^2} \leq \rho, \quad (2.3)$$

ponieważ z uwagi na założenie 2.2

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

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

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

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

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

Przed wypowiedzeniem głównego twierdzenia zauważmy jeszcze, że zawsze zachodzi związek

$$M \leq N,$$

gdzie N oznacza licznosc rodziny Λ .

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

Twierdzenie 2.1. *Niech spełnione będą założenia 2.1– 2.4. Wtedy dla dowolnego $\theta \in l^2$, dla dowolnego $B > B_0$ i dla estymatora liniowego θ^* z filtrem wybranym zgodnie z (2.2) zachodzi*

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

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

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

Twierdzenie 2.2. *Niech spełnione będą założenia 2.1– 2.4. 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 (2.2) zachodzi*

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

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

Zanim przejdziemy do dowodu powyższych twierdzeń podamy wniosek z twierdzenia 2.2, który pozwoli wnioskować o asymptotycznej dokładności podanych nierówności wyroczni postaci (1.5).

Wniosek 2.1. *Niech spełnione będą założenia 2.1– 2.4. Ponadto niech zachodzi $\lim_{\epsilon \rightarrow 0} \rho^2 \ln(NS) = 0$. 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*

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

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

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

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

gdyż $M \leq N$. □

Zatem warunek $\lim_{\epsilon \rightarrow 0} \rho^2 \ln(NS) = 0$ jest warunkiem wystarczającym do otrzymania dokładnych asymptotycznie nierówności wyroczni postaci (1.5).

Zanim przejdziemy do dowodu twierdzeń 2.1 i 2.2 podamy trzy lematy z których będziemy korzystać, a których dowody znajdują się w rozdziale 5.

Lemat 2.1. *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 takim, że przyjmuje on wartości w skończonym zbiorze $V \subset l^2$ o liczności $N > 1$. Wtedy dla dowolnego $K \geq 1$ zachodzi*

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

Lemat 2.2. *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 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*

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

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

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

Lemat 2.3. *Niech $\hat{\theta}_i = \hat{\lambda}_i(X) X_i$ będzie estymatorem liniowym z wagami z przedziału $[-1, 1]$ przyjmującymi wartości w zbiorze Λ . Oznaczmy*

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

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

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

Różnica w stałej $1 + 4B^{-1}$ w porównaniu do pracy [7] jest wynikiem innego doboru stałych w dowodzie twierdzeń 2.1 i 2.2 z uwagi na niewłaściwe dobraną stałą w pracy [7] w nierówności (2.4), gdzie z prawej strony stała przyjmowała wartość 2 zamiast 4, jednak w takim przypadku, z uwagi na dopuszczenie ujemnych wartości dla filtrów λ , nierówność ta nie byłaby prawdziwa. Różnica ta nie ma wpływu na ostateczny wydzźwięk tych twierdzeń i wnioski z nich.

Przejdziemy teraz do dowodu twierdzenia 2.1, a później do dowodu twierdzenia 2.2.

Dowód twierdzenia 2.1. Niech $\lambda, \lambda^* \in \Lambda$. Wtedy korzystając ze wzoru na różnicę kwadratów i szacowania $(a+b)^2 \leq 2(a^2+b^2)$, można pokazać, że

$$[(1-\lambda_i)^2 - (1-\lambda_i^*)^2]^2 \leq 4[(1-\lambda_i)^2 + (1-\lambda_i^*)^2][\lambda_i^2 + \lambda_i^{*2}]. \quad (2.4)$$

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 (2.2) i konsekwentnie związany z nim estymator przez θ^* . W rozpatrywanym modelu (1.2) $x_i = \theta_i + \epsilon \sigma_i \xi_i$, 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 (2.1) mamy

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

Obliczając wartość oczekiwaną powyższego wyrażenia dostajemy

$$\mathbb{E}_{\theta} U[\lambda^*, X] = \mathbb{E}_{\theta} \mathcal{R}(\theta^*, \theta) - \sum_{i=1}^{\infty} \theta_i^2 + 2\epsilon \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i + \epsilon^2 \mathbb{E}_{\theta} \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1). \quad (2.5)$$

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

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

Korzystając z lematu 2.1 z $K = S$ i $v_i = [(1 - \lambda_i^*)^2 - (1 - \tilde{\lambda}_i)^2]\sigma_i\theta_i$ dostajemy

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

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

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

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

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

Analogicznie postępując z drugim wyrażeniem dostajemy

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

Zauważmy, że skoro

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

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

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

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

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

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

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

Zachodzi również

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

Przy oznaczeniu $\Delta^\epsilon[\lambda] = \epsilon^2 L_\Lambda \sup_i \sigma_i^2 \lambda_i^2$ jak w lemacie 2.3 prowadzi to do oszacowania postaci

$$\begin{aligned} &-6B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) - 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} [(1 - \lambda_i^*)^2 + (1 - \tilde{\lambda}_i)^2] \theta_i^2 \\ &\geq -6B\epsilon^2 \ln(NS) \mathbb{E}_\theta \left(\sup_i \lambda_i^{*2} \sigma_i^2 + \sup_i \tilde{\lambda}_i^2 \sigma_i^2 \right) - 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 \\ &\quad - 2B^{-1} \sum_{i=1}^{\infty} (1 - \tilde{\lambda}_i)^2 \theta_i^2 \geq -2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 - 2B^{-1} \mathcal{R}(\tilde{\theta}, \theta) \\ &\quad - 6B \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 6B \mathbb{E}_\theta \Delta^\epsilon[\tilde{\lambda}]. \end{aligned} \tag{2.6}$$

Znajdziemy teraz oszacowanie dla składnika $\epsilon^2 \mathbb{E}_\theta \sum_{i=1}^\infty (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1)$. Za-
uważmy na początek, że z uwagi na to, że $|\lambda_i| \leq 1$ dla dowolnego $\lambda \in \Lambda$ z założenia 2.2,
zachodzi $\lambda_i^2 \leq (\lambda_i^2 - 2\lambda_i)^2 \leq 9\lambda_i^2$. Do oszacowania analizowanego wyrażenia posłużymy
się lematem 2.2 z $K = S$, $q = 3$ i $v_i = (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2$. Zatem

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

Oszacujemy teraz niektóre z elementów pojawiających się powyżej:

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

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

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

Stąd dostajemy, że

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

a stąd

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

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

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

z uwagi na to, że $(\lambda_i^{*2} - 2\lambda_i^*)^2 \leq 9\lambda_i^{*2}$ oraz założenie 2.4.

W dalszej części skorzystamy z faktu, że $\forall \omega \in \Omega \quad \min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2 \leq \sup_i \lambda_i^{*2} \sigma_i^2 = \sup_i \lambda_i^{*2}(\omega) \sigma_i^2$, a stąd $\min_{\lambda \in \Lambda} \sup_i \lambda_i^2 \sigma_i^2 \leq \mathbb{E}_\theta \sup_i \lambda_i^{*2} \sigma_i^2$. Analogicznie jak poprzednio

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

Podobnie dla drugiego wyrażenia

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

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

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

Podobnie dla drugiego wyrażenia

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

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

$$\begin{aligned} \epsilon \mathbb{E}_\theta \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) &\geq C \epsilon^2 B \mathbb{E}_\theta L_\Lambda \sup_i \sigma_i^2 \lambda_i^{*2} - 4\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \\ &= -CB \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 4\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2}. \end{aligned} \quad (2.7)$$

Łącząc wyrażenia (2.6) oraz (2.7) dostajemy

$$\begin{aligned} 2\epsilon \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \sigma_i \theta_i \xi_i + \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} (\lambda_i^{*2} - 2\lambda_i^*) \sigma_i^2 (\xi_i^2 - 1) &\geq -4B^{-1} \mathcal{R}(\tilde{\theta}, \theta) \\ &\quad - 4B^{-1} \mathbb{E}_\theta \left[\sum_{i=1}^{\infty} (1 - \lambda_i^*)^2 \theta_i^2 + \epsilon^2 \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^{*2} \right] - 12B \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 12B \Delta^\epsilon[\tilde{\lambda}] \\ &= -4B^{-1} \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) - 4B^{-1} \mathcal{R}(\tilde{\theta}, \theta) - CB \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - CB \Delta^\epsilon[\tilde{\lambda}]. \end{aligned} \quad (2.8)$$

Możemy teraz przejść do ostatecznego szacowania ryzyka badanego estymatora. Zgodnie z (2.5) mamy wykorzystując oszacowania (2.6) i (2.7) (bądź od razu (2.8))

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

Zatem zachodzi

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

Jednak, jako że filtr λ^* był zdefiniowany w (2.2) jako argument minimalizujący nieobciążony estymator ryzyka (2.1), musi zachodzić

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

Dochodzimy zatem do oszacowania postaci

$$(1 - 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\theta^*, \theta) \leq (1 + 4B^{-1}) \mathcal{R}(\tilde{\theta}, \theta) + CB \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CB \Delta^\epsilon[\tilde{\lambda}]. \quad (2.9)$$

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

$$\begin{aligned}\sup_i \sigma_i^2 \lambda_i^2 &= \sup_i \sigma_i^2 \lambda_i^2 \mathbf{1} \left\{ x \sup_i \sigma_i^2 \lambda_i^2 < \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \right\} + \sup_i \sigma_i^2 \lambda_i^2 \mathbf{1} \left\{ x \sup_i \sigma_i^2 \lambda_i^2 \geq \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \right\} \\ &\leq \frac{1}{x} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 + \max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2 \mathbf{1} \left\{ x \sup_i \sigma_i^2 \lambda_i^2 \geq \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 \right\} \\ &\leq \frac{1}{x} \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2 + \omega(x) \leq \frac{1}{x \epsilon^2} \mathcal{R}(\hat{\theta}, \theta) + \omega(x) \quad \forall \lambda \in \Lambda.\end{aligned}$$

Stąd dostajemy oszacowanie na składnik

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

Wykorzystamy teraz powyższe nierówności dla nierówności (2.9).

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

Mamy stąd

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

Korzystając z lematu 2.3 mamy ponadto

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

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

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

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

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

Nierówność ta kończy dowód twierdzenia 2.1. □

Dowód twierdzenia 2.2. Będziemy stosować te same oznaczenia na estymator i wyrocznie jak w dowodzie twierdzenia 2.1. Zauważmy na początek, że z uwagi (2.3) mamy

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

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

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

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

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

Łącząc te dwie nierówności dostajemy

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

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

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

Następnie analogicznie jak w dowodzie twierdzenia 2.1 możemy dostać nierówność postaci (oszacowanie (2.9))

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

To prowadzi do nierówności

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

a stąd dostajemy

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

Ponownie korzystając z lematu 2.3 dostajemy

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

Łącząc te dwie nierówności mamy

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

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

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

co, rozumując analogicznie do dowodu twierdzenia 2.2, prowadzi do nierówności

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

która kończy dowód. □

Nierówności wyrocznie w ogólnym modelu

Rezultaty zaprezentowane w tym rozdziale są modyfikacjami rezultatów uzyskanych w pracy [6]. Podejmiemy próbę zbudowania analogicznych jak w rozdziale poprzednim nierówności wyroczni w przypadku, gdy w rozważanym problemie

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

operator A jest dowolnym operatorem liniowym i ograniczonym, ale niekoniecznie zwartym.

Będziemy potrzebowali pojęcia operatora unitarnego.

Definicja 3.1. *Niech H, G będą ośrodkowymi przestrzeniami Hilberta. Operator $U: H \rightarrow G$ nazwiemy operatorem unitarnym, jeżeli jest ciągły, liniowy, bijektywny i zachodzi $U^* = U^{-1}$.*

Uwaga 3.1. *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, tak jak w rozdziale 1, zdefiniować działanie operatora T na ξ w następujący sposób

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

Ponadto, jeżeli operator kowariancji ξ oznaczmy przez \mathbf{Cov}_ξ , to operator kowariancji $T\xi$ jest postaci $\mathbf{Cov}_{T\xi} = T\mathbf{Cov}_\xi T^*$ ([5], str. 2615). Wprowadźmy także następujące definicje

Definicja 3.2 ([25], str. 102). *Niech $\xi: H \rightarrow L_2(\Omega, \mathcal{F}, \mathbb{P})$ będzie losowym szumem. Powiemy, że ξ ma skończony słaby drugi moment, jeżeli dla dowolnego $f \in H$ zachodzi $\int_W |\langle \xi, f \rangle|^2 d\mathbb{P} < \infty$. Ponadto jeżeli ξ ma trajektorie w przestrzeni H to powiemy, że ξ ma skończony silny drugi moment, jeżeli $\int_W \|\xi\|^2 d\mathbb{P} < \infty$. Jeżeli proces posiada silny drugi moment, to posiada także słaby drugi moment.*

Poniżej przypomnimy definicję operatora śladowego wprowadzoną w rozdziale 1. Jeżeli operator kowariancji procesu jest operatorem śladowym, to wtedy proces ten posiada trajektorie w przestrzeni ([25], str. 175).

Definicja 3.3 ([26], str. 285). *Operator $T: H \rightarrow H$ nazwiemy operatorem śladowym na przestrzeni Hilberta H z miarą μ , jeżeli istnieje baza $\{x_j\}_{j \in J}$ przestrzeni H taka, że zachodzi*

$$\sum_{j \in J} \langle (T^*T)^{1/2} x_j, x_j \rangle < \infty.$$

Niech $T: H \rightarrow H$ będzie liniowym i ograniczonym operatorem na przestrzeni H . Możemy wprowadzić wtedy następującą wielkość nazywaną śladem operatora T

$$\text{Tr}T = \sum_{j \in J} \langle T x_j, x_j \rangle.$$

Na mocy Propozycji 4.7.6 z [26], str. 289, mamy że jeżeli T jest operatorem śladowym, to $\text{Tr}T < \infty$. Niech teraz T będzie ponadto samosprężony i nieujemny i taki, że jego ślad jest skończony. Wtedy $T = (T^2)^{1/2} = (T^*T)^{1/2}$ oraz

$$\text{Tr}T = \sum_{j \in J} \langle T x_j, x_j \rangle = \sum_{j \in J} \langle (T^*T)^{1/2} x_j, x_j \rangle < \infty,$$

czyli operator ten jest operatorem śladowym.

Okazuje się, że dla operatorów samosprężonych można uogólnić twierdzenie o reprezentacji według wartości singularnych do następującej postaci ([24], str. 97)

Twierdzenie 3.1. *Niech A będzie operatorem samosprężonym na ośrodkowej przestrzeni Hilberta H . Wtedy istnieją σ -zwarta przestrzeń mierzalna (S, \mathcal{S}, μ) z miarą Radona μ , rzeczywista funkcja mierzalna b określona na S i operator unitarny $U: H \rightarrow L_2(S, \mathcal{S}, \mu)$, 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 samosprężone 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

$$A^*Y = A^*Af + \epsilon A^*\xi. \quad (3.1)$$

Podstawową korzyścią z takiego przekształcenia jest to, że w dalszej części możemy zajmować się wyłącznie operatorami samosprężonymi i dodatnimi. Fakt, że operator A^*A jest dodatni implikuje, że funkcja b z twierdzenia 3.1 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 3.1 możemy rozważany model (3.1)

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

przekształcić do następującej postaci stosując operator U

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

co można zapisać jako

$$Z = b\theta + \epsilon\eta, \quad (3.2)$$

gdzie $Z = UA^*Y$, $\theta = Uf$ oraz $\eta = UA^*\xi$. Podobnie jak w przypadku operatorów zwartych możemy zapisać to wyrażenie w postaci analogicznej do (1.2)

$$X = \theta + \epsilon\sigma\eta, \quad (3.3)$$

gdzie $\sigma\eta$ oznacza $M_{b^{-1}}UA^*\xi$. Przejście z równości (3.2) do równości (3.3) rozumiane jest jako następujące przekształcenie równania operatorowego. Skoro (3.2) jest postaci $UA^*Y = M_bUf + \epsilon UA^*\xi$, zatem działając lewostronnie na to równanie operatorem $M_{b^{-1}}$ 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 (3.3), gdyż $M_{b^{-1}}M_b$ jest operatorem identycznościowym. Natomiast równanie (3.3) 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 3.1 mówi tylko, że funkcja b jest mierzalną funkcją rzeczywistą nie mówiąc nic więcej o jej zachowaniu. O źle postawionym problemie będziemy mogli mówić tylko wtedy, gdy $b \rightarrow 0$ w jakimś sensie.

Przed dalszym rozumowaniem, udowodnimy czym jest η . Okazuje się, że gdy ξ jest gaussowskim białym szumem na H , to η jest gaussowskim kolorowym szumem na przestrzeni $L_2(S, \mathcal{S}, \mu) = UH$.

Twierdzenie 3.2. Niech η będzie określona jako $\eta = UA^*\xi$ i niech ξ będzie gaussowskim białym szumem na przestrzeni H . Wtedy $\sqrt{\sigma}\eta$ jest gaussowskim białym szumem na przestrzeni $L_2(S, \mathcal{S}, \mu) = UH$.

Dowód. Niech $f \in H$ będzie dowolnym elementem. Oznaczmy przez $\theta = Uf$. Wtedy

$$\langle \sqrt{\sigma}\eta, \theta \rangle = \langle \xi, AU^{-1}\sqrt{\sigma}Uf \rangle \sim \mathcal{N}\left(0, \|AU^{-1}\sqrt{\sigma}Uf\|^2\right),$$

ponieważ σ jest dodatnią funkcją rzeczywistą. Ponadto zachodzi

$$\begin{aligned} \|AU^{-1}\sqrt{\sigma}Uf\|^2 &= \|A(A^*A)^{-1/2}f\|^2 = \langle A(A^*A)^{-1/2}f, A(A^*A)^{-1/2}f \rangle \\ &= \langle (A^*A)^{1/2}f, (A^*A)^{-1/2}f \rangle = \langle (A^*A)^{-1/2}(A^*A)^{1/2}f, f \rangle = \|f\|^2. \end{aligned}$$

Zatem $\langle \sqrt{\sigma}\eta, \theta \rangle \sim \mathcal{N}(0, \|f\|^2)$. Niech teraz dodatkowo $\nu = Uh$, $h \in H$. Wtedy

$$\begin{aligned} \mathbb{E}(\langle \sqrt{\sigma}\eta, \theta \rangle, \langle \sqrt{\sigma}\eta, \nu \rangle) &= \mathbb{E}(\langle \xi, AU^{-1}\sqrt{\sigma}Uf \rangle \langle \xi, AU^{-1}\sqrt{\sigma}Uh \rangle) \\ &= \langle AU^{-1}\sqrt{\sigma}Uf, AU^{-1}\sqrt{\sigma}Uh \rangle = \langle f, h \rangle = \langle \theta, \nu \rangle. \end{aligned}$$

Pokazuje to, że $\sqrt{\sigma}\eta = \sqrt{\sigma}UA^*\xi$ jest gaussowskim białym szumem na przestrzeni $UH = L_2(S, \mathcal{S}, \mu)$. \square

Niech teraz $\hat{\theta}$ oznacza estymator elementu θ w modelu (3.3) 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 [6] pracowano przy założeniu, że obserwacje zaburzone są przez błąd pochodzący z białego szumu. Jednak z uwagi na to, że w przestrzeni nieskończonej 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 operator kowariancji pewnego szumu jest operatorem śladowym, to wtedy szum ten posiada silny drugi moment ([25], str. 175) i w konsekwencji ma też trajektorie w rozważanej przestrzeni. Dlatego za [5] i korzystając z wcześniejszego prekondycjonowania będziemy zakładać o operatorze A i szumie ξ , że zachodzą następujące warunki gwarantujące skończoność drugich momentów regularyzowanego rozwiązania

$$\forall g \in G \quad \mathbb{E}\langle \xi, g \rangle = 0 \text{ oraz } \|\mathbf{Cov}_\xi\| \leq 1, \quad (3.4)$$

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

Jeżeli spełnione są powyższe dwa warunki, możliwe jest takie dobranie miary μ i operatora U w twierdzeniu 3.1, by zachodziło

$$\forall s \in S \quad \mathbf{Var}(UA^*\xi(s)) \leq b(s) \quad (3.5)$$

oraz $b \in L_1(S, \mathcal{S}, \mu)$ ([5], str. 2616). W pozostałej części pracy będziemy zakładać, że miara μ i operator U zostały dobrane w taki właśnie sposób. Przykłady modeli spełniających powyższe warunki znajdują się w rozdziale 4 oraz w [5], 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 ([5], str. 2613, [18], str. 1426)

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

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

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

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

$$\hat{\theta} = \lambda X, \quad (3.6)$$

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 typu obciążenia spektralnego funkcja Φ_α przyjmuje postać

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

Funkcję λ będziemy nazywać wagą lub filtrem. W dalszej części ograniczymy się tylko do rozważania rzeczywistych filtrów i przestrzeni. Możemy teraz wyznaczyć górne oszacowanie ryzyka estymatora postaci (3.6). Warunki (3.4) narzucone na postać szumu i operator A gwarantują skończoność ryzyka.

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

W dalszej części będziemy omijać argumenty w funkcjach podcałkowych. Warto zauważyć, że w powyższym wyrażeniu w składniku $\epsilon^2 \int_S \lambda^2 \sigma d\mu$ czynnik σ związany ze złym uwarunkowaniem problemu występuje w potęgce o jeden niższej niż w analogicznym wyrażeniu (1.3), jednak trzeba zauważyć, że oba modele różnią się charakterem szumu.

Wprowadzimy następujące oznaczenie

$$\Psi(\lambda, \theta) = \int_S (1 - \lambda)^2 \theta^2 d\mu + \epsilon^2 \int_S \lambda^2 \sigma d\mu \quad (3.7)$$

na wyrażenie będące odpowiednikiem wyrażenia (1.3) w modelu rozważanym w rozdziale 2.

Zauważmy, że z uwagi na to, że dysponujemy jedynie oszacowaniem na wariancję szumu η postaci (3.5) nie jest możliwe znalezienie nieobciążonego estymatora ryzyka jak w rozważanym wcześniej problemie z operatorem zwartym. Możliwe jest jednak wskazanie pewnego oszacowania. Z uwagi na (3.3) i (3.5) zachodzi

$$\begin{aligned} \mathbb{E}(X^2 - \epsilon^2 \sigma) &= \mathbb{E}(\theta^2 + 2\epsilon\sigma\theta\eta + \epsilon^2\sigma^2\eta^2 - \epsilon^2\sigma) \\ &= \theta^2 + \epsilon^2\sigma^2 \mathbb{E}\eta^2 - \epsilon^2\sigma \leq \theta^2. \end{aligned}$$

We wzorze (3.7) zastąpimy więc θ^2 przez $X^2 - \epsilon^2\sigma$. W konsekwencji zamiast $\Psi(\lambda, \theta)$ będziemy rozważać jego przybliżenie postaci

$$\psi(\lambda, X) = \int_S (\lambda^2 - 2\lambda)X^2 d\mu + 2\epsilon^2 \int_S \lambda\sigma d\mu. \quad (3.8)$$

Powyższa wartość jest skończona przy założeniu (3.1). Wartość oczekiwana wyrażenia $\psi(\lambda, X)$ jest dolnym oszacowaniem na wyrażenie $\Psi(\lambda, \theta) - \int_S \theta^2 d\mu$, czyli dla dowolnego filtra λ zachodzi

$$\mathbb{E}\psi(\lambda, X) \leq \Psi(\lambda, \theta) - \int_S \theta^2 d\mu.$$

Niech teraz badane filtry należą do pewnej skończonej rodziny $\Lambda = \{\lambda^1, \dots, \lambda^N\}$. 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 oszacowanie $\psi(\lambda, X)$ były skończone. Założymy także skończoność wyrażeń występujących później w nierównościach wyroczniach dla badanych wyrażeń.

Założenie 3.1. *Założmy, że*

$$\begin{aligned} \forall \lambda \in \Lambda \quad 0 < \int_S \sigma^2 \lambda^2 d\mu < \infty, \\ \max_{\lambda \in \Lambda} \|\lambda\|_\infty &\leq 1, \\ \forall \lambda \in \Lambda \quad \int_S \lambda \sigma^2 d\mu &< \infty, \\ \forall \lambda \in \Lambda \quad \int_S \sigma \lambda^2 d\mu &< \infty \\ \forall \lambda \in \Lambda \quad \int_S \sigma \lambda d\mu &< \infty \\ \exists C_1 > 0 \quad \forall \lambda \in \Lambda \quad \int_S \sigma^4 \lambda^2 d\mu &\leq C_1 \int_S \sigma^3 \lambda^4 d\mu, \\ \exists C_2 > 0 \quad \forall \lambda \in \Lambda \quad \int_S \sigma^2 |\lambda| d\mu &\leq C_2 \int_S \sigma \lambda^2 d\mu, \\ \exists C_3 > 0 \quad \forall \lambda \in \Lambda \quad \int_S \sigma |\lambda| d\mu &\leq C_3 \int_S \sigma^2 \lambda^2 d\mu, \end{aligned}$$

Powyższe założenia są odpowiednimi modyfikacjami założeń występujących w rozdziale 2. Różnice wynikają z postaci górnego oszacowania na ryzyko estymatora w obecnie rozważanym modelu, gdzie rząd σ jest o jeden niższy w porównaniu do 1.3. Wprowadzimy także potrzebne oznaczenia

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

Interpretacje powyższych wyrażeń przenoszą się z rozważanego wcześniej przypadku z operatorem zwartym. Zauważmy teraz, że zachodzą następujące związki

$$\int_S \sigma^4 \lambda^4 d\mu \leq C_2 \|\sigma \lambda\|_\infty \int_S \sigma^2 \lambda^2 d\mu$$

skąd mamy

$$\frac{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}}{\int_S \sigma^2 \lambda^2 d\mu} \leq C_2 \rho(\lambda) \leq C_2 \rho.$$

Mamy także

$$\forall \lambda \in \Lambda \quad \rho(\lambda) = \frac{\|\sigma \lambda\|_\infty}{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}} \leq \frac{(\int_S \sigma^2 \lambda^2 d\mu)^{1/2}}{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}} \leq C_3.$$

Mając już odpowiednie założenia możemy zdefiniować poszukiwany estymator naśladujący ryzyko najlepszego estymatora w klasie Λ . Ponownie wykorzystamy w tym celu zdefiniowane wcześniej wyrażenie $\psi(\lambda, X)$, które w pewien sposób przybliży górne ograniczenie na ryzyko, które chcielibyśmy zminimalizować.

Definicja 3.4. Niech funkcjonal $\psi(\lambda, X)$ będzie zdefiniowany jak w (3.8). Poszukiwanym filtrem jest element minimalizujący względem $\lambda \in \Lambda$ funkcjonal $\psi(\lambda, X)$, czyli

$$\lambda^* = \arg \min_{\lambda \in \Lambda} \psi(\lambda, X). \quad (3.9)$$

Przedstawimy teraz uogólnione wersje lematów 2.1, 2.2 i 2.3 na rozważany obecnie przypadek. Dowody znajdują się w części 5.

Lemat 3.1. 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 dla dowolnego elementu $v \in V$ zachodzi $\|T^*v\|_2 \neq 0$. Wtedy dla dowolnego $K \geq 1$ zachodzi

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

Lemat 3.2. 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 co najmniej słabym 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

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

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

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

Lemat 3.3. *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 co najmniej słabym 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*

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

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

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

Możemy teraz przejść do wypowiedzenia dwóch twierdzeń będących zarazem głównym wynikiem tego rozdziału i uogólnieniem twierdzeń 2.1 i 2.2 na przypadek, gdy badany problem może być modelowany jako (3.1).

Twierdzenie 3.3. *Niech będzie spełnione założenie 3.1. 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 (3.9) zachodzi*

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

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

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

Twierdzenie 3.4. *Niech będzie spełnione założenie 3.1. 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 (3.9) zachodzi*

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

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 3.3 możemy zapisać w postaci

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

dla pewnej funkcji $\text{pen}_A: \Lambda \rightarrow \mathbb{R}_+$, 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, [15], str. 181, [15], str. 668, [11], str. 37).

Dowód twierdzenia 3.3. Niech $T = UA^*$, wtedy TT^* jest operatorem kowariancji 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 (3.9) i związany z nim estymator przez θ^* . Oznaczmy ponadto $\max\{1, C_2\} = C'$ oraz $\max\{1, \|T\|^2\} = C''$. W rozpatrywanym modelu (3.3) $X = \theta + \epsilon\sigma\eta$. Dostajemy stąd

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

A stąd mamy

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

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

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

Korzystając z lematu 3.1 z $K = S$ i $v = [(1 - \lambda^*)^2 - (1 - \tilde{\lambda})^2] \sigma \theta$ dostajemy

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

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

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

Postępując podobnie z drugim wyrażeniem otrzymujemy

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

Korzystając z definicji wielkości S mamy

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

A stąd mamy oszacowanie postaci

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

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

$$\begin{aligned}
& \epsilon \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 \sigma \theta \eta d\mu \geq \\
& \geq -2B^{-1} \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 d\mu - 2B^{-1} \int_S (1 - \tilde{\lambda})^2 d\mu - 6B \|T\|^2 \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - 6B \|T\|^2 \mathbb{E}_\theta \Delta^\epsilon[\tilde{\lambda}].
\end{aligned}$$

Znajdziemy teraz oszacowanie dla składnika $\epsilon^2 \mathbb{E}_\theta \left| \int_S (\lambda^{*2} - 2\lambda^*) \sigma (\sigma \eta^2 - 1) d\mu \right|$. Za-uważmy, że dla dowolnego $\lambda \in \Lambda$ zachodzi $\lambda^2 \leq (\lambda^2 - 2\lambda)^2 \leq 9\lambda^2$ rozumiane jako nierówność funkcyjna zachodząca dla wszystkich argumentów funkcji λ . Posłużymy się teraz lematem 3.2 z $K = S$, $q = 3C_2$ i $v = (\lambda^{*2} - 2\lambda^*)\sigma$.

$$\begin{aligned}
& \epsilon^2 \mathbb{E}_\theta \left| \int_S (\lambda^{*2} - 2\lambda^*) \sigma^2 (\eta^2 - 1) d\mu \right| \leq \\
& \leq \epsilon^2 \|T\|^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3C_2)) \right) \left(\mathbb{E}_\theta \|v\| + \sqrt{\mathbb{E}_\theta \|v\|^2 / S} \right).
\end{aligned}$$

Analogicznie jak w rozważanym wcześniej przypadku z operatorem zwartym mamy

$$m(v) = \frac{\|(\lambda^{*2} - 2\lambda^*)\sigma\|_\infty}{\sqrt{\int_S (\lambda^{*2} - 2\lambda^*)^2 \sigma^2 d\mu}} \leq 3C_2 \rho(\lambda),$$

$$M(3C_2) \leq M, \quad m_V \leq 3C_2 \max_{\lambda \in \Lambda} \rho(\lambda) = 3C_2 \rho.$$

Stąd dostajemy, że

$$\sqrt{\ln(NS)} + m_V \ln(SM(3C_2)) \leq C\sqrt{L_\Lambda},$$

gdzie C jest pewną stałą zależną tylko od C_1 i C_2 . Zgodnie z twierdzeniem 3.2 $\sqrt{\sigma}\eta$ jest białym szumem, więc jego operator kowariancji ma postać $I = II^*$. Możemy teraz kontynuować szacowanie analizowanego wyrażenia.

$$\begin{aligned} & \epsilon^2 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*)\sigma(\sigma\eta^2 - 1)d\mu \\ & \geq -\epsilon^2 D \left(\sqrt{\ln(NS)} + m_V \ln(SM(3C_2)) \right) \left(\mathbb{E}_\theta \|v\| + \sqrt{\mathbb{E}_\theta \|v\|^2 / S} \right) \\ & \geq -\epsilon^2 C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\int_S (\lambda^{*2} - 2\lambda^*)^2 \sigma^4 d\mu \right)^{1/2} - \epsilon^2 C \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*)^2 \sigma^4 d\mu \right)^{1/2} \\ & \geq -\epsilon^2 C \sqrt{L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^{*4} \sigma^3 d\mu \right)^{1/2} - \epsilon^2 C \sqrt{L_\Lambda} S^{-1/2} \left(\mathbb{E}_\theta \int_S \lambda^{*4} \sigma^3 d\mu \right)^{1/2}. \end{aligned}$$

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

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

Podobnie dla drugiego wyrażenia

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

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

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

Możemy także zapisać

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

Łącząc oba te oszacowania dostajemy

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

Powyższe oszacowania pozwalają zapisać dalej następującą nierówność

$$\begin{aligned} \mathbb{E}_\theta \psi[\lambda^*, X] &= \mathbb{E}_\theta \Psi(\lambda^*, \theta) - \int_S \theta^2 d\mu + 2\epsilon \mathbb{E}_\theta \int_S (1 - \lambda^*)^2 \sigma \theta_i \eta d\mu \\ &\quad + \epsilon^2 \mathbb{E}_\theta \int_S (\lambda^{*2} - 2\lambda^*)\sigma(\sigma\eta^2 - 1)d\mu \geq \mathbb{E}_\theta \Psi(\lambda^*, \theta) \\ &\quad - \int_S \theta^2 d\mu - 4B^{-1} \mathbb{E}_\theta \Psi(\lambda^*, \theta) - 4B^{-1} \Psi(\tilde{\lambda}, \theta) \\ &\quad - CBC'' \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] - CBC'' \Delta^\epsilon[\tilde{\lambda}]. \end{aligned}$$

Zatem zachodzi

$$(1 - 4B^{-1})\mathbb{E}_\theta \Psi(\lambda^*, \theta) \leq \mathbb{E}_\theta \psi[\lambda^*, X] + \int_S \theta^2 d\mu \\ + 4B^{-1}\Psi(\tilde{\lambda}, \theta) + CBC''\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CBC''\Delta^\epsilon[\tilde{\lambda}].$$

Jednak, jako że filtr λ^* był zdefiniowany jako argument minimalizujący wyrażenie ψ , zatem musi zachodzić

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

Otrzymujemy zatem

$$(1 - 4B^{-1})\mathbb{E}_\theta \Psi(\lambda^*, \theta) \leq (1 + 4B^{-1})\Psi(\tilde{\lambda}, \theta) \\ + CBC''\mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + CBC''\Delta^\epsilon[\tilde{\lambda}]. \quad (3.10)$$

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

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

Stąd dostajemy

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

Wykorzystamy teraz powyższe nierówności do wyrażenia (3.10).

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

Mamy stąd

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

Korzystając z lematu 3.3 mamy ponadto

$$\begin{aligned} \frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1 + 4B^{-1} C'') \mathbb{E}_\theta \Psi(\lambda^*, \theta) + CB \|T\|^2 \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] \\ &\leq \left(1 + 4B^{-1} C'' + CB \|T\|^2 \frac{L_\Lambda}{x} \right) \mathbb{E}_\theta \Psi(\lambda^*, \theta) + CB \|T\|^2 \epsilon^2 L_\Lambda \omega(x) \\ &\leq \left(1 + 4B^{-1} C'' + CBC'' \frac{L_\Lambda}{x} \right) \mathbb{E}_\theta \Psi(\lambda^*, \theta) + CB \|T\|^2 \epsilon^2 L_\Lambda \omega(x). \end{aligned}$$

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

$$\begin{aligned} \frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq \frac{(1 + \gamma B^{-1} C'')^2}{1 - \gamma B^{-1} C''} \Psi(\tilde{\lambda}, \theta) \\ &+ \frac{1 + \gamma B^{-1} C''}{1 - \gamma B^{-1} C''} C B C'' \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) + C B \|T\|^2 \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda). \end{aligned}$$

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

$$\frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq (1 + \gamma_1 B^{-1} C'') \Psi(\tilde{\lambda}, \theta) + \gamma_2 B C'' \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda),$$

która kończy dowód twierdzenia 3.3. □

Dowód twierdzenia 3.4. Będziemy stosować te same oznaczenia na estymator, wyroczenie i stałe jak w dowodzie twierdzenia 3.3. Wyrażenie $\frac{\|\sigma^2 \lambda^2\|_\infty}{\int_S \sigma^2 \lambda^2 d\mu}$ możemy oszacować w następujący sposób

$$\frac{\|\sigma^2 \lambda^2\|_\infty}{\int_S \sigma^2 \lambda^2 d\mu} \leq C_2 \rho \frac{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}}{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}} \leq C_2 \rho.$$

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

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

Skąd dostajemy

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

Nierówność ta prowadzi nas do oszacowania

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

W dowodzie twierdzenia 3.3 uzyskaliśmy nierówność następującej postaci

$$\begin{aligned} (1 - 4B^{-1}) \mathbb{E}_\theta \Psi(\lambda^*, \theta) &\leq (1 + 4B^{-1}) \Psi(\tilde{\lambda}, \theta) \\ &+ C B C'' \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] + C B C'' \Delta^\epsilon[\tilde{\lambda}]. \end{aligned}$$

Wykorzystując wyprowadzone oszacowanie mamy

$$\begin{aligned} (1 - 4B^{-1}) \mathbb{E}_\theta \Psi(\lambda^*, \theta) &\leq (1 + 4B^{-1}) \Psi(\tilde{\lambda}, \theta) + C B C_2^2 C'' \rho L_\Lambda \Psi(\tilde{\lambda}, \theta) \\ &+ C B C_2^2 C'' \rho L_\Lambda \mathbb{E}_\theta \Psi(\lambda^*, \theta), \end{aligned}$$

co prowadzi do

$$\mathbb{E}_\theta \Psi(\lambda^*, \theta) \leq \frac{1 + 4B^{-1} + C B C_2^2 C'' \rho L_\Lambda}{1 - 4B^{-1} - C B C_2^2 C'' \rho L_\Lambda} \Psi(\tilde{\lambda}, \theta).$$

Ponownie korzystając z lematu 3.3 dostajemy

$$\begin{aligned} \frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 &\leq (1 + 4B^{-1}C'') \mathbb{E}_\theta \Psi(\lambda^*, \theta) + CB \|T\|^2 \mathbb{E}_\theta \Delta^\epsilon[\lambda^*] \\ &\leq (1 + 4B^{-1}C'' + CB \|T\|^2 C_2^2 \rho L_\Lambda) \Psi(\tilde{\lambda}, \theta). \end{aligned}$$

Łącząc te dwie nierówności mamy

$$\frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq \frac{(1 + 4B^{-1}C'' + CBC_2^2 C'' \rho L_\Lambda)^2}{1 - 4B^{-1}C'' - CBC_2^2 C'' \rho L_\Lambda} \Psi(\tilde{\lambda}, \theta).$$

Zauważmy teraz, że istnieje taka stała $\gamma_4 > 0$, że jeśli tylko zachodzi $C_2^2 C'' \rho L_\Lambda \leq \gamma_4$, to wybór B jako $(C_2^2 \rho L_\Lambda)^{-1/2}$ prowadzi do nierówności $4B^{-1}C'' + CBC_2^2 C'' \rho L_\Lambda \leq 1/2$, a stąd $1 - 4B^{-1}C'' - CBC_2^2 C'' \rho L_\Lambda \geq 1/2$, czyli jest odcięte od zera. Wtedy wybór $B = (C_2^2 \rho L_\Lambda)^{-1/2}$ prowadzi do nierówności

$$\frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq \frac{(1 + CC'' \sqrt{\rho L_\Lambda})^2}{1 - CC'' \sqrt{\rho L_\Lambda}} \Psi(\tilde{\lambda}, \theta),$$

która z kolei prowadzi nas do postulowanej na początku nierówności

$$\frac{1}{C'} \mathbb{E}_\theta \|\theta^* - \theta\|^2 \leq (1 + \gamma_3 C'' \sqrt{\rho L_\Lambda}) \Psi(\tilde{\lambda}, \theta),$$

która kończy dowód twierdzenia 3.4. □

Wniosek 3.1. *Niech będzie spełnione założenie 3.1, ponadto niech zachodzi $\lim_{\epsilon \rightarrow 0} \rho \ln(NS) = 0$. Wtedy istnieją stałe $\mathbb{C}_2 > 0, \mathbb{C}_3 > 0$ zależące tylko od stałej C_2 i operatora $T = UA^*$, takie że dla $\rho \ln(NS) < \mathbb{C}_2$ i dla dowolnego $\theta \in L_2(S, \mathbb{S}, \mu)$ zachodzi*

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

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

Rozdział 4

Przykład

W tym rozdziale rozważymy problem dekonwolucji na prostej rzeczywistej. Pokażemy, że w tym przypadku można tak zdefiniować problem, by spełniał on narzucone warunki (3.4) oraz korzystając z nierówności wyroczeni wykazanych w twierdzeniach 3.3 i 3.4 udowodnimy, że estymator wybierany zaproponowaną metodą (3.9) może osiągać optymalne tempo zbieżności.

Naszym celem jest estymacja gęstości f zmiennej losowej Z przyjmującej wartości na prostej rzeczywistej \mathbb{R} . Niech jednak dostępne obserwacje będą próbą prostą z rozkładu zmiennej losowej postaci

$$Y = Z + W,$$

czyli są zaburzone przez pewną zmienną losową W . Niech zmienne X i W będą stochastycznie niezależne oraz niech h oznacza gęstość zmiennej W . Wtedy gęstość g zmiennej losowej Y ma postać

$$g = \int_{\mathbb{R}} h(\cdot - x)f(x)dx =: Kf, \quad (4.1)$$

gdzie K jest operatorem konwolucji na prostej. Zakładać będziemy, że gęstość h , a co za tym idzie operator K , są znane, h jest całkowalna z kwadratem, dodatnia prawie wszędzie oraz symetryczna względem zera. Niech \mathcal{F} oznacza unitarny operator transformaty Fouriera, który będzie dokładniej opisany poniżej. Operator $K: L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$ nie jest operatorem zwartym, ponieważ operator K może być zwarty wtedy i tylko wtedy, gdy operator $\mathcal{F}K\mathcal{F}^{-1} = M_b$ jest zwarty, gdzie $b = \mathcal{F}h$, ponieważ \mathcal{F} jest operatorem unitarnym. Zatem operator K jest zwarty wtedy i tylko wtedy, gdy operator mnożenia przez funkcję b jest zwarty. Jednak operator M_b nie jest zwarty, ponieważ można pokazać, że spektrum operatora mnożenia jest równe domknięciu obrazu całej przestrzeni przez funkcję, czyli w rozważanym przypadku będzie to zbiór nieprzeliczalny. Jednak operator zwarty może mieć co najwyżej przeliczalne spektrum. Zatem operator M_h nie jest zwarty, a co za tym idzie również operator K .

W rozdziale 3 dalsze rozważania poprzedzone były prekondukcjonowaniem, zatem niech

$$q = K^*g = \int_{\mathbb{R}} h(\cdot - x)g(x)dx$$

z uwagi na fakt, że h jest symetryczną funkcją rzeczywistą. Nieobciążonym estymatorem q jest

$$\hat{q}(y) = \frac{1}{n} \sum_{i=1}^n h(Y_i - y), \quad (4.2)$$

gdzie n jest wielkością próby, ponieważ

$$\mathbb{E}\hat{q}(y) = \mathbb{E}\frac{1}{n} \sum_{i=1}^n h(Y_i - y) = \int_{\mathbb{R}} h(z - y)g(z)dz = q(y).$$

By wpisać powyższe zagadnienie w ogólną postać problemu odwrotnego (1.1), potrzeba zdefiniować w odpowiedni sposób szum. W pracy [5], str. 2621 pokazano, że szum $\tilde{\epsilon}: L_2(\mathbb{R}) \rightarrow L_2(\Omega, \mathcal{F}, \mathbb{P})$ zdefiniowany w następujący sposób

$$\forall \phi \in L_2(\mathbb{R}) \quad \langle \tilde{\epsilon}, \phi \rangle = \frac{1}{n} \sum_{i=1}^n \phi(Y_i) - \langle g, \phi \rangle. \quad (4.3)$$

pozwała rozważane zagadnienie zapisać w poszukiwanej postaci $\hat{q} = q + K * \tilde{\epsilon} = K^*f + K^*\tilde{\epsilon}$, ponieważ

$$\langle K^*\tilde{\epsilon}, \phi \rangle = \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}} h(X_i - z)\phi(z)dz - \langle K^*q, \phi \rangle = \langle \hat{q} - q, \phi \rangle$$

dla dowolnego $\phi \in L_2(\mathbb{R})$. Ponadto pokazano (w nieco ogólniejszej postaci) następujące twierdzenie

Twierdzenie 4.1. *Niech operator zdefiniowany w (4.1) będzie iniektywny oraz niech $\|K\|_{2,2} < \infty$, $\|K\|_{2,\infty} < \infty$, gdzie norma $\|K\|_{p,q}$ rozumiana jest jako norma operatora $K: L_p(\mathbb{R}) \rightarrow L_q(\mathbb{R})$. Niech ponadto \hat{q} i $\tilde{\epsilon}$ będą zdefiniowane przez (4.2) i (4.3). Oznaczmy $\sigma = n^{-1/2} (\|g\|_{\infty} + \|g\|_2)^{1/2}$. Wtedy szum zdefiniowany jako*

$$\epsilon = \tilde{\epsilon}/\sigma$$

spełnia warunki (3.4).

Powyższe twierdzenie pokazuje, że do badanego problemu można zastosować metodologię z rozdziału 3. Operatorem unitarnym gwarantującym w tym przypadku zachodzenie warunku

$$\forall x \in \mathbb{R} \quad \mathbf{Var}(UK^*\xi(x)) \leq b(x) \quad (4.4)$$

jest operator Fouriera zdefiniowany jako ciągle przedłużenie na $L_2(\mathbb{R})$ operatora określonego jako

$$(\mathcal{F}f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(itx) f(t) dt,$$

z dziedziną $D(\mathcal{F}) = L_1(\mathbb{R}) \cap L_2(\mathbb{R})$ ([18], str. 1435). Funkcją b jest w tym przypadku kwadrat transformacji Fouriera $(\mathcal{F}h)^2$ funkcji h , ponieważ

$$(\mathcal{F}K^*K)f = \mathcal{F}(h * f * h) = \mathcal{F}f \cdot \mathcal{F}h \cdot \mathcal{F}h = (\mathcal{F}h)^2 (\mathcal{F}f),$$

gdzie $*$ oznacza operację splotu.

Założmy teraz, że istnieją stałe $m, M > 0$ takie, że zachodzi

$$mx^{-2a} \leq (\mathcal{F}h)^2(x) \leq Mx^{-2a}.$$

Wprowadźmy następującą klasę funkcji

$$\mathbb{H}_v = \left\{ f \in L_2(\mathbb{R}) : \int_{-\infty}^{\infty} (1+x^2)^v (\mathcal{F}f(x))^2 dx < \infty \right\} \subset L_2(\mathbb{R}).$$

Będziemy zakładać, że poszukiwany element $f \in \Theta_v \subset \mathbb{H}_v$, jednak żeby wskazać konkretne tempo zbieżności zawężymy f do podmodelu postaci

$$\Theta_s = \left\{ f \in \Theta_v : \int_{-\infty}^{\infty} (1+x^2)^s (\mathcal{F}f(x))^2 dx \leq C \right\}$$

dla $s > v$ ([18], str. 1427). Wtedy w [18], str. 1438 pokazano, że dolne oszacowanie na tempo zbieżności jest postaci

$$\inf_{\hat{f}} \sup_{f \in \Theta_s} \mathbb{E} \left\| \hat{f} - f \right\|_v^2 \geq cn^{-2(s-v)/(2s+5)},$$

czyli

$$\inf_{\hat{f}} \sup_{f \in \Theta_s} \Psi(\hat{\lambda}, \theta) \geq cn^{-2(s-v)/(2s+5)},$$

gdzie norma $\|\cdot\|_v$ rozumiana jest jako norma implikowana przez iloczyn skalarny określony

$$\langle f, g \rangle_v = \langle L^v f, L^v g \rangle$$

z operatorem L zdefiniowanym jako

$$L = \mathcal{F}^{-1} M_l \mathcal{F}, \quad l(x) = (1+x^2)^{1/2}.$$

Rozważmy teraz rodzinę Λ estymatorów rzutowych postaci $\lambda^i(x) = \mathbf{1}\{|x| \leq w_i\}$, $w_i \in \{w_1, w_2, \dots, w_N\}$, $1 \leq w_1 < w_2 < \dots < w_N$ ([6], str. 399). Wtedy założenia (3.1) są spełnione oraz możemy dokładniej oszacować poszczególne składniki występujące w twierdzeniu 3.3. Przez C będziemy rozumieć dodatnią stałą niezależną od $\lambda, \theta, \epsilon$

$$S = \frac{\max_{\lambda \in \Lambda} \|\sigma^2 \lambda^2\|_{\infty}}{\min_{\lambda \in \Lambda} \|\sigma^2 \lambda^2\|_{\infty}} \leq C \left(\frac{w_N}{w_1} \right)^{4a}$$

$$\rho = \max_{\lambda \in \Lambda} \rho(\lambda) = \max_{\lambda \in \Lambda} \frac{\|\sigma \lambda\|_{\infty}}{(\int_S \sigma^4 \lambda^4 d\mu)^{1/2}} \leq C \max_{\lambda \in \Lambda} w_{\lambda}^{-1/4} = C w_1^{-1/4}$$

$$M = \sum_{\lambda \in \Lambda} \exp(-1/\rho(\lambda)) \leq N$$

$$L_{\Lambda} = \ln(NS) + \rho^2 \ln(MS) \leq C \ln \left(\frac{Nw_N}{w_1} \right) + C w_1^{-1/2} \ln \left(\frac{Nw_N}{w_1} \right) \leq C \ln \left(\frac{Nw_N}{w_1} \right)$$

$$\begin{aligned}
\omega(x) &= \max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2 \mathbf{1} \left\{ \int_S \sigma \lambda^2 d\mu \leq x \left\| \sigma^2 \lambda^2 \right\|_\infty \right\} \right\|_\infty \\
&\leq \max_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2 \mathbf{1} \left\{ \int_S \sigma^2 \lambda d\mu \leq \frac{x}{C_2} \left\| \sigma^2 \lambda^2 \right\|_\infty \right\} \right\|_\infty \\
&\leq C \left\| s^{4a} \mathbf{1} \left\{ s^{4a+1} \leq C x s^{4a} \right\} \right\|_\infty \leq C x^{4a}
\end{aligned}$$

Założmy teraz, że $N \sim n^d$, $w_N/w_1 \sim n^e$ oraz niech $\ln(n)/w_1 \rightarrow 0$, gdy $n \rightarrow \infty$, gdzie przez $a \sim b$ rozumiemy, że $\lim_{n \rightarrow \infty} a/b = C > 0$ ([6], str. 399). Ponadto niech $\epsilon \sim n^{-1/2}$. Wtedy z twierdzenia 3.3 dostajemy

$$\begin{aligned}
\sup_{f \in \Theta_s} \mathbb{E} \|f^* - f\|_v^2 &\leq \max\{1, C_2\} (1 + \gamma_1 B^{-1} \max\{1, \|A\|^2\}) \min_{\lambda \in \Lambda} \sup_{f \in \Theta_s} \Psi(\lambda, \theta) \\
&\quad + \max\{1, C_2\} \gamma_2 B \max\{1, \|A\|^2\} \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda) \\
&\leq \max\{1, C_2\} \left((1 + \gamma_3 B^{-1}) \min_{\lambda \in \Lambda} \sup_{f \in \Theta_s} \Psi(\lambda, \theta) \right) + C B n^{-1} (\ln n)^{4a+1}.
\end{aligned}$$

Zauważmy, że człon $C B n^{-1} (\ln n)^{4a+1}$ jest zaniedbywalny przy $c n^{-2(s-v)/(2s+5)}$ dla $s > v$. Niech teraz $B = \ln n$, wtedy powyższa nierówność redukuje się do postaci

$$\sup_{f \in \Theta_s} \mathbb{E} \|f^* - f\|_v^2 \leq C(1 + o(1)) \min_{\lambda \in \Lambda} \sup_{f \in \Theta_s} \Psi(\lambda, \theta).$$

Zatem jeżeli w rodzinie Λ istnieje estymator dla którego $\sup_{f \in \Theta_s} \Psi(\lambda, \theta) = O(n^{-2(s-v)/(2s+5)})$, to wtedy także

$$\sup_{f \in \Theta_s} \mathbb{E} \|f^* - f\|_v^2 = O(n^{-2(s-v)/(2s+5)}).$$

Lematy pomocnicze

W rozdziale tym podamy dowody pomocniczych lematów użytych w pracy.

Lemat 5.1. *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 $L_2 - \lim_{n \rightarrow \infty} \sum_{i=1}^n X_i = X$.*

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

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

bo $\sum_{i=1}^n X_i \sim \mathcal{N}(0, \sum_{i=1}^n s_i^2)$ i 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 $L_2 - \lim_{n \rightarrow \infty} \sum_{i=1}^n X_i = X$ oraz z uwagi na niezależność zmiennych X_i , $\sum_{i=1}^n X_i \sim \mathcal{N}(0, \sum_{i=1}^n s_i^2)$, więc $\sum_{i=1}^n X_i \xrightarrow{d} \mathcal{N}(0, S)$, bo zbieżność w normie L_2 implikuje słabą zbieżność. \square

Lemat 5.2. *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 $L_2 - \lim_{n \rightarrow \infty} \sum_{i=1}^n X_i^2 = Y$.*

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

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

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

Lemat 5.3. Niech $\eta: H \rightarrow L_2(T, \mathbb{T}, \tau)$ 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

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

Dowód. Zachodzenie powyższego lematu pokazano, gdy v jest funkcją prostą w [14], str. 1325. Jako że η ma skończone drugie momenty i v jest ograniczona, możemy zastosować metodę komplikacji by uzyskać żadaną tezę. \square

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

Dowód lematu 2.1. Na początek zauważmy pewien użyteczny fakt. Niech $X \sim \mathcal{N}(0, 1)$ będzie zmienną losową. Wtedy zachodzi oszacowanie

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

Istotnie, możemy napisać, że

$$\begin{aligned} \mathbb{E} X^2 \mathbf{1}_{\{|X| > a\}} &= 2 \int_a^\infty \frac{1}{\sqrt{2\pi}} x^2 e^{-x^2/2} dx = \frac{2}{\sqrt{2\pi}} \left[-x e^{-x^2/2} \Big|_a^\infty + \int_a^\infty e^{-x^2/2} dx \right] \\ &\leq \frac{2}{\sqrt{2\pi}} \left[a e^{-a^2/2} + \frac{1}{a} e^{-a^2/2} \right], \end{aligned}$$

gdzie skorzystaliśmy z nierówności $1 - \Phi(x) \leq x^{-1} \phi(x)$ zachodzącej dla dowolnego $x > 0$, gdzie Φ, ϕ oznaczają odpowiednio dystrybuantę i gęstość standardowego rozkładu normalnego ([10], str. 175). Przejdziemy teraz do dowodu właściwej części lematu.

Zauważmy, że jeżeli $\|v\| = 0$, to $\mathbb{E} |\sum_{k=1}^\infty v_k \xi_k| = 0$ oraz $\sqrt{2 \ln(NK)} \mathbb{E} \|v\| + \sqrt{2 \ln(NK)} (\mathbb{E} \|v\|^2 \cdot \frac{1}{K})^{1/2} = 0$ dla dowolnych $N, K > 0$, czyli teza lematu jest spełniona. Możemy zatem założyć, że $\|v\| \neq 0$. Oznaczmy przez $\zeta_v = \frac{1}{\|v\|} \sum_{k=1}^\infty v_k \xi_k$. Szereg ten jest zbieżny w normie L_2 , gdyż ciąg sum częściowych szeregu $\sum_{k=1}^\infty v_k \xi_k$ jest ciągiem Cauchy'ego w przestrzeni zupełnej, ponieważ gdy $n > m > 0$, to

$$\begin{aligned} \left\| \sum_{k=1}^n v_k \xi_k - \sum_{k=1}^m v_k \xi_k \right\|_2^2 &= \left\| \sum_{k=m+1}^n v_k \xi_k \right\|_2^2 = \sum_{i=1}^N \mathbb{E} \left[\mathbf{1}_{\{v=v^i\}} \left(\sum_{k=m+1}^n v_k \xi_k \right)^2 \right] \\ &= \sum_{i=1}^N \mathbb{E} \left[\mathbf{1}_{\{v=v^i\}} \left(\sum_{k=m+1}^n v_k^i \xi_k \right)^2 \right] \leq \sum_{i=1}^N \mathbb{E} \left(\sum_{k=m+1}^n v_k^i \xi_k \right)^2 \rightarrow 0, \text{ gdy } n, m \rightarrow \infty, \end{aligned}$$

z uwagi na lemat 5.1 i fakt, że dla dowolnego $1 \leq i \leq N$ $v^i \in l^2$ i szereg $\sum_{k=1}^\infty v_k^2$ jest zbieżny także w normie L_2 .

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

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

gdzie w ostatniej nierówności skorzystaliśmy z nierówności Cauchy'ego–Schwarza. Rozważmy teraz funkcję $F(t) = t^2 \mathbf{1}_{\{t > \sqrt{2 \ln(NK)}\}}$. Z uwagi na monotoniczność funkcji kwadratowej dla dodatnich argumentów zachodzi

$$F(\max_{u \in V} |\zeta_u|) = \max_{u \in V} F(|\zeta_u|) \leq \sum_{u \in V} F(|\zeta_u|).$$

Ponownie na mocy lematu 5.1 i niezależności zmiennych losowych $\{\xi_i\}$ mamy, że zmienne losowe ζ_u mają takie same rozkłady normalne $\mathcal{N}(0, 1)$ dla każdego $u \in V$. Zatem możemy napisać

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

gdzie v^1 jest dowolnym ustalonym elementem V . Następnie korzystając z oszacowania pokazanego na początku dowodu z $a = \sqrt{2 \ln(NK)}$ dostajemy

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

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

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

co kończy dowód. □

Dowód lematu 2.2. Zauważmy na początek, że z nierówności Markowa dostajemy, że dla dowolnego $t > 0$ zachodzi, że $P(X > \epsilon) \leq e^{-t\epsilon} \mathbb{E} e^{tX}$. Policzmy pomocniczo następującą wartość oczekiwaną, 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 $\eta_v = \frac{1}{\sqrt{2\|v\|}} \sum_{i=1}^{\infty} v_i(\xi_i^2 - 1)$. Szereg ten jest zbieżny w normie L_2 , gdyż ciąg sum częściowych szeregu $\sum_{i=1}^{\infty} v_i(\xi_i^2 - 1)$ jest ciągiem Cauchy'ego w przestrzeni zupełnej, ponieważ gdy $n > m > 0$, to

$$\begin{aligned} \left\| \sum_{i=1}^n v_i(\xi_i^2 - 1) - \sum_{i=1}^m v_i(\xi_i^2 - 1) \right\|_2^2 &= \sum_{k=1}^N \mathbb{E} \left[\mathbf{1}_{\{v=v^k\}} \left(\sum_{i=m+1}^n v_i(\xi_i^2 - 1) \right)^2 \right] \\ &= \sum_{k=1}^N \mathbb{E} \left[\mathbf{1}_{\{v=v^k\}} \left(\sum_{i=m+1}^n v_i^k(\xi_i^2 - 1) \right)^2 \right] \leq \sum_{k=1}^N \mathbb{E} \left(\sum_{i=m+1}^n v_i^k(\xi_i^2 - 1) \right)^2 \xrightarrow{n,m \rightarrow \infty} 0 \end{aligned}$$

korzystając z lematu 5.2 i faktu, że dla dowolnego $1 \leq k \leq N$ $v^k \in l^2$ i szereg $\sum_{k=1}^{\infty} v_k^2$ jest zbieżny w normie L_2 . Korzystając z powyższych faktów możemy napisać dla dowolnego $t > 0$ i dla ustalonego $u \in V$

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

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

Stąd

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

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

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

Zauważmy następnie, że zachodzi następująca zależność

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

Zatem powyższe oszacowanie sprowadza się do postaci

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

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

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

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

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

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

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

gdy $Q \leq \frac{1}{\sqrt{2}m(u)}$. 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 $C \exp\left(-\frac{Q^2}{4}\right) \geq 0$ prowadzi do tego samego oszacowania. Następnie możemy zapisać, że

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

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

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

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

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

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

dla pewnej stałej D zależnej tylko od q . Nierówność ta kończy dowód lematu. \square

Dowód lematu 2.3. Przypomnijmy, że w naszym modelu (1.2) 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

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

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

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

gdzie skorzystaliśmy z nierówności $2ab \leq Ba^2 + B^{-1}b^2$ spełnionej dla dowolnego $B > 0$. Zauważmy, że skoro $S = \frac{\max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}{\min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}$ to $\frac{\max_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2}{S} = \min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2$, a stąd powyższe wyrażenie redukuje się do postaci

$$\begin{aligned} 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 + \frac{1}{2} \epsilon^2 B \ln(NS) \mathbb{E}_\theta \sup_i \sigma_i^2 \lambda_i^2(X) + \epsilon^2 B \ln(NS) \min_{\lambda \in \Lambda} \sup_i \sigma_i^2 \lambda_i^2 \\ \leq 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 + \frac{1}{2} B \mathbb{E}_\theta \left(\epsilon^2 \ln(NS) \sup_i \sigma_i^2 \lambda_i^2(X) \right) \\ + B \mathbb{E}_\theta \left(\epsilon^2 \ln(NS) \sup_i \sigma_i^2 \lambda_i^2(X) \right) \\ \leq 2B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 + \frac{3}{2} B \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)]. \end{aligned}$$

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

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

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

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

Zauważmy, że zachodzi następujący związek

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

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

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

Łącząc te oszacowania dostajemy, że

$$\begin{aligned}
\mathbb{E}_\theta \|\hat{\theta} - \theta\|^2 & \leq \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) - 2\epsilon \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X)) \theta_i \lambda_i(X) \sigma_i \xi_i + \epsilon^2 \mathbb{E}_\theta \sum_{i=1}^{\infty} \lambda_i^2(X) \sigma_i^2 (\xi_i^2 - 1) \\
& \leq \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + 4B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} (1 - \lambda_i(X))^2 \theta_i^2 + \frac{3}{2} B \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)] \\
& \quad + 4\epsilon^2 B^{-1} \mathbb{E}_\theta \sum_{i=1}^{\infty} \sigma_i^2 \lambda_i^2(X) + \frac{B D^2}{2} \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)] \\
& = (1 + 4B^{-1}) \mathbb{E}_\theta \mathcal{R}(\hat{\theta}, \theta) + C B \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)],
\end{aligned}$$

co kończy dowód lematu. □

Dowód lematu 3.1. W dowodzie lematu ponownie skorzystamy z oszacowania

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

zachodzącego dla zmiennych losowych o standardowym rozkładzie normalnym oraz z własności

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

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

Oznaczmy przez $\zeta_v = \frac{\langle \eta, v \rangle}{\|T^* v\|_2}$. 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

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

gdzie v^1 jest dowolnym ustalonym elementem V . Korzystając ze wspomnianego na początku oszacowania z $a = \sqrt{2 \ln(NK)}$ dostajemy

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

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

Dowód lematu 3.2. 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 5.3 z $v = \frac{v}{\|T^*v\|_2}$. Z nierówności Markowa dostajemy, że dla dowolnego $t > 0$ i $x > 0$ zachodzi i ustalonego $u \in V$

$$\begin{aligned}
\ln P(\zeta_u > x) &\leq -tx + \ln \mathbb{E} \exp \left(\frac{t}{\|T^*u\|_2^2} \langle \eta^2 - 1, u \rangle \right) \\
&\leq -tx + \frac{\left(\frac{t}{\|T^*u\|_2} \right)^2 \|u\|_2^2}{1 - \frac{t}{\|T^*u\|_2} \|u\|_\infty} \leq -tx + \frac{\left(\frac{t}{\|T^*u\|_2} \right)^2 \|u\|_2^2}{1 - \frac{t}{\|T^*u\|_2} \|u\|_2 m(u)}.
\end{aligned}$$

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

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

Dostajemy zatem następującą nierówność

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

o ile $t < \frac{\|T^*u\|}{\sqrt{2}m(u)\|u\|_2}$. Zauważmy, że nierówność ta sprowadza się do analogicznej nierówności z dowodu lematu 2.2, 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

$$\begin{aligned} & \exp\left(-\frac{\|T^*u\|_2 x}{\sqrt{2}m(u)\|u\|_2} + \frac{1}{m^2(u)} \ln\left(1 + \frac{\|T^*u\|_2 m(u)x}{\sqrt{2}\|u\|_2}\right)\right) \\ &= \exp\left(\frac{1}{m^2(u)} \left(\ln\left(1 + \frac{\|T^*u\|_2 m(u)}{\sqrt{2}\|u\|_2} x\right) - \frac{\|T^*u\|_2 m(u)}{\sqrt{2}\|u\|_2}\right)\right) \\ &\leq \exp\left(\frac{-\frac{\|T^*u\|_2^2 x^2}{\|u\|_2^2}}{1 + \frac{\|T^*u\|_2 m(u)}{\sqrt{2}\|u\|_2} x}\right), \end{aligned} \quad (5.1)$$

gdzie ponownie skorzystaliśmy z nierówności $\ln(1+z) - z \leq -\frac{z^2}{2(1+z)}$ 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 ([17], str. 216)

$$c\|u\|_2 \leq \|Tu\|_2 \leq C\|u\|_2.$$

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 (5.1) 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 .

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

Ostatecznie prowadzi nas to do następującego oszacowania

$$P(\zeta_u > x) \leq \begin{cases} \exp\left(\frac{-\|t\|^2 x^2}{2}\right), & \text{gdy } 2^{-1/2} \|T\| m(u)x < 1, \\ \exp\left(\frac{-\|t\|^2 x}{\sqrt{2}\|T\| m(u)}\right), & \text{gdy } 2^{-1/2} \|T\| m(u)x \geq 1. \end{cases}$$

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

$$P(|\zeta_u| > x) \leq 2 \begin{cases} \exp\left(\frac{-\|t\|^2 x^2}{2}\right), & \text{gdy } 2^{-1/2} \|T\| m(u)x < 1, \\ \exp\left(\frac{-\|t\|^2 x}{\sqrt{2}\|T\| m(u)}\right), & \text{gdy } 2^{-1/2} \|T\| m(u)x \geq 1. \end{cases}$$

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

$$\begin{aligned} \mathbb{E}\zeta_u^2 \mathbf{1}\{|\zeta_u| > Q\} &= 2 \int_Q^\infty z P(|\zeta_u| > z) dz \leq 4 \int_Q^\infty z \exp\left(\frac{-\|t\|^2 z^2}{1 + 2^{-1/2} \|T\| m(u) z}\right) dt \\ &\leq 4 \int_Q^{\frac{\sqrt{2}}{\|T\|m(u)}} z \exp\left(\frac{-\|t\|^2 z^2}{2}\right) dz + 4 \int_Q^\infty z \exp\left(\frac{-\|t\|^2 z}{\sqrt{2} \|T\| m(u)}\right) dz \\ &\leq C \exp\left(\frac{-\|t\|^2 Q^2}{2}\right) + CQ \exp\left(\frac{-\|t\|^2 Q}{\sqrt{2} \|T\| m(u)}\right). \end{aligned}$$

Podobnie jak w dowodzie lematu 2.2 nierówność tą możemy uogólnić dla dowolnego $Q > 0$. Zauważmy następnie, że zachodzi

$$\begin{aligned} \sum_{u \in V} \exp\left(\frac{-\|t\|^2 Q}{\sqrt{2} \|T\| m(u)}\right) &= \sum_{u \in V} \exp\left(-\frac{q}{m(u)}\right) \exp\left(-\frac{\|t\| Q/(\sqrt{2} \|T\|) - q}{m(u)}\right) \\ &\leq M(q) \exp\left(-\frac{\|t\| Q/(\sqrt{2} \|T\|) - q}{m_V}\right), \end{aligned}$$

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

$$\begin{aligned} \mathbb{E}|\langle \eta^2 - 1, u \rangle| &= \mathbb{E}\|T^*u\|_2^2 |\zeta_u| \leq \|T\|^2 \mathbb{E}\|u\|_2^2 |\zeta_u| \leq \|T\|^2 \mathbb{E}\|u\|_2^2 \max_{u \in V} |\zeta_u| \\ &\leq \|T\|^2 \mathbb{E}\|u\|_2^2 \max_{u \in V} |\zeta_u| \mathbf{1}\{\max_{u \in V} |\zeta_u| \leq Q\} \\ &\quad + \|T\|^2 \mathbb{E}\|u\|_2^2 \max_{u \in V} |\zeta_u| \mathbf{1}\{\max_{u \in V} |\zeta_u| > Q\} \\ &\leq \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 (\mathbb{E}\|v\|_2^4)^{1/2} \left(\mathbb{E} \max_{u \in V} |\zeta_u|^2 \mathbf{1}\{\max_{u \in V} |\zeta_u| > Q\}\right)^{1/2} \\ &\leq \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 (\mathbb{E}\|v\|_2^4)^{1/2} \left(\sum_{u \in V} \mathbb{E}|\zeta_u|^2 \mathbf{1}\{\max_{u \in V} |\zeta_u| > Q\}\right)^{1/2} \\ &\leq \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 (\mathbb{E}\|v\|_2^4)^{1/2} \cdot \\ &\quad \cdot \left(\sum_{u \in V} \left(C \exp\left(\frac{-\|t\|^2 Q^2}{2}\right) + CQ \exp\left(\frac{-\|t\|^2 Q}{\sqrt{2} \|T\| m(u)}\right)\right)\right)^{1/2} \\ &\leq \|T\|^2 \mathbb{E}\|u\|_2^2 Q + \|T\|^2 (\mathbb{E}\|v\|_2^4)^{1/2} \cdot \\ &\quad \cdot \left(NC \exp\left(\frac{-\|t\|^2 Q^2}{2}\right) + CQM(q) \exp\left(-\frac{\|t\| Q/(\sqrt{2} \|T\|) - q}{m_V}\right)\right)^{1/2} \end{aligned}$$

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. \square

Dowód lematu 3.3. Oznaczmy przez $C' = \max\{1, C_2\}$. Przypomnijmy, że w naszym modelu (3.3) mamy, że $X = \theta + \epsilon\sigma\eta$, a stąd $X^2 = \theta^2 + \epsilon^2\sigma^2\eta^2 + 2\epsilon\theta\sigma\eta$.

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

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

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

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

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

$$\begin{aligned}
&2B^{-1}C' \mathbb{E}_\theta \int_S (1 - \lambda(X))^2 \theta^2 d\mu + \frac{1}{2C'} \epsilon^2 \|T\|^2 B \ln(NS) \mathbb{E}_\theta \left\| \sigma^2 \lambda^2(X) \right\|_\infty \\
&\quad + \epsilon^2 B \frac{1}{C'} \|T\|^2 \ln(NS) \min_{\lambda \in \Lambda} \left\| \sigma^2 \lambda^2(X) \right\|_\infty \\
&\leq 2B^{-1}C' \mathbb{E}_\theta \int_S (1 - \lambda(X))^2 \theta^2 d\mu + \frac{3}{2C'} B \|T\|^2 \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)].
\end{aligned}$$

Następnie będziemy szacować drugie z wyrażeń, czyli $\epsilon^2 \mathbb{E}_\theta \left| \int_S \lambda^2(X) \sigma^2(\eta^2 - 1) d\mu \right|$ korzystając z lematu 3.2. Zauważmy na początek

$$m(v) = \frac{\|v\|_\infty}{\|v\|_2} \leq C_2 \rho(\lambda),$$

a stąd $m_V \leq C_2 \rho$ oraz $M(C_2) \leq M$ i możemy szacować wyrażenie $\sqrt{\ln(NS)} + m_V \ln(M(C_2)S)$ przez $\sqrt{2L_\Lambda}$. Zachodzi zatem

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

Analogicznie do poprzednich rozważań zachodzi

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

Dostajemy stąd następujące oszacowanie

$$\begin{aligned} & \epsilon^2 D \sqrt{C_2} \|T\|^2 \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^4(X) \sigma^3 d\mu \right)^{1/2} + \frac{\epsilon^2 D \sqrt{C_2} \|T\|^2 \sqrt{2L_\Lambda}}{\sqrt{S}} \left(\mathbb{E}_\theta \int_S \lambda^4(X) \sigma^3 d\mu \right)^{1/2} \\ & \leq \epsilon^2 D \sqrt{C_2} \|T\|^2 \sqrt{2L_\Lambda} \mathbb{E}_\theta \left(\int_S \lambda^4(X) \sigma^3 d\mu \right)^{1/2} \\ & \quad + \epsilon^2 D \sqrt{C_2} \|T\|^2 \sqrt{2L_\Lambda} \left(\mathbb{E}_\theta \left\| \sigma^2 \lambda^2(X) \right\|_\infty \mathbb{E}_\theta \int_S \sigma \lambda^2(X) d\mu \right)^{1/2} \\ & \leq \mathbb{E}_\theta \frac{\epsilon^2 B D^2 C_2 \|T\|^4 L_\Lambda}{4C'} \left\| \sigma^2 \lambda^2(X) \right\|_\infty + 2\epsilon^2 B^{-1} C' \|T\|^2 \mathbb{E}_\theta \int_S \sigma \lambda^2(X) d\mu \\ & \quad + \mathbb{E}_\theta \frac{\epsilon^2 B D^2 C_2 \|T\|^4 L_\Lambda}{4C'} \left\| \sigma^2 \lambda^2(X) \right\|_\infty + 2\epsilon^2 B^{-1} C' \|T\|^2 \mathbb{E}_\theta \int_S \sigma \lambda^2(X) d\mu \\ & = 4\epsilon^2 B^{-1} C' \|T\|^2 \mathbb{E}_\theta \int_S \sigma \lambda^2(X) d\mu + \frac{\epsilon^2 B D^2 C_2 \|T\|^2}{2C'} \mathbb{E}_\theta L_\Lambda \left\| \sigma^2 \lambda^2(X) \right\|_\infty \\ & \leq 4\epsilon^2 B^{-1} C' \|T\|^2 \mathbb{E}_\theta \int_S \sigma \lambda^2(X) d\mu + \frac{B D^2 C_2 \|T\|^2}{2C'} \mathbb{E}_\theta \Delta^\epsilon[\lambda(X)]. \end{aligned}$$

Łącząc te oszacowania dostajemy tezę lematu. □

Podsumowanie

W pracy zaprezentowano metodę wyboru estymatora ze skończonego zbioru naśladowującego ryzyko najlepszego estymatora w tej rodzinie w stochastycznym problemie odwrotnym postaci

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

Głównym rezultatem jest zaprezentowanie odpowiednich nieasymptotycznych nierówności wyroczni ujętych w twierdzeniach 2.1, 2.2, 3.3 i 3.4 pozwalających kontrolować ryzyko wybranego estymatora przez ryzyko najlepszego estymatora w danej rodzinie.

Rozważania zostały podzielone na przypadek, gdy operator A jest operatorem zwartym oraz przypadek ogólny dla dowolnego operatora liniowego i ograniczonego. Pierwsza część rozważań została zawarta w rozdziale 2 i jest prezentacją wyników uzyskanych w pracy [7]. Dodatkowo w pracy sformułowano w sposób jawny założenie 2.3, poprawiono stałą (2.4) występującą w dowodzie twierdzeń 2.1 i 2.2 oraz uzasadniono zbieżność pojawiających się szeregów. Druga część rozważań została ujęta w rozdziale 3. Zaprezentowane tam rezultaty są istotnymi modyfikacjami wyników uzyskanych w [6]. W artykule tym pracowano przy założeniu, że obserwacje zaburzone są przez błąd pochodzący z białego szumu, co prowadziło jednak do tego, że rozpatrywany proces Y nie miał trajektorii w rozważanej przestrzeni. Dlatego w rozdziale 3 wprowadzono odpowiednie modyfikacje gwarantujące skończoność drugich momentów rozwiązania. Wprowadzone zmiany wymagały głębokich zmian w dowodach twierdzeń 3.3 i 3.4, a także lematów 3.1, 3.3 i szczególnie 3.2, uwzględniających fakt, że źródłem zaburzenia stał się szum kolorowy. W rezultacie niemożliwa stała się praca z nieobciążonym estymatorem ryzyka, a jedynie z pewnym oszacowaniem ryzyka. Dodatkowo w dowodzie lematu 3.2 zastosowano odmienną technikę do uzyskania odpowiedniej nierówności koncentracyjnej.

W rozdziale 4 zaprezentowano ponadto autorski przykład zastosowania wyników z rozdziału 3 w problemie dokonwolucji na prostej. Pokazano w nim, że zaproponowana metoda pozwala wybrać estymator z optymalnym tempem zbieżności na pewnej klasie funkcji pod warunkiem, że rozważana rodzina estymatorów zawiera taki element.

Bibliografia

- [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] L. Birge, *Model selection via testing: an alternative to (penalized) maximum likelihood estimators*, Ann. I. H. Poincaré, PR 42, 2006, pp. 273–325,
- [4] L. Birge, *Statistical estimation with model selection*, arXiv, 2006, The Brouwer Lecture, 2005,
- [5] N. Bissantz, T. Hohage, 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,
- [6] L. Cavalier, *Inverse problems with non-compact operators*, Journal of Statistical Planning and Inference, 136, 2006, pp. 390–400,
- [7] 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,
- [8] L. Cavalier, G. K. Golubev, *Risk hull method and regularization by projections of ill-posed inverse problems*, The Annals of Statistics, Vol. 34, No. 4, 2006, pp. 1653–1677,
- [9] H. Engl, M. Hanke, A. Neubauer *Regularization of Inverse Problems*, Kluwer Academic Publishing, 1996,
- [10] W. Feller, *An Introduction to Probability Theory and Its Applications. Volume 1*, John Wiley and Sons, 1968,
- [11] C. Giraud, *Introduction to High-Dimensional Statistics*, CRC Press, 2015,
- [12] P. R. Halmos, *What does the spectral theorem say?*, The American Mathematical Monthly, Vol. 70, No. 3, 1963, pp. 241–247,
- [13] T. Hida, *Brownian Motion*, Springer, 1980,
- [14] 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,
- [15] J.-M. Loubes, C. Ludena, *Penalized estimators for non linear inverse problems*, ESAIM: PS Vol. 14, 2010, pp. 173–191,
- [16] J.-M. Loubes, C. Ludena, *Adaptive complexity regularization for linear inverse problems*, Electronic Journal of Statistics, Vol. 2, 2008, pp. 661–677,
- [17] L. A. Lusternik, V. J. Sobolew, *Elements of functional analysis*, John Wiley and Sons, 1974,

- [18] B. A. Mair, F. H. Ruymgaart, *Statistical inverse estimation in Hilbert scales*, SIAM J. Appl. Math, Vol. 56, No. 5, 1996, pp. 1424–1444,
- [19] J. Kaipio, E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2004,
- [20] Maniglia S., Ruandi A., *Gaussian measures on Hilbert spaces*, Quaderni dal Dipartimento di Matematica dell' Università del Salento, 1(1), 2004, pp. 1–24,
- [21] C. Mitchell, S. van de Geer, *General oracle inequalities for model selection*, Electron. J. Statist, 3 (2009), pp. 176–204,
- [22] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Springer-Science+Business Media, B.Y., 1986,
- [23] Z. Szkutnik, *Statystyczne problemy odwrotne*, nieopublikowane notatki do wykładu wygłoszonego podczas XXXII Konferencji Statystyka Matematyczna, Wisła, 2006,
- [24] M. E. Taylor, *Partial Differential Equations II. Qualitative Studies of Linear Equations*, Springer Science+Business Media, 2011,
- [25] N. N. Vakhania, V. I. Tarieladze, *Probability Distributions on Banach Spaces*, D. Reidel Publishing Company, 1987,
- [26] H. Lal Vasudeva, *Elements of Hilbert Spaces and Operator Theory*, Springer Verlag, 2017,
- [27] L. Wasserman, *All of Nonparametric Statistics*, Springer Science+Business Media, Inc., 2006,