UNIWERSYTET WARMIŃSKO-MAZURSKI
W OLSZTYNIE

WYDZIAŁ MATEMATYKI I INFORMATYKI

Probabilistyczne struktury regresji

Agnieszka Szumera

Rozprawa doktorska
napisana pod kierunkiem
dr. hab. Dariusza Partyki,
prof. nadzw. KUL JPII

OLSZTYN 2022

Serdeczne podziękowania składam
Panu dr. hab. Dariuszowi Partyce
za wysiłek włożony w mój rozwój,
cenne sugestie, poświęcony mi czas
i okazaną pomoc a przede wszystkim
za ogromną cierpliwość i wyrozumiałość.

Spis treści

Wstęp **7**

1 Zagadnienie regresji **11**

2 Całkowanie funkcji zespolonych **17**

3 Probabilistyczne struktury regresji **27**

3.1 Podstawowe własności funkcji regresji 32

3.2 Regresje pierwszego rodzaju jako szczególny przypadek klasy regresji względem struktury probabilistycznej 37

3.3 Jednoznaczność funkcji regresji 40

3.4 Rzeczywiste probabilistyczne struktury regresji 44

4 Obliczanie funkcji regresji **49**

4.1 Struktura klasy regresji 50

4.2 Procedura wyznaczania funkcji regresji 57

4.3 Przypadek jednowymiarowy 61

4.4 Przypadek dwuwymiarowy 62

4.5 Przypadek trójwymiarowy 69

4.6 Przypadek n-wymiarowy, przegląd podstawowych typów modeli teoretycznych 71

4.7 Szczególny przypadek funkcji skalującej 76

5 Przykłady zastosowań i uwagi uzupełniające **79**

Załącznik **93**

Bibliografia **103**

Wstęp

Postrzeganie zależności jest podwaliną cywilizacji. Rozumiejąc, jak pewne zjawiska zależą od innych, uczymy się przewidywać konsekwencje naszych działań i wpływać na nasze środowisko. Większość znanych nam zależności opiera się na obserwacjach empirycznych. Chociaż niektóre z tych zależności są efektem rozważań teoretycznych, to zazwyczaj same teorie zostały pierwotnie uzyskane z obserwacji empirycznych. Dowolną metodę dopasowywania przebiegów teoretycznych w postaci zadanych funkcji do danych można nazwać regresją. Takie funkcje są przydatne w celu przewidywania i oceny siły relacji opisujących zależności teoretyczne. Metody regresji stały się niezbędne w wielu dziedzinach, w tym w ekonomii, naukach społecznych, inżynierii, badaniach medycznych i biznesie. Spośród różnych metod regresji najszerzej stosowana jest metoda najmniejszych kwadratów, która jest powszechnie używana w zastosowaniach statystycznych.

Znaczący rozwój metod regresyjnych nastąpił w obrębie ekonometrii, gdzie zasadniczy kierunek badań dotyczy stochastycznego charakteru modelowanych procesów ekonomicznych. Termin „regresja" pojawił się po raz pierwszy w XIX wieku. Został on użyty przez F. Galtona (1822 - 1911) do opisu związków zachodzących w procesie dziedziczenia; zob. [8], [31]. Z czasem metody i techniki regresji znalazły zastosowanie w różnych dziedzinach nauki do opisu kształtowania się poziomu pewnych zjawisk w czasie oraz do opisu zależności zachodzących między badanymi wielkościami. Analiza regresji skupiła się głównie na estymacji parametrów równania regresji, które w sposób teoretyczny opisuje istniejący związek między rozważanymi procesami lub zjawiskami; zob. [8]. Za prekursora stosowania metod ilościowych w statystyce uznaje się W. Petty'ego (1623 - 1687). Według niego statystyka była metodą wnioskowania na podstawie danych liczbowych, umożliwiającą wykrycie określonych prawidłowości wśród chaotycznych zjawisk masowych. Na uwagę zasługują również badania prowadzone przez G. Kinga (1648 - 1712), który jako pierwszy podjął próbę ilościowego opisu zależności pomiędzy zmianą cen kukurydzy a wielkością jej zbiorów przy pomocy funkcji liniowej. Z kolei G. U. Yule (1871 - 1951), w swoich pracach opublikowanych w latach 1895 i 1896, proponował zastosowanie analizy korelacji w ekonomii do badania związku pomiędzy ubóstwem a metodami przeciwdziałania temu zjawisku. R. H. Hooker (1867 - 1944) w 1901 roku wykorzystał tę samą technikę do badania związku między zmiennymi, analizując korelację

pomiędzy współczynnikiem ilości zawieranych małżeństw i poziomem koniunktury w Wielkiej Brytanii. Prowadząc swoje badania był jednak świadomy ograniczeń jakie niesie za sobą wykorzystanie analizy korelacji, szczególnie w przypadku jej stosowania do badania danych w postaci szeregów czasowych. W 1907 roku R. Benini (1862 - 1956) jako pierwszy wykorzystał w ekonomii metodę regresji wielorakiej, zaś w 1914 roku H. L. Moore (1869 - 1958) zasłynął wprowadzeniem na grunt ekonomii statystycznej estymacji parametrów ekonomicznych. W 1900 roku L. Bachelier (1870 - 1946), wykorzystując szeregi czasowe cen akcji paryskiej giełdy, zauważył losowy wpływ na ich wielkość, co stało się w późniejszym okresie podstawą do rozważań na temat efektywności rynku; zob. [6], [8], [18], [31].

Klasycznie pojmowane zagadnienie regresji polega na wyznaczaniu wielomianów W stopnia nieprzekraczającego zadanej liczby $n \in \mathbb{N}$, które są optymalnie dopasowane do zadanych skończonych ciągów liczbowych x i y, reprezentujących dane empiryczne. Jako kryterium optymalizacji przyjmuje się tutaj odchylenie kwadratowe. Zatem wielomiany W są wyznaczane metodą najmniejszych kwadratów, na podstawie danych empirycznych. Wielomiany regresyjne odgrywają istotną rolę w analizie empirycznych danych liczbowych. Najprostsze do wyznaczenia są wielomiany regresyjne stopnia nie przekraczającego 1, czyli tzw. regresje liniowe. Zależność liniowa jest również bardzo intuicyjna. Z tego powodu regresje liniowe są najczęściej stosowane w praktyce. Z drugiej strony takie podejście niesie za sobą nadmierne uproszczenie szukanej zależności funkcyjnej pomiędzy danymi. Trudno jest w ten sposób wychwycić bardziej złożone zależności.

Rozważania dotyczące regresji na gruncie rachunku prawdopodobieństwa, jakie są przedstawione w tej pracy, prowadzą do wyodrębnienia szczególnego typu struktur regresji określonych w Definicji 1.2. Są to synchroniczne struktury regresji (SRS) postaci $\mathfrak{R} := (A, B, \delta; x, y)$ zdefiniowane dla zadanej przestrzeni probabilistycznej $(\Omega, \mathcal{A}, P)$, gdzie kryterium odchylenia δ spełnia równość (1.4) z podstawieniem $\nu := P$. O modelu funkcyjnym $\mathcal{F}$ zakłada się, że tworzy on rozmaitość liniową w przestrzeni funkcji $(A \to B)$ ze standardową strukturą dodawania funkcji i mnożenia funkcji przez stałą.

Rozdział pierwszy zawiera zarys historyczny teorii regresji oraz syntetyczny opis asynchronicznych (ARS) i synchronicznych (SRS) struktur regresji. Pojęcia te zostały zaczerpnięte z artykułu *Generalized approach to the problem of regression* (zob. [25]) będącego próbą ujednolicenia teorii regresji. Ukazał się on w 2015 roku i stał się inspiracją do napisania niniejszej pracy doktorskiej. W rozdziale drugim przedstawiono pomocnicze fakty dotyczące całkowania funkcji zespolonych, wykorzystywane w dalszej części dysertacji. Udowodniono między innymi twierdzenia

o całkowaniu przez podstawianie oraz o aproksymacji zespolonych funkcji całkowalnych przez funkcje proste. Rozdział trzeci stanowi zasadniczą część pracy, w której zdefiniowano probabilistyczną strukturę regresji $\mathfrak{P} := (A, B, \delta; x, y)$ nad przestrzenią probabilistyczną $\mathcal{P} := (\Omega, \mathcal{A}, P)$. Głównym wynikiem jest tutaj twierdzenie 3.8 charakteryzujące klasę funkcji regresji. Wynika z niego szereg podstawowych własności funkcji regresji opisanych w dalszej części pracy. W szczególności podano warunki zapewniające jednoznaczność funkcji regresji. Opisano również regresję pierwszego rodzaju w terminach probabilistycznych struktur regresji. Poruszono także zagadnienie zależności pomiędzy rzeczywistymi i zespolonymi probabilistycznymi strukturami regresji. W szczególności zaprezentowany został przykład uzasadniający rozważanie zespolonych probabilistycznych struktur regresji. Rozdział czwarty przedstawia zagadnienie efektywnego wyznaczenia funkcji regresji w terminach zadanej skończonej bazy rozmaitości liniowej $\mathcal{F}$. Została tutaj omówiona struktura klasy regresji oraz procedura wyznaczania funkcji regresji. Ponadto rozwiązano problem regresji dla jednowymiarowych, dwuwymiarowych i trójwymiarowych modeli teoretycznych $\mathcal{F}$. Został poruszony także przypadek ogólny rozwiązania problemu regresji dla wielowymiarowych liniowych modeli teoretycznych. Powyższe rozważania zostały zilustrowane przykładami. Poruszono również zagadnienie transformacji funkcji regresji spowodowane zastąpieniem funkcji próbkującej y przez złożenie $g \circ y$, gdzie g jest wielomianem stopnia pierwszego. Rozdział piąty przedstawia ilustrację praktyczną rozważań teoretycznych. Na podstawie twierdzenia 4.10 opracowano metodę numeryczną obliczania funkcji regresji dla dowolnie zadanego układu funkcji bazowych. Zaprezentowano również realne przykłady wyznaczania funkcji regresji oparte na implementacji tej metody w arkuszu kalkulacyjnym MS Excel przy użyciu języka programowania VBA. Kod programu został zamieszczony w załączniku. Ponadto znajdują się tam wynikowe zrzuty ekranu z wykonanych obliczeń numerycznych dla realnych przykładów zamieszczonych w rozdziale piątym.

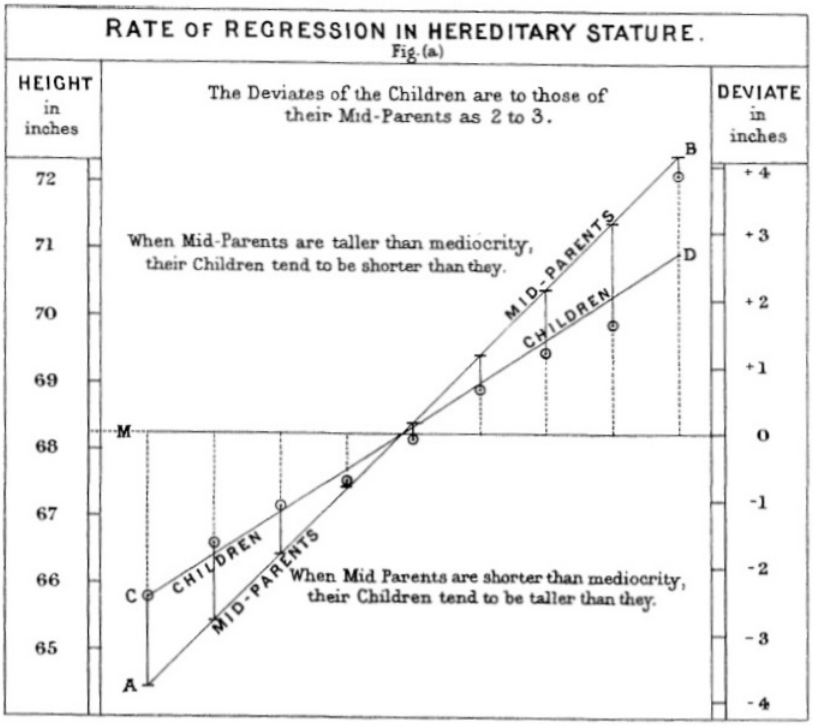
W całej pracy przyjmujemy oznaczenia $\mathbb{Z}_{p,q} := \{n \in \mathbb{Z} \colon p \leqslant n \leqslant q\}$ i $\mathbb{Z}_p := \{n \in \mathbb{Z} \colon p \leqslant n\}$ dla $p, q \in \mathbb{Z}$. W szczególności $\mathbb{N} = \mathbb{Z}_1$. Ponadto przyjmujemy, że wyrażenia typu $\int_\Omega X \mathrm{d}P$ i $\int_\Omega X(\omega) \mathrm{d}P(\omega)$ oznaczają tę samą wartość całki funkcji X na zbiorze Ω względem miary P.

Rozdział 1

Zagadnienie regresji

Analiza regresji zajmuje się badaniem zależności pomiędzy zmiennymi, mając na celu konstrukcję modelu, który dobrze odzwierciedla tę zależność. Takiego badania dokonujemy na podstawie przeprowadzenia doświadczeń i analizy ich wyników. Warto zaznaczyć, że słowo *regresja* w tłumaczeniu z łaciny oznacza *cofanie*. Stosowanie w nazwie tej dziedziny, jak też w nazwie kilku innych pojęć, słowa regresja jest historyczne i raczej niefortunne. Nazwa ta została prawdopodobnie po raz pierwszy użyta w 1885 r. przez angielskiego naukowca Sir F. Galtona (ucznia K. Darwina) podczas badania zależności wzrostu potomstwa od wzrostu rodziców. Galton był pionierem w stosowaniu metod statystycznych do pomiarów zjawisk w wielu gałęziach nauki oraz w badaniu danych dotyczących ludzkich różnic indywidualnych i dziedziczenia inteligencji oraz wprowadził kwestionariusze i ankiety do zbierania danych na temat społeczności ludzkich. Prowadził on między innymi badania empiryczne dotyczące względnych rozmiarów rodziców i ich potomstwa u różnych gatunków roślin i zwierząt. Jako pierwszy w 1886 roku opisał zjawisko regresji w kierunku średniej, które dotyczyło związku między wartością wzrostu dzieci a wzrostem ich rodziców. Wykazał on bowiem, że niezwykle wysocy rodzice (znacznie wyżsi powyżej średniej), mają dzieci niższego wzrostu, natomiast rodzice o wzroście znacznie niższym, niż przeciętny, mają dzieci wyższe od nich; zob. [6]. Galton zdał sobie sprawę, że potomstwo ma tendencję do powrotu do średniej wielkości i dlatego nazwał to zjawisko cofaniem w kierunku przeciętności (*regression towards mediocrity* - RTM); zob. [13]. Zjawisko to jest zobrazowane na rysunku 1.1. Krótkie poziome linie na skali z lewej strony wykresu odnoszą się do średniego wzrostu rodziców od $64,5$ do $72,5$ cali ze skalą podziałki jednego cala. Małe kółka na rysunku pokazują średni wzrost dzieci grupy rodziców o zadanym wzroście. Linia AB obrazuje średnie wzrosty rodziców, a linia CD przedstawia „wygładzone" wyniki odpowiadające średniemu wzrostowi ich dzieci. Stosunek długości odcinka CM do długości odcinka AM wynosi $2/3$. Punkt przecięcia linii AB i CD znajduje się na poziomie średnim $M = 68,25$ cala. Dystans pomiędzy liniami AB i CD obrazuje pojęcie *regression towards mediocrity*.

To samo dotyczy praktycznie każdego pomiaru fizycznego, a w przypadku ludzi – większości pomiarów zdolności poznawczych i fizycznych, które można wykonać na



Rysunek 1.1: Zjawisko regresji w kierunku średniej (RTM) rozumiane przez Galtona; zob. [13].

rodzicach i ich potomstwie; zob. [12], [31]. Od tego czasu opisano wiele przypadków występowania RTM. Wykazano, iż regresja w kierunku średniej ma miejsce zawsze wtedy, gdy dobór grupy badanej nie jest losowy, co powoduje, iż obserwowana cecha nie odzwierciedla stanu rzeczywistego w populacji. Brak losowości skutkuje tym, że średnia wartość ocenianej cechy jest albo wyższa niż w populacji, albo jest nieprawdziwie niska; zob. [3], [23]. Wśród ciekawych, wartych przytoczenia przykładów efektu RTM znajdują się obserwacje Streinera, który zauważył, że dzieci profesorów akademickich osiągają gorsze wyniki na egzaminach wstępnych niż ich rodzice, a dzieci milionerów, którzy sami osiągnęli sukces, odnoszą mniejsze sukcesy niż rodzice. Podkreśla ponadto, że osoby sprawujące najwyższe stanowiska po zakończeniu swojej kadencji wypalają się zawodowo i osiągają mniejsze sukcesy. Niewątpliwie w każdym z wymienionych przypadków grupa, w stosunku do której odnoszą się rezultaty, cechuje się wartością średnią badanej cechy wyższą niż jej rzeczywista wartość w populacji; zob. [23].

W istocie dziedzina zajmująca się poszukiwaniem zależności na podstawie prowadzenia doświadczeń jest o wiele starsza. Tak na przykład matematycy francuscy (szczególnie P. S. Laplace) w XVIII wieku przeprowadzali analizy, które nazwalibyśmy dziś analizą regresji. Pierwszą formą regresji była regresja liniowa otrzy-

mana metodą najmniejszych kwadratów. Po raz pierwszy w publikacji metody tej użył A. M. Legendre w dodatku do swojej książki *„Nouvelles méthods pour la détermination des orbites des cométes"*. Książkę tę wraz z suplementem opublikowano w 1805 roku; zob. [22]. Tymczasem w wydanej cztery lata później *„Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium"* (zob. [16]) G. F. Gauss stwierdził, że używa metody najmniejszych kwadratów od roku 1795; zob. [6]. Zarówno Legendre, jak i Gauss zastosowali tę metodę do rozwiązania problemu wynikającego z obserwacji astronomicznych, a dotyczącego orbit ciał poruszających się wokół Słońca – głównie komet, ale także odkrytych później mniejszych planet; zob. [32]. Gauss opublikował swoje wyniki dotyczące teorii najmniejszych kwadratów w 1821 r. (zob. [17]) w tym pierwszą wersję znanego dzisiaj twierdzenia Gaussa-Markowa.

Dla Galtona regresja miała znaczenie tylko biologiczne[1] (zob. [14], [15]), ale jego praca została później rozszerzona przez Yule'a (zob. [34]) i Pearsona (zob. [26]) na bardziej ogólny kontekst statystyczny. W pracach Yule'a i Pearsona zakłada się, że łączny rozkład zmiennych objaśnianych (zależnych) i zmiennych objaśniających (niezależnych) jest rozkładem Gaussa. To założenie zostało osłabione przez Fishera w jego pracach z 1922 i 1925 r.; zob. [1], [9], [10]. Fisher założył, że rozkład warunkowy zmiennej objaśnianej jest gaussowski, ale rozkład łączny nie musi nim być. Pod tym względem założenie Fishera jest bliższe sformułowaniu Gaussa z 1821 r.

W latach 50-tych i 60-tych XX wieku ekonomiści, wraz z wprowadzeniem na rynek komputerów, używali ich do wyznaczania regresji. Przed 1970 r. obliczanie jednej regresji czasem trwało do 24 godzin; zob. [28].

Metody regresji są nadal obszarem aktywnych badań. Zwykle w problemach regresji jedna ze zmiennych (zmienna zależna) oznaczana najczęściej przez y jest „kształtowana" przez zmienne oznaczane najczęściej przez $x_1, \ldots, x_k$ zwane zmiennymi niezależnymi, które mają wyjaśnić zachowanie zmiennej y. Jeśli dane sugerują zależność pomiędzy y i $x_1, \ldots, x_k$ to chcemy określić tę relację za pomocą pewnej funkcji f,

$$y = f(x_1, \ldots, x_k), \tag{1.1}$$

co pomoże przewidzieć y dla danych $x_1, \ldots, x_k$. Taka zależność może nie być spełniona ze względu na wahania przypadkowe czy błędy pomiaru. Często znana jest postać matematyczna zależności (1.1), za wyjątkiem parametrów, lub też zależność ta jest określona przez proces fizyczny czy pewne prawo. Innymi słowy związek (1.1)

[1]Galton używa terminu „reversion", gdy omawia wielkość groszku lub terminu „regresja", gdy omawia wzrost ludzi

przybiera postać

$$y = f(x_1, \ldots, x_k; \alpha),$$

gdzie α jest ciągiem parametrów rzeczywistych. Celem jest znalezienie takich ciągów α, by model jak najlepiej pasował do obserwacji. Jeśli funkcja f jest liniowa względem parametru α, to regresja (odp. model) nazywa się liniową (odp. liniowym), w przeciwnym przypadku – nieliniową (odp. nieliniowym). Najprostsza funkcja regresji liniowej jednej zmiennej, to $\mathbb{R} \ni x \mapsto f(x;\alpha) = \alpha_0 + \alpha_1 x$. Jak wśród wszystkich możliwych takich funkcji wybrać optymalną? Rozwiązanie tego zagadnienia opiera się na metodzie najmniejszych kwadratów, wprowadzonej jeszcze na początku XIX wieku przez A. M. Legendre'a i C. F. Gaussa. Polega ona na tym, że mając obserwacje $(x_i, y_i), i = 1, \ldots, n$, wybieramy α_1 i α_0 tak, by wartość

$$\sum_{i=1}^{n} \Big(y_i - (\alpha_0 + \alpha_1 x_i)\Big)^2$$

była jak najmniejsza. Okazuje się, że taki problem, o ile wszystkie wartości x_i nie są jednakowe, ma jednoznaczne rozwiązanie, wyrażone wzorami

$$\alpha_1 = \frac{\sum\limits_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum\limits_{i=1}^{n}(x_i - \bar{x})^2} \qquad \text{i} \qquad \alpha_0 = \bar{y} - \alpha_1 \bar{x},$$

gdzie $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$ oraz $\bar{y} = \frac{1}{n}\sum\limits_{i=1}^{n} y_i$; zob. [6], [25].

Przedmiotem badań w tej rozprawie jest opracowanie matematycznie spójnej i możliwie ogólnej teorii regresji, stosowanej w rachunku prawdopodobieństwa, gdzie wyznaczanie najlepiej dopasowanych przebiegów teoretycznych do danych empirycznych jest oparte na koncepcji metody najmniejszych kwadratów.

Niniejsza praca doktorska zainspirowana została artykułem *Generalized approach to the problem of regression*, napisanym przez D. Partykę i J. Zająca; zob. [25]. W pracy tej jest omówiony ogólny problem regresji. Zostało tam wprowadzone pojęcie *struktury regresji* $\mathfrak{R} := (A, B, \delta; x, y)$, która spełnia następujące warunki:

I.1 A, B są niepustymi zbiorami;

I.2 funkcje $x\colon \Omega_1 \to A$ i $y\colon \Omega_2 \to B$ są określone na niepustych zbiorach parametrów Ω_1 i Ω_2; są to tzw. funkcje próbkujące, czyli funkcje interpretowane jako dane eksperymentalne modelu regresji;

I.3 funkcja $\delta\colon (\Omega_1 \to B) \times (\Omega_2 \to B) \to \overline{\mathbb{R}}$ interpretowana jest jako kryterium odchylenia (dopasowania) funkcji teoretycznej od danych empirycznych, gdzie $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ jest rozszerzonym zbiorem liczb rzeczywistych.

Dla zadanej struktury regresji $\mathfrak{R}$ rozważano rodzinę funkcji $\mathcal{F}$ będącą podzbiorem rodziny wszystkich funkcji odwzorowujących zbiór A w zbiór B, tzn. $\mathcal{F} \subset (A \to B)$. Rodzinę $\mathcal{F}$ interpretuje się jako teoretyczny model funkcyjny rozważanego zjawiska, tzn. składa się ona ze wszystkich funkcji opisujących teoretycznie rozważane zjawisko. Może to być np. rodzina rozwiązań pewnego równania różniczkowego. Empirycznych wartości argumentów i wartości tych funkcji dostarczają odpowiednio funkcje próbkujące x i y. Funkcja δ pełni rolę kryterium dopasowania modelu teoretycznego $\mathcal{F}$ do danych empirycznych reprezentowanych przez funkcje x i y. Rodzina $\mathcal{F}$ jest nazywana *modelem funkcyjnym struktury regresji* $\mathfrak{R}$.

Naturalnym zagadnieniem dla zadanej struktury regresji $\mathfrak{R}$ jest badanie i wyznaczanie optymalnych funkcji modelu funkcyjnego $\mathcal{F}$, które są w sensie kryterium δ, najlepiej dopasowane do danych empirycznych reprezentowanych przez funkcje x i y. Ujmując to w języku matematycznym rozważano problem ekstremalny polegający na wyznaczeniu i badaniu klasy $\mathrm{Reg}(\mathcal{F}, \mathfrak{R})$ wszystkich funkcji $f_0 \in \mathcal{F}$ minimalizujących funkcjonał

$$\mathcal{F} \ni f \to F(f) := \delta(f \circ x, y) \in \overline{\mathbb{R}}, \tag{1.2}$$

czyli funkcji $f_0 \in \mathcal{F}$ spełniających warunek

$$F(f) \geqslant F(f_0), \qquad f \in \mathcal{F}.$$

Tak postawiony problem ekstremalny nazwano *problemem regresji* w klasie $\mathcal{F}$ dla struktury $\mathfrak{R}$. Każdą funkcję $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$ nazywa się *funkcją regresji* w klasie $\mathcal{F}$ względem struktury $\mathfrak{R}$.

W pracy [25] wyróżniono dwa typy struktur regresji opartych na koncepcji metody najmniejszych kwadratów: *asynchroniczne struktury regresji* oraz *synchroniczne struktury regresji*. Są one określone jak następuje.

Rozważmy rodzinę struktur regresji $\mathfrak{R} = (A, B, \delta; x, y)$ spełniającą następujące warunki:

II.1 $B = \mathbb{R}$ lub $B = \mathbb{C}$;

II.2 istnieje σ-ciało $\mathcal{B}$ podzbiorów iloczynu kartezjańskiego $\Omega_1 \times \Omega_2$ oraz miara $\mu\colon \mathcal{B} \to [0; +\infty]$ takie, że funkcja $\delta\colon (\Omega_1 \to B) \times (\Omega_2 \to B) \to \overline{\mathbb{R}}$ spełnia następującą równość

$$\delta(u, v) = \int_{\Omega_1 \times \Omega_2} |u(t_1) - v(t_2)|^2 \mathrm{d}\mu(t_1, t_2), \tag{1.3}$$

o ile funkcje $\Omega_1 \times \Omega_2 \ni (t_1, t_2) \mapsto u(t_1)$ i $\Omega_1 \times \Omega_2 \ni (t_1, t_2) \mapsto v(t_2)$ są $\mathcal{B}$-mierzalne oraz $\delta(u, v) = +\infty$ w przeciwnym przypadku.

Przy tych warunkach problem regresji rozumiany jest jako problem ekstremalny polegający na wyznaczeniu wszystkich funkcji $f_0 \in \mathcal{F}$ minimalizujących funkcjonał F, który z uwagi na (1.2) i (1.3) spełnia następującą równość

$$F(f) = \int_{\Omega_1 \times \Omega_2} |f \circ x(t_1) - y(t_2)|^2 \mathrm{d}\mu(t_1, t_2), \quad f \in \mathcal{F}.$$

Definicja 1.1. Każdą strukturę regresji $\mathfrak{R}$ spełniającą warunki II.1 i II.2 nazywamy *asynchroniczną strukturą regresji* (ARS).

W pracy [25] problem regresji dla asynchronicznej struktury regresji został szeroko omówiony i rozwiązany w przypadku, gdy model funkcyjny $\mathcal{F}$ jest liniowy.

Rozważmy teraz przypadek gdy funkcje próbkujące x i y zależą od tego samego parametru t, czyli $t_1 = t_2 = t$. Rozważamy wtedy rodzinę struktur regresji $\mathfrak{R} = (A, B, \delta; x, y)$ spełniającą następujące warunki:

III.1 $\Omega_1 = \Omega_2$;

III.2 $B = \mathbb{R}$ lub $B = \mathbb{C}$;

III.3 istnieje σ-ciało $\mathcal{A}$ podzbiorów zbioru $\Omega := \Omega_1$ oraz miara $\nu \colon \mathcal{A} \to [0; +\infty]$ takie, że funkcja $\delta \colon (\Omega \to B) \times (\Omega \to B) \to \overline{\mathbb{R}}$ spełnia następującą równość

$$\delta(u, v) = \int_{\Omega} |u(t) - v(t)|^2 \mathrm{d}\nu(t), \tag{1.4}$$

o ile $u, v \colon \Omega \to B$ są funkcjami $\mathcal{A}$-mierzalnymi oraz $\delta(u, v) = +\infty$ w przeciwnym przypadku.

Przy tych warunkach problem regresji oznacza problem ekstremalny polegający na wyznaczeniu wszystkich funkcji $f_0 \in \mathcal{F}$ minimalizujący funkcjonał F spełniający z uwagi na (1.2) i (1.4) następującą równość

$$F(f) = \int_{\Omega} |f \circ x(t) - y(t)|^2 \mathrm{d}\nu(t), \quad f \in \mathcal{F}.$$

Definicja 1.2. Każdą strukturę regresji $\mathfrak{R}$ spełniającą warunki III.1, III.2 i III.3 nazywamy *synchroniczną strukturą regresji* (SRS).

W pracy [25] wyniki dla asynchronicznych struktur regresji zostały przeniesione na synchroniczne struktury regresji za pomocą tzw. *metody diagonalnej* redukcji ARS do SRS.

Niniejsza rozprawa dotyczy synchronicznych struktur regresji określonych za pomocą przestrzeni probabilistycznych. W odróżnieniu od pracy [25] uzyskano szereg wyników w alternatywny sposób , bez odwoływania się do asynchronicznych struktur regresji. To pozwoliło uprościć rozważania oraz uzyskać nowe rezultaty precyzyjniej charakteryzujące klasy funkcji regresji dla probabilistycznych struktur regresji.

Rozdział 2

Całkowanie funkcji zespolonych

Niniejsza dysertacja dotyczy w znacznej mierze funkcji zespolonych, więc w tym rozdziale przypomnimy pojęcie całki Lebesque'a funkcji zespolonej oraz jej podstawowe własności.

Dla zadanego niepustego zbioru Ω i σ-ciała $\mathcal{B}$ jego podzbiorów, oznaczamy przez $\mathrm{L}(\Omega,\mathcal{B})$ rodzinę wszystkich funkcji $f\colon \Omega \to \mathbb{C}$, mierzalnych względem $\mathcal{B}$, tzn. $f^{-1}(A) \in \mathcal{B}$ dla każdego zbioru $A \in \mathrm{Bor}(\mathbb{C})$. Symbol $\mathrm{Bor}(\mathbb{C})$ oznacza klasę wszystkich zbiorów borelowskich płaszczyzny zespolonej. Ponieważ funkcje $\mathbb{C} \ni z \mapsto \operatorname{Re} z \in \mathbb{R}$ i $\mathbb{C} \ni z \mapsto \operatorname{Im} z \in \mathbb{R}$ są ciągłe, więc dla każdej funkcji $f\colon \Omega \to \mathbb{C}$, $f \in \mathrm{L}(\Omega,\mathcal{B})$ wtedy i tylko wtedy, gdy $\operatorname{Re} f$ i $\operatorname{Im} f$ są funkcjami borelowskimi w sensie rzeczywistym.

Niech $\mathbf{L}(\Omega,\mathcal{B})$ oznacza zespoloną przestrzeń liniową o nośniku $\mathrm{L}(\Omega,\mathcal{B})$ ze standardowymi operacjami dodawania funkcji i mnożenia funkcji przez stałą, tzn. $\mathbf{L}(\Omega,\mathcal{B}) := (\mathrm{L}(\Omega,\mathcal{B});+,\cdot)$. Dla zadanej miary $\mu\colon \mathcal{B} \to [0;+\infty]$ i $p \in [1;+\infty)$, niech $\mathrm{L}^p(\Omega,\mathcal{B},\mu)$ oznacza klasę wszystkich funkcji $f \in \mathrm{L}(\Omega,\mathcal{B})$ takich, że

$$\|f\|_p := \left(\int_\Omega |f|^p \mathrm{d}\mu\right)^{1/p} < +\infty; \tag{2.1}$$

por. [30, Definition 3.6]. Wiadomo, że dla każdego $p \geqslant 1$, klasa $\mathrm{L}^p(\Omega,\mathcal{B},\mu)$ jest zbiorem liniowym w $\mathbf{L}(\Omega,\mathcal{B})$ (por. [20, §C.5]) i $\|\cdot\|_p$ jest pseudo-normą w przestrzeni liniowej $(\mathrm{L}^p(\Omega,\mathcal{B},\mu);+,\cdot)$ spełniającą następujący warunek

$$\|f\|_p = 0 \;\Leftrightarrow\; \mu(\{\omega \in \Omega\colon f(\omega) \neq 0\}) = 0.$$

Zatem struktura

$$\mathbf{L}^p(\Omega,\mathcal{B},\mu) := (\mathrm{L}^p(\Omega,\mathcal{B},\mu);+,\cdot,\|\cdot\|_p)$$

jest zespoloną przestrzenią pseudo-Banacha, tzn. zespoloną przestrzenią pseudo-unormowaną, która jest zupełna. Przestrzenie pseudo-Banacha są naturalnym uogólnieniem przestrzeni Banacha (por. np. [2, §2.4], [30, Theorem 3.11]), gdzie różnica polega na zastąpieniu normy przez pseudo-normę; por. [5, Chap. III, Definition 1.1].

Definicja 2.1. Funkcję zespoloną $f \in \mathrm{L}(\Omega,\mathcal{B})$ nazywamy *μ-całkowalną* wtedy i tylko wtedy, gdy $\operatorname{Re} f$ i $\operatorname{Im} f$ są funkcjami μ-całkowalnymi w sensie rzeczywistym, tzn. $\operatorname{Re} f$ i $\operatorname{Im} f$ są funkcjami $\mathcal{B}$-mierzalnymi oraz

$$\int_\Omega |\operatorname{Re} f|\mathrm{d}\mu < +\infty \qquad \text{i} \qquad \int_\Omega |\operatorname{Im} f|\mathrm{d}\mu < +\infty.$$

Uwaga 2.2. Dla każdego $f \in \mathrm{L}(\Omega, \mathcal{B})$, f jest funkcją μ-całkowalną wtedy i tylko wtedy, gdy $f \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu)$. Z warunku $|z| \leqslant |\operatorname{Re} z| + |\operatorname{Im} z|$ oraz z definicji 2.1 mamy bowiem

$$\int_\Omega |f| \mathrm{d}\mu \leqslant \int_\Omega \big(|\operatorname{Re} f| + |\operatorname{Im} f|\big) \mathrm{d}\mu < +\infty,$$

a więc $f \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu)$. Na odwrót, jeśli $f \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu)$, to

$$\int_\Omega |\operatorname{Re} f| \mathrm{d}\mu \leqslant \int_\Omega |f| \mathrm{d}\mu < +\infty \qquad \text{i} \qquad \int_\Omega |\operatorname{Im} f| \mathrm{d}\mu \leqslant \int_\Omega |f| \mathrm{d}\mu < +\infty,$$

a więc f jest funkcją μ-całkowalną. W konsekwencji dla każdego $f \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu)$ można określić całkę z funkcji f względem miary μ następującym wzorem

$$\int_\Omega f \mathrm{d}\mu := \int_\Omega \operatorname{Re} f \mathrm{d}\mu + \mathrm{i} \int_\Omega \operatorname{Im} f \mathrm{d}\mu;$$

por. [2, Definition 2.4.1].

Niech $(\Omega, \mathcal{B})$ i $(\Omega', \mathcal{B}')$ będą przestrzeniami mierzalnymi i niech $x \colon \Omega \to \Omega'$ będzie odwzorowaniem spełniającym warunek

$$x^{-1}(A') \in \mathcal{B}, \qquad A' \in \mathcal{B}'. \tag{2.2}$$

Wtedy dla dowolnej funkcji $\mu \colon \mathcal{B} \to [0; +\infty]$, funkcja $\mu' \colon \mathcal{B}' \to [0; +\infty]$ jest dobrze określona wzorem

$$\mu'(A') := \mu(x^{-1}(A')), \qquad A' \in \mathcal{B}'. \tag{2.3}$$

Jeśli μ jest miarą na $\mathcal{B}$, to na mocy wzoru (2.3) funkcja μ' jest miarą na $\mathcal{B}'$; por. [4, (13.7)]. W dalszych rozważaniach będziemy korzystać z następującego faktu – o całkowaniu przez podstawianie.

Lemat 2.3. *Niech $(\Omega, \mathcal{B})$ i $(\Omega', \mathcal{B}')$ będą przestrzeniami mierzalnymi. Wtedy dla każdej funkcji $x \colon \Omega \to \Omega'$ spełniającej warunek (2.2) i każdej funkcji $\mathcal{B}'$-mierzalnej $f \colon \Omega' \to \mathbb{C}$,*

$$f \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu') \;\Leftrightarrow\; f \circ x \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu) \tag{2.4}$$

oraz

$$\int_\Omega f \circ x \mathrm{d}\mu = \int_{\Omega'} f \mathrm{d}\mu', \quad f \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu'), \tag{2.5}$$

gdzie μ' jest miarą na $\mathcal{B}'$ określoną wzorem (2.3).

Dowód. Dowody własności (2.4) i (2.5) w przypadku rzeczywistym, czyli gdy $f \colon \Omega' \to \mathbb{R}$, można znaleźć na przykład w [2, Theorem 1.6.12] lub [4, Theorem 16.13]. Pozostaje zatem wykazać własności (2.4) i (2.5) w przypadku zespolonym, czyli gdy

$f\colon \Omega' \to \mathbb{C}$. Ponieważ $\mathbb{C} \ni z \mapsto |z|$ jest funkcją ciągłą, więc dla dowolnie ustalonej funkcji $\mathcal{B}'$-mierzalnej $f\colon \Omega' \to \mathbb{C}$ funkcja $|f|$ jest $\mathcal{B}'$-mierzalna i ma wartości rzeczywiste. Możemy więc wykorzystać własność (2.4) dla przypadku rzeczywistego. W rezultacie dostajemy równoważność

$$|f| \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu') \;\Leftrightarrow\; |f \circ x| \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu). \tag{2.6}$$

Ponieważ

$$|f| \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu') \;\Leftrightarrow\; f \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu')$$

oraz

$$|f \circ x| \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu) \;\Leftrightarrow\; f \circ x \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu),$$

więc z równoważności (2.6) wynika równoważność (2.4) w przypadku zespolonym. Ponadto dla dowolnie zadanego $f \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu')$

$$\int_{\Omega'} |\operatorname{Re} f| \mathrm{d}\mu' \leqslant \int_{\Omega'} |f| \mathrm{d}\mu' < +\infty \quad \text{i} \quad \int_{\Omega'} |\operatorname{Im} f| \mathrm{d}\mu' \leqslant \int_{\Omega'} |f| \mathrm{d}\mu' < +\infty,$$

skąd $\operatorname{Re} f, \operatorname{Im} f \in \mathrm{L}^1(\Omega', \mathcal{B}', \mu')$. Ponieważ $\operatorname{Re} f$ i $\operatorname{Im} f$ są funkcjami rzeczywistymi więc stosując (2.5) dla przypadku rzeczywistego dostajemy

$$\begin{aligned} \int_{\Omega} f \circ x \mathrm{d}\mu &= \int_{\Omega} \operatorname{Re}(f \circ x) \mathrm{d}\mu + \mathrm{i} \int_{\Omega} \operatorname{Im}(f \circ x) \mathrm{d}\mu \\ &= \int_{\Omega'} \operatorname{Re} f \mathrm{d}\mu' + \mathrm{i} \int_{\Omega'} \operatorname{Im} f \mathrm{d}\mu' = \int_{\Omega'} f \mathrm{d}\mu', \end{aligned}$$

co dowodzi własności (2.5) w przypadku zespolonym. □

Niech $(\Omega'', \mathcal{B}'')$ będzie przestrzenią mierzalną i niech $y\colon \Omega \to \Omega''$ będzie odwzorowaniem spełniającym warunek

$$y^{-1}(A'') \in \mathcal{B}, \qquad A'' \in \mathcal{B}''.$$

Wtedy funkcja $\mu\colon \mathcal{B} \to [0; +\infty]$ generuje funkcję $\mu''\colon \mathcal{B}'' \to [0; +\infty]$ określoną wzorem

$$\mu''(A'') := \mu(y^{-1}(A'')), \qquad A'' \in \mathcal{B}''. \tag{2.7}$$

Jeżeli μ jest miarą na $\mathcal{B}$, to ze wzoru (2.7) wynika, że funkcja μ'' jest miarą na $\mathcal{B}''$.

Oznaczmy przez $\pi_{x,y}$ miarę określoną na σ-ciele $\mathcal{B}' \otimes \mathcal{B}''$ spełniającą warunek

$$\pi_{x,y}(I \times J) = \mu(x^{-1}(I) \cap y^{-1}(J)), \quad I \in \mathcal{B}',\ J \in \mathcal{B}'', \tag{2.8}$$

gdzie $\mathcal{B}' \otimes \mathcal{B}''$ jest najmniejszym, w sensie inkluzji, σ-ciałem zawierającym wszystkie zbiory $I \times J$ dla $I \in \mathcal{B}'$ i $J \in \mathcal{B}''$. Istnienie i jednoznaczność miary $\pi_{x,y}$ jest konsekwencją [20, §C.3.].

Niech $\mathbb{1}_A$ oznacza funkcję charakterystyczną zbioru A względem Ω; por. [2, Sect. 1.5.1], [30, Sect. 1.9], tzn.

$$\Omega \ni \omega \mapsto \mathbb{1}_A(\omega) := \begin{cases} 1, & \text{gdy } \omega \in A \\ 0, & \text{gdy } \omega \in \Omega \backslash A. \end{cases}$$

Następujący lemat jest uogólnieniem lematu [25, Lemma 9.1].

Lemat 2.4. *Niech $(\Omega, \mathcal{B})$, $(\Omega', \mathcal{B}')$ i $(\Omega'', \mathcal{B}'')$ będą przestrzeniami mierzalnymi. Wtedy dla dowolnych funkcji $f\colon \Omega' \to \mathbb{C}$ oraz $g\colon \Omega'' \to \mathbb{C}$, jeśli f jest $\mathcal{B}'$-mierzalna i g jest $\mathcal{B}''$-mierzalna, to funkcja $(f\circ x)(g\circ y)$ jest całkowalna względem miary μ wtedy i tylko wtedy gdy $\Omega' \times \Omega'' \ni (t_1, t_2) \to f(t_1)g(t_2)$ jest funkcją całkowalną względem miary $\pi_{x,y}$. Ponadto, jeśli jedna z tych funkcji jest całkowalna to zachodzi równość*

$$\int_\Omega (f \circ x(\omega))(g \circ y(\omega))\, \mathrm{d}\mu(\omega) = \int_{\Omega' \times \Omega''} f(t_1)g(t_2)\, \mathrm{d}\pi_{x,y}(t_1, t_2). \tag{2.9}$$

Dowód. Załóżmy najpierw, że f i g są funkcjami prostymi postaci

$$f = \sum_{k=1}^{n} \lambda_k \mathbb{1}_{I_k} \qquad \text{i} \qquad g = \sum_{l=1}^{m} \mu_l \mathbb{1}_{J_l}$$

dla pewnych $n, m \in \mathbb{N}$ oraz ciągów $\mathbb{Z}_{1,n} \ni k \mapsto \lambda_k \in \mathbb{C}$, $\mathbb{Z}_{1,n} \ni k \mapsto I_k \in \mathcal{B}'$, $\mathbb{Z}_{1,m} \ni l \mapsto \mu_l \in \mathbb{C}$ i $\mathbb{Z}_{1,m} \ni l \mapsto J_l \in \mathcal{B}''$, przy czym $I_k \cap I_l = \emptyset$ i $J_k \cap J_l = \emptyset$, gdy $k \neq l$. Z definicji całki Lebesgue'a oraz (2.8) mamy

$$\begin{aligned}
\int_{\Omega' \times \Omega''} f(t_1)g(t_2)\, \mathrm{d}\pi_{x,y}(t_1, t_2) &= \int_{\Omega' \times \Omega''} \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\, \mathbb{1}_{I_k}(t_1)\ \mu_l\, \mathbb{1}_{J_l}(t_2)\, \mathrm{d}\pi_{x,y}(t_1, t_2) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l \int_{\Omega' \times \Omega''} \mathbb{1}_{I_k}(t_1)\, \mathbb{1}_{J_l}(t_2)\, \mathrm{d}\pi_{x,y}(t_1, t_2) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l \int_{\Omega' \times \Omega''} \mathbb{1}_{I_k \times J_l}(t_1, t_2)\, \mathrm{d}\pi_{x,y}(t_1, t_2) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l\ \pi_{x,y}(I_k \times J_l) = \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l\ \mu\left(x^{-1}(I_k) \cap y^{-1}(J_l)\right) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l \int_\Omega \mathbb{1}_{x^{-1}(I_k)\cap y^{-1}(J_l)}(\omega)\, \mathrm{d}\mu(\omega) \\
&= \int_\Omega \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l\, \mathbb{1}_{x^{-1}(I_k)\cap y^{-1}(J_l)}(\omega)\, \mathrm{d}\mu(\omega) \\
&= \int_\Omega \sum_{k=1}^{n}\sum_{l=1}^{m} \lambda_k\mu_l\, \mathbb{1}_{x^{-1}(I_k)}(\omega)\, \mathbb{1}_{y^{-1}(J_l)}(\omega)\, \mathrm{d}\mu(\omega) \\
&= \int_\Omega \left(\sum_{k=1}^{n} \lambda_k\, \mathbb{1}_{x^{-1}(I_k)}(\omega)\right)\left(\sum_{l=1}^{m} \mu_l\, \mathbb{1}_{y^{-1}(J_l)}(\omega)\right) \mathrm{d}\mu(\omega) \\
&= \int_\Omega \Big(f \circ x(\omega)\Big)\Big(g \circ y(\omega)\Big)\, \mathrm{d}\mu(\omega),
\end{aligned}$$

gdyż

$$f \circ x(\omega) = \sum_{k=1}^{n} \lambda_k \, \mathbb{1}_{x^{-1}(I_k)}(\omega) \quad \text{i} \quad g \circ y(\omega) = \sum_{l=1}^{m} \mu_l \, \mathbb{1}_{y^{-1}(J_l)}(\omega), \quad \omega \in \Omega.$$

Zatem lemat jest prawdziwy dla funkcji prostych f i g.

Załóżmy teraz, że $f\colon \Omega' \to \mathbb{C}$ i $g\colon \Omega'' \to \mathbb{C}$ są dowolnie ustalonymi funkcjami odpowiednio $\mathcal{B}'$ i $\mathcal{B}''$-mierzalnymi. Wówczas $|f|$ jest nieujemną funkcją $\mathcal{B}'$-mierzalną, zaś $|g|$ jest nieujemną funkcją $\mathcal{B}''$-mierzalną. Istnieją więc niemalejące ciągi funkcji prostych $\mathbb{N} \ni n \mapsto \big(\Omega' \ni t \mapsto f_n(t) \geqslant 0\big)$ i $\mathbb{N} \ni n \mapsto \big(\Omega'' \ni t \mapsto g_n(t) \geqslant 0\big)$ takie, że $f_n \to |f|$ na Ω' i $g_n \to |g|$ na Ω'' gdy $n \to +\infty$. Z twierdzenia o iloczynie ciągów zbieżnych wnosimy, że ciąg $\mathbb{N} \ni n \mapsto \big(\Omega' \times \Omega'' \ni (t_1, t_2) \mapsto f_n(t_1) g_n(t_2)\big)$ jest niemalejącym ciągiem funkcji prostych i dla każdego $(t_1, t_2) \in \Omega' \times \Omega''$,

$$f_n(t_1) g_n(t_2) \to |f(t_1)||g(t_2)| = |f(t_1) g(t_2)|, \quad \text{gdy} \quad n \to +\infty.$$

Wtedy również $\mathbb{N} \ni n \mapsto \big(\Omega \ni \omega \mapsto (f_n \circ x(\omega))(g_n \circ y(\omega)\big)$ jest niemalejącym ciągiem funkcji prostych i dla każdego $\omega \in \Omega$,

$$(f_n \circ x(\omega))(g_n \circ y(\omega)) \to |f \circ x(\omega)||g \circ y(\omega)| = |(f \circ x(\omega))(g \circ y(\omega))|, \quad \text{gdy} \quad n \to +\infty.$$

Stąd na mocy definicji całki Lebesgue'a mamy

$$\int_{\Omega' \times \Omega''} |f(t_1) g(t_2)| \mathrm{d}\pi_{x,y}(t_1, t_2) = \lim_{n \to +\infty} \int_{\Omega' \times \Omega''} f_n(t_1) g_n(t_2) \mathrm{d}\pi_{x,y}(t_1, t_2) \tag{2.10}$$

oraz

$$\int_{\Omega} |(f \circ x(\omega))(g \circ y(\omega))| \, \mathrm{d}\mu(\omega) = \lim_{n \to +\infty} \int_{\Omega} (f_n \circ x(\omega))(g_n \circ y(\omega)) \, \mathrm{d}\mu(\omega). \tag{2.11}$$

Korzystając z udowodnionej już części lematu widzimy, że

$$\int_{\Omega' \times \Omega''} f_n(t_1) g_n(t_2) \, \mathrm{d}\pi_{x,y}(t_1, t_2) = \int_{\Omega} (f_n \circ x(\omega))(g_n \circ y(\omega)) \, \mathrm{d}\mu(\omega), \quad n \in \mathbb{N}. \tag{2.12}$$

Równości (2.10), (2.11) i (2.12) implikują łącznie równość

$$\int_{\Omega} |f \circ x(\omega)||g \circ y(\omega)| \, \mathrm{d}\mu(\omega) = \int_{\Omega' \times \Omega''} |f(t_1)||g(t_2)| \, \mathrm{d}\pi_{x,y}(t_1, t_2), \tag{2.13}$$

z której wynika równoważność w lemacie.

Załóżmy teraz, że funkcja $(f \circ x)(g \circ y)$ jest całkowalna względem miary μ. Rozważmy najpierw przypadek, gdy obie funkcje są rzeczywiste. Przyjmując:

$$f_+ := \frac{1}{2}(|f| + f), \quad f_- := \frac{1}{2}(|f| - f), \quad g_+ := \frac{1}{2}(|g| + g), \quad g_- := \frac{1}{2}(|g| - g)$$

mamy $f_+, f_- \geqslant 0$ na Ω' i $g_+, g_- \geqslant 0$ na Ω'' oraz

$$f_+ - f_- = f, \quad f_+ + f_- = |f|, \quad g_+ - g_- = g, \quad g_+ + g_- = |g|. \tag{2.14}$$

Stąd

$$|(f_\alpha \circ x)\,(g_\beta \circ y)| = (f_\alpha \circ x)(g_\beta \circ y) \leqslant |f \circ x||g \circ y|$$

na Ω dla $\alpha, \beta \in \{+, -\}$, i w konsekwencji

$$\int_\Omega |f_\alpha \circ x(\omega)|\,|g_\beta \circ y(\omega)|\,\mathrm{d}\mu(\omega) \leqslant \int_\Omega |f \circ x(\omega)||g \circ y(\omega)|\,\mathrm{d}\mu(\omega) < +\infty, \quad \alpha, \beta \in \{+, -\}.$$

Stosując więc równość (2.13) odpowiednio z podstawieniem $f := f_\alpha$ i $g := g_\beta$ dostajemy

$$\begin{aligned}\int_\Omega \Big(f_\alpha \circ x(\omega)\Big)\Big(g_\beta \circ y(\omega)\Big)\,\mathrm{d}\mu(\omega) &= \int_\Omega |f_\alpha \circ x(\omega)||g_\beta \circ y(\omega)|\,\mathrm{d}\mu(\omega)\\ &= \int_{\Omega' \times \Omega''} |f_\alpha(t_1)||g_\beta(t_2)|\,\mathrm{d}\pi_{x,y}(t_1, t_2)\\ &= \int_{\Omega' \times \Omega''} f_\alpha(t_1) g_\beta(t_2)\,\mathrm{d}\pi_{x,y}(t_1, t_2), \quad \alpha, \beta \in \{+, -\}.\end{aligned}$$

To łącznie z (2.14) daje

$$\begin{aligned}&\int_\Omega \Big(f \circ x(\omega)\Big)\Big(g \circ y(\omega)\Big)\,\mathrm{d}\mu(\omega)\\ &\qquad= \int_\Omega \Big(f_+ \circ x(\omega) - f_- \circ x(\omega)\Big)\Big(g_+ \circ y(\omega) - g_- \circ y(\omega)\Big)\,\mathrm{d}\mu(\omega)\\ &\qquad= \sum_{\alpha,\beta \in \{+,-\}} \mathrm{sn}(\alpha, \beta) \int_\Omega \Big(f_\alpha \circ x(\omega)\Big)\Big(g_\beta \circ y(\omega)\Big)\,\mathrm{d}\mu(\omega)\\ &\qquad= \sum_{\alpha,\beta \in \{+,-\}} \mathrm{sn}(\alpha, \beta) \int_{\Omega' \times \Omega''} f_\alpha(t_1) g_\beta(t_2)\,\mathrm{d}\pi_{x,y}(t_1, t_2)\\ &\qquad= \int_{\Omega' \times \Omega''} (f_+ - f_-)(t_1)(g_+ - g_-)(t_2)\,\mathrm{d}\pi_{x,y}(t_1, t_2)\\ &\qquad= \int_{\Omega' \times \Omega''} f(t_1) g(t_2)\,\mathrm{d}\pi_{x,y}(t_1, t_2),\end{aligned}$$

gdzie

$$\mathrm{sn}(\alpha, \beta) := \begin{cases} 1, & \text{gdy } \alpha = \beta\\ -1, & \text{gdy } \alpha \neq \beta.\end{cases}$$

To dowodzi równości (2.9) w przypadku gdy obie funkcje są rzeczywiste.
W ogólnym przypadku funkcji zespolonych rozważamy funkcje:

$$f_1 := \mathrm{Re}\, f, \quad f_2 := \mathrm{Im}\, f, \quad g_1 := \mathrm{Re}\, g, \quad g_2 := \mathrm{Im}\, g.$$

Ponieważ

$$|(f_\alpha \circ x)(g_\beta \circ y)| = |f_\alpha \circ x|\,|g_\beta \circ y| \leqslant |f \circ x|\,|g \circ y|, \quad \alpha, \beta \in \mathbb{Z}_{1,2},$$

więc $(f_\alpha \circ x)(g_\beta \circ y) \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu) \cap (\Omega \to \mathbb{R})$ dla $\alpha, \beta \in \mathbb{Z}_{1,2}$. Stosując zatem równość (2.9) dla przypadku funkcji rzeczywistych otrzymujemy

$$\begin{aligned}
\int_\Omega \big(f \circ x(\omega)\big)\big(g \circ y(\omega)\big)\mathrm{d}\mu(\omega) &= \int_\Omega \big((f_1 + \mathrm{i}f_2) \circ x(\omega)\big)\big((g_1 + \mathrm{i}g_2) \circ y(\omega)\big)\,\mathrm{d}\mu(\omega) \\
&= \int_\Omega \big(f_1 \circ x(\omega)\big)\big(g_1 \circ y(\omega)\big)\mathrm{d}\mu(\omega) - \int_\Omega \big(f_2 \circ x(\omega)\big)\big(g_2 \circ y(\omega)\big)\mathrm{d}\mu(\omega) \\
&\quad + \mathrm{i}\int_\Omega \big(f_1 \circ x(\omega)\big)\big(g_2 \circ y(\omega)\big)\mathrm{d}\mu(\omega) + \mathrm{i}\int_\Omega \big(f_2 \circ x(\omega)\big)\big(g_1 \circ y(\omega)\big)\mathrm{d}\mu(\omega) \\
&= \int_{\Omega' \times \Omega''} f_1(t_1)g_1(t_2)\mathrm{d}\pi_{x,y}(t_1, t_2) - \int_{\Omega' \times \Omega''} f_2(t_1)g_2(t_2)\mathrm{d}\pi_{x,y}(t_1, t_2) \\
&\quad + \mathrm{i}\int_{\Omega' \times \Omega''} f_1(t_1)g_2(t_2)\mathrm{d}\pi_{x,y}(t_1, t_2) + \mathrm{i}\int_{\Omega' \times \Omega''} f_2(t_1)g_1(t_2)\mathrm{d}\pi_{x,y}(t_1, t_2) \\
&= \int_{\Omega' \times \Omega''} (f_1 + \mathrm{i}f_2)(t_1)(g_1 + \mathrm{i}g_2)(t_2)\mathrm{d}\pi_{x,y}(t_1, t_2) \\
&= \int_{\Omega' \times \Omega''} f(t_1)g(t_2)\mathrm{d}\pi_{x,y}(t_1, t_2),
\end{aligned}$$

co daje równość (2.9) w ogólnym przypadku funkcji zespolonych f i g. □

Lemat 2.5. *Dla dowolnych $p \in [1; +\infty)$ i $f\colon \Omega \to \mathbb{R}$, jeśli $f \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$, to istnieje ciąg funkcji prostych $\mathbb{N} \ni n \mapsto f_n \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ taki, że $f_n(\Omega) \subset \mathbb{R}$ dla $n \in \mathbb{N}$ i*

$$\int_\Omega |f - f_n|^p \,\mathrm{d}\mu \to 0, \quad \textit{gdy} \;\; n \to +\infty. \tag{2.15}$$

Dowód. Ustalmy dowolnie $p \in [1; +\infty)$ i funkcję rzeczywistą $f \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$. Załóżmy najpierw, że $f(\omega) \geqslant 0$ dla $\omega \in \Omega$. Wtedy dla każdego $n \in \mathbb{N}$ funkcja

$$\Omega \ni \omega \mapsto f_n(\omega) := \sum_{k=1}^{4^n} \frac{k-1}{2^n}\, \mathbb{1}_{A_{n,k}} \tag{2.16}$$

jest prosta, gdzie dla każdego $n \in \mathbb{N}$,

$$A_{n,k} := f^{-1}\left(\left[\frac{k-1}{2^n}; \frac{k}{2^n}\right)\right), \qquad k \in \mathbb{Z}_{1,4^n}. \tag{2.17}$$

Ze wzoru (2.16) wynika, że $f_n(\Omega) \subset [0; +\infty)$ dla $n \in \mathbb{N}$. Ze wzorów (2.16) i (2.17) wnioskujemy ponadto, że

$$0 \leqslant f_n(\omega) \leqslant f(\omega), \qquad n \in \mathbb{N}, \;\; \omega \in \Omega \tag{2.18}$$

oraz

$$f_n(\omega) \to f(\omega), \quad \text{gdy} \;\; n \to +\infty, \;\; \omega \in \Omega. \tag{2.19}$$

Z (2.18) dostajemy

$$|f - f_n|^p \leqslant (|f| + |f_n|)^p = (f + f_n)^p \leqslant 2^p f^p \;\; \text{w} \;\; \Omega, \;\; n \in \mathbb{N}.$$

Z (2.19) wynika zaś, że $|f(\omega) - f_n(\omega)|^p \to 0$, gdy $n \to +\infty$, dla $\omega \in \Omega$. Ponieważ $2^p f^p \in \mathrm{L}^1(\Omega, \mathcal{B}, \mu)$, więc stosując twierdzenie Lebesque'a o ograniczonym przechodzeniu do granicy pod znakiem całki (por. np. [2, §1.6.9]) dostajemy własność (2.15). Ponadto na mocy nierówności (2.18),

$$\int_\Omega |f_n|^p \mathrm{d}\mu = \int_\Omega f_n^p \mathrm{d}\mu \leqslant \int_\Omega f^p \mathrm{d}\mu = \int_\Omega |f|^p \mathrm{d}\mu < +\infty, \quad n \in \mathbb{N},$$

a więc $f_n \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ dla $n \in \mathbb{N}$. Tym samym lemat został wykazany przy dodatkowym założeniu, że $f \geqslant 0$ w Ω. Załóżmy teraz, że funkcja f nie spełnia tego założenia. Wtedy $f = f_+ - f_-$, gdzie

$$f_+ := \frac{1}{2}(|f| + f) \geqslant 0 \quad \text{i} \quad f_- := \frac{1}{2}(|f| - f) \geqslant 0 \ \text{ w } \Omega. \tag{2.20}$$

Stąd

$$\begin{aligned}\int_\Omega |f_+|^p \mathrm{d}\mu + \int_\Omega |f_-|^p \mathrm{d}\mu &= \int_\Omega f_+^p \mathrm{d}\mu + \int_\Omega f_-^p \mathrm{d}\mu = \int_\Omega (f_+^p + f_-^p)\mathrm{d}\mu \\ &= \int_\Omega (f_+ + f_-)^p \mathrm{d}\mu = \int_\Omega |f| \mathrm{d}\mu < +\infty.\end{aligned}$$

Zatem $f_+, f_- \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$. To łącznie z (2.20) oznacza, że do funkcji rzeczywistych f_+ i f_- można zastosować lemat 2.5 przy założeniu nieujemności funkcji f. W konsekwencji istnieją ciągi funkcji rzeczywistych prostych $\mathbb{N} \ni n \mapsto f_{+,n} \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ i $\mathbb{N} \ni n \mapsto f_{-,n} \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ takich, że

$$\int_\Omega |f_+ - f_{+,n}|^p \mathrm{d}\mu \to 0 \quad \text{i} \quad \int_\Omega |f_- - f_{-,n}|^p \mathrm{d}\mu \to 0, \quad \text{gdy} \ \ n \to +\infty. \tag{2.21}$$

Każda z funkcji $f_n := f_{+,n} - f_{-,n}$, $n \in \mathbb{N}$, jest $\mathcal{B}$-mierzalną funkcją rzeczywistą prostą. Ponadto $f_n \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ dla $n \in \mathbb{N}$ i z (2.21) dostajemy

$$\begin{aligned}\int_\Omega |f - f_n|^p \mathrm{d}\mu &= \|f - f_n\|_p^p = \|(f_+ - f_-) - (f_{+,n} - f_{-,n})\|_p^p \\ &= \|(f_+ - f_{+,n}) + (f_{-,n} - f_-)\|_p^p \leqslant \big(\|f_+ - f_{+,n}\|_p + \|f_- - f_{-,n}\|_p\big)^p \\ &= \Big(\Big(\int_\Omega |f_+ - f_{+,n}|^p \mathrm{d}\mu\Big)^{\frac{1}{p}} + \Big(\int_\Omega |f_- - f_{-,n}|^p \mathrm{d}\mu\Big)^{\frac{1}{p}}\Big)^p \to 0, \quad \text{gdy} \ \ n \to +\infty,\end{aligned}$$

co kończy dowód. □

Twierdzenie 2.6. *Dla wszystkich $p \in [1; +\infty)$, $f \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ i $\varepsilon > 0$ istnieją $n \in \mathbb{N}$ oraz ciągi $\mathbb{Z}_{1,n} \ni k \mapsto A_k \in \mathcal{B}$, $\mathbb{Z}_{1,n} \ni k \mapsto a_k \in \mathbb{C}$ takie, że*

$$\bigcup_{k=1}^n A_k = \Omega \quad i \quad k \neq l \ \Rightarrow \ A_k \cap A_l = \emptyset, \quad k, l \in \mathbb{Z}_{1,n} \tag{2.22}$$

oraz

$$\int_\Omega \Big| f - \sum_{k=1}^n a_k \, \mathbb{1}_{A_k} \Big|^p \mathrm{d}\mu < \varepsilon. \tag{2.23}$$

Dowód. Ustalmy dowolnie $p \in [1;+\infty)$, funkcję $f \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ i $\varepsilon > 0$. Ponieważ $|\operatorname{Re} f| \leqslant |f|$ i $|\operatorname{Im} f| \leqslant |f|$, więc

$$\int_\Omega |\operatorname{Re} f|^p \mathrm{d}\mu \leqslant \int_\Omega |f|^p \mathrm{d}\mu < +\infty \quad \text{i} \quad \int_\Omega |\operatorname{Im} f|^p \mathrm{d}\mu \leqslant \int_\Omega |f|^p \mathrm{d}\mu < +\infty.$$

Stąd na mocy lematu 2.5 istnieją $n', n'' \in \mathbb{N}$ oraz ciągi $\mathbb{Z}_{1,n'} \ni k \mapsto A'_k \in \mathcal{B}$, $\mathbb{Z}_{1,n'} \ni k \mapsto a'_k \in \mathbb{R}$, $\mathbb{Z}_{1,n''} \ni k \mapsto A''_k \in \mathcal{B}$, $\mathbb{Z}_{1,n''} \ni k \mapsto a''_k \in \mathbb{R}$ takie, że

$$\begin{aligned} &\bigcup_{k=1}^{n'} A'_k = \Omega \quad \text{i} \quad k \neq l \;\Rightarrow\; A'_k \cap A'_l = \emptyset, \quad k, l \in \mathbb{Z}_{1,n'} \\ &\bigcup_{k=1}^{n''} A''_k = \Omega \quad \text{i} \quad k \neq l \;\Rightarrow\; A''_k \cap A''_l = \emptyset, \quad k, l \in \mathbb{Z}_{1,n''} \end{aligned} \tag{2.24}$$

oraz

$$\int_\Omega \Big| \operatorname{Re} f - \sum_{k=1}^{n'} a'_k \cdot \mathbb{1}_{A'_k} \Big|^p \mathrm{d}\mu < \frac{\varepsilon}{2^p} \quad \text{i} \quad \int_\Omega \Big| \operatorname{Im} f - \sum_{k=1}^{n''} a''_k \cdot \mathbb{1}_{A''_k} \Big|^p \mathrm{d}\mu < \frac{\varepsilon}{2^p}. \tag{2.25}$$

Z (2.24) wynika, że

$$(A'_{k'} \cap A''_{k''}) \cap (A'_{l'} \cap A''_{l''}) = (A'_{k'} \cap A'_{l'}) \cap (A''_{k''} \cap A''_{l''}) = \emptyset,$$

o ile $(k', k'') \neq (l', l'')$ dla $k', l' \in \mathbb{Z}_{1,n'}$ i $k'', l'' \in \mathbb{Z}_{1,n''}$ oraz

$$\bigcup_{k=1}^{n'} \bigcup_{l=1}^{n''} A'_k \cap A''_l = \bigcup_{k=1}^{n'} \left(A'_k \cap \bigcup_{l=1}^{n''} A''_l \right) = \bigcup_{k=1}^{n'} A'_k \cap \Omega = \bigcup_{k=1}^{n'} A'_k = \Omega.$$

Ponieważ funkcja $\mathbb{Z}_{1,n'} \times \mathbb{Z}_{1,n''} \ni (k,l) \mapsto \sigma(k,l) := k + n'(l-1)$ jest bijekcją zbioru $\mathbb{Z}_{1,n'} \times \mathbb{Z}_{1,n''}$ na zbiór $\mathbb{Z}_{1,n}$ gdzie $n := n' \cdot n''$, więc przyjmując $A_{\sigma(k,l)} := A'_k \cap A''_l$ dla $k \in \mathbb{Z}_{1,n'}$ i $l \in \mathbb{Z}_{1,n''}$ widzimy, że ciąg $\mathbb{Z}_{1,n} \ni k \mapsto A_k \in \mathcal{B}$ spełnia warunek (2.22). Funkcja $[0;+\infty) \ni t \mapsto t^p$ jest wypukła, gdyż $p \geqslant 1$. Dlatego

$$(x+y)^p \leqslant 2^{p-1}(x^p + y^p), \qquad x, y \in [0;+\infty).$$

Stąd dla każdego $h \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$,

$$\begin{aligned} \int_\Omega |h|^p \mathrm{d}\mu &= \int_\Omega |\operatorname{Re} h + \mathrm{i} \operatorname{Im} h|^p \mathrm{d}\mu \leqslant \int_\Omega \Big(|\operatorname{Re} h| + |\operatorname{Im} h| \Big)^p \mathrm{d}\mu \\ &\leqslant 2^{p-1} \left(\int_\Omega \Big(|\operatorname{Re} h|^p + |\operatorname{Im} h|^p \Big) \mathrm{d}\mu \right) = 2^{p-1} \left(\int_\Omega |\operatorname{Re} h|^p \mathrm{d}\mu + \int_\Omega |\operatorname{Im} h|^p \mathrm{d}\mu \right). \end{aligned} \tag{2.26}$$

Przyjmując $a_{\sigma(k,l)} := a'_k + \mathrm{i} a''_l$ dla $k \in \mathbb{Z}_{1,n'}$ i $l \in \mathbb{Z}_{1,n''}$ dostajemy na mocy (2.26)

i (2.25),

$$\begin{aligned}
\int_\Omega \Big| f - \sum_{k=1}^{n} a_k \cdot \mathbb{1}_{A_k} \Big|^p \mathrm{d}\mu &= \int_\Omega \Big| f - \sum_{k=1}^{n'} \sum_{l=1}^{n''} (a'_k + \mathrm{i}a''_l) \cdot \mathbb{1}_{A'_k \cap A''_l} \Big|^p \mathrm{d}\mu \\
&\leqslant 2^{p-1} \Big(\int_\Omega \Big| \operatorname{Re} f - \sum_{k=1}^{n'} \sum_{l=1}^{n''} a'_k \cdot \mathbb{1}_{A'_k \cap A''_l} \Big|^p \mathrm{d}\mu + \int_\Omega \Big| \operatorname{Im} f - \sum_{k=1}^{n'} \sum_{l=1}^{n''} a''_l \cdot \mathbb{1}_{A'_k \cap A''_l} \Big|^p \mathrm{d}\mu \Big) \\
&= 2^{p-1} \Big(\int_\Omega \Big| \operatorname{Re} f - \sum_{k=1}^{n'} \Big(a'_k \sum_{l=1}^{n''} \mathbb{1}_{A'_k \cap A''_l} \Big) \Big|^p \mathrm{d}\mu + \int_\Omega \Big| \operatorname{Im} f - \sum_{l=1}^{n''} \Big(a''_l \sum_{k=1}^{n'} \mathbb{1}_{A'_k \cap A''_l} \Big) \Big|^p \mathrm{d}\mu \Big) \\
&= 2^{p-1} \Big(\int_\Omega \Big| \operatorname{Re} f - \sum_{k=1}^{n'} a'_k \cdot \mathbb{1}_{A'_k \cap \Omega} \Big|^p \mathrm{d}\mu + \int_\Omega \Big| \operatorname{Im} f - \sum_{l=1}^{n''} a''_l \cdot \mathbb{1}_{\Omega \cap A''_l} \Big|^p \mathrm{d}\mu \Big) \\
&= 2^{p-1} \Big(\int_\Omega \Big| \operatorname{Re} f - \sum_{k=1}^{n'} a'_k \cdot \mathbb{1}_{A'_k} \Big|^p \mathrm{d}\mu + \int_\Omega \Big| \operatorname{Im} f - \sum_{l=1}^{n''} a''_l \cdot \mathbb{1}_{A''_l} \Big|^p \mathrm{d}\mu \Big) < \varepsilon.
\end{aligned}$$

Zatem ciągi $\mathbb{Z}_{1,n} \ni k \mapsto A_k$ i $\mathbb{Z}_{1,n} \ni k \mapsto a_k$ spełniają warunek (2.23), co kończy dowód. □

Wniosek 2.7. *Dla dowolnych $p \in [1; +\infty)$ i $f \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ istnieje ciąg $\mathbb{N} \ni n \mapsto f_n \colon \Omega \to \mathbb{C}$ funkcji prostych $\mathcal{B}$-mierzalnych takich, że $\|f_n - f\|_p \to 0$, gdy $n \to +\infty$.*

Dowód. Ustalmy dowolnie $p \in [1; +\infty)$ i $f \in \mathrm{L}^p(\Omega, \mathcal{B}, \mu)$. Z twierdzenia 2.6 wynika, że dla każdego $n \in \mathbb{N}$ istnieje funkcja prosta $\mathcal{B}$-mierzalna $f_n \colon \Omega \to \mathbb{C}$ o własności

$$\|f - f_n\|_p^p = \int_\Omega |f - f_n|^p \mathrm{d}\mu < \frac{1}{n}.$$

Zatem ciąg $\mathbb{N} \ni n \mapsto f_n$ spełnia warunek

$$\|f - f_n\|_p < \frac{1}{n^{1/p}} \to 0, \quad \text{gdy} \quad n \to +\infty,$$

co daje tezę wniosku. □

Rozdział 3

Probabilistyczne struktury regresji

Zgodnie z ogólną koncepcją struktury regresji przedstawionej w pracy [25, Definition 2.1 i Definition 7.1], wprowadzimy szczególny typ struktur regresji oparty na przestrzeniach probabilistycznych.

Definicja 3.1. *Probabilistyczną strukturą regresji* nad przestrzenią probabilistyczną $\mathcal{P} := (\Omega, \mathcal{A}, P)$ nazywamy każdą strukturę regresji $\mathfrak{P} := (A, B, \delta; x, y)$ spełniającą następujące warunki:

IV.1. A jest zbiorem niepustym oraz zachodzi jedna z równości $B = \mathbb{R}$ lub $B = \mathbb{C}$;

IV.2. $x\colon \Omega \to A$ i $y\colon \Omega \to B$;

IV.3. funkcja $\delta\colon (\Omega \to B) \times (\Omega \to B) \to \overline{\mathbb{R}}$ spełnia równość

$$\delta(u, v) = \int_\Omega |u(\omega) - v(\omega)|^2 \,\mathrm{d}P(\omega), \tag{3.1}$$

o ile funkcje $u, v\colon \Omega \to B$ są $\mathcal{A}$-mierzalne, zaś $\delta(u, v) = +\infty$ w przeciwnym przypadku.

Przy tych warunkach rozważmy niepusty zbiór $\mathcal{F} \subset (A \to B)$, który nazywamy *modelem funkcyjnym* probabilistycznej struktury regresji $\mathfrak{P}$. Problem regresji w klasie $\mathcal{F}$ względem probabilistycznej struktury regresji $\mathfrak{P}$ oznacza problem ekstremalny wyznaczenia wszystkich funkcji $f_0 \in \mathcal{F}$ minimalizujących funkcjonał F, spełniający z uwagi na (1.2) i (3.1) następującą równość

$$F(f) = \delta(f \circ x, y) = \int_\Omega |f \circ x(\omega) - y(\omega)|^2 \,\mathrm{d}P(\omega), \qquad f \in \mathcal{F}, \tag{3.2}$$

to znaczy wszystkich funkcji $f_0 \in \mathcal{F}$ spełniających warunek

$$F(f) \geqslant F(f_0), \qquad f \in \mathcal{F}. \tag{3.3}$$

Klasę wszystkich funkcji $f_0 \in \mathcal{F}$ spełniających warunek (3.3) będziemy oznaczać przez $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$.

Definicja 3.2. *Funkcją regresji w klasie $\mathcal{F}$ względem probabilistycznej struktury regresji* $\mathfrak{P}$ nazywamy każdą funkcję z klasy $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$.

Niech $\mathfrak{P} := (A, B, \delta; x, y)$ będzie probabilistyczną strukturą regresji nad przestrzenią probabilistyczną $\mathcal{P} := (\Omega, \mathcal{A}, P)$. Wtedy dla zadanej funkcji $g\colon B \to B$,

$$\mathfrak{P}_g := (A, B, \delta; x, g \circ y) \tag{3.4}$$

jest także probabilistyczną strukturą regresji nad przestrzenią probabilistyczną $\mathcal{P}$. Funkcja g jest interpretowana jako *funkcja skalująca* dane reprezentowane przez funkcję y.

Definicja 3.3. Probabilistyczną strukturę regresji $\mathfrak{P}$ nazywamy *zespoloną*, gdy $B = \mathbb{C}$, zaś *rzeczywistą*, gdy $B = \mathbb{R}$.

W dalszym ciągu rozważania ograniczymy do przypadku, gdy model funkcyjny $\mathcal{F} \subset (A \to B)$ struktury $\mathfrak{P}$ jest rozmaitością liniową ze względu na dodawanie funkcji i mnożenie funkcji przez stałą, tzn.

$$f + g, \quad \lambda \cdot f \in \mathcal{F} \qquad f, g \in \mathcal{F}, \quad \lambda \in B.$$

W szczególności będziemy rozważać modele skończenie wymiarowe, tzn. $\mathcal{F} = \operatorname{lin}(\{h_k\colon k \in \mathbb{Z}_{1,p}\})$ dla pewnych $p \in \mathbb{N}$ i ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k\colon A \to B$. Symbol $\operatorname{lin}(S)$ oznacza otoczkę liniową niepustego zbioru $S \subset (A \to B)$, czyli najmniejszą w sensie inkluzji rozmaitość liniową w przestrzeni $(A \to B; +, \cdot)$ zawierającą zbiór S. W przypadku, gdy $S = \{h_k\colon k \in \mathbb{Z}_{1,p}\}$, mamy

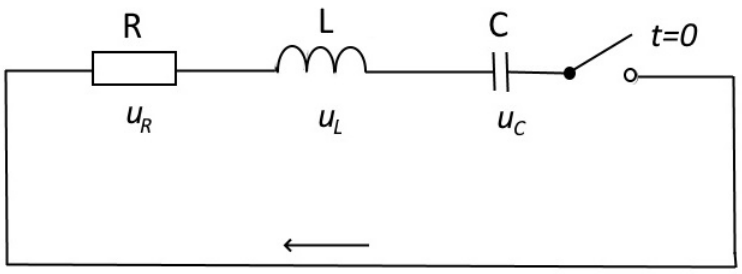
$$\operatorname{lin}(S) = \left\{ \sum_{k=1}^{n} \lambda_k h_k\colon \mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in B \right\}.$$

Przypadek rzeczywistych probabilistycznych struktur regresji będzie przedmiotem rozważań w podrozdziale 3.4. Do tego momentu będziemy stale zakładać, że $\mathfrak{P}$ jest zespoloną probabilistyczną strukturą regresji, czyli że $B = \mathbb{C}$.

Przykład 3.4. Przykład ten dotyczy zespolonej probabilistycznej struktury regresji związanej z przepływem prądu w obwodzie elektrycznym przedstawionym na rysunku 3.1.

Rozważmy obwód elektryczny, w którym znajdują się: naładowany do napięcia początkowego u_0 (woltów) kondensator o pojemności C (faradów), cewka o indukcyjności L (henrów) oraz rezystor o oporności R (omów). Zależność napięcia $t \mapsto u_C(t)$ na okładkach kondensatora od czasu $t > 0$ opisuje równanie różniczkowe dla obwodu RLC,

$$\frac{\mathrm{d}^2 u_C}{\mathrm{d}t^2} + \frac{R}{L}\frac{\mathrm{d}u_C}{\mathrm{d}t} + \frac{1}{LC}u_C = 0; \tag{3.5}$$



Rysunek 3.1: Przepływ prądu w obwodzie elektrycznym RLC.

por. [24, I§9]. Ogólne rozwiązanie równania różniczkowego (3.5) można wyrazić jako kombinację liniową $u_C(t) = \lambda_1 h_1(t) + \lambda_2 h_2(t)$ dla $t \in \mathbb{R}$, gdzie $\lambda_1, \lambda_2 \in \mathbb{C}$ są dowolnie zadane, zaś $h_1, h_2 \colon \mathbb{R} \to \mathbb{C}$ są bazowymi rozwiązaniami równania różniczkowego (3.5) wyznaczonymi przez rozwiązania r_1 oraz r_2 równania charakterystycznego $r^2 + \frac{R}{L}r + \frac{1}{LC} = 0$ zmiennej $r \in \mathbb{C}$ w następujący sposób. Jeśli $\Delta := \frac{R^2}{L^2} - \frac{4}{LC} \neq 0$ to $h_1(t) = \mathrm{e}^{r_1 t}$ i $h_2(t) = \mathrm{e}^{r_2 t}$ dla $t \in \mathbb{R}$, jeśli zaś $\Delta = 0$ to $r_1 = r_2$ oraz $h_1(t) = \mathrm{e}^{r_1 t}$ i $h_2(t) = t\mathrm{e}^{r_1 t}$ dla $t \in \mathbb{R}$.

Rozważmy model funkcyjny $\mathcal{F} := \mathrm{lin}(\{h_1, h_2\}) \subset (A \to \mathbb{C})$, gdzie $A := \mathbb{R}$. Na podstawie ciągu pomiaru czasu $\mathbb{Z}_{1,n} \ni k \mapsto x_k$ oraz ciągu pomiarów napięcia $\mathbb{Z}_{1,n} \ni k \mapsto y_k$, $n \in \mathbb{N}$, określamy probabilistyczną strukturę regresji $\mathfrak{P} := (A, \mathbb{C}, \delta; x, y)$ nad przestrzenią probabilistyczną $\mathcal{P} := (\Omega, \mathcal{A}, P)$ w następujący sposób. Dla dowolnie zadanego $n \in \mathbb{N}$ rozważmy zbiór $\Omega := \mathbb{Z}_{1,n} \times \mathbb{R}$ oraz funkcje próbkujące $x \colon \Omega \to [0; +\infty)$ i $y \colon \Omega \to \mathbb{R}$, gdzie x oznacza czas pomiaru, zaś y oznacza pomiar napięcia w chwili x na okładkach kondensatora. W tym modelu zakładamy, że pomiary czasu są dokładne oraz $x(t_1, t_2) = x_k$ i $y(t_1, t_2) = t_2$, gdy $t_1 = k$ i $t_2 \in \mathbb{R}$, dla $k \in \mathbb{Z}_{1,n}$. Dokonując pomiaru spotykamy się z zagadnieniem niepewności pomiaru. Dotyczy ono pewnego rozrzutu wartości, które mierzona wielkość może przyjmować. Z reguły zakłada się, że do obliczenia podanej niepewności w k-tym pomiarze zastosowano rozkład normalny z parametrami $\mu := y_k$ dla $k \in \mathbb{Z}_{1,n}$ oraz σ zależnym od urządzenia pomiarowego; por. [7, §4.2.1, §4.3.4]. Zakładając, że dokonujemy niezależnych pomiarów napięcia przy użyciu tego samego przyrządu przyjmujemy $\mathcal{A} := \mathcal{B}' \otimes \mathcal{B}''$, gdzie $\mathcal{B}'$ oznacza rodzinę wszystkich podzbiorów zbioru $\mathbb{Z}_{1,n}$, zaś $\mathcal{B}''$ oznacza rodzinę wszystkich podzbiorów borelowskich zbioru $\mathbb{R}$. Wtedy istnieje dokładnie jedna miara $P \colon \mathcal{A} \to \mathbb{R}$ spełniająca warunek

$$P\big(\{k\} \times I\big) = \frac{1}{n} \frac{1}{\sigma\sqrt{2\pi}} \int_I e^{\frac{-(\tau - y_k)^2}{2\sigma^2}} \mathrm{d}\tau, \qquad k \in \mathbb{Z}_{1,n},\ I \in \mathcal{B}''. \tag{3.6}$$

Zatem $\mathfrak{P}$ jest zespoloną probabilistyczną strukturą regresji, gdzie kryterium δ jest funkcją spełniającą warunek (3.1). Klasa funkcji regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$ będzie wyzna-

czona w podrozdziale 4.4, przykład 4.14.

Dla zadanej probabilistycznej struktury regresji $\mathfrak{P}$ definiujemy

$$\mathcal{A}_x := \Big\{V \in 2^A \colon x^{-1}(V) \in \mathcal{A}\Big\} \tag{3.7}$$

oraz

$$\mathcal{A}_x \ni V \mapsto P_x(V) := P\Big(x^{-1}(V)\Big). \tag{3.8}$$

Rodzina $\mathcal{A}_x$ jest σ-ciałem na A i P_x jest miarą probabilistyczną na σ-ciele $\mathcal{A}_x$.

Uwaga 3.5. Ze wzorów (3.7) i (3.8) wynika, że dla każdej przestrzeni probabilistycznej $(\Omega, \mathcal{A}, P)$ i każdej funkcji $x\colon \Omega \to A$, struktura $(A, \mathcal{A}_x, P_x)$ jest także przestrzenią probabilistyczną. Dodatkowo z lematu 2.3 (z $f := u,\ \Omega' := A,\ \mathcal{B}' := \mathcal{A}_x,\ \mu' := P_x,$ $\mathcal{B} := \mathcal{A},\ \mu := P$) wynika, że dla każdej funkcji $\mathcal{A}_x$-mierzalnej $u\colon A \to \mathbb{C}$,

$$u \in \mathrm{L}^1(A, \mathcal{A}_x, P_x) \ \Leftrightarrow\ u \circ x \in \mathrm{L}^1(\Omega, \mathcal{A}, P)$$

oraz

$$\int_\Omega u \circ x(\omega)\mathrm{d}P(\omega) = \int_A u(t)\mathrm{d}P_x(t), \quad u \in \mathrm{L}^1(A, \mathcal{A}_x, P_x). \tag{3.9}$$

Uwaga 3.6. Z nierówności $|u\bar{v}| \leqslant \frac{1}{2}\Big(|u|^2 + |v|^2\Big)$ dla $u, v \in \mathbb{C}$ wynika, że funkcja

$$\mathrm{L}^2(A, \mathcal{A}_x, P_x) \times \mathrm{L}^2(A, \mathcal{A}_x, P_x) \ni (u, v) \mapsto \langle u|v\rangle := \int_A u(t) \cdot \overline{v(t)}\mathrm{d}P_x(t) \tag{3.10}$$

jest dobrze określona. Ze wzoru (3.10) i własności całki wnioskujemy, że dla wszystkich $\lambda_1, \lambda_2 \in \mathbb{C}$ i $u, v, w \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ zachodzą następujące własności (por. [30, Definition 4.1])

$$\begin{aligned} \langle \lambda_1 u + \lambda_2 v|w\rangle &= \lambda_1\langle u|w\rangle + \lambda_2\langle v|w\rangle; \\ \langle u|v\rangle &= \overline{\langle v|u\rangle}; \\ \langle u|u\rangle &\geqslant 0. \end{aligned} \tag{3.11}$$

Definiując

$$\mathrm{L}^2(A, \mathcal{A}_x, P_x) \ni u \mapsto \|u\| := \sqrt{\langle u|u\rangle} = \left(\int_A |u(t)|^2\mathrm{d}P_x(t)\right)^{1/2} \tag{3.12}$$

i korzystając z własności (3.11) stwierdzamy, że (por. [30, §4.3])

$$\|\lambda u\| = |\lambda| \cdot \|u\| \quad \text{i} \quad \|u + v\| \leqslant \|u\| + \|v\| \tag{3.13}$$

oraz

$$\|u\| = 0 \ \Leftrightarrow\ P_x(\{t \in A : u(t) \neq 0\}) = 0$$

dla wszystkich $\lambda \in \mathbb{C}$ i $u, v \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. Dlatego, $\|\cdot\|$ jest pseudo-normą w przestrzeni liniowej $(\mathrm{L}^2(A, \mathcal{A}_x, P_x); +, \cdot)$; por. [5, Chap. III, Definition 1.1].
Z własności (3.11) wynika w standardowy sposób nierówność Schwarz'a

$$|\langle u|v\rangle| \leqslant \|u\| \cdot \|v\|, \quad u, v \in \mathrm{L}^2(A, \mathcal{A}_x, P_x); \tag{3.14}$$

por. [21, §28], [30, §4.2].
Ponieważ przestrzeń $\mathbf{L}^2(A, \mathcal{A}_x, P_x)$ jest przestrzenią zupełną (por. [2, §2.4.11]), więc struktura $\mathrm{H}(\mathfrak{P}) := (\mathrm{L}^2(A, \mathcal{A}_x, P_x); +, \cdot, \langle\cdot|\cdot\rangle)$ jest zespoloną przestrzenią pseudo-Hilberta, tzn., struktura $(\mathrm{L}^2(A, \mathcal{A}_x, P_x); +, \cdot, \|\cdot\|)$ jest zespoloną przestrzenią pseudo-Banacha.

Podobnie jak w przypadku wzorów (3.7) i (3.8) określamy

$$\mathcal{A}_y := \Big\{V \in 2^{\mathbb{C}} \colon y^{-1}(V) \in \mathcal{A}\Big\} \tag{3.15}$$

oraz

$$\mathcal{A}_y \ni V \mapsto P_y(V) := P\Big(y^{-1}(V)\Big). \tag{3.16}$$

Rodzina $\mathcal{A}_y$ jest σ-ciałem na $\mathbb{C}$, zaś P_y jest miarą probabilistyczną na $\mathcal{A}_y$.

Uwaga 3.7. Zauważmy, że dla dowolnie zadanych funkcji $u \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ mamy $|u|^2 \in \mathrm{L}^1(A, \mathcal{A}_x, P_x)$ i $|g|^2 \in \mathrm{L}^1(\mathbb{C}, \mathcal{A}_y, P_y)$. Ponieważ $|u|^2 \circ x = |u \circ x|^2$ i $|g|^2 \circ y = |g \circ y|^2$, więc z (3.7), (3.8), (3.15), (3.16) oraz z uwagi 3.5 wynika, że $u \circ x,\ g \circ y \in \mathrm{L}^2(\Omega, \mathcal{A}, P)$ i

$$\int_\Omega |u \circ x(\omega)|^2 \mathrm{d}P(\omega) = \int_A |u(t)|^2 \mathrm{d}P_x(t), \quad u \in \mathrm{L}^2(A, \mathcal{A}_x, P_x).$$

Stąd

$$M_g := \left(\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega)\right)^{1/2} < +\infty. \tag{3.17}$$

Stosując więc nierówność Schwarz'a dla całki Lebesgue'a otrzymujemy

$$\begin{aligned} \int_\Omega |u \circ x(\omega) \cdot \overline{g \circ y(\omega)}| \mathrm{d}P(\omega) & \\ \leqslant \left(\int_\Omega |u \circ x(\omega)|^2 \mathrm{d}P(\omega)\right)^{1/2} & \cdot \left(\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega)\right)^{1/2} \\ = M_g \left(\int_A |u(t)|^2 \mathrm{d}P_x(t)\right)^{1/2} & = M_g \|u\|. \end{aligned} \tag{3.18}$$

W związku z tym dla każdej funkcji $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ funkcjonał $g^* \colon \mathrm{L}^2(A, \mathcal{A}_x, P_x) \to \mathbb{C}$ jest dobrze określony następującym wzorem

$$\mathrm{L}^2(A, \mathcal{A}_x, P_x) \ni u \mapsto g^*(u) := \int_\Omega u \circ x(\omega) \overline{g \circ y(\omega)} \mathrm{d}P(\omega). \tag{3.19}$$

Ponadto z (3.18) otrzymujemy

$$|g^*(u)| \leqslant M_g\|u\|, \quad u \in \mathrm{L}^2(A, \mathcal{A}_x, P_x). \tag{3.20}$$

W konsekwencji funkcjonał g^* jest liniowy a także ograniczony w przestrzeni $(\mathrm{L}^2(A, \mathcal{A}_x, P_x); +, \cdot, \|\cdot\|)$ dla każdego $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$.

3.1 Podstawowe własności funkcji regresji

Od tej pory będzie rozważany problem regresji dla modelu funkcyjnego $\mathcal{F}$ w odniesieniu do zespolonej probabilistycznej struktury regresji $\mathfrak{P}_g$ danej wzorem (3.4), gdzie $g\colon \mathbb{C} \to \mathbb{C}$ jest stosownie zadaną funkcją skalującą oraz $\mathcal{F}$ jest liniowym modelem funkcyjnym ze standardowymi operacjami dodawania funkcji i mnożenia funkcji przez stałą.

Następujące twierdzenie jest odpowiednikiem [25, Lemma 3.1]. Jest to podstawowe twierdzenie charakteryzujące klasę funkcji regresji. Za jego pomocą w dalszym ciągu zostaną pokazane podstawowe własności funkcji regresji.

Twierdzenie 3.8. *Jeśli $\mathcal{F} \neq \emptyset$ jest rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, to dla każdego $f \in \mathcal{F}$ zachodzi następujący warunek*

$$f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \;\Leftrightarrow\; \langle h|f\rangle = g^*(h), \quad h \in \mathcal{F}.$$

Dowód. Dla zadanej funkcji $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ definiujemy funkcjonał

$$(A \to \mathbb{C}) \ni f \mapsto F_g(f) := \delta(f \circ x, g \circ y).$$

Z własności IV.3 definicji 3.1 wynika

$$F_g(f) = \int_\Omega |f \circ x(\omega) - g \circ y(\omega)|^2 \,\mathrm{d}P(\omega), \quad f \in \mathcal{F}. \tag{3.21}$$

Ustalmy dowolnie $f, h \in \mathcal{F}$ i $\lambda \in \mathbb{C}$. Korzystając z (3.21) oraz z własności całki Lebesgue'a otrzymujemy

$$\begin{aligned}
F_g(f + \lambda h) &= \int_\Omega \Big|(f + \lambda h) \circ x(\omega) - g \circ y(\omega)\Big|^2 \mathrm{d}P(\omega) \\
&= \int_\Omega \Big|f \circ x(\omega) + \lambda h \circ x(\omega) - g \circ y(\omega)\Big|^2 \mathrm{d}P(\omega) \\
&= \int_\Omega \Big(\Big|f \circ x(\omega) - g \circ y(\omega)\Big|^2 + 2\,\mathrm{Re}\Big[(f \circ x(\omega) - g \circ y(\omega))\overline{\lambda h \circ x(\omega)}\Big] \\
&\qquad + |\lambda|^2 |h \circ x(\omega)|^2\Big) \mathrm{d}P(\omega) \\
&= \int_\Omega \Big|f \circ x(\omega) - g \circ y(\omega)\Big|^2 \mathrm{d}P(\omega) \\
&\qquad + 2\int_\Omega \mathrm{Re}\Big[(f \circ x(\omega) - g \circ y(\omega))\overline{\lambda h \circ x(\omega)}\Big]\, \mathrm{d}P(\omega) \\
&\qquad + |\lambda|^2 \int_\Omega |h \circ x(\omega)|^2 \mathrm{d}P(\omega).
\end{aligned}$$

Stąd na mocy (3.21), (3.12) i (3.9), dostajemy

$$F_g(f+\lambda h) = F_g(f) + |\lambda|^2\|h\|^2 + 2\,\mathrm{Re}\int_\Omega f\circ x(\omega)\overline{\lambda h\circ x(\omega)}\mathrm{d}P(\omega) \\ - 2\,\mathrm{Re}\int_\Omega g\circ y(\omega)\overline{\lambda h\circ x(\omega)}\mathrm{d}P(\omega).$$

To łącznie z (3.10), (3.19) i (3.9) implikuje

$$F_g(f+\lambda h) = F_g(f) + |\lambda|^2\|h\|^2 + 2\,\mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right].$$

Dlatego dla wszystkich $\lambda\in\mathbb{C}$ i $f,h\in\mathcal{F}$, mamy

$$F_g(f+\lambda h) - F_g(f) = 2\,\mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right] + |\lambda|^2\|h\|^2. \tag{3.22}$$

Ustalmy dowolnie $f\in\mathcal{F}$ spełniający równość $\langle h|f\rangle = g^*(h),\ h\in\mathcal{F}$. Przyjmując $\lambda := 1$ w (3.22) otrzymujemy

$$F_g(f+h) - F_g(f) = \|h\|^2 \geqslant 0$$

w związku z czym

$$F_g(f+h) \geqslant F_g(f), \qquad h\in\mathcal{F},$$

co oznacza, że $f\in\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$. Na odwrót załóżmy, że $f\in\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$. Wtedy z (3.22) otrzymujemy

$$2\,\mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right] + |\lambda|^2\|h\|^2 \geqslant 0, \quad h\in\mathcal{F},\ \lambda\in\mathbb{C}. \tag{3.23}$$

Zastępując h przez $(-h)$ w (3.23) dostajemy

$$-2\,\mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right] + |\lambda|^2\|h\|^2 \geqslant 0, \quad h\in\mathcal{F},\ \lambda\in\mathbb{C}. \tag{3.24}$$

Własności (3.23) i (3.24) dają łącznie

$$-\frac{1}{2}|\lambda|^2\|h\|^2 \leqslant \mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right] \leqslant \frac{1}{2}|\lambda|^2\|h\|^2, \quad h\in\mathcal{F},\ \lambda\in\mathbb{C}.$$

Ustalając dowolnie $h\in\mathcal{F},\ \alpha\in\mathbb{R}$ i zakładając, że $\lambda = |\lambda|\mathrm{e}^{\mathrm{i}\alpha}\neq 0$ otrzymujemy

$$-\frac{1}{2}|\lambda|\|h\|^2 \leqslant \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}\alpha}\Big(\langle h|f\rangle - g^*(h)\Big)\right] \leqslant \frac{1}{2}|\lambda|\|h\|^2.$$

Przechodząc teraz do granicy z $|\lambda|\to 0$ wnioskujemy, że

$$\mathrm{Re}\left[\mathrm{e}^{\mathrm{i}\alpha}\Big(\langle h|f\rangle - g^*(h)\Big)\right] = 0, \quad h\in\mathcal{F},\ \alpha\in\mathbb{R}.$$

W szczególności powyższa równość zachodzi dla $\alpha\in\{0,\frac{\pi}{2}\}$. Dlatego $\langle h|f\rangle - g^*(h) = 0$ dla $h\in\mathcal{F}$, co kończy dowód. □

Z własności pseudo-normy (3.13) wynika, że zbiór

$$\Theta := \{h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)\colon \|h\| = 0\}$$

jest liniowy. Nazywamy go *zbiorem zerowym* przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$. Zbiór Θ jest w istocie kulą domkniętą o promieniu 0 i środku w funkcji zerowej $A \ni t \mapsto \theta(t) := 0$.

Standardowe operacje dodawania funkcji i mnożenia funkcji przez stałą można rozszerzyć na dowolne zbiory $F_1, F_2 \subset (A \to \mathbb{C})$ jak następuje:

$$\begin{aligned} F_1 + F_2 :=& \{f_1 + f_2\colon f_1 \in F_1, f_2 \in F_2\}, \\ \lambda \cdot F_1 :=& \{\lambda f_1\colon f_1 \in F_1\}, \quad \lambda \in \mathbb{C}, \\ f + F_1 :=& \{f\} + F_1 \quad \text{i} \quad F_1 + f := F_1 + \{f\}, \quad f \in (A \to \mathbb{C}). \end{aligned}$$

Wniosek 3.9. *Jeśli $\mathcal{F} \neq \emptyset$ jest rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, to*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g). \tag{3.25}$$

Jeśli dodatkowo $\mathcal{F} \subset \Theta$ to $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$.

Dowód. Ustalmy dowolnie $f, h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. Jeśli $\|h\| = 0$ to z nierówności Schwarz'a (3.14) i (3.20) wynika, że

$$|\langle h|f\rangle| \leqslant \|h\|\|f\| = 0 \quad \text{i} \quad |g^*(h)| \leqslant \left(\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega)\right)^{1/2} \|h\| = 0.$$

Stąd

$$\langle h|f\rangle = 0 = g^*(h), \quad f \in \mathrm{L}^2(A, \mathcal{A}_x, P_x),\ h \in \Theta. \tag{3.26}$$

Załóżmy, że $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ i $h \in \Theta + \mathcal{F}$. Wtedy $h = h_0 + h_1$ dla pewnych $h_0 \in \Theta$ i $h_1 \in \mathcal{F}$. Stosując teraz (3.26) i twierdzenie 3.8 widzimy, że

$$\langle h|f\rangle = \langle h_0|f\rangle + \langle h_1|f\rangle = 0 + g^*(h_1) = g^*(h_0) + g^*(h_1) = g^*(h), \quad h \in \Theta + \mathcal{F}.$$

Ponieważ $f \in \mathcal{F} \subset \Theta + \mathcal{F}$, więc z twierdzenia 3.8 wynika, że $f \in \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g)$. To dowodzi inkluzji

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g). \tag{3.27}$$

Na odwrót załóżmy, że $f \in \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g)$ i $h \in \mathcal{F}$. Wtedy $h \in \Theta + \mathcal{F}$, skąd na mocy twierdzenia 3.8,

$$\langle h|f\rangle = g^*(h), \quad h \in \mathcal{F}.$$

Stosując ponownie twierdzenie 3.8 stwierdzamy, że $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, i w konsekwencji

$$\mathcal{F} \cap \mathrm{Reg}(\Theta \cap \mathcal{F}, \mathfrak{P}_g) \subset \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g).$$

Inkluzja ta łącznie z inkluzją (3.27) daje równość (3.25).
Ponieważ $\Theta \subset \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, więc równości w (3.26) zachodzą dla wszystkich $f, h \in \Theta$. Z twierdzenia 3.8 wynika inkluzja $\mathrm{Reg}(\Theta, \mathfrak{P}_g) \supset \Theta$. Inkluzja odwrotna jest bezpośrednią konsekwencją określenia klasy $\mathrm{Reg}(\Theta, \mathfrak{P}_g)$. Dlatego $\mathrm{Reg}(\Theta, \mathfrak{P}_g) = \Theta$. Jeśli teraz $\mathcal{F} \subset \Theta$ to $\Theta + \mathcal{F} = \Theta$ i tym samym równość (3.25) przyjmuje postać $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$, co kończy dowód. □

Przez $S^\perp$ oznaczmy dopełnienie ortogonalne zbioru niepustego $S \subset \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$, tzn.,

$$S^\perp := \{f \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \colon \langle h|f\rangle = 0 \ \text{ dla } \ h \in S\}.$$

Twierdzenie 3.10. *Jeśli $\mathcal{F} \neq \emptyset$ jest domkniętą rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, to $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$ i $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + f$ dla każdego $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Ponadto, jeśli $\mathcal{F} \subset S := (g^*)^{-1}(0)$, to $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta$. W przeciwnym wypadku $(\mathcal{F} \cap S)^\perp \cap \mathcal{F} \setminus \Theta \neq \emptyset$ i*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + \frac{\overline{g^*(h)}}{\|h\|^2} h, \quad h \in (\mathcal{F} \cap S)^\perp \cap \mathcal{F} \setminus \Theta. \tag{3.28}$$

Dowód. Załóżmy, że $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$ i ustalmy dowolnie $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Z twierdzenia 3.8 wynika, że

$$\langle h|f\rangle = g^*(h), \quad h \in \mathcal{F}. \tag{3.29}$$

Stąd w szczególności dla każdego $f' \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, $\langle h|f'\rangle = g^*(h)$ dla $h \in \mathcal{F}$, a więc przyjmując $h := f' - f$ wnioskujemy z (3.29), że

$$\|h\|^2 = \langle h|f' - f\rangle = \langle h|f'\rangle - \langle h|f\rangle = g^*(h) - g^*(h) = 0.$$

Dlatego $f' \in \Theta + f$ dla $f' \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, więc $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset \Theta + f$. Na odwrót, ustalmy dowolnie $f' \in \Theta + f$. Wtedy, z nierówności Schwarz'a (3.14) otrzymujemy dla każdego $h \in \mathcal{F}$,

$$|\langle h|f'\rangle - \langle h|f\rangle| = |\langle h|f' - f\rangle| \leqslant \|h\| \cdot \|f' - f\| = 0.$$

Stąd i z (3.29) mamy $\langle h|f'\rangle = \langle h|f\rangle = g^*(h)$ dla $h \in \mathcal{F}$. Ponieważ $\mathcal{F}$ jest domkniętą rozmaitością liniową w przestrzeni $\mathrm{H}(\mathfrak{P})$, więc $\Theta \subset \mathcal{F}$. Dlatego $\Theta + f \subset \mathcal{F}$. Stosując twierdzenie 3.8 mamy $f' \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ dla $f' \in \Theta + f$, a więc $\Theta + f \subset \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$.

Ta inkluzja razem z inkluzją do niej odwrotną prowadzą do równości $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) = \Theta + f$, o ile $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) \neq \emptyset$. Zatem otrzymujemy następującą implikację

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) \neq \emptyset \;\Rightarrow\; \mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) = \Theta + f. \tag{3.30}$$

Załóżmy teraz, że $\mathcal{F} \subset S$. Wtedy

$$\langle h|\theta\rangle = 0 = g^*(h), \quad h \in \mathcal{F},$$

skąd na mocy twierdzenia 3.8, $\theta \in \mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$. To łącznie z (3.30) daje $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) = \Theta + \theta = \Theta$. Pozostaje rozważyć przypadek, gdy inkluzja $\mathcal{F} \subset S$ nie zachodzi. Wtedy $\mathcal{F} \cap S \neq \mathcal{F}$. Na mocy założenia, $\mathcal{F}$ jest domkniętym zbiorem w przestrzeni $\mathrm{H}(\mathfrak{P})$. Ponieważ $g \in \mathrm{L}^2(\mathbb{C},\mathcal{A}_y,P_y)$, więc g^* na podstawie (3.20) jest funkcjonałem ciągłym w $\mathrm{H}(\mathfrak{P})$. Zatem S jest także zbiorem domkniętym w $\mathrm{H}(\mathfrak{P})$. Wobec powyższego $\mathcal{F}\cap S$ jest zbiorem domkniętym w przestrzeni $\mathrm{H}(\mathfrak{P})$. W konsekwencji,

$$\Theta \subset \mathcal{F} \cap S \neq \mathcal{F}, \tag{3.31}$$

więc $\mathcal{F}\backslash(\mathcal{F}\cap S) \neq \emptyset$. Ustalmy dowolnie $h \in \mathcal{F}\backslash(\mathcal{F}\cap S)$. Ponieważ zbiór $\mathcal{F}\cap S$ jest domknięty w $\mathrm{H}(\mathfrak{P})$, więc h posiada rzut ortogonalny h_S na $\mathcal{F}\cap S$, tzn.,

$$h_S \in \mathcal{F}\cap S \quad \text{i} \quad \langle h - h_S|h'\rangle = 0, \quad h' \in \mathcal{F}\cap S. \tag{3.32}$$

Dlatego $h - h_S \in (\mathcal{F}\cap S)^\perp \cap \mathcal{F}$. Jeśli $h - h_S \in \Theta$, wtedy z (3.31) i (3.32) wynika, że $h = h_S + (h - h_S) \in \mathcal{F}\cap S + \Theta = \mathcal{F}\cap S$, co jest niemożliwe. W związku z tym $h - h_S \notin \Theta$, co daje $h - h_S \in (\mathcal{F}\cap S)^\perp \cap \mathcal{F}\backslash\Theta$. Zatem $(\mathcal{F}\cap S)^\perp \cap \mathcal{F}\backslash\Theta \neq \emptyset$. Załóżmy, że $h \in (\mathcal{F}\cap S)^\perp \cap \mathcal{F}\backslash\Theta$. Wtedy $\|h\| \neq 0$ i $g^*(h) \neq 0$, skąd dla każdego $h' \in \mathcal{F}$,

$$h'_S := h' - \frac{g^*(h')}{g^*(h)}h \in \mathcal{F}\cap S \quad \text{i} \quad h' - h'_S = \frac{g^*(h')}{g^*(h)}h \in (\mathcal{F}\cap S)^\perp \cap \mathcal{F}. \tag{3.33}$$

Ponieważ

$$\frac{\overline{g^*(h)}}{\|h\|^2}h \in (\mathcal{F}\cap S)^\perp \cap \mathcal{F},$$

więc z (3.33) wynika, że

$$\begin{aligned}\left\langle h'\middle|\frac{\overline{g^*(h)}}{\|h\|^2}h\right\rangle &= \left\langle h' - h'_S\middle|\frac{\overline{g^*(h)}}{\|h\|^2}h\right\rangle = \left\langle \frac{g^*(h')}{g^*(h)}h\middle|\frac{\overline{g^*(h)}}{\|h\|^2}h\right\rangle \\ &= \frac{g^*(h')}{g^*(h)}\overline{\left(\frac{\overline{g^*(h)}}{\|h\|^2}\right)}\langle h|h\rangle = g^*(h'), \quad h' \in \mathcal{F}.\end{aligned}$$

Stosując twierdzenie 3.8 widzimy, że

$$f := \frac{\overline{g^*(h)}}{\|h\|^2}h \in \mathrm{Reg}(\mathcal{F},\mathfrak{P}_g), \quad h \in (\mathcal{F}\cap S)^\perp \cap \mathcal{F}\backslash\Theta. \tag{3.34}$$

Dlatego $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) \neq \emptyset$. Własności (3.34) i (3.30) dają łącznie równość (3.28) o ile inkluzja $\mathcal{F} \subset S$ nie zachodzi. W obu przypadkach $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) \neq \emptyset$, co kończy dowód. □

3.2 Regresje pierwszego rodzaju jako szczególny przypadek klasy regresji względem struktury probabilistycznej

Związek pomiędzy regresją a warunkową wartością oczekiwaną podany jest w [29, Ch. V, §6, Theorem 4]. Może on posłużyć do zdefiniowania warunkowej wartości oczekiwanej za pomocą funkcji regresji, czyli w wyniku procesu optymalizacji.

Rozważmy następujące uogólnienie regresji pierwszego rodzaju. Niech $\mathcal{A}' \subset \mathcal{A}_x$ będzie dowolnie zadanym σ-ciałem i niech $\mathcal{F}'$ będzie klasą przebiegów teoretycznych złożoną ze wszystkich funkcji $f\colon A \to \mathbb{C}$, które są $\mathcal{A}'$ mierzalne, czyli

$$f^{-1}(V) \in \mathcal{A}', \quad V \in \mathrm{Bor}(\mathbb{C}).$$

Wtedy dla każdego $f \in \mathcal{F}'$

$$(f \circ x)^{-1}(V) = x^{-1}\Big(f^{-1}(V)\Big) \in \mathcal{A}, \quad V \in \mathrm{Bor}(\mathbb{C}),$$

a więc $f \circ x$ jest zmienną losową.

Twierdzenie 3.11. *Dla dowolnych funkcji $f\colon A \to \mathbb{C}$ oraz $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, $f \in \mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$ wtedy i tylko wtedy, gdy $f \in \mathcal{F} := \mathcal{F}' \cap \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ i*

$$\int_U f \circ x(\omega)\mathrm{d}P(\omega) = \int_U g \circ y(\omega)\mathrm{d}P(\omega), \quad U \in \mathcal{A}'', \tag{3.35}$$

gdzie $\mathcal{A}'' := \{x^{-1}(V)\colon V \in \mathcal{A}'\}$.

Dowód. Ustalmy dowolnie $f\colon A \to \mathbb{C}$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$. Załóżmy, że $f \in \mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$. Wtedy

$$\delta(h \circ x, g \circ y) \geqslant \delta(f \circ x, g \circ y), \quad h \in \mathcal{F}'. \tag{3.36}$$

Z określenia klasy $\mathcal{F}'$ wynika, że $h \circ x$ jest funkcją $\mathcal{A}$-mierzalną dla każdego $h \in \mathcal{F}'$, a w szczególności $f \circ x$ jest funkcją $\mathcal{A}$-mierzalną. Ponieważ g jest również funkcją $\mathcal{A}_y$-mierzalną, więc nierówność (3.36) przyjmie postać

$$\int_\Omega |h \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega) \geqslant \int_\Omega |f \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega), \quad h \in \mathcal{F}'.$$

Przyjmując w szczególności $h := 0 \cdot \mathbb{1}_\Omega \in \mathcal{F}'$, otrzymujemy

$$\int_\Omega |f \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega) \leqslant \int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega) < +\infty,$$

gdyż $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$. Stąd

$$\begin{aligned}\int_\Omega |f \circ x(\omega)|^2 \mathrm{d}P(\omega) &\leqslant 2\int_\Omega |f \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega) + 2\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega) \\ &\leqslant 4\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega) < +\infty,\end{aligned}$$

a więc $f\circ x\in \mathrm{L}^2(\Omega,\mathcal{A},P)$. Z uwagi 3.5 wynika więc, że $f\in \mathrm{L}^2(A,\mathcal{A}_x,P_x)$ i tym samym $f\in\mathcal{F}$. Ponieważ $\mathcal{F}\subset\mathcal{F}'$ więc z (3.36) wynika, że $f\in \mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$. Możemy więc zastosować charakteryzację klasy regresji $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$ z twierdzenia 3.8. W rezultacie dostajemy

$$\langle h|f\rangle = g^*(h),\quad h\in\mathcal{F}. \tag{3.37}$$

Zadając dowolnie $U\in\mathcal{A}''$ przyjmijmy $h:=\mathbb{1}_{x(U)}$. Ponieważ $U=x^{-1}(V)$ dla pewnego $V\in\mathcal{A}'$, więc $V=x(U)$, i w konsekwencji $h=\mathbb{1}_V$. Stąd h jest funkcją $\mathcal{A}'$-mierzalną, a więc $h\in\mathcal{F}$. Ponadto dla każdego $\omega\in U$, $h\circ x(\omega)=\mathbb{1}_V(x(\omega))=1$. Dla każdego $\omega\in\Omega\backslash U$ mamy zaś, $\omega\in x^{-1}(A)\backslash x^{-1}(V)=x^{-1}(A\backslash V)$, co daje $x(\omega)\in A\backslash V$, i w konsekwencji $h\circ x(\omega)=\mathbb{1}_V(x(\omega))=0$. Dlatego $h\circ x=\mathbb{1}_U$, i wobec (3.9), (3.10), (3.37) oraz (3.19) dostajemy

$$\begin{aligned}\int_U \overline{f\circ x(\omega)}\mathrm{d}P(\omega)&=\int_\Omega h\circ x(\omega)\overline{f\circ x(\omega)}\mathrm{d}P(\omega)=\int_A h(t)\overline{f(t)}\mathrm{d}P_x(t)=\langle h|f\rangle\\&=g^*(h)=\int_\Omega h\circ x(\omega)\overline{g\circ y(\omega)}\mathrm{d}P(\omega)=\int_U\overline{g\circ y(\omega)}\mathrm{d}P(\omega).\end{aligned}$$

Stąd wynika równość w (3.35), co dowodzi twierdzeniu w stronę ($\Rightarrow$).

Na odwrót, załóżmy, że $f\in\mathcal{F}$ i zachodzi warunek (3.35). Niech K oznacza zbiór wszystkich funkcji prostych $h\colon A\to\mathbb{C}$, które są $\mathcal{A}'$-mierzalne. Dla dowolnie zadanej funkcji $h\in K$ istnieją $n\in\mathbb{N}$ oraz ciągi $\mathbb{Z}_{1,n}\ni k\mapsto\lambda_k\in\mathbb{C}$ i $\mathbb{Z}_{1,n}\ni k\mapsto V_k\in\mathcal{A}'$ takie, że $h=\sum\limits_{k=1}^n\lambda_k\,\mathbb{1}_{V_k}$. Wtedy $U_k:=x^{-1}(V_k)\in\mathcal{A}''\subset\mathcal{A}$ dla $k\in\mathbb{Z}_{1,n}$ oraz

$$h\circ x=\Big(\sum_{k=1}^n\lambda_k\,\mathbb{1}_{V_k}\Big)\circ x=\sum_{k=1}^n\lambda_k\Big(\,\mathbb{1}_{V_k}\circ x\Big)=\sum_{k=1}^n\lambda_k\,\mathbb{1}_{U_k}.$$

Stąd $h\circ x$ jest funkcją prostą $\mathcal{A}$-mierzalną, i tym samym $h\circ x\in\mathrm{L}^2(\Omega,\mathcal{A},P)$. Z warunku (3.35) dostajemy

$$\begin{aligned}\int_\Omega h\circ x(\omega)\overline{f\circ x(\omega)}\mathrm{d}P(\omega)&=\int_\Omega\Big(\sum_{k=1}^n\lambda_k\,\mathbb{1}_{U_k}(\omega)\Big)\overline{f\circ x(\omega)}\mathrm{d}P(\omega)\\&=\sum_{k=1}^n\lambda_k\int_\Omega\mathbb{1}_{U_k}(\omega)\overline{f\circ x(\omega)}\mathrm{d}P(\omega)=\sum_{k=1}^n\lambda_k\int_{U_k}\overline{f\circ x(\omega)}\mathrm{d}P(\omega)\\&=\sum_{k=1}^n\lambda_k\int_{U_k}\overline{g\circ y(\omega)}\mathrm{d}P(\omega)=\sum_{k=1}^n\lambda_k\int_\Omega\mathbb{1}_{U_k}(\omega)\overline{g\circ y(\omega)}\mathrm{d}P(\omega)\\&=\int_\Omega\Big(\sum_{k=1}^n\lambda_k\,\mathbb{1}_{U_k}(\omega)\Big)\overline{g\circ y(\omega)}\mathrm{d}P(\omega)=\int_\Omega h\circ x(\omega)\overline{g\circ y(\omega)}\mathrm{d}P(\omega).\end{aligned}$$

Korzystając teraz z własności (3.10), (3.9) i (3.19) mamy

$$\begin{aligned}\langle h|f\rangle&=\int_A h(t)\overline{f(t)}\mathrm{d}P_x(t)=\int_\Omega h\circ x(\omega)\overline{f\circ x(\omega)}\mathrm{d}P(\omega)\\&=\int_\Omega h\circ x(\omega)\overline{g\circ y(\omega)}\mathrm{d}P(\omega)=g^*(h),\quad h\in K.\end{aligned} \tag{3.38}$$

Z wniosku 2.7, wynika, że dla dowolnie ustalonej funkcji $h \in \mathrm{L}^2(A, \mathcal{A}', P_x)$ istnieje ciąg $\mathbb{N} \ni n \mapsto h_n \in K$ taki, że

$$\int_A |h_n(t) - h(t)|^2 \mathrm{d}P_x(t) \to 0, \quad \text{gdy} \;\; n \to +\infty,$$

skąd na podstawie (3.12)

$$\|h_n - h\| \to 0, \quad \text{gdy} \;\; n \to +\infty. \tag{3.39}$$

Łącząc (3.39) z nierównością (3.14) dostajemy

$$|\langle h_n|f\rangle - \langle h|f\rangle| = |\langle h_n - h|f\rangle| \leqslant \|h_n - h\| \cdot \|f\| \to 0, \quad \text{gdy} \;\; n \to +\infty.$$

Łącząc zaś (3.39) z nierównością (3.20) mamy

$$|g^*(h_n) - g^*(h)| = |g^*(h_n - h)| \leqslant M_g \|h_n - h\| \to 0, \quad \text{gdy} \;\; n \to +\infty.$$

Z własności (3.38) wynika zatem, że

$$\langle h|f\rangle = \lim_{n\to+\infty} \langle h_n|f\rangle = \lim_{n\to+\infty} g^*(h_n) = g^*(h).$$

Dlatego $\langle h|f\rangle = g^*(h)$ dla $h \in \mathrm{L}^2(A, \mathcal{A}', P_x)$. To wobec inkluzji $\mathcal{F} \subset \mathrm{L}^2(A, \mathcal{A}', P_x)$ implikuje warunek (3.37). Stosując twierdzenie 3.8 widzimy, że $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, czyli

$$\delta(h \circ x, g \circ y) \geqslant \delta(f \circ x, g \circ y), \quad h \in \mathcal{F}. \tag{3.40}$$

Ustalając dowolnie $h \in \mathcal{F}' \setminus \mathcal{F}$ widzimy, że

$$\begin{aligned}\int_\Omega |h \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega) + \int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega) \\ \geqslant \frac{1}{2} \int_\Omega |h \circ x(\omega)|^2 \mathrm{d}P(\omega) = +\infty.\end{aligned}$$

Ponieważ $g \circ y \in \mathrm{L}^2(\Omega, \mathcal{A}, P)$, więc

$$\int_\Omega |h \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega) = +\infty > \int_\Omega |f \circ x(\omega) - g \circ y(\omega)|^2 \mathrm{d}P(\omega),$$

co łącznie z (3.40) prowadzi do warunku (3.36). Dlatego $f \in \mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$, co dowodzi twierdzenia w stronę przeciwną ($\Leftarrow$). □

Uwaga 3.12. Z twierdzenia 3.11 wynika, że każda funkcja $f \in \mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$ spełnia warunek (3.35), a więc $f \circ x \in \mathcal{E}(g \circ y|\mathcal{A}'')$, gdzie $\mathcal{E}(\xi|\mathcal{M})$ oznacza zbiór wszystkich warunkowych wartości oczekiwanych zmiennej losowej $\xi\colon \Omega \to \mathbb{C}$ względem σ-ciała $\mathcal{M} \subset \mathcal{A}$. Określenie warunkowej wartości oczekiwanej zmiennej losowej względem

σ-ciała można znaleźć np. w [2, §6.4], [20, §6.3]. Na odwrót, z twierdzenia 3.11 wynika również, że dla każdej funkcji $f \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, jeśli $f \circ x \in \mathcal{E}(g \circ y|\mathcal{A}'')$ to $f \in \mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$. W związku z powyższym warunkowe wartości oczekiwane zmiennych losowych względem σ-ciała można scharakteryzować poprzez funkcje regresji. W konsekwencji warunkowe wartości oczekiwane zmiennych losowych względem σ-ciała można zdefiniować alternatywnie jako rozwiązania zagadnienia optymalizacji.

Uwaga 3.13. Zakładając, że $\mathcal{A}' = \mathcal{A}_x$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ otrzymujemy równość $\mathcal{F} := \mathcal{F}' \cap \mathrm{L}^2(A, \mathcal{A}_x, P_x) = \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, gdzie $\mathcal{F}'$ jest rodziną wszystkich funkcji $f\colon A \to \mathbb{C}$, które są $\mathcal{A}_x$-mierzalne. Stąd na mocy twierdzenia 3.10, $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$, i w konsekwencji $\mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g) \neq \emptyset$; por. dowód twierdzenia 3.11. Funkcje klasy $\mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$ nazywane są zazwyczaj *regresjami typu* (alt. *rodzaju*) *pierwszego zmiennej losowej $g \circ y$ względem zmiennej losowej x*, o ile $A = \mathbb{R}$ i $g \circ y\colon \Omega \to \mathbb{R}$; por. [11, §3.7]. Zatem funkcje klasy $\mathrm{Reg}(\mathcal{F}', \mathfrak{P}_g)$ możemy interpretować jako uogólnienie regresji typu pierwszego i nazywać je regresjami typu pierwszego zmiennej losowej $g \circ y\colon \Omega \to \mathbb{C}$ względem funkcji $x\colon \Omega \to A$. Zagadnienie efektywnego wyznaczania takich regresji nie będzie w tej pracy poruszane. Przypadek, gdy $A = \mathbb{R}$, $g \circ y\colon \Omega \to \mathbb{R}$ i obie zmienne losowe $g \circ y$ i x są jednocześnie typu dyskretnego bądź typu ciągłego, jest rozważany np. w [11, §3.7], [20, §6.3].

3.3 Jednoznaczność funkcji regresji

Z twierdzenia 3.10 wynika, że klasa regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ może składać się z wielu funkcji gdy $\Theta \neq \{\theta\}$. Powstaje więc naturalny problem: kiedy klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest jednoelementowa? Odpowiedzi na to pytanie dostarcza wniosek 3.16, który poprzedzimy następującymi rozważaniami.

Lemat 3.14. *Jeśli $\mathcal{F} \neq \emptyset$ jest rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, to dla każdego $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ zachodzi następujący warunek*

$$f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \;\Leftrightarrow\; \|f - f_0\| = 0, \quad f \in \mathcal{F}. \tag{3.41}$$

Dowód. Ustalmy dowolnie $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ i $f \in \mathcal{F}$. Załóżmy, że $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Z twierdzenia 3.8 wynika, że

$$\langle h|f_0\rangle = g^*(h) \quad \text{i} \quad \langle h|f\rangle = g^*(h), \quad h \in \mathcal{F}.$$

Stąd przyjmując $h := f - f_0$ mamy

$$0 = \langle h|f\rangle - \langle h|f_0\rangle = \langle h|f - f_0\rangle = \|f - f_0\|^2,$$

a więc $\|f - f_0\| = 0$. To dowodzi równoważności w (3.41) w stronę ($\Rightarrow$). Na odwrót załóżmy, że $\|f - f_0\| = 0$. Stosując ponownie twierdzenie 3.8 oraz nierówność Schwarz'a (3.14) stwierdzamy, że dla każdego $h \in \mathcal{F}$,

$$\begin{aligned}|\langle h|f\rangle - g^*(h)| &= |\langle h|f - f_0 + f_0\rangle - g^*(h)| = |\langle h|f - f_0\rangle + \langle h|f_0\rangle - g^*(h)| \\ &= |\langle h|f - f_0\rangle| \leqslant \|h\| \cdot \|f - f_0\| = 0.\end{aligned}$$

Stąd $\langle h|f\rangle = g^*(h)$ dla $h \in \mathcal{F}$, i na mocy twierdzenia 3.8, $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. To dowodzi równoważności w (3.41) w stronę ($\Leftarrow$). Obie implikacje dają łącznie równoważność w (3.41). □

Twierdzenie 3.15. *Jeśli $\mathcal{F} \neq \emptyset$ jest rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, to dla każdego $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ zachodzi równość*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = (\Theta \cap \mathcal{F}) + f_0. \tag{3.42}$$

Dowód. Ustalmy dowolnie $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Z lematu 3.14 wynika, że dla każdej funkcji regresji $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, $h := f - f_0 \in \Theta$, skąd $f = h + f_0 \in (\Theta \cap \mathcal{F}) + f_0$, gdyż $f, f_0 \in \mathcal{F}$ i $\mathcal{F}$ jest rozmaitością liniową w $\mathrm{H}(\mathfrak{P})$. To dowodzi inkluzji

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset (\Theta \cap \mathcal{F}) + f_0. \tag{3.43}$$

Na odwrót, ustalmy dowolnie $f \in (\Theta \cap \mathcal{F}) + f_0$. Wtedy $f - f_0 \in (\Theta \cap \mathcal{F}) \subset \Theta$, skąd $\|f - f_0\| = 0$, i wobec lematu 3.14, $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. To dowodzi inkluzji $(\Theta \cap \mathcal{F}) + f_0 \subset \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, która łącznie z inkluzją (3.43) daje równość (3.42). □

Następny wniosek pokazuje warunek na istnienie dokładnie jednej funkcji regresji w $\mathcal{F}$ względem probabilistycznej struktury regresji $\mathfrak{P}_g$.

Wniosek 3.16. *Jeśli $\mathcal{F} \neq \emptyset$ jest rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$, $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$, to klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest jednoelementowa wtedy i tylko wtedy, gdy $\Theta \cap \mathcal{F} = \{\theta\}$.*

Dowód. Ustalmy dowolnie $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i załóżmy, że istnieje $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Jeśli $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest klasą jednoelementową, to $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{f_0\}$, skąd na mocy twierdzenia 3.15, $\{f_0\} = (\Theta \cap \mathcal{F}) + f_0$. Z równości tej wynika, że dla każdego $h \in \Theta \cap \mathcal{F}$, $h + f_0 = f_0$, czyli $h = \theta$, i tym samym $\Theta \cap \mathcal{F} \subset \{\theta\}$. Ponieważ Θ i $\mathcal{F}$ są rozmaitościami liniowymi w $\mathrm{H}(\mathfrak{P})$, więc $\theta \in \Theta \cap \mathcal{F}$. Dlatego $\Theta \cap \mathcal{F} = \{\theta\}$, o ile klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ składa się z jednej funkcji. Na odwrót, jeśli $\Theta \cap \mathcal{F} = \{\theta\}$, to na mocy twierdzenia 3.15,

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\theta\} + f_0 = \{\theta + f_0\} = \{f_0\},$$

czyli klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ składa się z jednego elementu. □

Zagadnienie jednoznaczności funkcji regresji zilustrujemy przykładami. Dla dowolnie zadanego $p \in \mathbb{N}$ rozważmy klasę $\mathcal{F}_p$ złożoną ze wszystkich wielomianów $h\colon \mathbb{C} \to \mathbb{C}$ stopnia $\deg(h) \leqslant p$. Wtedy $\mathcal{F}_p = \operatorname{lin}(\{h_k\colon k \in \mathbb{Z}_{0,p}\})$, gdzie

$$\mathbb{C} \ni z \mapsto h_k(z) := z^k \in \mathbb{C}, \quad k \in \mathbb{Z}_{0,p}.$$

Definicja 3.17. Dla dowolnego $p \in \mathbb{N}$, *regresją wielomianową stopnia p względem probabilistycznej struktury regresji* $\mathfrak{P}_g$ nazywamy każdą funkcję $f \in \operatorname{Reg}(\mathcal{F}_p, \mathfrak{P}_g)$.

Szczególnym przypadkiem regresji wielomianowych są regresje drugiego rodzaju.

Definicja 3.18. *Regresją drugiego rodzaju względem probabilistycznej struktury regresji* $\mathfrak{P}_g$ nazywamy każdą funkcję $f \in \operatorname{Reg}(\mathcal{F}_1, \mathfrak{P}_g)$.

Regresje wielomianowe będą przedmiotem rozważań w rozdziale czwartym jako szczególny przypadek regresji z klasy $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$, gdzie $\mathcal{F}$ jest jest skończenie wymiarową rozmaitością liniową w przestrzeni $(A \to \mathbb{C}; +, \cdot)$. Teraz skupimy się na przykładzie, który ilustruje zagadnienie jednoznaczności funkcji regresji.

Przykład 3.19. Dla dowolnie zadanego $n \in \mathbb{N}$ rozważmy przestrzeń probabilistyczną $\mathcal{P}$, gdzie $\Omega := \mathbb{Z}_{0,n}$, $\mathcal{A} := 2^{\Omega}$. Przyjmując $p_k := P(\{k\})$ dla $k \in \Omega$ oraz $A := \mathbb{C}$ otrzymujemy dla dowolnie zadanych funkcji $h\colon \mathbb{C} \to \mathbb{C}$ oraz $g\colon \mathbb{C} \to \mathbb{C}$,

$$\int_{\Omega} |h \circ x(k) - g \circ y(k)|^2 \mathrm{d}P(k) = \sum_{k=0}^{n} |h(x_k) - g(y_k)|^2 p_k,$$

gdzie $x_k := x(k)$ i $y_k := y(k)$ dla $k \in \Omega$. Stąd na mocy równości (3.1) mamy

$$\delta(h \circ x, g \circ y) = \sum_{k=0}^{n} |h(x_k) - g(y_k)|^2 p_k, \quad h\colon \mathbb{C} \to \mathbb{C}. \tag{3.44}$$

Ponadto dla każdej funkcji $h\colon \mathbb{C} \to \mathbb{C}$,

$$\|h\|^2 = \int_{\Omega} |h \circ x(k)|^2 \mathrm{d}P(k) = \sum_{k=0}^{n} |h(x_k)|^2 p_k,$$

a więc

$$\Theta = \Big\{h\colon \mathbb{C} \to \mathbb{C}\colon \sum_{k=0}^{n} |h(x_k)|^2 p_k = 0\Big\}. \tag{3.45}$$

Załóżmy dla uproszczenia, że funkcja x jest różnowartościowa i $p_k > 0$ dla $k \in \Omega$. Z twierdzenia o interpolacji wielomianowej wynika istnienie dokładnie jednego wielomianu f_0 stopnia $\deg(f_0) \leqslant n$ spełniającego równości

$$f_0(x_k) = g(y_k), \quad k \in \Omega.$$

To łącznie z (3.44) daje

$$\delta(h \circ x, g \circ y) \geqslant 0 = \delta(f_0 \circ x, g \circ y), \quad h \in \mathcal{F}_p.$$

Załóżmy, że $n \leqslant p$. Wtedy $f_0 \in \mathrm{Reg}(\mathcal{F}_p, \mathfrak{P}_g)$. Ponadto z (3.45) wynika, że

$$h \in \Theta \;\Leftrightarrow\; \Big(h(x_k) = 0, \quad k \in \Omega\Big), \quad h\colon \mathbb{C} \to \mathbb{C}, \tag{3.46}$$

skąd

$$\Theta \cap \mathcal{F}_p = \Big\{h \in \mathcal{F}_p\colon h(x_k) = 0, \quad k \in \Omega\Big\}. \tag{3.47}$$

Rozważmy teraz dwa przypadki.
Przypadek I, gdy $n < p$. Ponieważ funkcja

$$\mathbb{C} \ni z \mapsto h_0(z) := \prod_{k=0}^{n} (z - x_k)$$

jest wielomianem stopnia $\deg(h_0) = n+1$ i $n < p$, więc $h_0 \in \mathcal{F}_p$. Ponadto z równości (3.47) wynika, że dla każdego $h \in \Theta \cap \mathcal{F}_p$, $h = \tilde{h} \cdot h_0$ dla pewnego wielomianu $\tilde{h}$ stopnia $\deg(\tilde{h}) \leqslant p - n - 1$. Dlatego

$$\Theta \cap \mathcal{F}_p = \Big\{\tilde{h} \cdot h_0\colon \tilde{h} \in \mathcal{F}_{p-n-1}\Big\}. \tag{3.48}$$

Korzystając z twierdzenia 3.15 otrzymujemy równość

$$\mathrm{Reg}(\mathcal{F}_p, \mathfrak{P}_g) = \Theta \cap \mathcal{F}_p + f_0 = \Big\{f_0 + \tilde{h} \cdot h_0\colon \tilde{h} \in \mathcal{F}_{p-n-1}\Big\}. \tag{3.49}$$

Z równości (3.48) wynika, że $\Theta \cap \mathcal{F}_p$ jest rozmaitością liniową wymiaru $\dim(\Theta \cap \mathcal{F}_p) = p - n \geqslant 1$. To łącznie z (3.49) oznacza, że $\mathrm{Reg}(\mathcal{F}_p, \mathfrak{P}_g)$ jest rozmaitością afiniczną wymiaru $p - n$. W szczególności klasa $\mathrm{Reg}(\mathcal{F}_p, \mathfrak{P}_g)$ zawiera nieskończenie wiele funkcji.
Przypadek II, gdy $n = p$. Wtedy na mocy (3.46) każdy $h \in \Theta \cap \mathcal{F}_p$ ma $p + 1$ parami różnych miejsc zerowych. Z drugiej strony $\deg(h) \leqslant p$ dla $h \in \mathcal{F}_p$. Zatem $\Theta \cap \mathcal{F}_p = \{\theta\}$, i na mocy twierdzenia 3.15,

$$\mathrm{Reg}(\mathcal{F}_p, \mathfrak{P}_g) = \Theta \cap \mathcal{F}_p + f_0 = \{f_0\},$$

czyli klasa $\mathrm{Reg}(\mathcal{F}_p, \mathfrak{P}_g)$ jest jednoelementowa i składa się z wielomianu interpolacyjnego f_0.

Przykład 3.20. W nawiązaniu do przykładu 3.19 przyjmijmy $\Omega := \mathbb{Z}_{0,3}$, $A := \mathbb{C}$, $\mathbb{C} \ni z \mapsto g(z) := z$ oraz

$$\Omega \ni k \mapsto x(k) := k\ , \quad \Omega \ni k \mapsto y(k) := k^2 - 2k.$$

Przyjmując $p_k := 1/3$ dla $k \in \mathbb{Z}_{0,2}$ i $p_3 := 0$ oraz $\mathcal{F} := \mathcal{F}_3$ dostajemy

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta \cap \mathcal{F} + \{f_0\} = \Big\{\mathbb{C} \ni z \mapsto f_0(z) + \lambda z(z-1)(z-2) \colon \lambda \in \mathbb{C}\Big\},$$

gdzie

$$\mathbb{C} \ni z \mapsto f_0(z) := z^2 - 2z \quad \text{i} \quad \Theta = \Big\{h : \mathbb{C} \to \mathbb{C} \colon h(k) = 0, \quad k \in \mathbb{Z}_{0,2}\Big\}.$$

Zauważmy, że w porównaniu z przykładem 3.19 opuszczamy tutaj założenie $p_k > 0$ dla $k \in \Omega$. Skutkuje to poszerzeniem klasy $\Theta \cap \mathcal{F}$ i w konsekwencji klasy regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Zauważmy również, że przedefiniowanie funkcji x i y dla argumentu $k := 3$ nie ma wpływu na klasy Θ i $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$.

3.4 Rzeczywiste probabilistyczne struktury regresji

Rozważamy tutaj probabilistyczne struktury regresji $\mathfrak{P}$ nad przestrzenią probabilistyczną $\mathcal{P}$ określone w definicji 3.1 z $B := \mathbb{R}$. Podobnie jak w przypadku zespolonym, ze strukturą regresji $\mathfrak{P}$ można powiązać strukturę regresji $\mathfrak{P}_g := (A, \mathbb{R}, \delta; x, g \circ y)$ dla dowolnie zadanej funkcji skalującej $g \colon \mathbb{R} \to \mathbb{R}$. O modelu funkcyjnym $\mathcal{F}$ zakładamy zaś, że $\mathcal{F} \subset (A \to \mathbb{R})$. Ponieważ $\mathbb{R} \subset \mathbb{C}$, więc $\mathcal{F}$ jest również modelem funkcyjnym zespolonej probabilistycznej struktury regresji $\mathfrak{P}_{\mathbb{C}} := (A, \mathbb{C}, \delta_{\mathbb{C}}; x, y)$, gdzie funkcja $\delta_{\mathbb{C}} \colon (\Omega \to \mathbb{C}) \times (\Omega \to \mathbb{C}) \to \overline{\mathbb{R}}$ spełnia równość

$$\delta_{\mathbb{C}}(u, v) = \int_{\Omega} |u(\omega) - v(\omega)|^2 \, \mathrm{d}P(\omega), \tag{3.50}$$

o ile funkcje $u, v \colon \Omega \to \mathbb{C}$ są $\mathcal{A}$-mierzalne, zaś $\delta_{\mathbb{C}}(u, v) = +\infty$ w przeciwnym przypadku. Wtedy również $\mathcal{F}$ jest modelem funkcyjnym w przeskalowanej zespolonej probabilistycznej strukturze regresji $\mathfrak{P}_{\mathbb{C},g} := (A, \mathbb{C}, \delta_{\mathbb{C}}; x, g \circ y)$.

Uwaga 3.21. Porównując równości (3.1) z (3.50) stwierdzamy, że dla każdej funkcji $g \colon \mathbb{R} \to \mathbb{R}$ i każdej niepustej klasy $\mathcal{F} \subset (A \to \mathbb{R})$,

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_{\mathbb{C},g}). \tag{3.51}$$

Ustalając bowiem dowolnie $g \colon \mathbb{R} \to \mathbb{R}$, $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ i $f \in \mathcal{F}$ mamy

$$\delta(f \circ x, g \circ y) \geqslant \delta(f_0 \circ x, g \circ y). \tag{3.52}$$

Jeśli $\delta_{\mathbb{C}}(f \circ x, g \circ y) = +\infty$ to oczywiście

$$\delta_{\mathbb{C}}(f \circ x, g \circ y) \geqslant \delta_{\mathbb{C}}(f_0 \circ x, g \circ y).$$

W przeciwnym przypadku $\delta_{\mathbb{C}}(f\circ x, g\circ y) < +\infty$. Wtedy obie funkcje $f\circ x$ i $g\circ y$ są $\mathcal{A}$-mierzalne, i wobec (3.52) mamy

$$\delta(f_0\circ x, g\circ y) \leqslant \delta(f\circ x, g\circ y) = \delta_{\mathbb{C}}(f\circ x, g\circ y) < +\infty.$$

Dlatego funkcja $f_0\circ x$ jest $\mathcal{A}$-mierzalna, i w konsekwencji

$$\delta_{\mathbb{C}}(f_0\circ x, g\circ y) = \delta(f_0\circ x, g\circ y) \leqslant \delta_{\mathbb{C}}(f\circ x, g\circ y).$$

Zatem

$$\delta_{\mathbb{C}}(f\circ x, g\circ y) \geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y), \quad f\in\mathcal{F}.$$

Stąd $f_0 \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_{\mathbb{C},g})$, co dowodzi inkluzji

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_{\mathbb{C},g}).$$

Zamieniając w powyższym rozumowaniu δ z $\delta_{\mathbb{C}}$ oraz $\mathfrak{P}_g$ z $\mathfrak{P}_{\mathbb{C},g}$ można wykazać inkluzję przeciwną

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_{\mathbb{C},g}) \subset \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g).$$

Obie inkluzje dają łącznie równość (3.51).

Dla dowolnej klasy $\mathcal{F} \subset (A\to\mathbb{R})$ rozważmy rodzinę funkcji $\mathcal{F}_{\mathbb{C}} := \{f\in(A\to\mathbb{C})\colon \operatorname{Re} f\in\mathcal{F}\}$. Wtedy $\mathcal{F}\subset\mathcal{F}_{\mathbb{C}}\subset(A\to\mathbb{C})$. Można postawić pytanie o zależność pomiędzy klasami funkcji regresji $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$ i $\operatorname{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g})$.

Twierdzenie 3.22. *Dla dowolnych $\mathcal{F}\subset(A\to\mathbb{R})$ i $g\colon\mathbb{R}\to\mathbb{R}$ zachodzi inkluzja*

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset \operatorname{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g}). \tag{3.53}$$

Co więcej, dla każdego $f_0\in\operatorname{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g})$, jeśli $\delta_{\mathbb{C}}(f_0\circ x, g\circ y) < +\infty$, to ma miejsce równość

$$P\Big(\{\omega\in\Omega\colon \operatorname{Im} f_0\circ x(\omega)\neq 0\}\Big) = 0. \tag{3.54}$$

Dowód. Załóżmy, że $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)\neq\emptyset$ i ustalmy dowolnie $f_0\in\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Wtedy

$$\delta(f\circ x, g\circ y) \geqslant \delta(f_0\circ x, g\circ y), \quad f\in\mathcal{F}. \tag{3.55}$$

Załóżmy, że $g\circ y$ jest funkcją $\mathcal{A}$-mierzalną. Ustalmy dowolnie funkcję $f\in\mathcal{F}_{\mathbb{C}}$. Zakładając, że $f\circ x$ jest funkcją $\mathcal{A}$-mierzalną (wtedy $\operatorname{Re} f\circ x$ i $\operatorname{Im} f\circ x$ są funkcjami $\mathcal{A}$-mierzalnymi) mamy na mocy (3.55),

$$\begin{aligned}
\delta_{\mathbb{C}}(f\circ x, g\circ y) &= \int_\Omega |f\circ x - g\circ y|^2 \mathrm{d}P = \int_\Omega |\operatorname{Re} f\circ x - g\circ y + \mathrm{i}\operatorname{Im} f\circ x|^2 \mathrm{d}P \\
&= \int_\Omega |\operatorname{Re} f\circ x - g\circ y|^2 \mathrm{d}P + \int_\Omega |\operatorname{Im} f\circ x|^2 \mathrm{d}P \\
&\geqslant \int_\Omega |\operatorname{Re} f\circ x - g\circ y|^2 \mathrm{d}P = \delta(\operatorname{Re} f\circ x, g\circ y) \\
&\geqslant \delta(f_0\circ x, g\circ y).
\end{aligned}$$

Jeśli więc $\delta_{\mathbb{C}}(f\circ x, g\circ y)<+\infty$ to $\delta(f_0\circ x, g\circ y)<+\infty$ i tym samym $f_0\circ x$ jest funkcją $\mathcal{A}$-mierzalną co daje

$$\delta_{\mathbb{C}}(f\circ x, g\circ y)\geqslant \delta(f_0\circ x, g\circ y)=\delta_{\mathbb{C}}(f_0\circ x, g\circ y).$$

Jeśli zaś $\delta_{\mathbb{C}}(f\circ x, g\circ y)=+\infty$ to

$$\delta_{\mathbb{C}}(f\circ x, g\circ y)\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y).$$

Dlatego

$$\delta_{\mathbb{C}}(f\circ x, g\circ y)\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y),\quad f\in\mathcal{F}_{\mathbb{C}}. \tag{3.56}$$

W przypadku, gdy jedna z funkcji $f\circ x$ lub $g\circ y$ nie jest $\mathcal{A}$-mierzalna to

$$\delta_{\mathbb{C}}(f\circ x, g\circ y)=+\infty\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y),\quad f\in\mathcal{F}_{\mathbb{C}}.$$

To łącznie z (3.56) oznacza, że $f_0\in \mathrm{Reg}(\mathcal{F}_{\mathbb{C}},\mathfrak{P}_{\mathbb{C},g})$, co dowodzi inkluzji (3.53). Ustalmy teraz dowolnie $f_0\in \mathrm{Reg}(\mathcal{F}_{\mathbb{C}},\mathfrak{P}_{\mathbb{C},g})$. Wtedy

$$\delta_{\mathbb{C}}(f\circ x, g\circ y)\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y),\quad f\in\mathcal{F}_{\mathbb{C}}. \tag{3.57}$$

Ponieważ $\operatorname{Re} f_0\in\mathcal{F}\subset\mathcal{F}_{\mathbb{C}}$, więc z (3.57) wynika, że

$$\delta_{\mathbb{C}}(\operatorname{Re} f_0\circ x, g\circ y)\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y).$$

Załóżmy, że $\delta_{\mathbb{C}}(f_0\circ x, g\circ y)<+\infty$. Wtedy $f_0\circ x$ i $g\circ y$ są funkcjami $\mathcal{A}$-mierzalnymi oraz

$$\begin{aligned}\delta_{\mathbb{C}}(f_0\circ x, g\circ y)&=\int_\Omega |f_0\circ x-g\circ y|^2\mathrm{d}P=\int_\Omega |\operatorname{Re} f_0\circ x-g\circ y+\mathrm{i}\operatorname{Im} f_0\circ x|^2\mathrm{d}P\\&=\int_\Omega |\operatorname{Re} f_0\circ x-g\circ y|^2\mathrm{d}P+\int_\Omega |\operatorname{Im} f_0\circ x|^2\mathrm{d}P\\&=\delta_{\mathbb{C}}(\operatorname{Re} f_0\circ x, g\circ y)+\int_\Omega |\operatorname{Im} f_0\circ x|^2\mathrm{d}P\\&\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y)+\int_\Omega |\operatorname{Im} f_0\circ x|^2\mathrm{d}P.\end{aligned}$$

Stąd

$$\int_\Omega |\operatorname{Im} f_0\circ x|^2\mathrm{d}P\leqslant 0,$$

i w konsekwencji zachodzi równość (3.54). □

Uwaga 3.23. Zauważmy, że założenie "$\delta_{\mathbb{C}}(f_0\circ x, g\circ y)<+\infty$" w twierdzeniu 3.22 jest istotne. Aby to wykazać załóżmy, że tak nie jest. Wtedy na mocy założenia $f_0\in\mathrm{Reg}(\mathcal{F}_{\mathbb{C}},\mathfrak{P}_{\mathbb{C},g})$ mamy

$$\delta_{\mathbb{C}}(f\circ x, g\circ y)\geqslant \delta_{\mathbb{C}}(f_0\circ x, g\circ y)=+\infty,\quad f\in\mathcal{F}_{\mathbb{C}},$$

a więc $\delta_{\mathbb{C}}(f\circ x, g\circ y) = +\infty$ dla $f \in \mathcal{F}_{\mathbb{C}}$. Stąd $\mathcal{F}_{\mathbb{C}} \subset \mathrm{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g})$ i w konsekwencji

$$\mathrm{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g}) = \mathcal{F}_{\mathbb{C}}, \tag{3.58}$$

gdyż zawsze $\mathrm{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g}) \subset \mathcal{F}_{\mathbb{C}}$. Dobierając klasę $\mathcal{F}$ w taki sposób, aby klasa $\mathcal{F}_{\mathbb{C}}$ zawierała funkcję f_0 taką, że $f_0 \circ x$ jest funkcją $\mathcal{A}$-mierzalną i spełniającą warunek

$$P\Big(\{\omega \in \Omega\colon\ \mathrm{Im}\, f_0 \circ x(\omega) \neq 0\}\Big) > 0, \tag{3.59}$$

widzimy, że warunek (3.54) nie zachodzi. Dla przykładu rozważmy probabilistyczną strukturę regresji $\mathfrak{P} := (A, B, \delta; x, y)$ nad przestrzenią probabilistyczną $\mathcal{P}$, gdzie $A := \mathbb{R}$, $B := \mathbb{R}$, $\Omega := \mathbb{R}$, $\Omega \ni \omega \mapsto x(\omega) := \omega$, $\Omega \ni \omega \mapsto y(\omega) := \omega$, $\mathcal{A}$ jest σ-ciałem zbiorów borelowskich na $\mathbb{R}$, zaś P jest miarą probabilistyczną określoną wzorem

$$\mathcal{A} \ni S \mapsto P(S) := \begin{cases} 1, & \text{gdy } \ 0 \in S, \\ 0, & \text{gdy } \ 0 \notin S. \end{cases} \tag{3.60}$$

Przyjmijmy $\mathcal{F} := (\mathbb{R} \to \mathbb{R})$ i $g := \mathbb{1}_K$, gdzie K jest dowolnie zadanym podzbiorem zbioru $\mathbb{R}$, który nie jest borelowski, czyli $K \in 2^{\mathbb{R}} \setminus \mathcal{A}$. Wtedy $\mathcal{F}_{\mathbb{C}} = (\mathbb{R} \to \mathbb{C})$ i g nie jest funkcją $\mathcal{A}$-mierzalną. Ponieważ $g \circ y = g$, więc $g \circ y$ nie jest funkcją $\mathcal{A}$-mierzalną i tym samym $\delta_{\mathbb{C}}(f_0 \circ x, g \circ y) = +\infty$. To zaś implikuje równość (3.58), czyli $\mathrm{Reg}(\mathcal{F}_{\mathbb{C}}, \mathfrak{P}_{\mathbb{C},g}) = (\mathbb{R} \to \mathbb{C})$. Przyjmując $S := [-1; 1]$ i $f_0 := \mathrm{i} \cdot \mathbb{1}_S$ widzimy, że $f_0 \circ x$ jest funkcją $\mathcal{A}$-mierzalną, gdyż $f_0 \circ x = f_0$. Ponadto na mocy wzoru (3.60) dostajemy

$$P\Big(\{\omega \in \Omega\colon\ \mathrm{Im}\, f_0 \circ x(\omega) \neq 0\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathrm{Im}(\mathrm{i} \cdot \mathbb{1}_S)(\omega) \neq 0\}\Big) = P(S) = 1 > 0,$$

a więc zachodzi nierówność (3.59).

Uwaga 3.24. Równość (3.54) jest równoważna każdej z równości

$$P\Big(\{\omega \in \Omega\colon\ f_0 \circ x(\omega) \neq \mathrm{Re}\, f_0 \circ x(\omega)\}\Big) = 0$$

oraz

$$P\Big(\{\omega \in \Omega\colon\ f_0 \circ x(\omega) \in \mathbb{R}\}\Big) = 1.$$

Rozdział 4

Obliczanie funkcji regresji

Ustalmy dowolnie przestrzeń probabilistyczną $\mathcal{P} := (\Omega, \mathcal{A}, P)$, zbiór $A \neq \emptyset$ oraz funkcje $x\colon \Omega \to A$ i $y\colon \Omega \to \mathbb{C}$. Przypomnijmy, że funkcje x i y generują σ-ciała $\mathcal{A}_x$ i $\mathcal{A}_y$ określone wzorami (3.7) i (3.15) odpowiednio. Załóżmy, że klasa przebiegów teoretycznych $\mathcal{F} \subset (A \to \mathbb{C})$ jest rozmaitością liniową w przestrzeni $((A \to \mathbb{C}); +, \cdot)$. Przypomnijmy, że operator wartości oczekiwanej E nad przestrzenią probabilistyczną $(\Omega, \mathcal{A}, P)$ jest określony wzorem

$$\mathrm{E}(f) := \int_\Omega f(\omega)\mathrm{d}P(\omega), \quad f \in \mathrm{L}^1(\Omega, \mathcal{A}, P); \tag{4.1}$$

por. [29, §11]. Porównując równości (3.1) i (4.1) możemy wyrazić kryterium dopasowania δ za pomocą operatora wartości oczekiwanej E następująco

$$\delta(u, v) = \int_\Omega |u(\omega) - v(\omega)|^2 \mathrm{d}P(\omega) = \mathrm{E}(|u - v|^2), \tag{4.2}$$

o ile funkcje $u, v\colon \Omega \to B$ są $\mathcal{A}$-mierzalne, zaś $\delta(u, v) = +\infty$ w przeciwnym przypadku. Funkcja δ określa zespoloną probabilistyczną strukturę regresji $\mathfrak{P} := (A, \mathbb{C}, \delta; x, y)$. W dalszym ciągu będziemy rozważać problem regresji w klasie funkcji $\mathcal{F}$ dla struktury $\mathfrak{P}$. Polega on na wyznaczeniu zbioru $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$ złożonego z funkcji regresji w klasie $\mathcal{F}$ względem struktury $\mathfrak{P}$. Minimalizują one, na podstawie (3.2) i (4.2), funkcjonał

$$\mathcal{F} \ni f \mapsto F(f) = \mathrm{E}(|f \circ x - y|^2),$$

gdy $y \in \mathrm{L}(\Omega, \mathcal{A})$ i $\mathcal{F} \subset \mathrm{L}(A, \mathcal{A}_x)$, czyli gdy każda funkcja $f \in \mathcal{F}$ jest $\mathcal{A}_x$-mierzalna. Innymi słowy szukamy funkcji $f_0 \in \mathcal{F}$ spełniających, na podstawie (3.3), warunek

$$\mathrm{E}(|f_0 \circ x - y|^2) \leqslant \mathrm{E}(|f \circ x - y|^2), \quad f \in \mathcal{F}.$$

Ustalmy dowolnie funkcje $u, v \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. Z (3.12), (3.10) otrzymujemy na mocy uwagi 3.5 i wzoru (4.1) równości

$$\|u\|^2 = \int_A |u(t)|^2 \mathrm{d}P_x(t) = \int_\Omega |u \circ x(\omega)|^2 \mathrm{d}P(\omega) = \mathrm{E}(|u \circ x|^2) \tag{4.3}$$

oraz

$$\langle u|v\rangle = \int_A u(t) \cdot \overline{v(t)} \mathrm{d}P_x(t) = \int_\Omega (u \circ x(\omega))\overline{v \circ x(\omega)} \mathrm{d}P(\omega) = \mathrm{E}((u \circ x)(\overline{v \circ x})), \tag{4.4}$$

gdzie P_x jest miarą probabilistyczną na σ-ciele $\mathcal{A}_x$ określoną wzorem (3.8). Ponadto, dla każdego $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ otrzymujemy na mocy wzorów (3.19) i (4.1) równości

$$g^*(u) = \int_\Omega (u \circ x(\omega))\overline{(g \circ y(\omega))}\,\mathrm{d}P(\omega) = \mathrm{E}((u \circ x)\overline{(g \circ y)}), \tag{4.5}$$

gdzie P_y jest miarą probabilistyczną na σ-ciele $\mathcal{A}_y$ określoną wzorem (3.16).

Załóżmy, że $\mathcal{F}$ jest dowolnie zadaną rozmaitością liniową w przestrzeni pseudo-Hilberta $\mathrm{H}(\mathfrak{P})$ i dana jest funkcja skalująca $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$. Jeśli dodatkowo $\mathcal{F}$ jest zbiorem domkniętym w $\mathrm{H}(\mathfrak{P})$ to na mocy twierdzenia 3.10, $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$ oraz zachodzi równość (3.28). Umożliwia ona wyznaczenie funkcji regresji w modelu funkcyjnym $\mathcal{F}$ względem struktury regresji $\mathfrak{P}_g$ pod warunkiem, że możemy określić zbiór liniowy $(\mathcal{F} \cap S)^\perp \cap \mathcal{F}$. Ogólnie jest to zadanie trudne. Jednak w przypadku, gdy klasa $\mathcal{F}$ jest skończenie wymiarowa możemy efektywnie wyznaczyć funkcje regresji w terminach zadanej bazy rozmaitości liniowej $\mathcal{F}$. Zagadnienie to jest szczególnie istotne z praktycznego punktu widzenia i zostanie omówione w dalszym ciągu tego rozdziału. Odtąd zawsze zakładamy, że $\mathfrak{P}$ jest zadaną probabilistyczną strukturą regresji nad przestrzenią probabilistyczną $\mathcal{P}$.

4.1 Struktura klasy regresji

Z twierdzenia 3.15 wynika, że $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = (\Theta \cap \mathcal{F}) + f_0$ dla dowolnie wybranej funkcji $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$, o ile taka funkcja istnieje. Geometrycznie rzecz ujmując klasa regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest przesunięciem w przestrzeni $\mathrm{H}(\mathfrak{P})$ zbioru $\Theta \cap \mathcal{F}$ o wektor f_0. Dlatego struktura klasy $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest opisywana, z dokładnością do przesunięcia, przez własności zbioru $\Theta \cap \mathcal{F}$, które będą przedmiotem rozważań w tym podrozdziale. W tym celu użyjemy zmodyfikowanej metody ortogonalizacji Gramma-Schmidta ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ rozpinającego rozmaitość liniową $\mathcal{F}$ dla pewnego $p \in \mathbb{N}$. Polega ona na wyznaczeniu ciągu określonego rekurencyjnie w następujący sposób

$$h'_1 := h_1 \quad \text{i} \quad h'_n := h_n - \sum_{k=1}^{n-1} \langle h_n | h'_k \rangle_* \cdot h'_k, \quad n \in \mathbb{Z}_{2,p}, \tag{4.6}$$

gdzie symbol $\langle \cdot | \cdot \rangle_*$ zdefiniowany jest przez

$$\mathrm{L}^2(A, \mathcal{A}_x, P_x) \times \mathrm{L}^2(A, \mathcal{A}_x, P_x) \ni (u, v) \mapsto \langle u | v \rangle_* := \begin{cases} \dfrac{\langle u | v \rangle}{\|v\|^2}, & \text{gdy } \|v\| > 0, \\ 0, & \text{gdy } \|v\| = 0. \end{cases} \tag{4.7}$$

Dla dowolnych $u, v \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ przyjmujemy standardowe oznaczenie $u \perp v$ wtedy i tylko wtedy, gdy $\langle u | v \rangle = 0$, tzn. funkcje u i v są ortogonalne w $\mathrm{H}(\mathfrak{P})$.

Uwaga 4.1. Ze wzoru (4.6) wynika, że dla każdego $p \in \mathbb{N}$ i każdego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ zachodzi równość

$$\operatorname{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p}\}) = \operatorname{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$$

oraz $h'_k \perp h'_l$ dla $k, l \in \mathbb{Z}_{1,p}$, o ile $k \neq l$.

Uwaga 4.2. Ze wzoru (3.12), (4.6) i uwagi 4.1 wynika, że dla każdego $p \in \mathbb{N}$ i każdego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$,

$$\|h'_n\|^2 = \langle h_n | h'_n \rangle, \quad n \in \mathbb{Z}_{1,p}. \tag{4.8}$$

Ponieważ $h'_1 = h_1$, więc równość w (4.8) jest oczywista dla $n = 1$. Jeśli zaś $n \in \mathbb{Z}_{2,p}$, to na mocy (3.12), (4.6) i uwagi 4.1,

$$\begin{aligned} \|h'_n\|^2 = \langle h'_n | h'_n \rangle &= \left\langle h_n - \sum_{k=1}^{n-1} \langle h_n | h'_k \rangle_* h'_k \middle| h'_n \right\rangle \\ &= \langle h_n | h'_n \rangle - \sum_{k=1}^{n-1} \langle h_n | h'_k \rangle_* \langle h'_k | h'_n \rangle = \langle h_n | h'_n \rangle. \end{aligned}$$

W szczególności z (4.8) wynika, że

$$\|h'_n\| > 0 \;\Rightarrow\; \|h_n\| > 0, \qquad n \in \mathbb{Z}_{1,p}. \tag{4.9}$$

Za pomocą ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ definiujemy następujące zbiory

$$E_0 := \{k \in \mathbb{Z}_{1,p} \colon \|h'_k\| = 0\} \quad \text{i} \quad E := \{k \in \mathbb{Z}_{1,p} \colon \|h'_k\| > 0\}. \tag{4.10}$$

Symbolem $\overline{\overline{S}}$ będziemy oznaczać ilość elementów zbioru skończonego S.

Twierdzenie 4.3. *Dla dowolnie zadanych $p \in \mathbb{N}$ i $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ niech $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ będzie ciągiem zdefiniowanym wzorem* (4.6)*. Wtedy*

$$\Theta \cap \mathcal{F} = \mathcal{F}_0 := \operatorname{lin}(\{h'_k \colon k \in E_0\}), \tag{4.11}$$

gdzie $\mathcal{F} := \operatorname{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$, oraz

$$\dim(\mathcal{F}_0) = \dim(\mathcal{F}) - \overline{\overline{E}}. \tag{4.12}$$

Dowód. Ustalmy dowolnie $p \in \mathbb{N}$ i ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$.

Rozważmy najpierw przypadek, gdy $E_0 = \emptyset$. Wtedy $\mathcal{F}_0 = \operatorname{lin}(\emptyset) = \{\theta\}$, skąd $\dim(\mathcal{F}_0) = 0$. Ponieważ $E = \mathbb{Z}_{1,p} \setminus E_0 = \mathbb{Z}_{1,p}$, więc $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$. Z uwagi 4.1 wynika, że dla dowolnie zadanego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$,

$$\left\| \sum_{k=1}^{p} \lambda_k h'_k \right\|^2 = \left\langle \sum_{k=1}^{p} \lambda_k h'_k \middle| \sum_{l=1}^{p} \lambda_l h'_l \right\rangle = \sum_{k=1}^{p} \sum_{l=1}^{p} \lambda_k \overline{\lambda_l} \langle h'_k | h'_l \rangle = \sum_{k=1}^{p} |\lambda_k|^2 \|h'_k\|^2. \tag{4.13}$$

Jeśli więc $\sum_{k=1}^{p} \lambda_k h'_k = \theta$ to $\sum_{k=1}^{p} |\lambda_k|^2 \|h'_k\|^2 = 0$, skąd $\lambda_k = 0$ dla $k \in \mathbb{Z}_{1,p}$. Zatem ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ jest liniowo niezależny, skąd wobec uwagi 4.1,

$$\dim(\mathcal{F}) = \dim\Big(\mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})\Big) = \dim\Big(\mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p}\})\Big) = p.$$

Dlatego

$$\dim(\mathcal{F}_0) = 0 = p - p = \dim(\mathcal{F}) - \overline{\overline{E}},$$

co dowodzi równości (4.12). Ponadto dla każdego $h \in \Theta \cap \mathcal{F}$, $h = \sum_{k=1}^{p} \lambda_k h'_k$ dla pewnego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$, skąd na mocy (4.13) mamy $0 = \|h\|^2 = \sum_{k=1}^{p} |\lambda_k|^2 \|h'_k\|^2$. Dlatego $\lambda_k = 0$ dla $k \in \mathbb{Z}_{1,p}$, i tym samym $h = \theta$. To dowodzi inkluzji $\Theta \cap \mathcal{F} \subset \{\theta\}$, która łącznie z oczywistą inkluzją przeciwną $\{\theta\} \subset \Theta \cap \mathcal{F}$ daje równość $\Theta \cap \mathcal{F} = \{\theta\}$. Ponieważ $\mathcal{F}_0 = \{\theta\}$, więc zachodzi równość (4.11).

Załóżmy teraz, że $E = \emptyset$. Wtedy $E_0 = \mathbb{Z}_{1,p} \backslash E = \mathbb{Z}_{1,p}$, i wobec uwagi 4.1, $\mathcal{F} = \mathcal{F}_0$. Stąd

$$\dim(\mathcal{F}_0) = \dim(\mathcal{F}) - 0 = \dim(\mathcal{F}) - \overline{\overline{E}},$$

co daje równość (4.12). Ponieważ $\|h'_k\| = 0$ dla $k \in \mathbb{Z}_{1,p}$, więc dla każdego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$,

$$\Big\| \sum_{k=1}^{p} \lambda_k h'_k \Big\| \leqslant \sum_{k=1}^{p} |\lambda_k| \|h'_k\| = 0,$$

czyli $\sum_{k=1}^{p} \lambda_k h'_k \in \Theta \cap \mathcal{F}$. Dlatego $\mathcal{F} \subset \Theta \cap \mathcal{F}$. Ponieważ $\Theta \cap \mathcal{F} \subset \mathcal{F}$, więc $\Theta \cap \mathcal{F} = \mathcal{F} = \mathcal{F}_0$, i tym samym zachodzi równość (4.11).

Pozostaje rozważyć przypadek, gdy $E_0 \neq \emptyset \neq E$. Wykażemy najpierw równość (4.11). Dla dowolnego $v \in \mathcal{F}_0$ istnieje $\lambda \colon E_0 \to \mathbb{C}$ takie, że $v = \sum_{k \in E_0} \lambda_k h'_k$, i w konsekwencji

$$\|v\| \leqslant \sum_{k \in E_0} \|\lambda_k h'_k\| = \sum_{k \in E_0} |\lambda_k| \|h'_k\| = 0,$$

czyli $v \in \Theta$. Dlatego $\mathcal{F}_0 \subset \Theta$. Z uwagi 4.1 wynika, że

$$\mathcal{F}_0 = \mathrm{lin}(\{h'_k \colon k \in E_0\}) \subset \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p}\}) = \mathcal{F}.$$

Zatem $\mathcal{F}_0 \subset \Theta \cap \mathcal{F}$. W celu wykazania inkluzji przeciwnej ustalmy dowolnie $v \in \Theta \cap \mathcal{F}$. Korzystając ponownie z uwagi 4.1 widzimy, że $v \in \Theta \cap \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p}\})$. Istnieje więc ciąg $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$ taki, że $v = \sum_{k=1}^{p} \lambda_k h'_k$. Ponadto ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ jest ortogonalny, a więc $\sum_{k=1}^{p} |\lambda_k|^2 \|h'_k\|^2 = \|v\|^2 = 0$, gdyż $v \in \Theta$. Stąd $|\lambda_k| \|h'_k\| = 0$ dla $k \in \mathbb{Z}_{1,p}$. Ponieważ $\|h'_k\| > 0$ dla $k \in E$, więc $\lambda_k = 0$ dla $k \in E$. Zatem

$$v = \sum_{k=1}^{p} \lambda_k h'_k = \sum_{k \in E} \lambda_k h'_k + \sum_{k \in E_0} \lambda_k h'_k = \sum_{k \in E_0} \lambda_k h'_k \in \mathcal{F}_0.$$

To dowodzi inkluzji $\Theta \cap \mathcal{F} \subset \mathcal{F}_0$, która wraz z wykazaną już inkluzją przeciwną daje równość (4.11).

Pozostaje wykazać równość (4.12). Przyjmijmy $\mathcal{F}' := \mathrm{lin}(\{h'_k \colon k \in E\})$. Na mocy uwagi 4.1, $h'_k \perp h'_l$ dla $k, l \in E$, o ile $k \neq l$. Ponieważ $\|h'_k\| > 0$ dla $k \in E$, więc zbiór wektorów $\{h'_k \colon k \in E\}$ jest liniowo niezależny (por. fragment dowodu po (4.13)), i w konsekwencji $\dim(\mathcal{F}') = \overline{\overline{E}}$.

Wykażemy, że $\mathcal{F}_0 + \mathcal{F}' = \mathcal{F}$ i $\mathcal{F}_0 \cap \mathcal{F}' = \{\theta\}$.

Dla dowolnie zadanego $v \in \mathcal{F}_0 + \mathcal{F}'$ istnieją $v' \in \mathcal{F}_0$ i $v'' \in \mathcal{F}'$ takie, że $v = v' + v''$. Z definicji zbiorów $\mathcal{F}_0$ i $\mathcal{F}'$ wynika istnienie funkcji $\lambda' \colon E_0 \to \mathbb{C}$ i $\lambda'' \colon E \to \mathbb{C}$ spełniających równości $v' = \sum\limits_{k \in E_0} \lambda'_k h'_k$ i $v'' = \sum\limits_{k \in E} \lambda''_k h'_k$. Przyjmując

$$\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k := \begin{cases} \lambda'_k, & \text{gdy } k \in E_0 \\ \lambda''_k, & \text{gdy } k \in E \end{cases}$$

stwierdzamy, na mocy uwagi 4.1, że

$$v = v' + v'' = \sum_{k \in E_0} \lambda'_k h'_k + \sum_{k \in E} \lambda''_k h'_k = \sum_{k \in E_0 \cup E} \lambda_k h'_k = \sum_{k=1}^{p} \lambda_k h'_k \in \mathcal{F}.$$

Dlatego $\mathcal{F}_0 + \mathcal{F}' \subset \mathcal{F}$. Na odwrót, ustalmy dowolnie $v \in \mathcal{F}$. Z uwagi 4.1 wynika, że $v = \sum\limits_{k=1}^{p} \lambda_k h'_k$ dla pewnego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$. Przyjmując $v' := \sum\limits_{k \in E_0} \lambda_k h'_k$ i $v'' := \sum\limits_{k \in E} \lambda_k h'_k$ widzimy, że $v' \in \mathcal{F}_0$, $v'' \in \mathcal{F}'$ oraz

$$v' + v'' = \sum_{k \in E_0} \lambda_k h'_k + \sum_{k \in E} \lambda_k h'_k = \sum_{k \in E_0 \cup E} \lambda_k h'_k = \sum_{k=1}^{p} \lambda_k h'_k = v.$$

To dowodzi inkluzji $\mathcal{F} \subset \mathcal{F}_0 + \mathcal{F}'$, która łącznie z inkluzją przeciwną daje równość $\mathcal{F}_0 + \mathcal{F}' = \mathcal{F}$.

Ustalmy teraz dowolnie $v \in \mathcal{F}_0 \cap \mathcal{F}'$. Z równości (4.11) wynika, że $v \in \Theta$, czyli $\|v\| = 0$. Z drugiej strony $v \in \mathcal{F}'$, a więc $v = \sum\limits_{k \in E} \lambda_k h'_k$ dla pewnego $\lambda \colon E \to \mathbb{C}$. Na mocy uwagi 4.1, $h'_k \perp h'_l$ dla $k, l \in \mathbb{Z}_{1,p}$, o ile $k \neq l$. Stąd

$$0 = \|v\|^2 = \Big\| \sum_{k \in E} \lambda_k h'_k \Big\|^2 = \sum_{k \in E} |\lambda_k|^2 \|h'_k\|^2,$$

i w konsekwencji $\lambda_k = 0$ dla $k \in E$. Dlatego $v = \sum\limits_{k \in E} \lambda_k h'_k = \theta$. Stąd $\mathcal{F}_0 \cap \mathcal{F}' \subset \{\theta\}$, i tym samym $\mathcal{F}_0 \cap \mathcal{F}' = \{\theta\}$. Ponieważ jednocześnie $\mathcal{F}_0 + \mathcal{F}' = \mathcal{F}$, więc korzystając z [19, Twierdzenie 9.10] otrzymujemy

$$\dim(\mathcal{F}) = \dim(\mathcal{F}_0) + \dim(\mathcal{F}') - \dim(\mathcal{F}_0 \cap \mathcal{F}') = \dim(\mathcal{F}_0) + \overline{\overline{E}},$$

co dowodzi równości (4.12), i tym samym kończy dowód. □

Wniosek 4.4. *Przy założeniach* twierdzenia 4.3*, jeśli* $\dim(\mathcal{F}) = p$*, to* $\dim(\mathcal{F}_0) = \overline{\overline{E_0}}$.

Dowód. Z założeń wniosku wynika na mocy twierdzenia 4.3, że

$$\dim(\mathcal{F}_0) = \dim(\mathcal{F}) - \overline{\overline{E}} = p - \overline{\overline{E}}.$$

Ponieważ $E \cup E_0 = \mathbb{Z}_{1,p}$ i $E \cap E_0 = \emptyset$, więc $\overline{\overline{E}} + \overline{\overline{E_0}} = \overline{\overline{E \cup E_0}} = \overline{\overline{(\mathbb{Z}_{1,p})}} = p$. Dlatego $\dim(\mathcal{F}_0) = p - \overline{\overline{E}} = \overline{\overline{E_0}}$, co kończy dowód. □

Wniosek 4.5 (por. [25, Lemma 5.2])**.** *Przy założeniach* twierdzenia 4.3 *ciąg* $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ *jest liniowo niezależny i* $\Theta \cap \mathcal{F} = \{\theta\}$ *wtedy i tylko wtedy, gdy* $\|h'_k\| > 0$ *dla* $k \in \mathbb{Z}_{1,p}$.

Dowód. Ustalmy dowolnie $p \in \mathbb{N}$ i ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. Załóżmy, że ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ jest liniowo niezależny i $\Theta \cap \mathcal{F} = \{\theta\}$. Wtedy $\dim(\mathcal{F}) = p$ i $\dim(\Theta \cap \mathcal{F}) = 0$. Na mocy twierdzenia 4.3 zachodzą równości (4.11) i (4.12), skąd

$$p - \overline{\overline{E}} = \dim(\mathcal{F}) - \overline{\overline{E}} = \dim(\mathcal{F}_0) = \dim(\Theta \cap \mathcal{F}) = 0.$$

Dlatego $E = \mathbb{Z}_{1,p}$, i w konsekwencji $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$.
Na odwrót, załóżmy, że $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$. Stąd $E = \mathbb{Z}_{1,p}$. Stosując ponownie twierdzenie 4.3 widzimy, że zachodzą równości (4.11) i (4.12), skąd

$$\dim(\Theta \cap \mathcal{F}) = \dim(\mathcal{F}_0) = \dim(\mathcal{F}) - \overline{\overline{E}} \leqslant p - p = 0.$$

Dlatego $\dim(\Theta \cap \mathcal{F}) = 0$, i tym samym $\dim(\mathcal{F}) = \overline{\overline{E}} = p$. W konsekwencji $\Theta \cap \mathcal{F} = \{\theta\}$ i ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ jest liniowo niezależny, co kończy dowód. □

Następujący lemat pokazuje, że klasa regresji jest niepusta.

Lemat 4.6. *Dla dowolnych* $p \in \mathbb{N}$*, ciągu* $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ *i* $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$, $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$*, gdzie* $\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$.

Dowód. Ustalając dowolnie $p \in \mathbb{N}$ i ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ widzimy, że $\Theta + \mathcal{F}$ jest rozmaitością liniową w $\mathrm{H}(\mathfrak{P})$. Wykażemy, że $\Theta + \mathcal{F}$ jest zbiorem domkniętym w $\mathrm{H}(\mathfrak{P})$. Z uwagi 4.1 wynika, że

$$\mathcal{F} = \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p}\}), \tag{4.14}$$

gdzie $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ jest ciągiem określonym wzorem (4.6). Zakładając, że $E = \emptyset$ mamy $\|h'_k\| = 0$ dla $k \in \mathbb{Z}_{1,p}$. Stąd $\mathcal{F} \subset \Theta$, i tym samym $\Theta + \mathcal{F} = \Theta$. Dlatego $\Theta + \mathcal{F}$ jest zbiorem domkniętym w $\mathrm{H}(\mathfrak{P})$, gdyż zbiór Θ taki jest.

Możemy więc ograniczyć się do przypadku, gdy $E \neq \emptyset$. Ustalmy dowolnie $f \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ oraz ciąg $\mathbb{N} \ni n \mapsto f_n \in \Theta + \mathcal{F}$ taki, że

$$\|f_n - f\| \to 0, \quad \text{gdy } n \to +\infty. \tag{4.15}$$

Rozważmy ciąg

$$\mathbb{N} \ni n \mapsto \tilde{f}_n := \sum_{k=1}^{p} \langle f_n | h'_k \rangle_* \cdot h'_k. \tag{4.16}$$

Przyjmując

$$\tilde{f} := \sum_{k=1}^{p} \langle f | h'_k \rangle_* \cdot h'_k \tag{4.17}$$

widzimy, że $\tilde{f} \in \mathcal{F}$ oraz na mocy (4.7) i (3.14),

$$\begin{aligned}
\|\tilde{f}_n - \tilde{f}\| &= \left\| \sum_{k=1}^{p} \langle f_n | h'_k \rangle_* \cdot h'_k - \sum_{k=1}^{p} \langle f | h'_k \rangle_* \cdot h'_k \right\| \\
&= \left\| \sum_{k \in E} \langle f_n | h'_k \rangle_* \cdot h'_k - \sum_{k \in E} \langle f | h'_k \rangle_* \cdot h'_k \right\| \\
&= \left\| \sum_{k \in E} \langle f_n - f | h'_k \rangle_* \cdot h'_k \right\| \leqslant \sum_{k \in E} |\langle f_n - f | h'_k \rangle_*| \cdot \|h'_k\| \\
&\leqslant \sum_{k \in E} \frac{\|f_n - f\| \cdot \|h'_k\|}{\|h'_k\|^2} \|h'_k\| \leqslant \sum_{k \in E} \|f_n - f\| \leqslant p \|f_n - f\|, \quad n \in \mathbb{N}.
\end{aligned}$$

Stąd i z (4.15) wynika, że

$$\|\tilde{f}_n - \tilde{f}\| \to 0, \quad \text{gdy } n \to +\infty. \tag{4.18}$$

Ponieważ $f_n \in \Theta + \mathcal{F}$ dla $n \in \mathbb{N}$, więc istnieją ciągi $\mathbb{N} \ni n \mapsto u_n \in \Theta$ i $\mathbb{N} \ni n \mapsto v_n \in \mathcal{F}$ takie, że $f_n = u_n + v_n$ dla $n \in \mathbb{N}$. Stąd na mocy wzoru (4.16),

$$\begin{aligned}
\tilde{f}_n &= \sum_{k=1}^{p} \langle f_n | h'_k \rangle_* \cdot h'_k = \sum_{k \in E} \langle f_n | h'_k \rangle_* \cdot h'_k = \sum_{k \in E} \langle u_n + v_n | h'_k \rangle_* \cdot h'_k \\
&= \sum_{k \in E} \langle u_n | h'_k \rangle_* \cdot h'_k + \sum_{k \in E} \langle v_n | h'_k \rangle_* \cdot h'_k,
\end{aligned}$$

i wobec nierówności (3.14),

$$\tilde{f}_n - \sum_{k \in E} \langle v_n | h'_k \rangle_* \cdot h'_k = \sum_{k \in E} \langle u_n | h'_k \rangle_* \cdot h'_k = \theta. \tag{4.19}$$

Ponieważ $v_n \in \mathcal{F}$ dla $n \in \mathbb{N}$, więc na mocy (4.14) dla dowolnie ustalonego $n \in \mathbb{N}$ istnieje ciąg $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$ taki, że $v_n = \sum_{k=1}^{p} \lambda_k h'_k$. Wówczas na mocy uwagi 4.1 wnioskujemy, że dla każdego $l \in \mathbb{Z}_{1,p}$,

$$\langle v_n | h'_l \rangle = \left\langle \sum_{k=1}^{p} \lambda_k h'_k \middle| h'_l \right\rangle = \sum_{k=1}^{p} \lambda_k \langle h'_k | h'_l \rangle = \lambda_l \|h'_l\|^2,$$

i w konsekwencji $\lambda_l = \langle v_n|h'_l\rangle_*$ dla $l \in E$. Jeśli więc $E = \mathbb{Z}_{1,p}$, to

$$v_n - \sum_{k\in E}\langle v_n|h'_k\rangle_* \cdot h'_k = v_n - \sum_{k=1}^{p}\lambda_k h'_k = \theta \in \Theta.$$

W przeciwnym przypadku $E_0 \neq \emptyset$ i

$$v_n - \sum_{k\in E}\langle v_n|h'_k\rangle_* \cdot h'_k = \sum_{k=1}^{p}\lambda_k h'_k - \sum_{k\in E}\langle v_n|h'_k\rangle_* \cdot h'_k = \sum_{k\in E_0}\lambda_k h'_k \in \Theta.$$

To łącznie z (4.19) daje $\tilde{f}_n - v_n \in \Theta$, skąd

$$\tilde{f}_n - f_n = \tilde{f}_n - (u_n + v_n) = (\tilde{f}_n - v_n) - u_n \in \Theta.$$

Dlatego $\|\tilde{f}_n - f_n\| = 0$ dla $n \in \mathbb{N}$. To wobec (4.15) i (4.18) daje

$$\begin{aligned}\|\tilde{f} - f\| &= \|\tilde{f} - \tilde{f}_n + \tilde{f}_n - f_n + f_n - f\| \leqslant \|\tilde{f} - \tilde{f}_n\| + \|\tilde{f}_n - f_n\| + \|f_n - f\| \\ &= \|\tilde{f} - \tilde{f}_n\| + \|f_n - f\| \to 0, \quad \text{gdy } n \to +\infty.\end{aligned}$$

Stąd $\tilde{f} - f \in \Theta$. Ponieważ $\tilde{f} \in \mathcal{F}$, więc $f = f - \tilde{f} + \tilde{f} \in \Theta + \mathcal{F}$. W ten sposób wykazaliśmy, że $\Theta + \mathcal{F}$ jest zbiorem domkniętym w przestrzeni $\mathrm{H}(\mathfrak{P})$. Zatem $\Theta + \mathcal{F}$ jest domkniętą rozmaitością liniową w przestrzeni $\mathrm{H}(\mathfrak{P})$, skąd na mocy twierdzenia 3.10 istnieje $f \in \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g)$. Ponieważ klasa $\mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g) \subset \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, więc korzystając z poprzedniej części dowodu widzimy, że funkcja $\tilde{f}$ określona wzorem (4.17) spełnia warunki

$$\tilde{f} \in \mathcal{F} \quad \text{oraz} \quad \|\tilde{f} - f\| = 0. \tag{4.20}$$

Z twierdzenia 3.8 wynika, że

$$\langle h|f\rangle = g^*(h), \quad h \in \Theta + \mathcal{F}. \tag{4.21}$$

Stosując nierówność (3.14) wnioskujemy z (4.20) i (4.21), że dla każdego $h \in \mathcal{F}$,

$$\begin{aligned}|\langle h|\tilde{f}\rangle - g^*(h)| &= |\langle h|f + (\tilde{f} - f)\rangle - g^*(h)| = |\langle h|f\rangle - g^*(h) + \langle h|\tilde{f} - f\rangle| \\ &= |\langle h|\tilde{f} - f\rangle| \leqslant \|h\| \cdot \|\tilde{f} - f\| = 0.\end{aligned}$$

Stąd $\langle h|\tilde{f}\rangle = g^*(h)$ dla $h \in \mathcal{F}$. Ponieważ $\tilde{f} \in \mathcal{F}$, więc korzystając ponownie z twierdzenia 3.8 stwierdzamy, że $\tilde{f} \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Dlatego $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$, czego należało dowieść. □

Twierdzenie 4.7. *Dla dowolnie zadanych $p \in \mathbb{N}$ i $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ niech $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ będzie ciągiem zdefiniowanym wzorem* (4.6)*. Wtedy dla każdego $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest jednoelementowa i $\dim(\mathcal{F}) = p$ wtedy i tylko wtedy, gdy $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$, gdzie $\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$.*

Dowód. Ustalmy dowolnie $p \in \mathbb{N}$, ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ i funkcję g spełniającą założenia twierdzenia. Załóżmy, że $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest klasą jednoelementową i $\dim(\mathcal{F}) = p$. Wtedy $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$, skąd na mocy wniosku 3.16, $\Theta \cap \mathcal{F} = \{\theta\}$. Ponieważ $\dim(\mathcal{F}) = p$, więc ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ jest liniowo niezależny, skąd na mocy wniosku 4.5, $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$. Na odwrót, załóżmy, że $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$. Stąd na mocy wniosku 4.5, $\Theta \cap \mathcal{F} = \{\theta\}$ i ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ jest liniowo niezależny, przez co $\dim(\mathcal{F}) = p$. Z lematu 4.6 wynika, że $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$. Korzystając z wniosku 3.16 widzimy, że klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest jednoelementowa, co kończy dowód. □

Uwaga 4.8. Przy założeniach twierdzenia 4.7 klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest jednoelementowa wtedy i tylko wtedy, gdy $\Theta \cap \mathcal{F} = \{\theta\}$. Równoważność ta jest bezpośrednią konsekwencją wniosku 3.16, gdyż na mocy lematu 4.6, $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$.

Uwaga 4.9. Przy założeniach twierdzenia 4.7 klasa $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest rozmaitością afiniczną. Z lematu 4.6 wynika bowiem, że istnieje $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Stąd na mocy twierdzenia 3.15, $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta \cap \mathcal{F} + f_0$. Dlatego $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ jest rozmaitością afiniczną jako przesunięcie rozmaitości liniowej $\Theta \cap \mathcal{F}$ o wektor f_0.

4.2 Procedura wyznaczania funkcji regresji

Dla zadanych $p, q \in \mathbb{Z}$, $p \leqslant q$ i ciągu $\mathbb{Z}_{p,q} \ni k \mapsto \mathcal{F}_k$ niepustych zbiorów przestrzeni $\mathrm{H}(\mathfrak{P})$ oznaczmy przez $\sum_{k=p}^{q} \mathcal{F}_k$ klasę wszystkich funkcji postaci $\sum_{k=p}^{q} u_k$ gdzie $\mathbb{Z}_{p,q} \ni k \mapsto u_k \in \mathcal{F}_k$. Oczywiście $\sum_{k=1}^{2} \mathcal{F}_k = \mathcal{F}_1 + \mathcal{F}_2$.

Następujące twierdzenie dotyczy wyznaczania funkcji regresji, w przypadku gdy model teoretyczny $\mathcal{F}$ jest rozmaitością liniową generowaną przez dowolnie zadany ciąg funkcji $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ w przestrzeni $\mathrm{H}(\mathfrak{P})$, gdzie $p \in \mathbb{N}$.

Twierdzenie 4.10. *Dla dowolnie zadanych $p \in \mathbb{N}$ i $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ niech $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ będzie ciągiem zdefiniowanym wzorami* (4.6)*. Jeśli $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i*

$$\|h'_k\| > 0, \qquad k \in \mathbb{Z}_{1,p}, \tag{4.22}$$

to

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \sum_{k=1}^{p} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \right\}, \tag{4.23}$$

gdzie $\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$.

Dowód. Ustalmy dowolnie $p \in \mathbb{N}$, ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ i $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ spełniające założenia twierdzenia. Z twierdzenia 4.7 wynika, że dla pewnego $f \in \mathcal{F}$,

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{f\}. \tag{4.24}$$

Na mocy uwagi 4.1, $\mathcal{F} = \operatorname{lin}(\{h'_k\colon k \in \mathbb{Z}_{1,p}\})$. Istnieje więc ciąg $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$ o tej własności, że

$$f = \sum_{k=1}^{p} \lambda_k \cdot h'_k. \tag{4.25}$$

Korzystając z twierdzenia 3.8 stwierdzamy, że dla każdego $l \in \mathbb{Z}_{1,p}$,

$$\overline{g^*(h'_l)} = \langle f|h'_l\rangle = \left\langle \sum_{k=1}^{p} \lambda_k \cdot h'_k \middle| h'_l \right\rangle = \sum_{k=1}^{p} \lambda_k \langle h'_k|h'_l\rangle = \lambda_l \|h'_l\|^2,$$

gdyż na mocy uwagi 4.1, $h'_k \perp h'_l$ dla $k \in \mathbb{Z}_{1,p}\backslash\{l\}$. Stąd wobec (4.22),

$$\lambda_l = \frac{\overline{g^*(h'_l)}}{\|h'_l\|^2}, \qquad l \in \mathbb{Z}_{1,p}.$$

To łącznie z równościami (4.24) i (4.25) daje równość (4.23), czego należało dowieść. □

Z wniosku 4.5 wynika, że warunek (4.22) zachodzi wtedy i tylko wtedy, gdy ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ jest liniowo niezależny i $\mathcal{F} \cap \Theta = \{\theta\}$. W szczególności nierówności (4.22) zachodzą pod warunkiem, że ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ jest liniowo niezależny i funkcjonał $\|\cdot\|$ jest normą w $(\mathrm{L}^2(A, \mathcal{A}_x, P_x); +, \cdot)$. Warunek (4.22) gwarantuje nam, że klasa regresji jest jednoelementowa; por. [33, Remark 3.3]. Poniżej omówimy przypadek ogólny, gdy warunek (4.22) nie musi być spełniony. Zachodzi wówczas następujące twierdzenie, które opisuje ogólne rozwiązanie problemu regresji.

Twierdzenie 4.11. *Dla dowolnie zadanych $p \in \mathbb{N}$ i ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ niech $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ będzie ciągiem zdefiniowanym wzorem* (4.6). *Wtedy dla każdego $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ zachodzi równość*

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \operatorname{lin}(\{h'_k\colon k \in \mathbb{Z}_{1,p}\backslash E\}) + \sum_{k\in E} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \tag{4.26}$$

z $\mathcal{F} := \operatorname{lin}(\{h_k\colon k \in \mathbb{Z}_{1,p}\})$ i $E := \{k \in \mathbb{Z}_{1,p}\colon \|h'_k\| > 0\}$, gdy $E \neq \emptyset$; ponadto $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$, *gdy $E = \emptyset$.*

Dowód. Przy założeniach i oznaczeniach twierdzenia rozważmy najpierw przypadek, gdy $E = \emptyset$. Wtedy $\|h'_k\| = 0$ dla $k \in \mathbb{Z}_{1,p}$, skąd na mocy uwagi 4.1, $\mathcal{F} \subset \Theta$. Korzystając z wniosku 3.9 dostajemy równość $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$. Załóżmy teraz, że $E = \mathbb{Z}_{1,p}$. Wtedy $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$, skąd na mocy twierdzenia 4.10 zachodzi równość (4.23). Ponadto $\operatorname{lin}(\{h'_k\colon k \in \mathbb{Z}_{1,p}\backslash E\}) = \operatorname{lin}(\emptyset) = \{\theta\}$, co łącznie z równością (4.23) implikuje równość (4.26). Pozostaje zatem rozważyć przypadek, gdy $\emptyset \neq E \neq \mathbb{Z}_{1,p}$. Wzorując się równością (4.23) rozważmy funkcję

$$f_0 := \sum_{k\in E} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k. \tag{4.27}$$

Wykażemy, że $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P})$. Przyjmując $E_0 := \mathbb{Z}_{1,p} \setminus E$,

$$\mathcal{F}' := \mathrm{lin}(\{h'_k \colon k \in E\}) \qquad \text{i} \qquad \mathcal{F}_0 := \mathrm{lin}(\{h'_k \colon k \in E_0\})$$

widzimy, że $\mathcal{F} = \mathcal{F}' + \mathcal{F}_0$; por. dowód twierdzenia 4.3. Stąd dla dowolnie zadanego $h \in \mathcal{F}$ zachodzi rozkład $h = h' + h''$ dla pewnych $h' \in \mathcal{F}'$ i $h'' \in \mathcal{F}_0$. Istnieje zatem ciąg $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in \mathbb{C}$ taki, że

$$h' = \sum_{k \in E} \lambda_k h'_k \qquad \text{i} \qquad h'' = \sum_{k \in E_0} \lambda_k h'_k. \tag{4.28}$$

Z równości (4.27) i (4.28) stwierdzamy na mocy uwagi 4.1, że

$$\begin{aligned}
\langle h'|f_0\rangle &= \Big\langle \sum_{l \in E} \lambda_l h'_l \Big| \sum_{k \in E} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \Big\rangle = \sum_{l \in E} \sum_{k \in E} \Big\langle \lambda_l h'_l \Big| \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \Big\rangle \\
&= \sum_{l \in E} \sum_{k \in E} \lambda_l \frac{g^*(h'_k)}{\|h'_k\|^2} \langle h'_l | h'_k \rangle = \sum_{k \in E} \lambda_k \frac{g^*(h'_k)}{\|h'_k\|^2} \|h'_k\|^2 \\
&= \sum_{k \in E} \lambda_k g^*(h'_k) = g^*\Big(\sum_{k \in E} \lambda_k h'_k \Big) = g^*(h')
\end{aligned}$$

oraz

$$\begin{aligned}
\langle h''|f_0\rangle &= \Big\langle \sum_{l \in E_0} \lambda_l h'_l \Big| \sum_{k \in E} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \Big\rangle = \sum_{l \in E_0} \sum_{k \in E} \Big\langle \lambda_l h'_l \Big| \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \Big\rangle \\
&= \sum_{l \in E_0} \sum_{k \in E} \lambda_l \frac{g^*(h'_k)}{\|h'_k\|^2} \langle h'_l | h'_k \rangle = 0.
\end{aligned}$$

Dlatego

$$\langle h|f_0\rangle = \langle h' + h''|f_0\rangle = \langle h'|f_0\rangle + \langle h''|f_0\rangle = g^*(h'). \tag{4.29}$$

Z nierówności (3.20) dostajemy ponadto

$$\begin{aligned}
|g^*(h'')| &= \Big| g^*\Big(\sum_{k \in E_0} \lambda_k h'_k \Big) \Big| = \Big| \sum_{k \in E_0} \lambda_k g^*(h'_k) \Big| \leqslant \sum_{k \in E_0} |\lambda_k| |g^*(h'_k)| \\
&\leqslant \sum_{k \in E_0} |\lambda_k| M_g \|h'_k\| = 0.
\end{aligned}$$

To łącznie z (4.29) daje

$$\langle h|f_0\rangle = g^*(h') + g^*(h'') = g^*(h' + h'') = g^*(h).$$

Stąd $\langle h|f_0\rangle = g^*(h)$ dla $h \in \mathcal{F}$, i wobec twierdzenia 3.8, $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Stosując twierdzenie 3.15 otrzymujemy równość $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = (\Theta \cap \mathcal{F}) + f_0$. Z twierdzenia 4.3 wynika zaś, że $(\Theta \cap \mathcal{F}) = \mathcal{F}_0$. Zatem $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}_0 + f_0$, czyli zachodzi równość (4.26). □

Z praktycznego punktu widzenia do wyznaczenia funkcji regresji wygodnie będzie operować funkcjonałem $\mathrm{E}(\cdot|\cdot)$ określonym wzorem

(4.30)
$$\mathrm{L}^2(\Omega,\mathcal{A},P)\times\mathrm{L}^2(\Omega,\mathcal{A},P)\ni(u,v)\mapsto\mathrm{E}(u|v):=\begin{cases}\dfrac{\mathrm{E}(u\overline{v})}{\mathrm{E}(|v|^2)}, & \text{gdy } \mathrm{E}(|v|^2)>0\\ 0, & \text{gdy } \mathrm{E}(|v|^2)=0.\end{cases}$$

Korzystając ze wzoru (4.30) oraz zależności (4.3) i (4.4) otrzymujemy pierwszą równość w (4.31),

$$\mathrm{E}(f\circ x|h\circ x)=\frac{\langle f|h\rangle}{\|h\|^2}=\frac{\displaystyle\int_A f(t)\overline{h(t)}\mathrm{d}P_x(t)}{\displaystyle\int_A|h(t)|^2\mathrm{d}P_x(t)},\tag{4.31}$$
$$f\in\mathrm{L}^2(A,\mathcal{A}_x,P_x),\ h\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)\backslash\Theta.$$

Druga równość w (4.31) wynika bezpośrednio ze wzorów (3.10) i (3.12).
Oznaczmy przez $P_{x,y}$ miarę probabilistyczną na σ-ciele $\mathcal{A}_x\otimes\mathcal{A}_y$ spełniającą warunek

$$P_{x,y}(I\times J)=P\Big(x^{-1}(I)\cap y^{-1}(J)\Big),\qquad I\in\mathcal{A}_x,\quad J\in\mathcal{A}_y,\tag{4.32}$$

gdzie $\mathcal{A}_x\otimes\mathcal{A}_y$ jest najmniejszym, w sensie inkluzji, σ-ciałem zawierającym wszystkie zbiory $I\times J$ dla $I\in\mathcal{A}_x$ i $J\in\mathcal{A}_y$. Istnienie i jednoznaczność miary $P_{x,y}$ jest konsekwencją [20, §C.3]. Porównując warunek (4.32) z warunkiem (2.8) widzimy, że $P_{x,y}=\pi_{x,y}$. Stosując lemat 2.4 z $\Omega':=A$, $\Omega'':=\mathbb{C}$, $\mathcal{B}':=\mathcal{A}_x$, $\mathcal{B}''=\mathcal{A}_y$ i $\mu:=P$ otrzymujemy dla dowolnie zadanych $h\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)$ i $g\in\mathrm{L}^2(\mathbb{C},\mathcal{A}_y,P_y)$ równość

$$\int_\Omega(h\circ x(\omega))(g\circ y(\omega))\mathrm{d}P(\omega)=\int_{A\times\mathbb{C}}h(t_1)g(t_2)\mathrm{d}P_{x,y}(t_1,t_2).$$

To łącznie z pierwszą równością w (4.5) daje

$$g^*(h)=\int_{A\times\mathbb{C}}h(t_1)\overline{g(t_2)}\mathrm{d}P_{x,y}(t_1,t_2).$$

Stąd na podstawie wzoru (3.12) dostajemy drugą równość w (4.33),

$$\mathrm{E}(g\circ y|h\circ x)=\frac{\overline{g^*(h)}}{\|h\|^2}=\frac{\displaystyle\int_{A\times\mathbb{C}}\overline{h(t_1)}g(t_2)\mathrm{d}P_{x,y}(t_1,t_2)}{\displaystyle\int_A|h(t)|^2\mathrm{d}P_x(t)},\tag{4.33}$$
$$g\in\mathrm{L}^2(\mathbb{C},\mathcal{A}_y,P_y),\ h\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)\backslash\Theta.$$

Pierwsza równość w (4.33) wynika bezpośrednio ze wzoru (4.30) oraz zależności (4.3) i (4.5). Stosując powyższe formuły można, w ślad za [25, Corollary 9.2], przeformułować twierdzenie 4.11 opisujące ogólne rozwiązanie problemu regresji jak następuje.

Twierdzenie 4.12. *Dla dowolnie zadanych* $p \in \mathbb{N}$ *i ciągu* $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ *niech* $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ *będzie ciągiem zdefiniowanym wzorem*

$$h'_1 := h_1 \quad i \quad h'_n := h_n - \sum_{k=1}^{n-1} \mathrm{E}(h_n \circ x | h'_k \circ x) h'_k, \qquad n \in \mathbb{Z}_{2,p}. \tag{4.34}$$

Wtedy dla każdego $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ *zachodzi równość*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p} \setminus E\}) + \sum_{k \in E} \mathrm{E}(g \circ y | h'_k \circ x) h'_k \tag{4.35}$$

$z\ \mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$ *i* $E := \{k \in \mathbb{Z}_{1,p} \colon \|h'_k\| > 0\}$, *gdy* $E \neq \emptyset$; *ponadto* $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$, *gdy* $E = \emptyset$.

Uwaga 4.13. Z uwagi 3.21 wynika, że w przypadku rzeczywistej funkcji skalującej g i klasy $\mathcal{F}$ rzeczywistych funkcji przebiegów teoretycznych zachodzi równość (3.51). Zatem wyznaczanie klasy regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ można sprowadzić do wyznaczenia klasy $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_{\mathbb{C},g})$ względem zespolonej probabilistycznej struktury regresji $\mathfrak{P}_{\mathbb{C},g}$. W szczególności zachodzą twierdzenia 4.11 i 4.12 dla ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \cap (A \to \mathbb{R})$ i funkcji skalującej $g \in \mathrm{L}^2(\mathbb{R}, \mathcal{A}_y, P_y) \cap (\mathbb{R} \to \mathbb{R})$. Wtedy równość (4.26) przyjmuje postać

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p} \setminus E\}) + \sum_{k \in E} \frac{g^*(h'_k)}{\|h'_k\|^2} h'_k,$$

natomiast równość (4.35) pozostaje w niezmienionej postaci. Zauważmy, że na mocy wzorów (4.6) i (4.7), $h'_k \colon A \to \mathbb{R}$ dla $k \in \mathbb{Z}_{1,p}$.

W kolejnych podrozdziałach przedstawimy rozwiązania problemu regresji dla szczególnych przypadków modelu teoretycznego $\mathcal{F}$.

4.3 Przypadek jednowymiarowy

Rozważmy przypadek, w którym model funkcyjny $\mathcal{F}$ rozpięty jest przez jedną dowolnie ustaloną funkcję $h_1 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, tzn., $\mathcal{F} = \mathrm{lin}(\{h_1\})$. Stosując twierdzenie 4.12 widzimy, że $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathrm{lin}(\{h_1\})$, gdy $\|h_1\| = 0$. Jeśli zaś $\|h_1\| > 0$, to

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Big\{ \mathrm{E}(g \circ y | h_1 \circ x) \cdot h_1 \Big\}, \tag{4.36}$$

czyli, na podstawie (4.30),

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \frac{\mathrm{E}((g \circ y)(\overline{h_1 \circ x}))}{\mathrm{E}(|h_1 \circ x|^2)} \cdot h_1 \right\}. \tag{4.37}$$

Ustalając dowolnie $\alpha \in \mathbb{Z}_0$ załóżmy, że $A \subset \mathbb{C}$, $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i $h_1 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, gdzie $h_1(t) := t^\alpha$ dla $t \in A$ i $g(t) := t$ dla $t \in \mathbb{C}$. Wtedy

$$\mathrm{E}((g \circ y)(\overline{h_1 \circ x})) = \mathrm{E}(\overline{x}^\alpha \cdot y) \quad \text{i} \quad \mathrm{E}(|h_1 \circ x|^2) = \mathrm{E}(|x|^{2\alpha}),$$

skąd

$$\mathrm{E}(g \circ y | h_1 \circ x) = \frac{\mathrm{E}(\overline{x}^\alpha \cdot y)}{\mathrm{E}(|x|^{2\alpha})}.$$

Z równości (4.36) otrzymujemy więc równość

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}) = \left\{ A \ni t \mapsto \frac{\mathrm{E}(\overline{x}^\alpha \cdot y)}{\mathrm{E}(|x|^{2\alpha})} \cdot t^\alpha \right\}, \tag{4.38}$$

o ile funkcja x nie jest równa 0 p.n. w Ω.
Jeśli x jest zmienną losową rzeczywistą, to przyjmując za $\alpha := 1$ w (4.38) stwierdzamy, że $\mathrm{E}(xy)$ może być wyrażone za pomocą funkcji z klasy regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$ i $\mathrm{E}(x^2)$. Kładąc $\alpha := 0$ w (4.38) otrzymujemy

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}) = \big\{ A \ni t \mapsto \mathrm{E}(y) \big\}. \tag{4.39}$$

Zauważmy, że równość (4.39) zachodzi nawet jeśli A jest dowolnym zbiorem niepustym i $h_1(t) = 1$ dla $t \in A$.
Załóżmy teraz, że $h_1 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, gdzie $A \ni t \mapsto h_1(t) := at + b$ dla dowolnie zadanych $a, b \in \mathbb{C}$. Wtedy

$$\begin{aligned} \mathrm{E}(y(\overline{h_1 \circ x})] &= \mathrm{E}(y(\bar{a}\bar{x} + \bar{b})) = \bar{a}\,\mathrm{E}(\bar{x} \cdot y) + \bar{b}\,\mathrm{E}(y), \\ \mathrm{E}(|h_1 \circ x|^2) &= \mathrm{E}(|ax + b|^2) = |a|^2\,\mathrm{E}(|x|^2) + |b|^2 + 2\mathrm{Re}(a\bar{b}\,\mathrm{E}(x)). \end{aligned}$$

Stosując równość (4.37) dostajemy

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}) = \left\{ A \ni t \mapsto \frac{\bar{a}\,\mathrm{E}(\overline{x} \cdot y) + \bar{b}\,\mathrm{E}(y)}{|a|^2\,\mathrm{E}(|x|^2) + |b|^2 + 2\mathrm{Re}(a\bar{b}\,\mathrm{E}(x))} \cdot (at + b) \right\},$$

o ile funkcja h_1 nie jest równa 0 p.n. w Ω.

4.4 Przypadek dwuwymiarowy

Załóżmy, że model funkcyjny $\mathcal{F}$ jest dwuwymiarową rozmaitością liniową rozpiętą przez dwie dowolnie ustalone funkcje $h_1, h_2 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, tzn., $\mathcal{F} = \mathrm{lin}(\{h_1, h_2\})$. Mogą zajść trzy przypadki ze względu na $\overline{\overline{E}} \in \mathbb{Z}_{0,2}$, gdzie $E := \{k \in \mathbb{Z}_{1,2} \colon \|h_k'\| > 0\}$; por. (4.10).

Przypadek I, gdy $\overline{\overline{E}} = 2$. Wtedy $\|h'_1\| > 0$ i $\|h'_2\| > 0$, gdzie $\mathbb{Z}_{1,2} \ni k \mapsto h'_k$ jest ciągiem zdefiniowanym wzorem (4.6). Stosując twierdzenie 4.11 otrzymujemy

$$\begin{aligned}\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) &= \operatorname{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,2} \setminus E\}) + \sum_{k \in E} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \\ &= \left\{ \frac{\overline{g^*(h'_1)}}{\|h'_1\|^2} h'_1 + \frac{\overline{g^*(h'_2)}}{\|h'_2\|^2} h'_2 \right\}\end{aligned} \tag{4.40}$$

gdzie, zgodnie z (4.6),

$$h'_1 := h_1 \quad \text{i} \quad h'_2 := h_2 - \langle h_2 | h_1 \rangle_* \cdot h_1. \tag{4.41}$$

Stąd $h'_2 \perp h_1$, i w konsekwencji

$$\begin{aligned}\|h'_2\|^2 = \langle h'_2 | h'_2 \rangle &= \langle h'_2 | h_2 - \langle h_2 | h_1 \rangle_* h_1 \rangle = \langle h'_2 | h_2 \rangle \\ &= \langle h_2 - \langle h_2 | h_1 \rangle_* h_1 | h_2 \rangle = \|h_2\|^2 - \langle h_2 | h_1 \rangle_* \langle h_1 | h_2 \rangle\end{aligned} \tag{4.42}$$

oraz

$$g^*(h'_2) = g^*(h_2 - \langle h_2 | h_1 \rangle_* h_1) = g^*(h_2) - \langle h_2 | h_1 \rangle_* g^*(h_1).$$

Zatem z twierdzenia 4.11 oraz (4.41) i (4.42) wynika, że

$$\begin{aligned}\sum_{k \in E} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k &= \frac{\overline{g^*(h'_1)}}{\|h'_1\|^2} h'_1 + \frac{\overline{g^*(h'_2)}}{\|h'_2\|^2} h'_2 \\ &= \frac{\overline{g^*(h_1)}}{\|h_1\|^2} h_1 + \frac{\overline{g^*(h_2)} - \overline{\langle h_2 | h_1 \rangle}_* \cdot \overline{g^*(h_1)}}{\|h_2\|^2 - \langle h_2 | h_1 \rangle_* \langle h_1 | h_2 \rangle} \Big(h_2 - \langle h_2 | h_1 \rangle_* h_1 \Big).\end{aligned}$$

Kładąc

$$\alpha_2 := \frac{\overline{g^*(h_2)} \|h_1\|^2 - \overline{g^*(h_1) \langle h_2 | h_1 \rangle}}{\|h_2\|^2 \|h_1\|^2 - |\langle h_2 | h_1 \rangle|^2} \quad \text{i} \quad \alpha_1 := \frac{\overline{g^*(h_1)} - \langle h_2 | h_1 \rangle \alpha_2}{\|h_1\|^2} \tag{4.43}$$

wnioskujemy z (4.40), że

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\alpha_1 h_1 + \alpha_2 h_2\}. \tag{4.44}$$

Korzystając z równości (4.31) i (4.33) możemy stałe α_2 i α_1 wyrazić także w następującej postaci

$$\begin{aligned}\alpha_2 &= \frac{\frac{\overline{g^*(h_2)}}{\|h_2\|^2} - \frac{\overline{g^*(h_1)}}{\|h_1\|^2} \cdot \frac{\overline{\langle h_2 | h_1 \rangle}}{\|h_2\|^2}}{1 - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} \cdot \frac{\langle h_1 | h_2 \rangle}{\|h_2\|^2}} = \frac{\frac{\overline{g^*(h_2)}}{\|h_2\|^2} - \frac{\overline{g^*(h_1)}}{\|h_1\|^2} \cdot \langle h_1 | h_2 \rangle_*}{1 - \langle h_2 | h_1 \rangle_* \cdot \langle h_1 | h_2 \rangle_*} \\ &= \frac{\mathrm{E}(g \circ y | h_2 \circ x) - \mathrm{E}(g \circ y | h_1 \circ x) \, \mathrm{E}(h_1 \circ x | h_2 \circ x)}{1 - \mathrm{E}(h_2 \circ x | h_1 \circ x) \, \mathrm{E}(h_1 \circ x | h_2 \circ x)}\end{aligned} \tag{4.45}$$

oraz

(4.46)

$$\begin{aligned}\alpha_1 &= \frac{\overline{g^*(h_1)}}{\|h_1\|^2} - \frac{\langle h_2|h_1\rangle}{\|h_1\|^2}\alpha_2 = \frac{\overline{g^*(h_1)}}{\|h_1\|^2} - \frac{\langle h_2|h_1\rangle}{\|h_1\|^2}\cdot\frac{\overline{g^*(h_2)}\|h_1\|^2 - \overline{g^*(h_1)\langle h_2|h_1\rangle}}{\|h_2\|^2\|h_1\|^2 - |\langle h_2|h_1\rangle|^2}\\ &= \frac{\frac{\overline{g^*(h_1)}}{\|h_1\|^2}\Big(\|h_2\|^2\|h_1\|^2 - \langle h_2|h_1\rangle\langle h_1|h_2\rangle\Big) - \overline{g^*(h_2)}\langle h_2|h_1\rangle + \frac{\overline{g^*(h_1)}}{\|h_1\|^2}\langle h_1|h_2\rangle\langle h_2|h_1\rangle}{\|h_2\|^2\|h_1\|^2 - |\langle h_2|h_1\rangle|^2}\\ &= \frac{\overline{g^*(h_1)}\|h_2\|^2 - \overline{g^*(h_2)}\langle h_2|h_1\rangle}{\|h_2\|^2\|h_1\|^2 - \langle h_2|h_1\rangle\langle h_1|h_2\rangle} = \frac{\frac{\overline{g^*(h_1)}}{\|h_1\|^2} - \frac{\overline{g^*(h_2)}}{\|h_2\|^2}\frac{\langle h_2|h_1\rangle}{\|h_1\|^2}}{1 - \frac{\langle h_2|h_1\rangle}{\|h_1\|^2}\frac{\langle h_1|h_2\rangle}{\|h_2\|^2}}\\ &= \frac{\mathrm{E}(g\circ y|h_1\circ x) - \mathrm{E}(g\circ y|h_2\circ x)\,\mathrm{E}(h_2\circ x|h_1\circ x)}{1 - \mathrm{E}(h_2\circ x|h_1\circ x)\,\mathrm{E}(h_1\circ x|h_2\circ x)}.\end{aligned}$$

W szczególności przypuśćmy, że $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i $h_1, h_2 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, gdzie $g(t) := t$ dla $t \in \mathbb{C}$ oraz $h_1(t) := 1$ i $h_2(t) := t$ dla $t \in A$. Stosując równości (4.3), (4.4) oraz (4.5) możemy przepisać formuły (4.43) na współczynniki α_2 i α_1, w następujący sposób

$$\alpha_2 = \frac{\mathrm{E}(\overline{x}\cdot y) - \mathrm{E}(\overline{x})\cdot\mathrm{E}(y)}{\mathrm{E}(|x|^2) - (\mathrm{E}(x))^2} \quad \text{i} \quad \alpha_1 = \mathrm{E}(y) - \mathrm{E}(x)\cdot\alpha_2 \tag{4.47}$$

o ile funkcja x nie jest stała p.n. w Ω. Wtedy, na podstawie (4.44), otrzymujemy

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{A \ni t \mapsto \alpha_1 + \alpha_2 t\}.$$

Współczynniki α_2 i α_1 zadane przez równości (4.47) pokrywają się ze współczynnikami klasycznej regresji liniowej w przypadku rzeczywistych zmiennych losowych; por. [6, §1.7], [11, §3.8], [31, §1.4].

Przypadek II, gdy $\overline{\overline{E}} = 1$. Wtedy $E = \{1\}$ lub $E = \{2\}$. Jeśli $E = \{2\}$ to $h_1' = h_1$ i $h_2' = h_2$. Jeśli $E = \{1\}$ to $h_1' = h_1$ i $h_2' = h_2 - \mathrm{E}(h_2\circ x|h_1\circ x)h_1$. Stosując twierdzenie 4.12 otrzymujemy

$$\begin{aligned}\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) &= \mathrm{lin}(\{h_k' \colon k \in \mathbb{Z}_{1,2}\backslash E\}) + \sum_{k\in E}\mathrm{E}(g\circ y|h_k'\circ x)h_k'\\ &= \begin{cases}\mathrm{lin}(\{h_2 - \mathrm{E}(h_2\circ x|h_1\circ x)h_1\}) + \mathrm{E}(g\circ y|h_1\circ x)h_1, & \text{gdy } E = \{1\}\\ \mathrm{lin}(\{h_1\}) + \mathrm{E}(g\circ y|h_2\circ x)h_2, & \text{gdy } E = \{2\}.\end{cases}\end{aligned}$$

W szczególności, jeśli $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i $h_1, h_2 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, gdzie $g(t) := t$ dla $t \in \mathbb{C}$ oraz $h_1(t) := 1$ i $h_2(t) := t$ dla $t \in A$, to

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \begin{cases}\Big\{A \ni t \mapsto \lambda(t - \mathrm{E}(x)) + \mathrm{E}(y) \colon \lambda \in \mathbb{C}\Big\}, & \text{gdy } E = \{1\}\\ \Big\{A \ni t \mapsto \lambda + \frac{\mathrm{E}(\overline{x}\cdot y)}{\mathrm{E}(|x|^2)}t \colon \lambda \in \mathbb{C}\Big\}, & \text{gdy } E = \{2\}.\end{cases}$$

Przypadek III, gdy $\overline{\overline{E}} = 0$. Wtedy stosując twierdzenie 4.12 otrzymujemy $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$.

Przykład 4.14. Nawiązując do przykładu 3.4 wyznaczania regresji dla obwodu elektrycznego, wyznaczymy klasę funkcji regresji $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Z warunku (3.6) wynika, że

$$P_{x,y}(\{t_1\} \times (-\infty, t_2)) = P(x = t_1 \wedge y < t_2) = \frac{1}{n} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t_2} e^{\frac{-(\tau - y_k)^2}{2\sigma^2}} \mathrm{d}\tau, \qquad t_1 \in \mathbb{Z}_{1,n},\ t_2 \in \mathbb{R}. \tag{4.48}$$

Aby wyznaczyć klasę funkcji regresji (4.44) możemy zastosować wzory (4.45) i (4.46) dla funkcji skalującej $\mathbb{C} \ni t \mapsto g(t) := t$, gdyż $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$.
Załóżmy najpierw, że $\Delta \neq 0$. Z równości (4.31) mamy

$$\mathrm{E}(h_2 \circ x | h_1 \circ x) = \frac{\int_A h_2(t)\overline{h_1(t)}\mathrm{d}P_x(t)}{\int_A |h_1(t)|^2 \mathrm{d}P_x(t)} = \frac{\sum\limits_{k=1}^{n} h_2(x_k)\overline{h_1(x_k)} \cdot P_x(\{x_k\})}{\sum\limits_{k=1}^{n} |h_1(x_k)|^2 \cdot P_x(\{x_k\})} = \frac{\sum\limits_{k=1}^{n} \mathrm{e}^{r_2 x_k} \cdot \mathrm{e}^{\overline{r_1 x_k}} \cdot \frac{1}{n}}{\sum\limits_{k=1}^{n} |\mathrm{e}^{r_1 x_k}|^2 \cdot \frac{1}{n}} = \frac{\sum\limits_{k=1}^{n} \mathrm{e}^{(r_2 + \overline{r_1}) x_k}}{\sum\limits_{k=1}^{n} \mathrm{e}^{2x_k \operatorname{Re} r_1}}$$

oraz podobnie

$$\mathrm{E}(h_1 \circ x | h_2 \circ x) = \frac{\sum\limits_{k=1}^{n} \mathrm{e}^{(r_1 + \overline{r_2}) x_k}}{\sum\limits_{k=1}^{n} \mathrm{e}^{2x_k \operatorname{Re} r_2}}.$$

Definiując

$$c := \sum_{k=1}^{n} \mathrm{e}^{(r_2 + \overline{r_1}) x_k}, \qquad c_1 := \sum_{k=1}^{n} \mathrm{e}^{2x_k \operatorname{Re} r_1}, \qquad c_2 := \sum_{k=1}^{n} \mathrm{e}^{2x_k \operatorname{Re} r_2}$$

otrzymujemy

$$\mathrm{E}(h_2 \circ x | h_1 \circ x) = \frac{c}{c_1}$$

oraz

$$\mathrm{E}(h_1 \circ x | h_2 \circ x) = \frac{\overline{c}}{c_2}.$$

Z równości (4.33) mamy zaś

$$\mathrm{E}(g \circ y | h_1 \circ x) = \frac{\int_{A \times \mathbb{C}} t_2 \cdot \overline{h_1(t_1)} \mathrm{d}P_{x,y}(t_1, t_2)}{\int_A |h_1(t)|^2 \mathrm{d}P_x(t)} = \frac{\sum\limits_{k=1}^{n} \overline{h_1(x_k)} \int_{\mathbb{C}} t_2 \mathrm{d}P_{x,y}(x_k, t_2)}{\sum\limits_{k=1}^{n} |h_1(x_k)|^2 \cdot P_x(\{x_k\})}.$$

Korzystając z (4.48), dostajemy

$$\int_{\mathbb{C}} t_2 \mathrm{d}P_{x,y}(x_k,t_2) = \int_{\mathbb{R}} t_2 \cdot \frac{1}{n} \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot \mathrm{e}^{\frac{-(t_2-y_k)^2}{2\sigma^2}} \mathrm{d}t_2 = \frac{1}{n\sqrt{\pi}} \int_{\mathbb{R}} t_2 \cdot \mathrm{e}^{\frac{-(t_2-y_k)^2}{2\sigma^2}} \frac{\mathrm{d}t_2}{\sigma\sqrt{2}}.$$

Całkując przez podstawianie z nową zmienną $\tau := \frac{t_2-y_k}{\sqrt{2}\sigma}$ mamy

(4.49)
$$\int_{\mathbb{C}} t_2 \mathrm{d}P_{x,y}(x_k,t_2) = \frac{1}{n\sqrt{\pi}} \int_{\mathbb{R}} (\sqrt{2}\sigma\tau + y_k) \cdot \mathrm{e}^{-\tau^2} \mathrm{d}\tau = \frac{y_k}{n} \cdot \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\tau^2} \mathrm{d}\tau = \frac{y_k}{n},$$

gdyż

$$\mathrm{d}\tau = \frac{\mathrm{d}t_2}{\sqrt{2}\sigma}, \qquad \sqrt{2}\sigma \int_{\mathbb{R}} \tau \cdot \mathrm{e}^{-\tau^2} \mathrm{d}\tau = 0 \quad \text{i} \quad \int_{\mathbb{R}} \mathrm{e}^{-\tau^2} \mathrm{d}\tau = \sqrt{\pi}.$$

Dowód ostatniej równości można znaleźć np. w [27, §3.1]. Zatem

$$\mathrm{E}(g \circ y | h_1 \circ x) = \frac{\sum\limits_{k=1}^{n} \overline{h_1(x_k)} y_k \cdot \frac{1}{n}}{\sum\limits_{k=1}^{n} |h_1(x_k)|^2 \cdot \frac{1}{n}} = \frac{\sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_1 x_k}}}{\sum\limits_{k=1}^{n} |\mathrm{e}^{r_1 x_k}|^2} = \frac{\sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_1} x_k}}{\sum\limits_{k=1}^{n} \mathrm{e}^{2x_k \operatorname{Re} r_1}} = \frac{\sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_1} x_k}}{c_1}$$

oraz podobnie

$$\mathrm{E}(g \circ y | h_2 \circ x) = \frac{\sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_2} x_k}}{c_2}.$$

Podstawiając wyznaczone wielkości do (4.45) i (4.46) dostajemy odpowiednio

$$\alpha_2 = \frac{\frac{\sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_2} x_k}}{c_2} - \frac{\sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_1} x_k}}{c_1} \cdot \frac{\overline{c}}{c_2}}{1 - \frac{c}{c_1} \cdot \frac{\overline{c}}{c_2}} = \frac{c_1 \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_2} x_k} - \overline{c} \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_1} x_k}}{c_1 c_2 - |c|^2}$$

oraz podobnie

$$\alpha_1 = \frac{c_2 \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_1} x_k} - c \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{\overline{r_2} x_k}}{c_1 c_2 - |c|^2}.$$

W szczególności, jeśli $\Delta > 0$, to $c = \overline{c},\ r_1 := -\frac{R}{2L} + \frac{\sqrt{\Delta}}{2},\ r_2 := -\frac{R}{2L} - \frac{\sqrt{\Delta}}{2},$ i

$$\alpha_2 = \frac{c_1 \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{r_2 x_k} - c \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{r_1 x_k}}{c_1 c_2 - c^2}$$

oraz

$$\alpha_1 = \frac{c_2 \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{r_1 x_k} - c \sum\limits_{k=1}^{n} y_k \cdot \mathrm{e}^{r_2 x_k}}{c_1 c_2 - c^2}$$

Zatem

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\left\{\mathbb{R}\ni t\mapsto\left(\alpha_1\mathrm{e}^{\frac{\sqrt{\Delta}}{2}t}+\alpha_2\mathrm{e}^{-\frac{\sqrt{\Delta}}{2}t}\right)\mathrm{e}^{-\frac{R}{2L}t}\right\}.$$

Jeśli zaś $\Delta<0$, to $r_1:=-\frac{R}{2L}+\mathrm{i}\frac{\sqrt{-\Delta}}{2}$, $r_2=\overline{r_1}$, i $c_1=c_2=\sum\limits_{k=1}^{n}\mathrm{e}^{2x_k\,\mathrm{Re}\,r_1}$ oraz $c=\sum\limits_{k=1}^{n}\mathrm{e}^{2r_2x_k}$. Otrzymujemy zatem

$$\alpha_2=\frac{c_1\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{r_1x_k}-\overline{c}\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{r_2x_k}}{c_1^2-|c|^2}$$

oraz

$$\alpha_1=\frac{c_1\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{r_2x_k}-c\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{r_1x_k}}{c_1^2-|c|^2}=\overline{\alpha_2}. \tag{4.50}$$

W konsekwencji $\alpha_2h_2+\alpha_1h_1=\overline{\alpha_1}\overline{h_1}+\alpha_1h_1=2\,\mathrm{Re}(\alpha_1h_1)$, a więc

$$\begin{aligned}\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)&=\{2\,\mathrm{Re}(\alpha_1h_1)\}=\left\{\mathbb{R}\ni t\mapsto 2\,\mathrm{Re}(\alpha_1\mathrm{e}^{r_1t})\right\}\\&=\left\{\mathbb{R}\ni t\mapsto 2|\alpha_1|\,\mathrm{Re}\left(\mathrm{e}^{r_1t}\frac{\alpha_1}{|\alpha_1|}\right)\right\}\\&=\left\{\mathbb{R}\ni t\mapsto 2|\alpha_1|\mathrm{e}^{-\frac{R}{2L}t}\cos\left(\frac{\sqrt{-\Delta}}{2}t+\varphi\right)\right\},\end{aligned}$$

gdzie φ jest dowolnym argumentem liczby α_1, tzn. liczbą rzeczywistą spełniającą równość $\alpha_1=|\alpha_1|\mathrm{e}^{\mathrm{i}\varphi}$.

Pozostaje rozważyć przypadek, gdy $\Delta=0$. Wtedy $r_1=r_2=r:=-\frac{R}{2L}$. Z równości (4.31) mamy

$$\begin{aligned}\mathrm{E}(h_2\circ x|h_1\circ x)&=\frac{\int_A h_2(t)\overline{h_1(t)}\mathrm{d}P_x(t)}{\int_A|h_1(t)|^2\mathrm{d}P_x(t)}=\frac{\sum\limits_{k=1}^{n}h_2(x_k)\overline{h_1(x_k)}\cdot P_x(\{x_k\})}{\sum\limits_{k=1}^{n}|h_1(x_k)|^2\cdot P_x(\{x_k\})}\\&=\frac{\sum\limits_{k=1}^{n}x_k\mathrm{e}^{rx_k}\mathrm{e}^{rx_k}\cdot\frac{1}{n}}{\sum\limits_{k=1}^{n}|\mathrm{e}^{rx_k}|^2\cdot\frac{1}{n}}=\frac{\sum\limits_{k=1}^{n}x_k\mathrm{e}^{2rx_k}}{\sum\limits_{k=1}^{n}\mathrm{e}^{2rx_k}}\end{aligned}$$

oraz podobnie

$$\mathrm{E}(h_1\circ x|h_2\circ x)=\frac{\sum\limits_{k=1}^{n}x_k\mathrm{e}^{2rx_k}}{\sum\limits_{k=1}^{n}|x_k|^2\mathrm{e}^{2rx_k}}.$$

Z równości (4.33) oraz (4.49) mamy zaś

$$\mathrm{E}(g\circ y|h_1\circ x)=\frac{\int_{A\times\mathbb{C}} t_2\cdot\overline{h_1(t_1)}\mathrm{d}P_{x,y}(t_1,t_2)}{\int_A|h_1(t)|^2\mathrm{d}P_x(t)}=\frac{\sum\limits_{k=1}^{n}\overline{h_1(x_k)}\int_{\mathbb{C}} t_2\mathrm{d}P_{x,y}(x_k,t_2)}{\sum\limits_{k=1}^{n}|h_1(x_k)|^2\cdot P_x(\{x_k\})}$$
$$=\frac{\sum\limits_{k=1}^{n}\overline{h_1(x_k)}y_k\cdot\frac{1}{n}}{\sum\limits_{k=1}^{n}|h_1(x_k)|^2\cdot\frac{1}{n}}=\frac{\sum\limits_{k=1}^{n}y_k\mathrm{e}^{rx_k}}{\sum\limits_{k=1}^{n}\mathrm{e}^{2rx_k}}$$

oraz podobnie

$$\mathrm{E}(g\circ y|h_2\circ x)=\frac{\sum\limits_{k=1}^{n}y_k\cdot x_k\mathrm{e}^{rx_k}}{\sum\limits_{k=1}^{n}|x_k|^2\mathrm{e}^{2rx_k}}.$$

Oznaczając

$$d:=\sum_{k=1}^{n}x_k\mathrm{e}^{2rx_k},\qquad d_1:=\sum_{k=1}^{n}\mathrm{e}^{2rx_k},\qquad d_2:=\sum_{k=1}^{n}|x_k|^2\mathrm{e}^{2rx_k}$$

otrzymujemy

$$\mathrm{E}(h_2\circ x|h_1\circ x)=\frac{d}{d_1}$$
$$\mathrm{E}(h_1\circ x|h_2\circ x)=\frac{d}{d_2}$$
$$\mathrm{E}(g\circ y|h_1\circ x)=\frac{\sum\limits_{k=1}^{n}y_k\mathrm{e}^{rx_k}}{d_1}$$
$$\mathrm{E}(g\circ y|h_2\circ x)=\frac{\sum\limits_{k=1}^{n}y_k\cdot x_k\mathrm{e}^{rx_k}}{d_2}.$$

Podstawiając wyznaczone wielkości do (4.45) i (4.46) dostajemy

$$\alpha_2=\frac{\frac{\sum\limits_{k=1}^{n}y_k\cdot x_k\mathrm{e}^{rx_k}}{d_2}-\frac{\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{rx_k}}{d_1}\cdot\frac{d}{d_2}}{1-\frac{d}{d_1}\cdot\frac{d}{d_2}}=\frac{d_1\sum\limits_{k=1}^{n}y_k\cdot x_k\mathrm{e}^{rx_k}-d\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{rx_k}}{d_1d_2-d^2}$$

oraz

$$\alpha_1=\frac{d_2\sum\limits_{k=1}^{n}y_k\cdot\mathrm{e}^{rx_k}-d\sum\limits_{k=1}^{n}y_k\cdot x_k\mathrm{e}^{rx_k}}{d_1d_2-d^2}.$$

Zatem

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\left\{\mathbb{R}\ni t\mapsto(\alpha_1+\alpha_2t)\mathrm{e}^{-\frac{R}{2L}t}\right\}.$$

Uwaga 4.15. Zauważmy, że równość $\alpha_1 = \overline{\alpha_2}$ w (4.50) ma charakter ogólniejszy. Załóżmy, że $g \circ y(\Omega) \subset \mathbb{R}$ i dowolnie zadane funkcje $h_1, h_2 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ są wzajemnie sprzężone, tzn. $h_2 = \overline{h_1}$. Wtedy $\alpha_2 = \overline{\alpha_1}$, i

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\alpha_1 h_1 + \overline{\alpha_1}\overline{h_1}\} = \{2\,\mathrm{Re}(\alpha_1 h_1)\},$$

a więc funkcja regresji jest rzeczywista. Ze wzorów (4.43) mamy bowiem

$$\begin{aligned}
\overline{\alpha_1} &= \frac{g^*(h_1) - \langle h_1|h_2\rangle\overline{\alpha_2}}{\|h_1\|^2} \\
&= \frac{1}{\|h_1\|^2}\left(g^*(h_1) - \langle h_1|h_2\rangle\frac{g^*(h_2)\|h_1\|^2 - g^*(h_1)\langle h_2|h_1\rangle}{\|h_2\|^2\|h_1\|^2 - |\langle h_2|h_1\rangle|^2}\right) \\
&= \frac{g^*(h_1)\|h_2\|^2\|h_1\|^2 - g^*(h_2)\|h_1\|^2\langle h_1|h_2\rangle}{\|h_1\|^2\left(\|h_2\|^2\|h_1\|^2 - |\langle h_2|h_1\rangle|^2\right)} \\
&= \frac{\overline{g^*(h_2)}\|h_1\|^2 - \overline{g^*(h_1)}\cdot\overline{\langle h_2|h_1\rangle}}{\|h_2\|^2\|h_1\|^2 - |\langle h_2|h_1\rangle|^2} = \alpha_2,
\end{aligned}$$

gdyż $\|h_1\| = \|h_2\|$ i $g^*(h_1) = \overline{g^*(h_2)}$.

4.5 Przypadek trójwymiarowy

Rozważmy przypadek, w którym model funkcyjny $\mathcal{F}$ rozpięty jest przez trzy ustalone funkcje $h_1, h_2, h_3 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, tzn., $\mathcal{F} = \mathrm{lin}(\{h_1, h_2, h_3\})$. Załóżmy, że $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,3}$, gdzie $\mathbb{Z}_{1,3} \ni k \mapsto h'_k$ jest ciągiem zdefiniowanym przez (4.6). Stosując twierdzenie 4.10 otrzymujemy

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{\sum_{k=1}^{3}\frac{\overline{g^*(h'_k)}}{\|h'_k\|^2}h'_k\right\},$$

gdzie funkcje h'_1, h'_2, h'_3 są wyznaczone na podstawie wzoru (4.6) z $p := 3$.
Z uwagi 4.1 wynika, że

$$h'_3 \perp \mathrm{lin}(\{h'_2, h'_1\}) \quad \text{oraz} \quad \mathrm{lin}(\{h'_2, h'_1\}) = \mathrm{lin}(\{h_2, h_1\}). \tag{4.51}$$

Stąd $h'_3 \perp \mathrm{lin}(\{h_2, h_1\})$ oraz istnieją $\nu_{3,2}, \nu_{3,1} \in \mathbb{C}$ takie, że

$$h'_3 = h_3 - \langle h_3|h'_2\rangle_* \cdot h'_2 - \langle h_3|h'_1\rangle_* \cdot h'_1 = h_3 - \nu_{3,2}h_2 - \nu_{3,1}h_1.$$

Dlatego

$$\nu_{3,2}\langle h_2|h_l\rangle + \nu_{3,1}\langle h_1|h_l\rangle = \langle h_3 - h'_3|h_l\rangle = \langle h_3|h_l\rangle, \quad l \in \mathbb{Z}_{1,2}. \tag{4.52}$$

Ponieważ $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,3}$, więc na mocy uwagi 4.2, $\|h_k\| > 0$, $k \in \mathbb{Z}_{1,3}$. Ponadto na mocy wzoru (4.6),

$$\langle h_2|h_2\rangle\langle h_1|h_1\rangle - \langle h_2|h_1\rangle\langle h_1|h_2\rangle = \det\left[\langle h_k|h_l\rangle\right]_{2\times 2} = \|h'_2\|^2\|h'_1\|^2 > 0.$$

Z równości (4.52) wynika więc, że

$$\nu_{3,k}=\frac{\langle h_3|h_k\rangle\langle h_l|h_l\rangle-\langle h_3|h_l\rangle\langle h_l|h_k\rangle}{\langle h_2|h_2\rangle\langle h_1|h_1\rangle-\langle h_2|h_1\rangle\langle h_1|h_2\rangle}=\frac{\langle h_3|h_k\rangle_*-\langle h_3|h_l\rangle_*\langle h_l|h_k\rangle_*}{1-\langle h_2|h_1\rangle_*\langle h_1|h_2\rangle_*},\quad k\in\mathbb{Z}_{1,2},\ l\in\mathbb{Z}_{1,2}\backslash\{k\}. \tag{4.53}$$

Z równości $\mathcal{F}=\operatorname{lin}(\{h'_1,h'_2,h'_3\})$ wynika, że

$$\sum_{k=1}^{3}\frac{\overline{g^*(h'_k)}}{\|h'_k\|^2}h'_k=\sum_{k=1}^{3}\alpha_k h_k \tag{4.54}$$

dla pewnego ciągu $\mathbb{Z}_{1,3}\ni k\mapsto\alpha_k\in\mathbb{C}$, skąd na podstawie (4.51)

$$\alpha_3\langle h_3|h'_3\rangle=\left\langle\sum_{k=1}^{3}\alpha_k h_k\Big|h'_3\right\rangle=\left\langle\sum_{k=1}^{3}\frac{\overline{g^*(h'_k)}}{\|h'_k\|^2}h'_k\Big|h'_3\right\rangle=\overline{g^*(h'_3)}.$$

Z równości (4.51) wynika, że $\langle h_3|h'_3\rangle=\|h'_3\|>0$, i tym samym $\langle h'_3|h_3\rangle=\overline{\langle h_3|h'_3\rangle}=\langle h_3|h'_3\rangle$. Zatem

$$\alpha_3=\frac{\overline{g^*(h'_3)}}{\langle h'_3|h_3\rangle}=\frac{\overline{g^*(h_3-\nu_{3,2}h_2-\nu_{3,1}h_1)}}{\langle h_3-\nu_{3,2}h_2-\nu_{3,1}h_1|h_3\rangle}=\frac{\overline{g^*(h_3)}-\overline{\nu_{3,2}}\overline{g^*(h_2)}-\overline{\nu_{3,1}}\overline{g^*(h_1)}}{\langle h_3|h_3\rangle-\nu_{3,2}\langle h_2|h_3\rangle-\nu_{3,1}\langle h_1|h_3\rangle}.$$

Dzieląc więc licznik i mianownik przez $\|h_3\|^2$ dostajemy

$$\alpha_3=\frac{\overline{\frac{g^*(h_3)}{\|h_3\|^2}}-\overline{\nu_{3,2}}\frac{\|h_2\|^2}{\|h_3\|^2}\overline{\frac{g^*(h_2)}{\|h_2\|^2}}-\overline{\nu_{3,1}}\frac{\|h_1\|^2}{\|h_3\|^2}\overline{\frac{g^*(h_1)}{\|h_1\|^2}}}{1-\nu_{3,2}\langle h_2|h_3\rangle_*-\nu_{3,1}\langle h_1|h_3\rangle_*}=\frac{\overline{\frac{g^*(h_3)}{\|h_3\|^2}}-\nu_{2,3}\overline{\frac{g^*(h_2)}{\|h_2\|^2}}-\nu_{1,3}\overline{\frac{g^*(h_1)}{\|h_1\|^2}}}{1-\nu_{3,2}\langle h_2|h_3\rangle_*-\nu_{3,1}\langle h_1|h_3\rangle_*} \tag{4.55}$$

gdzie

$$\nu_{k,3}:=\overline{\nu_{3,k}}\frac{\|h_k\|^2}{\|h_3\|^2}=\frac{\langle h_k|h_3\rangle\langle h_l|h_l\rangle-\langle h_k|h_l\rangle\langle h_l|h_3\rangle}{\langle h_2|h_2\rangle\langle h_1|h_1\rangle-\langle h_2|h_1\rangle\langle h_1|h_2\rangle}\cdot\frac{\|h_k\|^2}{\|h_3\|^2}=\frac{\langle h_k|h_3\rangle_*-\langle h_k|h_l\rangle_*\langle h_l|h_3\rangle_*}{1-\langle h_2|h_1\rangle_*\langle h_1|h_2\rangle_*},\qquad k\in\mathbb{Z}_{1,2},\ l\in\mathbb{Z}_{1,2}\backslash\{k\}. \tag{4.56}$$

Porównując równości (4.53) i (4.56) stwierdzamy, że

$$\nu_{3,k}=\mu_{3,k}\quad\text{i}\quad\nu_{k,3}=\mu_{k,3},\qquad k\in\mathbb{Z}_{1,2}, \tag{4.57}$$

gdzie dla dowolnych $n\in\mathbb{Z}_{1,3}$, $k\in\mathbb{Z}_{1,3}\backslash\{n\}$ i $l\in\mathbb{Z}_{1,3}\backslash\{n,k\}$,

$$\begin{aligned}\mu_{n,k}&:=\frac{\langle h_n|h_k\rangle_*-\langle h_n|h_l\rangle_*\langle h_l|h_k\rangle_*}{1-\langle h_k|h_l\rangle_*\langle h_l|h_k\rangle_*},\\ \mu_{k,n}&:=\frac{\langle h_k|h_n\rangle_*-\langle h_k|h_l\rangle_*\langle h_l|h_n\rangle_*}{1-\langle h_k|h_l\rangle_*\langle h_l|h_k\rangle_*}.\end{aligned} \tag{4.58}$$

Korzystając więc z równości (4.55) i (4.57) dostajemy

$$\alpha_3=\frac{\overline{\frac{g^*(h_3)}{\|h_3\|^2}}-\mu_{2,3}\overline{\frac{g^*(h_2)}{\|h_2\|^2}}-\mu_{1,3}\overline{\frac{g^*(h_1)}{\|h_1\|^2}}}{1-\mu_{3,2}\langle h_2|h_3\rangle_*-\mu_{3,1}\langle h_1|h_3\rangle_*}. \tag{4.59}$$

Ponieważ $\mathcal{F} = \operatorname{lin}(\{h_1, h_2, h_3\}) = \operatorname{lin}(\{h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}\})$ dla dowolnej permutacji $\sigma\colon \mathbb{Z}_{1,3} \to \mathbb{Z}_{1,3}$, więc tworząc ciąg $\mathbb{Z}_{1,3} \ni n \mapsto h'_n$ na podstawie ciągu spermutowanego $\mathbb{Z}_{1,3} \ni n \mapsto h_{\sigma(n)}$ w miejsce ciągu $\mathbb{Z}_{1,3} \ni n \mapsto h_n$ (por. wzór (4.6)) możemy wyprowadzić z równości (4.54) i (4.59) ogólną zależność

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \sum_{n=1}^{3} \alpha_n h_n \right\},$$

gdzie

$$\alpha_n := \frac{\overline{\frac{g^*(h_n)}{\|h_n\|^2}} - \mu_{k,n} \overline{\frac{g^*(h_k)}{\|h_k\|^2}} - \mu_{l,n} \overline{\frac{g^*(h_l)}{\|h_l\|^2}}}{1 - \mu_{n,k}\langle h_k | h_n \rangle_* - \mu_{n,l}\langle h_l | h_n \rangle_*}, \quad n \in \mathbb{Z}_{1,3},\ k, l \in \mathbb{Z}_{1,3} \setminus \{n\},\ k \neq l.$$

Stąd

$$\begin{aligned} \alpha_n &= \frac{\overline{\frac{g^*(h_n)}{\|h_n\|^2}} - \sum\limits_{k \in \mathbb{Z}_{1,3}\setminus\{n\}} \mu_{k,n} \overline{\frac{g^*(h_k)}{\|h_k\|^2}}}{1 - \sum\limits_{k \in \mathbb{Z}_{1,3}\setminus\{n\}} \mu_{n,k}\langle h_k | h_n \rangle_*} \\ &= \frac{\mathrm{E}(g \circ y | h_n \circ x) - \sum\limits_{k \in \mathbb{Z}_{1,3}\setminus\{n\}} \mu_{k,n} \mathrm{E}(g \circ y | h_k \circ x)}{1 - \sum\limits_{k \in \mathbb{Z}_{1,3}\setminus\{n\}} \mu_{n,k} \mathrm{E}(h_k \circ x | h_n \circ x)}, \qquad n \in \mathbb{Z}_{1,3} \end{aligned} \tag{4.60}$$

przy czym na mocy (4.58) dla dowolnych parami różnych $n, k, l \in \mathbb{Z}_{1,3}$,

$$\begin{aligned} \mu_{n,k} &= \frac{\mathrm{E}(h_n \circ x | h_k \circ x) - \mathrm{E}(h_n \circ x | h_l \circ x)\,\mathrm{E}(h_l \circ x | h_k \circ x)}{1 - \mathrm{E}(h_k \circ x | h_l \circ x)\,\mathrm{E}(h_l \circ x | h_k \circ x)}, \\ \mu_{k,n} &= \frac{\mathrm{E}(h_k \circ x | h_n \circ x) - \mathrm{E}(h_k \circ x | h_l \circ x)\,\mathrm{E}(h_l \circ x | h_n \circ x)}{1 - \mathrm{E}(h_k \circ x | h_l \circ x)\,\mathrm{E}(h_l \circ x | h_k \circ x)}. \end{aligned}$$

W szczególności wzór (4.60) można wykorzystać do wyznaczenia funkcji regresji w klasie $\mathcal{F} := \left\{\mathbb{R} \ni t \mapsto \lambda_1 + \lambda_2 \cos\left(\frac{\pi}{6}t\right) + \lambda_3 \sin\left(\frac{\pi}{6}t\right)\colon \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}\right\}$ względem zadanej struktury $\mathfrak{P}$. Model teoretyczny $\mathcal{F}$ odzwierciedla np. zużycie energii elektrycznej w danym mieście, które waha się okresowo w ciągu roku.

4.6 Przypadek n-wymiarowy, przegląd podstawowych typów modeli teoretycznych

W podrozdziałach 4.3 - 4.5 rozważaliśmy przypadki, gdy model funkcyjny $\mathcal{F}$ był rozmaitością liniową rozpiętą przez jedną, dwie lub trzy funkcje odpowiednio. W praktyce czasem zachodzi potrzeba rozważania dokładniejszych modeli z użyciem większej ilości funkcji bazowych. Załóżmy więc, że $\mathcal{F} = \operatorname{lin}(\{h_k\colon k \in \mathbb{Z}_{1,p}\})$ dla dowolnie zadanych $p \in \mathbb{N}$ i ciągu $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. Wtedy klasę

$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$ można wyznaczyć na podstawie twierdzenia 4.11 bądź twierdzenia 4.12 z uwzględnieniem w razie potrzeby uwagi 4.13. Załóżmy, że $\|h'_k\|>0$ dla $k\in\mathbb{Z}_{1,p}$. Wtedy na podstawie twierdzenia 4.7, $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\{f\}$ dla pewnego $f\in\mathcal{F}$. Stąd $f=\sum\limits_{k=1}^{p}\alpha_k h_k$ dla pewnego ciągu $\mathbb{Z}_{1,p}\ni k\mapsto\alpha_k\in\mathbb{C}$. Z twierdzenia 3.8 wynika, że $\langle h|f\rangle=g^*(h)$ dla $h\in\mathcal{F}$. W szczególności $\langle h_k|f\rangle=g^*(h_k)$ dla $k\in\mathbb{Z}_{1,p}$. Ponieważ dla każdego $k\in\mathbb{Z}_{1,p}$, $\Big\langle h_k\Big|\sum\limits_{l=1}^{p}\alpha_l h_l\Big\rangle=\sum\limits_{l=1}^{p}\overline{\alpha_l}\overline{\langle h_l|h_k\rangle}$, więc $\sum\limits_{l=1}^{p}\alpha_l\langle h_l|h_k\rangle=\overline{g^*(h_k)}$. Ponadto na mocy uwagi 4.2, $\|h_k\|>0$ dla $k\in\mathbb{Z}_{1,p}$. Zatem

$$\sum_{l=1}^{p}\alpha_l\langle h_l|h_k\rangle_*=\frac{\overline{g^*(h_k)}}{\|h_k\|^2},\qquad k\in\mathbb{Z}_{1,p}.$$

Powyższe równości można wyrazić w postaci macierzowej następująco

$$H\cdot\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{bmatrix}=\begin{bmatrix}\frac{\overline{g^*(h_1)}}{\|h_1\|^2}\\ \frac{\overline{g^*(h_2)}}{\|h_2\|^2}\\ \vdots\\ \frac{\overline{g^*(h_p)}}{\|h_p\|^2}\end{bmatrix},\tag{4.61}$$

gdzie

$$H:=[\mathbb{Z}_{1,p}\times\mathbb{Z}_{1,p}\ni(k,l)\mapsto\langle h_l|h_k\rangle_*]=[\mathbb{Z}_{1,p}\times\mathbb{Z}_{1,p}\ni(k,l)\mapsto\mathrm{E}(h_l\circ x|h_k\circ x)]\,.$$

Ponieważ $\|h'_k\|>0$ i $\|h_k\|>0$ dla $k\in\mathbb{Z}_{1,p}$, więc korzystając z metody dowodu twierdzeń 11.8 i 11.9 w [19] otrzymujemy $\det H=\prod\limits_{k=1}^{p}\frac{\|h'_k\|^2}{\|h_k\|^2}>0$. Dlatego istnieje macierz odwrotna H^{-1}. Mnożąc stronami równość (4.61) przez macierz H^{-1} dostajemy równości

$$\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{bmatrix}=H^{-1}\cdot\begin{bmatrix}\frac{\overline{g^*(h_1)}}{\|h_1\|^2}\\ \frac{\overline{g^*(h_2)}}{\|h_2\|^2}\\ \vdots\\ \frac{\overline{g^*(h_p)}}{\|h_p\|^2}\end{bmatrix}=H^{-1}\cdot\begin{bmatrix}\mathrm{E}(g\circ y|h_1\circ x)\\ \mathrm{E}(g\circ y|h_2\circ x)\\ \vdots\\ \mathrm{E}(g\circ y|h_p\circ x)\end{bmatrix}.\tag{4.62}$$

Równości (4.62) mają znaczenie teoretyczne. W praktyce obliczeń numerycznych, dla większych wartości parametru p będziemy stosować twierdzenie 4.11 bądź twierdzenie 4.12. Stosowne przykłady będą podane w rozdziale piątym. W pewnych sytuacjach twierdzenia te są łatwiejsze w użyciu niż równości (4.62), co pokazuje następujący przykład.

Przykład 4.16. Załóżmy, że $x\colon\Omega\to\mathbb{R}$ i $y\colon\Omega\to\mathbb{R}$ są rzeczywistymi zmiennymi losowymi o rozkładzie łącznym z funkcją gęstości postaci

$$\begin{aligned}f(x,y)&=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\Bigg(-\frac{1}{2(1-\rho^2)}\Big(\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{x-\mu_1}{\sigma_1}\frac{y-\mu_2}{\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\Big)\Bigg)\\&=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(\frac{-\big(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\big)^2}{2(1-\rho^2)\sigma_2^2}\right)\exp\Big(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\Big),\end{aligned}$$

gdzie $\rho \in (-1;1)$ jest współczynnikiem korelacji; por. [11, Definition 5.11.1], [20, §5.10]. Wtedy rozkłady brzegowe wektora losowego (x,y), czyli rozkłady rzeczywistych zmiennych losowych x i y są rozkładami normalnymi odpowiednio $N(\mu_1,\sigma_1)$ i $N(\mu_2,\sigma_2)$. Ustalmy dowolnie $p \in \mathbb{Z}_3$. Przyjmując $A := \mathbb{R}$ mamy $h_k \in \mathrm{L}^2(A,\mathcal{A}_x,P_x)$ dla $k \in \mathbb{Z}_{1,p}$ i $g \in \mathrm{L}^2(\mathbb{C},\mathcal{A}_y,P_y)$, gdzie $g(t) := t$, $h_k(t) := t^{k-1}$, $k \in \mathbb{Z}_{1,p}$ dla $t \in \mathbb{C}$. Wyznaczymy klasę regresji $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$, gdzie $\mathcal{F} := \mathrm{lin}(\{h_k\colon k \in \mathbb{Z}_{1,p}\})$. Korzystając z formuły (3.10), z uwagi 3.5 oraz wzoru (4.1) otrzymujemy

$$\begin{aligned}\langle h_n|h_k\rangle &= \int_{\mathbb{R}} h_n(t)\overline{h_k(t)}\,\mathrm{d}P_x(t) = \int_{\mathbb{R}} t^{n+k-2}\,\mathrm{d}P_x(t)\\ &= \int_\Omega x^{n+k-2}(\omega)\,\mathrm{d}P(\omega) = \mathrm{E}\left(x^{n+k-2}\right), \qquad n,k \in \mathbb{Z}_{1,p},\end{aligned}$$

skąd

$$\begin{aligned}&\langle h_1|h_1\rangle = 1,\\ &\langle h_2|h_1\rangle = \langle h_1|h_2\rangle = \mathrm{E}(x) = \mu_1,\\ &\langle h_2|h_2\rangle = \mathrm{E}(x^2) = \mathrm{E}(((x-\mu_1)+\mu_1)^2)\\ &\qquad = \mathrm{E}((x-\mu_1)^2) + 2\mu_1\,\mathrm{E}(x-\mu_1) + \mu_1^2 = \sigma_1^2 + \mu_1^2.\end{aligned}$$

Z (4.6) otrzymujemy

$$\begin{aligned}h_1' &= h_1,\\ h_2' &= h_2 - \langle h_2|h_1'\rangle_* \cdot h_1 = h_2 - \mu_1 h_1.\end{aligned}$$

Zatem

$$\langle h_n|h_1'\rangle_* = \frac{\langle h_n|h_1\rangle}{\|h_1\|^2} = \langle h_n|h_1\rangle = \mathrm{E}(x^{n-1}), \qquad n \in \mathbb{Z}_{1,p}$$

oraz na mocy uwagi 4.2,

$$\|h_2'\|^2 = \langle h_2|h_2'\rangle = \langle h_2|h_2-\mu_1 h_1\rangle = \langle h_2|h_2\rangle - \mu_1\langle h_2|h_1\rangle = \sigma_1^2 + \mu_1^2 - \mu_1^2 = \sigma_1^2,$$

skąd

$$\langle h_n|h_2'\rangle_* = \frac{\langle h_n|h_2\rangle - \mu_1\langle h_n|h_1\rangle}{\|h_2'\|^2} = \frac{\mathrm{E}(x^n) - \mu_1\,\mathrm{E}(x^{n-1})}{\sigma_1^2}, \qquad n \in \mathbb{Z}_{1,p}.$$

Ponieważ x i y są zmiennymi losowymi, wnioskujemy ze wzoru (3.19), że dla $n \in \mathbb{Z}_{1,p}$

$$g^*(h_n) = \int_\Omega h_n \circ x(\omega)\cdot\overline{g\circ y(\omega)}\mathrm{d}P(\omega) = \int_\Omega x^{n-1}(\omega)\cdot y(\omega)\mathrm{d}P(\omega) = \mathrm{E}\left(x^{n-1}y\right),$$

co daje

$$\begin{aligned}g^*(h_n) = \mathrm{E}(x^{n-1}y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^{n-1}yf(x,y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty} x^{n-1}\left(\int_{-\infty}^{+\infty} yf(x,y)\mathrm{d}y\right)\mathrm{d}x\\ &= \int_{-\infty}^{+\infty}\left(x^{n-1}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\Big(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\Big)\right.\\ &\qquad\left.\cdot\int_{-\infty}^{+\infty} y\exp\Big(\frac{-(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))^2}{2(1-\rho^2)\sigma_2^2}\Big)\mathrm{d}y\right)\mathrm{d}x.\end{aligned}$$

Obliczając całkę wewnętrzną przez podstawienie $s := \frac{y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)}{\sqrt{2}\sqrt{1-\rho^2}\sigma_2}$ mamy

$$\begin{aligned}
&\int_{-\infty}^{+\infty} y \exp\left(\frac{-\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2}{2(1-\rho^2)\sigma_2^2}\right)\mathrm{d}y \\
&\quad= \int_{-\infty}^{+\infty}\left(\sqrt{2}\sqrt{1-\rho^2}\sigma_2\cdot s+\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)\mathrm{e}^{-s^2}\cdot\sqrt{2}\sqrt{1-\rho^2}\sigma_2\mathrm{d}s \\
&\quad= 2(1-\rho^2)\sigma_2^2\int_{-\infty}^{+\infty} s\mathrm{e}^{-s^2}\mathrm{d}s+\sqrt{2}\sqrt{1-\rho^2}\sigma_2\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)\int_{-\infty}^{+\infty}\mathrm{e}^{-s^2}\mathrm{d}s \\
&\quad= 2(1-\rho^2)\sigma_2^2\cdot 0+\sqrt{2}\sqrt{1-\rho^2}\sigma_2\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)\sqrt{\pi} \\
&\quad= \sqrt{2\pi}\sqrt{1-\rho^2}\sigma_2\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right).
\end{aligned}$$

Stąd

$$\begin{aligned}
g^*(h_n) &= \int_{-\infty}^{+\infty}\Bigg(x^{n-1}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\big(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\big) \\
&\qquad\qquad\cdot\left(\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_2\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)\right)\Bigg)\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\left(\mu_2-\rho\frac{\sigma_2}{\sigma_1}\mu_1\right)\int_{-\infty}^{+\infty}x^{n-1}\exp\big(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\big)\mathrm{d}x \\
&\qquad+\frac{\rho\frac{\sigma_2}{\sigma_1}}{\sqrt{2\pi}\sigma_1}\int_{-\infty}^{+\infty}x^n\exp\big(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\big)\mathrm{d}x \\
&= \left(\mu_2-\rho\frac{\sigma_2}{\sigma_1}\mu_1\right)\mathrm{E}(x^{n-1})+\rho\frac{\sigma_2}{\sigma_1}\mathrm{E}(x^n)=\mu_2\,\mathrm{E}(x^{n-1})+\rho\frac{\sigma_2}{\sigma_1}\Big(\mathrm{E}(x^n)-\mu_1\,\mathrm{E}(x^{n-1})\Big),
\end{aligned}$$

a w szczególności

$$\begin{aligned}
g^*(h_1') &= g^*(h_1)=\mu_2 \\
g^*(h_2') &= g^*(h_2)-\mu_1 g^*(h_1)=\mu_1\mu_2+\rho\sigma_1\sigma_2-\mu_1\cdot\mu_2=\rho\sigma_1\sigma_2.
\end{aligned}$$

Zatem, dla każdego $n\in\mathbb{Z}_{3,p}$, mamy

(4.63)

$$\begin{aligned}
g^*(h_n-\langle h_n|h_2'\rangle_* h_2'-\langle h_n|h_1'\rangle_* h_1') &= g^*(h_n)-\langle h_n|h_2'\rangle_* g^*(h_2')-\langle h_n|h_1'\rangle_* g^*(h_1') \\
&= \mu_2\,\mathrm{E}(x^{n-1})+\rho\frac{\sigma_2}{\sigma_1}\Big(\mathrm{E}(x^n)-\mu_1\,\mathrm{E}(x^{n-1})\Big) \\
&\qquad-\rho\sigma_1\sigma_2\frac{1}{\sigma_1^2}\Big(\mathrm{E}(x^n)-\mu_1\,\mathrm{E}(x^{n-1})\Big)-\mu_2\,\mathrm{E}(x^{n-1})=0.
\end{aligned}$$

Pokażemy, za pomocą indukcji zupełnej, że

$$g^*(h_n')=0,\qquad n\in\mathbb{Z}_{3,p}, \tag{4.64}$$

o ile $p\geqslant 3$. Stosując (4.63) z $n:=3$ mamy $g^*(h_3')=0$. Ustalmy dowolnie $n\in\mathbb{Z}_{3,p-1}$ i załóżmy, że równość w (4.64) zachodzi dla wszystkich $n\in\mathbb{Z}_{3,l}$. Wtedy na mocy

(4.63) mamy

$$\begin{aligned} g^*(h'_{l+1}) &= g^*\left(h_{l+1} - \sum_{k=1}^{l}\langle h_{l+1}|h'_k\rangle_* h'_k\right) \\ &= g^*(h_{l+1}) - \sum_{k=3}^{l}\langle h_{l+1}|h'_k\rangle_* g^*(h'_k) - \langle h_{l+1}|h'_2\rangle_* g^*(h'_2) - \langle h_{l+1}|h'_1\rangle_* g^*(h'_1) \\ &= g^*(h_{l+1}) - \langle h_{l+1}|h'_2\rangle_* g^*(h'_2) - \langle h_{l+1}|h'_1\rangle_* g^*(h'_1) = 0. \end{aligned}$$

Stąd na mocy indukcji zupełnej dostajemy własność (4.64). Ponadto dla dowolnie zadanej funkcji $h \in \Theta \cap \mathcal{F}$,

$$0 = \|h\|^2 = \int_\Omega |h \circ x(\omega)|^2 \mathrm{d}P(\omega) = \int_{\mathbb{R}} |h(t)|^2 \mathrm{d}P_x(t) = \frac{1}{\sqrt{2\pi}\sigma_1}\int_{\mathbb{R}} |h(t)|^2 \mathrm{e}^{\frac{-(t-\mu_1)^2}{2\sigma_1^2}} \mathrm{d}t.$$

Stąd $h(t) = 0$ dla p.w. $t \in \mathbb{R}$. Ponieważ jednocześnie $h \in \mathcal{F}$, więc $h(t) = 0$ dla $t \in \mathbb{R}$. Dlatego $\Theta \cap \mathcal{F} = \{\theta\}$. Ponadto $\dim(\mathcal{F}) = p$. Z wniosku 4.5 wynika więc, że $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$. Stąd na mocy twierdzenia 4.10 i (4.64) dostajemy $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{f\}$, gdzie

$$\begin{aligned} f :&= \sum_{k=1}^{p} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k = \frac{\overline{g^*(h'_2)}}{\|h'_2\|^2} h'_2 + \frac{\overline{g^*(h'_1)}}{\|h'_1\|^2} h'_1 = \frac{\rho\sigma_1\sigma_2}{\sigma_1^2}(h_2 - \mu_1 h_1) + \mu_2 h_1 \\ &= \rho\frac{\sigma_2}{\sigma_1}(h_2 - \mu_1 h_1) + \mu_2 h_1 = \rho\frac{\sigma_2}{\sigma_1} h_2 + \left(\mu_2 - \mu_1\rho\frac{\sigma_2}{\sigma_1}\right) h_1. \end{aligned}$$

W ten sposób otrzymujemy klasyczną równość

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{\mathbb{R} \ni t \mapsto \left(\mu_2 - \mu_1\rho\frac{\sigma_2}{\sigma_1}\right) + \rho\frac{\sigma_2}{\sigma_1} t\right\}.$$

Jest to alternatywny dowód w stosunku do dowodu opartego na równości regresji pierwszego i drugiego rodzaju dla dwuwymiarowego rozkładu normalnego; por. [20, §6.4], [11, §3.8].

Odnotujmy, że rozważana w tej pracy teoria regresji jest stosowalna do modelu teoretycznego $\mathcal{F}$ będącego rozmaitością liniową rozpiętą przez układ dowolnie zadanych funkcji h_k, $k \in \mathbb{Z}_{1,p}$. Tym samym teoria ta jest niezależna od wyboru takich funkcji, co nadaje jej charakter uniwersalności. Odpowiednie modele teoretyczne $\mathcal{F}$ tworzymy na podstawie wiedzy teoretycznej i praktycznej o badanych zależnościach. Wstępnego dopasowania typu modelu do badanego zjawiska można również dokonać oceniając wzrokowo graficzną ilustrację zależności pomiędzy stosowanymi zmiennymi. Najczęściej w praktyce wykorzystywane są modele utworzone przez standardowe funkcje elementarne jak: wielomiany, funkcje potęgowe, funkcje wykładnicze, funkcje trygonometryczne, itp. W ten sposób można sklasyfikować odpowiednio modele na wielomianowe, potęgowe, wykładnicze, trygonometryczne, itp. Można również tworzyć modele mieszane wykorzystując w tym celu funkcje różnych typów. Klasycznym

typem teoretycznych modeli liniowych są modele wielomianowe i odpowiadające im regresje wielomianowe opisane w podrozdziale 3.3. Konkretne zagadnienie wyznaczenia regresji wielomianowych zostało opisane w przykładzie 4.16. W praktyce – z uwagi na złożoność obliczeniową – do wyznaczenia regresji stosuje się wspomaganie komputerowe. Zagadnienie to zostanie zilustrowane w części piątej niniejszej pracy.

4.7 Szczególny przypadek funkcji skalującej

Z dotychczasowych rozważań wynika, że funkcje regresji zależą od funkcji skalującej $g \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$. Zależność tę można łatwo wyrazić, gdy funkcja g jest wielomianem stopnia pierwszego. Oznaczmy przez $\mathrm{Pr}_{\mathcal{F}}(h)$ klasę projekcji ortogonalnych funkcji $h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ na $\mathcal{F}$, tzn. klasę wszystkich funkcji $h' \in \mathcal{F}$ takich, że $h - h' \perp v$ dla $v \in \mathcal{F}$. Następujące twierdzenie opisuje zależność pomiędzy klasą $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$ a klasą $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$, gdzie $\mathfrak{P}$ jest dowolnie ustaloną zespoloną probabilistyczną strukturą regresji.

Twierdzenie 4.17. *Dla dowolnie zadanych $p \in \mathbb{N}$ i $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ niech $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ będzie ciągiem zdefiniowanym wzorem* (4.34)*. Wtedy dla dowolnych $a, b \in \mathbb{C}$, jeśli $g = a \cdot \mathrm{id}_{\mathbb{C}} + b$, $\mathrm{id}_{\mathbb{C}} \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ i $\mathrm{Pr}_{\mathcal{F}}(\mathbb{1}_A) \neq \emptyset$, to dla wszystkich $f_1 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P})$ i $f_0 \in \mathrm{Pr}_{\mathcal{F}}(\mathbb{1}_A)$ zachodzi równość*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p} \setminus E\}) + a \cdot f_1 + b \cdot f_0, \tag{4.65}$$

gdzie $\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$ i $E := \{k \in \mathbb{Z}_{1,p} \colon \|h'_k\| > 0\}$.

Dowód. Załóżmy, że funkcja skalująca $g \colon \mathbb{C} \to \mathbb{C}$ jest postaci $g := a \cdot \mathrm{id}_{\mathbb{C}} + b$ dla pewnych $a, b \in \mathbb{C}$ oraz $\mathrm{id}_{\mathbb{C}} \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$. Wtedy dla $h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \setminus \Theta$ mamy

$$\begin{aligned}
\mathrm{E}(g \circ y | h \circ x) &= \frac{\mathrm{E}((g \circ y)\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} = \frac{\mathrm{E}((ay + b)\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} \\
&= a \frac{\mathrm{E}(y\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} + b \frac{\mathrm{E}(\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} \\
&= a \frac{\mathrm{E}((\mathrm{id}_{\mathbb{C}} \circ y)\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} + b \frac{\mathrm{E}((\mathbb{1}_{\mathbb{C}} \circ y)\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} \\
&= a\, \mathrm{E}(\mathrm{id}_{\mathbb{C}} \circ y | h \circ x) + b\, \mathrm{E}(\mathbb{1}_{\mathbb{C}} \circ y | h \circ x).
\end{aligned}$$

Załóżmy najpierw, że $E \neq \emptyset$. Stosując twierdzenie 4.12 dostajemy

$$\begin{aligned}
\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) &= \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p} \setminus E\}) + \sum_{k \in E} \mathrm{E}(g \circ y | h'_k \circ x) h'_k \\
&= \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p} \setminus E\}) \\
&\qquad + \sum_{k \in E} \Big(a\, \mathrm{E}(\mathrm{id}_{\mathbb{C}} \circ y | h'_k \circ x) + b\, \mathrm{E}(\mathbb{1}_{\mathbb{C}} \circ y | h'_k \circ x) \Big) h'_k.
\end{aligned}$$

Zatem

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\mathrm{lin}(\{h'_k\colon k\in\mathbb{Z}_{1,p}\setminus E\})\\ +a\sum_{k\in E}\mathrm{E}(\mathrm{id}_{\mathbb{C}}\circ y|h'_k\circ x)h'_k+b\sum_{k\in E}\mathrm{E}(\mathbb{1}_{\mathbb{C}}\circ y|h'_k\circ x)h'_k.$$

Ponadto dla każdego $k\in E$,

$$\mathrm{E}(\mathbb{1}_{\mathbb{C}}\circ y|h'_k\circ x)=\frac{\mathrm{E}((\mathbb{1}_{\mathbb{C}}\circ y)\overline{(h'_k\circ x)})}{\mathrm{E}(|h'_k\circ x|^2)}=\frac{\mathrm{E}((\mathbb{1}_A\circ x)\overline{(h'_k\circ x)})}{\mathrm{E}(|h'_k\circ x|^2)}=\frac{\langle\mathbb{1}_A\,|h'_k\rangle}{\|h'_k\|^2}.$$

W konsekwencji

$$\tilde{f}_0:=\sum_{k\in E}\mathrm{E}(\mathbb{1}_{\mathbb{C}}\circ y|h'_k\circ x)h'_k=\sum_{k\in E}\frac{\langle\mathbb{1}_A\,|h'_k\rangle}{\|h'_k\|^2}h'_k\in\mathrm{Pr}_{\mathcal{F}}(\mathbb{1}_{\mathrm{A}}),$$

gdyż $\tilde{f}_0\in\mathcal{F}$, $|\langle\mathbb{1}_A-\tilde{f}_0|h'_l\rangle|\leqslant\|\,\mathbb{1}_A-\tilde{f}_0\|\cdot\|h'_l\|=0$ dla $l\in\mathbb{Z}_{1,p}\setminus E$ oraz

$$\langle\mathbb{1}_A-\tilde{f}_0|h'_l\rangle=\langle\mathbb{1}_A\,|h'_l\rangle-\langle\tilde{f}_0|h'_l\rangle=\langle\mathbb{1}_A\,|h'_l\rangle-\sum_{k\in E}\frac{\langle\mathbb{1}_A\,|h'_k\rangle}{\|h'_k\|^2}\langle h'_k|h'_l\rangle=0,\quad l\in E.$$

Stosując ponownie twierdzenie 4.12 z $g:=\mathrm{id}_{\mathbb{C}}$, stwierdzamy, że

$$\tilde{f}_1:=\sum_{k\in E}\mathrm{E}(\mathrm{id}_{\mathbb{C}}\circ y|h'_k\circ x)h'_k\in\mathrm{Reg}(\mathcal{F},\mathfrak{P}).$$

Zatem

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\mathrm{lin}(\{h'_k\colon k\in\mathbb{Z}_{1,p}\setminus E\})+a\cdot\tilde{f}_1+b\cdot\tilde{f}_0. \tag{4.66}$$

Ustalmy dowolnie $f_1\in\mathrm{Reg}(\mathcal{F},\mathfrak{P})$ i $f_0\in\mathrm{Pr}_{\mathcal{F}}(\mathbb{1}_A)$. Wtedy $f_0-\tilde{f}_0\in\mathcal{F}$ oraz $\mathbb{1}_A-f_0\perp v$ i $\mathbb{1}_A-\tilde{f}_0\perp v$ dla $v\in\mathcal{F}$. Stąd

$$\begin{aligned}\|f_0-\tilde{f}_0\|^2&=\langle f_0-\tilde{f}_0|f_0-\tilde{f}_0\rangle=\langle(\mathbb{1}_A-\tilde{f}_0)-(\mathbb{1}_A-f_0)|f_0-\tilde{f}_0\rangle\\&=\langle\mathbb{1}_A-\tilde{f}_0|f_0-\tilde{f}_0\rangle-\langle\mathbb{1}_A-f_0|f_0-\tilde{f}_0\rangle=0,\end{aligned}$$

czyli $f_0-\tilde{f}_0\in\Theta\cap\mathcal{F}$. Z twierdzenia 3.15 wynika zaś, że $f_1-\tilde{f}_1\in\Theta\cap\mathcal{F}$. Zatem $(af_1+bf_0)-(a\tilde{f}_1+b\tilde{f}_0)\in\Theta\cap\mathcal{F}$, i w konsekwencji

$$\Theta\cap\mathcal{F}+af_1+bf_0=\Theta\cap\mathcal{F}+a\tilde{f}_1+b\tilde{f}_0. \tag{4.67}$$

Ponadto na mocy twierdzenia 4.3, $\mathrm{lin}(\{h'_k\colon k\in\mathbb{Z}_{1,p}\setminus E\})=\Theta\cap\mathcal{F}$. To łącznie z równościami (4.67) i (4.66) implikuje równość (4.65), gdy $E\neq\emptyset$.

Załóżmy na koniec, że $E=\emptyset$. Wtedy na mocy uwagi 4.1, $\mathcal{F}=\mathrm{lin}(\{h'_k\colon k\in\mathbb{Z}_{1,p}\setminus E\})$. Ponadto z twierdzenia 4.12 wynika, że $\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\mathcal{F}=\mathrm{Reg}(\mathcal{F},\mathfrak{P})$. Stąd dla dowolnych $f_1\in\mathrm{Reg}(\mathcal{F},\mathfrak{P})$ i $f_0\in\mathrm{Pr}_{\mathcal{F}}(\mathbb{1}_A)$, $af_1+bf_0\in\mathcal{F}$. Dlatego

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\mathcal{F}=\mathcal{F}+af_1+bf_0=\mathrm{lin}(\{h'_k\colon k\in\mathbb{Z}_{1,p}\setminus E\})+af_1+bf_0,$$

czyli zachodzi równość (4.65), gdy $E=\emptyset$. To kończy dowód twierdzenia. □

Wniosek 4.18. *Dla dowolnie zadanych* $p \in \mathbb{N}$ *i* $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ *niech* $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ *będzie ciągiem zdefiniowanym wzorem* (4.34). *Wtedy dla dowolnych* $a, b \in \mathbb{C}$, *jeśli* $g = a \cdot \mathrm{id}_{\mathbb{C}} + b$, $\mathrm{id}_{\mathbb{C}} \in \mathrm{L}^2(\mathbb{C}, \mathcal{A}_y, P_y)$ *i* $\mathbb{1}_A \in \mathcal{F}$, *to*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathrm{lin}(\{h'_k \colon k \in \mathbb{Z}_{1,p} \setminus E\}) + a \cdot f + b \cdot \mathbb{1}_A, \quad f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}), \tag{4.68}$$

gdzie $\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$ *i* $E := \{k \in \mathbb{Z}_{1,p} \colon \|h'_k\| > 0\}$.

Dowód. Wniosek jest bezpośrednią konsekwencją twierdzenia 4.17, gdyż na mocy założenia $\mathbb{1}_A \in \mathcal{F}$, $\mathbb{1}_A \in \mathrm{Pr}_{\mathcal{F}}(\mathbb{1}_A)$.

□

Rozdział 5

Przykłady zastosowań i uwagi uzupełniające

Opierając się na twierdzeniu 4.10 wyprowadzimy metodę numeryczną obliczania funkcji regresji dla dowolnie zadanego układu funkcji bazowych $\mathbb{Z}_{1,p} \ni k \mapsto h_k$. W dalszej części podamy kilka realnych przykładów wyznaczania funkcji regresji opartych na implementacji tej metody w arkuszu kalkulacyjnym MS Excel przy użyciu języka programowania VBA.
Ustalając dowolnie $p \in \mathbb{N}$ i ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ rozważmy macierze

$$\mathbb{Z}_{1,p} \times \mathbb{Z}_{1,p} \ni (k,l) \mapsto \mathrm{R}_{k,l} := \langle h_k | h'_l \rangle \tag{5.1}$$

oraz

$$\mathbb{Z}_{1,p} \ni k \mapsto \mathrm{Q}_k := g^*(h'_k), \tag{5.2}$$

gdzie ciąg $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ jest określony wzorem (4.6). Załóżmy, że $\|h'_k\| > 0$ dla $k \in \mathbb{Z}_{1,p}$. Ponieważ $h'_1 = h_1$, więc ze wzorów (5.1) i (5.2) wynika, że

$$\mathrm{R}_{k,1} = \langle h_k | h_1 \rangle \quad \text{i} \quad \mathrm{Q}_1 = g^*(h_1), \qquad k \in \mathbb{Z}_{1,p}. \tag{5.3}$$

Ponadto na mocy wzoru (4.6) widzimy, że dla dowolnych $k \in \mathbb{Z}_{1,p}$ i $l \in \mathbb{Z}_{2,p}$.

$$\begin{aligned}\mathrm{R}_{k,l} = \langle h_k | h'_l \rangle &= \left\langle h_k \middle| h_l - \sum_{j=1}^{l-1} \langle h_l | h'_j \rangle_* h'_j \right\rangle = \langle h_k | h_l \rangle - \sum_{j=1}^{l-1} \overline{\langle h_l | h'_j \rangle_*} \langle h_k | h'_j \rangle \\ &= \langle h_k | h_l \rangle - \sum_{j=1}^{l-1} \frac{\overline{\langle h_l | h'_j \rangle} \langle h_k | h'_j \rangle}{\|h'_j\|^2} = \langle h_k | h_l \rangle - \sum_{j=1}^{l-1} \frac{\overline{\mathrm{R}_{l,j}} \mathrm{R}_{k,j}}{\mathrm{R}_{j,j}},\end{aligned}$$

gdyż na mocy uwagi 4.2,

$$\|h'_j\|^2 = \langle h_j | h'_j \rangle = \mathrm{R}_{j,j}, \quad j \in \mathbb{Z}_{2,p}. \tag{5.4}$$

To prowadzi do formuły rekurencyjnej

$$\mathrm{R}_{k,l} = \langle h_k | h_l \rangle - \sum_{j=1}^{l-1} \frac{\mathrm{R}_{k,j}}{\mathrm{R}_{j,j}} \overline{\mathrm{R}_{l,j}}, \qquad k \in \mathbb{Z}_{1,p},\ l \in \mathbb{Z}_{2,p}. \tag{5.5}$$

Stosując ponownie wzór (4.6) dostajemy dla każdego $k \in \mathbb{Z}_{2,p}$,

$$g^*(h'_k) = g^*\left(h_k - \sum_{j=1}^{k-1} \langle h_k | h'_j \rangle_* h'_j\right) = g^*(h_k) - \sum_{j=1}^{k-1} \langle h_k | h'_j \rangle_* g^*(h'_j),$$

co wobec (5.1), (5.2) i (5.4) prowadzi do zależności rekurencyjnej

$$\mathrm{Q}_k = g^*(h_k) - \sum_{j=1}^{k-1} \frac{\mathrm{R}_{k,j}}{\mathrm{R}_{j,j}} \mathrm{Q}_j, \qquad k \in \mathbb{Z}_{2,p}. \tag{5.6}$$

Z uwagi 4.1 wynika ponadto, że

$$\mathrm{R}_{k,l} = 0, \qquad k \in \mathbb{Z}_{1,p-1},\ l \in \mathbb{Z}_{k+1,p}, \tag{5.7}$$

gdyż $h'_l \perp h_k$ dla $k \in \mathbb{Z}_{1,l-1}$. Na mocy twierdzenia 4.10 dostajemy

$$f := \sum_{k=1}^{p} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g),$$

gdzie $\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})$. Z uwagi 4.1 oraz z (5.2) i (5.4) wynika, że

$$f = \sum_{k=1}^{p} \frac{\overline{\mathrm{Q}_k}}{\mathrm{R}_{k,k}} h'_k = \sum_{k=1}^{p} \alpha_k h_k,$$

dla pewnego ciągu $\mathbb{Z}_{1,p} \ni k \mapsto \alpha_k \in \mathbb{C}$. Stąd dla każdego $l \in \mathbb{Z}_{1,p}$,

$$\overline{\mathrm{Q}_l} = \left\langle \sum_{k=1}^{p} \frac{\overline{\mathrm{Q}_k}}{\mathrm{R}_{k,k}} h'_k \middle| h'_l \right\rangle = \left\langle \sum_{k=1}^{p} \alpha_k h_k \middle| h'_l \right\rangle = \sum_{k=1}^{p} \alpha_k \langle h_k | h'_l \rangle = \sum_{k=1}^{p} \alpha_k \mathrm{R}_{k,l}.$$

Uwzględniając zależności (5.7) dostajemy ostatecznie

$$\alpha_p = \frac{\overline{\mathrm{Q}_p}}{\mathrm{R}_{p,p}} \qquad \text{oraz} \qquad \alpha_l = \frac{1}{\mathrm{R}_{l,l}} \left(\overline{\mathrm{Q}_l} - \sum_{k=l+1}^{p} \alpha_k \mathrm{R}_{k,l}\right), \qquad l \in \mathbb{Z}_{1,p-1}. \tag{5.8}$$

Stosując wzory początkowe (5.3) oraz formuły rekurencyjne (5.5), (5.6) i (5.8) możemy bezpośrednio wyliczyć współczynniki α_k, $k \in \mathbb{Z}_{1,p}$, funkcji regresji f w bazie $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ modelu funkcyjnego $\mathcal{F}$. W oparciu o formuły (5.3), (5.5), (5.6) i (5.8) można zaprojektować program komputerowy obliczający funkcję regresji. Na potrzeby ilustracji działania tego algorytmu został opracowany przykładowy program komputerowy dla probabilistycznych struktur regresji $\mathfrak{P} := (A, \mathbb{R}, \delta; x, y)$ nad przestrzenią probabilistyczną $\mathcal{P} := (\Omega, \mathcal{A}, P)$, gdzie $\Omega := \mathbb{Z}_{1,n}$ dla pewnego $n \in \mathbb{N}$, $\mathcal{A} := 2^\Omega$, $x \colon \mathbb{Z}_{1,n} \to A \subset \mathbb{R}$, $y \colon \mathbb{Z}_{1,n} \to \mathbb{R}$ i kryterium dopasowania δ spełnia równość (3.1) dla wszystkich $u, v \colon \mathbb{Z}_{1,n} \to \mathbb{R}$. Wówczas dla dowolnych funkcji $h \colon A \to \mathbb{R}$ i $g \colon \mathbb{R} \to \mathbb{R}$ zachodzi równość

$$\delta(h \circ x, g \circ y) = \sum_{k=1}^{n} |h(x_k) - g(y_k)|^2 p_k,$$

gdzie $x_k := x(k)$, $y_k := y(k)$ i $p_k := P(\{k\})$ dla $k \in \mathbb{Z}_{1,n}$. Tego typu struktury regresji mają zastosowanie w analizie szeregów czasowych, co pokazują przykłady 5.1 – 5.4. Prawdopodobieństwo p_k jest tutaj interpretowane jako waga istotności k-tego pomiaru. Program komputerowy został zrealizowany w języku programowania VBA interpretowanym z poziomu arkusza kalkulacyjnego MS Excel, który sam w sobie jest wygodnym narzędziem do przetwarzania danych. Tak przygotowany program komputerowy posłużył do wyznaczenia funkcji regresji w przykładach 5.1 – 5.4. Kod programu znajduje się w załączniku na końcu pracy. Są tam również zamieszczone wyniki obliczeń programu MS Excel dla tych przykładów. Pojawiające się w przykładach równości należy traktować w sposób numeryczny, czyli z ewentualnym błędem zaokrągleń. Wyniki obliczeń podawane są z dokładnością do 10^{-5}.

Przykład 5.1. Załóżmy, że mamy dane umowne dla 5 obserwacji przedstawione w tabeli 5.1.

Tabela 5.1: Przykładowe dane umowne

p_k	0,2	0,2	0,2	0,2	0,2
x_k	-2	-1	0	1	2
y_k	1	0	1	4	9

Rozważmy teoretyczny model funkcyjny

$$\mathcal{F} := \mathrm{lin}(\{h_0,\ h_1,\ h_2\}), \tag{5.9}$$

gdzie $\mathbb{R} \ni t \mapsto h_l(t) := t^{l-1}$, $l \in \mathbb{Z}_{1,3}$. Stosując zaimplementowany program komputerowy otrzymujemy następujące wyniki (por. rysunek 6.1):

- Macierz iloczynów pseudo-skalarnych

$$\begin{bmatrix} \langle h_1|h_1\rangle & \langle h_1|h_2\rangle & \langle h_1|h_3\rangle \\ \langle h_2|h_1\rangle & \langle h_2|h_2\rangle & \langle h_2|h_3\rangle \\ \langle h_3|h_1\rangle & \langle h_3|h_2\rangle & \langle h_3|h_3\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 6{,}8 \end{bmatrix};$$

- Macierz wartości dla funkcji skalującej $\mathbb{R} \ni t \mapsto g(t) := t$,

$$\begin{bmatrix} g^*(h_1) & g^*(h_2) & g^*(h_3) \end{bmatrix} = \begin{bmatrix} 3 & 4 & 8{,}8 \end{bmatrix};$$

- Macierz R

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & 0 & 2{,}8 \end{bmatrix};$$

- Macierz Q

$$\begin{bmatrix} Q_1 & Q_2 & Q_3 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 2{,}8 \end{bmatrix};$$

- Macierz współczynników regresji $\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$.

Zatem

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto 1 + 2t + t^2\}, \tag{5.10}$$

czyli funkcja regresji pokrywa się z wielomianem interpolacyjnym co najwyżej drugiego stopnia dla danych z tabeli 5.1.

Jeżeli zmodyfikujemy przykład zamieniając rozkład jednostajny prawdopodobieństwa przez rozkład zadany w tabeli 5.2, otrzymamy przy takim samym modelu

Tabela 5.2: Przykładowe dane umowne

p_k	0,01	0,05	0,4	0,3	0,24
x_k	-2	-1	0	1	2
y_k	1	0	1	4	9

teoretycznym (5.9) następujące wyniki (por. rysunek 6.2):

- Macierz iloczynów pseudo-skalarnych

$$\begin{bmatrix} \langle h_1|h_1\rangle & \langle h_1|h_2\rangle & \langle h_1|h_3\rangle \\ \langle h_2|h_1\rangle & \langle h_2|h_2\rangle & \langle h_2|h_3\rangle \\ \langle h_3|h_1\rangle & \langle h_3|h_2\rangle & \langle h_3|h_3\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0{,}71 & 1{,}35 \\ 0{,}71 & 1{,}35 & 2{,}09 \\ 1{,}35 & 2{,}09 & 4{,}35 \end{bmatrix};$$

- Macierz wartości dla funkcji skalującej $\mathbb{R} \ni t \mapsto g(t) := t$,

$$\begin{bmatrix} g^*(h_1) & g^*(h_2) & g^*(h_3) \end{bmatrix} = \begin{bmatrix} 3{,}77 & 5{,}5 & 9{,}88 \end{bmatrix};$$

- Macierz R

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0{,}71 & 0{,}8459 & 0 \\ 1{,}35 & 1{,}1315 & 1{,}01397 \end{bmatrix};$$

- Macierz Q

$$\begin{bmatrix} Q_1 & Q_2 & Q_3 \end{bmatrix} = \begin{bmatrix} 3{,}77 & 2{,}8233 & 1{,}01397 \end{bmatrix};$$

- Macierz współczynników regresji $\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$.

Zatem

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto 1 + 2t + t^2\}, \tag{5.11}$$

czyli w obu przypadkach funkcje regresji pokrywają się z wielomianem interpolacyjnym dla ciągów x i y; por. równość (5.10). Wyjaśnienie tego faktu znajduje się w przykładzie 3.19.

Przyjmijmy teraz funkcję skalującą postaci $\mathbb{R} \ni t \mapsto g(t) := t^2$ dla danych przedstawionych w tabeli 5.2. Nie zmieniając teoretycznego modelu funkcyjnego (5.9), otrzymujemy wyniki (por. rysunek 6.3):

- Macierz iloczynów pseudo-skalarnych

$$\begin{bmatrix} \langle h_1|h_1\rangle & \langle h_1|h_2\rangle & \langle h_1|h_3\rangle \\ \langle h_2|h_1\rangle & \langle h_2|h_2\rangle & \langle h_2|h_3\rangle \\ \langle h_3|h_1\rangle & \langle h_3|h_2\rangle & \langle h_3|h_3\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0{,}71 & 1{,}35 \\ 0{,}71 & 1{,}35 & 2{,}09 \\ 1{,}35 & 2{,}09 & 4{,}35 \end{bmatrix};$$

- Macierz wartości dla funkcji skalującej g,

$$\begin{bmatrix} g^*(h_1) & g^*(h_2) & g^*(h_3) \end{bmatrix} = \begin{bmatrix} 24{,}65 & 43{,}66 & 82{,}6 \end{bmatrix};$$

- Macierz R

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0{,}71 & 0{,}8459 & 0 \\ 1{,}35 & 1{,}1315 & 1{,}01397 \end{bmatrix};$$

- Macierz Q

$$\begin{bmatrix} Q_1 & Q_2 & Q_3 \end{bmatrix} = \begin{bmatrix} 24{,}65 & 26{,}1585 & 14{,}33214 \end{bmatrix};$$

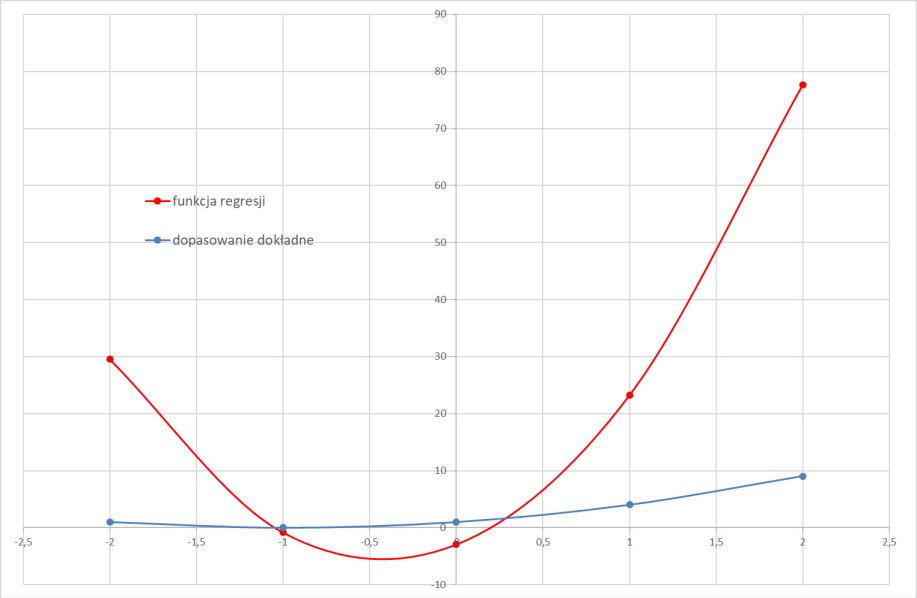
- Macierz współczynników regresji

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} = \begin{bmatrix} -2{,}96381 & 12{,}01698 & 14{,}13464 \end{bmatrix}.$$

Zatem

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto -2{,}96381 + 12{,}01698t + 14{,}13464t^2\}, \tag{5.12}$$

czyli funkcja regresji znacznie odbiega od wielomianu interpolacyjnego dla danych z tabeli 5.2, por. rysunek 5.1.



Rysunek 5.1: Wielomian interpolacyjny (dopasowanie dokładne) z zaznaczoną krzywą regresji (5.12).

Przykład 5.2 (Dane giełdowe)**.** Wyznaczymy regresję dla danych dotyczących cotygodniowego otwarcia giełdy papierów wartościowych WIG w Warszawie od początku jej notowań, czyli od 21 kwietnia 1991 roku do 05 września 2021 roku. Dane obejmują 1584 tygodni i pochodzą ze strony `https://stooq.pl/q/d/?s=wig`. Zakładamy, że wszystkie notowania są jednakowo istotne, tzn. $p_k := \frac{1}{1584}$, $k \in \mathbb{Z}_{1,1584}$, zaś funkcja skalująca jest postaci $\mathbb{R} \ni t \mapsto g(t) := t$. Rozważmy na początek teoretyczny model funkcyjny postaci $\mathcal{F}_1 := \mathrm{lin}(\{h_l \colon l \in \mathbb{Z}_{1,2}\})$, gdzie $\mathbb{R} \ni t \mapsto h_l(t) := t^{l-1}$, $l \in \mathbb{Z}_{1,2}$, czyli rodzina $\mathcal{F}_1$ składa się z wielomianów rzeczywistych stopnia co najwyżej pierwszego. Stosując zaimplementowany program komputerowy wyznaczamy klasę funkcji regresji (por. rysunek 6.4)

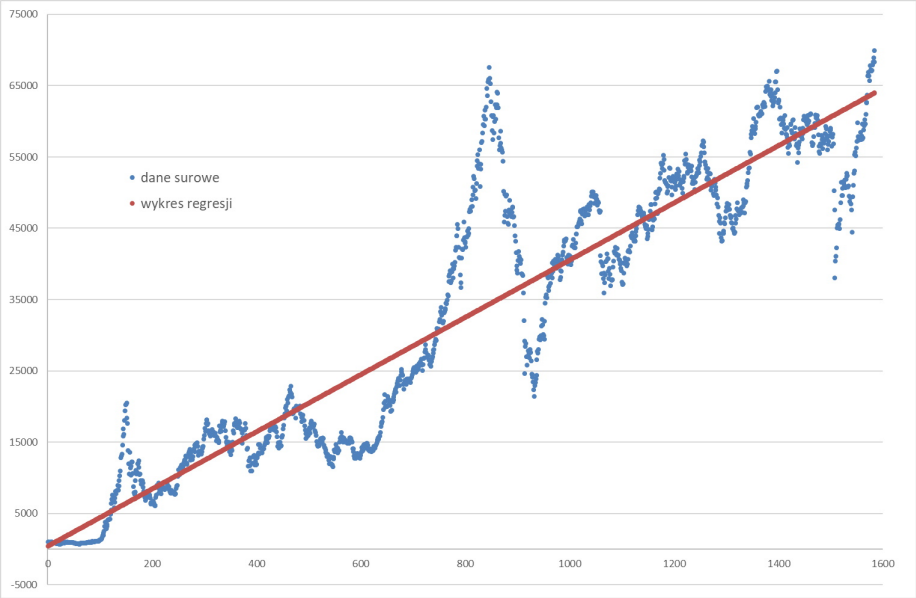
$$\mathrm{Reg}(\mathcal{F}_1, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto 406{,}84816 + 40{,}14188t\}. \tag{5.13}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.2.

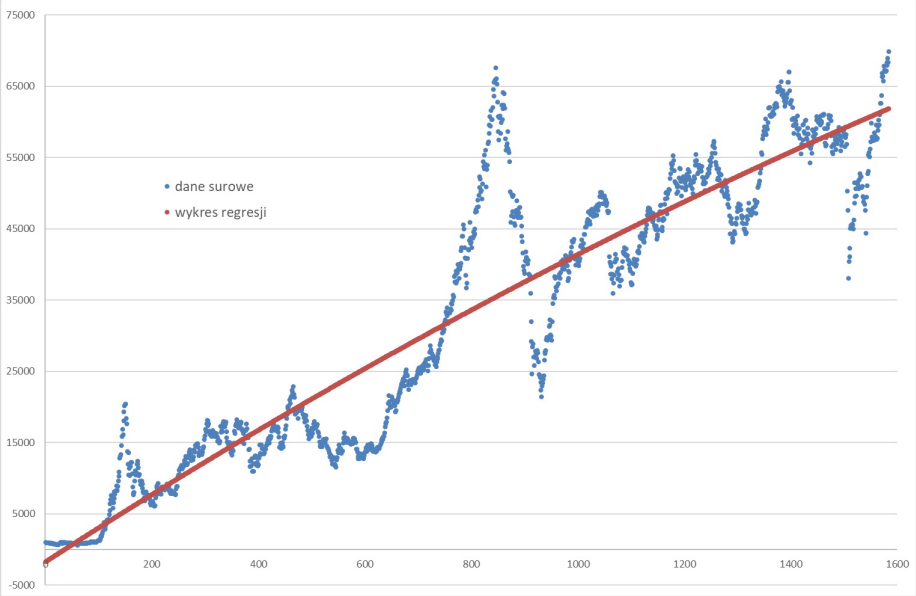
Przyjmijmy teraz $\mathcal{F}_2 := \mathrm{lin}(\{h_l \colon l \in \mathbb{Z}_{1,3}\})$, gdzie $\mathbb{R} \ni t \mapsto h_l(t) := t^{l-1}$ dla $l \in \mathbb{Z}_{1,3}$, czyli model teoretyczny $\mathcal{F}_2$ składa się z wielomianów rzeczywistych stopnia co najwyżej drugiego. Wtedy klasa funkcji regresji (por. rysunek 6.5) jest postaci

$$\mathrm{Reg}(\mathcal{F}_2, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto -1740{,}74105 + 48{,}26643t - 0{,}00513t^2\}. \tag{5.14}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.3.



Rysunek 5.2: Dane surowe z zaznaczoną krzywą regresji (5.13).



Rysunek 5.3: Dane surowe z zaznaczoną krzywą regresji (5.14).

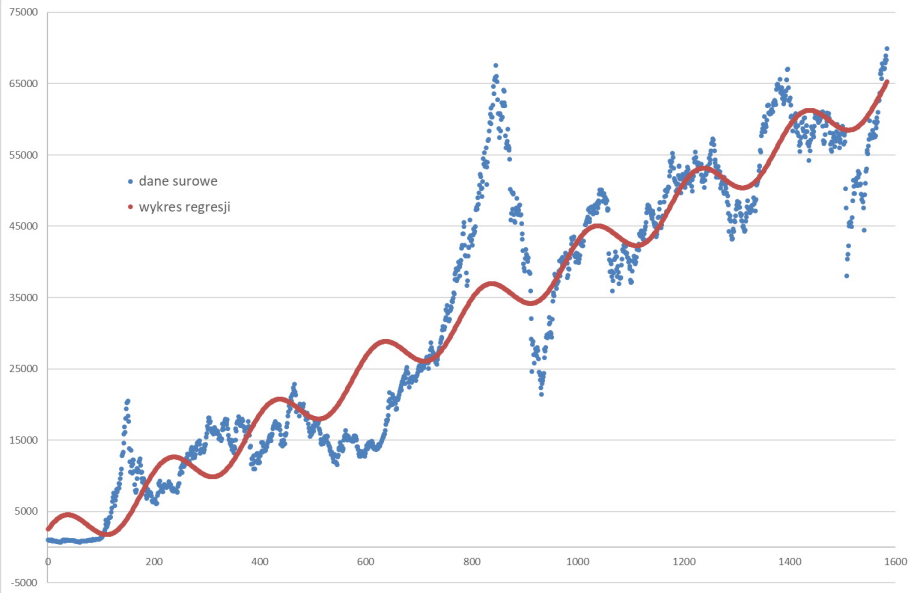
Można zauważyć, że trend notowań giełdy WIG jest rosnący, jednakże dynamika tego wzrostu jest malejąca.

Dane giełdowe cechują się również wahaniami okresowymi. Załóżmy więc, że model teoretyczny oprócz liniowego trendu wielomianowego posiada także trend okresowy. Przyjmijmy, że okres liczy 200 tygodni, czyli około cztery lata. Wtedy teoretyczny model funkcyjny ma postać $\mathcal{F}_3 := \mathrm{lin}\,(\{h_k\colon k \in \mathbb{Z}_{1,4}\})$, gdzie $h_1(t) := 1$, $h_2(t) := t$, $h_3(t) := \sin\left(\frac{\pi}{100}t\right)$, $h_4(t) := \cos\left(\frac{\pi}{100}t\right)$ dla $t \in \mathbb{R}$. Wówczas klasa funkcji regresji jest następująca (por. rysunek 6.6)

(5.15)

$$\mathrm{Reg}(\mathcal{F}_3, \mathfrak{P}_g) = \Big\{\mathbb{R} \ni t \mapsto 109{,}60675 + 40{,}53825t + 2192{,}7395 \sin\left(\frac{\pi}{100}t\right) + 2263{,}31969 \cos\left(\frac{\pi}{100}t\right)\Big\}.$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.4.



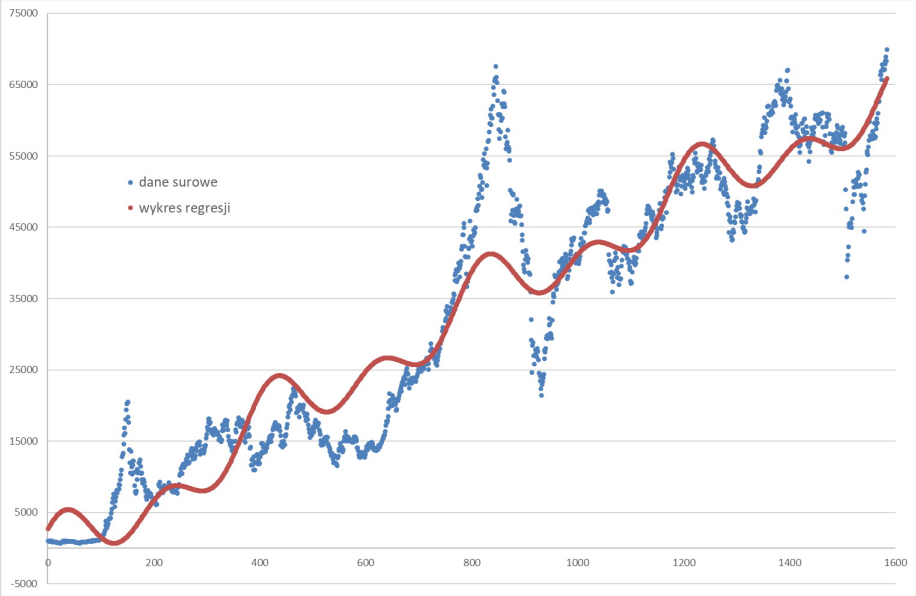
Rysunek 5.4: Dane surowe z zaznaczoną krzywą regresji (5.15).

Rozważmy teraz bardziej złożony teoretyczny model funkcyjny $\mathcal{F}_4 := \mathrm{lin}(\{h_k\colon k \in \mathbb{Z}_{1,7}\})$, gdzie $h_1(t) := 1$, $h_2(t) := t$, $h_3(t) := t^2$, $h_4(t) := \sin\left(\frac{\pi}{100}t\right)$, $h_5(t) := \cos\left(\frac{\pi}{100}t\right)$, $h_6(t) := \sin\left(\frac{\pi}{200}t\right)$, $h_7(t) := \cos\left(\frac{\pi}{200}t\right)$ dla $t \in \mathbb{R}$. W stosunku do poprzedniego modelu został uwzględniony trend wielomianowy drugiego stopnia oraz dodatkowo okres 400 tygodniowy. Wówczas klasa funkcji regresji jest następująca (por. rysunek 6.7)

(5.16)

$$\mathrm{Reg}(\mathcal{F}_4, \mathfrak{P}_g) = \Big\{\mathbb{R} \ni t \mapsto -2521{,}95916 + 49{,}50141t - 0{,}00531t^2 + 2219{,}01987 \sin\left(\frac{\pi}{100}t\right) + 2300{,}0164 \cos\left(\frac{\pi}{100}t\right) + 1456{,}70906 \sin\left(\frac{\pi}{200}t\right) + 2801{,}64332 \cos\left(\frac{\pi}{200}t\right)\Big\}.$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.5.



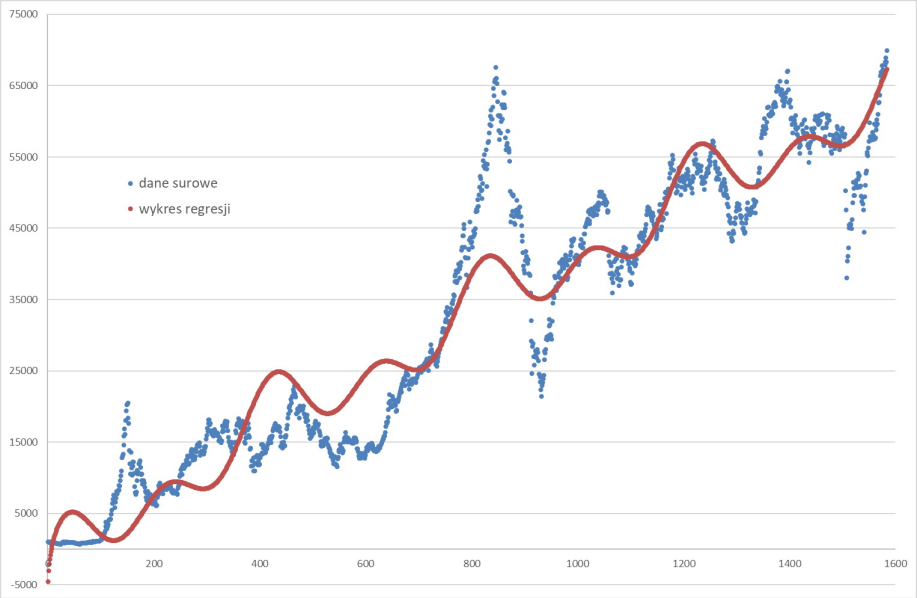
Rysunek 5.5: Dane surowe z zaznaczoną krzywą regresji (5.16).

Zmodyfikujmy teraz model teoretyczny $\mathcal{F}_4$ przez zastąpienie trendu kwadratowego trendem logarytmicznym. Wtedy teoretyczny model funkcyjny $\mathcal{F}_5 := \operatorname{lin}(\{h_k \colon k \in \mathbb{Z}_{1,7}\})$ jest zadany przez następujące funkcje bazowe $h_1(t) := 1$, $h_2(t) := t$, $h_3(t) := \ln t$, $h_4(t) := \sin\left(\frac{\pi}{100}t\right)$, $h_5(t) := \cos\left(\frac{\pi}{100}t\right)$, $h_6(t) := \sin\left(\frac{\pi}{200}t\right)$, $h_7(t) := \cos\left(\frac{\pi}{200}t\right)$ dla $t \in A := (0; +\infty)$. W tym przypadku klasa funkcji regresji spełnia równość (por. rysunek 6.8)

(5.17)

$$\begin{aligned}\operatorname{Reg}(\mathcal{F}_5, \mathfrak{P}_g) = \Big\{ A \ni t \mapsto\ & -10132{,}43738 + 37{,}37466t + 2005{,}83088 \ln t \\ & + 2339{,}39068 \sin\left(\frac{\pi}{100}t\right) + 2424{,}97173 \cos\left(\frac{\pi}{100}t\right) \\ & + 1597{,}76804 \sin\left(\frac{\pi}{200}t\right) + 2983{,}52543 \cos\left(\frac{\pi}{200}t\right) \Big\}.\end{aligned}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.6.



Rysunek 5.6: Dane surowe z zaznaczoną krzywą regresji (5.17).

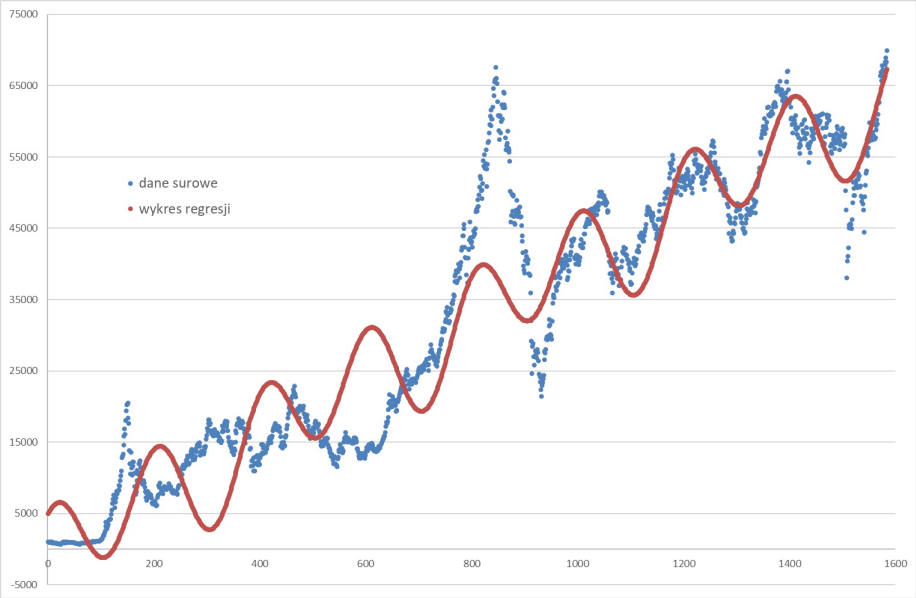
Stosując przedstawioną teorię możemy rozważać także inny rozkład prawdopodobieństwa niż rozkład jednostajny. Każdemu tygodniowi w poprzednich przypadkach przyporządkowaliśmy jednakową wagę istotności. Przyjmijmy teraz dla odmiany inny rozkład prawdopodobieństwa. Załóżmy znikomy wpływ danych historycznych, czyli prawdopodobieństwo dla początkowych 936 tygodni jest bardzo małe i równe 0,01. Załóżmy także duży wpływ ostatnich 24 tygodni przyjmując prawdopodobieństwo równe 0,5. Pozostałe tygodnie dzielimy na trzy równe okresy po 208 tygodni przyjmując prawdopodobieństwo równe odpowiednio 0,04, 0,15 i 0,3 w kolejności historycznej. Precyzyjniej rzecz ujmując rozkład prawdopodobieństwa jest postaci

$$p_k := \begin{cases} \frac{0{,}01}{936}, & k \in \mathbb{Z}_{1,936}, \\ \frac{0{,}04}{208}, & k \in \mathbb{Z}_{937,1144}, \\ \frac{0{,}15}{208}, & k \in \mathbb{Z}_{1145,1352}, \\ \frac{0{,}3}{208}, & k \in \mathbb{Z}_{1353,1560}, \\ \frac{0{,}5}{24}, & k \in \mathbb{Z}_{1561,1584}. \end{cases}$$

Wtedy klasa funkcji regresji dla np. modelu teoretycznego $\mathcal{F}_4$ spełnia równość (por. rysunek 6.9)

$$\begin{aligned} \mathrm{Reg}(\mathcal{F}_4, \mathfrak{P}_g) = \Big\{ \mathbb{R} \ni t \mapsto\ & -1234{,}61264 + 42{,}49254t - 0{,}00097t^2 \\ & + 2281{,}38964 \sin\left(\frac{\pi}{100}t\right) + 6393{,}33686 \cos\left(\frac{\pi}{100}t\right) \\ & + 2231{,}2263 \sin\left(\frac{\pi}{200}t\right) - 316{,}35312 \cos\left(\frac{\pi}{200}t\right) \Big\}. \end{aligned} \tag{5.18}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.7.



Rysunek 5.7: Dane surowe z zaznaczoną krzywą regresji (5.18).

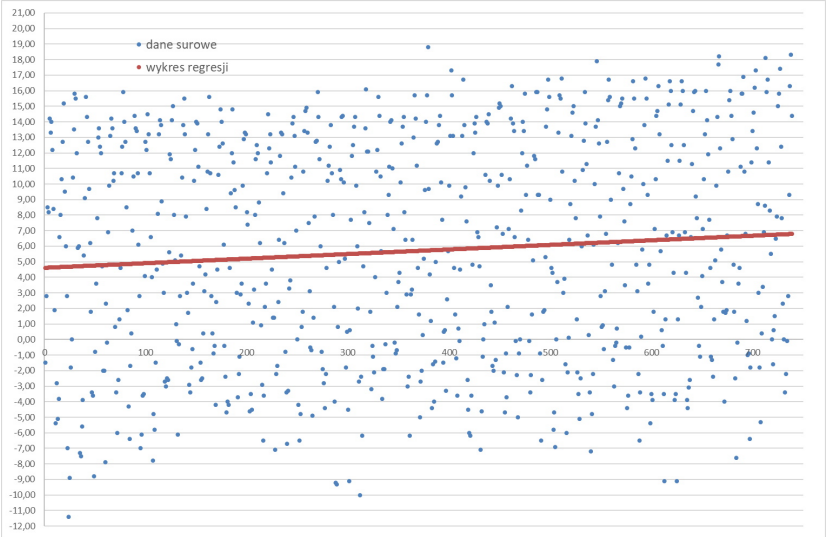
Porównując rysunki 5.4, 5.5 i 5.7 można zauważyć, że funkcja regresji w (5.18) jest zgodnie z oczekiwaniami lepiej dopasowana do danych niezbyt odległych w czasie.

Przykład 5.3 (Dane meteorologiczne - temperatura w miesiącach)**.** Wyznaczymy regresję dla danych meteorologicznych dotyczących średnich miesięcznych temperatur w Zakopanem. Dane pochodzą ze strony `https://dane.imgw.pl/data/dane_pomiarowo_obserwacyjne/dane_meteorologiczne/miesieczne/synop/` i dotyczą okresu od lutego 1961 roku do sierpnia 2021 roku, czyli obejmują 739 miesięcy. Przyjmujemy, że wszystkie pomiary są jednakowo istotne, tzn. obserwacje są traktowane z jednakową wagą prawdopodobieństwa $p_k := \frac{1}{739}$, $k \in \mathbb{Z}_{1,739}$, zaś funkcja skalująca jest postaci $\mathbb{R} \ni t \mapsto g(t) := t$. Rozważmy teoretyczny model funkcyjny postaci $\mathcal{F}_1 := \operatorname{lin}(\{h_j \colon l \in \mathbb{Z}_{1,2}\})$, gdzie $\mathbb{R} \ni t \mapsto h_l(t) := t^{l-1}$ dla $l \in \mathbb{Z}_{1,2}$. Stosując zaimplementowany program komputerowy otrzymujemy klasę funkcji regresji (por. rysunek 6.10)

$$\operatorname{Reg}(\mathcal{F}_1, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto 4{,}61255 + 0{,}00295t\}. \tag{5.19}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.8.

Średnia miesięczna temperatura cechuje się także wahaniami okresowymi. Załóżmy więc, że model teoretyczny oprócz wielomianów rzeczywistych stopnia co najwyżej drugiego składa się także z dziesięcioletniego trendu okresowego. Zatem teoretyczny model funkcyjny jest postaci $\mathcal{F}_2 := \operatorname{lin}\{h_k \colon k \in \mathbb{Z}_{1,5}\}$, gdzie $h_1(t) := 1$, $h_2(t) := t$, $h_3(t) := t^2$, $h_4(t) := \sin\left(\frac{\pi}{60}t\right)$, $h_5(t) := \cos\left(\frac{\pi}{60}t\right)$ dla $t \in \mathbb{R}$. Klasa funkcji

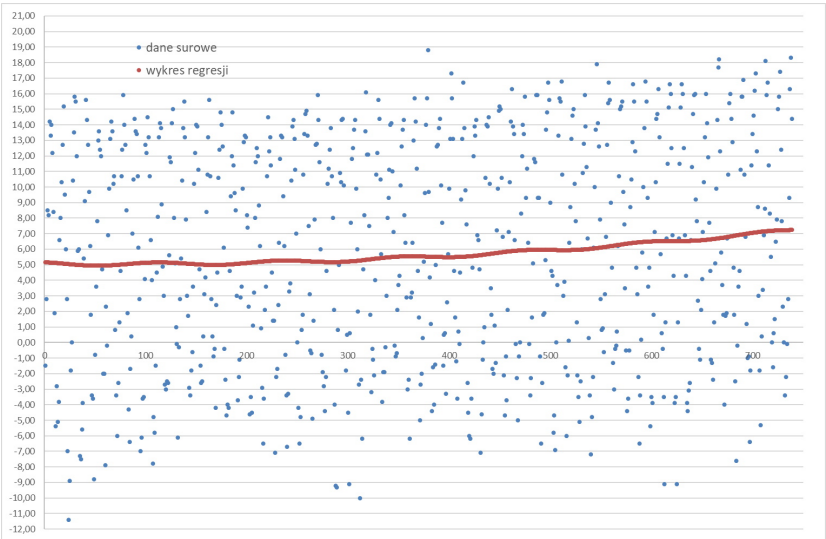


Rysunek 5.8: Dane surowe z zaznaczoną krzywą regresji (5.19).

regresji spełnia wówczas równość (por. rysunek 6.11)

$$\text{(5.20)}\qquad \operatorname{Reg}(\mathcal{F}_2,\mathfrak{P}_g)=\Big\{\mathbb{R}\ni t\mapsto 5{,}06818-0{,}0007t+4{,}88\cdot 10^{-6}t^2 - 0{,}0415\sin\left(\frac{\pi}{60}t\right)+0{,}09326\cos\left(\frac{\pi}{60}t\right)\Big\}.$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.9.



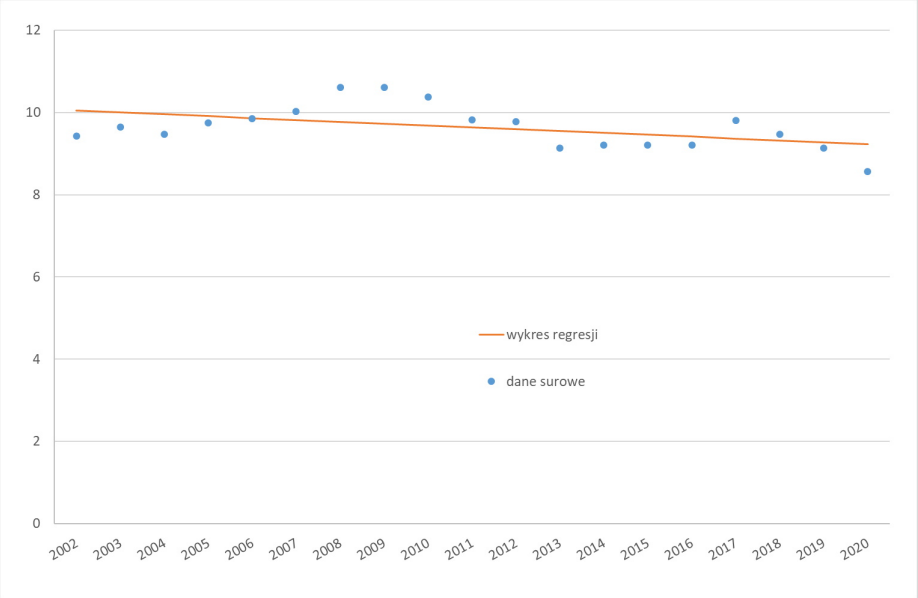
Rysunek 5.9: Dane surowe z zaznaczoną krzywą regresji (5.20).

Z przeprowadzonych obserwacji można zauważyć, że temperatura w ciągu ostatnich 50 lat w Zakopanem wzrosła średnio o 2°C, co świadczy o ociepleniu w tym rejonie kraju.

Przykład 5.4 (Współczynnik urodzeń)**.** Wyznaczymy regresję dla rocznego współczynnika urodzeń na 1000 mieszkańców w województwie lubelskim w latach 2002 - 2020. Współczynnik urodzeń jest to stosunek liczby urodzeń żywych w danym okresie (roku) do średniej liczby ludności z tego okresu - wyrażony w ‰ (tj. na 1000 ludności). Dane pochodzą ze strony Głównego Urzędu Statystycznego `https://bdl.stat.gov.pl/BDL/dane/podgrup/tablica` i dotyczą ostatnich 19 lat. Przyjmujemy, że wszystkie dane są jednakowo ważne, tzn. dane są traktowane z jednakową wagą prawdopodobieństwa $p_k := \frac{1}{19}$, $k \in \mathbb{Z}_{1,19}$, zaś funkcja skalująca jest postaci $\mathbb{R} \ni t \mapsto g(t) := t$. Rozważmy teoretyczny model funkcyjny $\mathcal{F}_1 := \mathrm{lin}(\{h_j \colon l \in \mathbb{Z}_{1,2}\})$, gdzie $\mathbb{R} \ni t \mapsto h_l(t) := t^{l-1}$ dla $l \in \mathbb{Z}_{1,2}$. Stosując zaimplementowany program komputerowy wyznaczamy klasę funkcji regresji (por. rysunek 6.12)

$$\mathrm{Reg}(\mathcal{F}_1, \mathfrak{P}_g) = \{\mathbb{R} \ni t \mapsto 10{,}0893 - 0{,}04519t\}. \tag{5.21}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.10.

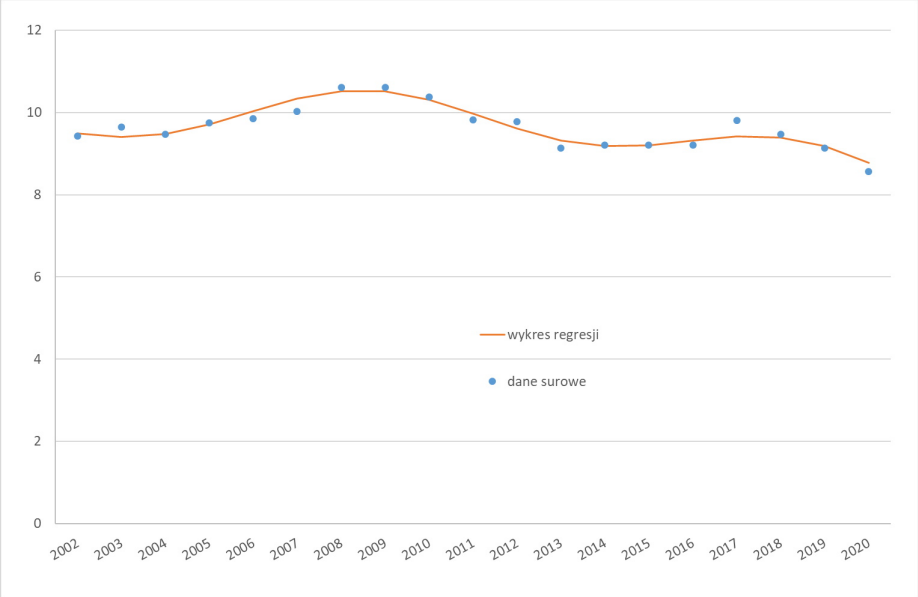


Rysunek 5.10: Dane surowe z zaznaczoną krzywą regresji (5.21).

Na podstawie wykresu danych surowych możemy zauważyć, że współczynnik urodzeń cechuje się także wahaniami okresowymi. Załóżmy więc, że model teoretyczny oprócz wielomianów rzeczywistych stopnia co najwyżej drugiego składa się także z dziesięcioletniego trendu okresowego. Zatem teoretyczny model funkcyjny jest postaci $\mathcal{F}_2 := \operatorname{lin}\{h_l \colon l \in \mathbb{Z}_{1,5}\}$, gdzie $h_1(t) := 1$, $h_2(t) := t$, $h_3(t) := t^2$, $h_4(t) := \sin\left(\frac{\pi}{5}t\right)$, $h_5(t) := \cos\left(\frac{\pi}{5}t\right)$ dla $t \in \mathbb{R}$. Klasa funkcji regresji spełnia wówczas równość (por. rysunek 6.13)

$$\operatorname{Reg}(\mathcal{F}_2, \mathfrak{P}_g) = \Big\{\mathbb{R} \ni t \mapsto 9{,}62476 + 0{,}13583t - 0{,}01034t^2 - 0{,}47856 \sin\left(\frac{\pi}{5}t\right) + 0{,}02055 \cos\left(\frac{\pi}{5}t\right)\Big\}. \tag{5.22}$$

Wykres danych surowych z zaznaczoną krzywą regresji przedstawia rysunek 5.11.



Rysunek 5.11: Dane surowe z zaznaczoną krzywą regresji (5.22).

Załącznik

Poniższy kod programu został napisany w języku programowania VBA w środowisku arkusza kalkulacyjnego MS Excel. Jest on implementacją metody obliczania regresji opisanej w rozdziale piątym. W zależności od przykładów były zmieniane: wartość stałej `WymReg` oraz kod dwóch funkcji `FunkBaz` i `FunkSkal`.

```
Public Const Ark_Dane = "Dane"
Public Const Ark_Oblicz = "Obliczenia"
Public Const Ark_Wyk = "Wykresy"
Public Const WymReg = 3 'Liczba funkcji bazowych $h_k$.
Dim MacIloczSkal(1 To WymReg, 1 To WymReg) As Double 'Macierz iloczynów skalarnych $h_k$ przez $h_l$.
Dim MacWartFunk(1 To WymReg) As Double 'Macierz wartości funkcji $g^*(h_k)$.
Dim MacR(1 To WymReg, 1 To WymReg) As Double 'Macierz iloczynów skalarnych $h_k$ przez $h_l'$.
Dim MacQ(1 To WymReg) As Double 'Macierz wartości funkcji $g^*(h_k')$.
Dim MacWspReg(1 To WymReg) As Double
   'Macierz współczynników regresji w bazie $h_k$, $k=1,2,...,p$ ($p=WymReg$).

Private Function FunkBaz(ByVal NrFunk As Integer, ByVal Arg As Double) As Double
Select Case NrFunk
    Case 1
        FunkBaz = 1
    Case 2
        FunkBaz = Arg
    Case 3
        FunkBaz = Arg ^ 2
End Select
End Function

Private Function FunkSkal(ByVal Arg As Double) As Double
FunkSkal = Arg
End Function

Private Sub ObliczMacIloczSkal()
Dim IloczSkal As Double
Worksheets(Ark_Dane).Activate
For NrWier = 1 To WymReg
    For NrKol = 1 To WymReg
        IloczSkal = 0
        Cells(1, 1).Activate
        Do
            ActiveCell.Offset(1, 0).Activate
            Arg = ActiveCell.Offset(0, 1).Value
            Wart = ActiveCell.Value
            Wart = Wart * FunkBaz(NrWier, Arg)
            Wart = Wart * FunkBaz(NrKol, Arg)
            IloczSkal = IloczSkal + Wart
        Loop Until ActiveCell.Offset(1, 0).Value = Empty
        MacIloczSkal(NrWier, NrKol) = IloczSkal
    Next NrKol
Next NrWier
End Sub

Private Sub ObliczMacWartFunk()
Dim WartFunk As Double
Worksheets(Ark_Dane).Activate
For NrKol = 1 To WymReg
    WartFunk = 0
    Cells(1, 1).Activate
    Do
        ActiveCell.Offset(1, 0).Activate
        Wart = ActiveCell.Value
```

```
        Arg = ActiveCell.Offset(0, 1).Value
        Wart = Wart * FunkBaz(NrKol, Arg)
        Arg = ActiveCell.Offset(0, 2).Value
        Wart = Wart * FunkSkal(Arg)
        WartFunk = WartFunk + Wart
    Loop Until ActiveCell.Offset(1, 0).Value = Empty
    MacWartFunk(NrKol) = WartFunk
Next NrKol
End Sub

Private Sub ObliczMacR()
For NrWier = 1 To WymReg
    MacR(NrWier, 1) = MacIloczSkal(NrWier, 1)
Next NrWier
For NrKol = 2 To WymReg
    For NrWier = 1 To WymReg
        Suma = MacIloczSkal(NrWier, NrKol)
        For NrKolPom = 1 To NrKol - 1
            Suma = Suma - (MacR(NrWier, NrKolPom) * MacR(NrKol, NrKolPom) / MacR(NrKolPom, NrKolPom))
        Next NrKolPom
        MacR(NrWier, NrKol) = Suma
    Next NrWier
Next NrKol
End Sub

Private Sub ObliczMacQ()
MacQ(1) = MacWartFunk(1)
For NrKol = 2 To WymReg
        Suma = MacWartFunk(NrKol)
        For NrKolPom = 1 To NrKol - 1
            Suma = Suma - (MacR(NrKol, NrKolPom) * MacQ(NrKolPom) / MacR(NrKolPom, NrKolPom))
        Next NrKolPom
        MacQ(NrKol) = Suma
Next NrKol
End Sub

Private Sub ZapiszOblicz()
Worksheets(Ark_Oblicz).Select
Cells.Delete
Cells(1, 1).Value = "Macierz iloczynów pseudo-skalarnych"
Cells(WymReg + 3, 1).Value = "Macierz wartości funkcji"
Cells(WymReg + 6, 1).Value = "Macierz R"
Cells(2 * WymReg + 8, 1).Value = "Macierz Q"
Cells(2 * WymReg + 11, 1).Value = "Macierz współczynników regresji"
For NrKol = 1 To WymReg
    Cells(WymReg + 4, NrKol).Value = MacWartFunk(NrKol)
    Cells(2 * WymReg + 9, NrKol).Value = MacQ(NrKol)
    Cells(2 * WymReg + 12, NrKol).Value = MacWspReg(NrKol)
    For NrWier = 1 To WymReg
        Cells(NrWier + 1, NrKol).Value = MacIloczSkal(NrWier, NrKol)
        Cells(WymReg + NrWier + 6, NrKol).Value = MacR(NrWier, NrKol)
        Next NrWier
Next NrKol
End Sub

Public Sub ObliczMacWspReg()
Dim NrWier, NrKol, NrKolPom As Integer
Dim Arg, Wart As Double
Dim Suma As Double
Call ObliczMacIloczSkal
Call ObliczMacWartFunk
Call ObliczMacR
Call ObliczMacQ
MacWspReg(WymReg) = MacQ(WymReg) / MacR(WymReg, WymReg)
For NrKol = WymReg - 1 To 1 Step -1
    Suma = MacQ(NrKol)
        For NrKolPom = NrKol + 1 To WymReg
            Suma = Suma - MacWspReg(NrKolPom) * MacR(NrKolPom, NrKol)
        Next NrKolPom
        MacWspReg(NrKol) = Suma / MacR(NrKol, NrKol)
Next NrKol
Call ZapiszOblicz
End Sub
```

Następujące rysunki przedstawiają zrzuty ekranów z wynikami obliczeń powyższego programu dla przykładów zaprezentowanych w rozdziale piątym.

	A	B	C	D
1	Macierz iloczynów pseudo-skalarnych			
2	1,00000	0,00000	2,00000	
3	0,00000	2,00000	0,00000	
4	2,00000	0,00000	6,80000	
5				
6	Macierz wartości funkcji			
7	3,00000	4,00000	8,80000	
8				
9	Macierz R			
10	1,00000	0,00000	0,00000	
11	0,00000	2,00000	0,00000	
12	2,00000	0,00000	2,80000	
13				
14	Macierz Q			
15	3,00000	4,00000	2,80000	
16				
17	Macierz współczynników regresji			
18	1,00000	2,00000	1,00000	
19				

Rysunek 6.1: Zrzut ekranu wyników obliczeń dla funkcji (5.10).

	A	B	C	D
1	Macierz iloczynów pseudo-skalarnych			
2	1,00000	0,71000	1,35000	
3	0,71000	1,35000	2,09000	
4	1,35000	2,09000	4,35000	
5				
6	Macierz wartości funkcji			
7	3,77000	5,50000	9,88000	
8				
9	Macierz R			
10	1,00000	0,00000	0,00000	
11	0,71000	0,84590	0,00000	
12	1,35000	1,13150	1,01397	
13				
14	Macierz Q			
15	3,77000	2,82330	1,01397	
16				
17	Macierz współczynników regresji			
18	1,00000	2,00000	1,00000	
19				

Rysunek 6.2: Zrzut ekranu wyników obliczeń dla funkcji (5.11).

	A	B	C	D
1	Macierz iloczynów pseudo-skalarnych			
2	1,00000	0,71000	1,35000	
3	0,71000	1,35000	2,09000	
4	1,35000	2,09000	4,35000	
5				
6	Macierz wartości funkcji			
7	24,65000	43,66000	82,60000	
8				
9	Macierz R			
10	1,00000	0,00000	0,00000	
11	0,71000	0,84590	0,00000	
12	1,35000	1,13150	1,01397	
13				
14	Macierz Q			
15	24,65000	26,15850	14,33214	
16				
17	Macierz współczynników regresji			
18	-2,96381	12,01698	14,13464	

Rysunek 6.3: Zrzut ekranu wyników obliczeń dla funkcji (5.12).

	A	B
1	Macierz iloczynów pseudo-skalarnych	
2	1,00000	792,50000
3	792,50000	837144,16667
4		
5	Macierz wartości funkcji	
6	32219,28538	33926965,01648
7		
8	Macierz R	
9	1,00000	0,00000
10	792,50000	209087,91667
11		
12	Macierz Q	
13	32219,28538	8393181,35379
14		
15	Macierz współczynników regresji	
16	406,84816	40,14188
17		

Rysunek 6.4: Zrzut ekranu wyników obliczeń dla funkcji (5.13).

	A	B	C
1	Macierz iloczynów pseudo-skalarnych		
2	1,00000	792,50000	837144,16667
3	792,50000	837144,16667	994841100,00000
4	837144,16667	994841100,00000	1261060410931,17000
5			
6	Macierz wartości funkcji		
7	32219,28538	33926965,01648	40096105237,13970
8			
9	Macierz R		
10	1,00000	0,00000	0,00000
11	792,50000	209087,91667	0,00000
12	837144,16667	331404347,91668	34974163699,23020
13			
14	Macierz Q		
15	32219,28538	8393181,35379	-179274017,64489
16			
17	Macierz współczynników regresji		
18	-1740,74105	48,26643	-0,00513

Rysunek 6.5: Zrzut ekranu wyników obliczeń dla funkcji (5.14).

	A	B	C	D
1	Macierz iloczynów pseudo-skalarnych			
2	1,00000	792,50000	0,00233	-0,00972
3	792,50000	837144,16667	-28,44047	-14,97441
4	0,00233	-28,44047	0,50431	0,00220
5	-0,00972	-14,97441	0,00220	0,49569
6				
7	Macierz wartości funkcji			
8	32219,28538	33926965,01648	-41,86812	518,61461
9				
10	Macierz R			
11	1,00000	0,00000	0,00000	0,00000
12	792,50000	209087,91667	0,00000	0,00000
13	0,00233	-30,28968	0,49992	0,00000
14	-0,00972	-7,27191	0,00117	0,49534
15				
16	Macierz Q			
17	32219,28538	8393181,35379	1098,83645	1121,10435
18				
19	Macierz współczynników regresji			
20	109,60675	40,53825	2192,73950	2263,31969
21				

Rysunek 6.6: Zrzut ekranu wyników obliczeń dla funkcji (5.15).

	A	B	C	D	E	F	G
1	Macierz iloczynów pseudo-skalarnych						
2	1,00000	792,50000	837144,16667	0,00233	-0,00972	0,00118	-0,01000
3	792,50000	837144,16667	994841100,00000	-28,44047	-14,97441	-62,42131	-15,42786
4	837144,16667	994841100,00000	1261060410931,17000	-45542,90264	-21798,59285	-99893,51452	-16378,17966
5	0,00233	-28,44047	-45542,90264	0,50431	0,00220	-0,00038	0,00230
6	-0,00972	-14,97441	-21798,59285	0,00220	0,49569	0,00112	-0,00963
7	0,00118	-62,42131	-99893,51452	-0,00038	0,00112	0,50486	0,00117
8	-0,01000	-15,42786	-16378,17966	0,00230	-0,00963	0,00117	0,49514
9							
10	Macierz wartości funkcji						
11	32219,28538	33926965,01648	40096105237,13970	-41,86812	518,61461	-1822,07127	720,35681
12							
13	Macierz R						
14	1,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
15	792,50000	209087,91667	0,00000	0,00000	0,00000	0,00000	0,00000
16	837144,16667	331404347,91668	34974163699,23020	0,00000	0,00000	0,00000	0,00000
17	0,00233	-30,28968	512,85874	0,49991	0,00000	0,00000	0,00000
18	-0,00972	-7,27191	-2136,21039	0,00120	0,49521	0,00000	0,00000
19	0,00118	-63,35974	-459,62068	-0,00955	-0,00108	0,48547	0,00000
20	-0,01000	-7,49912	3883,31552	0,00118	-0,00975	-0,00104	0,49414
21							
22	Macierz Q						
23	32219,28538	8393181,35379	-179274017,64489	1101,46531	1110,07915	704,26658	1384,41251
24							
25	Macierz współczynników regresji						
26	-2521,95916	49,50141	-0,00531	2219,01987	2300,01640	1456,70906	2801,64332

Rysunek 6.7: Zrzut ekranu wyników obliczeń dla funkcji (5.16).

	A	B	C	D	E	F	G
1	Macierz iloczynów pseudo-skalarnych						
2	1,00000	792,50000	6,37061	0,00233	-0,00972	0,00118	-0,01000
3	792,50000	837144,16667	5442,90959	-28,44047	-14,97441	-62,42131	-15,42786
4	6,37061	5442,90959	41,56358	-0,07311	-0,09992	-0,14411	-0,13239
5	0,00233	-28,44047	-0,07311	0,50431	0,00220	-0,00038	0,00230
6	-0,00972	-14,97441	-0,09992	0,00220	0,49569	0,00112	-0,00963
7	0,00118	-62,42131	-0,14411	-0,00038	0,00112	0,50486	0,00117
8	-0,01000	-15,42786	-0,13239	0,00230	-0,00963	0,00117	0,49514
9							
10	Macierz wartości funkcji						
11	32219,28538	33926965,01648	221207,95283	-41,86812	518,61461	-1822,07127	720,35681
12							
13	Macierz R						
14	1,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
15	792,50000	209087,91667	0,00000	0,00000	0,00000	0,00000	0,00000
16	6,37061	394,19767	0,23566	0,00000	0,00000	0,00000	0,00000
17	0,00233	-30,28968	-0,03087	0,49588	0,00000	0,00000	0,00000
18	-0,00972	-7,27191	-0,02429	-0,00201	0,49283	0,00000	0,00000
19	0,00118	-63,35974	-0,03220	-0,01377	-0,00445	0,48084	0,00000
20	-0,01000	-7,49912	-0,05452	-0,00591	-0,01563	-0,00885	0,48143
21							
22	Macierz Q						
23	32219,28538	8393181,35379	127,47495	1115,53502	1141,34207	741,86800	1436,35786
24							
25	Macierz współczynników regresji						
26	-10132,43738	37,37466	2005,83088	2339,39068	2424,97173	1597,76804	2983,52543
27							

Rysunek 6.8: Zrzut ekranu wyników obliczeń dla funkcji (5.17).

	A	B	C	D	E	F	G
1	Macierz iloczynów pseudo-skalarnych						
2	1,00000	1456,78000	2154629,78000	-0,38928	0,31840	-0,30044	0,38290
3	1456,78000	2154629,78000	3217490843,98000	-609,62029	486,86085	-474,50454	610,79093
4	2154629,78000	3217490843,98000	4840361215022,90000	-956160,94335	750004,07883	-746849,88579	973721,13301
5	-0,38928	-609,62029	-956160,94335	0,54416	-0,22565	0,14528	-0,37195
6	0,31840	486,86085	750004,07883	-0,22565	0,45584	-0,07150	0,23761
7	-0,30044	-474,50454	-746849,88579	0,14528	-0,07150	0,34080	-0,19464
8	0,38290	610,79093	973721,13301	-0,37195	0,23761	-0,19464	0,65920
9							
10	Macierz wartości funkcji						
11	58932,37590	87104110,46922	130000306458,01300	-24255,01285	21731,70155	-18370,78608	24563,96702
12							
13	Macierz R						
14	1,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
15	1456,78000	32421,81160	0,00000	0,00000	0,00000	0,00000	0,00000
16	2154629,78000	78669273,07160	7046201255,75537	0,00000	0,00000	0,00000	0,00000
17	-0,38928	-42,52472	-14223,01445	0,30814	0,00000	0,00000	0,00000
18	0,31840	23,02790	8102,94513	-0,05514	0,31892	0,00000	0,00000
19	-0,30044	-36,82604	-10151,89334	-0,04047	0,05474	0,17937	0,00000
20	0,38290	52,99714	20130,26687	-0,11275	0,03473	-0,01117	0,32272
21							
22	Macierz Q						
23	58932,37590	1252603,90305	-16502001,71730	295,80846	2150,10869	403,74922	-102,09259
24							
25	Macierz współczynników regresji						
26	-1234,61264	42,49254	-0,00097	2281,38964	6393,33686	2231,22630	-316,35312

Rysunek 6.9: Zrzut ekranu wyników obliczeń dla funkcji (5.18).

	A	B	C
1	Macierz iloczynów pseudo-skalarnych		
2	1,00000	370,00000	
3	370,00000	182410,00000	
4			
5	Macierz wartości funkcji		
6	5,70298	2244,22422	
7			
8	Macierz R		
9	1,00000	0,00000	
10	370,00000	45510,00000	
11			
12	Macierz Q		
13	5,70298	134,12273	
14			
15	Macierz współczynników regresji		
16	4,61255	0,00295	

Rysunek 6.10: Zrzut ekranu wyników obliczeń dla funkcji (5.19).

	A	B	C	D	E
1	Macierz iloczynów pseudo-skalarnych				
2	1,00000	370,00000	182410,00000	0,01233	0,02136
3	370,00000	182410,00000	101169100,00000	-9,56608	16,06128
4	182410,00000	101169100,00000	59851603078,00000	-6771,93887	12417,01898
5	0,01233	-9,56608	-6771,93887	0,49458	0,00939
6	0,02136	16,06128	12417,01898	0,00939	0,50542
7					
8	Macierz wartości funkcji				
9	5,70298	2244,22422	1147791,94628	0,01645	0,20446
10					
11	Macierz R				
12	1,00000	0,00000	0,00000	0,00000	0,00000
13	370,00000	45510,00000	0,00000	0,00000	0,00000
14	182410,00000	33677400,00000	1656918977,99986	0,00000	0,00000
15	0,01233	-14,12930	1434,07596	0,48880	0,00000
16	0,02136	8,15755	2483,89242	0,00951	0,49959
17					
18	Macierz Q				
19	5,70298	134,12273	8261,08972	-0,01940	0,04659
20					
21	Macierz współczynników regresji				
22	5,06818	-0,00070	4,88E-06	-0,04150	0,09326

Rysunek 6.11: Zrzut ekranu wyników obliczeń dla funkcji (5.20).

	A	B	C	D
1	Macierz iloczynów pseudo-skalarnych			
2	1,00000	10,00000		
3	10,00000	130,00000		
4				
5	Macierz wartości funkcji			
6	9,63737	95,01789		
7				
8	Macierz R			
9	1,00000	0,00000		
10	10,00000	30,00000		
11				
12	Macierz Q			
13	9,63737	-1,35579		
14				
15	Macierz współczynników regresji			
16	10,08930	-0,04519		

Rysunek 6.12: Zrzut ekranu wyników obliczeń dla funkcji (5.21).

	A	B	C	D	E
1	Macierz iloczynów pseudo-skalarnych				
2	1,00000	10,00000	130,00000	0,00000	-0,05263
3	10,00000	130,00000	1900,00000	-1,61983	-0,52632
4	130,00000	1900,00000	29614,00000	-32,39667	-5,01467
5	0,00000	-1,61983	-32,39667	0,52632	0,00000
6	-0,05263	-0,52632	-5,01467	0,00000	0,47368
7					
8	Macierz wartości funkcji				
9	9,63737	95,01789	1218,39474	-0,13681	-0,51646
10					
11	Macierz R				
12	1,00000	0,00000	0,00000	0,00000	0,00000
13	10,00000	30,00000	0,00000	0,00000	0,00000
14	130,00000	600,00000	714,00000	0,00000	0,00000
15	0,00000	-1,61983	0,00000	0,43885	0,00000
16	-0,05263	0,00000	1,82744	0,00000	0,46624
17					
18	Macierz Q				
19	9,63737	-1,35579	-7,34737	-0,21002	0,00958
20					
21	Macierz współczynników regresji				
22	9,62476	0,13583	-0,01034	-0,47856	0,02055
23					

Rysunek 6.13: Zrzut ekranu wyników obliczeń dla funkcji (5.22).

Bibliografia

[1] Aldrich J., *Fisher and Regression*, Statistical Science, 2005, 20(4):401–417. DOI: 10.1214/088342305000000331. JSTOR 20061201

[2] Ash R. B., *Real Analysis and Probability,* Probability and Mathematical Statistics, Academic Press, Inc., New York and London, 1972.

[3] Barnett A. G., Van der Pols J. C., Dobson A. J., *Regression to the mean: what it is and how to deal with it*, International Journal of Epidemiology 2005;34:215–220, DOI:10.1093/ije/dyh299.

[4] Billingsley P., *Probability and Measure,* John Wiley & Sons, New York 2012.

[5] Conway J. B., *A Course in Functional Analysis*, Springer, 1990.

[6] Draper N. R., Smith H., *Applied Regression Analysis,* John Wiley & Sons, New York 1998.

[7] JCGM 100:2008, *Ewaluacja danych pomiarowych, Przewodnik wyrażania niepewności pomiaru,* `https://www.gum.gov.pl/`

[8] Fałda B., Zając J., *Zagadnienie regresji w naukach ekonomicznych,* Metody ilościowe w badaniach ekonomicznych, tom XIII/1, SGGW, Warszawa, 2012.

[9] Fisher R. A., *The goodness of fit of regression formulae, and the distribution of regression coefficients*. Journal of the Royal Statistical Society. 1922, 85 (4): 597–612. DOI:10.2307/2341124. JSTOR 2341124

[10] Fisher R. A., *Statistical Methods for Research Workers*, Edinburgh, 1954: Oliver and Boyd.

[11] Fisz M., *Probability Theory and Mathematical Statistic,* John Wiley & Sons, Inc., New York 1967

[12] Galton F., *Kinship and Correlation (reprinted 1989)*, Statistical Science 4(2):80–86. DOI: 10.1214/ss/1177012581. JSTOR 2245330

[13] Galton F., *Regression towards mediocrity in hereditary stature*, Anthropological Miscellanea, Journal of the Anthropological Institute 15, 1886. `http://galton.org/essays/1880-1889/galton-1886-jaigi-regression-stature.pdf`

[14] Galton F., *Typical laws of heredity*, Nature 15 (April 5), 1877. `https://galton.org/essays/1870-1879/galton-1877-roy-soc-typical-laws-heredity.pdf`

[15] Galton F., *Presidential address*, Section H, Anthropology, 1885. `https://galton.org/essays/1880-1889/galton-1885-rba-address.pdf`

[16] Gauss C. F., *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium*, 1809, `oai:www.bibliotekacyfrowa.pl:96383`

[17] Gauss C. F., *Theoria combinationis observationum erroribus minimis obnoxiae*, 1821/1823, e-book - books.google.pl

[18] Geweke J. F., Horowitz J. L., Pesaran M. H., *Econometrics: A Bird's Eye View*, IZA, Discussion Paper Series No. 2458, 2006.

[19] Gleichgewicht B., *Algebra*, PWN, Warszawa 1983.

[20] Jakubowski J., Sztencel R., *Wstęp do teorii prawdopodobieństwa*, Script, Warszawa 2001.

[21] Kołodziej W., *Wybrane rozdziały analizy matematycznej*, Biblioteka Matematyczna, PWN, Warszawa 1982.

[22] Legendre A. M., *Nouvelles méthodes pour la détermination des orbites des cométes*, Firmin Didot, Paris 1805. "Sur la Méthode des moindres quarrés" (dodatek), e-book - books.google.pl.

[23] Krzych Ł. J., Kowalska M., Zejda J. E., *Regresja w kierunku wartości średniej w badaniach ciśnienia tętniczego*, `https://journals.viamedica.pl/arterial_hypertension/article/viewFile/12442/10279`.

[24] Leksiński W., Żakowski W., *Matematyka, część IV*, Wydawnictwa Naukowo - Techniczne, Warszawa 1995.

[25] Partyka D., Zając J., *Generalized approach to the problem of regression*, Anal. Math. Phys, 2015, 5:249-320. DOI 10.1007/s13324-014-0096-7.

[26] Pearson K., Yule G. U., Blanchard N., Lee A., *The Law of Ancestral Heredity*, Biometrika, 1903, 2(2):211–236. DOI:10.1093/biomet/2.2.211. JSTOR 2331683

[27] Pituch J., Szumera A., *Matematyka dla inżynierów, cz. 2*, Państwowa Wyższa Szkoła Zawodowa w Chełmie, Chełm 2014.

[28] Ramcharan R., *Regressions: Why Are Economists Obessessed with Them?*, March 2006.

[29] Rényi A., *Probability Theory,* Akadémiai Kiadó, Budapest 1970.

[30] Rudin W., *Real and complex analysis,* McGRAF-HILL, Inc., New York 1970.

[31] Sen A. K., Srivastava M. S., *Regression analysis: theory, methods and applications,* Springer-Verlag, New York, 1990.

[32] Stigler S.M., *Gauss and the Invention of Least Squares*, The Annals of Statistics 9, 1981. DOI: 10.1214/aos/1176345451.

[33] Szumera A., *Probabilistic regression structures,* Bulletin De La Société des Sciences et des Lettres De Łódź, vol. 68, no. 3 (2018). DOI: 10.26485/0459-6854/2018/68.3/2.

[34] Yule G. U., *On the Theory of Correlation*, Journal of the Royal Statistical Society, 1897. 60(4):812–54. DOI: 10.2307/2979746.