Politechnika Gdańska
Wydział Fizyki Technicznej i Matematyki Stosowanej

Anna Wieżel
Nr albumu: 132540

Funkcjonalne Modele Liniowe

Praca magisterska
na kierunku MATEMATYKA
w zakresie MATEMATYKA FINANSOWA

Praca wykonana pod kierunkiem
dra hab. Karola Dziedziula
Katedra Analizy Matematycznej i Numerycznej

Wrzesień 2015

Oświadczenie kierujcego prac

Potwierdzam, że niniejsza praca została przygotowana pod moim kierunkiem i kwalifikuje si do przedstawienia jej w postpowaniu o nadanie tytułu zawodowego.

Data Podpis kierujcego prac

Oświadczenie autora (autorów) pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowizujcymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur zwizanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załczon wersj elektroniczn.

Data Podpis autora (autorów) pracy

Streszczenie

The paper's motivation is to contribute to popularization of mathematical statistics on infinite dimensional function Hilbert spaces. The author presents the fully functional linear model in form $Y = \beta X + \varepsilon$ and its significance test proposed by Kokoszka et al. The test detects nullity of operator β which implies the lack of linear dependence between X and Y. Using the principal component decomposition it is concluded with test statistic convergent by distribution to chi-squared.
The test is further used for magnetic field data collected in some stations in different latitudes. The results show linear dependence between horizontal intensities of the magnetic field in mid- and low-latitude stations with high-latitude station data with a day or two delay but they contradict the linear dependence between data with more than a two-day lag.

Słowa kluczowe

analiza danych funkcjonalnych, dane funkcjonalne, funkcjonalny model liniowy, test istotności

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.1 Matematyka
11.2 Statystyka

Klasyfikacja tematyczna

62 Statistics
62-07 Data analysis
62J12 Generalized linear models

Tytuł pracy w jzyku angielskim

Functional Linear Models

Spis treści

Wstęp

Przykłady danych funkcjonalnych
Odpowiednik testu istotności dla prostego modelu regresji = F-test (+ t-test) [patrz: artykuł]

Rozdział 1

Preliminaria

Przestrzenią funkcyjną E nazywać będziemy przestrzeń liniową funkcji z dowolnego zbioru A do zbioru B.

Definicja 1.0.1 *[Ferraty, Vieu]*
*Zmienną losową X nazywamy **zmienną funkcjonalną** (ang. functional variable) wtedy i tylko wtedy, gdy przyjmuje wartości w nieskończenie wymiarowej przestrzeni (przestrzeni funkcyjnej). Obserwację χ zmiennej X nazywamy **daną funkcjonalną** (ang. functional data).*

Jeśli zmienna funkcjonalna X (odpowiednio obserwacja χ) jest krzywą, to możemy przedstawić X w następującej postaci $X = \{X(t),\ t \in T\}$ (odp. $\chi = \{\chi(t),\ t \in T\}$), gdzie zbiór indeksów $T \subset \mathbb{R}$. Taką zmienną funkcjonalną możemy zatem utożsamiać z procesem stochastycznym z nieskończenie wymiarową przestrzenią stanów. W szczególności, zmienna funkcjonalna może być powierzchnią, czyli dwuwymiarowym wektorem krzywych - wtedy, analogicznie, T będzie dwuwymiarowym zbiorem indeksów tj. $T \subset \mathbb{R}^2$ - lub dowolnie wymiarowym wektorem krzywych.
W niniejszej pracy skupimy się na zmiennych funkcjonalnych przyjmujących postać krzywych.
[przykłady? czy tylko we wstępie?]
[tu: próba = punkty - ostatecznie: funkcja gładka?]

Aby zbudować pojęcie operatora kowariancji dla zmiennych funkcjonalnych wprowadzimy niezbędne pojęcia z dziedziny operatorów liniowych.

1.1. Klasyfikacja operatorów liniowych

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną, Ω jest zatem zbiorem scenariuszy ω, $\mathcal{F}$ jest σ-algebrą podzbiorów Ω, a P miarą prawdopodobieństwa nad $\mathcal{F}$. Dla uproszczenia zakładamy zupełność zadanej przestrzeni probabilistycznej. Rozważmy proces stochastyczny z czasem ciągłym $X = \{X_t,\ t \in T\}$, gdzie T jest przedziałem w $\mathbb{R}$, zdefiniowany na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$, taki, że $X_t(\omega)$ należy do przestrzeni funkcyjnej E dla wszystkich $\omega \in \Omega$.
W pracy rozważać będziemy zmienne funkcjonalne przyjmujące wartości w przestrzeni Hilberta.

Rozważmy ośrodkową nieskończenie wymiarową przestrzeń Hilberta H z iloczynem skalarnym $\langle \cdot, \cdot \rangle$ zadającym normę $\|\cdot\|$ i oznaczmy przez $\mathcal{L}$ przestrzeń ciągłych (ograniczonych) operato-

rów liniowych w H z normą

$$\|\Psi\|_{\mathcal{L}} := \sup\{\|\Psi(x)\| : \ \|x\| \leqslant 1\}.$$

Definicja 1.1.1 *[Horváth, Kokoszka]*
*Operator $\Psi \in \mathcal{L}$ nazywamy **operatorem zwartym**, jeśli istnieją dwie ortonormalne bazy $\{\nu_j\}_{j=1}^{\infty}$ i $\{f_j\}_{j=1}^{\infty}$, oraz ciąg liczb rzeczywistych $\{\lambda_j\}_{j=1}^{\infty}$ zbieżny do zera, takie że*

$$\Psi(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, \nu_j \rangle f_j, \quad x \in H. \tag{1.1}$$

Bez straty ogólności możemy założyć, że w przedstawionej reprezentacji λ_j są wartościami dodatnimi, w razie konieczności wystarczy f_j zamienić na $-f_j$.
Równoważną definicją operatora zwartego jest spełnienie przez Ψ następującego warunku: zbieżność $\langle y, x_n \rangle \to \langle y, x \rangle$ dla każdego $y \in H$ implikuje $\|\Psi(x_n) - \Psi(x)\| \to 0$.

Inną klasą operatorów są operatory Hilberta-Schmidta, którą oznaczać będziemy przez $\mathcal{S}$.

Definicja 1.1.2 *[Bosq]*
***Operatorem Hilberta-Schmidta** nazywamy taki operator zwarty $\Psi \in \mathcal{L}$, dla którego ciąg $\{\lambda_j\}_{j=1}^{\infty}$ w reprezentacji (1.1) spełnia $\sum_{j=1}^{\infty} \lambda_j^2 < \infty$.*

Uwaga 1.1.1 *[Bosq], [Horváth, Kokoszka]*
Klasa $\mathcal{S}$ jest przestrzenią Hilberta z iloczynem skalarnym

$$\langle \Psi_1, \Psi_2 \rangle_{\mathcal{S}} := \sum_{j=1}^{\infty} \langle \Psi_1(e_j), \Psi_2(e_j) \rangle,$$

gdzie $\{e_j\}_{j=1}^{\infty}$ jest dowolną bazą ortonormalną w H.
Powyższy iloczyn skalarny zadaje normę $\|\Psi\|_{\mathcal{S}} := \left(\sum_{j=1}^{\infty} \lambda_j^2\right)^{1/2}$.

Definicja 1.1.3 *[Bosq]*
*Operator liniowy nazywamy **operatorem śladowym** (ang. nuclear operator), jeśli równość (1.1) spełniona jest dla ciągu $\{\lambda_j\}_{j=1}^{\infty}$ takiego, że $\sum_{j=1}^{\infty} |\lambda_j| < \infty$.*

Uwaga 1.1.2 *[Bosq]*
Klasa operatorów śladowych $\mathcal{N}$ z normą $\|\Psi\|_{\mathcal{N}} := \sum_{j=1}^{\infty} |\lambda_j|$ jest przestrzenią Banacha.

Definicja 1.1.4 *[Horváth, Kokoszka]*
*Operator $\Psi \in \mathcal{L}$ nazywamy **symetrycznym**, jeśli*

$$\langle \Psi(x), y \rangle = \langle x, \Psi(y) \rangle, \quad x, y \in H,$$

*oraz **nieujemnie określonym** (połowicznie pozytywnie określonym, ang. positive semidefinite), jeśli*

$$\langle \Psi(x), x \rangle \geqslant 0, \quad x \in H.$$

Uwaga 1.1.3 *[Horváth, Kokoszka]*
Symetryczny nieujemnie określony operator Hilberta-Schmidta Ψ możemy przedstawić w reprezentacji

$$\Psi(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, \nu_j \rangle \nu_j, \quad x \in H, \tag{1.2}$$

*gdzie ortonormalne ν_j są **funkcjami własnymi** Ψ, tj. $\Psi(\nu_j) = \lambda_j \nu_j$. Funkcje ν_j mogą być rozszerzone do bazy, przez dopełnienie ortogonalne podprzestrzeni rozpiętej przez oryginalne ν_j. Możemy zatem założyć, że funkcje ν_j w (1.2) tworzą bazę, a pewne wartości λ_j mogą być równe zero.*

1.2. Przestrzeń L^2

Przestrzeń $L^2 = L^2(K, \mathcal{A}, \mu)$ nad pewną przestrzenią liniową K jest zbiorem mierzalnych funkcji rzeczywistych całkowalnych z kwadratem określonych na K, tj.

$$x \in L^2(K) \iff x : K \to \mathbb{R} \ \wedge \ \int_K x^2(t)dt < \infty.$$

Przestrzeń L^2 jest ośrodkową przestrzenią Hilberta z iloczynem skalarnym

$$\langle x, y \rangle := \int_K x(t)y(t)dt.$$

Normę zaś wyznacza wzór

$$\|x\|^2 = \langle x, x \rangle = \int x^2(t)dt.$$

Tak jak zwyczajowo zapisujemy L^2 zamiast $L^2(K)$, tak w przypadku symbolu całki bez wskazania obszaru całkowania będziemy mieć na myśli całkowanie po całej przestrzeni K. Jeśli $x, y \in L^2$, równość $x = y$ zawsze oznaczać będzie $\int [x(t) - y(t)]^2 \, dt = 0$.

Ważną klasę operatorów liniowych na przestrzeni L^2 stanowią operatory całkowe.

Definicja 1.2.1 ***Operatorem całkowym*** *nazywamy operator liniowy Ψ dający się przedstawić w formie*

$$\Psi(x)(t) = \int \psi(t, s)x(s)ds, \quad x \in L^2,$$

gdzie ψ stanowi ***jądro całkowe*** *operatora Ψ.*

Uwaga 1.2.1 *[Horváth, Kokoszka]*
Operatory całkowe są operatorami Hilberta-Schmidta wtedy i tylko wtedy, gdy

$$\iint \psi^2(t, s)dtds < \infty.$$

Ponadto zachodzi

$$\| \Psi \|_{\mathcal{S}}^2 = \iint \psi^2(t, s)dtds.$$

Uwaga 1.2.2 *(Twierdzenie Mercera) [Horváth, Kokoszka]*
Jeśli operator spełnia również $\psi(s, t) = \psi(t, s)$ oraz $\iint \psi(t, s)x(t)x(s)dtds \geqslant 0$, to operator całkowy Ψjest symetryczny i nieujemnie określony, zatem z uwagi 1.1.3 mamy

$$\psi(t, s) = \sum_{j=1}^{\infty} \lambda_j \nu_j(t)\nu_j(s) \quad w \ L^2(K) \times L^2(K).$$

Jeżeli funkcja ψ jest ciągła, powyższe rozwinięcie jest prawdziwe dla wszystkich $s, t \in K$ i szereg jest zbieżny jednostajnie.

1.3. Zmienne funkcjonalne w L^2. Pojęcie średniej i operatora kowariancji

Rozważmy zmienną funkcjonalną $X = \{X(t),\ t \in T\}$ będącą krzywą ($T \subset \mathbb{R}$) jako element losowy z przestrzeni $L^2(T)$ zaopatrzonej w σ-algebrę borelowskich podzbiorów T.
Mówimy, że zmienna X jest **całkowalna**, jeśli $\mathbb{E}\,\|X\| = \mathbb{E}\left[\int X^2(t)dt\right]^{1/2} < \infty$. Jeśli X jest całkowalna, to istnieje jedyna funkcja $\mu \in L^2$ taka, że $\mathbb{E}\langle y, X\rangle = \langle y, \mu\rangle$ dla dowolnej funkcji $y \in L^2$. Zachodzi $\mu(t) = \mathbb{E}[X(t)]$ dla prawie wszystkich $t \in T$ i funkcję μ nazywać będziemy **funkcją średniej**.

Definicja 1.3.1 *[Bosq]*
Operator kowariancji *całkowalnej zmiennej funkcjonalnej X o średniej μ_X przyjmującej wartości w przestrzeni funkcyjnej L^2 spełniającej $\mathbb{E}\,\|X\|^2 < \infty$ definiujemy następująco*

$$C_X(x) := \mathbb{E}[\langle X - \mu_X, x\rangle(X - \mu_X)], \quad x \in L^2.$$

Jeśli Y jest zmienną funkcjonalną o średniej μ_Y spełniającą powyższe warunki, wtedy operator kowariancji między zmiennymi X i Y przedstawiamy jako

$$C_{X,Y}(x) := \mathbb{E}\Big[\langle (X - \mu_X), x\rangle(Y - \mu_Y)\Big], \quad x \in L^2$$

oraz

$$C_{Y,X}(x) := \mathbb{E}\Big[\langle (Y - \mu_Y), x\rangle(X - \mu_X)\Big], \quad x \in L^2.$$

Operator kowariancji jest operatorem całkowym, czyli

$$C_X(x)(t) = \int c(t,s)x(s)ds,$$

gdzie jądro całkowe $c(t,s)$ zdefiniowane następująco

$$c(t,s) = \mathbb{E}\Big[\big(X(t) - \mu(t)\big)\big(X(s) - \mu(s)\big)\Big]$$

nazywać będziemy **funkcją kowariancji**. Oczywistym jest, że $c(t,s) = c(s,t)$ i mamy

$$\begin{aligned}\iint c(t,s)x(t)x(s)dtds &= \iint \mathbb{E}[\big(X(t) - \mu(t)\big)\big(X(s) - \mu(s)\big)]x(t)x(s)dtds \\ &= \mathbb{E}\left[\left(\int X(t)x(t)dt\right)^2\right] \geqslant 0.\end{aligned}$$

Zatem operator kowariancji C_X jest symetryczny oraz nieujemnie określony. Wartości własne λ_j operatora C_X są dodatnie i spełniony jest warunek $\sum_{j=1}^{\infty}\lambda_j = \mathbb{E}\,\|X\|^2 < \infty$. C_X jest operatorem Hilberta-Schmidta (a nawet operatorem śladowym) i posiada on następującą reprezentację

$$C_X(x) = \sum_{j=1}^{\infty}\lambda_j\langle x, \nu_j\rangle\nu_j, \quad x \in L^2.$$

1.4. Estymacja średniej, funkcji kowariancji i operatora kowariancji

Naturalnym problemem pojawiającym się przy danych funkcjonalnych jest wnioskowanie o obiektach nieskończenie wymiarowych na podstawie skończonej próbki danych.
Obserwujemy zatem N krzywych $X_1, ..., X_N$, które może traktować jako realizacje losowej funkcji X lub obserwacje zmiennej funkcjonalnej X z przestrzeni L^2.

Założenie 1.4.1 *[Horváth, Kokoszka]*
Zakładamy, że $X_1, ..., X_N$ są niezależnymi zmiennymi losowymi w L^2 o jednakowym rozkładzie jak zmienna $X \in L^2$.

Definiujemy szukane parametry

funkcja średniej:	$\mu(t) = \mathbb{E}[X(t)]$;
funkcja kowariancji:	$c(t,s) = \mathbb{E}[(X(t) - \mu(t)(X(s) - \mu(s))]$;
operator kowariancji:	$C = \mathbb{E}[\langle (X - \mu), \cdot \rangle (X - \mu)]$.

Funkcję średniej μ estymujemy średnią z funkcji z próby

$$\widehat{\mu}(t) = \frac{1}{N} \sum_{n=1}^{N} X_n(t),$$

funkcję kowariancji ze wzoru

$$\hat{c}(t,s) = \frac{1}{N} \sum_{n=1}^{N} \big(X_n(t) - \hat{\mu}(t)\big)\big(X_n(s) - \hat{\mu}(s)\big),$$

zaś operator kowariancji estymujemy

$$\widehat{C}(x) = \frac{1}{N} \sum_{n=1}^{N} \langle X_n - \hat{\mu}, x \rangle (X_n - \hat{\mu}), \quad x \in L^2.$$

1.5. Estymacja wartości własnych i funkcji własnych

1.6. Funkcjonalny model liniowy

Standardowy model liniowy dla par zmiennych skalarnych Y_n i wektorów $\mathbf{X}_n$ (tworzonych przez p skalarnych zmiennych X_{ni}, $i = 1, ..., p$), przy założeniu $\mathbb{E}Y_n = 0$, $\mathbb{E}\mathbf{X}_n = \mathbf{0}$[1] (gdzie $n = 1, ..., N$), przyjmuje postać

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1.3}$$

gdzie
$\mathbf{Y}$ jest wektorem zmiennych objaśnianych długości N,
$\mathbf{X}$ jest macierzą zmiennych objaśniających wymiaru $N \times p$,
$\boldsymbol{\beta}$ jest wektorem parametrów długości p,
$\boldsymbol{\varepsilon}$ jest wektorem błędów losowych długości N.

[Mając dane realizacje zmiennych $\mathbf{Y}$ oraz $\mathbf{X}$ poszukiwany wektor współczynników modelu $\boldsymbol{\beta}$ znajdujemy metodą najmniejszych kwadratów.]

[1] przenieść tę uwagę/wytłumaczenie do przypisu?

Poza narzuconym już założeniem o scentrowanych zmiennych losowych $\mathbf{Y}$ i $\mathbf{X}$ (tu: jedynie aby uniknąć uwzględniania wyrazu wolnego[2]) najważniejszymi założeniami powyższego modelu liniowego są wymagania, aby zmienna losowa $\boldsymbol{\varepsilon}$ opisująca błąd modelu również spełniała $\mathbb{E}[\boldsymbol{\varepsilon}] = 0$ oraz aby nie była skorelowana ze zmiennymi X_n.

Rozważać będziemy odpowiednik modelu liniowego dla zmiennych funkcjonalnych. Dla uproszczenia (podobnie jak wyżej) zakładać będziemy, że zmienne objaśniane i objaśniające mają średnie równe zero. **Pełen model funkcjonalny** (ang. *fully functional model*) przyjmuje postać

$$Y_n = \Psi X_n + \varepsilon_n, \quad n = 1, ..., N, \tag{1.4}$$

gdzie krzywe Y_n, X_n oraz nieobserwowalny błąd ε_n należą do przestrzeni Hilberta $L^2(T)$. Operator $\Psi : L^2 \to L^2$ jest ograniczonym operatorem liniowym, który w szczególności jest operatorem całkowym. Jądro całkowe $\psi(t, s)$ operatora Ψ jest funkcją całkowalną z kwadratem na $T \times T$. Równość (1.4) rozumiemy zatem następująco

$$Y_n(t) = \int \psi(t, s) X_n(s) ds + \varepsilon_n(t), \quad n = 1, ..., N.$$

[Nazwa powyższego modelu wynika z faktu, że zarówno zmienne objaśniane Y_n jak i zmienne objaśniające X_n są zmiennymi funkcjonalnymi. Niewielkim uproszczeniem są pozostałe typy funkcjonalnych modeli liniowych, tj.

- model z odpowiedzią skalarną (ang. *scalar response model*)

$$Y_n = \int \psi(s) X_n(s) ds + \varepsilon_n, \quad n = 1, ..., N,$$

  w którym tylko zmienne objaśniające X_n są zmiennymi funkcjonalnymi,

- model z odpowiedzią funkcyjną (ang. *functional response model*)

$$Y_n(t) = \psi(t) X_n + \varepsilon_n(t), \quad n = 1, ..., N,$$

  w którym zmienne objaśniające X_n są skalarami.]

Naturalnym problemem pojawiającym się przy funkcjonalnym modelu liniowym jest estymacja operatora Ψ należącego do nieskończenie wymiarowej przestrzeni na podstawie skończonej próbki danych. Możliwym jest znalezienie operatora, który daje idealne dopasowanie do danych (dla którego wszystkie różnice od próbki są równe zero) nie narzucając dodatkowych założeń, ale jego interpretacja jest często problemowa i nie funkcjonalna. Problem ten rozwiązuje się przez poszukiwanie operatora należącego do podprzestrzeni generowanej przez funkcje własne operatora kowariancji danych z próby, nazywane **empirycznymi funkcjonalnymi głównymi składowymi** (*ang. empirical functional principal components, EFPC's*).

[2] przenieść tę uwagę/wytłumaczenie do przypisu?

Rozdział 2

Test istotności w funkcjonalnym modelu liniowym

2.1. Procedura testowa

Jednym z podstawowych testów na efektywność modelu jest test istotności zmiennych objaśniających. Jak w przypadku modelu liniowego dla zmiennych skalarnych (postaci (1.3)) testuje się hipotezę o zerowaniu się wektora $\boldsymbol{\beta}$, tak w przypadku funkcjonalnego modelu liniowego badamy zerowanie się operatora Ψ, tj. hipotezy

$$\mathrm{H}_0: \quad \Psi = 0 \quad \text{przeciw} \quad \mathrm{H}_A: \quad \Psi \neq 0.$$

Zauważmy, że przyjęcie H_0 nie oznacza braku związku między zmienną objaśnianą a objaśniającą. Prowadzi jedynie do stwierdzenia braku zależności liniowej.

Zakładamy, że zmienna objaśniana Y_n, zmienne objaśniające X_n i błędy ε_n są scentrowanymi zmiennymi losowymi przyjmującymi wartości w przestrzeni Hilberta L^2. Oznaczając przez X (analogicznie Y) zmienną funkcjonalną o tym samym rozkładzie co X_n (Y_n) wprowadzamy operatory

$$C(x) = \mathbb{E}[\langle X, x\rangle\, X], \quad \Gamma(x) = \mathbb{E}[\langle Y, x\rangle\, Y], \quad \Delta(x) = \mathbb{E}[\langle X, x\rangle\, Y].$$

Przez $\widehat{C}$, $\widehat{\Gamma}$, $\widehat{\Delta}$ oznaczamy ich estymatory, np.

$$\widehat{C}(x) = \frac{1}{N}\sum_{n=1}^{N} \langle X_n, x\rangle\, X_n.$$

Definiujemy również wartości i wektory własne C i Γ

$$C(\upsilon_k) = \lambda_k \upsilon_k, \quad \Gamma(u_j) = \gamma_j u_j,$$

których estymatory będziemy oznaczać $(\widehat{\lambda}_k, \widehat{\upsilon}_k)$, $(\widehat{\gamma}_j, \widehat{u}_j)$.
Test obejmuje obcięcie powyższych operatorów na podprzestrzenie skończenie wymiarowe. Podprzestrzeń $\mathcal{V}_p = \mathrm{span}\{\upsilon_1, ..., \upsilon_p\}$ zawiera najlepsze przybliżenia X_n, które są liniowymi kombinacjami pierwszych p głównych składowych (ang, *Functional Principal Components, FPC*). Metodą głównych składowych wyznaczamy p największych wartości własnych operatora $\widehat{C}$ tak, że $\widehat{\mathcal{V}}_p = \mathrm{span}\{\widehat{\upsilon}_1, ..., \widehat{\upsilon}_p\}$ zawiera najlepsze przybliżenie X_n. Analogicznie $\mathcal{U}_q = \mathrm{span}\{u_1, ..., u_q\}$ zawiera przybliżenia $\mathrm{span}\{Y_1, ..., Y_N\}$.

Z równości

$$Y(t) = \int \psi(s,t)X(s)ds + \boldsymbol{\varepsilon}(t)$$

wynika $\Delta = \psi C$ i dla $k \leqslant p$ mamy

$$\psi(\upsilon_k) = \lambda_k^{-1}\Delta(\upsilon_k).$$

Stąd, ψ zeruje się na $\text{span}\{\upsilon_1, ..., \upsilon_p\}$ wtedy i tylko wtedy, gdy $\Delta(\upsilon_k) = 0$ dla każdego $k = 1, ..., p$. Zauważmy, że

$$\Delta(\upsilon_k) \approx \widehat{\Delta}(\upsilon_k) = \frac{1}{N}\sum_{n=1}^{N} \langle X_n, \upsilon_k\rangle Y_n.$$

Skoro zatem $\text{span}\{Y_1, ..., Y_N\}$ są dobrze aproksymowane przez $\mathcal{U}_q$, to możemy ograniczyć się do sprawdzania czy

$$\left\langle \widehat{\Delta}(\upsilon_k), u_j \right\rangle = 0, \quad k = 1, ..., p, \quad j = 1, ..., q. \tag{2.1}$$

Jeśli H_0 jest prawdziwa, to dla każdego $x \in \mathcal{V}_p$, $\psi(x)$ nie należy do $\mathcal{U}_q$. Co znaczy, że żadna funkcja Y_n nie może być opisana jako liniowa kombinacja X_n, $n = 1, ..., N$. Statystyka testowa powinna zatem sumować kwadraty iloczynów skalarnych (2.1). Poniższe twierdzenia prowadzą do wyznaczenia statystyki

$$\widehat{T}_N(p,q) = N\sum_{k=1}^{p}\sum_{j=1}^{q} \widehat{\lambda}_k^{-1}\widehat{\gamma}_j^{-1}\left\langle \widehat{\Delta}(\widehat{\upsilon}_k), \widehat{u}_j \right\rangle^2, \tag{2.2}$$

która zbiega według rozkładu do rozkładu χ^2 z pq stopniami swobody.
Przy czym

$$\left\langle \widehat{\Delta}(\widehat{\upsilon}_k), \widehat{u}_j \right\rangle = \left\langle \frac{1}{N}\sum_{n=1}^{N} \langle X_n, \widehat{\upsilon}_k\rangle Y_n, \widehat{u}_j \right\rangle = \frac{1}{N}\sum_{n=1}^{N} \langle X_n, \widehat{\upsilon}_k\rangle \langle Y_n, \widehat{u}_j\rangle$$

oraz $\lambda_k = \mathbb{E}\langle X, \upsilon_k\rangle^2$ i $\gamma_j = \mathbb{E}\langle Y, u_j\rangle^2$.

Uwaga 2.1.1 *Oczywistym jest, że jeśli odrzucamy H_0, to $\psi(\upsilon_k) \neq 0$ dla pewnego $k \geqslant 1$. Jednak ograniczając się do p największych wartości własnych, test jest skuteczny tylko jeśli ψ nie zanika na którymś wektorze υ_k, $k = 1, ..., p$. Aczkolwiek takie ograniczenie jest intuicyjnie niegroźne, ponieważ test ma za zadanie sprawdzić czy główne źródła zmienności Y mogą być opisane przez główne źródła zmienności zmiennych X.*

Schemat przebiegu testu

1. Sprawdzamy założenie o liniowości metodą *FPC score predictor-response plots.*
2. Wybieramy liczbę głównych składowych p i q metodami *scree test* oraz *CPV.*
3. Wyliczamy wartość statystyki $\widehat{T}_N(p,q)$ (2.2).
4. Jeśli $\widehat{T}_N(p,q) > \chi^2_{pq}(1-\alpha)$, to odrzucamy hipotezę zerową o braku liniowej zależności. W przeciwnym razie nie mamy podstaw do odrzucenia H_0.

...

2.2. Formalne podstawy

Założenie 2.2.1 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Trójka $(Y_n, X_n, \varepsilon_n)$ tworzy ciąg niezależnych elementów losowych o jednakowym rozkładzie, takich że ε_n jest niezależne od X_n oraz

$$\mathbb{E}X_n = 0, \quad \mathbb{E}\varepsilon_n = 0,$$

$$\mathbb{E}\|X_n\|^4 < \infty \quad i \quad \mathbb{E}\|\varepsilon_n\|^4 < \infty.$$

Założenie 2.2.2 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Wartości własne operatorów C oraz Γ spełniają, dla pewnych $p > 0$ i $q > 0$

$$\lambda_1 > \lambda_2 > ... > \lambda_p > \lambda_{p+1}, \quad \gamma_1 > \gamma_2 > ... > \gamma_q > \gamma_{q+1}.$$

Twierdzenie 2.2.1 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Jeśli spełnione są H_0 i powyższe Założenia 2.2.1, 2.2.2, to $\widehat{T}_N(p,q) \xrightarrow{d} \chi^2_{pq}$ przy $N \to \infty$.

Twierdzenie 2.2.2 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Przy Założeniach 2.2.1, 2.2.2 oraz jeśli $\langle \psi(\upsilon_k), u_j \rangle \neq 0$ dla pewnych $k \leqslant p$ oraz $j \leqslant q$, to $\widehat{T}_N(p,q) \xrightarrow{P} \chi^2_{pq}$ przy $N \to \infty$.

Dowody powyższych twierdzeń rozbijemy w krokach na kolejne lematy i wnioski.

Lemat 2.2.1 *[Kokoszka et al. (2008)], [Bosq]*
Przy powyższych Założeniach spełnione są nierówności

$$\limsup_{N\to\infty} N\mathbb{E}\left\|\nu_k - \widehat{\nu}_k\right\|^2 < \infty, \quad \limsup_{N\to\infty} N\mathbb{E}\left\|u_j - \widehat{u}_j\right\|^2 < \infty,$$

$$\limsup_{N\to\infty} N\mathbb{E}\left[\left|\gamma_k - \widehat{\gamma}_k\right|^2\right] < \infty, \quad \limsup_{N\to\infty} N\mathbb{E}\left[\left|\lambda_j - \widehat{\lambda}_j\right|^2\right] < \infty,$$

dla $k \leqslant p$ oraz $j \leqslant q$.

Lemat 2.2.2 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Jeśli spełnione są H_0 i powyższe Założenia, to dla $j \leqslant q$, $k \leqslant p$

$$\sqrt{N}\langle \widehat{\Delta}\nu_k, u_j \rangle \xrightarrow{d} \eta_{kj}\sqrt{\gamma_k \lambda_j},$$

gdzie $\eta_{kj} \sim N(0,1)$. Przy czym $\eta_{k,j}$ oraz $\eta_{k'j'}$ są niezależne dla $(k,j) \neq (k',j')$.

Dowód. Przy H_0

$$\sqrt{N}\langle \widehat{\Delta}\nu_k, u_j \rangle = N^{-1/2} \sum_{n=1}^{N} \langle X_n, \nu_k \rangle \langle \varepsilon_n, u_j \rangle.$$

... Aby udowodnić niezależność między η_{kj} i $\eta_{k'j'}$ dla $(k,j) \neq (k',j')$, wystarczy pokazać, że $\sqrt{N}(\widehat{\Delta}(\nu_k), u_j)$ i $\sqrt{N}(\widehat{\Delta}(\nu_{k'}), u_{j'})$ są nieskorelowane

$$\begin{aligned}
&\mathbb{E}\left[\sqrt{N}\langle\widehat{\Delta}(\nu_k), u_j\rangle, \sqrt{N}\langle\widehat{\Delta}(\nu_{k'}), u_{j'}\rangle\right] \\
&= \frac{1}{N}\mathbb{E}\left[\sum_{n=1}^{N}\langle X_n, \nu_k\rangle\langle\varepsilon_n, u_j\rangle \sum_{n'=1}^{N}\langle X_{n'}, \nu_{k'}\rangle\langle\varepsilon_{n'}, u_{j'}\rangle\right] \\
&= \frac{1}{N}\sum_{n,n'=1}^{N}\mathbb{E}\left[\langle X_n, \nu_k\rangle\langle X_{n'}, \nu_{k'}\rangle\right]\mathbb{E}\left[\langle\varepsilon_n, u_j\rangle\langle\varepsilon_{n'}, u_{j'}\rangle\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[\langle X_n, \nu_k\rangle\langle X_n, \nu_{k'}\rangle\right]\mathbb{E}\left[\langle\varepsilon_n, u_j\rangle\langle\varepsilon_n, u_{j'}\rangle\right] \\
&= \langle C(\nu_k), \nu_{k'}\rangle\langle\Gamma u_j, u_{j'}\rangle = \gamma_k\delta_{kk'}\gamma_j\delta jj'.
\end{aligned}$$

□

Przypomnijmy, że norma Hilberta-Schmidta operatora Hilberta-Schmidta S zdefiniowana jest wzorem $\|S\|_{\mathcal{S}}^2 = \sum_{j=1}^{\infty}\|S(e_j)\|^2$, gdzie ciąg $\{e_1, e_2, ...\}$ stanowi bazę ortonormalną oraz, że norma ta jest nie mniejsza od normy operatorowej, tj. $\|S\|_{\mathcal{L}}^2 \leqslant \|S\|_{\mathcal{S}}^2$.

Lemat 2.2.3 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Przy założeniach Twierdzenia 2.2.1 mamy

$$\mathbb{E}\left\|\widehat{\Delta}\right\|_{\mathcal{S}}^2 = N^{-1}\mathbb{E}\|X\|^2\,\mathbb{E}\|\varepsilon_1\|^2.$$

Dowód. Zauważmy, że

$$\left\|\widehat{\Delta}(e_j)\right\|^2 = N^{-2}\sum_{n,n'=1}^{N}\langle X_n, e_j\rangle\langle X_{n'}, e_j\rangle\langle Y_n, Y_{n'}\rangle.$$

Stąd mamy

$$\begin{aligned}
\mathbb{E}\left\|\widehat{\Delta}\right\|_{\mathcal{S}}^2 &= N^{-2}\sum_{j=1}^{\infty}\sum_{n,n'=1}^{N}\mathbb{E}\left[\langle X_n, e_j\rangle\langle X_{n'}, e_j\rangle\langle\varepsilon_n, \varepsilon_{n'}\rangle\right] \\
&= N^{-2}\sum_{j=1}^{\infty}\sum_{n,n'=1}^{N}\mathbb{E}\langle X_n, e_j\rangle^2\,\mathbb{E}\|\varepsilon_n\|^2 \\
&= N^{-1}\mathbb{E}\|\varepsilon_1\|^2\sum_{j=1}^{\infty}\langle X, e_j\rangle^2 = N^{-1}\mathbb{E}\|\varepsilon_1\|^2\|X\|^2.
\end{aligned}$$

□

Lemat 2.2.4 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Załóżmy, że $\{U_n\}_{n=1}^{\infty}$ oraz $\{V_n\}_{n=1}^{\infty}$ są ciągami elementów losowych z przestrzeni Hilberta takich, że $\|U_n\| \xrightarrow{P} 0$ i $\|V_n\| = O_P(1)$, tj.

$$\lim_{C\to\infty}\limsup_{n\to\infty}P(\|V_n\| > C) = 0.$$

Wtedy zachodzi

$$\langle U_n, V_n\rangle \xrightarrow{P} 0.$$

Dowód. Prawdziwość lematu wynika z analogicznej własności dla losowych ciągów liczb rzeczywistych i nierówności $|\langle U_n, V_n\rangle| \leqslant \|U_n\|\|V_n\|$. □

Lemat 2.2.5 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Przy założeniach Twierdzenia 2.2.1, dla $j \leqslant q$, $k \leqslant p$ zachodzi

$$\sqrt{N}\langle\widehat{\Delta}(\hat{\nu}_k), \hat{u}_j\rangle \xrightarrow{d} \eta_{kj}\sqrt{\gamma_k\lambda_j},$$

gdzie η_{kj} definiowane są jak w Lemacie 2.2.2.

Dowód. Na mocy Lematu 2.2.2, wystarczy pokazać

$$\sqrt{N}\langle\widehat{\Delta}(\hat{\nu}_k), \hat{u}_j\rangle - \sqrt{N}\langle\widehat{\Delta}(\nu_k), u_j\rangle \xrightarrow{P} 0. \tag{2.3}$$

Równość (2.3) wynika z

$$\sqrt{N}\langle\widehat{\Delta}(\hat{\nu}_k), \hat{u}_j - u_j\rangle \xrightarrow{P} 0 \tag{2.4}$$

i

$$\sqrt{N}\langle\widehat{\Delta}(\hat{\nu}_k - \nu_k), \hat{u}_j\rangle \xrightarrow{P} 0. \tag{2.5}$$

Aby udowodnić równość (2.4), zauważmy, że $\sqrt{N}(\hat{u}_j - u_j) = O_P(1)$ oraz, na mocy Lematu 2.2.3, $\mathbb{E}\left\|\widehat{\Delta}(\nu_k)\right\| \leqslant \mathbb{E}\left\|\widehat{\Delta}\right\|_{\mathcal{S}} = O(N^{-1/2})$. Stąd równość (2.4) wynika z Lematu 2.2.4... □

Wniosek 2.2.1 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Przy założeniach Twierdzenia 2.2.1, dla $j \leqslant q$, $k \leqslant p$ zachodzi

$$\sqrt{N}\langle\hat{\lambda}_k^{-1/2}\hat{\gamma}_j^{-1/2}\widehat{\Delta}(\hat{\nu}_k), \hat{u}_j\rangle \xrightarrow{d} \eta_{kj},$$

gdzie η_{kj} definiowane są jak w Lemacie 2.2.2.

Dowód Twierdzenia 2.2.1...

Lemat 2.2.6 *[Kokoszka et al. (2008)], [Horváth, Kokoszka] Jeśli $\{Y_n\}_{n\geqslant 1}$ są elementami losowymi o jednakowych rozkładach, to zachodzi $\mathbb{E}\|\widehat{\Delta}\| \leqslant \mathbb{E}\|Y\|^2$.*

Dowód. Dla dowolnego $u \in L^2$ takiego, że $\|u\| \leqslant 1$, mamy

$$\left\|\widehat{\Delta}u\right\| \leqslant N^{-1}\sum_{n=1}^{N}|\langle Y_n, u\rangle|\|Y_n\| \leqslant N^{-1}\sum_{n=1}^{N}\|Y_n\|^2.$$

Co ze względu na założenie, że Y_n mają jednakowy rozkład, jest równoważne tezie lematu. □

Twierdzenie 2.2.3 *Mocne Prawo Wielkich Liczb [Bosq]*
Niech $\{X_n\}_{n\geqslant 1}$ będzie ciągiem zmiennych funkcjonalnych o jednakowym rozkładzie przyjmujących wartości w ośrodkowej przestrzeni Hilberta takich, że $\mathbb{E}\|X_n\|^2 < \infty$. Niech $m = \mathbb{E}X_n$, wtedy mamy

$$N^{-1}\sum_{n=1}^{N}X_n \xrightarrow{p.n.} m.$$

Lemat 2.2.7 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Jeżeli spełnione jest Założenie 2.2.1, wtedy dla dowolnych funkcji $\nu, u \in L^2$

$$\langle\widehat{\Delta}(\nu), u\rangle \xrightarrow{P} \langle\Delta(\nu), u\rangle.$$

Dowód. Tezę otrzymujemy korzystając z Prawa Wielkich Liczb zauważając

$$\langle \widehat{\Delta}(\nu), u\rangle = \frac{1}{N}\sum_{n=1}^{N}\langle X_n, \nu\rangle\langle Y_n, u\rangle$$

oraz

$$\mathbb{E}[\langle X_n, \nu\rangle\langle Y_n, u\rangle] = \mathbb{E}[\langle\langle X_n, \nu\rangle Y_n, u\rangle] = \langle \Delta(\nu), u\rangle.$$

□

Lemat 2.2.8 *[Kokoszka et al. (2008)], [Horváth, Kokoszka]*
Jeżeli spełnione są Założenia 2.2.1 i 2.2.2, to

$$\langle \widehat{\Delta}(\hat{\nu}_k), \hat{u}_j\rangle \xrightarrow{P} \langle \Delta(\nu_k), u_j\rangle, \quad \textit{dla } k \leqslant p,\ j \leqslant q.$$

Dowód. Na mocy Lematu 2.2.7 wystarczy pokazać

$$\langle \widehat{\Delta}(\nu_k), \hat{u}_j - u_j\rangle \xrightarrow{P} 0$$

i

$$\langle \widehat{\Delta}(\hat{\nu}_k) - \widehat{\Delta}(\nu_k), \hat{u}_j\rangle \xrightarrow{P} 0.$$

Relacje te wynikają z Lematu 2.2.4 oraz Lematu 2.2.6. □

Dowód Twierdzenia 2.2.2. Wprowadźmy oznaczenie

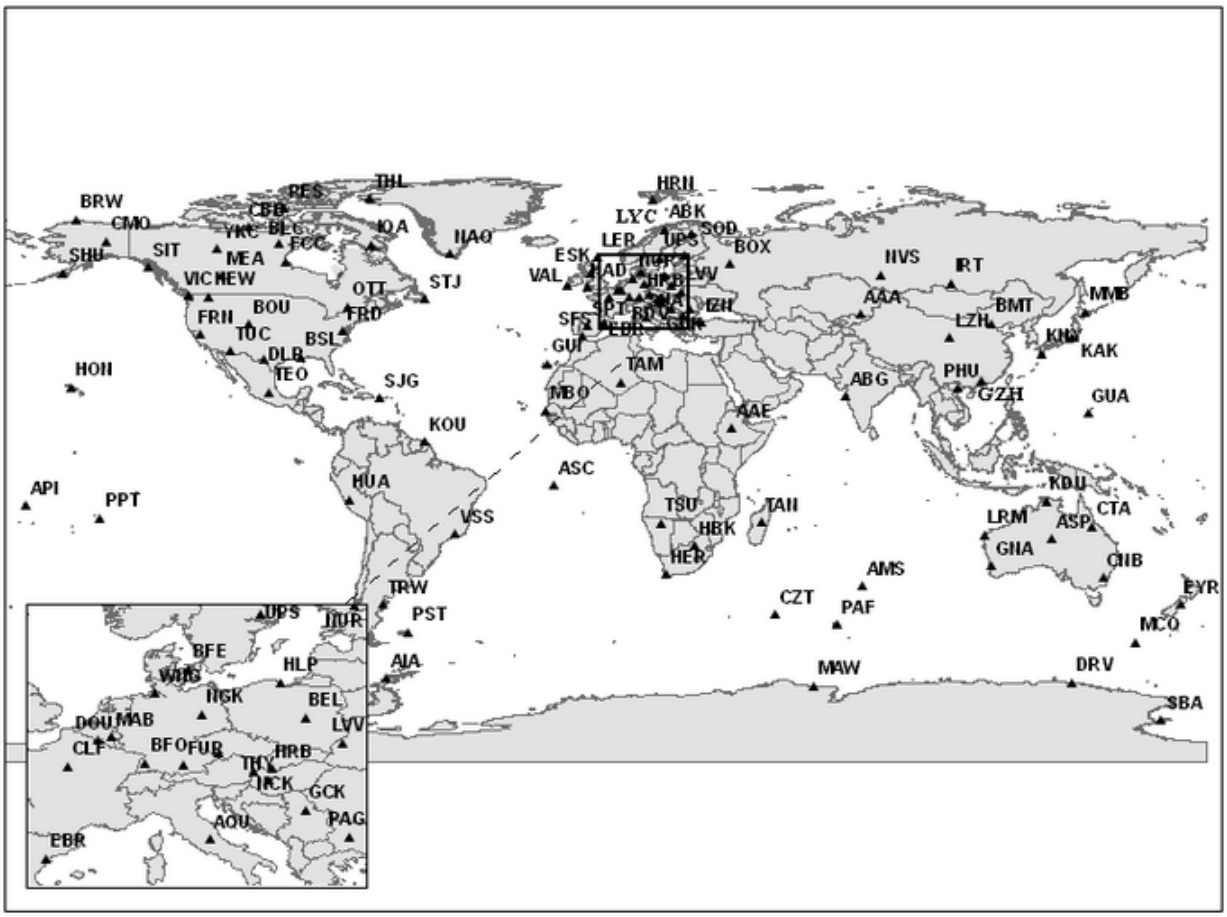
$$\widehat{S}_N(p,q) = \sum_{k=1}^{p}\sum_{j=1}^{q}\hat{\lambda}_k^{-1}\hat{\gamma}_j^{-1}\langle \widehat{\Delta}(\hat{\nu}_k), \hat{u}_j\rangle^2.$$

Na mocy Lematu 2.2.8 ... □

Rozdział 3

Przykład zastosowania

Magnetometer data... dostępne na stronie INTERMAGNET [I]



Rysunek 3.1: Mapa stacji meteorologicznych należących do programu INTERMAGNET

Korzystając z dostępnego pakietu *fda* ([R: fda])

Dodatek A

Kod w R

...

Bibliografia

[Bosq] D. Bosq, *Linear Processes in Function Spaces*. Springer, 2000.

[Ferraty, Vieu] F. Ferraty, P. Vieu, *Nonparametric Functional Data Analysis. Theory and practice*. Springer, 2006.

[Horváth, Kokoszka] L. Horváth, P. Kokoszka, *Interference for Functional Data with Applications*. Springer, 2012.

[I] INTERMAGNET `http://www.intermagnet.org/index-eng.php`

[Kokoszka et al. (2008)] P. Kokoszka, I. Maslova, J. Sojka, L. Zhu, *Testing for lack of dependence in the functional linear model*. Canadian Journal of Statistics, 2008, 36, 207-222.

[R: fda] J. O. Ramsay, H. Wickham, S. Graves, G. Hooker, *Package 'fda'*. On-line: `https://cran.r-project.org/web/packages/fda/fda.pdf`

[Ramsay, Silverman] J. O. Ramsay, B. W. Silverman, *Functional Data Analysis*. Springer, 2005.