Analiza korelacji kanonicznych

Mirosław Krzyśko, Łukasz Waszak

Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza

Wielowymiarowe metody statystyczne
Poznań 09.01.2013

W klasycznej analizie korelacji kanonicznych (CCA) interesuje nas zależność między dwoma wektorami losowami $\boldsymbol{Y}$ i $\boldsymbol{X}$. Poszukujemy wektorów wag $\boldsymbol{l}$ i $\boldsymbol{m}$ takich, żeby kombinacje liniowe $\boldsymbol{l}'\boldsymbol{Y}$ i $\boldsymbol{m}'\boldsymbol{X}$, zwane zmiennymi kanonicznymi, były ze sobą maksymalnie skorelowane.

- Współczynnik korelacji Pearsona

$$(Y, X) \longrightarrow \rho = \frac{Cov(Y, X)}{\sqrt{Var(Y)Var(X)}} \in [-1, 1]$$

- Współczynnik korelacji wielokrotnej (wielorakiej)

$$(Y, \boldsymbol{X}) \longrightarrow \rho = \max_{\boldsymbol{m} \in \mathbb{R}^q} \frac{Cov(Y, \boldsymbol{m}'\boldsymbol{X})}{\sqrt{Var(Y)Var(\boldsymbol{m}'\boldsymbol{X})}} \in [0, 1]$$

Niech

$$Var\begin{pmatrix} Y \\ \boldsymbol{X} \end{pmatrix} = \begin{bmatrix} \sigma_Y^2 & \sigma'_{Y\mathbf{X}} \\ \sigma_{Y\mathbf{X}} & \boldsymbol{\Sigma} \end{bmatrix}, \qquad \sigma_{Y\mathbf{X}} = \begin{bmatrix} Cov(Y, X_1) \\ \dots \\ Cov(Y, X_q) \end{bmatrix}.$$

Stąd

$$\rho = \left(\frac{\sigma'_{Y\mathbf{X}} \boldsymbol{\Sigma}^{-1} \sigma_{Y\mathbf{X}}}{\sigma_Y^2} \right)^{\frac{1}{2}}$$

Współczynnik korelacji

- Współczynnik korelacji kanonicznej

$$(\boldsymbol{Y}, \boldsymbol{X}) \longrightarrow \rho = \max_{\boldsymbol{l}\in\mathbb{R}^p, \boldsymbol{m}\in\mathbb{R}^q} \frac{Cov(\boldsymbol{l}'\boldsymbol{Y}, \boldsymbol{m}'\boldsymbol{X})}{\sqrt{Var(\boldsymbol{l}'\boldsymbol{Y})Var(\boldsymbol{m}'\boldsymbol{X})}} \in (0, 1]$$

Niech

$$Var\left(\begin{array}{c} \boldsymbol{Y} \\ \boldsymbol{X} \end{array}\right) = \left[\begin{array}{cc} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}' & \boldsymbol{\Sigma}_{22} \end{array}\right].$$

Załóżmy, że $p \leq q$ oraz $rz\boldsymbol{\Sigma}_{12} = s \leq p$.

Klasyczna analiza korelacji kanonicznych

Niech

$$U = \boldsymbol{l}'\boldsymbol{Y} = l_1 Y_1 + l_2 Y_2 + l_p Y_p,$$

$$V = \boldsymbol{m}'\boldsymbol{X} = m_1 X_1 + m_2 X_2 + m_q X_q.$$

Wówczas

$$Var(U) = \boldsymbol{l}'\boldsymbol{\Sigma}_{11}\boldsymbol{l}, \qquad Var(V) = \boldsymbol{m}'\boldsymbol{\Sigma}_{22}\boldsymbol{m},$$

$$cov(U, V) = \boldsymbol{l}'\boldsymbol{\Sigma}_{12}\boldsymbol{m}$$

oraz

$$\rho = \max_{\boldsymbol{l}\in\mathbb{R}^p, \boldsymbol{m}\in\mathbb{R}^q} \frac{\boldsymbol{l}'\boldsymbol{\Sigma}_{12}\boldsymbol{m}}{\sqrt{\boldsymbol{l}'\boldsymbol{\Sigma}_{11}\boldsymbol{l}\boldsymbol{m}'\boldsymbol{\Sigma}_{22}\boldsymbol{m}}}.$$

Ponieważ $Corr(U,V) = Corr(c_1 U, c_2 V)$, gdzie $c_1, c_2 \in \mathbb{R}$, to wektory $\boldsymbol{l}$ i $\boldsymbol{m}$ unormujemy tak, by $Var(\boldsymbol{l}'\boldsymbol{Y}) = \boldsymbol{l}'\boldsymbol{\Sigma}_{11}\boldsymbol{l} = 1$ i $Var(\boldsymbol{m}'\boldsymbol{X}) = \boldsymbol{m}'\boldsymbol{\Sigma}_{22}\boldsymbol{m} = 1$. Wówczas

$$\rho = \max_{\boldsymbol{l}\in\mathbb{R}^p, \boldsymbol{m}\in\mathbb{R}^q} \boldsymbol{l}'\boldsymbol{\Sigma}_{12}\boldsymbol{m},$$

przy warunku ograniczającym $\boldsymbol{l}'\boldsymbol{\Sigma}_{11}\boldsymbol{l} = 1$ i $\boldsymbol{m}'\boldsymbol{\Sigma}_{22}\boldsymbol{m} = 1$. Stosując metodę mnożników Lagrange'a otrzymujemy:

$$\begin{cases} -\rho\Sigma_{11}\boldsymbol{l} + \Sigma_{12}\boldsymbol{m} & = & 0 \\ \Sigma_{21}\boldsymbol{l} - \rho\Sigma_{22}\boldsymbol{m} & = & 0 \end{cases}$$

Po przekształceniu otrzymujemy dwa układy równań jednorodnych:

$$(\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \rho_k^2 \boldsymbol{I}_p)\boldsymbol{l} = \boldsymbol{0}, \qquad (\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \rho_k^2 \boldsymbol{I}_q)\boldsymbol{m} = \boldsymbol{0},$$

Klasyczna analiza korelacji kanonicznych

Zajmiemy się tylko układem $(\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} - \rho_k^2 \boldsymbol{I}_p)\boldsymbol{l} = \mathbf{0}$. Warunkiem koniecznym i dostatecznym na to, by istniało niezerowe rozwiązanie tego układu jest warunek:

$$|\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} - \rho_k^2 \boldsymbol{I}_p| = 0.$$

Równanie to ma s niezerowych rozwiązań:

$$\rho_1^2 \geq \rho_2^2 \geq ... \geq \rho_s^2, \qquad s = rz\mathbf{\Sigma}_{12}.$$

Rozwiązaniom tym odpowiada s wektorów własnych:

$$\boldsymbol{l}_1, \boldsymbol{l}_2, ..., \boldsymbol{l}_s.$$

Wtedy wektory własne $\boldsymbol{m}_1, \boldsymbol{m}_2, ..., \boldsymbol{m}_s$ otrzymujemy z równości:

$$\boldsymbol{m}_i = \rho_i^{-1}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\boldsymbol{l}_i, \qquad i = 1, 2, ..., s.$$

- Niezerowe wartości $\rho_1 \geq \rho_2 \geq ... \geq \rho_s$ nazywane są współczynnikami korelacji kanonicznej.
- Zmienne $U_i = \boldsymbol{l}_i'\boldsymbol{Y}$ nazywane są zmiennymi kanonicznymi przestrzeni $\boldsymbol{Y}$, natomiast zmienne $V_i = \boldsymbol{m}_i'\boldsymbol{X}$ nazywane są zmiennymi kanonicznymi przestrzeni $\boldsymbol{X}$.

Zachodzą następujące związki:

$$Corr(U_i, V_j) = \begin{cases} \rho_i, & \text{jeżeli } i = j, \\ 0, & \text{jeżeli } i \neq j, \end{cases}$$

$$Corr(U_i, U_j) = 0, \qquad Corr(V_i, V_j) = 0, \text{ jeżeli } i \neq j.$$

Zatem pierwsze zmienne kanoniczne U_1 i V_1 są kombinacjami liniowymi składowych wektorów $\boldsymbol{Y}$ i $\boldsymbol{X}$, maksymalnie ze sobą skorelowanymi (współczynnik korelacji jest równy ρ_1) i mającymi jednostkowe wariancje. Druga para zmiennych kanonicznych (U_2, V_2) jest tą parą, spośród wszystkich kombinacji liniowych składowych wektorów $\boldsymbol{Y}$ i $\boldsymbol{X}$, która nie jest skorelowana z parą (U_1, V_1), posiada współczynnik korelacji ρ_2 oraz posiada jednostkowe wariancje.

Jeżeli zmienne $Y_1, Y_2, ..., Y_p$ oraz $X_1, X_2, ..., X_q$ zestandaryzujemy, to wówczas macierzom kowariancji $\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}, \mathbf{\Sigma}_{12}$ odpowiadać będą macierze korelacji $\boldsymbol{R}_{11}, \boldsymbol{R}_{22}, \boldsymbol{R}_{12}$. Po standaryzacji, współczynniki korelacji kanonicznej nie ulegną zmianie, a wektory własne dla danych standaryzowanych wyrażą się za pomocą wektorów niestandaryzowanych następująco:

$$\widetilde{\boldsymbol{l}}_i = (diag \mathbf{\Sigma}_{11})^{\frac{1}{2}} \boldsymbol{l}_i, \qquad \widetilde{\boldsymbol{m}}_i = (diag \mathbf{\Sigma}_{22})^{\frac{1}{2}} \boldsymbol{m}_i, \qquad i = 1, 2, ..., s.$$

Przykład 1.

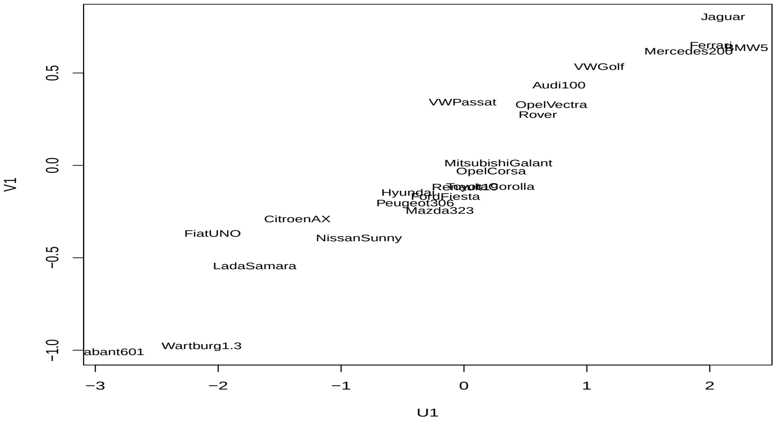
Przykład pochodzi z książki *W. Härdle, L.Simar*, *Applied Multivariate Statistical Analysis, Springer 2007*. Dane są uśrednioną oceną 24 marek samochodów ($N = 24$) uzyskaną na podstawie ankiety 40 osób. Każda z cech została oceniona w skali od 1 (bardzo dobra) do 6 (bardzo zła). Dane zostały podzielone na dwa zbiory zmiennych:

X$-$*Cena, Stabilność wartości*

Y$-$*Ekonomiczność, Serwis, Konstrukcja, Charakter sportowy, Bezpieczeństwo, Łatwość obsługi*

Współczynniki korelacji: **0.979197**, 0.885122

Rysunek: Rzut na przestrzeń 2 pierwszych zmiennych kanonicznych



Odwzorowujemy przestrzeń $\mathbb{R}^p$ do przestrzeni Hilberta z jądrem reprodukującym $H(k_y), \varphi : \mathbb{R}^p \to H(k_y)$, a przestrzeń $\mathbb{R}^q$ do przestrzeni Hilberta z jądrem reprodukującym $H(k_x), \psi : \mathbb{R}^q \to H(k_x)$.
Załóżmy, że $\widetilde{\mathbf{K}}$ jest macierzą jądrową dla danych niescentrowanych.
Wówczas macierz jądrowa dla danych scentrowanych jest postaci:

$$\mathbf{K} = \mathbf{P}\widetilde{\mathbf{K}}\mathbf{P},$$

gdzie $\mathbf{P} = \mathbf{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$, $\mathbf{1}_n$ jest n-wymiarowym wektorem złożonym z samych jedynek.
Formułujemy nowe zagadnienie CCA na zbiorach $\varphi(\mathbf{Y})$ oraz $\psi(\mathbf{X})$ (KCCA) i korzystamy z twierdzenia Moore'a-Aronszajna.

Ocenom macierzy kowariancji $\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}, \mathbf{\Sigma}_{12}$ odpowiadają następujące macierze w wersji jądrowej:

$$\hat{\mathbf{\Sigma}}_{11} \to \mathbf{K}_y\mathbf{K}_y = \mathbf{K}_y^2$$

$$\hat{\mathbf{\Sigma}}_{22} \to \mathbf{K}_x\mathbf{K}_x = \mathbf{K}_x^2$$

$$\hat{\mathbf{\Sigma}}_{12} \to \mathbf{K}_y\mathbf{K}_x$$

Stosując metodę mnożników Lagrange'a otrzymujemy:

$$\begin{cases} -\rho\mathbf{K}_y^2\boldsymbol{a} + \mathbf{K}_y\mathbf{K}_x\boldsymbol{b} & = & 0 \\ \mathbf{K}_x\mathbf{K}_y\boldsymbol{a} - \rho\mathbf{K}_x^2\boldsymbol{b} & = & 0 \end{cases}$$

Powyższe uogólnione zagadnienie własne można zapisać w postaci macierzowej:

$$\mathbf{K}_1\mathbf{c} = \rho\mathbf{K}_2\mathbf{c}, \tag{1}$$

$$\mathbf{K}_1 = \begin{bmatrix} \mathbf{0} & \mathbf{K}_x\mathbf{K}_y \\ \mathbf{K}_y\mathbf{K}_x & \mathbf{0} \end{bmatrix}, \mathbf{K}_2 = \begin{bmatrix} \mathbf{K}_x^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_y^2 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} \mathbf{a} = \boldsymbol{\alpha}^T\varphi(\mathbf{X}) \\ \mathbf{b} = \boldsymbol{\beta}^T\psi(\mathbf{Y}) \end{bmatrix},$$

gdzie $\mathbf{K}_y = (k_y(y_i, y_j))$ oraz $\mathbf{K}_x = (k_x(x_i, x_j))$, $i, j = 1, ..., N$, są macierzami jądrowymi, których elementami są funkcje jądrowe.

Macierz $\mathbf{K}_2$ jest nieujemnie określona, mogą pojawić się więc osobliwości. Najprostrzym sposobem (wykorzystywanym również przy konstrukcji jądrowych zmiennych dyskryminacyjnych) jest regularyzacja macierzy $\mathbf{K}_2$:

$$\mathbf{K}_2 \longmapsto \mathbf{K}_2 + \epsilon \mathbf{I},$$

$\epsilon = 10^{-5}$.

Jądrowe korelacje kanoniczne

i-ta wartość własna ρ_i macierzy $\mathbf{K}_1(\mathbf{K}_2 + \epsilon \mathbf{I})^{-1}$ wyznacza i-ty ***jądrowy współczynnik korelacji kanonicznej***, a odpowiadający jej wektor własny **c** i-tą parę ***jądrowych zmiennych kanonicznych***.

Przykład 1.

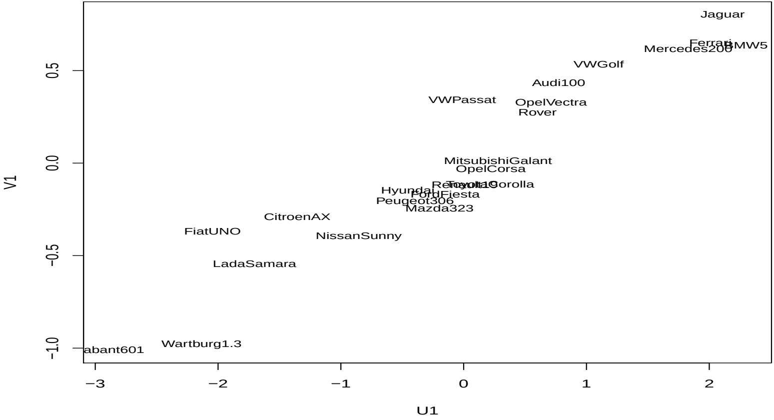
Przykład pochodzi z książki *W. Härdle, L.Simar*, *Applied Multivariate Statistical Analysis, Springer 2007*. Dane są uśrednioną oceną 24 marek samochodów ($N = 24$) uzyskaną na podstawie ankiety 40 osób. Każda z cech została oceniona w skali od 1 (bardzo dobra) do 6 (bardzo zła). Dane zostały podzielone na dwa zbiory zmiennych:

X−*Cena, Stabilność wartości*

Y−*Ekonomiczność, Serwis, Konstrukcja, Charakter sportowy, Bezpieczeństwo, Łatwość obsługi*

0.979197 0.885122

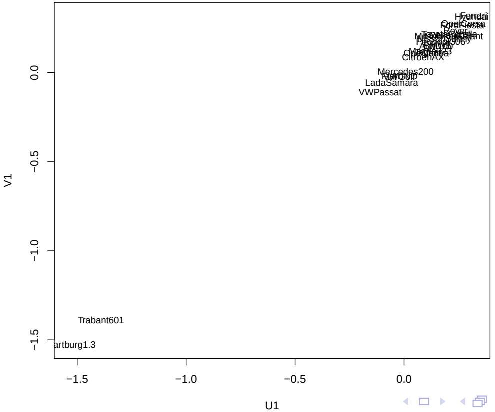
Rysunek: Rzut na przestrzeń 2 pierwszych klasycznych zmiennych kanonicznych



Jądrowe korelacje kanoniczne

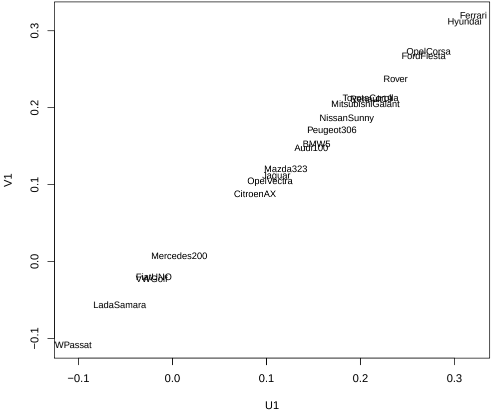
1.000000 0.999998 0.999980 0.999385 0.998763 0.000002

Rysunek: Rzut na przestrzeń 2 pierwszych jądrowych zmiennych kanonicznych (KCCA)



Jądrowe korelacje kanoniczne

Rysunek: Rzut na przestrzeń 2 pierwszych jądrowych zmiennych kanonicznych (KCCA) - górna prawa część wykresu

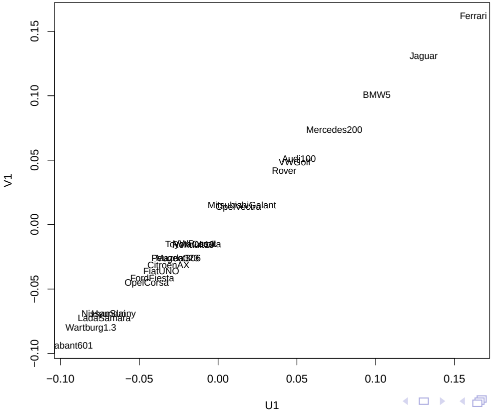


Innym podejściem jest odwzorowanie tylko przestrzeni $\mathbb{R}^q$ do przestrzeni Hilberta z jądrem reprodukującym. Pozostałe rozważania pozostają bez zmian (Zheng et al (2006)) (Q-KCCA).

Jądrowe korelacje kanoniczne

0.99995 0.99935

Rysunek: Rzut na przestrzeń 2 pierwszych jądrowych zmiennych kanonicznych (Q-KCCA)



W przypadku funkcjonalnym (FCCA), interesuje nas zależność między dwoma procesami losowymi $Y(t)$ i $X(t)$. Poszukujemy funkcji wagowych $u(t)$ i $v(t)$ takich, by $\int uY$ i $\int vX$ były ze sobą maksymalnie skorelowane. Problematyka ta została zapoczątkowana pracą Leurgansa, Moyeeda i Silvermana (1993), a w bardziej dojrzałej formie została przedstawiona w monografii Ramsay'a i Silvermana (2005).
Przedstawimy nową metodę konstrukcji funkcjonalnych zmiennych kanonicznych odmienną od metody Ramesay'a i Silvermana.

Niech $\{y_{ij}, x_{ij}\}$ oznacza zaobserwowaną wartość pary cech statystycznych Y i X na i-tej jednostce doświadczalnej w j-tym momencie czasowym, gdzie $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, J_i$. Wówczas nasze dane składają się z N trójek $\{t_{ij}, y_{ij}, x_{ij}\}$, gdzie $t_{ij} \in [0, T]$, $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, J_i$. Jednakże w wielu przypadkach wygodniej jest posługiwać się ciągłymi funkcjami czasu $y(t)$ i $x(t)$, $t \in [0, T]$, tj. danymi funkcjonalnymi. W takim przypadku dane dyskretne $\{t_{ij}, y_{ij}, x_{ij}\}$ przekształcamy na dane funkcjonalne $\{y_i(t), x_i(t), t \in [0, T]\}$.
Proces przekształcania danych jest identyczny dla wszystkich $y_i(t)$ oraz $x_i(t)$, stąd nasze dalsze rozważania dotyczyć będą pojedynczych funkcji $y(t)$ oraz $x(t)$, $t \in [0, T]$.

Jednym ze sposobów wygładzania (Ramsay i Silverman (2005)) jest przedstawienie odpowiednio funkcji $y(t)$ oraz $x(t)$ jako kombinacji liniowych odpowiednio $K_1 + 1$ ortonormalnych funkcji bazowych φ_k oraz $K_2 + 1$ ortonormalnych funkcji bazowych ψ_k:

$$y(t) = \sum_{k=0}^{K_1} \alpha_k \varphi_k(t), \ t \in [0, T_1], \qquad x(t) = \sum_{l=0}^{K_2} \beta_l \psi_l(t), \ t \in [0, T_2].$$

Współczynniki tych kombinacji liniowych dobrane metodą najmniejszych kwadratów są postaci:

$$\hat{\alpha} = \left(\mathbf{\Phi}'\mathbf{\Phi}\right)^{-1} \mathbf{\Phi}'\boldsymbol{y}, \qquad \hat{\boldsymbol{\beta}} = \left(\mathbf{\Psi}'\mathbf{\Psi}\right)^{-1} \mathbf{\Psi}'\boldsymbol{x}.$$

gdzie $\boldsymbol{y} = (y_1, y_2, \ldots, y_{J_1})'$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_{J_2})'$, $\alpha = (\hat{\alpha_0}, \hat{\alpha_1}, \ldots, \hat{\alpha_{K_1}})'$, $\beta = (\hat{\beta_0}, \hat{\beta_1}, \ldots, \hat{\beta_{K_2}})'$ oraz $\mathbf{\Phi}$ jest macierzą wymiaru $J_1 \times (K_1 + 1)$ zawierającą wartości $\varphi_k(t_j)$, a $\mathbf{\Psi}$ jest macierzą wymiaru $J_2 \times (K_2 + 1)$ zawierającą wartości $\psi_l(t_j)$.

Stopień gładkości funkcji $y(t)$ zależy od K_1 (małe K_1 powodują większe wygładzenie krzywych). Optymalną wartość K_1 wybieramy dla każdej funkcji $y_i(t)$ za pomocą bayesowskiego kryterium informacyjnego (BIC, patrz Shmueli(2010)), a następnie z wartości K_1 odpowiadającym wszystkim funkcjom wybierana jest wartość modalna (wspólna dla wszystkich $y_i(t)$).
Wartość BIC dla funkcji $y(t)$ wyraża się wzorem:

$$BIC(y(t)) = ln\left(\frac{\boldsymbol{e}'\boldsymbol{e}}{2}\right) + (K_1+1)\left(\frac{lnJ}{J}\right),$$

gdzie $\boldsymbol{e} = (e_1, ..., e_J)'$,$e_j = y_j - \sum_{k=0}^{K_1} \hat{\alpha}_k \varphi_k(t_j)$,$j = 1, 2, \ldots, J$.
Analogicznie wyliczamy BIC dla funkcji $x(t)$.

Załóżmy, że obserwujemy realizacje dwuwymiarowego procesu losowego (Y, X), gdzie $Y(t) \in L_2(I_1)$, $X(t) \in L_2(I_2)$. Tutaj $L_2(I_1)$ oraz $L_2(I_2)$ są przestrzeniami funkcji całkowalnych z kwadratem na przedziale odpowiednio I_1 oraz I_2 z iloczynem skalarnym

$$< u, v >= \int u(t)v(t)dt.$$

Zakładamy również, że

$$E(< Z, Z >)) = E\left[\int (Z(t))^2 dt\right] < \infty,$$

gdzie $Z = Y$ lub $Z = X$.

Funkcjonalne korelacje kanoniczne

Korelacje kanoniczne dla skończenie wymiarowych wektorów losowych $\boldsymbol{Y} \in \mathbb{R}^{K1+1}$, $\boldsymbol{X} \in \mathbb{R}^{K2+1}$ i procesów stochastycznych $Y(t) \in L_2(I_1)$, $X(t) \in L_2(I_2)$, definiujemy następująco. Niech $H_i = \mathbb{R}^{K_i+1}$ w przypadku wektorowym oraz $H_i = L_2(I_i)$ w przypadku funkcjonalnym.
Wówczas pierwsza korelacja kanoniczna ρ_1 i związane z nią wektory u_1 i v_1 lub funkcje wagowe $u_1(t)$ i $v_1(t)$ jest definiowana następująco:

$$\rho_1 = \sup_{u \in H_1, v \in H_2} Cov(< u, Y >, < v, X >) = Cov(< u_1, Y >, < v_1, X >)$$

przy czym

$$Var(< u, Y >) = 1, \qquad Var(< v, X >) = 1.$$

Ogólnie k-ta korelacja kanoniczna ρ_k i związane z nią wektory u_k i v_k lub funkcje wagowe $u_k(t)$ i $v_k(t)$ jest definiowana następująco:

$$\rho_k = \sup_{u \in H_1, v \in H_2} Cov(< u, Y >, < v, X >) = Cov(< u_k, Y >, < v_k, X >)$$

przy czym $Var(< u, Y >) = 1, \qquad Var(< v, X >) = 1$ oraz k-ta para zmiennych kanonicznych (U_k, V_k) nie jest skorelowana z $k-1$ pierwszymi parami $\{(U_i, V_i), i = 1, 2, \ldots, k-1\}$, gdzie

$$U_k =< u_k, Y >, \qquad V_k =< v_k, X > .$$

Wyrażenie (ρ_k, u_k, v_k) nazywać będziemy k-tą składową kanoniczną.

Rozpatrując przypadek, gdy X i Y są procesami stochastycznymi, zakładać będziemy, że X i Y mogą być reprezentowane przez skończoną liczbę ortonormalnych funkcji bazowych, tj. można je zapisać w postaci:

$$Y(t) = \sum_{k=0}^{K_1} \alpha_k \varphi_k(t), \;\; t \in I_1, \qquad X(t) = \sum_{l=0}^{K_2} \beta_l \psi_l(t), \;\; t \in I_2,$$

gdzie $\{\varphi_k\}$ i $\{\psi_k\}$ są $(K_i + 1)$ pierwszymi elementami baz przestrzeni $L_2(I_i)$ oraz $\{\alpha_k\}$ i $\{\beta_k\}$ są zmiennymi losowymi z zerowymi wartościami oczekiwanymi oraz skończonymi wariancjami.
Funkcjonalna analiza kanoniczna jest równoważna zwykłej analizie kanonicznej dla wektorów losowych złożonych ze współczynników α_k i β_k.

Niech:

$$\boldsymbol{\varphi}(t) = (\varphi_0(t), \varphi_1(t), \dots, \varphi_{K_1}(t))', \qquad \boldsymbol{\psi}(t) = (\psi_0(t), \psi_1(t), \dots, \psi_{K_2}(t))',$$

$$\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{K_1})', \qquad \boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{K_2})',$$

gdzie

$$E(\boldsymbol{\alpha}) = \mathbf{0}, \qquad E(\boldsymbol{\beta}) = \mathbf{0},$$

$$Var(\boldsymbol{\alpha}) = \boldsymbol{\Sigma}_{11}, \qquad Var(\boldsymbol{\beta}) = \boldsymbol{\Sigma}_{22}, \qquad Cov(\boldsymbol{\alpha}, \boldsymbol{\beta}) = E(\boldsymbol{\alpha}\boldsymbol{\beta}') = \boldsymbol{\Sigma}_{12}.$$

Wówczas procesy $Y(t)$ $X(t)$ możemy zapisać w postaci:

$$Y(t) = \boldsymbol{\alpha}'\boldsymbol{\varphi}(t), \qquad X(t) = \boldsymbol{\beta}'\boldsymbol{\psi}(t).$$

Twierdzenie

k-ta składowa kanoniczna $(\rho_k, \boldsymbol{u}_k, \boldsymbol{v}_k)$ pary wektorów losowych $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ związana jest z k-tą składową kanoniczną $(\gamma_k, u_k(t), v_k(t))$ pary procesów losowych $(X(t), Y(t))$ zależnościami:

$$\gamma_k = \rho_k, \qquad u_k(t) = \boldsymbol{u}_k' \boldsymbol{\varphi}(t), \qquad v_k(t) = \boldsymbol{v}_k' \boldsymbol{\psi}(t).$$

Niezerowe wartości własne ρ_k^2 oraz odpowiadające im wektory charakterystyczne $\boldsymbol{u}_k$ wyznaczamy z równania:

$$(\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \rho_k^2 \boldsymbol{I}_{p_1})\boldsymbol{u}_k = \boldsymbol{0},$$

przy czym $1 \leq k \leq min(K_1 + 1, K_2 + 1)$. Natomiast wektory $\boldsymbol{v}_k$ wyznaczamy z równości:

$$\boldsymbol{v}_k = \rho_k^{-1}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{u}_k, \qquad i = 1, 2, ..., s,$$

(Krzyśko (2009)).

Analiza korelacji kanonicznych dla wektorów losowych $\boldsymbol{\alpha}$ i $\boldsymbol{\beta}$ bazuje na macierzach $\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}$ i $\mathbf{\Sigma}_{12}$. W praktyce macierze te nie są znane. Oceniamy je na podstawie N niezależnych realizacji tych wektorów. Mając wyznaczone wektory własne $\hat{\boldsymbol{u}}_k$ i $\hat{\boldsymbol{v}}_k$ wyznaczamy funkcje wagowe

$$\hat{u}_k(t) = \hat{\boldsymbol{u}}'_k \boldsymbol{\varphi}(t), \qquad \hat{v}_k(t) = \hat{\boldsymbol{v}}'_k \boldsymbol{\psi}(t), \qquad t \in [0, T].$$

Stąd współrzędna rzutu i-tej realizacji $y_i(t)$ procesu $Y(t)$ na j-tą funkcjonalną zmienną kanoniczną jest równa

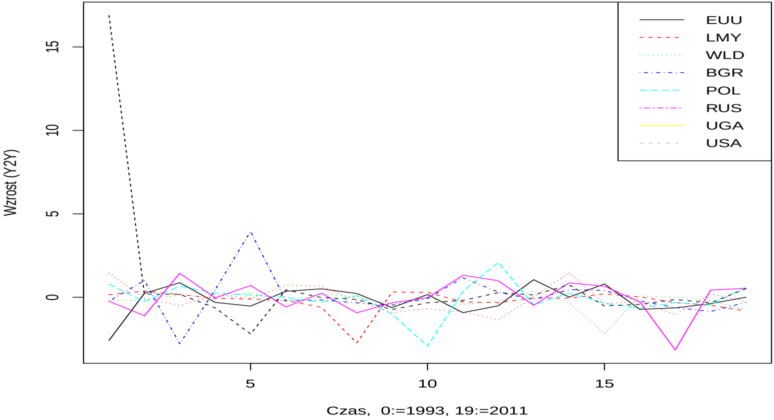
$$\hat{U}_{ij} =< \hat{u}_j(t), y_i(t) >= \int \hat{u}_j(t) y_i(t) dt = \sum_{k=0}^{K_1} \hat{\alpha}_{ik} \hat{u}_{jk} = \hat{\boldsymbol{\alpha}}_i' \hat{\boldsymbol{u}}_j,$$

gdzie $i = 1, 2, \ldots, N, \qquad j = 1, \ldots, min(K_1 + 1, K_2 + 1)$. Analogicznie współrzędna rzutu i-tej realizacji $x_i(t)$ procesu $X(t)$ na j-tą funkcjonalną zmienną kanoniczną jest równa

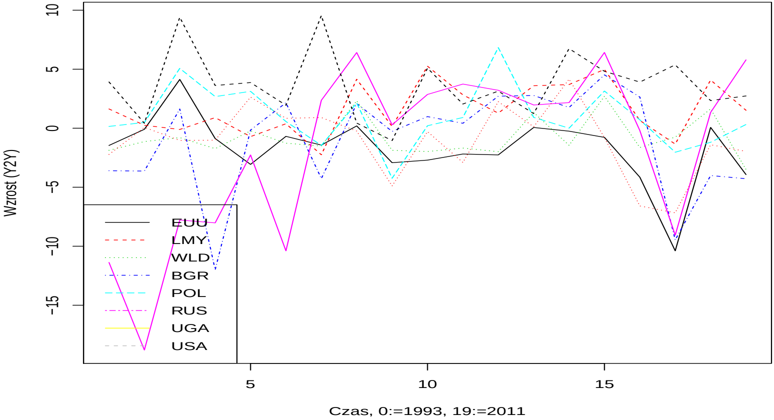
$$\hat{V}_{ij} = \hat{\boldsymbol{\beta}}_i' \hat{\boldsymbol{v}}_j.$$

Dane pochodzą z internetowej bazy danych Banku Światowego (http://data.worldbank.org/).
Do analizy wybrano 8 obszarów świata ($n = 8$)(Unia Europejska (EUU), Kraje o niskich i średnich dochodach (LMY), Cały Świat (WLD), Bułgaria (BGR), Polska (POL), Rosja (RUS), Uganda (UGA), Stany Zjednoczone (USA)) scharakteryzowanych dwiema zmiennymi - **Wzrost PKB** (Y) oraz **Stopa wzrostu bezpośrednich inwestycji zagranicznych** (X) badanymi w latach 1993-2011 ($J = 19$).
Szeregi zostały scentrowane, a do obliczeń użyta została baza Legendre'a w $L_2([-1, 1])$. Optymalne wartość K_1 oraz K_2 wybrane za pomocą kryterium BIC wynoszą dla obu szeregów 2 ($K_1 = 2 = K_2$).

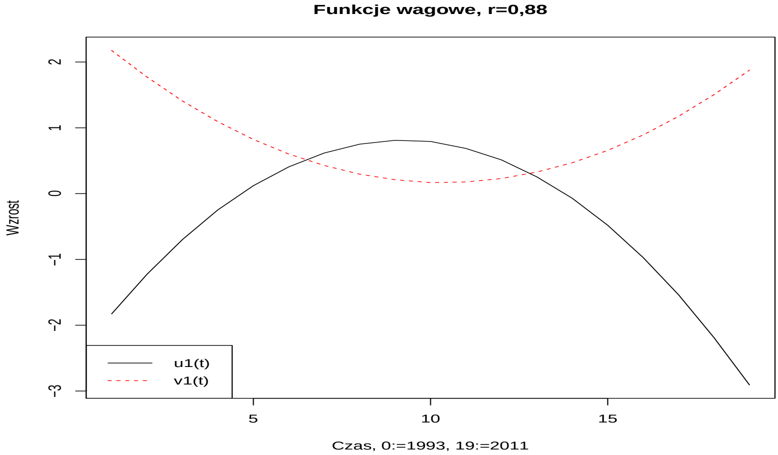
Rysunek: Szereg czasowy wzrostu PKB (Y)



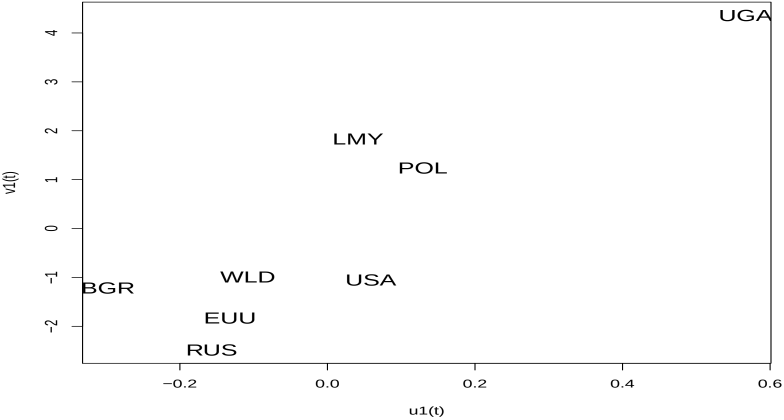
Rysunek: Szereg czasowy stóp wzrostu bezpośrednich inwestycji zagranicznych (X)



Rysunek: Funkcje wagowe $(u_1(t), v_1(t))$



Rysunek: Rzut 8 obszarów świata na płaszczyznę $(u_1(t), v_1(t))$



Literatura

Krzyśko, M. (2009): Podstawy wielowymiarowego wnioskowania statystycznego, Wydawnictwo Naukowe UAM, Poznań.

Krzyśko, M., Waszak., Ł. (2012): Methods of representation for kernel canonical correlation analysis. STATISTICS IN TRANSITION new series - An International Journal of the Polish Statistical Association, Volume 13, Number 2, Summer 2012, 301-311

Leurgans, S.E., Moyeed, R.A. and Silverman B.W. (1993): Canonical correlation analysis when tha data are curves, Journal of the Royal Statistical Society B 55, No. 3, 725-740.

Ramsay, J.O. and Silverman, B.W.: Functional Data Analysis, Second Edition, Springer 2005.

Zheng, W., Zhou, X., Zou, C., Zhao, L. (2006): Facial Expression Recognition Using Kernel Canonical Correlation Analysis, IEEE Transaction on Neural Networks 17(1), 233