Extending Correlation and Regression from Multivariate to Functional Data

By He, Müller & Wang

Presented by Javier Zapata

Preliminaries

- Let $X \in \mathbb{R}^p$ and Y $\in \mathbb{R}^q$

A multivariate linear regression can be defined as

$$\boldsymbol{Y} = \boldsymbol{\alpha} + \boldsymbol{\beta}_0^T \boldsymbol{X} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{\epsilon} \in \mathbf{R^q}$, with $E[\boldsymbol{\epsilon}] = 0$, and $\boldsymbol{\beta}_0 \in \mathbf{R^{p\times q}}$ is the parameter matrix.

In multivariate analysis, we seek the solution of a linear regression model (2.1) by finding the parameter matrix $\beta_0^* \in \mathbf{R^{p\times q}}$ which minimizes the squared distance $E\|Y - \beta X\|^2$. When the covariance matrix of X is invertible, by classical least squares theory (see, *e.g.*, Anderson, 1984), the unique minimizer can be found as

$$\beta_0^* = R_{XX}^{-1} R_{XY}. \tag{2.10}$$

L_2-processes

For a stochastic process with support T, on a probability space Ω, $X(t) = \{X(t,\omega); \omega \in \Omega, t \in T\}$, it holds that $X \in L_2(T)$, if $E \int_T X^2(t)dt < \infty$. For convenience, we will always assume that T and T_1, T_2 below are compact intervals. We note that $L_2(T)$ is a Hilbert space if equipped with the inner product $\langle f, g \rangle = \int_T f(t)g(t)dt$, for $f, g \in L_2(T)$, where dt is the Lebesgue measure. The results can be easily extended to cover more general measures μ and scalar products in spaces $L_2(T; \mu)$.

Covariance Operators and Functions

the infinite-dimensional case, covariance matrices are generalized to *covariance operators*. Specifically, the covariance operator $R_{XX}\colon L_2(T_1) \to L_2(T_1)$, is given by

$$R_{XX}u(s) = \int_{T_1} r_{XX}(s,t)u(t)dt, \quad u \in L_2(T_1). \tag{2.4}$$

where

$$r_{XX}(s,t) = \mathrm{Cov}\big[X(s), X(t)\big], \quad s,t \in T_1,$$

is the covariance function of process X. Similarly we can define covariance functions

$$r_{YY}(s,t) = \mathrm{cov}\big[Y(s), Y(t)\big], \quad s,t \in T_2,$$

$$\text{and} \quad r_{XY}(s,t) = \mathrm{Cov}\big[X(s), Y(t)\big], \quad s \in T_1,\ t \in T_2.$$

The covariance operators$R_{YY}\colon L_2(T_2) \to L_2(T_2)$, $R_{XY}\colon L_2(T_2) \to L_2(T_1)$, and $R_{YX}\colon L_2(T_1) \to L_2(T_2)$ are defined in complete analogy to R_{XX}. We

Karhunen-Loève decomposition for $X(s)$ and $Y(t)$

Using the Karhunen-Loève decomposition, X and Y may be expanded as

$$X(s) = E[X(s)] + \sum_{i=1}^{\infty} \xi_i \theta_i(s), \quad s \in T_1,$$

$$Y(t) = E[Y(t)] + \sum_{i=1}^{\infty} \zeta_i \phi_i(t), t \in T_2, \tag{2.5}$$

with a sequence of uncorrelated random variables ξ_i with $E(\xi_i) = 0$, and a sequence of uncorrelated random variables ζ_i with $E(\zeta_i) = 0$. Here, $\lambda_{Xi} = E[\xi_i^2]$, $\lambda_{Yi} = E[\zeta_i^2]$, $\sum_{i=1}^{\infty} \lambda_{X_i} < \infty$, $\Sigma_{i=1}^{\infty} \lambda_{Y_i} < \infty$, and $\{(\lambda_i, \theta_i)\}$, $\{(\zeta_j, \phi_j)\}$ are the eigenvalues and eigenfunctions of the covariance operators

- We can see that:

$$r_{XY}(s,t) = Cov\big(X(s), Y(t)\big) = \sum_{i,j \geq 1} \mathrm{E}[\xi_i \zeta_j] \theta_i(s) \phi_j(t)$$

The functional linear model for L_2-processes

Consider L_2-processes $X \in L_2(T_1)$, $Y \in L_2(T_2)$. The functional linear regression model is defined as

$$Y(t) = \alpha(t) + \int_{T_1} X(s)\beta_0(s,t)ds + \epsilon(t), \tag{2.7}$$

where $\beta_0 \in L_2(T_1 \times T_2)$ is a parameter function, $\alpha \in L_2(T_2)$ is an intercept function, and $\epsilon \in L_2(T_2)$ is a random error process, with the assumption that X and ϵ are uncorrelated, and that $E[\epsilon(t)] = 0$, for all t.

By assuming, without loss of generality, that $EX(t) = 0$ and $EY(s) = 0$, for all t, s, one may simplify the linear model (2.7) to

$$Y(t) = \int_{T_1} X(s)\beta_0(s,t)ds + \epsilon(t). \tag{2.8}$$

Integral Operator $\mathcal{L}_X$ and its adjoint $\mathcal{L}_X^*$

Define a random integral operator $\mathcal{L}_X \colon L_2(T_1 \times T_2) \to L_2(T_2)$ by

$$(\mathcal{L}_X \beta)(t) = \int_{T_1} X(s)\beta(s,t)ds, \quad \text{for } \beta \in L_2(T_1 \times T_2).$$

It is easy to see that the adjoint operator of $\mathcal{L}_X$ is $\mathcal{L}_X^* \colon L_2(T_2) \to L_2(T_1 \times T_2)$, defined by

$$(\mathcal{L}_X^* z)(s,t) = X(s)z(t), \quad \text{for all } z \in L_2(T_2).$$

Note that (2.8) can be rewritten as

$$Y(t) = (\mathcal{L}_X \beta_0)(t) + \epsilon(t). \tag{2.9}$$

How to achieve a normal equation?

- Let $X \in \mathbb{R}^p$ and Y $\in \mathbb{R}^q$

A multivariate linear regression can be defined as

$$\boldsymbol{Y} = \boldsymbol{\alpha} + \boldsymbol{\beta}_0^T \boldsymbol{X} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{\epsilon} \in \mathbf{R}^{\mathbf{q}}$, with $E[\boldsymbol{\epsilon}] = 0$, and $\boldsymbol{\beta}_0 \in \mathbf{R}^{\mathbf{p} \times \mathbf{q}}$ is the parameter matrix.

- For class LS, if the covariance matrix of X is invertible:

$$\beta_0^* = R_{XX}^{-1} R_{XY}. \tag{2.10}$$

- For the functional linear model, we would like an operator Γ_{XX} to obtain:

$$\beta_0^* = \Gamma_{XX}^{-1} r_{XY}$$

- The model:

$$Y(t) = (\mathcal{L}_X \beta_0)(t) + \epsilon(t). \tag{2.9}$$

For the functional linear model (2.9), we seek a parameter function β_0^* such that

$$\beta_0^* = \arg\,\min_{\beta \in L_2(T_1 \times T_2)} E\|Y - \mathcal{L}_X \beta\|^2. \tag{2.11}$$

PROPOSITION 4.1. *Let β_0 be a solution of the linear regression model (2.9). Then*

$$\beta_0 \in \underset{\beta \in L_2(T_1 \times T_2)}{\arg\,\min}\ E\|Y - \mathcal{L}_X \beta\|^2.$$

- Apply $E\left[\mathcal{L}_X^*(\cdot)\right]$ on both sides of (2.9) to get:

$$\Gamma_{XX}\beta = r_{XY} \qquad \text{for } \beta \in L_2(T_1 \times T_2)$$

- What is Γ_{XX} ? Can we invert it?

Integral Operator Γ_{XX}

Motivated by the form of the least squares solution (2.10) for the multivariate linear model, we define a linear integral operator $\Gamma_{XX}\colon L_2(T_1 \times T_2) \to L_2(T_1 \times T_2)$ as

$$(\Gamma_{XX}\beta)(s,t) = \int_{T_1} r_{XX}(s,w)\beta(w,t)dw.$$

- Some properties of Γ_{XX}
  - *i.* $\Gamma_{XX} = E[\mathcal{L}_X^* \mathcal{L}_X]$
  - *ii.* Γ_{XX} is self-adjoint
  - *iii.* Γ_{XX} is nonnegative Hilbert-Schmidt operator (a class of compact operators)

PROPOSITION 4.2. *Let $\beta \in L_2(T_1 \times T_2)$. Then*

$$\beta \in \underset{\beta \in L_2(T_1 \times T_2)}{\arg\min}\ E\|Y - \mathcal{L}_X\beta\|^2 \textit{ if and only if } \mathrm{P}_{\mathrm{R}(\Gamma_{XX})}\beta = \mathrm{P}_{\mathrm{R}(\Gamma_{XX})}\beta_0,$$

where $\mathrm{P}_{\mathrm{R}(\Gamma_{XX})}$ is the projection from $L_2(T_1 \times T_2)$ to $\mathrm{R}(\Gamma_{XX})$.

Functional normal equation

- The model: $$Y(t) = (\mathcal{L}_X\beta_0)(t) + \epsilon(t). \tag{2.9}$$
- Apply $E[\mathcal{L}_X^*(\cdot)]$ on both sides of (2.9) to get:

$$\Gamma_{XX}\beta = r_{XY} \qquad \text{for } \beta \in L_2(T_1 \times T_2)$$

- What is Γ_{XX} ? Can we invert it?
- To obtain an inverse Γ_{XX}^{-1}, we impose further structure on the problem:

Condition 2.2. L_2-processes X and Y with the expansion (2.5) satisfy

$$\sum_{i,j=1}^{\infty} \frac{E^2[\xi_i \zeta_i]}{\lambda_{Xi}^2} < \infty. \tag{2.12}$$

- More details in Conway (1990), Corollary 5.4

Main result

THEOREM 4.3. *Let X and Y be L_2-processes with the expansion (2.5) which satisfy Condition 2.2. Then,*

(a) $\beta_0^* = \Gamma_{XX}^{-1} r_{XY}$ *exists and is the unique solution of (4.1) in* $R(\Gamma_{XX})$;

(b) β_0^* *has the representation*

$$\beta_0^*(s,t) = \sum_{i,j=1}^{\infty} \frac{E[\xi_i \zeta_i]}{\lambda_{Xi}} \theta_i(s) \phi_j(t);$$

(c) *The set of the solutions of (4.1) is:*

$$\beta_0^* + \ker(\Gamma_{XX}) := \left\{ \beta_0^* + h | h \in \ker(\Gamma_{XX}) \right\},$$

where $\ker(\Gamma_{XX})$ *is the kernel space of* Γ_{XX}, *i.e.,* $\ker(\Gamma_{XX}) = \{h \in L(T_1 \times T_2)\colon \Gamma_{XX} h = 0\}$.

Main result

THEOREM 4.4. *Assume condition 2.4 holds for X and Y. Then*

$$\arg\min_{\beta} E\|Y - \mathcal{L}_X\beta\|^2 = \beta_0^* + \ker(\Gamma_{XX}),$$

and this coincides with the set of solutions for the functional normal equation (4.1).

Bibliography

- Conway – A Course in Functional Analysis, 2nd edition , 1990
  - Chapter 5* - The Diagonalization of Compact Self-Adjoint Operators
- He, Mueller & Wang - Extending correlation and regression from multivariate to functional data , 2000.