Multi-Response Functional Linear Regression

黃敏嘉

中興大學　統研所

Model

Assume that we have n groups of observed functions $(Y_{1,i}, \ldots, Y_{R,i})$, each with associated predictor functions $(X_{1,i}, \ldots, X_{P,i})$.

The $X_{p,i}$ are functions on interval $\mathcal{T}_{X_p}$, and the $Y_{r,i}$ are functions on interval $\mathcal{T}_{Y_r}$, for $p = 1, \ldots, P$; $r = 1, \ldots, R$.

Then the linear model becomes

$$\hat{Y}_{r,i}(t) = \beta_{r0}(t) + \sum_{p=1}^{P} \int_{T_{X_p}} \beta_{rp}(t,s) X_{p,i}(s)\mathrm{d}s; \ \textit{for } r = 1, \ldots, R \quad (1)$$

Estimate intercept functions

One simple approach to the estimation of β_{r0} is to centre the observed $Y_{r,i}$ and the given $X_{p,i}$ by subtracting their sample mean functions $\hat{\mu}_{Y_r}$ and $\hat{\mu}_{X_p}$,then the model becomes

$$\hat{Y}_{r,i}(t) - \hat{\mu}_{Y_r}(t) = \sum_{p=1}^{P} \int_{T_{X_p}} \beta_{rp}(t,s)[X_{p,i}(s) - \hat{\mu}_{X_p}(s)]\mathrm{d}s \tag{2}$$

This constrains the value of $\beta_{r0}(t)$ to be equal to $\hat{\mu}_{Y_r}(t) - \sum_{p=1}^{P} \int_{T_{X_p}} \beta_{rp}(t,s)\hat{\mu}_{X_p}(s)\mathrm{d}s$

Definition

Define

$$\begin{aligned}\Gamma_{X_p,X_{p'}}(s_p, s_{p'}) &= \operatorname{cov}\left(X_p(s_p), X_{p'}(s_{p'})\right)\\ \Gamma_{X_p,Y_r}(s_p, t_r) &= \operatorname{cov}\left(X_p(s_p), Y_r(t_r)\right)\end{aligned}$$

to be the covariances of $(X_p, X_{p'})$ and (X_p, Y_r), respectively.

Definition

Next, we choose $(s_{p1}, \ldots, s_{pM_p})$ and $(t_{r1}, \ldots, t_{rN_r})$ to be the regular dense grids on X_p and Y_r, respectively, for some values M_p, N_r.
Then, we can define

$$\boldsymbol{\Sigma}_{X_p,X_{p'}} = \begin{pmatrix} \Gamma_{X_p,X_{p'}}(s_{p1}, s_{p'1}) & \ldots & \Gamma_{X_p,X_{p'}}(s_{p1}, s_{p'M_{p'}}) \\ \vdots & & \vdots \\ \Gamma_{X_p,X_{p'}}(s_{pM_p}, s_{p'1}) & \ldots & \Gamma_{X_p,X_{p'}}(s_{pM_p}, s_{p'M_{p'}}) \end{pmatrix},$$

$$\boldsymbol{\Sigma}_{X_p,Y_r} = \begin{pmatrix} \Gamma_{X_p,Y_r}(s_{p1}, t_{r1}) & \ldots & \Gamma_{X_p,Y_r}(s_{p1}, t_{rN_r}) \\ \vdots & & \vdots \\ \Gamma_{X_p,Y_r}(s_{pM_p}, t_{r1}) & \ldots & \Gamma_{X_p,Y_r}(s_{pM_p}, t_{rN_r}) \end{pmatrix}$$

Definition

We can further define

$$\boldsymbol{\Sigma}_{X,X} = \begin{pmatrix} \boldsymbol{\Sigma}_{X_1,X_1} & \cdots & \boldsymbol{\Sigma}_{X_1,X_P} \\ \vdots & & \vdots \\ \boldsymbol{\Sigma}_{X_P,X_1} & \cdots & \boldsymbol{\Sigma}_{X_P,X_P} \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{X,Y} = \begin{pmatrix} \boldsymbol{\Sigma}_{X_1,Y_1} & \cdots & \boldsymbol{\Sigma}_{X_1,Y_R} \\ \vdots & & \vdots \\ \boldsymbol{\Sigma}_{X_P,Y_1} & \cdots & \boldsymbol{\Sigma}_{X_P,Y_R} \end{pmatrix}$$

Definition

Let

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{11}(s_{11}, t_{11}) & \dots & \beta_{11}(s_{11}, t_{1N_1}) & \dots & \beta_{R1}(s_{11}, t_{RN_R}) \\ \vdots & & \vdots & & \vdots \\ \beta_{11}(s_{1M_1}, t_{11}) & \dots & \beta_{11}(s_{11}, t_{1N_1}) & \dots & \beta_{R1}(s_{1M_1}, t_{RN_R}) \\ \vdots & & \vdots & & \vdots \\ \beta_{1P}(s_{P1}, t_{11}) & \dots & \beta_{1P}(s_{P1}, t_{1N_1}) & \dots & \beta_{RP}(s_{P1}, t_{RN_R}) \\ \vdots & & \vdots & & \vdots \\ \beta_{1P}(s_{PM_P}, t_{11}) & \dots & \beta_{1P}(s_{PM_P}, t_{1N_1}) & \dots & \beta_{RP}(s_{PM_P}, t_{RN_R}) \end{pmatrix}$$

Estimate coefficient functions

The optimal choice of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\Delta_s)^{-1}(\boldsymbol{\Sigma}_{X,X})^{-1}\boldsymbol{\Sigma}_{X,Y}$$

where
$\Delta_s = \mathrm{diag}(\Delta_{s_1}, \ldots, \Delta_{s_P})$ is a block diagonal matrix, with
$\Delta_{s_p} = \mathrm{diag}(\delta_{s_p}, \ldots, \delta_{s_p})$ is a $(M_p \times M_p)$ matrix, and
$\delta_{s_p} = s_{p2} - s_{p1}$ is the width of grid points on $\mathcal{T}_{X_p}$

Prediction

If we have new observations $(X_1^*, \ldots, X_P^*)$, then the predictions $(Y_1^*, \ldots, Y_R^*)$ becomes

$$\hat{Y}_r^*(t) = \hat{\mu}_{Y_r}(t) + \sum_{p=1}^{P} \int_{T_{X_p}} \beta_{rp}(t,s)[X_p^*(s) - \hat{\mu}_{X_p}(s)]\mathrm{d}s, \quad r = 1, \ldots, R \tag{3}$$

Prediction

We can get the estimation $\hat{Y}^*_r$ on the grid points $(t_{r1}, \ldots, t_{rN_r})$, $r = 1, \ldots, R$ by

$$\begin{aligned}\hat{\boldsymbol{Y}}^* &= \hat{\boldsymbol{\mu}}_Y + \hat{\boldsymbol{\beta}}_r^T \, \Delta_s \, (\hat{\boldsymbol{X}}^* - \hat{\boldsymbol{\mu}}_X) \\ &= \hat{\boldsymbol{\mu}}_Y + \hat{\boldsymbol{\Sigma}}_{X,Y}^T (\hat{\boldsymbol{\Sigma}}_{X,X})^{-1} \, (\hat{\boldsymbol{X}}^* - \hat{\boldsymbol{\mu}}_X)\end{aligned}$$

where

$$\hat{\boldsymbol{Y}}^* = \begin{pmatrix} \hat{Y}_1^*(t_{11}) \\ \vdots \\ \hat{Y}_1^*(t_{1N_1}) \\ \vdots \\ \hat{Y}_R^*(t_{RN_R}) \end{pmatrix}, \hat{\boldsymbol{\mu}}_{Y_r} = \begin{pmatrix} \hat{\mu}_{Y_1}(t_{11}) \\ \vdots \\ \hat{\mu}_{Y_1}(t_{1N_1}) \\ \vdots \\ \hat{\mu}_{Y_R}(t_{RN_R}) \end{pmatrix}, \hat{\boldsymbol{X}}^* = \begin{pmatrix} \hat{X}_1(s_{11}) \\ \vdots \\ \hat{X}_1(s_{1M_1}) \\ \vdots \\ \hat{X}_P(s_{PM_P}) \end{pmatrix},$$

$$\hat{\boldsymbol{\mu}}_X = \begin{pmatrix} \hat{\mu}_{X_1}(s_{11}) \\ \vdots \\ \hat{\mu}_{X_1}(s_{1M_1}) \\ \vdots \\ \hat{\mu}_{X_P}(s_{PM_P}) \end{pmatrix}$$