Supplementary Document for the Manuscript entitled "Two-Dimensional Functional Principal Component Analysis for Image Feature Extraction"

Proofs of Theorems 1 and 2

Since $\Psi^0 := \{\psi_m^0 : m = 1, 2, \cdots, \infty\}$ is a complete orthonormal system in $L^2(\mathcal{T})$, we represent $\psi_1(\mathbf{t})$ as $\psi_1(\mathbf{t}) = \boldsymbol{\gamma}^\top \Psi^0$ for coefficient vector $\boldsymbol{\gamma} = (\gamma_m)$.

The following assumptions are required to complete the proof.

(A1) The $\widehat{\mu}(\mathbf{t})$ is a consistent estimator of $\mu(\mathbf{t})$.

(A2) There exists an $M > 0$ s.t. $\psi_m^0(\mathbf{t}) < M$ for all $\mathbf{t} \in \mathcal{T}$ and each $m = 1, 2, \cdots, \infty$. The sum $\sum_{m=1}^{\infty} E(\xi_m^4)$ is bounded.

(A3) The set $\Theta = \{((\alpha_{i1}), (\gamma_m)) \in \mathcal{C}_{00} \oplus \mathcal{B}_{\ell_2}\}$ is manageable (Pollard, 1989), where $\mathcal{C}_{00} = \{(c_i) : |c_i| < C$ for some constant C and $c_i = 0$ for $i \geq l$ for some $l\}$ and $\mathcal{B}_{\ell_2} = \{(c_i) : \sum_{i=1}^{\infty} c_i^2 \leq 1\}$.

We rely upon Pollard's uniform law of large number to complete the proofs of the theorems, which is given below (Pollard, 1989).

Theorem A.1. (Pollard's Uniform Law of Large Number). *Let $\{f_i(\omega, \theta) : \theta \in \Theta\}$ be a sequence of independent processes that are manageable for their envelopes $\{F_i(\omega) = \sup_{\theta \in \Theta} |f_i(\omega, \theta)|\}$. If*

$$\sum_i \frac{E\{F_i(\omega)\}}{i^2} < \infty,$$

then

$$\frac{1}{n} \sup_{\theta \in \Theta} |S_n(w, \theta) - E\{S_n(w, \theta)\}| \to 0 \quad \textit{almost surely.}$$

Proof of Theorem 1.

Denote the loss function as $L_n(\boldsymbol{\alpha}_1, \psi_1)$ and without loss of generality, assume $\mathcal{T} = [0,1]^2$. Since $y_{ij}^* = \sum_{m=1}^{\infty} \xi_m \psi_m^0(\mathbf{t}_{ij}) + \epsilon_{ij}$ and $\psi_1(\mathbf{t}_{ij}) = \sum_{m=1}^{\infty} \gamma_m \psi_m^0(\mathbf{t}_{ij})$, we have

$$\begin{aligned}
L_n(\boldsymbol{\alpha}_1, \psi_1) &\leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i} 2\Big[\big\{y_{ij}^* - \alpha_{i1}\psi_1(\mathbf{t}_{ij})\big\}^2 + \big\{\widehat{\mu}(\mathbf{t}_{ij}) - \mu(\mathbf{t}_{ij})\big\}^2\Big] \\
&= 2\Big[\frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{\infty}(\xi_{im} - \alpha_i\gamma_m)^2\rho_{imm} + \sum_{m\neq l}(\xi_{im} - \alpha_i\gamma_m)(\xi_{il} - \alpha_i\gamma_l)\rho_{iml} \\
&\quad - \frac{2}{n}\sum_{i=1}^{n}\sum_{m=1}^{\infty}(\xi_{im} - \alpha_i\gamma_m)\frac{1}{n_i}\sum_{j=1}^{n_i}\epsilon_{ij}\psi_m^0(\mathbf{t}_{ij}) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i}\epsilon_{ij}^2 + \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i}\big\{\widehat{\mu}(\mathbf{t}_{ij}) - \mu(\mathbf{t}_{ij})\big\}^2\Big],
\end{aligned}$$

where $\rho_{iml} = \frac{1}{n_i}\sum_{j=1}^{n_i}\psi_m^0(\mathbf{t}_{ij})\psi_l^0(\mathbf{t}_{ij})$. Since $\mathbf{t}_{ij}$ are uniformly drawn from $\mathcal{T}$, we have

$$E(\rho_{iml}) = \frac{1}{n_i}\sum_{j=1}^{n_i}\int_0^1 \{\psi_m^0(\mathbf{t}_{ij})\psi_l^0(\mathbf{t}_{ij})\}\, d\mathbf{t}_{ij} = \delta_{ml}.$$

We will show that

$$L_n(\boldsymbol{\alpha}_1, \psi_1) = \frac{2}{n}\sum_{i=1}^{n}\sum_{m=1}^{\infty}(\xi_{im} - \alpha_i\gamma_m)^2 + o_p(1) + \sigma^2,$$

where $o_p(1)$ is some random quantity uniformly small over the parameter set Θ in probability.

By the law of large number, it can be shown that

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i}\epsilon_{ij}^2 = \sigma^2 + o_p(1),$$

and based on assumption (A1),

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i}\big\{\widehat{\mu}(\mathbf{t}_{ij}) - \mu(\mathbf{t}_{ij})\big\}^2 = o_p(1).$$

In order to show

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{\infty}(\xi_{im} - \alpha_i\gamma_m)^2\rho_{imm} - \frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{\infty}(\xi_{im} - \alpha_i\gamma_m)^2 = o_p(1), \\
&\frac{1}{n}\sum_{i=1}^{n}\sum_{m\neq l}(\xi_{im} - \alpha_i\gamma_m)(\xi_{il} - \alpha_i\beta_l)\rho_{iml} = o_p(1), \\
&\frac{-2}{n}\sum_{i=1}^{n}\sum_{m=1}^{\infty}(\xi_{im} - \alpha_i\gamma_m)\frac{1}{n_i}\sum_{j=1}^{n_i}\epsilon_{ij}\psi_m^0(\mathbf{t}_{ij}) = o_p(1),
\end{aligned}$$

we check the conditions of Pollard's uniform law of large number, respectively,

$$\sum_i E\bigg\{\sup_\Theta \sum_{m=1}^\infty (\xi_{im}-\alpha_i\gamma_m)^2\rho_{imm}\bigg\}^2 \bigg/ i^2 < \infty,$$
$$\sum_i E\bigg\{\sup_\Theta \sum_{m\neq l} (\xi_{im}-\alpha_i\gamma_m)(\xi_{il}-\alpha_i\gamma_l)\rho_{iml}\bigg\}^2 \bigg/ i^2 < \infty,$$
$$\sum_i E\bigg\{\sup_\Theta \sum_{m=1}^\infty (\xi_{im}-\alpha_i\gamma_m)\frac{1}{n_i}\sum_{j=1}^{n_i}\epsilon_{ij}\psi_m^0(\mathbf{t}_{ij})\bigg\}^2 \bigg/ i^2 < \infty.$$

By the boundedness of the FPC ψ_m^0 and the Cauchy–Schwartz inequality, it suffices to check

$$E\bigg\{\sup_\Theta \sum_{m=1}^\infty (\xi_{im}-\alpha_i\gamma_m)^2\bigg\}^2 < \infty,$$

which follows from assumptions (A2)–(A3). If we denote

$$\widetilde{\boldsymbol{\alpha}}_n, \widetilde{\boldsymbol{\gamma}}_n = \arg\min \frac{1}{n}\sum_{i=1}^n\sum_{m=1}^\infty (\xi_{im}-\alpha_i\gamma_m)^2,$$
$$\widehat{\boldsymbol{\alpha}}_n, \widehat{\boldsymbol{\gamma}}_n = \arg\min \frac{1}{n}\sum_{i=1}^n\sum_{m=1}^\infty L_n(\boldsymbol{\alpha}_1, \psi_1),$$

then

$$\widehat{\boldsymbol{\alpha}}_n = \widetilde{\boldsymbol{\alpha}}_n + o_p(1), \qquad \widehat{\boldsymbol{\gamma}}_n = \widetilde{\boldsymbol{\gamma}}_n + o_p(1).$$

By the law of large number,

$$\frac{1}{n}\sum_{i=1}^n \xi_{im}\xi_{il} \to E(\xi_{im}\xi_{il}) = \lambda_m\delta_{ml}.$$

For fixed p, denote $A = (\frac{1}{\sqrt{n}}\xi_{im})_{i=1,\ldots,n}^{m=1,\cdots,p}$, we have

$$(A^\top A)_{ij} \to \lambda_m\delta_{ml}.$$

Consequently,

$$\|A^\top A - \mathrm{diag}[\lambda_1,\cdots,\lambda_p]\|_F \to 0.$$

Since for any two projection matrices P and Q, $\|P - Q\|_F \geq \sum_k \|u_k - v_k\|$, where (u_k) and (v_k) are eigenvectors of P and Q, the eigenvectors (u_k) of $A^\top A$ converges to those of $\mathrm{diag}[\lambda_1, \cdots, \lambda_p]$; and the eigenvectors v_k of $AA^\top$ are $v_k = A^\top u_k$, which will converges to the columns of A in the space $\mathcal{C}_{00}$ of sequences with finitely support equipped with the ℓ_2 norm. Hence the estimate $(\widetilde{\alpha}_i)$, the first eigenvector of $AA^\top$, will converges to the first column of A in ℓ_2 norm: $(\frac{1}{\sqrt{n}}\widetilde{\alpha}_i) - (\frac{1}{\sqrt{n}}\xi_{i1})$ converges to 0 in ℓ_2. Also, $\widetilde{\boldsymbol{\gamma}}$ converges to $\mathbf{e}_1$, the first unit vector. Hence $\widehat{\psi}_1(t) := \widehat{\boldsymbol{\gamma}}_n \Psi^0(t)$ converges to $\psi_1^0(t)$ in $L^2(\mathcal{T})$.

Proof of Theorem 2.

Without loss of generality, consider the case where $M = 2$. Since the estimates $\widehat{\alpha}_{i1}$ and $\hat{\psi}_1$ are such that

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i}\left\{\alpha_{i1}\psi_1^0(\mathbf{t}_{ij}) - \widehat{\alpha}_{i1}\widehat{\psi}_1(\mathbf{t}_{ij})\right\}^2 = o_p(1),$$

hence we can rewrite the model as

$$y_{ij}^* - \widehat{\alpha}_i\widehat{\psi}_1(\mathbf{t}_{ij}) = \sum_{m=2}^{\infty} a_{im}\psi_m^0(\mathbf{t}_{ij}) + o_p(1).$$

By the same argument as in the proof of Theorem 1, the proof is complete.