Supplementary Material for "M-Based Simultaneous Inference for the Mean Function of Functional Data"

By Italo R. Lima, Guanqun Cao and Nedret Billor
Auburn University

In this supplementary material, we give proofs to the Theorems in the Section 2.3. We begin by proving several useful lemmas.

Asymptotic Consistency

To prove Theorem 1, we need to introduce several lemmas firstly. The following lemma provides an initial approximation of $m(x)$ by a spline function, resulting in an approximation bias.

Lemma 1 *(Corollary 6.21, Schumaker (2007)) Define $R_n(x) = \boldsymbol{B}^{\mathrm{T}}(x)\boldsymbol{\beta}^* - m(x)$. If Assumption (A1) holds, then there exists $\boldsymbol{\beta}^* \in \mathbb{R}^{N_m+p}$ such that*

$$\sup_{x\in[0,1]} |R_n(x)| = \sup_{x\in[0,1]} |\boldsymbol{B}^{\mathrm{T}}(x)\boldsymbol{\beta}^* - m(x)| = O(N_m^{-p}).$$

The above results states that to prove Theorem 1, we can replace m by $m^*(\cdot) = \boldsymbol{B}^{\mathrm{T}}(\cdot)\boldsymbol{\beta}^*$ up to a order of N_m^{-p}. In particular, we just need to prove $\|\hat{m} - m^*\|_2^2 = O_P\left(n^{-1}N_m\right)$. Define

$$\mathbb{S}_n = \frac{n}{N}\sum_{j=1}^{N} \boldsymbol{B}\left(\frac{j}{N}\right)\boldsymbol{B}^{\mathrm{T}}\left(\frac{j}{N}\right), \quad \tilde{\boldsymbol{B}}\left(\frac{j}{N}\right) = \mathbb{S}_n^{-1/2}\boldsymbol{B}\left(\frac{j}{N}\right) \text{ and } \boldsymbol{\theta} = \mathbb{S}_n^{1/2}(\boldsymbol{\beta} - \boldsymbol{\beta}^*), \quad (1)$$

where $\boldsymbol{\beta}^*$ is defined in Lemma 1.

Lemma 2 *(Lemma A.3, Cao et al. (2012)) There are two positive constants M_1 and M_2, such that except on an event whose probability tends to zero, all the eigenvalues of $(N_m/n)\mathbb{S}_n$ fall between M_1 and M_2, and $\mathbb{S}_n$ is invertible consequently.*

By Lemma 2, and the definition of $\tilde{\boldsymbol{B}}(\cdot)$ in (1), we have $Y_{ij} - \boldsymbol{B}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\beta} = e_{ij} + R_{nj} - \tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}$. This implies that

$$\begin{aligned} &\min_{\boldsymbol{\beta}\in\mathbb{R}^{N_m+p}} \sum_{i=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\rho\left(Y_{ij} - \boldsymbol{B}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\beta}\right) \\ =&\min_{\boldsymbol{\theta}\in\mathbb{R}^{N_m+p}} \sum_{i=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\left[\rho\left(e_{ij} + R_{nj} - \tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right) - \rho\left(e_{ij} + R_{nj}\right)\right], \end{aligned} \quad (2)$$

where $\rho(e_{ij}+R_{nj})$ is a constant term with respect to $\boldsymbol{\theta}$, so it does not change the minimum. Let $\mathit{\Gamma}_n(\theta)$ be the operator defined in (2), that is,

$$\mathit{\Gamma}_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{1}{N} \sum_{j=1}^{N} \left[\rho \left(e_{ij} + R_{nj} - \tilde{\boldsymbol{B}}^{\mathrm{T}} \left(\frac{j}{N} \right) \boldsymbol{\theta} \right) - \rho \left(e_{ij} + R_{nj} \right) \right],$$

and define

$$\mathit{\Delta}_n(\boldsymbol{\theta}) = \mathit{\Gamma}_n(\boldsymbol{\theta}) - \mathbb{E}\left(\mathit{\Gamma}_n(\boldsymbol{\theta})\right) + \sum_{i=1}^{n} \frac{1}{N} \sum_{j=1}^{N} \left[\psi(e_{ij}) \tilde{\boldsymbol{B}}^{\mathrm{T}} \left(\frac{j}{N} \right) \boldsymbol{\theta} \right]. \tag{3}$$

Before proving Theorem 1, we need to obtain asymptotic upper bounds on $\mathit{\Delta}_n(\boldsymbol{\theta})$, $\mathit{\Gamma}_n(\boldsymbol{\theta})$ and $\mathbb{E}\left(\mathit{\Gamma}_n(\boldsymbol{\theta})\right)$ in the following three lemmas.

Lemma 3 *Under Assumptions (A1) - (A4) and (A6) and for a fixed constant $L > 1$,*

$$\sup_{\|\boldsymbol{\theta}\|_2 < 1} \left| \frac{1}{N_m} \mathit{\Delta}_n \left(N_m^{1/2} L\, \boldsymbol{\theta} \right) \right| = o_P(1). \tag{4}$$

Proof Bernstein's theorem is used to prove that $\mathbb{P}\left(\sup_{\|\boldsymbol{\theta}\|_2 \le 1} \frac{1}{N_m} \left| \mathit{\Delta}_n \left(N_m^{1/2} L\, \boldsymbol{\theta} \right) \right| \ge \varepsilon \right)$ $\to 0$, for any $\varepsilon > 0$. To do this, first define $\mathit{\Pi} = \left\{ \boldsymbol{\theta} \in \mathbb{R}^{N_m+p}; \|\boldsymbol{\theta}\|_2 \le 1 \right\}$, and find a decomposition $\mathit{\Pi} = \mathit{\Pi}_1 \cup \cdots \cup \mathit{\Pi}_{K_n}$, where $\{\mathit{\Pi}_k\}_{k=1}^{K_n}$ are pairwise disjoint sets and for any $1 \le k \le K_n$, $\mathrm{diam}(\mathit{\Pi}_k) = \max_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathit{\Pi}_k} \{ \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2 \} \le q_o = \varepsilon C^{-1} N_m n^{-1}$. Notice that we can find such decomposition with $K_n \le \left(2 \frac{\sqrt{N_m + p}}{q_0} + 1 \right)^{N_m+p}$. For each $1 \le k \le K_n$, select $\boldsymbol{\theta}_k \in \mathit{\Pi}_k$. Then, by (3) we have

$$\begin{aligned}
&\min_{1 \le k \le K_n} \frac{1}{N_m} \left| \mathit{\Delta}_n \left(N_m^{1/2} L\, \boldsymbol{\theta} \right) - \mathit{\Delta}_n \left(N_m^{1/2} L\, \boldsymbol{\theta}_k \right) \right| \\
&\le \min_{1 \le k \le K_n} \frac{1}{N_m} \left| \mathit{\Gamma}_n(N_m^{1/2} L\, \boldsymbol{\theta}) - \mathit{\Gamma}_n(N_m^{1/2} L\, \boldsymbol{\theta}_k) \right| \\
&\quad + \min_{1 \le k \le K_n} \frac{1}{N_m} \left| \mathbb{E} \left(\mathit{\Gamma}_n(N_m^{1/2} L\, \boldsymbol{\theta}) \right) - \mathbb{E} \left(\mathit{\Gamma}_n(N_m^{1/2} L\, \boldsymbol{\theta}_k) \right) \right| \\
&\quad + \min_{1 \le k \le K_n} \frac{1}{N_m} \left| N_m^{1/2} L \sum_{i=1}^{n} \frac{1}{N} \sum_{j=1}^{N} \left[\psi(e_{ij}) \cdot \tilde{\boldsymbol{B}}^{\mathrm{T}} \left(\frac{j}{N} \right) \boldsymbol{\theta} - \psi(e_{ij}) \cdot \tilde{\boldsymbol{B}}^{\mathrm{T}} \left(\frac{j}{N} \right) \boldsymbol{\theta}_k \right] \right| \\
&= I + II + III.
\end{aligned}$$

We proceed by obtaining an asymptotic upper bound for *I, II* and *III*. Using the definition of $\mathit{\Gamma}_n(\boldsymbol{\theta})$, we have

$$I = \min_{1\le k\le K_n} \frac{1}{N_m}\left|\sum_{i=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\left[\rho\left(e_{ij}+R_{nj}-N_m^{1/2}L\,\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right)-\rho\left(e_{ij}+R_{nj}\right)\right]\right.$$
$$\left.-\sum_{i=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\left[\rho\left(e_{ij}+R_{nj}-N_m^{1/2}L\,\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}_k\right)-\rho\left(e_{ij}+R_{nj}\right)\right]\right|,$$

and using the mean value theorem on $\rho(\cdot)$ and Assumption (A2), we have

$$I \le N_m^{-1}CnN_m^{1/2}L\max_{1\le j\le N}\left\|\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\right\|_2\min_{1\le k\le K_n}\|\boldsymbol{\theta}-\boldsymbol{\theta}_k\|_2.$$

Similarly, $\max\{II, III\} \le N_m^{-1}CnN_m^{1/2}L\max_{1\le j\le N}\left\|\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\right\|_2\min_{1\le k\le K_n}\|\boldsymbol{\theta}-\boldsymbol{\theta}_k\|_2$. Combining the previous results, we have

$$\begin{aligned}
&\min_{1\le k\le K_n}\frac{1}{N_m}\left|\mathit{\Delta}_n\left(N_m^{1/2}L\,\boldsymbol{\theta}\right)-\mathit{\Delta}_n\left(N_m^{1/2}L\,\boldsymbol{\theta}_k\right)\right|\\
&\le 3\,N_m^{-1}C\min_{1\le k\le K_n}\|\boldsymbol{\theta}-\boldsymbol{\theta}_k\|_2\,nN_m^{1/2}L\max_{1\le j\le N}\left\|\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\right\|_2\\
&\le 3N_m^{-1}Cq_0d_n = 3\varepsilon d_n,
\end{aligned}\tag{5}$$

where $d_n = N_m\left[\max_{1\le j\le N}\left(\left\|\tilde{\boldsymbol{B}}\left(\frac{j}{N}\right)\right\|_2 LN_m^{1/2}+|R_{nj}|\right)\right]$.

By Lemma 2 one has $N_m\left[\max_{1\le j\le N}\left\|\tilde{\boldsymbol{B}}\left(\frac{j}{N}\right)\right\|_2 LN_m^{1/2}\right] = O\left(N_m^{3/2}n^{-1/2}\right) = o(1)$. According to Lemma 1, one has $N_m\max_{1\le j\le N}|R_{nj}| = O\left(N_m^{1-p}\right) = o(1)$. Combining these two upper bounds we have $d_n = o(1)$. In particular, by (5) and choosing $d_n < 1/12$, we have

$$\min_{1\le k\le K_n}\frac{1}{N_m}\left|\mathit{\Delta}_n\left(N_m^{1/2}L\,\boldsymbol{\theta}\right)-\mathit{\Delta}_n\left(N_m^{1/2}L\,\boldsymbol{\theta}_k\right)\right| < \varepsilon/4. \tag{6}$$

For any $1\le i\le n$, $1\le j\le N$, and $\boldsymbol{\theta}\in\mathbb{R}^{N_m+p}$, define

$$\mathit{\Omega}_{ij}(\boldsymbol{\theta}) = \rho\left(e_{ij}+R_{nj}-N_m^{1/2}L\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right)-\rho\left(e_{ij}+R_{nj}\right)+\psi(e_{ij})\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}.$$

Using an argument similar to equation (5), we can prove that $\sup_{\|\boldsymbol{\theta}\|_2\le 1}|\mathit{\Omega}_{ij}(\boldsymbol{\theta})| = O(N_m^{-1})$, and consequently $\sup_{\|\boldsymbol{\theta}\|_2\le 1}|\mathit{\Omega}_{ij}(\boldsymbol{\theta})-\mathbb{E}(\mathit{\Omega}_{ij}(\boldsymbol{\theta}))| = O(N_m^{-1})$. Using the previous equation, we have $\sup_{\|\boldsymbol{\theta}\|_2\le 1}\sum_{i=1}^{n}\mathrm{Var}\left(\frac{1}{N}\sum_{j=1}^{N}\mathit{\Omega}_{ij}(\boldsymbol{\theta})\right) = O\left(nN^{-1}N_m^{-1}\right)$. Using the above three upper bounds, Bernstein's Inequality and Assumption (A1), we have

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{\|\boldsymbol{\theta}\|_2\leq 1}\frac{1}{N_m}\left|\Delta_n\left(N_m^{1/2}L\boldsymbol{\theta}\right)\right|\geq\varepsilon\right)\\
&\leq\sum_{k=1}^{K_n}\mathbb{P}\left(\left|\Delta_n\left(N_m^{1/2}L\boldsymbol{\theta}_k\right)\right|\geq\frac{\varepsilon N_m}{2}\right)\\
&\leq\sum_{k=1}^{K_n}\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{N}\sum_{j=1}^{N}\left|\Omega_{ij}(\boldsymbol{\theta}_k)-\mathbb{E}(\Omega_{ij}(\boldsymbol{\theta}_k))\right|\right]\geq\frac{\varepsilon N_m}{2n}\right)\\
&\leq K_n\exp\left(-\frac{Cn\left(\varepsilon N_m/2n\right)^2}{N^{-1}N_m^{-1}+N_m^{-1}(\varepsilon N_m/2n)}\right)\\
&\leq K_n\exp\left(-\frac{C\varepsilon^2NN_m^3}{n+\varepsilon NN_m}\right)=o(1).
\end{aligned}
$$

Therefore, the Lemma 3 is proved. □

The asymptotic bound for $\Gamma_n(\boldsymbol{\theta})$ is given by the following lemma.

Lemma 4 *Under the Assumptions (A1) and (A4), for a fixed constant $L>1$,*

$$
\sup_{\|\boldsymbol{\theta}\|_2<L}\left|N_m^{-1/2}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\left[\psi(e_{ij})\cdot\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right]\right|=o_P(1). \tag{7}
$$

Proof Notice that by Assumption (A4) and Lemma 2, one has

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{N_m^{1/2}}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\left[\psi(e_{ij})\cdot\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right]\right)\\
&\leq\frac{1}{N_m}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\mathbb{E}\left[\psi(e_{ij})\right]^2\left[\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right]^2\leq CN_m^{-1}\|\boldsymbol{\theta}\|_2.
\end{aligned}
$$

Using Tchebychev's Inequality, the lemma is proved. □

The last asymptotic bound needed is given by the following lemma.

Lemma 5 *Under Assumptions (A1) - (A5) and for a fixed constant $L>1$,*

$$
\mathbb{P}\left(\inf_{\|\boldsymbol{\theta}\|_2=L}\left|\frac{1}{N_m}\mathbb{E}\left[\Gamma_n\left(N_m^{1/2}\boldsymbol{\theta}\right)\right]\right|>0\right)\to 1.
$$

Proof By Lemma 2 we can assume that $\sup_{\|\boldsymbol{\theta}\|_2 \le L}\left(|R_{nj}| + N_m^{1/2}\left\|\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right\|_2^2\right) < C$. By Assumption (A4) we have

$$
\begin{aligned}
& N_m^{-1}\mathbb{E}\left(\Gamma_n\left(N_m^{1/2}L\boldsymbol{\theta}\right)\right) \\
&= N_m^{-1}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\int_{R_{nj}}^{R_{nj}-N_m^{1/2}\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}}\mathbb{E}\left(\psi(e_{ij}+u)\right)du \\
&= N_m^{-1}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\int_{R_{nj}}^{R_{nj}-N_m^{1/2}\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}}\delta\left(\frac{j}{N}\right)u + O(u^2)du \\
&= N_m^{-1}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\delta\left(\frac{j}{N}\right)\frac{1}{2}\left[\left(R_{nj} - N_m^{1/2}\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right)^2 - R_{nj}^2\right] + o(1) \\
&= \sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\delta\left(\frac{j}{N}\right)\left[\frac{1}{2}\left(\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right)^2 - N_m^{-1/2}R_{nj}\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right] + o(1) \\
&\ge Cn\inf_{x\in\mathbb{R}}\delta(x)\|\boldsymbol{\theta}\|_2^2 - \sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}N_m^{-1/2}R_{nj}\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta} + o(1) \\
&= CnL^2 - CnL + o(1),
\end{aligned} \tag{8}
$$

which is positive for large enough L. This finishes the proof of the lemma. □

The following lemma is standard in the spline approximation theory and we omit the proof here.

Lemma 6 (Theorem 5.4.2, DeVore and Lorentz (1993)) *There is a constant $C_p > 0$, such that for any spline $S(\cdot) = \sum_{J=1-p}^{N_m}\gamma_J B_J(\cdot)$ of order p, and for each $0 < p \le \infty$, $C_p N_m^{-1}\|\boldsymbol{\gamma}\|_2^2 \le \|S\|_2^2 \le N_m^{-1}\|\boldsymbol{\gamma}\|_2^2$, where $\boldsymbol{\gamma} = (\gamma_{1-p}, \ldots, \gamma_{N_m})^{\mathrm{T}}$.*

Proof (Proof of Theorem 1) Combining Lemmas 3, 4 and 5, and using the convexity of $\rho(\cdot)$ we have

$$
\mathbb{P}\left(\inf_{\|\boldsymbol{\theta}\|_2\ge L}\frac{1}{N_m}\Gamma(N_m^{1/2}\boldsymbol{\theta}) > 0\right) = \mathbb{P}\left(\inf_{\|\boldsymbol{\theta}\|_2 = L}\frac{1}{N_m}\Gamma(N_m^{1/2}\boldsymbol{\theta}) > 0\right) \to 1.
$$

This in turn implies

$$
\begin{aligned}
&\mathbb{P}\left(\inf_{\|\boldsymbol{\theta}\|_2\ge LN_m^{1/2}}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\rho\left(e_{ij} + R_{nj} - N_m^{1/2}L\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\boldsymbol{\theta}\right) > \sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\rho\left(e_{ij}+R_{nj}\right)\right) \\
&\to 1.
\end{aligned} \tag{9}
$$

Define $\hat{\boldsymbol{\theta}} = \boldsymbol{S}_n^{-1/2}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\right) = \arg\min_{\boldsymbol{\theta}\in\mathbb{R}^{N_m+p}} \mathit{\Gamma}_n(\boldsymbol{\theta})$. By equation (9) one has $\|\hat{\boldsymbol{\theta}}\|_2 = O_P(N_m^{1/2})$, and using Lemma 2, we obtain $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_2^2 = O_P\left(n^{-1}N_m^2\right)$. The approximation property of B-Splines implies that $\|\hat{m}(\cdot) - \boldsymbol{B}(\cdot)^{\mathrm{T}}\boldsymbol{\beta}^*\|_2^2 = \|\boldsymbol{B}(\cdot)^{\mathrm{T}}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)\|_2^2 = O(N_m^{-1})\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_2^2 = O_P\left(n^{-1}N_m\right)$, where the second-to-last equality comes from Lemma 6. Finally, by Lemma 1 and Assumption (A2), Theorem 1 is proved. □

Asymptotic Normality

In this section we will prove the asymptotic normality of the estimator $\hat{m}(x)$ for $0 \leq x \leq 1$. Let $\widetilde{\mathbb{W}}_n = \mathbb{S}_n^{-1/2}\mathbb{W}_n\mathbb{S}_n^{-1/2}$ and $\tilde{\boldsymbol{\theta}} = \widetilde{\mathbb{W}}_n^{-1}\sum_{i=1}^n \frac{1}{N}\sum_{j=1}^N \tilde{\boldsymbol{B}}\left(\frac{j}{N}\right)\psi(e_{ij})$, Where $\mathbb{W}_n$ was defined in equation (5). The first step is to obtain an asymptotic upper bound on the difference between $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$.

Lemma 7 *Under Assumptions (A1) - (A7), we have* $N_m^{-1/2}\|\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}\|_2 = o_P(1)$.

Proof By Assumption (A5), note that $\widetilde{\mathbb{W}}_n$ is invertible and, for all n, $\lambda_{\min}(\widetilde{\mathbb{W}}_n) > \tilde{\lambda}_0 > 0$ for some constant $\tilde{\lambda}_0$. We will use an to the proof of the Theorem 1, and first show that, for any fixed $\varepsilon > 0$, $\mathbb{P}\left(\inf_{N_m^{-1/2}\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|_2\geq\varepsilon} N_m^{-1}|\mathit{\Gamma}(\boldsymbol{\theta}) - \mathit{\Gamma}(\tilde{\boldsymbol{\theta}})| > 0\right) \to 1$. To prove the above result, using the convexity of $\rho(\cdot)$, we only need to show that

$$\mathbb{P}\left(\inf_{N_m^{-1/2}\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|_2=\varepsilon} N_m^{-1}|\mathit{\Gamma}(\boldsymbol{\theta}) - \mathit{\Gamma}(\tilde{\boldsymbol{\theta}})| > 0;\ N_m^{-1/2}\|\tilde{\boldsymbol{\theta}}\|_2 < L\right) \to 1. \tag{10}$$

Using (3) and the argument similar to show the bound in equation (8), we have

$$\begin{aligned} N_m^{-1}\mathit{\Gamma}_n(\boldsymbol{\theta}) &= N_m^{-1}\left[\mathit{\Delta}_n(\boldsymbol{\theta}) + \mathbb{E}(\mathit{\Gamma}_n(\boldsymbol{\theta})) - \sum_{i=1}^n \frac{1}{N}\sum_{j=1}^N \psi(e_{ij})\tilde{\boldsymbol{B}}^{\mathrm{T}}\left(\frac{j}{N}\right)\theta\right] \\ &= N_m^{-1}\left[\mathit{\Delta}_n(\boldsymbol{\theta}) + \frac{\boldsymbol{\theta}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\boldsymbol{\theta}}{2} - \tilde{\boldsymbol{\theta}}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\boldsymbol{\theta}\right] + o(1). \end{aligned} \tag{11}$$

Notice that $2\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\boldsymbol{\theta} = \boldsymbol{\theta}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\boldsymbol{\theta} + \tilde{\boldsymbol{\theta}}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\tilde{\boldsymbol{\theta}} - (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^{\mathrm{T}}\widetilde{\mathbb{W}}_n(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})$. Substituting this into equation (11) we obtain

$$\frac{1}{N_m}\mathit{\Gamma}_n(\boldsymbol{\theta}) = \frac{1}{N_m}\left[\frac{(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^{\mathrm{T}}\widetilde{\mathbb{W}}_n(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})}{2} - \frac{\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\tilde{\boldsymbol{\theta}}}{2} + \mathit{\Delta}_n(\boldsymbol{\theta})\right] + o(1). \tag{12}$$

In particular we have

$$\frac{1}{N_m}\mathit{\Gamma}_n(\tilde{\boldsymbol{\theta}}) = \frac{1}{N_m}\left[-\frac{\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\widetilde{\mathbb{W}}_n\tilde{\boldsymbol{\theta}}}{2} + \mathit{\Delta}_n(\tilde{\boldsymbol{\theta}})\right] + o(1). \tag{13}$$

Using Lemmas 2 and Assumption (A3) we have $\|\tilde{\boldsymbol{\theta}}\|_2 = O(N_m^{1/2})$, which implies that for a large enough constant $L > 0$, we can assume $N_m^{-1/2}\|\tilde{\boldsymbol{\theta}}\|_2 < L$. Notice that if $N_m^{-1/2}\|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_2 = \varepsilon$ and $N_m^{-1/2}\|\tilde{\boldsymbol{\theta}}\|_2 < L$, then $N_m^{-1/2}\|\boldsymbol{\theta}\|_2 \le L + \varepsilon$. Subtracting equation (13) from (12) we get

$$\begin{aligned}
&\inf_{N_m^{-1/2}\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|_2=\varepsilon, N_m^{-1/2}\|\tilde{\boldsymbol{\theta}}\|_2<L} \frac{1}{N_m}\left|\Gamma(\boldsymbol{\theta}) - \Gamma(\tilde{\boldsymbol{\theta}})\right| \\
&= N_m^{-1}\left[\frac{(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})^{\mathrm{T}}\widetilde{\mathbb{W}}_n(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})}{2} + \Delta_n(\boldsymbol{\theta}) - \Delta_n(\tilde{\boldsymbol{\theta}})\right] + o(1) \\
&\ge \frac{\tilde{\lambda}_0\varepsilon^2}{2} - 2\sup_{N_m^{-1/2}\|\boldsymbol{\theta}\|_2\le L+\varepsilon} N_m^{-1}|\Delta_n(\boldsymbol{\theta})| + o(1) = \frac{\tilde{\lambda}\varepsilon^2}{2} + o(1),
\end{aligned}$$

where the last equality comes from Lemma 3. This proves equation (10) and implies that $N_m^{-1/2}\|\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}\|_2 = o_P(1)$. Lemma 7 is proved. □

Proof of Theorem 2

Proof By Lemmas 2 and 7, we have

$$\begin{aligned}
\left\|(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) - \boldsymbol{S}_n^{-1/2}\tilde{\boldsymbol{\theta}}\right\|_2 = \left\|\boldsymbol{S}_n^{-1/2}\left[\boldsymbol{S}_n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) - \tilde{\boldsymbol{\theta}}\right]\right\|_2 &= \left\|S_n^{-1/2}\right\|_2 \left\|\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}\right\|_2 \\
&= o_P\left(N_m n^{-1/2}\right),
\end{aligned}$$

which implies that, for any vector $\boldsymbol{\gamma} \in \mathbb{R}^{N_m+p}$, with $\|\boldsymbol{\gamma}\| \le L$, for a fixed constant $L > 0$,

$$\begin{aligned}
\boldsymbol{\gamma}^{\mathrm{T}}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) &= \boldsymbol{\gamma}^{\mathrm{T}}\mathbb{S}_n^{-1/2}\tilde{\boldsymbol{\theta}} + o_P\left(N_m n^{-1/2}\right) \\
&= \boldsymbol{\gamma}^{\mathrm{T}}\mathbb{W}_n^{-1}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{B}\left(\frac{j}{N}\right)^{\mathrm{T}}\psi(e_{ij}) + o_P\left(N_m n^{-1/2}\right). \qquad (14)
\end{aligned}$$

We can rewrite

$$\boldsymbol{\gamma}^{\mathrm{T}}\mathbb{W}_n^{-1}\sum_{i=1}^{n}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{B}\left(\frac{j}{N}\right)^{\mathrm{T}}\psi(e_{ij}) = \sum_{i=1}^{n}\frac{1}{N}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{\psi}(\boldsymbol{e}_i), \qquad (15)$$

where $\boldsymbol{v} = \left(\boldsymbol{\gamma}^{\mathrm{T}}\mathbb{W}_n^{-1}\boldsymbol{B}\left(\frac{1}{N}\right), \cdots, \boldsymbol{\gamma}^{\mathrm{T}}\mathbb{W}_n^{-1}\boldsymbol{B}\left(\frac{N}{N}\right)\right)^{\mathrm{T}}$. Notice also that

$$\mathrm{Var}\left(\sum_{i=1}^{n} N^{-1}\boldsymbol{v}^{\mathrm{T}}\psi(\boldsymbol{e_i})\right) = \sum_{i=1}^{n} N^{-2}\boldsymbol{v}^{\mathrm{T}}\mathbb{G}_i\boldsymbol{v}.$$

Using this calculation, we can rewrite equation (15) as $\sum_{i=1}^{n} a_i\xi_i$, where $a_i^2 = N^{-2}\boldsymbol{v}^{\mathrm{T}}\mathbb{G}_i\boldsymbol{v}$, $1 \leq i \leq n$, and $\{\xi_i\}_{i=1}^{n}$ are independent with mean zero and unit variance. By Lindeberg's Central Limit Theorem, if $\max_{i=1,\ldots,n} a_i^2/\sum_{i=1}^{n} a_i^2 = o(1)$, then $\frac{\sum_{i=1}^{n} a_i\xi_i}{\sqrt{\sum_{i=1}^{n} a_i^2}}$ converges in distribution to $N(0,1)$. According to Assumption (A4) and Lemma 2, we have

$$\begin{aligned}\max_{i=1,\ldots,n} a_i^2 &\leq \max_{i=1,\ldots,n} N^{-2}\|v\|_2^2 \sum_{j=1}^{N} \mathbb{E}\left[\psi(e_{ij})\right]^2 \\ &\leq CN^{-1}\sum_{j=1}^{N}\left(\boldsymbol{\gamma}^{\mathrm{T}}\mathbb{W}_n^{-1}\boldsymbol{B}\left(\frac{j}{N}\right)\right)^2 = O(N_m^2 n^{-2}).\end{aligned}$$

We also have $\sum_{i=1}^{n} a_i^2 = \sum_{i=1}^{n} N^{-2}\boldsymbol{v}^{\mathrm{T}}\mathbb{G}_i\boldsymbol{v} \geq N^{-1}\lambda_0\boldsymbol{\gamma}^{\mathrm{T}}\mathbb{W}_n^{-1}\mathbb{S}_n\mathbb{W}_n^{-1}\boldsymbol{\gamma} = O(N_m n^{-1}N^{-1})$, where the inequality comes from the Assumption (A7), and the last equation from Lemma 2. Collecting the previous bounds we have

$$\max_{i=1,\ldots,n} a_i^2/\sum_{i=1}^{n} a_i^2 = O(N_m N n^{-1}) = o(1)$$

due to Assumption (A1). This proves that the condition of Lindeberg's central limit theorem is satisfied. Setting $\boldsymbol{\gamma} = \boldsymbol{B}(x)$ we obtain

$$\sum_{i=1}^{n} N^{-2}\boldsymbol{v}^{\mathrm{T}}\mathbb{G}_i\boldsymbol{v} = \boldsymbol{B}(x)^{\mathrm{T}}\,\mathbb{W}_n^{-1}\left(\sum_{i=1}^{n} N^{-2}\mathbb{B}^{\mathrm{T}}\mathbb{G}_i\mathbb{B}\right)\mathbb{W}_n^{-1}\boldsymbol{B}(x) = D_n(x),$$

and due to the Assumption (A1) and (14) we finish the proof of Theorem 2. □

References

Cao, G., Yang, L., Todem, D. (2012). Simultaneous inference for the mean function based on dense functional data. *Journal of Nonparametric Statistics 24*(2), 359–377.

DeVore, R. A., Lorentz, G. G. (1993). *Constructive Approximation*, Volume 303. Springer Science & Business Media.

Schumaker, L. (2007). *Spline Functions: Basic Theory*. Cambridge University Press, Cambridge.