Chapter 2 : Hilbert Space Model for Functional Data

2.1

- Let H be a separable Hilbert space with inner product $\langle \cdot , \cdot \rangle$ which generates the norm $\| \cdot \|$
- Let $\mathcal{L}$ be the space of bounded linear operators on H where
$$\|\Psi\|_{\mathcal{L}} = \sup\{\|\Psi(x)\| : \|x\| \le 1\}$$
- An operator $\Psi \in \mathcal{L}$ is compact if $\exists \{v_j\}, \{f_j\}$ (orthonormal bases) & $\{\lambda_j\} \in \mathbb{R}$ converging to 0 such that
$$\Psi(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle f_j \quad x \in H$$
(singular value decomposition)
- A compact operator is a Hilbert-Schmidt operator if $\sum_{j=1}^{\infty} \lambda_j^2 < \infty$
- The space S of H-S operators is a separable Hilbert space w/ scalar product
$$\langle \Psi_1, \Psi_2 \rangle_S = \sum_{i=1}^{\infty} \langle \Psi_1(e_i), \Psi_2(e_i) \rangle$$
$\{e_i\}$ = arbitrary orthonormal basis
- A symmetric, positive definite H-S operator can be decomposed as
$$\Psi(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle v_j \quad x \in H$$
v_j are orthonormal eigenfunctions of Ψ $(\Psi(v_j) = \lambda_j v_j)$

2.2 The L^2 Space

- $L^2 = L^2([0,1])$ is the set of measurable real-valued functions x defined on $[0,1]$ such that $\int_0^1 x^2(t)\,dt < \infty$
- L^2 is a separable Hilbert space w/
$$\langle x, y \rangle = \int x(t)\, y(t)\, dt$$

2.3 Random elements in L^2 & the covariance operator

Let $X = \{X(t) : t \in [0,1]\}$ be a random element of L^2

If $E\|X\|^2 = E\int X^2(t)\,dt < \infty$ & $EX = 0$ then the covariance operator of X is defined

$$C(y) = E[\langle X, y\rangle X] \quad y \in L^2$$

$$C(y)(t) = \int c(t,s)\,y(s)\,ds \quad \text{where } c(t,s) = E[X(t)X(s)]$$

- C is symmetric & positive-definite
- $C \in \mathcal{L}(L^2)$ is a covariance operator $\iff$ it is symmetric positive definite & its eigenvalues satisfy $\sum_{j=1}^{\infty} \lambda_j < \infty$

Thm 2.1 Suppose $\{X_n, n \geq 1\}$ is a sequence of iid mean 0 random elements in a separable Hilbert space such that $E\|X_i\|^2 < \infty$. Then:

$$\frac{1}{\sqrt{N}} \sum_{n=1}^{N} X_n \xrightarrow{D} Z$$

Z is a Gaussian random element w/ covariance operator

$$C(x) = E[\langle Z, x\rangle Z] = E[\langle X_1, x\rangle X_1]$$

$$Z \overset{D}{=} \sum_{j=1}^{\infty} \sqrt{\lambda_j}\, N_j v_j \qquad N_j \overset{iid}{\sim} (0,1)$$

Thm 2.2 Suppose $\{X_n, n \geq 1\}$ is a sequence of iid random elements in a separable Hilbert space such that $E\|X_i\|^2 < \infty$. Then $\mu = EX_i$ is uniquely defined by $\langle \mu, x\rangle = E\langle X, x\rangle$ and

$$\frac{1}{N} \sum_{n=1}^{N} X_n \xrightarrow{a.s.} \mu$$

2.4 Estimation of mean & covariance functions

$\mu(t) = E[X(t)]$ $\qquad\qquad$ $\hat{\mu}(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t)$

$c(t,s) = E[\{X(t)-\mu(t)\}\{X(s)-\mu(s)\}]$ $\qquad$ $\hat{c}(t,s) = \frac{1}{N}\sum_{i=1}^{N}\{X_i(t)-\hat{\mu}(t)\}\{X_i(s)-\hat{\mu}(s)\}$

$C = E[\langle (X-\mu), \cdot \rangle (X-\mu)]$ $\qquad$ $\hat{C}(x) = \frac{1}{N}\sum_{i=1}^{N}\langle X_i-\hat{\mu}, x\rangle (X_i-\hat{\mu}) \quad x \in L^2$

* Note: $E[\hat{c}(t,s)] = \frac{N}{N-1} c(t,s)$

Assumption 2.1

$X_1, \dots, X_N$ are iid in L^2 & have the same distribution X, which is square integrable

Thm 2.3 If Assumption 2.1 holds Then (i) $E\hat{\mu} = \mu$ & (ii) $E\|\hat{\mu}-\mu\|^2 = O(\frac{1}{N})$

Pf: (i) $E[\hat{\mu}] = E[\frac{1}{N}\sum_{i=1}^{N} X_i] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] \overset{(*)}{=} \frac{1}{N}\sum_{i=1}^{N}\mu = \mu$

(*) Because $\forall i$ & for almost every $t \in [0,1]$ $EX_i(t) = \mu(t)$

(ii) $E[\|\hat{\mu}-\mu\|^2] = E[\langle \hat{\mu}-\mu, \hat{\mu}-\mu\rangle] = E[\langle \frac{1}{N}\sum_{i=1}^{N}(X_i-\mu), \frac{1}{N}\sum_{j=1}^{N}(X_j-\mu)\rangle]$

$= \frac{1}{N}\frac{1}{N} E[\langle \sum_{i=1}^{N}(X_i-\mu), \sum_{j=1}^{N}(X_j-\mu)\rangle]$

$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E[\langle X_i-\mu, X_j-\mu\rangle]$

By Lemma 2.1 $\langle (X_i-\mu), (X_j-\mu)\rangle = 0$ if $i \neq j$ if X_i, X_j are independent square integrable & $EX_1 = 0$

$$= \frac{1}{N^2}\sum_{i=1}^{N} E[\langle X_i-\mu, X_i-\mu\rangle] = \frac{1}{N^2}\underbrace{\sum_{i=1}^{N} E\|X_i-\mu\|^2}_{= N E\|X-\mu\|^2 \text{ because } X_1,\dots,X_N \overset{iid}{\sim} X} = \frac{1}{N} E\|X-\mu\|^2 = O(\frac{1}{N})$$

2.5 Estimation of eigenvalues & eigenfunctions

Define the estimated eigenelements by

$$\int \hat{c}(t,s)\,\hat{v}_j(s)\,ds = \hat{\lambda}_j\,\hat{v}_j(t) \qquad j=1,2,\dots,N \qquad \text{Let } \hat{c}_j = \operatorname{sign}(\langle \hat{v}_j, v_j\rangle)$$

Thm 2.7 Suppose $E\|X\|^4 < \infty$, $EX = 0$, $X_1,\dots,X_N \overset{iid}{\sim} X$ & $E\|X\|^2 < \infty$

and $\lambda_1 > \lambda_2 > \dots > \lambda_p > \lambda_{p+1}$

Then for each $1 \le j \le p$

$$\limsup_{N\to\infty} N\,E\left[\|\hat{c}_j\hat{v}_j - v_j\|^2\right] < \infty \qquad \limsup_{N\to\infty} N\,E\left[|\lambda_j - \hat{\lambda}_j|^2\right] < \infty$$

* So under the regularity conditions, the population eigenfunctions can be consistently estimated by the empirical eigenfunctions

Define the eigenvalues & eigenfunctions of C_θ as

$$\int C_\theta(t,s)\,v_{j,\theta}(s)\,ds = \lambda_{j,\theta}\,v_{j,\theta}(t)$$

Also let

$$\hat{c}_{j,\theta} = \operatorname{sign}(\langle \hat{v}_j, v_{j,\theta}\rangle)$$

Thm 2.8 Suppose assumption 2.2 holds, & $\lim_{N\to\infty} \frac{k_N^*}{N} = \theta \quad 0 \le \theta \le 1$ for a sequence of integers k_N^* such that $1 \le k_N^* \le N$

and

$\lambda_{1,\theta} > \lambda_{2,\theta} > \dots > \lambda_{p,\theta} > \lambda_{p+1,\theta}$

Then for each $1 \le j \le p$

$$E\left[\|\hat{c}_{j,\theta}\hat{v}_j - v_{j,\theta}\|^2\right] \to 0 \qquad \& \qquad E\left[|\hat{\lambda}_j - \lambda_{j,\theta}|^2\right] \to 0$$

* Change point procedures in later chapters will use this Theorem

2.6 Asymptotic normality of the eigen functions

Define

$$Z_N(t,s) = \sqrt{N}\{\hat{c}(t,s) - c(t,s)\}$$

Thm 2.9 If Assumption 2.1 holds with $E[X(t)]=0$ & $E\|X\|^4<\infty$

Then $Z_N(t,s)$ converges weakly in $L^2([0,1]\times[0,1])$ to a Gaussian process $T(t,s)$ where

$$E[T(t,s)] = 0$$

$$E[T(t,s)\,T(t',s')] = E[X(t)X(s)X(t')X(s')] - c(t,s)\,c(t',s')$$

Pf:

$$Z_N(t,s) = \sqrt{N}\{\hat{c}(t,s) - c(t,s)\} = \sqrt{N}\Big\{\frac{1}{N}\sum_{n=1}^{N} X_n(t)X_n(s) - c(t,s)\Big\}$$

$$= \frac{1}{\sqrt{N}}\sum_{n=1}^{N} X_n(t)X_n(s) - c(t,s)$$

Thus we can apply Thm 2.1 provided $E\iint\{X(t)X(s)\}^2 dt\,ds < \infty$

$$E\iint\{X(t)X(s)\}^2 dt\,ds = E\underbrace{\int X^2(t)dt}_{\|X\|^2}\underbrace{\int X^2(s)ds}_{\|X\|^2} = E\|X\|^4 < \infty$$

(by assumption)

Thus $Z_N(t,s) \xrightarrow{D} T(t,s)$ where $T(t,s)$ is a Gaussian process with

$$E[T(t,s)] = E[X_n(t)X_n(s) - c(t,s)] = c(t,s) - c(t,s) = 0$$

$$E[T(t,s)T(t',s')] = E[\{X_n(t)X_n(s) - c(t,s)\}\{X_n(t')X_n(s') - c(t',s')\}]$$

$$= E[X_n(t)X_n(s)X_n(t')X_n(s')] - c(t',s')E[X_n(t)X_n(s)]$$

$$- c(t,s)E[X_n(t')X_n(s')] + c(t,s)c(t',s')$$

$$= E[X_n(t)X_n(s)X_n(t')X_n(s')] - c(t',s')c(t,s) - c(t,s)c(t',s') + c(t,s)c(t',s')$$

$$\overset{iid}{=} E[X(t)X(s)X(t')X(s')] - c(t,s)\,c(t',s')$$