Identifying the finite dimensionality of curve time series

Neil Bathia, Qiwei Yao and Flavio Ziegelmann

Department of Statistics, London School of Economics

email: n.bathia@lse.ac.uk

May 15, 2009

Finite dimensionality of functional data

DEFINITION: $X_t(\cdot)$ is d-dimensional if $\mathrm{Cov}\{X_t(u), X_t(v)\}$ has exactly d non-zero eigenvalues, i.e.

$$X_t(u) - \mu(u) = \sum_{j=1}^{d} \xi_{tj}\varphi_j(u), \qquad t = 1, \ldots, n,$$

where $\varphi_j(\cdot)$ are eigenfunctions of $\mathrm{Cov}\{X_t(u), X_t(v)\}$.

$$\boxed{\textbf{Task: } \text{Estimate } d \text{ and } \mathcal{M} = \mathrm{span}\{\varphi_1(\cdot), \cdots, \varphi_d(\cdot)\}.}$$

Simple if we see $X_t(\cdot)$; eigenelements of $\widehat{\mathrm{Cov}}\{X_t(u), X_t(v)\}$ are consistent for eigenelements of $\mathrm{Cov}\{X_t(u), X_t(v)\}$.

In practice we see

$$Y_t(u) = X_t(u) + \varepsilon_t(u),$$

$$\Longrightarrow \mathrm{Cov}\{Y_t(u), Y_t(v)\} \neq \mathrm{Cov}\{X_t(u), X_t(v)\}.$$

Characterizing $\mathcal{M}$ via serial dependence

Recall $Y_t(\cdot) = X_t(\cdot) + \varepsilon_t(\cdot)$. If $X_t(\cdot)$ is a time series and $\varepsilon_t(\cdot)$ is *white noise* uncorrelated with $X_t(\cdot)$, then

$$\text{Cov}\{Y_t(u), Y_{t+k}(v)\} = \text{Cov}\{X_t(u), X_{t+k}(v)\}, \qquad k \geq 1.$$

> Under some identifiability conditions, $\text{Cov}\{Y_t(u), Y_{t+k}(v)\}$ has d non-zero eigenvalues and it's eigenfunctions span $\mathcal{M}$.

Can overcome difficulties faced by *traditional* FPCA in estimating d and $\mathcal{M}$ by using $\widehat{\text{Cov}}\{Y_t(u), Y_{t+k}(v)\}$ rather than $\widehat{\text{Cov}}\{Y_t(u), Y_t(v)\}$.

Outline

- Methodology
  - Using dependence for dimension reduction
  - Practical implementation - duality
  - Testing hypothesis about d - white noise test
- Asymptotic results
- Numerical results
  - Monte Carlo
  - IBM densities