Factor Models for High-Dimensional Functional Time Series

Marc Hallin
Université libre de Bruxelles

joint with

Gilles Nisol
Université libre de Bruxelles

Shahin Tavakoli
University of Geneva

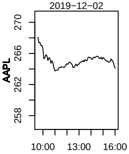
Gran Sasso National Laboratory November 30, 2023

Three challenges:

functional observations + high dimension + serial dependence

Why *functional* time series?

Intraday stock price (temperature, air pollution) (1 day)



This is a curve/function x

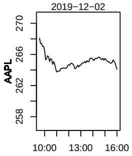
x : trading hours $\to \mathbb{R}$

$\tau \in$ trading hours $\mapsto x(\tau) \in \mathbb{R}$.

A univariate real-valued, continuous-time observed stochastic process (time series)

(stationarity, as a rule, does not hold)

Intraday stock price (1 day)



This is a curve/function x

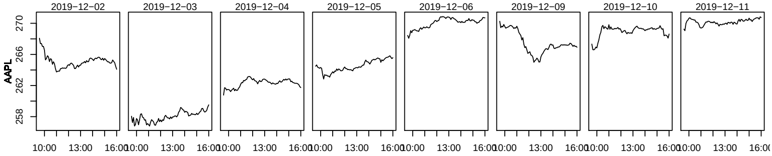
x : trading hours $\to \mathbb{R}$

$\tau \in$ trading hours $\mapsto x(\tau) \in \mathbb{R}$.

Traditionally: $x \in L^2([\tau_0, \tau_1], \mathbb{R})$—without loss of generality, $x \in L^2([0,1], \mathbb{R})$.

One univariate functional observation

Intraday stock price (several days)



Intraday price curve observed each trading day t,

Denote it by x_t, for each day $t = 1, 2, \ldots$ (say),

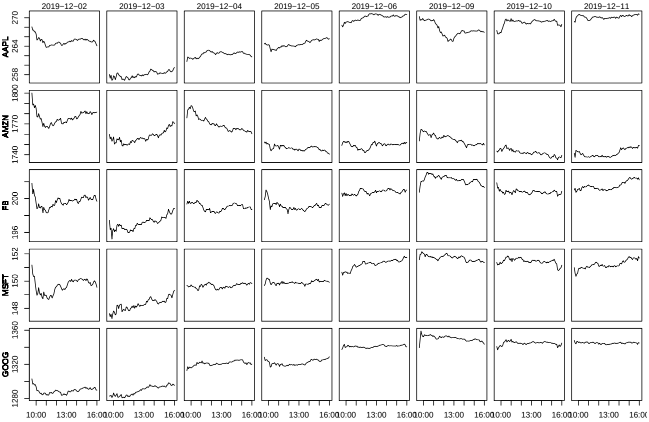
x_t : trading hours $\to \mathbb{R}$

$\tau \in$ trading hours $\mapsto x_t(\tau) \in \mathbb{R}$.

An observed univariate functional time series (FTS).

Depending on the problem, stationarity often holds

Intraday stock prices (several stocks; several days)



An observed low-dimensional ($N = 5$) multivariate functional time series (FTS); equivalently, an observed ($N = 5$) $\times$ ($T = 8$) panel of functional observations

Depending on the problem, stationarity may hold

Intraday stock prices ($N \to \infty =$"many stocks"; $T \to \infty =$"many days")

e.g., $N = 1000$ stocks observed over $T = 2000$ days

An observed high-dimensional functional time series (FTS)
equivalently,
An observed "large" panel of functional observations

Abstract Setting

$$\begin{array}{lcccccc}
\texttt{(AAPL intraday)} & x_{11} & x_{12} & \cdots & x_{1t} & \cdots & x_{1T} \\
\texttt{(AMZN intraday)} & x_{21} & x_{22} & \cdots & x_{2t} & \cdots & x_{2T} \\
\texttt{(FB intraday)} & x_{31} & x_{32} & \cdots & x_{3t} & \cdots & x_{3T} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\texttt{(GOOG intraday)} & x_{N1} & x_{N2} & \cdots & x_{Nt} & \cdots & x_{NT}
\end{array}$$

Each row is a time series of curves (order matters; (local) *stationarity* is a reasonable assumption)

Each column is a vector of curves (order is arbitrary/irrelevant; *exchangeability* is a reasonable assumption)

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

- Intra-day returns (FTS; τ is “intraday time”),

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

- Intra-day returns (FTS; τ is "intraday time"),
- Yield curves (FTS; here τ is "maturity," not "intraday time")

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

- Intra-day returns (FTS; τ is "intraday time"),
- Yield curves (FTS; here τ is "maturity," not "intraday time")
- Overnight returns (scalar time series),

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

- Intra-day returns (FTS; τ is "intraday time"),
- Yield curves (FTS; here τ is "maturity," not "intraday time")
- Overnight returns (scalar time series),
- Daily returns (scalar time series),
- macroeconomic indicators such as stock indices, exchange rates (vector time series),
- ...

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

- Intra-day returns (FTS; τ is "intraday time"),
- Yield curves (FTS; here τ is "maturity," not "intraday time")
- Overnight returns (scalar time series),
- Daily returns (scalar time series),
- macroeconomic indicators such as stock indices, exchange rates (vector time series),
- ...

Different τ's (no τ at all in case of a scalar series) but same t (e.g., daily observations—mixed frequencies are more delicate)

Mixed-Nature Panels

The rows (for each day t) could be of different nature, such as

- Intra-day returns (FTS; τ is "intraday time"),
- Yield curves (FTS; here τ is "maturity," not "intraday time")
- Overnight returns (scalar time series),
- Daily returns (scalar time series),
- macroeconomic indicators such as stock indices, exchange rates (vector time series),
- ...

Different τ's (no τ at all in case of a scalar series) but same t (e.g., daily observations—mixed frequencies are more delicate)

$\hookrightarrow$ importance of mixed-nature panels in applications

Mixed-Nature Panels: Abstract Setting I

`(AAPL intraday)`	x_{11}	x_{12}	$\cdots$	x_{1t}	$\cdots$	x_{1T}
`(Yield Curve)`	x_{21}	x_{22}	$\cdots$	x_{2t}	$\cdots$	x_{2T}
`(daily returns)`	x_{31}	x_{32}	$\cdots$	x_{3t}	$\cdots$	x_{3T}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
`(Overnight return)`	x_{i1}	x_{i2}	$\cdots$	x_{it}	$\cdots$	x_{iT}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	x_{N1}	x_{N2}	$\cdots$	x_{Nt}	$\cdots$	x_{NT}

Each row is a time series of curves, or a time series of numbers

Each column is a vector of curves & numbers

Each x_{it} takes values in a Hilbert Space H_i
(typically, $L^2([0,1],\mathbb{R})$ or $\mathbb{R}^{p_i}$ or $\mathbb{R}$)

Mixed-Nature Panels: Abstract Setting II

$$
\begin{array}{lcccccc}
\text{(elements of } H_1) & x_{11} & x_{12} & \cdots & x_{1t} & \cdots & x_{1T} \\
\text{(elements of } H_2) & x_{21} & x_{22} & \cdots & x_{2t} & \cdots & x_{2T} \\
\text{(elements of } H_3) & x_{31} & x_{32} & \cdots & x_{3t} & \cdots & x_{3T} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\text{(elements of } H_i) & x_{i1} & x_{i2} & \cdots & x_{it} & \cdots & x_{iT} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\text{(elements of } H_N) & x_{N1} & x_{N2} & \cdots & x_{Nt} & \cdots & x_{NT}
\end{array}
$$

Each row is a time series of curves, or a time series of numbers

Each column is a vector of curves & numbers

Mixed-Nature Panels: Abstract Setting III

Each x_{it} takes values in a real separable Hilbert space H_i (typically, $L^2([0,1],\mathbb{R})$ or $\mathbb{R}^{p_i}$) equipped with

- the inner product $\langle \cdot, \cdot \rangle_{H_i}$ and
- the norm $\|x_{it}\|_{H_i} := \langle x_{it}, x_{it} \rangle_{H_i}^{1/2}$, $i = 1, \ldots, N$.

The inner-product on $L^2([0,1],\mathbb{R})$ is (for $f, g \in L^2([0,1],\mathbb{R})$)

$$\langle f, g \rangle_{H_i} := \int_0^1 f(\tau) g(\tau) d\tau$$

Define

$$\boldsymbol{H}_N := H_1 \oplus H_2 \oplus \cdots \oplus H_N$$

(the direct sum of the Hilbert spaces $H_1, \ldots, H_N$): the elements of $\boldsymbol{H}_N$ are of the form $\boldsymbol{v} := (v_1, v_2, \ldots, v_N)^\mathsf{T}$ where $v_i \in H_i$, $i = 1, \ldots, N$. The space $\boldsymbol{H}_N$, naturally equipped with the inner product

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle_{\boldsymbol{H}_N} := \sum_{i=1}^{N} \langle v_i, w_i \rangle_{H_i},$$

is a real separable Hilbert space.

Mixed-Nature Panels: Analysis?

Some natural questions are:

- Joint model? typically impossible even for moderate N (curse of dimensionality)
- Underlying structure in the data? intricate cross-dependencies at all lags Better remain agnostic = nonparametric
- **Forecasting**? Arguably, the main problem in time series

Mixed-Nature Panels: Analysis?

Some natural questions are:

- Joint model? typically impossible even for moderate N (curse of dimensionality)
- Underlying structure in the data? intricate cross-dependencies at all lags Better remain agnostic = nonparametric
- **Forecasting**? Arguably, the main problem in time series

- In the scalar case ($H_i = \mathbb{R}$ for all i): (Dynamic) Factor Models—Marco Lippi's talk
- Extension needed to high-dimensional, mixed-nature, panels ...

Why **functional** series?

In practice, one never observes a function! Rather, the discretization of a function (e.g., intraday stock values recorded every minute). At the end of the day, thus, piling them up, ... a large- N panel of scalar or vector observations where traditional methods do apply!

In standard factor model methods, however, the cross-sectional ordering does not matter

- Here, after stacking the scalar values of discretized functional observations, cross-sectional ordering **does** matter: scalars originating from one given function are ordered, e.g., by intraday time τ
- Traditional methods, thus, **do not apply**—or then, fail to exploit the information related to the functional nature of observations—be they the discretized versions of unobservable functions.

Mixed-Nature Panels: Factor Models

$$\begin{array}{ccccccc}
(\text{elements of } H_1) & x_{11} & x_{12} & \cdots & x_{1t} & \cdots & x_{1T} \\
(\text{elements of } H_2) & x_{21} & x_{22} & \cdots & x_{2t} & \cdots & x_{2T} \\
(\text{elements of } H_3) & x_{31} & x_{32} & \cdots & x_{3t} & \cdots & x_{3T} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
(\text{elements of } H_i) & x_{i1} & x_{i2} & \cdots & x_{it} & \cdots & x_{iT} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
(\text{elements of } H_N) & x_{N1} & x_{N2} & \cdots & x_{Nt} & \cdots & x_{NT}
\end{array}$$

Factor model paradigm (scalar): for each t, decompose x_{it} into a sum

$$x_{it} = \chi_{it} + \xi_{it} =: \text{common}_{it} + \text{idiosyncratic}_{it}$$

where ...

The factor model paradigm (scalar case)

... decompose x_{it} into a sum

$$x_{it} = \chi_{it} + \xi_{it} =: \text{common}_{it} + \text{idiosyncratic}_{it}$$

where

- χ_{it}, the *common component*, takes values in the finite-dimensional space spanned by a finite (unspecified) number r of *factors*: $\chi_{it} = b_{i1}u_{1t} + \cdots + b_{ir}u_{rt}$—driven by a $q \leq r$-dimensional innovation (q unspecified)), formally N-dimensional but intrinsically r-dimensional time series with rank q
- ξ_{it}, the *idiosyncratic component*, is only *"mildly" cross-correlated*
- χ_{it} and ξ_{it} are mutually orthogonal

The factor model paradigm (scalar case)

... decompose x_{it} into a sum

$$x_{it} = \chi_{it} + \xi_{it} =: \text{common}_{it} + \text{idiosyncratic}_{it}$$

where

- χ_{it}, the *common component*, takes values in the finite-dimensional space spanned by a finite (unspecified) number r of *factors*: $\chi_{it} = b_{i1}u_{1t} + \cdots + b_{ir}u_{rt}$—driven by a $q \leq r$-dimensional innovation (q unspecified)), formally N-dimensional but intrinsically r-dimensional time series with rank q
- ξ_{it}, the *idiosyncratic component*, is only *"mildly" cross-correlated*
- χ_{it} and ξ_{it} are mutually orthogonal

Neither χ_{it} nor ξ_{it} (nor the factors u_{jt} nor the loadings b_{ij} ...) are observed; r (and q) are unspecified: to be recovered from the data.

The factor model paradigm (scalar case)

The various versions of factor models then differ by their definitions of *"mildly" cross-correlated* and the assumption of a finite r (finite-dimensional factor space).

The most general definition is the "General Dynamic Factor Model" one proposed in Forni-Hallin-Lippi-Reichlin (*Rev. Econ. & Statist.* 2000), Forni and Lippi (*Econometric Theory* 2001) and Forni-Hallin-Lippi-Zaffaroni (*JoE* 2015, 2017), where $q < \infty$ but $r < \infty$ is not required; there,

- ξ_{it} *idiosyncratic (mildly cross-correlated)* means: *the largest eigenvalues of* ξ_{it}*'s* $N \times N$ *spectral density matrices are bounded (all frequencies) as* $N \to \infty$
- χ_{it} *common (pervasively cross-correlated)* means: *the* q*th eigenvalues of* χ_{it}*'s* $N \times N$ *spectral density matrices are unbounded (all frequencies) as* $N \to \infty$ *but the* $(q+1)$*th ones are bounded (all frequencies)*

The factor model paradigm (scalar case)

The most popular definition is the one adopted in Bai and Ng (*Econometrica* 2002) and Stock and Watson (*JASA* 2002), where $r < \infty$ is required and

- ξ_{it} *idiosyncratic (mildly cross-correlated)* means: *the largest eigenvalues of ξ_{it}'s $N \times N$ lag-zero covariance matrices are bounded as $N \to \infty$*
- χ_{it} *common (pervasively cross-correlated)* means: *the rth eigenvalue of χ_{it}'s $N \times N$ lag-zero covariance matrices are unbounded as $N \to \infty$ but the $(r+1)$th one is bounded*

What the factor model is / is not

The factor model (in this high-dimensional time-series context)

- is not a data-generating process
- is not a dimension-reduction method
- is not a signal + noise model
- is not an approximate reduced rank model

Rather, the factor model

- is (under very general conditions, mostly under its GDFM form) the expression of a representation result—a mathematical fact rather than a "statistical model"
- is an operational decomposition aimed at a "divide and rule" strategy ...
- ... where χ_{it} and ξ_{it} are to be recovered, then handled (e.g., predicted) via drastically different methods ...
- ... then **put back together again**, e.g., to produce forecasts

Actually, the general dynamic factor model is not a "statistical model": beyond second-order stationarity and the existence of a spectrum, it does not place any restriction on the data-generating process—only requiring the number of exploding dynamic eigenvalues to be finite ... (which, in view of the fact that N in practice is fixed, is quite reasonable)

... an approach based on representation results that originates in Forni and Lippi, *Econometric Theory* (2001).

The factor model paradigm (Functional case)

A natural (and simple) functional extension of the scalar decomposition is

$$x_{it} = \underbrace{b_{i1}u_{1t} + \cdots + b_{ir}u_{rt}}_{\text{common component}} + \underbrace{\xi_{it}}_{\text{idiosyncratic component}}, \quad \forall i \in \mathbb{N}, \forall t \in \mathbb{N}.$$

where

- Factors $u_{1t}, \ldots, u_{rt} \in \mathbb{R}$ (**unobserved**, scalar),
- Factor loadings $b_{i1}, \ldots, b_{ir} \in H_i$ (**unobserved**, functional),
- Idiosyncratic component $\xi_{it} \in H_i$ (**unobserved**, functional).

The factor model paradigm (Functional case)

A natural (and simple) functional extension of the scalar decomposition is

$$x_{it} = \underbrace{b_{i1}u_{1t} + \cdots + b_{ir}u_{rt}}_{\text{common component}} + \underbrace{\xi_{it}}_{\text{idiosyncratic component}}, \quad \forall i \in \mathbb{N}, \forall t \in \mathbb{N}.$$

where

Factors $u_{1t}, \ldots, u_{rt} \in \mathbb{R}$ (**unobserved**, scalar),

Factor loadings $b_{i1}, \ldots, b_{ir} \in H_i$ (**unobserved**, functional),

Idiosyncratic component $\xi_{it} \in H_i$ (**unobserved**, functional).

<u>Scalar factor models are a</u> (very) <u>special case:</u> $\mathbb{R} = H_1 = H_2 = \cdots$.

The covariance operator

Let

$$\boldsymbol{X}_t^N := (X_{1t}, X_{2t}, \ldots, X_{Nt})^\mathsf{T}$$

denote an $\boldsymbol{H}_N$-valued random variable (to keep the presentation simple, the dependence on N does not explicitly appear below)

The covariance operator

$$C_N^X := \mathbb{E}\left[(\boldsymbol{X}_t - \mathbb{E}\,\boldsymbol{X}_t)\otimes(\boldsymbol{X}_t - \mathbb{E}\,\boldsymbol{X}_t)\right] \in \mathcal{L}(\boldsymbol{H}_N)$$

of $\boldsymbol{X}_t$ is mapping $\boldsymbol{y} \in \boldsymbol{H}_N$ to

$$C_N^X \boldsymbol{y} := \mathbb{E}\left[\langle(\boldsymbol{X}_t - \mathbb{E}\,\boldsymbol{X}_t), \boldsymbol{y}\rangle\,(\boldsymbol{X}_t - \mathbb{E}\,\boldsymbol{X}_t)\right] \in \boldsymbol{H}_N$$

Recall: for $u \in H_1$, $v \in H_2$, $u \otimes v$ is the operator (from H_2 to H_1)

$$f \in H_2 \mapsto (u \otimes v)(f) := \langle f, v\rangle\, u \in H_1.$$

For vectors $u \in \mathbb{R}^p, v \in \mathbb{R}^q$, $u \otimes v = uv^\mathsf{T}$ (a $p \times q$ matrix).
For $u = 1 \in \mathbb{R}$, $v \in H_2$, $u \otimes v$ is the operator (from H_2 to $\mathbb{R}$)

$$f \in H_2 \mapsto (u \otimes v)(f) := \langle f, v\rangle \in \mathbb{R}.$$

Eigendecomposition of the Covariance

Denote by $\lambda^X_{N,1}, \lambda^X_{N,2}, \ldots$ the eigenvalues, in decreasing order of magnitude, of this covariance operator.

Similarly denote by $\lambda^\chi_{N,1}, \lambda^\chi_{N,2}, \ldots$ and $\lambda^\xi_{N,1}, \lambda^\xi_{N,2}, \ldots$ the eigenvalues of the covariance operators C^χ_N and C^ξ_N

High-Dimensional Functional Factor Model

Definition. We say that a (second-order stationary in $t \in \mathbb{Z}$) functional zero-mean process

$$\mathcal{X} := \{x_{it} : i \in \mathbb{N}, t \in \mathbb{Z}\}$$

admits a *high-dimensional functional factor model representation* with r factors if

$$x_{it} = \underbrace{b_{i1}u_{1t} + \cdots + b_{ir}u_{rt}}_{:=\chi_{it}} + \xi_{it} = \chi_{it} + \xi_{it}, \quad i \in \mathbb{N}, t \in \mathbb{Z},$$

where

- $b_{ij} \in H_i$, (functional loadings; no dependence in t)
- $\boldsymbol{u}_t := (u_{1t}, \ldots, u_{rt})^\mathsf{T}$, with values in $\mathbb{R}^r$, is zero-mean second-order stationary, co-stationary with $\mathcal{X}$, and $\mathbb{E}\left[\boldsymbol{u}_t \boldsymbol{u}_t^\mathsf{T}\right]$ is positive definite (scalar factors),
- $\{\xi_{it}\}$, with values in H_i, is zero-mean second-order stationary, and $\mathbb{E}\left[u_{jt}\xi_{it}\right] = 0$ for all $j = 1, \ldots, r$ and $i \in \mathbb{N}$,
- $\sup_{N \geq 1} \lambda^{\chi}_{N,r} = \infty$, $\sup_{N \geq 1} \lambda^{\chi}_{N,r+1} < \infty$
- $\sup_{N \geq 1} \lambda^{\xi}_{N,1} < \infty$.

scalar factors, functional loadings ... allow the impact of a common shock to depend on τ_i in an item-specific way (recall that τ_{i_1} and τ_{i_2} may be of an entirely different nature) ...

... would not be possible with functional factors and scalar loadings (which, moreover, require $H_1 = H_2 = \ldots$, thus precluding the analysis of mixed-nature panels).

High-Dimensional Functional Factor Models

In "matrix" notation,

$$\boldsymbol{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \boldsymbol{B}_N \boldsymbol{u}_t + \boldsymbol{\xi}_t,$$

where

- $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})^{\mathsf{T}}$ is $\boldsymbol{H}_N$-valued,

High-Dimensional Functional Factor Models

In "matrix" notation,

$$\boldsymbol{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \boldsymbol{B}_N \boldsymbol{u}_t + \boldsymbol{\xi}_t,$$

where

- $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})^\mathsf{T}$ is $\boldsymbol{H}_N$-valued,
- $\boldsymbol{B}_N$ is $N \times r$ with $(\boldsymbol{B}_N)_{il} = b_{il} \in H_i$.

High-Dimensional Functional Factor Models

In "matrix" notation,

$$\boldsymbol{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \boldsymbol{B}_N \boldsymbol{u}_t + \boldsymbol{\xi}_t,$$

where

- $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})^{\mathsf{T}}$ is $\boldsymbol{H}_N$-valued,
- $\boldsymbol{B}_N$ is $N \times r$ with $(\boldsymbol{B}_N)_{il} = b_{il} \in H_i$.
- $\boldsymbol{u}_t$ is $\mathbb{R}^r$-valued, $\mathbb{E}\, \boldsymbol{u}_t = 0$,
  and $\mathbb{E}\left[\boldsymbol{u}_t \boldsymbol{u}_t^{\mathsf{T}}\right] = \boldsymbol{\Sigma}_{\boldsymbol{u}} \in \mathbb{R}^{r \times r}$ is positive definite,

High-Dimensional Functional Factor Models

In "matrix" notation,

$$\boldsymbol{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \boldsymbol{B}_N \boldsymbol{u}_t + \boldsymbol{\xi}_t,$$

where

- $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})^\mathsf{T}$ is $\boldsymbol{H}_N$-valued,
- $\boldsymbol{B}_N$ is $N \times r$ with $(\boldsymbol{B}_N)_{il} = b_{il} \in H_i$.
- $\boldsymbol{u}_t$ is $\mathbb{R}^r$-valued, $\mathbb{E}\, \boldsymbol{u}_t = 0$,
  and $\mathbb{E}\left[\boldsymbol{u}_t \boldsymbol{u}_t^\mathsf{T}\right] = \boldsymbol{\Sigma}_{\boldsymbol{u}} \in \mathbb{R}^{r \times r}$ is positive definite,
- $\boldsymbol{\xi}_t$ is $\boldsymbol{H}_N$-valued and $\mathbb{E}\, \boldsymbol{\xi}_t = 0$,
- $\mathbb{E}\, \boldsymbol{u}_t \otimes \boldsymbol{\xi}_t = 0$,

High-Dimensional Functional Factor Models

In "matrix" notation,

$$\boldsymbol{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \boldsymbol{B}_N \boldsymbol{u}_t + \boldsymbol{\xi}_t,$$

where

- $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})^{\mathsf{T}}$ is $\boldsymbol{H}_N$-valued,
- $\boldsymbol{B}_N$ is $N \times r$ with $(\boldsymbol{B}_N)_{il} = b_{il} \in H_i$.
- $\boldsymbol{u}_t$ is $\mathbb{R}^r$-valued, $\mathbb{E}\, \boldsymbol{u}_t = 0$,
  and $\mathbb{E}\left[\boldsymbol{u}_t \boldsymbol{u}_t^{\mathsf{T}}\right] = \boldsymbol{\Sigma}_{\boldsymbol{u}} \in \mathbb{R}^{r \times r}$ is positive definite,
- $\boldsymbol{\xi}_t$ is $\boldsymbol{H}_N$-valued and $\mathbb{E}\, \boldsymbol{\xi}_t = 0$,
- $\mathbb{E}\, \boldsymbol{u}_t \otimes \boldsymbol{\xi}_t = 0$,
- $\sup_{N \geq 1} \lambda^{\chi}_{N,r} = \infty$, $\sup_{N \geq 1} \lambda^{\chi}_{N,r+1} < \infty$
- $\sup_{N \geq 1} \lambda^{\xi}_{N,1} < \infty$.

Recall: for $u \in H_1$, $v \in H_2$, $u \otimes v : H_2 \to H_1$ is defined by

$$(u \otimes v)(f) = \langle f, v \rangle\, u.$$

For vectors $u \in \mathbb{R}^p, v \in \mathbb{R}^q$, $u \otimes v = uv^{\mathsf{T}}$.

Representation results (I)

Let (note that $\lambda_{N,j}^X$ is monotone increasing in N)

$$\lambda_j^X := \lim_{N\to\infty} \lambda_{N,j}^X, \quad j = 1, 2, \ldots$$

Representation results (I)

Let (note that $\lambda^X_{N,j}$ is monotone increasing in N)

$$\lambda^X_j := \lim_{N\to\infty} \lambda^X_{N,j}, \quad j = 1, 2, \ldots$$

Theorem (Existence: Tavakoli, Nisol and Hallin, 2020)

The process $\mathcal{X}$ admits a (high-dimensional) functional factor model representation with r factors if and only if

- $\lambda^X_r = \infty$, *and*
- $\lambda^X_{r+1} < \infty$.

Representation results (I)

Let (note that $\lambda^X_{N,j}$ is monotone increasing in N)

$$\lambda^X_j := \lim_{N\to\infty} \lambda^X_{N,j}, \quad j = 1, 2, \ldots$$

Theorem (Existence: Tavakoli, Nisol and Hallin, 2020)

The process $\mathcal{X}$ admits a (high-dimensional) functional factor model representation with r factors if and only if

- $\lambda^X_r = \infty$, *and*
- $\lambda^X_{r+1} < \infty$.

Except for the existence of a bounded eigenvalue (assumption: the number of exploding eigenvalues is finite), no specific factor model assumption! Moreover, recall that in practice N is fixed!

As in the scalar case [Chamberlain & Rothschild (1982); Forni & Lippi (2001); Hallin & Lippi (2013)] but we remove (our proof does not need it) the assumption $\text{var}(x_{it}) \geq \delta, \forall i$.

Representation results (II)

Theorem (Uniqueness: Tavakoli, Nisol and Hallin, 2020)

Let $x_{it} = \chi_{it} + \xi_{it}$, $i \in \mathbb{N}, t \in \mathbb{Z}$, *(functional factor model with* r *factors). Then,*

$$\chi_{it} = \mathrm{proj}_{H_i}(x_{it}|\mathcal{D}_t), \quad \forall i \in \mathbb{N}, t \in \mathbb{Z}$$

where

$$\mathcal{D}_t := \left\{ p \in L^2(\Omega) \,|\, p = \lim_{N\to\infty} \langle \boldsymbol{\alpha}_N, \boldsymbol{x}_t \rangle_{\boldsymbol{H}_N}, \boldsymbol{\alpha}_N \in \boldsymbol{H}_N, \|\boldsymbol{\alpha}_N\|_{\boldsymbol{H}_N} \stackrel{N\to\infty}{\longrightarrow} 0 \right\}$$

The common and idiosyncratic components, thus, are unique, and asymptotically identified.

Representation results (II)

Theorem (Uniqueness: Tavakoli, Nisol and Hallin, 2020)

Let $x_{it} = \chi_{it} + \xi_{it}$, $i \in \mathbb{N}, t \in \mathbb{Z}$, *(functional factor model with* r *factors). Then,*

$$\chi_{it} = \mathrm{proj}_{H_i}(x_{it} | \mathcal{D}_t), \quad \forall i \in \mathbb{N}, t \in \mathbb{Z}$$

where

$$\mathcal{D}_t := \left\{ p \in L^2(\Omega) \,|\, p = \lim_{N\to\infty} \langle \boldsymbol{\alpha}_N, \boldsymbol{x}_t \rangle_{\boldsymbol{H}_N}, \boldsymbol{\alpha}_N \in \boldsymbol{H}_N, \|\boldsymbol{\alpha}_N\|_{\boldsymbol{H}_N} \overset{N\to\infty}{\longrightarrow} 0 \right\}$$

The common and idiosyncratic components, thus, are unique, and asymptotically identified.

. . . As in the scalar case, however, for any invertible $\boldsymbol{Q}$,

$$\boldsymbol{B}_N \boldsymbol{u}_t = (\boldsymbol{B}_N \boldsymbol{Q})(\boldsymbol{Q}^{-1} \boldsymbol{u}_t),$$

hence loadings and factors are *jointly* but not *separately* identifiable.

Other Functional Factor Models?

Multivariate case (N fixed—genuine models, thus):

- Castellanos et al. (2015), White & Gelfand (2020): Functional factors, scalar loadings
- Kowal, Matteson & Ruppert (2017): scalar factors, functional loadings

Other Functional Factor Models?

Multivariate case (N fixed—genuine models, thus):

- Castellanos et al. (2015), White & Gelfand (2020): Functional factors, scalar loadings
- Kowal, Matteson & Ruppert (2017): scalar factors, functional loadings

High-dimensional case ($N \to \infty$):

- Gao, Shang & Yang (2019): univariate FPCA (with a dangerous preliminary dimension-reduction step which potentially may destroy all common components!) followed by separate factor models on scores.
- Tang, Shang & Yang (2021); Qiao, Guo, & Wang (2021): flexible loading schemes with $H_i = H_1 \ \forall i \geq 1$.

However,

- none of these alternative approaches is based on a representation result; the factor model structure they are based on, thus, may not be there!

- functional factors and scalar loadings are NOT a plus:

  – require $H_i = H_1$ for all i, which is extremely restrictive ...

  – preclude the possibility of τ-specific loadings

Estimation of factors and loadings

Given observations $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T \in \boldsymbol{H}_N$, under the assumptions of Theorem I with unspecified number r of factors, we need to estimate the factor loadings and the factors.

Therefore, we more generally consider, *for arbitrary k*, the solutions $\boldsymbol{B}_N^{(k)}$ and $\boldsymbol{U}_T^{(k)} = (\boldsymbol{u}_1^{(k)}, \dots, \boldsymbol{u}_T^{(k)})$ of the minimization problem

$$\min_{\boldsymbol{B}_N^{(k)} \in \mathcal{L}(\mathbb{R}^k, \boldsymbol{H}_N),\ \boldsymbol{U}_T^{(k)} \in \mathbb{R}^{k \times T}} P(\boldsymbol{B}_N^{(k)}, \boldsymbol{U}_T^{(k)}) := \sum_t \left\| \boldsymbol{x}_t - \boldsymbol{B}_N^{(k)} \boldsymbol{u}_t^{(k)} \right\|^2$$

(for $k = r$, the least-squares estimators) .

Estimation of factors and loadings

Given observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T \in \boldsymbol{H}_N$, under the assumptions of Theorem I with unspecified number r of factors, we need to estimate the factor loadings and the factors.

Therefore, we more generally consider, *for arbitrary k*, the solutions $\boldsymbol{B}_N^{(k)}$ and $\boldsymbol{U}_T^{(k)} = (\boldsymbol{u}_1^{(k)}, \ldots, \boldsymbol{u}_T^{(k)})$ of the minimization problem

$$\min_{\boldsymbol{B}_N^{(k)} \in \mathcal{L}(\mathbb{R}^k, \boldsymbol{H}_N),\ \boldsymbol{U}_T^{(k)} \in \mathbb{R}^{k \times T}} P(\boldsymbol{B}_N^{(k)}, \boldsymbol{U}_T^{(k)}) := \sum_t \left\| \boldsymbol{x}_t - \boldsymbol{B}_N^{(k)} \boldsymbol{u}_t^{(k)} \right\|^2$$

(for $k = r$, the least-squares estimators) .

Now, $\boldsymbol{X}_{NT} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ induces an operator

$$\mathsf{L}(\boldsymbol{X}_{NT}) : \mathbb{R}^T \to \boldsymbol{H}_N$$

while $\boldsymbol{U}_T^{(k)}$ is a $r \times T$ real matrix, hence can be viewed as a mapping $\boldsymbol{U}_T^{(k)} : \mathbb{R}^T \to \mathbb{R}^r$.

Estimation of factors and loadings

The minimization problem therefore can be rewritten as the minimization of

$$P(\boldsymbol{B}_N^{(k)}, \boldsymbol{U}_T^{(k)}) = \left|\left|\left| \mathsf{L}(\boldsymbol{X}_{NT}) - \boldsymbol{B}_N^{(k)} \boldsymbol{U}_T^{(k)} \right|\right|\right|_2^2,$$

where $|||\cdot|||_2$ is the Hilbert–Schmidt norm.

Estimation of factors, loadings

Since $\boldsymbol{B}_N^{(k)}\boldsymbol{U}_T^{(k)}$ is of rank k, by the **Eckart–Young–Mirsky** theorem, the minimum is achieved for

$$\boldsymbol{B}_N\boldsymbol{U}_T = \widetilde{\boldsymbol{B}}_N^{(k)}\widetilde{\boldsymbol{U}}_T^{(k)},$$

the rank k truncation of the singular value decomposition (SVD) of $\mathsf{L}(\boldsymbol{X}_{NT})$.

Details are skipped

Estimation of factors, loadings

Singular value decomposition

$$\mathsf{L}(\boldsymbol{X}_{NT}) = \sum_{l=1}^{N} \hat{\lambda}_l^{1/2} \hat{\boldsymbol{e}}_l \otimes \hat{\boldsymbol{f}}_l. \tag{1}$$

We could compute it either via

(A) Spectral decomposition of $\mathsf{L}(\boldsymbol{X}_{NT})\mathsf{L}(\boldsymbol{X}_{NT})^\star \in \mathcal{L}(\boldsymbol{H}_N, \boldsymbol{H}_N)$, or

(B) Spectral decomposition of $\mathsf{L}(\boldsymbol{X}_{NT})^\star\mathsf{L}(\boldsymbol{X}_{NT}) \in \mathcal{L}(\mathbb{R}^T, \mathbb{R}^T)$.

(B) is advantageous because no need for having basis functions of $\boldsymbol{H}_N$ and computing their inner products:

1. Compute $(\mathbf{F})_{st} = \langle \boldsymbol{x}_s, \boldsymbol{x}_t \rangle = \sum_{i=1}^{N} \langle x_{is}, x_{it} \rangle_{H_i}$ for $s, t = 1, \ldots, T$
2. compute the leading k eigenvalue/eigenvector pairs $(\tilde{\lambda}_l, \tilde{\boldsymbol{f}}_l)$ of $\mathbf{F}$, and set
$$\hat{\lambda}_l := T^{-1/2}\tilde{\lambda}_l \in \mathbb{R}, \quad \hat{\boldsymbol{f}}_l := T^{1/2}\tilde{\boldsymbol{f}}_l / \left| \tilde{\boldsymbol{f}}_l \right| \in \mathbb{R}^T;$$
3. compute $\hat{\boldsymbol{e}}_l := \hat{\lambda}_l^{-1/2} T^{-1} \sum_{t=1}^{T} (\hat{\boldsymbol{f}}_l)_t \boldsymbol{x}_t \in \boldsymbol{H}_N$;
4. set $\widetilde{\boldsymbol{U}}_T^{(k)} := (\hat{\boldsymbol{f}}_1, \ldots, \hat{\boldsymbol{f}}_k)^\mathsf{T} \in \mathbb{R}^{k \times T}$ and define $\widetilde{\boldsymbol{B}}_N^{(k)}$ as the operator in $\mathcal{L}(\mathbb{R}^k, \boldsymbol{H}_N)$ mapping the l-th canonical basis vector of $\mathbb{R}^k$ to $\hat{\lambda}_l^{1/2} \hat{\boldsymbol{e}}_l$, $l = 1, \ldots, k$.

Estimation of factors, loadings

Our method is of the FPCA type, but

- distinct from other multivariate FPCAs [Ramsay & Silverman (2005), Berrendero, Justel & Svarc (2011), Chiou, Chen and Yang (2014), Jacques and Preda (2014)]
- contrary to other FPCA methods, works for distinct H_i's,
- close to Happ & Greven (2018); however, no preliminary Karhunen–Loève dimension reduction for individual x_{it}'s prior to conducting the global PCA—*not* a good idea in our setting, as there is no guarantee that the common component will survive the individual Karhunen–Loève projections (which, actually, might well remove all common components)

Consistency results: average error bounds

(assuming $k = r$)

Let $C_{N,T} := \min\{\sqrt{N}, \sqrt{T}\}$. Assumptions A, B, ... are functional versions of classical assumptions on scalar factor models (Bai and Ng 2002, etc.)

Theorem (Tavakoli, Nisol and Hallin, 2020)

Under assumptions A, B, C, D,

$$\min_{\mathbf{R}\in\mathbb{R}^{r\times r}} \left|\left|\left| \widetilde{\boldsymbol{U}}_T^{(r)} - \mathbf{R}\boldsymbol{U}_T \right|\right|\right|_2 / \sqrt{T}, = O_{\mathrm{P}}(C_{N,T}^{-1}).$$

Theorem (Tavakoli, Nisol and Hallin, 2020)

Under Assumptions A, B, C, D, and E(α),

$$\min_{\mathbf{R}\in\mathbb{R}^{r\times r}} \left|\left|\left| \widetilde{\boldsymbol{B}}_N^{(r)} \hat{\boldsymbol{\Lambda}}^{-1/2} - \boldsymbol{B}_N \mathbf{R} \right|\right|\right|_2 / \sqrt{N} = O_{\mathrm{P}}\left(C_{N,T}^{-\frac{1+\alpha}{2}} \right).$$

Theorem (Tavakoli, Nisol and Hallin, 2020)

Under Assumptions A, B, C, D, and E(α), $\alpha \in [0, 1]$,

$$\frac{1}{\sqrt{NT}} \sqrt{\sum_{i=1}^{N} \sum_{t=1}^{T} \|\chi_{it} - \hat{\chi}_{it}\|^2} = O_{\mathrm{P}}\left(C_{N,T}^{-\frac{1+\alpha}{2}} \right).$$

Consistency results: uniform error bounds

(assuming $k = r$)

Let $\widetilde{\mathbf{R}} = \hat{\mathbf{\Lambda}}^{-1}\widetilde{\boldsymbol{U}}_T^{(r)}\boldsymbol{U}_T^{\star}\boldsymbol{B}_N^{\star}\boldsymbol{B}_N/(NT)$. Assumptions A, B, ... are functional versions of classical assumptions on scalar factor models (Bai and Ng 2002, etc.)

Theorem (Tavakoli, Nisol and Hallin, 2020)

Under Assumptions A, B, C, D and G(κ),

$$\max_{t=1,\dots,T}\left|\widetilde{\boldsymbol{u}}_t - \widetilde{\mathbf{R}}\boldsymbol{u}_t\right| = O_{\mathrm{P}}\left(\max\left\{\frac{1}{\sqrt{T}}, \frac{T^{1/(2\kappa)}}{\sqrt{N}}\right\}\right).$$

Theorem (Tavakoli, Nisol and Hallin, 2020)

Let Assumptions A, B, C, D, H(γ) hold. Then,

$$\max_{i=1,\dots,N}\left|\left|\left|\widetilde{\mathbf{b}}_i^{(r)} - \mathbf{b}_i\widetilde{\mathbf{R}}^{-1}\right|\right|\right|_2 = O_{\mathrm{P}}\left(\max\left\{\frac{1}{\sqrt{N}}, \frac{\log(N)\log(T)^{1/2\gamma}}{\sqrt{T}}\right\}\right).$$

Theorem (Tavakoli, Nisol and Hallin, 2020)

Under Assumptions A, B, C, D, G(κ), H(γ),

$$\max_{t=1,\dots T}\max_{i=1,\dots,N}\left\|\hat{\chi}_{it}^{(r)} - \chi_{it}\right\|_{H_i} = O_{\mathrm{P}}\Big(\max\Big\{\frac{T^{1/(2\kappa)}}{\sqrt{N}}, \frac{\log(N)\log(T)^{1/2\gamma}}{\sqrt{N}\,T^{(\kappa-1)/(2\kappa)}}, \frac{\log(N)\log(T)^{1/2\gamma}}{\sqrt{T}}\Big\}\Big).$$

Consistency if $N, T \to \infty$ such that $T = o(N^{\kappa})$ and $\log(N) = o(\sqrt{T}/\log(T)^{1/2\gamma})$.

Estimating the Number of Factors: Consistency

Estimate the number r of factor by (similar to Bai and Ng 2002)

$$\hat{r} := \arg \min_{k=1,\ldots,k_{\max}} V(k, \widetilde{\boldsymbol{U}}_T^{(k)}) + k\, g(N,T),$$

where $g(N,T)$ is a penalty function and

$$V(k, \widetilde{\boldsymbol{U}}_T^{(k)}) := \min_{\boldsymbol{B}_N^{(k)} \in \mathcal{L}(\mathbb{R}^k, \boldsymbol{H}_N)} \frac{1}{NT} \sum_{t=1}^{T} \left\| \boldsymbol{x}_t - \boldsymbol{B}_N^{(k)} \widetilde{\boldsymbol{u}}_t^{(k)} \right\|^2 \qquad (2)$$

Estimating the Number of Factors: Consistency

Estimate the number r of factor by (similar to Bai and Ng 2002)

$$\hat{r} := \arg \min_{k=1,\dots,k_{\max}} V(k, \widetilde{\boldsymbol{U}}_T^{(k)}) + k\, g(N,T),$$

where $g(N,T)$ is a penalty function and

$$V(k, \widetilde{\boldsymbol{U}}_T^{(k)}) := \min_{\boldsymbol{B}_N^{(k)} \in \mathcal{L}(\mathbb{R}^k, \boldsymbol{H}_N)} \frac{1}{NT} \sum_{t=1}^{T} \left\| \boldsymbol{x}_t - \boldsymbol{B}_N^{(k)} \widetilde{\boldsymbol{u}}_t^{(k)} \right\|^2 \tag{2}$$

Theorem (Tavakoli, Nisol and Hallin, 2020)

Under Assumptions A, B, C and D, if

$$g(N,T) \to 0 \quad \textit{and} \quad C_{N,T}\, g(N,T) \to \infty,$$

as $C_{N,T} := \min\{\sqrt{N}, \sqrt{T}\} \to \infty$, *then*

$$\mathbb{P}(\hat{r} = r) \longrightarrow 1, \quad \textit{as } C_{N,T} \to \infty.$$

Estimating the Number of Factors: Remarks

- the penalty should converge to zero slow enough that $C_{N,T}\, g(N,T) \to \infty$; this (which is consistent with Amengual and Watson (2007)) is stronger than Bai and Ng's condition that $C^2_{N,T}\, g(N,T) \to \infty$; but Bai and Ng require $\mathbb{E}[|\xi_{it}|^7] < \infty$. Since we have control over $g(N,T)$ but not on $\mathbb{E}[|\xi_{it}|^7]$, stronger conditions on $g(N,T)$ are preferable
- in the particular case $H_i = \mathbb{R}$, Bai and Ng also require $\mathbb{E}\,\|\boldsymbol{u}_t\|^4 < \infty$ and $\mathbb{E}\,|\xi_{it}|^8 < \infty$, which we do not need
- we also are weakening their assumption

$$\mathbb{E}\left|\sqrt{N}\left(\langle \boldsymbol{\xi}_t, \boldsymbol{\xi}_s \rangle / N - \nu_N(t-s)\right)\right|^4 < M < \infty, \quad \forall s, t, N \geq 1,$$

  on idiosyncratic cross-covariances into $\mathbb{E}\left|\sqrt{N}\left(\langle \boldsymbol{\xi}_t, \boldsymbol{\xi}_s \rangle / N - \nu_N(t-s)\right)\right|^2 < M$ thanks to a sharp use of Hölder inequalities between Schatten norms of compositions of operators
- in practice, we recommend combining the method considered here with the tuning device proposed in Hallin and Liška (2007) and Alessi, Barigozzi, and Capasso (2009)

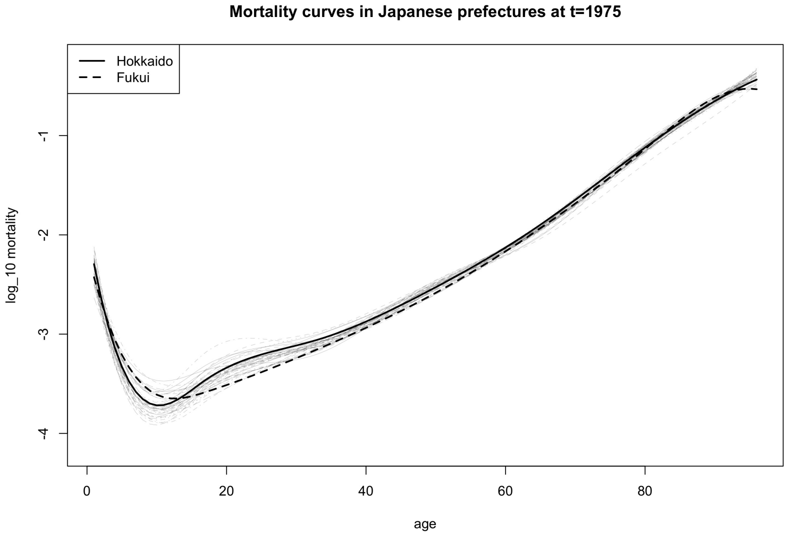
Application to forecasting mortality curves in Japan

Data

- 47 Japanese prefectures ($N = 47$),
- Yearly mortality curves from 1975 through 2016 ($T = 42$),
- Mortality curves by gender (female, male),
- Same dataset as Gao, Shang & Yang (2019).

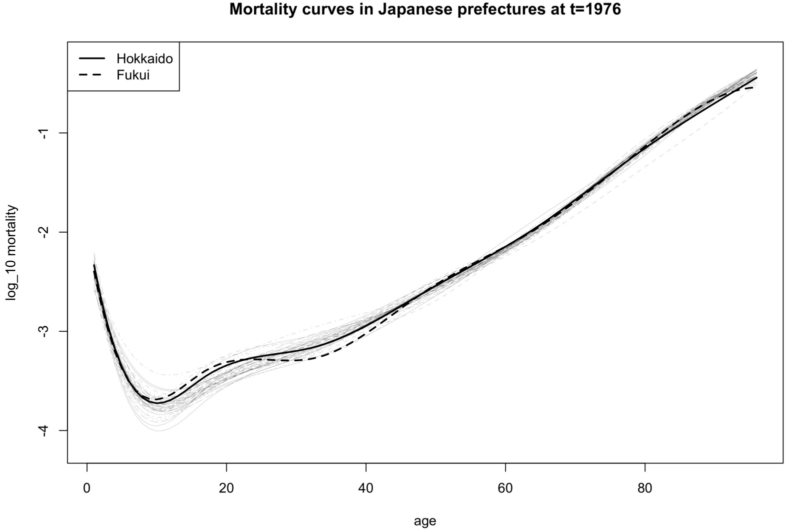
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



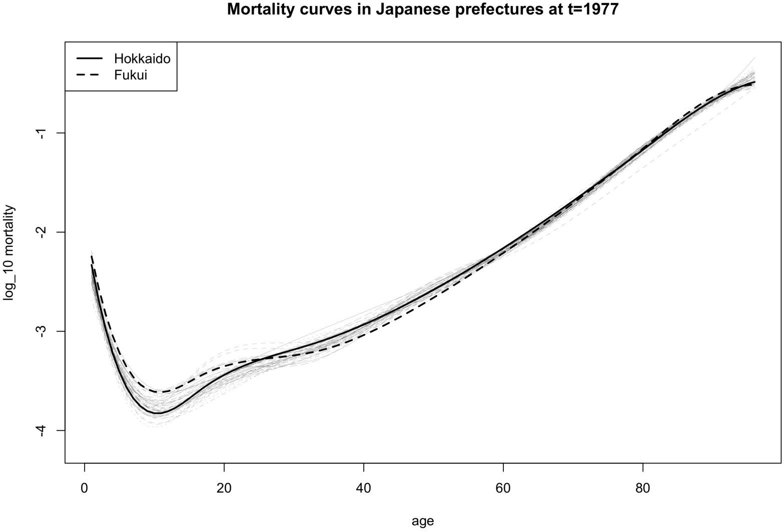
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



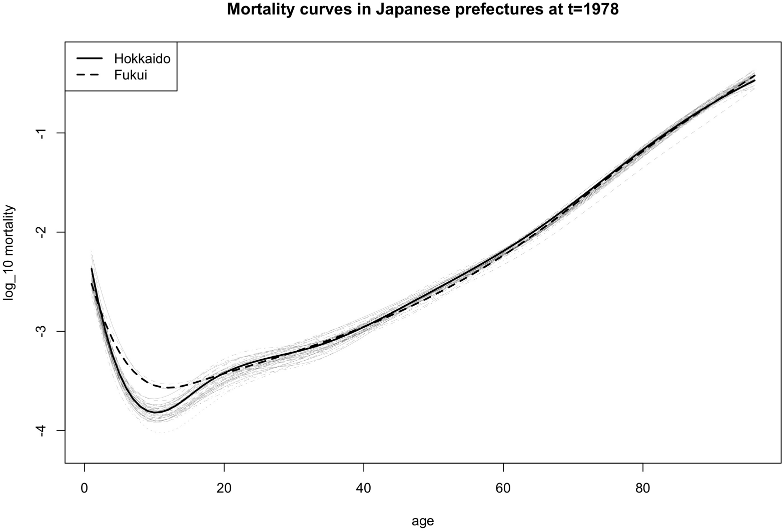
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



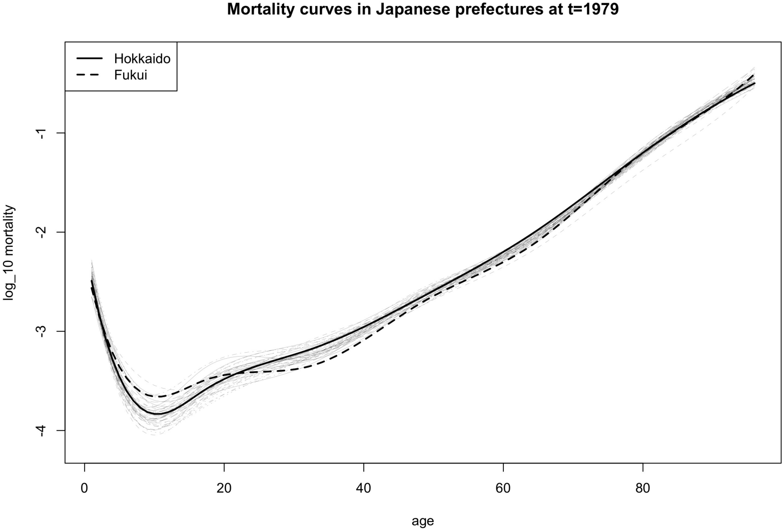
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



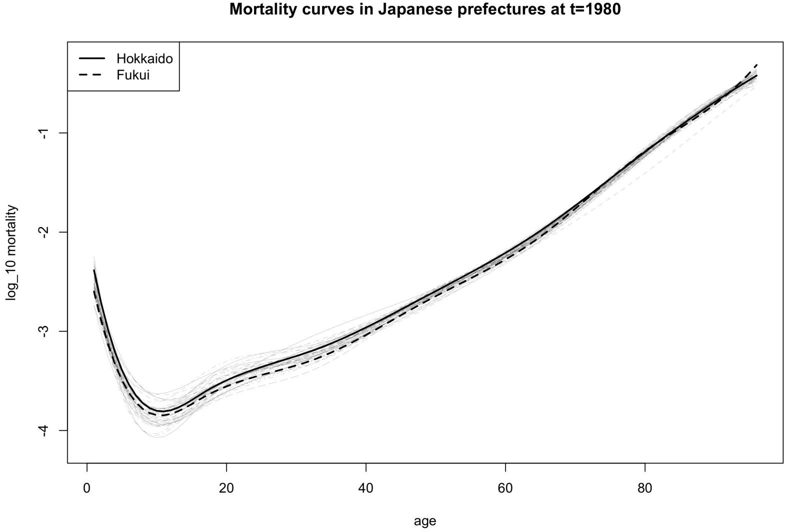
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



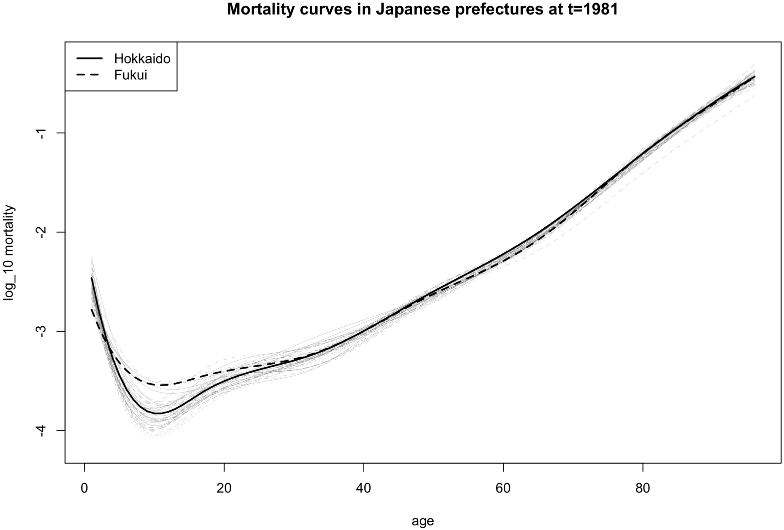
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



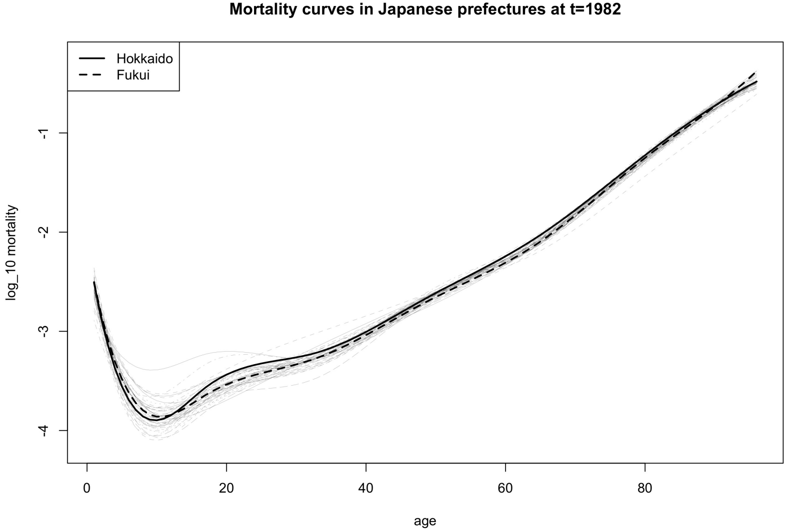
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



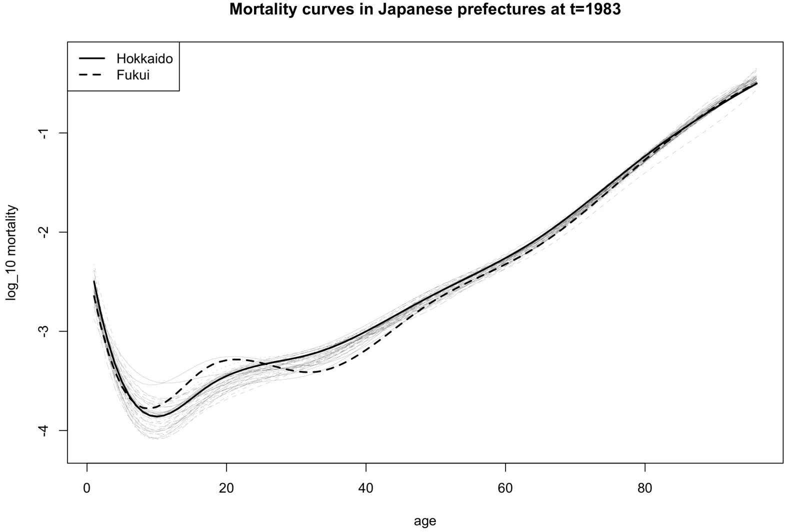
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



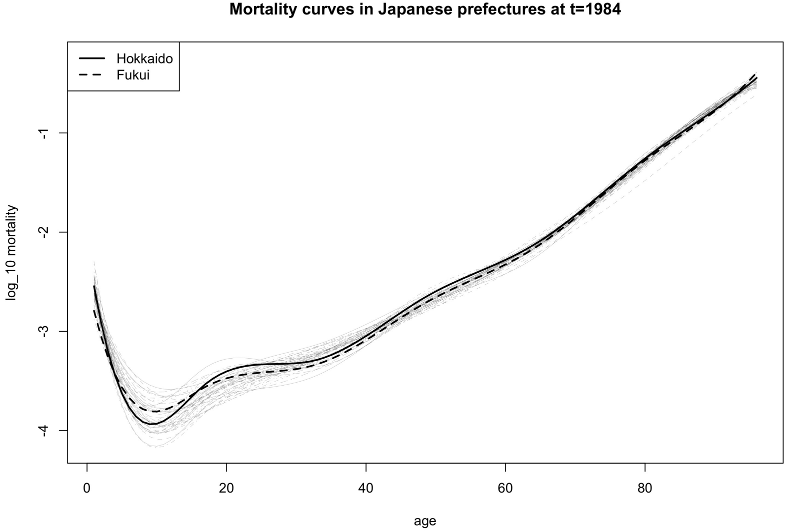
Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



Application to forecasting mortality curves in Japan

Yearly mortality curves of $N = 47$ Japanese prefectures for 1975–2016 ($T = 42$).



Application to forecasting mortality curves in Japan

Comparing 3 forecasting models

- GSY Method of Gao, Shan & Yang (2019), based on separate *scalar* factor models on the FPCA scores of each FTS (each i), with an ARMA model on the factors.
- CF Componentwise forecasting using ARIMA models on FPCA scores (Happ & Greven 2018).
- TNH Our method (identification of the number of factors yields $r = q = 1$), based on an ARIMA model on the estimated factor, and ARIMA models on idiosyncratics.

Measures of Performance

- MAFE Mean absolute forecasting error,
- MSFE Mean squared forecasting error.

Forecasting performance

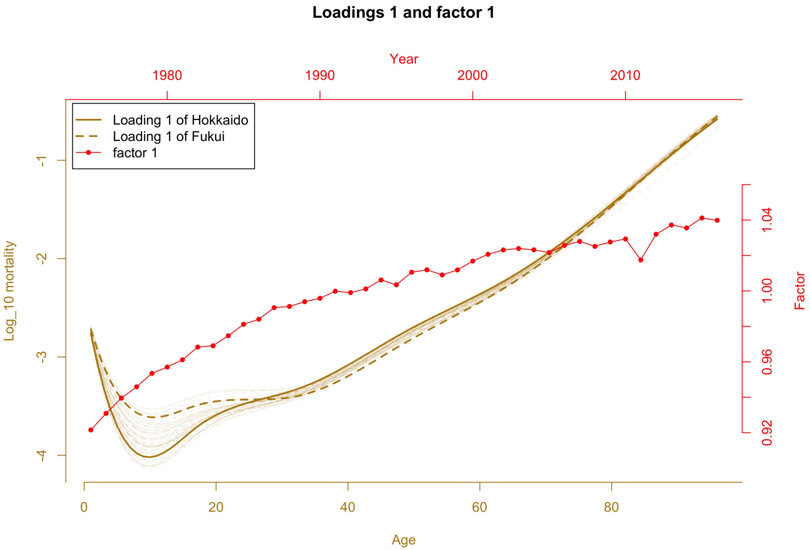
Forecasting Errors ($\times 1000$)

	Female						Male					
	MAFE			MSFE			MAFE			MSFE		
	GSY	CF	TNH	GSY	CF	TNH	GSY	CF	TNH	GSY	CF	TNH
$h = 1$	296	286	**250**	190	166	**143**	268	232	**221**	167	124	**122**
$h = 2$	295	294	**252**	187	171	**145**	271	243	**224**	171	131	**124**
$h = 3$	294	301	**254**	190	176	**148**	270	252	**227**	170	136	**126**
$h = 4$	300	305	**258**	195	178	**152**	274	259	**230**	177	141	**129**
$h = 5$	295	308	**259**	190	179	**154**	270	268	**233**	169	146	**131**
$h = 6$	295	313	**259**	194	181	**156**	271	278	**235**	169	152	**134**
$h = 7$	302	321	**263**	200	187	**161**	266	289	**240**	164	160	**138**
$h = 8$	298	329	**269**	192	193	**167**	266	302	**245**	161	168	**142**
$h = 9$	303	339	**275**	203	199	**172**	277	315	**251**	169	178	**148**
$h = 10$	308	347	**280**	209	205	**177**	283	327	**254**	174	186	**150**
Mean	299	314	**262**	195	183	**157**	272	277	**236**	169	152	**134**
Median	297	311	**259**	193	180	**155**	271	273	**234**	169	149	**133**

GSY = Gao, Shan & Yang (2019)
CF = Component-wise forecasting
TNH = our method

h is the number of steps ahead for forecasting
in red: **minimal prediction error amongst the 3 methods**

Flexibility of our method



Concluding Remarks

High-dimensional functional factor models:

- Mixed natured panels;
- Representation result: links between high-dimensional functional factor models and eigenvalues of covariance operator;
- $N, T \to \infty$ asymptotics (no cross-constraints);
- Estimation and consistency of factors, loadings, common component, and number of factors;
- Results inspired by the scalar case [Chamberlain & Rothschild, 1983; Forni et al. 2000; Bai & Ng 2002; Stock & Watson 2002; Fan et al. 2013, and many others] and reducing to scalar case results as a special case but with weaker assumptions;

References

[1] Hallin M., Tavakoli S., & Nisol G. (2023), 'High-dimensional functional factor models I: Representation results. *Journal of Time Series Analysis* **44**, 578–600.

[2] Tavakoli S., Nisol G., & Hallin M. (2023), 'High-dimensional functional factor models II: Estimation and forecasting. *Journal of Time Series Analysis* **44**, 601–621.