The ftsa Package

Han Lin Shang
Monash University

Abstract

Recent advances in computer recording and storing technology have tremendously increased the presence of functional data, whose graphical representation can be infinite-dimensional curve, image, or shape. When the same functional object is observed over a period of time, such data are known as functional time series. This article makes first attempt to describe several techniques for modeling and forecasting functional time series from a computational aspect, using a readily-available ℝ addon package. These methods are demonstrated using age-specific Australian fertility rates from 1921 to 2006, and monthly sea surface temperature from January 1950 to December 2011.

Keywords: functional time series visualization, functional principal component analysis, functional partial least squares regression.

Introduction

The aim of this article is to describe the ℝ codes that are readily-available in the **ftsa** package (Hyndman and Shang 2012), for modeling and forecasting functional time series. This article was motivated by recent advances in computer recording and storing technology that have enabled researchers to collect and store (ultra) high-dimensional data. When the high-dimensional data are repeatedly measured on the same object over a period of time, a time series of continuous functions is observed within a common bounded interval (Shen and Huang 2008).

Analyzing functional time series has received increasing attention in the functional data analysis literature (see for example, Hörmann and Kokoszka 2010; Horváth, Hušková, and Kokoszka 2010). Hyndman and Shang (2010) presented a rainbow plot for visualizing functional time series, where the distant past data are shown in red and most recent data are shown in purple. Aguilera, Ocana, and Valderrama (1999) proposed functional principal component regression (FPCR) to model and forecast functional time series.

Before reviewing the FPCR, we first define the problem more precisely. Let $y_t(x)$ denote a function, such as age-specific fertility rates for the continuous variable age x in year t, or sea surface temperature for the continuous variable time x in year t. We assume that there is an underlying smooth function $f_t(x)$ that observes with error at discretized grid points of x. In practice, we observe $\{x_i, y_t(x_i)\}$ for $t = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, p$, from which we extract a smooth function $f_t(x)$, given by

$$y_t(x_i) = f_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i}, \tag{1}$$

where $\varepsilon_{t,i}$ is an independent and identically distributed (iid) standard normal random variable,

$\sigma_t(x_i)$ allows the amount of noise to vary with x_i, and $\{x_1, x_2, \ldots, x_p\}$ is a set of discrete data points. Given a set of functional data $\boldsymbol{f}(x) = [f_1(x), f_2(x), \ldots, f_n(x)]^\top$, we are interested in finding underlying patterns using the FPCR, from which we obtain forecasts $f_{n+h}(x)$, where h denotes the forecast horizon.

This article proceeds as follows. Techniques for modeling and forecasting functional time series are reviewed and implemented using the **ftsa** package. Conclusions are given in the end.

Functional time series modeling and forecasting techniques

Functional principal component regression

The theoretical, methodological and practical aspects of functional principal component analysis (FPCA) have been extensively studied in the functional data analysis literature, since it allows finite dimensional analysis of a problem that is intrinsically infinite-dimensional (Hall and Hosseini-Nasab 2006). Numerous examples of using FPCA as an estimation tool in regression problem can be found in different fields of applications, such as breast cancer mortality rate modeling and forecasting (Erbas, Hyndman, and Gertig 2007), call volume forecasting (Shen and Huang 2008), climate forecasting (Shang and Hyndman 2011), demographical modeling and forecasting (Hyndman and Shang 2009), and electricity demand forecasting (Antoch, Prchal, De Rosa, and Sarda 2008).

At a population level, a stochastic process denoted by f can be decomposed into the mean function and the sum of the multiplications of orthogonal functional principal components and uncorrelated principal component scores. It can be expressed as

$$f = \mu + \sum_{k=1}^{\infty} \beta_k \phi_k,$$

where μ is the unobservable population mean, β_k is the k^{th} principal component scores, and ϕ_k is the k^{th} population functional principal component.

In practice, we can only observe n realizations of f evaluated at a compact interval $x \in [0, \tau]$, denoted by $f_t(x)$ for $t = 1, 2, \ldots, n$. At a sample level, the functional principal component decomposition can be written as

$$f_t(x) = \bar{f}(x) + \sum_{k=1}^{K} \hat{\beta}_{t,k} \hat{\phi}_k(x) + \hat{\varepsilon}_t(x), \tag{2}$$

where $\bar{f}(x) = \frac{1}{n}\sum_{t=1}^{n} f_t(x)$ is the estimated mean function, $\hat{\phi}_k(x)$ is the k^{th} estimated orthonormal eigenfunction of the empirical covariance matrix

$$\hat{\Gamma}(x) = \frac{1}{n}\sum_{t=1}^{n} [f_t(x) - \bar{f}(x)][f_t(x) - \bar{f}(x)],$$

the coefficient $\hat{\beta}_{t,k}$ is the k^{th} principal component score for year t, it is given by the projection of $f_t(x) - \bar{f}(x)$ in the direction of k^{th} eigenfunction $\hat{\phi}_k(x)$, that is, $\hat{\beta}_{t,k} = < f_t(x) - \bar{f}(x), \hat{\phi}_k(x) > = \int_x [f_t(x) - \bar{f}(x)]\hat{\phi}_k(x)dx$, $\hat{\varepsilon}_t(x)$ is the residual term, and K is the retained number of components.

The functional principal component decomposition is demonstrated using the age-specific Australian fertility rates between ages 15 and 49 observed from 1921 to 2006, obtained from the Australian Bureau of Statistics (Cat No, 3105.0.65.001, Table 38). A functional graphical display is given in Shang (2011).

Figure 1 presents the first two functional principal components and their associated principal component scores. The bottom panel of Figure 1 also plots the forecasted principal component scores, and their 80% prediction intervals (in yellow color), using an exponential smoothing state-space model (Hyndman, Koehler, Ord, and Snyder 2008).

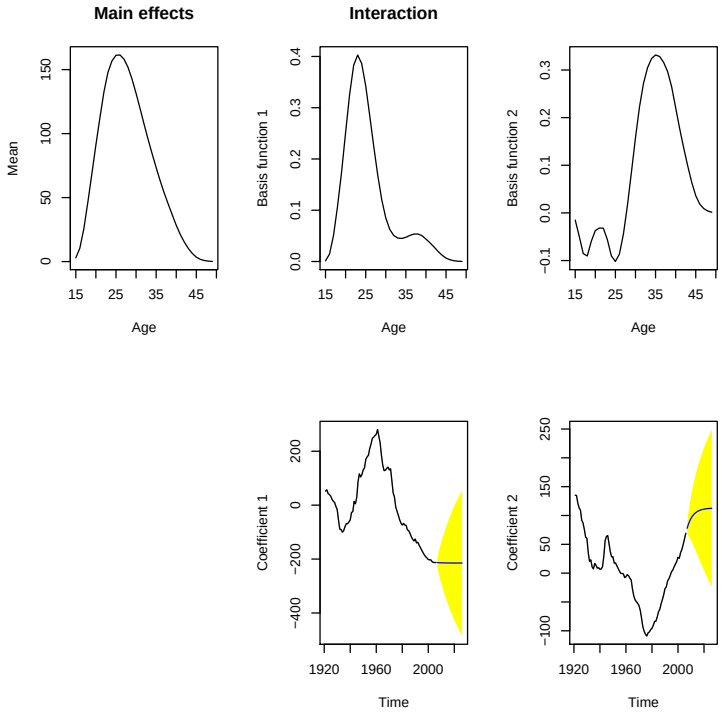

Figure 1: The first two functional principal components and their associated principal component scores for the Australian fertility data from 1921 to 2006.

Figure 1 was produced by the following code.

```
# load the package used throughout this article
library("ftsa")
# Fit and plot functional principal components
# order specifies the number of principal components
# h specifies the forecast horizon
plot(forecast(ftsm(Australiasmoothfertility, order=2), h = 20), "components")
```

By conditioning on the observed data $\boldsymbol{f}(x) = [f_1(x), f_2(x), \ldots, f_n(x)]^\top$ and the fixed functional principal components $\mathcal{B} = [\hat{\phi}_1(x), \hat{\phi}_2(x), \ldots, \hat{\phi}_K(x)]^\top$, the h-step-ahead forecasts of $y_{n+h}(x)$ can be

obtained as

$$\hat{y}_{n+h|n}(x) = \mathrm{E}[y_{n+h}(x)|\boldsymbol{f}(x), \mathcal{B}] = \bar{f}(x) + \sum_{k=1}^{K} \hat{\beta}_{n+h|n,k}\hat{\phi}_k(x),$$

where $\hat{\beta}_{n+h|n,k}$ denotes the h-step-ahead forecasts of $\beta_{n+h,k}$ using a univariate time series. Figure 2 shows the forecasts of Australian fertility rates from 2007 to 2026 highlighted in rainbow color, while the data used for estimation are grayed out. Forecasts exhibit a continuing shift to older ages of peak fertility rates, caused by a recent tendency to postpone child-bearing while pursuing careers.

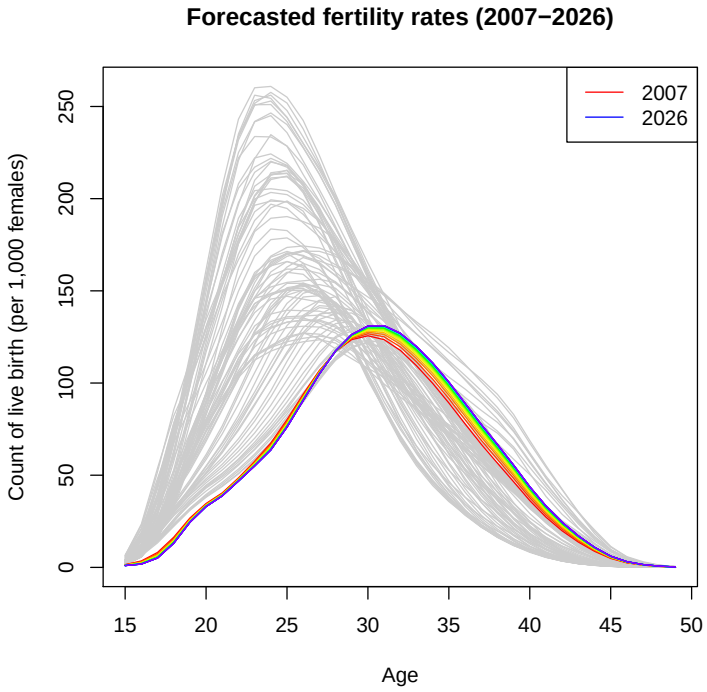

Figure 2: Multiple-step-ahead forecasts of the Australian fertility rates from 2007 to 2026, based on the first two functional principal components and their associated principal component scores as an illustration.

To construct a prediction interval, we calculate the forecast variance that follows from (1) and (2). Because of orthogonality, the forecast variance can be approximated by the sum of component variances

$$\begin{aligned}\xi_{n+h}(x) &= \mathrm{Var}[y_{n+h}(x)|\boldsymbol{f}(x), \mathcal{B}] \\ &= \hat{\sigma}^2_{\mu}(x) + \sum_{k=1}^{K} u_{n+h,k}\hat{\phi}^2_k(x) + v(x) + \sigma^2_{n+h}(x),\end{aligned}$$

where $u_{n+h,k} = \text{Var}(\beta_{n+h,k}|\beta_{1,k}, \beta_{2,k}, \ldots, \beta_{n,k})$ can be obtained from the time series model, and the model error variance $v(x)$ is estimated by averaging $\{\hat{\varepsilon}_1^2(x), \hat{\varepsilon}_2^2(x), \ldots, \hat{\varepsilon}_n^2(x)\}$ for each x, and $\hat{\sigma}_\mu^2(x)$ and $\sigma_{n+h}^2(x)$ can be obtained from the smoothing method used.

Based on the normality assumption, the $100(1-\alpha)\%$ prediction interval for $y_{n+h}(x)$ is constructed as $\hat{y}_{n+h|n}(x) \pm z_\alpha \sqrt{\xi_{n+h}(x)}$, where z_α is the $(1-\alpha/2)$ standard normal quantile.

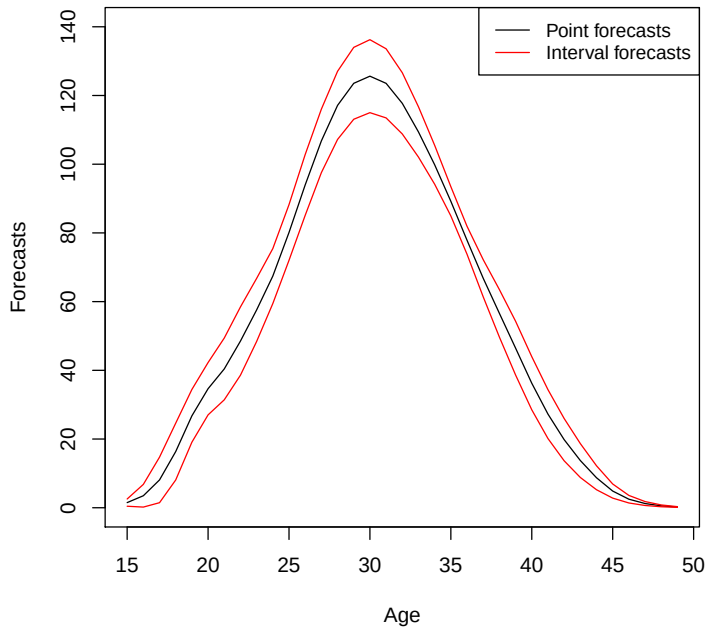

Figure 3: Forecasts of fertility rates in 2007, along with the 80% prediction interval.

Figure 3 displays the forecasts of fertility rates in 2007, along with the 80% prediction interval. It was created by the following code.

```
# Plot the point forecast
aus = forecast(ftsm(Australiasmoothfertility, order=2), h=1)
plot(aus, ylim=c(0,140))
# Plot the lower and upper bounds
lines(aus$lower, col=2)
lines(aus$upper, col=2)
# Add a legend to the plot
legend("topright", c("Point forecasts", "Interval forecasts"),col=c(1,2), lty=1, cex=0.9)
```

Updating point and interval forecasts

A special case of functional time series is when the continuous variable is also a time variable,

such as the monthly sea surface temperature from 1950 to 2011, obtained from National Oceanic and Atmospheric Administration (http://www.cpc.noaa.gov/data/indices/sstoi.indices). A similar type of functional graphical display is given in Shang (2011, Figure 2). Such data originate from a univariate seasonal time series. Let $\{Z_w, w \in [1, N]\}$ be a seasonal time series which has been observed at N equispaced times. We divide the observed time series into n trajectories, and then consider each trajectory of length p as a curve rather than p distinct data points. The functional time series is given by

$$y_t(x) = \{Z_w, w \in (p(t-1), pt]\}, \quad t = 1, 2, \ldots, n.$$

The problem of interest is to forecast $y_{n+h}(x)$, where h denotes forecast horizon. In the sea surface temperature data, we consider $\{Z_w\}$ be monthly sea surface temperatures from 1950 to 2011, so that $p = 12$ and $N = 62 \times 12 = 744$, and we are interested in forecasting sea surface temperatures in 2012 and beyond.

When $N = np$, all trajectories are complete, and forecasts can be obtained by the FPCR. However, when $N \neq np$, we revisited the block moving (BM) and penalized least squares (PLS) to update point and interval forecasts, when the most recent curve is partially observed.

When functional time series are segments of a univariate time series, the most recent trajectory is observed sequentially (Hyndman and Shang 2010). When we observe the first m_0 time period of $y_{n+1}(x_l)$, denoted by $\boldsymbol{y}_{n+1}(x_e) = [y_{n+1}(x_1), y_{n+1}(x_2), \ldots, y_{n+1}(x_{m_0})]^\top$, we are interested in forecasting the data in the remaining time period, denoted by $y_{n+1}(x_l)$ for $m_0 < l \leq p$. By using the FPCR, the partially observed data in the most recent curve are not incorporated into the forecasts of $y_{n+1}(x_l)$. Indeed, the point forecasts obtained from the FPCR can be expressed as

$$\begin{aligned}\hat{y}_{n+1|n}(x_l) &= \mathrm{E}[y_{n+1}(x_l)|\boldsymbol{f}(x_l), \mathcal{B}_l] \\ &= \bar{f}(x_l) + \sum_{k=1}^{K} \hat{\beta}_{n+1|n,k}\hat{\phi}_k(x_l),\end{aligned}$$

for $m_0 < l \leq p$, where $\boldsymbol{f}(x_l)$ denotes the historical data corresponding to the remaining time periods; $\bar{f}(x_l)$ is the mean function corresponding to the remaining time periods; and $\mathcal{B}_l = \{\hat{\phi}_1(x_l), \hat{\phi}_2(x_l), \ldots, \hat{\phi}_K(x_l)\}$ is a set of the estimated functional principal components corresponding to the remaining time periods.

In order to improve point forecast accuracy, it is desirable to dynamically update the point and interval forecasts for the rest of year $n + 1$ by incorporating the partially observed data. In what follows, I shall revisit two methods for updating point and interval forecasts.

Block moving (BM)

The BM method re-defines the start and end points of trajectories. Because time is a continuous variable, we can change the function support from $[1, p]$ to $[m_0 + 1, p] \bigcup [1, m_0]$. The re-defined functional time series forms a complete block, at the cost of losing some observations in the first year. With the complete data block, the FPCR can then be applied to update the point and interval forecasts.

The redefined data are shown diagrammatically in Figure 4, where the bottom box has moved to become the top box. The cyan colored region shows the data loss in the first year. The partially observed last trajectory under the old "year" completes the last trajectory under the new year.

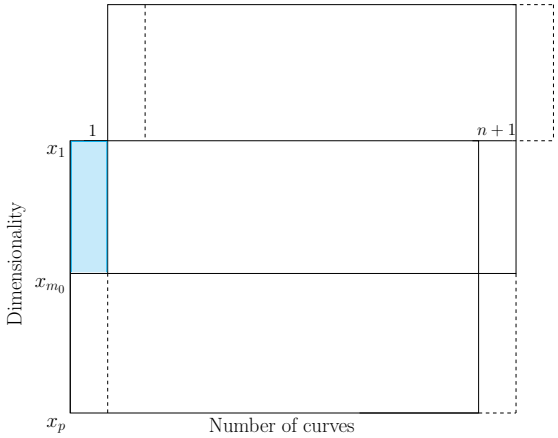

Figure 4: Dynamic update via the block moving approach. The colored region shows the data loss in the first year. The forecasts for the remaining months in year $n+1$ can be updated by the forecasts using the TS method applied to the upper block.

As an illustration, suppose we observe the monthly sea surface temperature from January to May 2011, we aim to update the point and interval forecasts from June to December. Figure 5 displays the point and interval forecasts for the remaining months of 2011, by using the BM method.

Figure 5 was created by the following code

```
# Name history to represent historical data,
history = ElNino2011smooth
# Name obs to represent partially observed data,
obs = ElNino2011smooth$y[1:5,62]
# Name fore to represent the forecasting period
fore = ElNino2011smooth$y[6:12,62]
int = dynupdate(data=history, newdata=obs, holdoutdata=fore, method="block", interval=TRUE)
bmupdate = dynupdate(data=history, newdata=obs, holdoutdata=fore, method="block", value=TRUE)
plot(6:12, fore, type="l", ylim=c(19,26), xlab="Month", ylab="Sea surface temperature")
lines(6:12, bmupdate, col=4)
lines(6:12, int$low$y, col=2)
lines(6:12, int$up$y, col=2)
legend("topright", c("True observations", "Point forecasts", "Interval forecasts"),
      col=c(1,4,2), lty=1, cex=0.8)
```

Penalized least squares (PLS)

We can also update the remaining part of the trajectory by using regression-based approaches. Let $\boldsymbol{F}_e$ be $m_0 \times K$ matrix, whose $(j,k)^{\text{th}}$ entry is $\hat{\phi}_k(x_j)$ for $1 \leq j \leq m_0$. Let $\boldsymbol{\beta}_{n+1} = [\beta_{n+1,1}, \beta_{n+1,2}, \dots, \beta_{n+1,K}]^\top$, $\bar{\boldsymbol{f}}(x_e) = [\bar{f}(x_1), \bar{f}(x_2), \dots, \bar{f}(x_{m_0})]^\top$, and $\boldsymbol{\epsilon}_{n+1}(x_e) = [\epsilon_{n+1}(x_1), \epsilon_{n+1}(x_2), \dots, \epsilon_{n+1}(x_{m_0})]^\top$. As the mean-adjusted $\hat{\boldsymbol{y}}^*_{n+1}(x_e) = \boldsymbol{y}_{n+1}(x_e) - \bar{\boldsymbol{f}}(x_e)$ become available, a regression can be expressed as

$$\hat{\boldsymbol{y}}^*_{n+1}(x_e) = \boldsymbol{F}_e \boldsymbol{\beta}_{n+1} + \boldsymbol{\epsilon}_{n+1}(x_e).$$

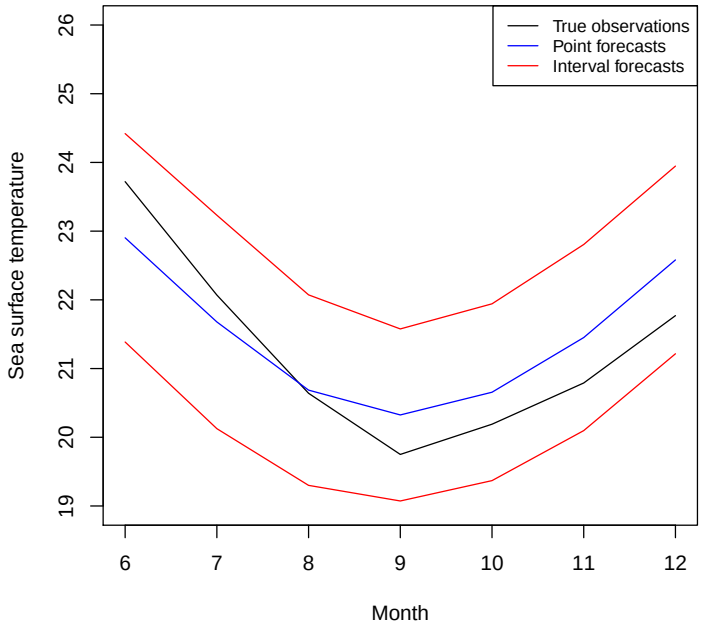

Figure 5: Prediction intervals of the sea surface temperatures between June and December 2011. By incorporating the sea surface temperatures between January and May, the prediction intervals can be updated using the BM method

.

The $\boldsymbol{\beta}_{n+1}$ can be estimated by ordinary least squares, assuming $(\boldsymbol{F}_e^\top \boldsymbol{F}_e)$ is invertible,

$$\hat{\boldsymbol{\beta}}_{n+1}^{\text{OLS}} = (\boldsymbol{F}_e^\top \boldsymbol{F}_e)^{-1} \boldsymbol{F}_e^\top \hat{\boldsymbol{y}}_{n+1}^*(x_e).$$

However, if $(\boldsymbol{F}_e^\top \boldsymbol{F}_e)$ is not invertible, then a regularized approach can be implemented, such as the ridge regression (RR) and penalized least squares (PLS). The regression coefficients of the RR and PLS are

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{n+1}^{\text{RR}} &= (\boldsymbol{F}_e^\top \boldsymbol{F}_e + \lambda \boldsymbol{I}_K)^{-1} \boldsymbol{F}_e^\top \hat{\boldsymbol{y}}_{n+1}(x_e), \\
\hat{\boldsymbol{\beta}}_{n+1}^{\text{PLS}} &= (\boldsymbol{F}_e^\top \boldsymbol{F}_e + \lambda \boldsymbol{I}_K)^{-1} (\boldsymbol{F}_e^\top \hat{\boldsymbol{y}}_{n+1}(x_e) + \lambda \hat{\beta}_{n+1|n}),
\end{aligned} \tag{3}$$

where $\hat{\boldsymbol{\beta}}_{n+1}^{\text{RR}} \to \boldsymbol{0}$ as $\lambda \to \infty$, and $\hat{\boldsymbol{\beta}}_{n+1}^{\text{RR}} \to \hat{\boldsymbol{\beta}}_{n+1}^{\text{OLS}}$ as $\lambda \to 0$. In contrast, the $\hat{\boldsymbol{\beta}}_{n+1}^{\text{PLS}} \to \hat{\boldsymbol{\beta}}_{n+1|n}$ as $\lambda \to \infty$, and $\hat{\boldsymbol{\beta}}_{n+1}^{\text{PLS}} \to \hat{\boldsymbol{\beta}}_{n+1}^{\text{OLS}}$ as $\lambda \to 0$.

The point forecasts of $y_{n+1}(x_l)$ obtained by the RR and PLS are given by

$$\hat{y}_{n+1}^{\text{RR}}(x_l) = \text{E}[y_{n+1}(x_l)|\boldsymbol{f}(x_l), \mathcal{B}_l] = \bar{f}(x_l) + \sum_{k=1}^{K} \hat{\boldsymbol{\beta}}_{n+1,k}^{\text{RR}} \hat{\phi}_k(x_l),$$

$$\hat{y}_{n+1}^{\text{PLS}}(x_l) = \text{E}[y_{n+1}(x_l)|\boldsymbol{f}(x_l), \mathcal{B}_l] = \bar{f}(x_l) + \sum_{k=1}^{K} \hat{\boldsymbol{\beta}}_{n+1,k}^{\text{PLS}} \hat{\phi}_k(x_l),$$

Among these regression-based approaches, the PLS method can also update the interval forecasts. Let the one-step-ahead forecast errors of the principal component scores be given by

$$\hat{\xi}_{j,k} = \hat{\beta}_{n-j+1,k} - \hat{\beta}_{n-j+1|n-j,k}, \quad \text{for} \quad j = 1, 2, \ldots, n-K.$$

$\{\hat{\xi}_{1,k}, \hat{\xi}_{2,k}, \ldots, \hat{\xi}_{n-K,k}\}$ can then be sampled with replacement to give a bootstrap sample of $\beta_{n+1|n,k}$:

$$\hat{\beta}^b_{n+1|n,k} = \hat{\beta}_{n+1|n,k} + \hat{\xi}^b_{*,k}, \quad \text{for} \quad b = 1, 2, \ldots, B,$$

where $\hat{\xi}^b_{*,k}$ denotes the bootstrap samples, and B is the number of bootstrap replications. Based on (3), the bootstrapped $\hat{\boldsymbol{\beta}}^b_{n+1|n}$ leads to the bootstrapped $\hat{\boldsymbol{\beta}}^{b,\text{PLS}}_{n+1}$, we obtain B replications of

$$\hat{y}^{b,\text{PLS}}_{n+1}(x_l) = \bar{f}(x_l) + \sum_{k=1}^{K} \hat{\beta}^{b,\text{PLS}}_{n+1,k} \hat{\phi}_k(x_l) + \hat{\epsilon}^b_{n+1}(x_l),$$

where $\hat{\epsilon}^b_{n+1}(x_l)$ is obtained by sampling with replacement from $\{\hat{\epsilon}_1(x_l), \hat{\epsilon}_2(x_l), \ldots, \hat{\epsilon}_n(x_l)\}$. Hence, the $100(1-\alpha)\%$ prediction intervals for the updated forecasts are defined as $\alpha/2$ and $(1-\alpha/2)$ quantiles of $\hat{y}^{b,\text{PLS}}_{n+1}(x_l)$.

Conclusion

This article described several techniques in the **ftsa** package, for modeling and forecasting functional time series. These methods centered on the FPCR, which is a common dimension reduction technique in the functional data analysis literature. FPCR reduces intrinsically infinite number of variables to several orthogonal regressors, which captures the main mode of variation in data. As illustrated by the Australian fertility rates, FPCR is able to model and forecast annual Australian fertility rates. When the continuous variable in a functional time series is also a time variable, a new observation arrives sequentially. As shown in the monthly sea surface temperature, the BM and PLS methods can update the point and interval forecasts based on the FPCR.

To sum up, the methods reviewed in this article focus on extracting patterns from a set of functional time series, and should be considered when the interest lies in modeling and forecasting the future realizations of a stochastic process.

References

Aguilera A, Ocana F, Valderrama M (1999). "Forecasting time series by functional PCA. Discussion of several weighted approaches." *Computational Statistics*, **14**(3), 443–467.

Antoch J, Prchal L, De Rosa MR, Sarda P (2008). "Functional linear regression with functional response: application to prediction of electricity consumption." In S Dabo-Niang, F Ferraty (eds.), *Functional and Operatorial Statistics*, pp. 23–29. Physica-Verlag, Heidelberg.

Erbas B, Hyndman RJ, Gertig DM (2007). "Forecasting age-specific breast cancer mortality using functional data models." *Statistics in Medicine*, **26**(2), 458–470.

Hall P, Hosseini-Nasab M (2006). "On properties of functional principal components analysis." *Journal of the Royal Statistical Society: Series B*, **68**(1), 109–126.

Hörmann S, Kokoszka P (2010). "Weakly dependent functional data." *The Annals of Statistics*, **38**(3), 1845–1884.

Horváth L, Hušková M, Kokoszka P (2010). "Testing the stability of the functional autoregressive process." *Journal of Multivariate Analysis*, **101**(2), 352–367.

Hyndman R, Shang HL (2012). *ftsa: Functional time series analysis*. R package version 3.1, URL http://CRAN.R-project.org/package=ftsa.

Hyndman RJ, Koehler AB, Ord JK, Snyder RD (2008). *Forecasting with exponential smoothing: the state space approach*. Springer, Berlin.

Hyndman RJ, Shang HL (2009). "Forecasting functional time series (with discussion)." *Journal of the Korean Statistical Society*, **38**(3), 199–221.

Hyndman RJ, Shang HL (2010). "Rainbow plots, bagplots, and boxplots for functional data." *Journal of Computational and Graphical Statistics*, **19**(1), 29–45.

Shang HL (2011). "rainbow: an R package for visualizing functional time series." *The R Journal*, **3**(2), 54–59.

Shang HL, Hyndman RJ (2011). "Nonparametric time series forecasting with dynamic updating." *Mathematics and Computers in Simulation*, **81**(7), 1310–1324.

Shen H, Huang JZ (2008). "Interday forecasting and intraday updating of call center arrivals." *Manufacturing & Service Operations Management*, **10**(3), 391–410.

Affiliation:

Han Lin Shang
Department of Econometrics & Business Statistics
Monash University
Melbourne, VIC, 3800,
E-mail: HanLin.Shang@monash.edu
URL: http://monashforecasting.com/index.php?title=User:Han