Time Series of Functional Data for Forecasting the Yield Curve and Electricity Prices

Rituparna Sen[1,*] and Claudia Klüppelberg[2]

[1]Applied Statistics Unit, Indian Statistical Institute, Chennai Center, India.
[2]Technical University of Munich
[*]Corresponding author. Email: rsen@isichennai.res.in

Abstract

We develop time series analysis of functional data, treating the whole curve as a random realization from a distribution on functions that evolve over time. The method consists of principal components analysis of functional data and subsequently modeling the principal component scores as vector ARMA. We carry out the estimation of VARMA parameters using the equivalent state space representation. We derive asymptotic properties of the estimators and the fits. We apply the method to two different data. For term structures of interest rates, this provides a unified framework for studying the time and maturity components of interest rates under one set-up with few parametric assumptions. We compare our forecasts to the parametric Diebold and Li (2006) model. Secondly, we apply this approach to hourly spot prices of electricity and obtain fits and forecasts that are better than those existing in the electricity literature.

Keywords: Functional Principal Components, Vector ARMA, Forecasting, Term Structure, Electricity spot prices.

1   Introduction

Functional data analysis (see Ramsay and Silverman (2005) for a comprehensive introduction to FDA methods) is an extension of multivariate data analysis to functional data. In this framework, each individual is characterized by one or more real valued functions, rather than by a vector in $\mathbb{R}^n$. An important tool of functional data analysis (FDA) is functional principal component analysis (FPCA, see Rice and Silverman (1991)). Functional processes can be characterized by their mean function and the eigenfunctions of the auto-covariance operator. This is a consequence of the Karhunen-Loéve representation of the functional process. This often leads to substantial dimension reduction.

Most of the development in FDA has been with independent and identical replications of data. However, in certain situations, it is unrealistic to assume that the functions across time are independent. Besse et al. (2000) develop an AR(1) model for FDA for forecasting climatic variations. Kargin and Onatski (2008) use an AR(1) model for forecasting Eurodollar futures. We develop the theory, where the functions follow a general ARMA(p, q) model. We start with dimension reduction using FPCA. Based on the time series of the first few significant principal components, we fit a VAR or VARMA model. We provide techniques for estimation of the model parameters and selection of the optimal model.

2   Theory

Consider a sample of n smooth random trajectories $(f_i(t))_{t \in \mathcal{T}}$ for $i = 1, \cdots, n$ generated from a process f. We denote the Hilbert space of such functions by $\mathcal{H}$. The observed measurements are available on a dense grid of support points t_{ij} on the domain $\mathcal{T} = [a_1, a_2]$ with additive white noise error W_{ij} which is independent of the underlying process. The measurements are:

$$\tilde{f}_i(t_{ij}) = f_i(t_{ij}) + W_{ij}, \quad i = 1, \cdots, n, \quad j = 1, \cdots, m \quad \mathrm{E}(W_{ij}) = 0, \mathrm{Var}(W_{ij}) = \sigma^2. \quad (1)$$

We represent the smooth functional f in terms of its decomposition into functional principal components, a common approach in FDA. For a domain $\mathcal{T}$, setting

$$\mathbf{G}_f(s,t) = \mathrm{Cov}(f(s), f(t)), \quad E(f(t)) = \mu_f(t), \quad s, t \in \mathcal{T}, \quad (2)$$

the functional principal components are the eigenfunctions of the auto-covariance operator $\mathbf{G}_f : L^2 \mapsto \mathbb{R}$. We denote the orthonormal eigenfunctions by ϕ_k, with associated eigenvalues λ_k for $k = 1, 2, \ldots$, such that $\lambda_1 \geq \lambda_2 \geq \cdots$ and $\sum_k \lambda_k < \infty$. The Karhunen-Loève theorem (see Rice and Silverman (1991)) provides a representation of individual random trajectories of the functional f, given by

$$f(t) = \mu_f(t) + \sum_{k=1}^{\infty} \xi_k \phi_k(t), \quad t \in \mathcal{T}, \quad (3)$$

where the ξ_k are uncorrelated random variables. Well-known procedures exist to infer eigenfunctions and eigenvalues (Rice and Silverman (1991); Müller et al. (2006)).

Processes f are then approximated by substituting estimates and using a judiciously chosen finite number K of terms in sum (3). This choice can be made using one-curve-leave-out cross-validation (Rice and Silverman (1991)), pseudo-AIC criteria (Yao et al. (2005)) or a scree plot, a tool from multivariate analysis, where one uses estimated eigenvalues to obtain a prespecified fraction of variance explained as a function of K or looks for a change-point.

We assume that the series of functions follows the ARMAH(p, q) model with mean $\mu \in \mathcal{H}$:

$$f_i(\cdot) - \mu = \theta_1(f_{i-1}(\cdot) - \mu) + \cdots + \theta_p(f_{i-p}(\cdot) - \mu) + \epsilon_i(\cdot). \quad (4)$$

where $\epsilon_i(\cdot) = \eta_i(\cdot) + \psi_1 \eta_{i-1}(\cdot) + \cdots + \psi_q \eta_{i-q}(\cdot)$, and $\eta_i(\cdot)$ is $\mathcal{H}$ white noise. $\theta_1, \cdots, \theta_p$ are linear functions. Combining (4) and (3) and using linearity of $\theta_1, \cdots, \theta_p$ and orthonormality of the columns of Φ, we can get:

$$\Xi_i = \Phi^T \tilde{\mu} + \Phi^T \theta_1(\Phi(\cdot))\Xi_{i-1} + \cdots + \Phi^T \theta_p(\Phi(\cdot))\Xi_{i-p} + \phi^T \epsilon(\cdot). \quad (5)$$

This implies a VARMA(p, q) structure on the vector of principal component scores Ξ_i.

3   Methodology

At the core of the estimation procedure is the principal analysis of random trajectories (PART), applied to the data $\tilde{f}_{ij}$ from (1), which is an algorithm to obtain mean and eigenfunctions, as well as FPC scores, from densely sampled functional data, as described in Müller et al. (2006).

The PART algorithm yields estimates of the individual FPC scores as

$$\hat{\xi}_{ik} = \sum_{j=2}^{m} (\tilde{f}_{ij} - \hat{\mu}_f(t_{ij}))(t_{ij} - t_{ij-1})\hat{\phi}_k(t_{ij}), \quad i = 1, \ldots, n, \, k = 1, 2, \ldots. \tag{6}$$

Individual trajectories can then be represented by an empirical version of the Karhunen-Loève expansion (3),

$$\hat{f}_i^{(K)}(t) = \hat{\mu}_f(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t). \tag{7}$$

The estimated principal component score vectors $\hat{\xi}_i = (\hat{\xi}_{i1}, \cdots, \hat{\xi}iK)$ form a vector time series of length n. The infinite dimension of the functional data has been reduced to a finite dimension K. We fit Vector Autoregressive Moving Average(VARMA) models of order p, q to the finite dimensional time series of estimated principal component scores $\hat{\xi}_i$. For details on VARMA models see Chapter 11 of Brockwell and Davis (2009).

Model selection and forecasts can be done conveniently by using the equivalent representation of VARMA using state space models proposed by Akaike (1976). For details consider Aoki and Havenner (1991). The main value of the state space approach is its capability to find the best model in terms of the Akaike information criterion.

Our primary aim is forecasting the curve for a future date based on the information available upto a certain point of time. The final VARMA(p, q) model is used to produce model forecasts $\tilde{\xi}_{ik}$ of future principal component scores. Plugging these into equation (7) we obtain the forecasts $\widehat{\tilde{f}}_i(t)$ of the original process f.

$$\widehat{\tilde{f}}_i(t) = \hat{\mu}_f(t) + \sum_{k=1}^{K} \tilde{\xi}_{ik}\hat{\phi}_k(t). \tag{8}$$

Diebold and Li (2006), henceforth referred to as DL, use parametric functions involving variations of Nelson-Siegel exponential components to model the yield curve and then use univariate AR(1) models componentwise to estimate and forecast the factors. We compare our results to their work.

4   Empirical Examples

Euribor (Euro Interbank Offered Rate) is the rate at which Euro interbank term deposits are being offered by one prime bank to another within the European Monetary Union. Historical data is available at www.euribor.org Since its launch, Euribor has become a reality on the derivatives markets and is the underlying rate of many derivatives transactions, both, over-the-counter and exchange-traded.

The electricity data was obtained on day-ahead spot prices from European Energy Exchange (EEX). With high trading volumes, the EEX is one of the largest and most important power exchanges in Europe. The EEX operates a day-ahead market for hourly and block electricity contracts. Hourly power contracts are traded daily for physical delivery in the next day's 24-hour period (midnight to midnight).

The initial fitting of functional data to obtain mean, covariance and principal components is done by employing the PACE package for functional data analysis written in Matlab. We use the Gaussian kernel. The package is available at

http://anson.ucdavis.edu/~wyang/PACE/

VAR model fitting and diagnostics is done using the econometrics toolbox in Matlab.

VARMA and related state space model computations are done using the Dynamic Systems Estimation (dse) package in R available at

http://cran.r-project.org/web/packages/dse/index.html.

For examples of some of the capabilities of the dse package see Gilbert (1993).

We separate the data into years because for long time horizons the stationarity assumption of the time series may not be valid. We present the results for the years 1999 and 2007. Together they are representative of the other years. In Figure 1 we present the distance between the observed functions and the predictions based on VARMA models for the principal component scores. In this case, at least for the year 1999, the FDA method outperforms the DL method. In general, the errors from the FDA method have smaller variability than those from the DL method.

Data are hourly spot prices for weekdays from Jan 1 to Sept 30 2008. The number of days is $n = 196$. As noted in section 2, electricity prices over long time horizons exhibit nonstationarity. We choose a period that looks like a stationary period from time series plots of the raw data and the principal component scores. The price is observed in one hour intervals over the whole day, so $\tau = 1, 2, \cdots, 24$. The data are of dimension 24 and we think of it as a time series of functions representing spot prices over each day.

The measure we use for evaluating the model is the percentage of variance explained compared to the simple invariant estimate given by the overall mean. This is like an R^2 criterion and is defined as:

$$1 - \frac{(y_{ij} - \hat{y}_{ij})^2}{(y_{ij} - \bar{y})^2}$$

This value equals 0.9503 for the fitted model and equals 0.7713 for the forecast using VARMA(1,0).

We compare the forecasts with the AR and iterated Hseih-Manski (IHMAR) models as described in Weron and Misiorek (2008). The forecasts are compared using the weekly-weighted mean absolute error(WMAE).

The WMAE for the 38 weeks are computed for the three models: pure AR, FDA and IHMAR. The summary statistics are presented in the table. These include the mean WMAE over all weeks, the number of times a given model was best and the mean deviation from the best model (m.d.f.b.) in each week. The latter measure indicates which approach is closest to the optimal model on the average.

It is seen that IHMAR does not give best forecast for any week. The FDA method does not dominate the AR model, but performs better on the average according to all three summary statistics.

References

[1] Akaike, H. (1976) "An information criterion (AIC)." Math. Sci., 14(153), 5-9.

[2] Aoki, M. and A. Havenner (1991) "State space modeling of multiple time series." Econometric Reviews, 10(1), 1Ů59.

[3] Besse, P.C., H. Cardot and D. B. Stephenson (2000) "Autoregressive Forecasting of Some Functional Climatic Variations." Scandinavian Journal of Statistics, 27(4), 673-687.

[4] Brockwell, J. and R. Davis (2009) "Time Series: Theory and Methods." Springer, New York, NY.

[5] Cao, R., J.D. Hart and A. Saavedra (2003) "Nonparametric maximum likelihood estimators for AR and MA time series." Journal of Statistical Computation and Simulation, 73(5), 347-360.

[6] Diebold, F.X. and C. Li (2006) "Forecasting the term structure of government bond yields." Journal of Econometrics, 130(2), 337-364.

[7] Fan, J. and I. Gijbels (1996) "Local Polynomial Modelling and its Applications." Chapman and Hall, London.

[8] Gilbert, P.D. (1993) "State Space and ARMA Models: An Overview of the Equivalence." Working Paper 93-4, Bank of Canada.

[9] Hamilton, J. D. (1994) "Time Series Analysis." Princeton University Press, Princeton, NJ.

[10] Kargin, V. and A. Onatski (2008) "Curve forecasting by functional autoregression." Journal of Multivariate Analysis, 99, 2508-2526.

[11] Ramsay, J.0. and B.W. Silverman (2005) "Functional Data Analysis." Springer Series in Statistics. New York, NY.

[12] Rice, J. A. and B. W. Silverman (1991) "Estimating the mean and covariance structure nonparametrically when the data are curves." Journal of Royal Statistical Society Series B, 53, 233-243.

[13] Weron1 Weron R. and A. Misiorek (2008) "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models" International Journal of Forecasting, 24, 744-763.

[14] Yao, F., H. G. Müller, A. J. Clifford, S. R. Dueker, J. Follett, Y. Lin, B. Buchholz and J. S. Vogel (2003) "Shrinkage estimation for functional principal component scores, with application to the population kinetics of plasma folate." Biometrics, 59, 676-685.

[15] Yao, F., H. G. Müller and J. L. Wang (2005) "Functional data analysis for sparse longitudinal data." Jornal of American Statistical Association, 100, 577-590.
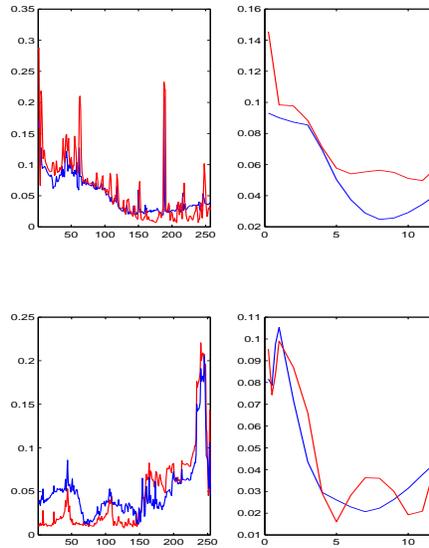
Figure 1: Distance between observed functions and predicted functions using VARMA modeling for Euribor data. The blue lines are predictions using principal components. The red lines are for predictions using Diebold and Li (2006) method. Left panel: Time series of RMSE plotted over days. Right panel: RMSE plotted over maturity. Top: 1999 order of ARMA is chosen to be (3,0), bottom: 2007 order of ARMA is chosen to be (1,0).

Week	AR	FDA	IHMAR
WMAE	0.1325	0.1203	0.1581
♯ best	11	27	0
m.d.f.b	0.5909	0.1280	1.5643

Table 1: WMAE errors for Electricity data. Measures of fit include the mean WMAE over all weeks, the number of times a given model was best and the mean deviation from the best model in each week (m.d.f.b.).