Modeling distributional time series by transformations

Zhicheng Chen* (October 2018)

Abstract

Probability distributions play a very important role in many applications. This paper describes a modeling approach for distributional time series. Probability density functions (PDFs) are approximated by real-valued vectors via successively applying the log-quantile-density (LQD) transformation and functional principal component analysis (FPCA); state-space models (SSMs) for real-valued time series are then applied to model the evolution of PCA scores, corresponding results are mapped back to the PDF space by the inverse LQD transformation.

Keywords: distribution; time series; state-space model

1. Introduction

Functional time series modeling is receiving increased interest in statistics, finance and other related fields. Unlike ordinary functions, probability density functions (PDFs) of distributions need to satisfy two important constrains: (1) non-negative; (2) unit integral [1]. For PDF-valued functional data, directly applying a statistical modeling method for general functions may encounter problems in insuring corresponding results automatically satisfy the two constrains of PDFs. Inspired by the seminal work in [1], this article describes an approach for modeling distributional time series using state-space models (SSMs) on the basis of converting PDFs to ordinary functions. Another related work is [2], where a SSM-based method was developed for forecasting ordinary functional data.

2. Modeling method

Consider a time series formed by probability density functions (of a stochastic distribution process), denoted by $\{f_t(x)\}_{t=1}^{n}$, which is termed the distributional time series in this study. Without loss of generality, all PDFs are assumed to be finitely supported on $[0,1]$; PDFs with general finite support can be easily tackled by scale transformation.

To release the constraints of PDFs, the newly proposed log-quantile-density (LQD) transformation [1] is employed to covert PDFs to ordinary functions, i.e.

$$\psi_t(s) = \log\big(q_t(s)\big) = -\log\big\{f_t\big(Q_t(s)\big)\big\},\ \ s \in [0,1] \tag{1}$$

where $Q_t(s)$ and $q_t(s)$ are, respectively, the quantile function and quantile density function corresponding to the PDF $f_t(x)$; $\psi_t(s) \in L^2[0,1]$ is an ordinary function, which can be converted back to the PDF form by the inverse LQD transformation [1] as follows:

$$f_t(x) = \theta_{\psi_t}\exp\big\{-\psi_t\big(F_t(x)\big)\big\},\ \ F_t^{-1}(s) = \theta_{\psi_t}^{-1}\int_0^s e^{\psi_t(\tau)}d\tau \tag{2}$$

---

* Doctor, Email: "13B933002@hit.edu.cn" or "hitchenzhicheng@163.com"

where $\theta_{\psi_t} = \int_0^1 e^{\psi_t(\tau)}\, d\tau$.

The functional time series $\{\psi_t(s)\}_{t=1}^n$ are further converted to real-valued time series by applying the principal component analysis (FPCA) [3]. FPCA is an effective dimension reduction method for functional data, the main point is approximating the continuous function by the truncated Karhunen–Loève representation [3], i.e.

$$\psi_t(s) \approx \mu_\psi(s) + \sum_{j=1}^{m} \xi_{t,j}\varphi_j(s)\ ,\ \ t = 1,2,\cdots,n \tag{3}$$

where $\mu_\psi(s)$ is the mean function of the functional dataset $\{\psi_t(s)\}_{t=1}^n$; φ_j are the eigenfunctions of the covariance operator estimated from the functional dataset $\{\psi_t(s)\}_{t=1}^n$; $\xi_{t,j}$ are PCA scores; m is the truncation order. By using Eqs. (1) and (3), the distributional time series can be converted to m real-valued time series denoted by $\{\xi_{t,j}\}_{t=1}^n$, $j = 1,2,\cdots,m$.

The state-space models (SSMs) [4,5] for the real-valued time series $\{\xi_{t,j}\}_{t=1}^n (j = 1,2,\cdots,m)$ can be generally defined as

$$\begin{cases} \xi_{t,j} = h_{t,j}(\alpha_{t,j}) + \varepsilon_j,\ \ \varepsilon_j \sim N\left(0, \sigma_{\varepsilon_j}^2\right) \\ \alpha_{t+1,j} = u_{t,j}(\alpha_{t,j}) + \eta_j,\ \ \eta_j \sim N\left(0, \sigma_{\eta_j}^2\right) \end{cases},\quad j = 1,2,\cdots,m \tag{4}$$

where $u_{t,j}$ are differentiable functions for characterizing the dynamic behavior of the latent state process (represented by $\{\alpha_{t,j}\}_{t=1}^n$); $h_{t,j}$ are differentiable functions for describing the observation process given the states; ε_j and η_j represent noise effects. As a special case, the linear SSMs [4,5] for the real-valued time series are rewritten as follows:

$$\begin{cases} \xi_{t,j} = Z_{t,j}\alpha_{t,j} + \varepsilon_j,\ \ \varepsilon_j \sim N\left(0, \sigma_{\varepsilon_j}^2\right) \\ \alpha_{t+1,j} = T_{t,j}\alpha_{t,j} + \eta_j,\ \ \eta_j \sim N\left(0, \sigma_{\eta_j}^2\right) \end{cases},\quad j = 1,2,\cdots,m \tag{5}$$

where $Z_{t,j}$ and $T_{t,j}$ are real matrixes or real numbers (only for univariate case).

On the basis of the SSMs given in Eqs. (4) and (5), filtering, smoothing and forecasting can be realized for distributional time series in the Kalman filtering (KF) setting (including its expanded version in considering nonlinearities, such as EKF, UKF, etc.) [4-6]. In addition, particle filtering algorithms [7] can also be employed to solve the inference problems for the above SSMs.

3. Simulation study

The following random walk model is used to simulate the distributional times series for investigation:

$$\begin{cases} a_t = a_{t-1} + \varepsilon_a,\ \varepsilon_a \sim N(0, \sigma_a^2)\ \text{and}\ a_0 = 14,\ \ \sigma_a = 0.6 \\ b_t = b_{t-1} + \varepsilon_b,\ \varepsilon_b \sim N(0, \sigma_b^2)\ \text{and}\ b_0 = 12,\ \ \sigma_b = 0.8 \\ f_t(x) = 0.5 g_t(x) + 0.5\ \ \text{where}\ \ g_t \sim \text{Beta}(a_t, b_t) \end{cases} \tag{6}$$

Note that parameters in the Beta distribution should be positive; if any element in series $\{a_t\}_{t=1}^n$ and $\{b_t\}_{t=1}^n$ takes negative value, the corresponding series is rejected and re-simulated. Given a_t and b_t, 300 random samples from each Beta distribution $\text{Beta}(a_t, b_t)$ are first generated, then

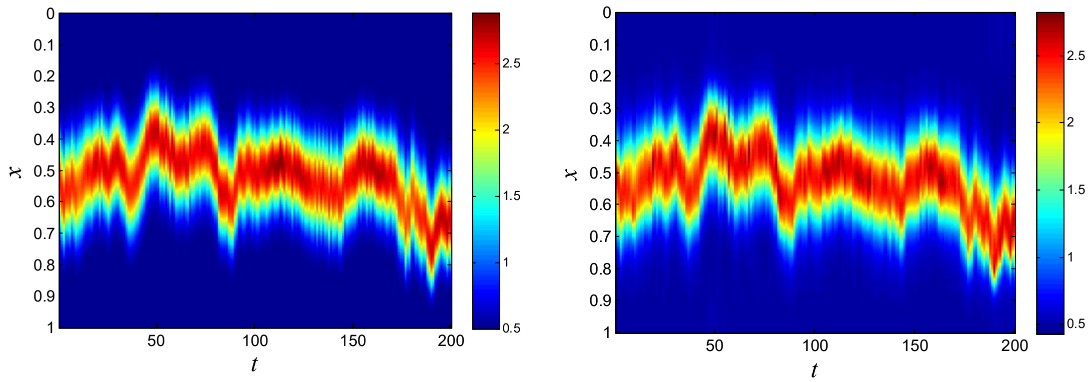
$g_t(x)$ is estimated by using the kernel density estimator, thus resulting in the observed distributional time series denoted as $\{\hat{f}_t(x)\}_{t=1}^{n}$ with $\hat{f}_t(x)$=$\hat{g}_t(x) + 0.5$ ($\hat{g}_t(x)$ is the kernel density estimate of $g_t(x)$).

The linear SSM given in Eq. (5) is employed to model the corresponding real-valued time series $\{\xi_{t,j}\}_{t=1}^{n}(j = 1,2,\cdots,m)$ by setting $Z_{t,j} = \xi_{t-1,j}$, $T_{t,j} = 1$ and $m = 10$. The remaining undetermined parameters σ_{ε_j} and σ_{η_j} are estimated by the likelihood method (see [4,5] for detailed descriptions).

For visual comparison, PDFs are organized as a matrix form, i.e.

$$\mathbf{A} = \begin{bmatrix} f_1(x_1) \cdots f_t(x_1) \cdots f_n(x_1) \\ f_1(x_2) \cdots f_t(x_2) \cdots f_n(x_2) \\ \vdots \qquad \vdots \qquad \vdots \\ f_1(x_p) \cdots f_t(x_p) \cdots f_n(x_p) \end{bmatrix} \tag{7}$$

where $0 \le x_1 < x_2 < \cdots < x_p \le 1$ are evenly spaced points within $[0,1]$. Using the graphical representation of matrix $\mathbf{A}$, Figure 1 illustrates the result comparison between the true PDF-series and the estimated PDF-series obtained by using the Kalman smoothing algorithm on the basis of the dynamic model given in Eq. (5). It can be seen from Figure 1 that the model-based estimates of PDFs agree well with true PDFs.



(a) (b)

Figure 1. Results comparison for PDF-series in terms of the image of matrix **A** defined in Eq. (7). (a) True PDFs. (b) Estimated PDFs by using the Kalman smoothing algorithm based on the model given in Eq. (5).

4. Conclusions

A dynamic modeling approach is introduced for distributional time series. By virtue of the log-quantile-density transformation and functional principal component analysis, distributional time series are modeled in the state-space modeling framework.

References

[1] Petersen A and Müller HG. Functional data analysis for density functions by transformation to a Hilbert space. The Annals of Statistics 2016; 44(1): 183-218.

[2] Nagbe K, Cugliari J and Jacques J. Short-term electricity demand forecasting using a functional state space model. Energies 2018; 11(5): 1120.

[3] Ramsay JO and Silverman BW. Functional data analysis (Second Edition). New York: Springer, 2005.

[4] Shumway RH and Stoffer DS. Time series analysis and its applications (Third Edition). New York: Springer, 2011.

[5] Durbin J and Koopman SJ. Time series analysis by state space methods (Second Edition). Oxford: Oxford University Press, 2012.

[6] Haykin S. Kalman filtering and neural networks. New York: John Wiley & Sons, 2001.

[7] Arulampalam MS, Maskell S, Gordon N and Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on signal processing 2002; 50(2): 174-188.