Granger Causality Analysis for Functional Time Series Data

Shubhangi Sikaria[1] and Rituparna Sen[2]

[1]Indian Institute of Technology Madras, Chennai, India
[2]Indian Statistical Institute, Bangalore, India

Abstract

This paper proposes a two-level hierarchical Gaussian process that is dynamically modeled as a multivariate functional autoregressive model (MFAR). Through MFAR, we capture the cross-correlation amongst the functional time series. We also investigate the causal connection between functional time series through the application of Granger causality. The performance of this model is illustrated through a series of simulation studies and applications to yield curve prediction of various countries. Our model provides improved forecast accuracy in most cases.
Keywords— Multivariate functional time series, dynamic linear model, Granger causality

1 Introduction

The collection of more complex data with function features encourages the study of multivariate functional data. Existing approaches ([6], [4], [1], [2]) rely on the functional principal component analysis (FPCA) to treat multivariate functional data. However, FPCA is designed for dense grids without calculation errors, leading to inadmissible predictions with large measurement errors in the case of a sparse setting. We later prove that the observable process with measurement error produces inefficient estimates even in a dense environment as the process is no longer MFAR(p). We propose a methodology that overcomes these concerns.

Kowal et al. [5] applied a functional autoregressive model (FAR) on the univariate functional time series. Working on similar lines, we also proposed a hierarchical model for multivariate functional time series. In multivariate cases, we seek to study the correlation among the data series. For computation, we first represent the hierarchical model as the dynamic linear model (DLM). We use observation equations to discretize the functional data and obtain measurement error, and evolution equations take into account of the correlation among the functional time series and define the process model. The latent process in our case is the multivariate functional autoregressive model of order p, written as MFAR(p), instead of FAR(p). A dynamic functional factor model further specifies the dynamic innovation process.

A question of great interest is whether there exists a directionality of information. Several works on multivariate time series based on vector autoregressive model (VAR) have considered this problem. We will be the first to address such a problem in the case of multivariate functional time series. Basically, we are interested in analyzing the causal relations among a set of variables. As discussed by Granger [3], the first time series X is said to have a causal influence on the second time series Y if the prediction of second series Y could be improved by the inclusion of past measurements from the first time series X. Later, this concept was extended to various nonlinear and time-varying cases. In this paper, we will investigate the Granger causality among the bond yields of different countries.

2 Multivariate Functional Autoregression Model

Suppose we have K functional time series defined as $\{Y_t^n\}$, $n = 1, \cdots, K$ on some compact index set $\mathscr{T} \subset \mathrm{R}^D$, typically $D = 1$. Consider $\{\boldsymbol{Y}_t\}$ as $K-$variate functional time series represented as: $\boldsymbol{Y}_t = [Y_t^1, \cdots, Y_t^K]'$ where $Y_t^n \in \mathrm{L}^2(T)$ for $n = 1, 2, \cdots, K$.

Multivariate functional autoregression model of order p, written as MFAR(p):

$$\boldsymbol{Y}_t - \boldsymbol{\mu} = \sum_{l=1}^{p} \boldsymbol{\Psi}_l(\boldsymbol{Y}_{t-l} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t. \tag{1}$$

where $\boldsymbol{\mu} = [\mu^1, \mu^2, \cdots, \mu^K]'$ is mean of $\boldsymbol{Y}_t$ under stationarity, $\boldsymbol{\varepsilon}_t = [\varepsilon_t^1, \varepsilon_t^2, \cdots, \varepsilon_t^K]'$ is a vector of innovation functions with each ε_t^k is a sequence of independent mean zero random function with $\mathrm{E}||\varepsilon_t^n||^2 < \infty$ for $n = 1, \cdots, K$ and $\boldsymbol{\Psi}_l = [\Psi_l^{i,j}]_{i,j=1}^K$ where each $\Psi_l^{i,j}$ is a bounded linear operator on $\mathrm{L}^2(\mathscr{T})$ for $i, j = 1, \cdots, K$ and $l = 1, 2, \cdots, p$.

We assume that each functions $Y_t^n(\tau^n)$, $n = 1, \cdots, K$ are observed at unequally-spaced points $\tau^n \in \mathscr{T}$. We restrict to $p = 1$ and consider only integral operators for better interpretability and computational convenience. In practice, the functional observations $\{Y_t\}$ are observed via discrete samples of each curve with some measurement error. Suppose we observe $y_{i,t}^n \in \mathrm{R}$ sampled with $v_{i,t}^n$ from $\{Y_t^n\} \in \mathrm{L}^2(\mathscr{T})$, $n = 1, \cdots, K$:

$$y_{i,t}^n = Y_t^n(\tau_{i,t}^n) + v_{i,t}^n \tag{2}$$

where $\tau_{i,t}^n$ for $i = 1, \cdots, m_t^n$ are observation points of Y_t^n and $v_{i,t}^n$ is a mean zero measurement error with finite variance. Assume $\boldsymbol{\alpha}_t = \boldsymbol{Y}_t - \boldsymbol{\mu}$, then two level hierarchical model is given as:

$$\begin{aligned} y_{i,t}^n &= \mu^n(\tau_{i,t}^n) + \alpha_t^n(\tau_{i,t}^n) + v_{i,t}^n, & i &= 1, \cdots, m_t^n, \\ \alpha_t^n &= \sum_{m=1}^{K} \int \psi^{n,m}(\tau^n, u)\alpha_{t-1}^n(u)du + \varepsilon_t^n(\tau^n), & &\forall \tau \in \mathscr{T} \text{ and } n = 1, \cdots, K, \end{aligned} \tag{3}$$

for $t = 2, \cdots, T$, where $\psi^{n,m}$ are integral operators and we assume that $\{\boldsymbol{v}_{i,t}\}$ and $\{\boldsymbol{\varepsilon}_t\}$ are mutually independent.

The presence of measurement error in model (3) increase the estimation error of $\boldsymbol{\Psi}$ and produce inefficient forecasts. The following proposition illustrates that MFAR(p) model for the observables is inappropriate.

Proposition 1: Let $\boldsymbol{Y}_t - \boldsymbol{\mu} = \sum_{l=1}^{p} \boldsymbol{\Psi}(\boldsymbol{Y}_{t-l} + \boldsymbol{\varepsilon}_t)$ and suppose we observe $\boldsymbol{y}_t = \boldsymbol{Y}_t + \boldsymbol{v}_t$ are independent white noise processes. Then the observable process $\{\boldsymbol{y}_t\}$ follows a multivariate functional autoregressive moving average (FARMA) process of order (p, p).

To overcome this model misspecification, we separate the observed data into functional process and measurement errors in the dynamic linear model (DLM).

2.1 Dynamic Linear Models for MFAR(p)

For practical implementation of model (3), we select a finite set of evaluation points for each of functional time series, $\mathscr{T}_e^n = \{\tau_1^n, \cdots, \tau_M^n\} \subset \mathscr{T}$ for $n = 1, \cdots, K$. We approximate the integral in (3) using quadrature methods with $Q^{n,m}$ for $n, m = 1, \cdots, K$, as a known quadrature weight matrix. Let Z_t^n be $m_t^n \times M$, $n = 1, \cdots, K$, incidence matrix that identifies the observation points observed at time t for n-th functional times series. We can formulate the dynamic linear model (DLM) corresponding to the hierarchical model (3) as:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{Z}_t\boldsymbol{\mu} + \boldsymbol{Z}_t\boldsymbol{\alpha}_t + \boldsymbol{v}_t, & [\boldsymbol{v}_t|\Sigma_v] &\overset{indep}{\sim} N(\boldsymbol{0}, \Sigma_v) \text{ for } t = 1, \cdots, T, \\ \boldsymbol{\alpha}_t &= \boldsymbol{\psi}\boldsymbol{Q}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\varepsilon}_t, & [\boldsymbol{\varepsilon}_t|\boldsymbol{K}_\varepsilon] &\overset{indep}{\sim} N(\boldsymbol{0}, \boldsymbol{K}_\varepsilon) \text{ for } t = 2, \cdots, T, \\ \boldsymbol{\alpha}_1 &\sim N(\boldsymbol{0}, \boldsymbol{K}_\varepsilon) \end{aligned} \tag{4}$$

where $\boldsymbol{y}_t = [y_{i,t}^n]$, is discrete sample of K-variate functional time series $\{\boldsymbol{Y}_t\}$ for $i = 1, \cdots, m_t^n$ and $n = 1, \cdots, K$. Here, $\boldsymbol{\mu} = [\mu^1(\tau^1), \cdots, \mu^K(\tau^K)]'$ and $\boldsymbol{\alpha}_t = [\alpha_t^1(\tau^1), \cdots, \alpha_t^K(\tau^K)]'$ where $\mu^i(\tau^i) = [\mu^i(\tau_1^i), \cdots, \mu^i(\tau_M^i)]$ and $\alpha_t^i(\tau^i) = [\alpha_t^i(\tau_1^i), \cdots, \alpha_t^i(\tau_M^i)]$ for $i = 1, \cdots, K$. Other matrices are defined as:

$$\boldsymbol{Z}_t = \begin{bmatrix} Z_t^1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & Z_t^K \end{bmatrix}, \boldsymbol{\psi} = \begin{bmatrix} \{\psi^{1,1}(\tau_i, \tau_j)\}_{i,j=1}^M & \cdots & \{\psi^{1,K}(\tau_i, \tau_j)\}_{i,j=1}^M \\ \vdots & \ddots & \vdots \\ \{\psi^{1,K}(\tau_i, \tau_j)\}_{i,j=1}^M & \cdots & \{\psi^{K,1}(\tau_i, \tau_j)\}_{i,j=1}^M \end{bmatrix} \text{ and } \boldsymbol{Q} = \begin{bmatrix} Q^{1,1} & \cdots & Q^{1,K} \\ \vdots & \ddots & \vdots \\ Q^{1,K} & \cdots & Q^{K,1} \end{bmatrix}.$$

2.2 Forecasting MFAR(p)

For $n = 1, \cdots, K$ we firstly initialize μ^n as smooth mean of $\{y_t^n\}_{t=1}^T$ and fit a spline to each $Y_t^n - \mu^n$ to estimate α_t^n. Using the estimates of μ^n and α_t^n we estimate the measurement error variance-covariance matrix $\Sigma_{\boldsymbol{v}}$ and also initialize the FAR kernels $\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_p$ as described in Kowal et al. [5]. We apply functional dynamic linear model (FDLM) to compute the implied covariance matrix for $\boldsymbol{\varepsilon}_t$. Later we propose the Gibbs sampling algorithm for model (4), which is an extension of Gibbs sampling algorithm for univariate functional autoregression model of order p [5].

2.3 Theoretical Results

Let $\mathscr{F}_t$ denotes the information set available at time t. Thus, $\mathscr{F}_t = \{\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \cdots, \boldsymbol{y}_1\} \cup \mathscr{F}_0$, where $\mathscr{F}_0$ is the information prior at $t = 1$.

Theorem 1 *For any finite set of evaluation points $\mathscr{T}_e \subset \mathscr{T}$, the unique best linear predictor of the conditional random matrix $\boldsymbol{\delta} \sim [\boldsymbol{\delta}|\boldsymbol{Y}, \boldsymbol{\Theta}]$, where $\boldsymbol{\delta}, \boldsymbol{Y} \subset \mathscr{F}_T \cup \{\boldsymbol{\alpha}_t(\tau) : \tau \in \mathscr{T}_e, t = 1, \cdots, T\}$ and $\Theta = \{\boldsymbol{\mu}, \Sigma_v, \boldsymbol{\psi}, \boldsymbol{K}_\varepsilon\}$, under the risk $\mathscr{R}_e$ and conditional on model (3) with integral approximation, is the conditional expectation $\hat{\boldsymbol{\delta}}(\boldsymbol{Y}|\boldsymbol{\Theta}) = \mathrm{E}[\boldsymbol{\delta}|\boldsymbol{Y}, \boldsymbol{\Theta}]$ as computed in model (4).*

Corollary 1.1 *The unique best linear predictor of $[\boldsymbol{\alpha}_t(\tau^*)|\mathscr{F}_s]$ for any time t, s and any point $\tau^* \in \mathscr{T}$ is the corresponding expectation under model (4).*

3 Granger Causality in Functional Autoregression Model

Here, we investigate the causal relationship between two functional time series data Y_t and X_t, $t \in T$ observed on unequally spaced points $\tau \in \mathscr{T}$ where $\mathscr{T} \subset \mathbf{R}$ is a compact index set. Firstly, we model Y_t as functional autoregressive model of order 1, written as FAR(1), and partition the linear projection of Y_t on Y_{t-1} and X_{t-1} to account for the Granger causality as

$$Y_t - \mu_Y = \Psi_Y(Y_{t-1} - \mu_Y) + \Psi_X(X_{t-1} - \mu_X) + \varepsilon_t \tag{5}$$

In second model, we model Y_t as linear projection of Y_{t-1} only as

$$Y_t - \mu_Y = \Psi(Y_{t-1} - \mu_Y) + \varepsilon_t' \tag{6}$$

where Y_t and $X_t \in \mathrm{L}^2(\mathscr{T})$, Ψ_Y, Ψ_X and Ψ are bounded linear operator on $\mathrm{L}^2(\mathscr{T})$, μ_Y and μ_X are mean of $\{Y_t\}$ and $\{X_t\}$ respectively and ε_t and $\varepsilon_t' \in \mathrm{L}^2(\mathscr{T})$ are zero mean independent random functions.

We perform the sampling for both the models and estimate the one-step forecasts $[y_{T+h}|y_{1:(T+h-1)}], h = 1, \cdots, R$ for both the models. Let $\hat{Y}_{T+h}$ and $\hat{Y}^*_{T+h}$ be the function of one-step forecast for first model (5) and second model (6) respectively for $h = 1, \cdots, R$. We compute the sum of square errors for both models $SSE_1 = \sum_{h=1}^R ||Y_{T+h} - \hat{Y}_{T+h}||^2$ and $SSE_2 = \sum_{h=1}^R ||Y_{T+h} - \hat{Y}^*_{T+h}||^2$ where $Y_{T+h} = (Y_{T+h}(\tau_1), \cdots, Y_{T+h}(\tau_M)))'$ are the observed values used to measure the one-step forecasting performance at the evaluation points. To investigate the effect of Granger causality of X_t on Y_t, we performed permutation test using the statistic called as F_{obs}-statistic and defined as:

$$F_{obs} = \frac{(SSE_2 - SSE_1)/(k_1 - k_2 + 1)}{SSE_1/(R - k_1 - 1)}. \tag{7}$$

where k_1 and k_2 be the number of parameters in the first model (5) and second model (6) respectively.

4 Simulations

We perform one-step and five-step forecasting and see how the associated performance varies with the sample size T and cross correlation between the series.We are particularly interested in the causal relationship between two functional time series. We will consider two cases based on model (5) and model (6) with varying sample size T. Functional time series $\{Y_t\}$ generated from MFAR(1) kernels is correlated to functional time series $\{X_t\}$ while $\{Y_t\}$ generated from FAR(1)

	h=1		
Data Generation	RMSFE (MFAR)	RMSFE (FAR)	F_{obs}
$\{Y_t\}$ is dependent on $\{X_t\}$	0.009160474	0.009312481	0.2572507
$\{Y_t\}$ is independent of $\{X_t\}$	0.009157562	0.00964527	0.4351136

Table 1: Simulation Results showing one-step ahead RMSFEs and F_{obs} statistic

	h=5		
Data Generation	RMSFE (MFAR)	RMSFE (FAR)	F_{obs}
$\{Y_t\}$ is dependent on $\{X_t\}$	0.01064169	0.0110255	0.1063456
$\{Y_t\}$ is independent of $\{X_t\}$	0.01069097	0.008994967	-0.4220476

Table 2: Simulation Results showing five-step ahead RMSFEs and F_{obs} statistic

kernel is independent of $\{X_t\}$. We use smooth Gaussian process for innovation process ε_t. In our experiment, we consider dense sampling design using $m_t = 25$ equally-spaced observation points on $[0,1]$ for all t, however the model works well with sparse sampling design also. Data sets generated using FAR(1) kernel and MFAR(1) kernels are applied to proposed model (MFAR) as well as to univariate model proposed by Kowal et. al. (FAR) [5]. Table 1 and Table 2 shows the one- and five-step root mean squared forecasting errors (RMSFEs) along with the F-statistics respectively. When $\{Y_t\}$ is dependent on $\{X_t\}$, we obtain lower root mean squared forecasting errors for MFAR model compared to FAR model. Also, we obtain positive F_{obs} for one-step and five-step forecasting which further illustrates that MFAR model perform better in presence of correlation among the functional time series.

5 An Empirical Illustration: Yield Curves

To assess the performance of the proposed model, we conducted an extensive forecasting study using yield curves of two countries: USA and UK. For our study we consider 5 month data starting from Jan 2019 till May 2019. Estimates of the yield curves are provided for maturities $\mathscr{T} = \{90, 180, 360, 720, 1080, 2520, 3600, 10800\}$ days. First four month are used for estimation to produce one-step ahead and five-step(one business week) ahead forecasts of next month. The one- and five-step root mean squared forecasting errors (RMSFEs) are calculated using the forecasted values and data for the fifth month. We compare the RMSFE values for USA and UK yield curves using MFAR(1) model and separately applied FAR(1) model to both the yield curves, presented in Table 3. Clearly, proposed model performs better than the univariate case as multivariate model accounts for cross correlation among the yield curves of various countries. To investigate whether UK yield curve Granger cause USA yield curve, we apply the permutation test described in Section 3. The p-value we obtain from the permutation test is 0.4333 implying there are not statistically significant result to say that UK yield curve Granger cause USA yield curve.

	h=1				h=5		
Country	MFAR(1)	FAR(1)	F_{obs}	Country	MFAR(1)	FAR(1)	F_{obs}
USA	0.06716552	0.1259286	9.740322	USA	0.1579645	0.1463301	-0.8707428
UK	0.05066788	0.05098459	0.6516207	UK	0.07166108	0.07252083	0.5886336

Table 3: h-step ahead RMSFEs

6 Conclusion

This paper introduces the two-level hierarchical Gaussian process for modeling multivariate functional data. Our model considers not only auto-correlations in the functional responses, but also considers cross-correlations among the series. Simulation analysis indicates that the proposed methodology for estimating MFAR(p) produces improve prediction. We further study the causal relationship among the functional responses, and the positive F-statistics in the simulation analysis strongly indicate that such causal relationship exists in correlated series. In the yield curve application, we obtained a better forecast using MFAR(p) model.

References

[1] José R Berrendero, Ana Justel, and Marcela Svarc. "Principal components for multivariate functional data". In: *Computational Statistics & Data Analysis* 55.9 (2011), pp. 2619–2634.

[2] Jeng-Min Chiou, Yu-Ting Chen, and Ya-Fang Yang. "Multivariate functional principal component analysis: A normalization approach". In: *Statistica Sinica* (2014), pp. 1571–1596.

[3] Clive WJ Granger. "Investigating causal relations by econometric models and cross-spectral methods". In: *Econometrica: journal of the Econometric Society* (1969), pp. 424–438.

[4] Julien Jacques and Cristian Preda. "Model-based clustering for multivariate functional data". In: *Computational Statistics & Data Analysis* 71 (2014), pp. 92–106.

[5] Daniel R Kowal, David S Matteson, and David Ruppert. "Functional autoregression for sparsely sampled data". In: *Journal of Business & Economic Statistics* 37.1 (2019), pp. 97–109.

[6] JO Ramsay and BW Silverman. "Principal components analysis for functional data". In: *Functional data analysis* (2005), pp. 147–172.