Synthetic prediction methods for minimizing post shock forecasting error

Daniel J. Eck and Jilei Lin and Dave Zhao

May 2019

Abstract

We seek to develop a forecasting methodology for time series data that is thought to have undergone a shock which has origins that have not been previously observed. We still can provide credible forecasts for a time series in the presence of such systematic shocks by drawing from disparate time series that have undergone similar shocks for which post-shock outcome data is recorded. These disparate time series are assumed to have mechanistic similarities to the time series under study but are otherwise independent. The inferential goal of our forecasting methodology is to supplement observed time series data with post-shock data from the disparate time series in order to minimize average forecast risk.

1 Setting

We will suppose that a researcher has time series data $(y_{i,t}, \mathbf{x}_{i,t})$, for $t = 1, \ldots, T_i$ and $i = 1, \ldots, n+1$, where $y_{i,t}$ is a scalar response and $\mathbf{x}_{i,t}$ is a vector of covariates that are revealed to the analyst prior to the observation of $y_{1,t}$. Suppose that the analyst is interested in forecasting $y_{1,t}$, the first time series in the collection. We will suppose that specific interest is in forecasting the response after the occurrence of a structural shock. To gauge the performance of forecasts, we consider forecast risk in the form of MSE,

$$R_T = \frac{1}{T}\sum_{t=1}^{T} \mathrm{E}(\hat{y}_{1,t} - y_{1,t})^2,$$

and root mean squared error (RMSE), given by $\sqrt{R_T}$, in our analyses. In this article, we focus on post-shock prediction where forecasts methods only differ at the next future time point. Thus the MSE reduces to the magnitude $\mathrm{E}(\hat{y}_{1,t} - y_{1,t})^2$.

1.1 Model Setup

In this section, we will describe the assumed dynamic panel models for which post-shock aggregated estimators are provided. The basic structures of these models are the same for all time-series in the analysis, the differences between them lie in the setup of the shock effect distribution.

Let $I(\cdot)$ be an indicator function, T_i be the time length of the time series i for $i = 1, \ldots, n+1$, and T_i^* be the time point just before the one when the shock is known to occur, with $T_i^* < T_i$. For $t = 1, \ldots, T_i$ and $i = 1, \ldots, n+1$, the model $\mathcal{M}_1$ is defined as

$$\mathcal{M}_1 \colon y_{i,t} = \eta_i + \alpha_i D_{i,t} + \phi_i y_{i,t-1} + \theta_i' \mathbf{x}_{i,t} + \varepsilon_{i,t} \tag{1}$$

where $D_{i,t} = I(t = T_i^* + 1)$ and $\mathbf{x}_{i,t} \in \mathbb{R}^p$ with $p \geq 1$. We assume that the $\mathbf{x}_{i,t}$'s are fixed. Let $|x|$ denote the absolute value of x for $x \in \mathbb{R}$. For $i = 1, \ldots, n+1$ and $t = 1, \ldots, T_i$, the random effects structure for $\mathcal{M}_1$ is:

$$\eta_i \overset{iid}{\sim} \mathcal{F}_\eta$$

$$
\begin{aligned}
\phi_i &\overset{iid}{\sim} \mathcal{F}_\phi \text{ where } |\mathcal{F}_\phi| < 1,\\
\theta_i &\overset{iid}{\sim} \mathcal{F}_\theta\\
\alpha_i &\overset{iid}{\sim} \mathcal{F}_\alpha\\
\varepsilon_{i,t} &\overset{iid}{\sim} \mathcal{F}_\varepsilon\\
\eta_i &\perp\!\!\!\perp \alpha_i \perp\!\!\!\perp \phi_i \perp\!\!\!\perp \theta_i \perp\!\!\!\perp \varepsilon_{i,t}.
\end{aligned}
$$

2 What comes next

We want to consider decision rules that are nonparametric in nature. For example, suppose that we posit a classical linear regression model for (1) (then $\varepsilon_{i,t} \sim N(0, \sigma_i^2)$).

1. We can estimate all α's using OLS, and provided that T_i is large enough, then $\hat{\alpha}_i|\alpha_i \approx N(0, \sigma_i^2)$. We could then consider estimating the distribution function of the α's using the noisy observations $\hat{\alpha}_i$, $i = 2, \ldots, n$. Dave suggested that we may consider deconvolution in order to handle the fact that $\alpha = \hat{\alpha} + \varepsilon$, where $\varepsilon \approx N(0, \sigma_i^2)$ and σ_i^2 is either known or estimated very well. Then we can use a two step approach to first estimate the distribution $\mathcal{F}_\alpha$ using a nonparametric MLE approach and then second obtain quantiles from this estimated distribution. Perhaps a one step approach exists, but we didn't think of one in the meeting (development of such an approach is advantageous and should be thought about).

2. We also discussed extreme-value distributions to handle tail events that are unlikely but potentially disastrous (think "COVID before COVID"). There are approaches that exist which "tack on" an extreme-value distribution to the tail of some data-generating process where the bulk of the data is fit by some other methodology.

3. We could also consider frameworks which allow for outcome models to be selected, we do not necessary need a linear regression model but we do need to have the ability to estimate a "shock effect."

References