Journal of Statistical Planning and Inference 138 (2008) 3927 – 3938

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

On robust forecasting in dynamic vector time series models

Christian Gagné[a], Pierre Duchesne[b,*]

[a] Statistics Canada
[b] Département de mathématiques et de statistique, Université de Montréal, C.P. 6128 Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7

ARTICLE INFO

Article history:
Received 31 January 2006
Received in revised form
12 November 2007
Accepted 17 February 2008
Available online 28 March 2008

MSC:
primary 62M20
secondary 62M10

Keywords:
Multivariate time series
Exogenous variables
Prediction
Robust estimators
Additive outliers

ABSTRACT

In this article, robust estimation and prediction in multivariate autoregressive models with exogenous variables (VARX) are considered. The conditional least squares (CLS) estimators are known to be non-robust when outliers occur. To obtain robust estimators, the method introduced in Duchesne [2005. Robust and powerful serial correlation tests with new robust estimates in ARX models. J. Time Ser. Anal. 26, 49–81] and Bou Hamad and Duchesne [2005. On robust diagnostics at individual lags using RA-ARX estimators. In: Duchesne, P., Rémillard, B. (Eds.), Statistical Modeling and Analysis for Complex Data Problems. Springer, New York] is generalized for VARX models. The asymptotic distribution of the new estimators is studied and from this is obtained in particular the asymptotic covariance matrix of the robust estimators. Classical conditional prediction intervals normally rely on estimators such as the usual non-robust CLS estimators. In the presence of outliers, such as additive outliers, these classical predictions can be severely biased. More generally, the occurrence of outliers may invalidate the usual conditional prediction intervals. Consequently, the new robust methodology is used to develop robust conditional prediction intervals which take into account parameter estimation uncertainty. In a simulation study, we investigate the finite sample properties of the robust prediction intervals under several scenarios for the occurrence of the outliers, and the new intervals are compared to non-robust intervals based on classical CLS estimators.

1. Introduction

Dynamic simultaneous equation models are frequently used for forecasting purposes. The statistical properties of these linear systems are discussed in Judge et al. (1985), Hannan and Deistler (1988) and Lütkepohl (1993), amongst others. A possible process encountered in economic and physical applications is the vector autoregressive model with stationary exogenous variables (VARX). Let the d-vectors $\mathbf{Y}_t = (Y_t(1), \ldots, Y_t(d))^\top$ and the m-vectors $\mathbf{X}_t = (X_t(1), \ldots, X_t(m))^\top$, $t = 1, \ldots, n$, be observations corresponding to a stationary VARX process of orders p and s, abbreviated VARX (p, s). Then, there exist $d \times d$ matrices $\mathbf{\Phi}_i$, $i = 1, \ldots, p$ and $d \times m$ matrices $\mathbf{V}_i$, $i = 0, 1, \ldots, s$, with $\mathbf{\Phi}_p \neq \mathbf{0}$ and $\mathbf{V}_s \neq \mathbf{0}$, such that

$$\mathbf{\Phi}(B)\mathbf{Y}_t = \boldsymbol{\theta} + \mathbf{V}(B)\mathbf{X}_t + \mathbf{a}_t, \quad (1)$$

where $\boldsymbol{\theta}$ denotes a $d \times 1$ constant vector, $\mathbf{\Phi}(B) = \mathbf{I}_d - \sum_{i=1}^{p} \mathbf{\Phi}_i B^i$, $\mathbf{I}_d$ being the $d \times d$ identity matrix, $\mathbf{V}(B) = \sum_{i=0}^{s} \mathbf{V}_i B^i$, B representing the usual backward shift operator. In representation (1), the random vector $\mathbf{a}_t = (a_t(1), \ldots, a_t(d))^\top$ corresponds to the error term, $t = 1, \ldots, n$. Let $\mathbf{X} = \{\mathbf{X}_t, t \in \mathbb{Z}\}$ and $\mathbf{a} = \{\mathbf{a}_t, t \in \mathbb{Z}\}$ be the exogenous and error stochastic processes, respectively. Process $\mathbf{a}$ is assumed to be a strong white noise, that is $\mathbf{a}_t$, $t \in \mathbb{Z}$, are identically and independently distributed (iid) random vectors with mean

* Corresponding author. Tel.: +1 514 343 7267; fax: +1 514 343 5700.
E-mail address: duchesne@dms.umontreal.ca (P. Duchesne).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2008.02.015

zero and regular covariance matrix $\mathbf{\Sigma}$. Without loss of generality the mean of process $\mathbf{X}$ is zero. Furthermore, we suppose that the autocovariance function of $\mathbf{X}$ defined by $\mathbf{\Gamma_X}(j) = E(\mathbf{X}_t\mathbf{X}_{t-j}^\top), j \in \mathbb{Z}$, is absolutely summable, that is $\sum_j \|\mathbf{\Gamma_X}(j)\| < \infty$, where $\|\mathbf{A}\|$ denotes the Euclidian norm of matrix $\mathbf{A}$. In the following, $\mathbf{X}$ and $\mathbf{a}$ are statistically independent stochastic processes. The VARX process $\mathbf{Y} = \{\mathbf{Y}_t, t \in \mathbb{Z}\}$ is supposed stationary and consequently all the roots of the equation $\det\{\mathbf{\Phi}(z)\} = 0$, solved in the variable $z \in \mathbb{C}$, lie outside the unit disk, where $\det(\mathbf{A})$ stands for the determinant of matrix $\mathbf{A}$.

In practical applications, the parameters of the VARX model can be estimated by conditional least squares (CLS) estimators or by alternative methods such as generalized least squares, estimated generalized least squares or maximum likelihood estimators, amongst others. See Lütkepohl (1993) for a description of these methods. However, the sensitivity of the CLS estimation procedure (or related techniques) to outliers is well-known. For example, CLS estimators may be strongly biased when the observations are not observed from the process $\{\mathbf{Y}_t\}$ satisfying (1), but for example from $\{\mathbf{Y}_t + \mathbf{W}_t\}$, where $\{\mathbf{Y}_t\}$ and $\{\mathbf{W}_t\}$ are independent, with $\{\mathbf{W}_t\}$ as an iid sequence that generates the outliers (typically the occurrence of outliers is relatively small, for example with probability less than 5%). The CLS method appears to be very sensitive to the kind of outliers which fall in this so-called additive category, since they affect the endogenous variable $\mathbf{Y}_t$ additively. Robust estimation techniques and robust diagnostics for time series models are introduced in Hampel et al. (1986) and Li (2004), amongst others.

The goal of this paper is to propose robust prediction intervals in VARX models. Whereas robust estimation has been widely studied in the time series literature, the construction of robust prediction intervals in dynamic models has received little attention, to the best of our knowledge. From a practical point of view, the empirical evaluation of existing prediction intervals to various kinds of outliers seems an important issue. Since outliers can cause bias problems for the CLS estimators of the parameters of the model, it is expected that the usual point predictions based on such estimation technique will also be biased. Furthermore, and maybe more importantly, the classical prediction confidence limits are calculated using the usual mean squared errors of prediction. Since these quantities are estimated in practice using non-robust estimation methods, they also display sensitivity to outliers. Thus, the usual prediction intervals are expected to be distorted by outliers, an undesirable outcome in the forecasting practice. In general, outliers may invalidate the predictions and the prediction confidence limits for practical purposes. These phenomena are well-known to occur in ARIMA models, see Ledolter (1989) and Chatfield (2001), who have discussed the effects of additive outliers on prediction intervals. These practical considerations call for the development of robust estimators in VARX models and for the construction of robust prediction intervals.

Since the CLS estimation method displays sensitivity to outliers, one of this paper's primary objectives consists in developing robust estimators in VARX models. The proposed method provides a multivariate generalization of the robust estimation technique elaborated in Duchesne (2005) and Bou Hamad and Duchesne (2005) for univariate autoregressive models with exogenous variables (ARX). The asymptotic distribution of the proposed estimators is studied. Using an argumentation similar to the one given in Bou Hamad and Duchesne (2005) but adapted to the multivariate case using the Cramer–Wold device, the proposed estimators are seen to converge in distribution towards the normal distribution. The asymptotic covariance matrix of the robust estimators is explicitly obtained, which is useful for constructing prediction intervals. Other estimation procedures for vector time series include the robust estimators used by Ben et al. (1999) in vector autoregressive moving average time series (VARMA) models. See also Li and Hui (1989) who developed alternative robust estimators in VARMA models.

Our new robust estimators allow us to develop new robust conditional prediction intervals in VARX models. The proposed robust prediction intervals are conditional because the coverage properties are obtained by conditioning on the lagged endogenous and exogenous processes. In time series practice, it is rather common to overlook parameter estimation uncertainty in constructing prediction intervals. However, in small and moderate samples, accounting for parameter estimation uncertainty can greatly improve the coverage properties of the prediction intervals. Consequently, the robust conditional prediction intervals being proposed here offer accuracy up to $\mathbf{O}_{\mathrm{P}}(n^{-1})$, and they take into account parameter estimation. In a seminal work on prediction in dynamic models, Schmidt (1977) considered the asymptotic distribution of dynamic simulation forecasts and provided some small sample evidence, concluding that a disregard for parameter estimation uncertainty may lead to substantial underestimation of forecast variance in small samples. For vector autoregressive (VAR) processes and VARMA processes, Baillie (1979), Reinsel (1980) and Yamamoto (1981) derived the asymptotic mean squared error of multi-step predictions. The asymptotic distribution of predictions for the more general simultaneous equation model with lagged endogenous variables and VAR errors has been studied in Baillie (1981). The optimal prediction scheme for multi-period predictions of a simultaneous equation autoregressive model with exogenous variables, with VAR errors or vector moving average errors, has been investigated in Yamamoto (1980). Ansley and Kohn (1986) have considered prediction mean squared error for state space models with estimated parameters as well as applications to ARMA models, while Reinsel and Lewis (1987) have investigated forecasting methods for non-stationary multivariate time series. Resampling techniques represent an alternative approach for constructing prediction intervals. Masarotto (1990) and Grigoletto (1998) have proposed bootstrap prediction intervals for autoregressive processes of unknown order p, when p can be estimated consistently. Cao (1999) has reviewed some resampling techniques for predicting time series. Alonso et al. (2002) have introduced the so-called sieve bootstrap for estimating prediction intervals. Their study concluded that empirical coverage comes closer to the nominal confidence level when the prediction intervals constructed incorporate the uncertainty due to parameter estimation, specially when the sample size is small. Prediction using estimated parameters in misspecified univariate models has been considered in Kunitomo and Yamamoto (1985), Stine (1987), de Luna (2000) and Kabaila and He (2004). Schorfheide (2005) has recently investigated forecasting under misspecification in VAR models.