Econometric Forecasting

Robert M. Kunst
robert.kunst@univie.ac.at

University of Vienna
and
Institute for Advanced Studies Vienna

November 26, 2012

Outline

Multivariate not necessarily better than univariate

Often, univariate prediction models outperform multivariate rival models in forecasting. Given that variables are typically highly interdependent, why?

- Larger parameter dimension: increased sampling variation due to estimation of more and sometimes poorly identified parameters;
- Specification search: more rival models available: longer search with reduced probability of finding the optimum candidate;
- Generalization of nonlinear models to higher dimension difficult: emphasis on linear structures may miss nonlinear features;
- Outliers in multivariate models are more difficult to identify and may have strong and unexpected effects.

Conditional forecasting

Economists are often interested in forecasts for x_{t+h} that assume $x_s, s \leq t$ as known as well as the values of other variables $y_s, s \leq t+h$. The solution appears to be $\mathrm{E}(x_{t+h}|x_{-\infty}^{t} \cup y_{-\infty}^{t+h})$.

- The assumed values $y_{t+1}, \ldots, y_{t+h}$ may be incorrect;
- Any dynamic model that views x_t as a function of $x_{-\infty}^{t-1}$ and $y_{-\infty}^{t}$ may miss the reaction of y to past x (feedback problem, open loop, weak and strong exogeneity);
- Changing the generation mechanism for y relative to the observations may affect the reaction of x (super exogeneity).

Static and dynamic forecasts

- A model-based forecast $\hat{x}_T(h)$ that estimates all model parameters from $t \leq T$ is also called a *dynamic forecast*;
- A model-based forecast $\hat{x}_{T+h-1}(1)$ that estimates parameters from $t \leq T$ (but assumes x_{T+1}^{T+h-1} as known) is called a *static forecast*;
- A forecast $\hat{x}_T(h)$ that estimates model parameters from any sample including x_{T+h} is called *ex-post forecast*. The two variants above are also called *ex-ante forecasts*.

Stationarity of a multivariate process

An n–variate process (x_t) is called (covariance) *stationary* iff

1. $\mathrm{E}x_t = \mu$ with time-constant n–vector μ;
2. $\mathrm{E}(x_t - \mu)(x_t - \mu)' = \Sigma$, with the time-constant and positive definite $n \times n$–matrix Σ;
3. $\mathrm{E}(x_t - \mu)(x_{t+k} - \mu)' = \Gamma(k)$, with generally non-symmetric $n \times n$–matrices of cross covariances $\Gamma(k)$.

The matrix-valued function $k \mapsto P(k), k \in \mathbb{Z}$, with cross covariances normalized to cross correlations, is called the *cross-correlation function* (CCF). Note $P(-k) = P'(k)$.

Single-equation models

An open-loop model is designed to create a forecast for a (possibly vector-valued) 'endogenous' variable $\hat{x}_t(h)$ that depends on a (usually vector-valued) 'exogenous' variable y. There are several definitions of 'exogeneity', the most important one is 'strong exogeneity' (ENGLE *et al.*). A variable y is said to be *strongly exogenous* for some parameters α iff

- y is weakly exogenous for the parameter α;
- x does not Granger-cause y.

Under strong exogeneity, open-loop modelling leads to efficient forecasting.

For reference: weak exogeneity

The definition according to ENGLE, HENDRY, RICHARD (1982):

Definition

y is *weakly exogenous* for the parameter α iff there is a *sequential cut*

$$f(x_t, y_t | X_1^{t-1}, Y_1^{t-1}; \alpha, \beta) = f(x_t | X_1^{t-1}, Y_1^t; \alpha) f(y_t | X_1^{t-1}, Y_1^{t-1}; \beta)$$

with (α, β) variation free, for all t.

In short, (α, β) are *variation free* if there are no cross restrictions on the admissible parameter space Θ. If both are one-dimensional, Θ is essentially a (possibly infinite) rectangle.

Classes of single-equation models

The two principal classes of open-loop models are

- Regression models;
- Transfer function models.

Regression models can be assumed as known.

Transfer-function models

A transfer-function model sees the *output* variable x as a function of lagged x, current and lagged *input* y, and current and lagged errors:

$$\delta(B) x_t = \omega(B) y_t + \theta(B) \varepsilon_t .$$

Lag orders of δ, ω, θ are determined as in ARMA models. A customary starting point in the specification search is an ARMA model for the input series. If $\omega_j = 0, j < d$, d is called the *delay*.

Vector autoregressions

The vector AR or VAR model

$$\Phi(B)\,x_t = \varepsilon_t$$

is sometimes generalized to the VARMA model $\Phi(B)\,x_t = \Theta(B)\varepsilon_t$. It is stable if all roots of $\det \Phi(B)$ are larger than one in modulus.

Granger causality in the VAR

Assume Φ is partitioned into submatrices corresponding to $x = (x_1', x_2')'$:

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}.$$

Then, x_1 does not Granger-cause x_2 iff $\Phi_{21} \equiv 0$. Similarly, x_2 does not Granger-cause x_1 iff $\Phi_{12} \equiv 0$.

Forecasting using a VAR

If x has been generated by a VAR

$$x_t = \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \varepsilon_t,$$

with ε_t a multivariate MDS, a feasible $\hat{x}_t(1)$ evolves from

$$\mathrm{E}(x_{t+1}|\mathfrak{I}_t) = \Phi_1 x_t + \ldots + \Phi_p x_{t-p+1} \approx \hat{\Phi}_1 x_t + \ldots + \hat{\Phi}_p x_{t-p+1},$$

as in the univariate case. At larger horizons, forecasts can be obtained by repeated substitution.

Integrated components in the VAR

If some of the roots of $\det \Phi(z)$ are 1 but all others are larger one, some of the components are not stationary but integrated. Cases of concern:

- If exactly n roots are 1 and some regularity conditions hold, all variables are I(1), there is no cointegration, and a VAR in Δx instead of x should be considered;
- If less than n roots are 1 (and s.r.c.), there is cointegration, and error correction models should be considered.

Error correction VAR models

The basic type of error-correction VAR models reads

$$\Delta x_t = \mu + \alpha\beta' x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t,$$

with the rank of $\alpha\beta'$ determined in a test sequence. Conditional on r and p, $\hat{\beta}$ follows from an eigenvalue problem, all other parameters are estimated by least squares.

Misspecification of integration orders

- If an I(2) variable has been included in the EC-VAR, most statistical results will be invalid, which may also affect forecasts;
- If a seasonally integrated variable has been included, most results will be invalid;
- If an I(0) variable has been included, there will be no problem. The corresponding unit vector will be in the cointegration space spanned by the columns of β.

Does error correction improve forecasts?

Starting with Engle & Yoo (1987), benefits of forecasts due to error correction have been studied in a rich literature. Traded wisdom is:

- At short horizons and in small samples, forecasts based on error correction models are often not as good as forecasts based on VAR models in levels or even misspecified VAR models in differences;
- In doubtful cases, ignoring cointegration is less detrimental than imposing cointegration. Too low r is often preferable to too large r;
- At large horizons and in large samples, careful cointegration modelling improves forecasts considerably.