Autoregressive-distributed lag models

by Kevin Kotzé

There are many interesting studies that are applied to time series variables that consider the use of multivariate methods. These techniques seek to describe the information that is incorporated within the temporal and cross-sectional dependence of these variables. In most cases, the goal of the analysis is to provide a better understanding of the dynamic relationship between variables and in certain cases these techniques may be used to improve the forecasting accuracy. The models that have been developed within this area of research may also be used in policy analysis purposes or for making specific inference about potential relationships.

Some of the early multivariate time series models fall within the class of linear distributed lag models. Early examples of these models include the polynomial and geometric distributed lag models. In addition, the autoregressive distributed lag (ARDL) model, which incorporates what have been termed the rational distributed lag model, continue to used in a number of studies that may be found in the current literature.

1 Polynomial distributed lag models

A general specification for the polynomial distributed lag (DL) model is:

$$\begin{aligned} y_t &= \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} + \varepsilon_t \\ &= \alpha + \beta(L) x_t + \varepsilon_t \end{aligned} \tag{1.1}$$

where ε_t is the part of the data that we are unable to explain and is assumed to be serially uncorrelated, while

$$\beta(L) = \beta_0 + \beta_1 L + \ldots + \beta_q L^q,$$

is a lag polynomial and L is the lag operator, defined by

$$L^i x_t = x_{t-i}, \; i = 0, 1, 2, \; \ldots \tag{1.2}$$

The lag coefficients are often restricted to lie on a polynomial of order $r \leq q$. In the case where $r = q$ the distributed lag model is unrestricted. A comprehensive early treatment of rational distributed lag models can be found in Dhrymes (1971).

2 Rational distributed lag models

A slight modification to the above model, which involves the use of additional parameters, would provide the specification for what is termed the rational distributed lag model:

$$y_t = \alpha + \frac{\beta(L)}{\lambda(L)} x_t + v_t \tag{2.1}$$

where $\beta(L)$ as defined above and

$$\lambda(L) = \lambda_0 + \lambda_1 L + \ldots + \lambda_p L^p$$

3 Autoregressive-distributed lag models

The above model may be extended to incorporate a number of autoregressive terms. The gives rise to the ARDL (p, q) model that may be specified as

$$y_t + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \ldots + \lambda_p y_{t-p} = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_q x_{t-q} + \varepsilon_t \tag{3.1}$$

or

$$\lambda(L) y_t = \alpha + \beta(L) x_t + \varepsilon_t \tag{3.2}$$

Note that the rational distributed lag model defined by (2.1) can also be written in the form of an ARDL $(p,\ q)$ model with moving average errors, namely

$$\lambda(L) y_t = \alpha\lambda(1) + \beta(L) x_t + \lambda(L)\varepsilon_t,$$

which represents equation (3.1) except that the error term is now given by

$$\varepsilon_t = \lambda(L) \upsilon_t$$

which takes the form of a moving average model.

When comparing these specifications, it is worth noting that recent developments in time series analysis focus on the application of the ARDL (p, q) specification since it is easier to work with and by selecting relatively large values for p and q, one can provide a reasonable approximation to the rational distributed lag specification if required.

Deterministic trends, or seasonal dummies, can be easily incorporated in the ARDL model. For example, we could have

$$y_t + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \ldots + \lambda_p y_{t-p} = \zeta_0 + \zeta_1 \cdot t + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_q x_{t-q} + \varepsilon_t \tag{3.3}$$

The model that is provided by equation (3.1) may be extended to the case of k regressors, each with a specific number of lags. Such a specification would take the form of an ARDL $(p, q_1, q_2, \ldots, q_k)$ model, which may be expressed as

$$\lambda(L, p) y_t = \sum_{j=1}^{k} \beta_j \left(L, q_j\right) x_{t,j} + \varepsilon_t$$

where

$$\begin{aligned} \lambda(L, p) &= 1 + \lambda_1 L + \lambda_2 L^2 + \ldots + \lambda_p L^p \\ \beta_j(L, q_j) &= \beta_{j,0} + \beta_{j,1} L + \ldots + \beta_{j,q_j} L^{q_j}, \quad j = 1, 2, \ldots,\ k \end{aligned} \tag{1}$$

See Hendry, Pagan, and Sargan (1984) for a comprehensive early review of ARDL models.

4 Conclusion

The framework for ARDL models make use of a relatively straightforward extension of the univariate autoregressive model, where we may include

5 References

Dhrymes, P. J. 1971. *Distributed Lags: Problems of Estimation and Formulation*. San Francisco: Holden-Day.

Hendry, David F., Adrian R. Pagan, and J.Denis Sargan. 1984. "Dynamic Specification." In *Handbook of Econometrics*, edited by Z. Grilichest and M. D. Intriligator, 2:1023–1100. Handbook of Econometrics. Elsevier.