variable is white noise (i.e., has P_2), which is clearly an interesting simplification of the full model. A further relevant reference on this topic is the article by Peña and Box (1987).

ADDITIONAL REFERENCE

Peña, D., and Box, G. E. P. (1987), "Identifying a Simplifying Structure in Time Series," *Journal of the American Statistical Association*, 82, 836–839.

Journal of Business & Economic Statistics, October 1993, Vol. 11, No. 4

Comment

Bruce E. Hansen
Department of Economics, University of Rochester, Rochester, New York 14627

The concept of *common features* introduced by Engle and Kozicki is useful for applied econometrics for two reasons. First, the presence of common features may be of direct economic interest. The authors present a good example of this with their analysis of the international business cycle. Second, common features are equivalent to a reduction in rank of certain coefficient matrices, and imposition of such restrictions will yield important gains in estimation efficiency.

Given that the phenomenon of common features will be of interest to econometricians, two inferential issues have to be explored. The first is developing tests for the presence of common features, which is the primary issue discussed in the article by Engle and Kozicki. The second issue is efficient estimation of models subject to common-features restrictions. The authors present an estimation method for the common-feature vector, but they neither present a distributional theory for the point estimate nor discuss estimation of the full model.

This silence is somewhat confusing because common-features restrictions are equivalent (as the authors point out) to reduced-rank restrictions, and the theory of estimation and inference in the latter is fairly well worked out. Reduced-rank methods are fairly easy to implement, yield estimates for the full model and not just the common-features vector, and have an associated asymptotic theory of inference.

The similarities and contrasts between the classic reduced-rank approach and the Engle–Kozicki approach can be best illustrated by working through a simple example. Suppose that y_t is an $n \times 1$ vector and z_t is $k \times 1$ (with $k \geq n$). These series are related by the linear regression equation

$$y_t = \Gamma z_t + \xi_t, \tag{1}$$

where ξ_t satisfies $E(\xi_t \mid z_t) = 0$.

The hypothesis of common features is that there exists an $n \times 1$ *common-features vector* δ such that the residual $u_t(\delta) = \delta' y_t$ is uncorrelated with z_t. This is equivalent to the reduced-rank restriction $\Gamma = \Lambda\Phi$, where Λ is $n \times (n - 1)$. These two expressions of the null hypothesis are related by the requirement that $\delta'\Phi = 0$.

Engle and Kozicki's proposal is to form a standard test statistic for orthogonality between $u_t(\delta)$ and z_t. Let Y and Z be the $T \times 1$ and $T \times k$ observation matrices for y_t and z_t, let $P_z = Z\,(Z'Z)^{-1}\,Z'$ be the projection matrix, and let $u(\delta)$ be the residual vector. Set

$$s(\delta) = \frac{u(\delta)' P_z u(\delta)}{\hat{\sigma}^2(\delta)},$$

where $\hat{\sigma}^2(\delta)$ equals either $(1/n)u(\delta)'u(\delta)$ for a Lagrange multiplier (LM) type test statistic or $(1/n)u(\delta)'(I - P_z)u(\delta)$ for a Wald-type statistic.The Engle–Kozicki statistic is then

$$\hat{s} = s(\hat{\delta}) = \inf_{\delta} s(\delta).$$

The authors suggest that $\hat{\delta}$ can be found by nonlinear optimization or can be approximated by an asymptotically equivalent two-stage least squares estimator. They also suggest that $\hat{\delta}$ be normalized by setting the first element of $\hat{\delta}$ equal to unity.

Let us take a closer look at their LM-type statistic. The preceding definitions can be rewritten to yield

$$\hat{s} = \inf_{\delta} n \frac{\delta Y' P_z Y \delta}{\delta' Y' Y \delta}. \tag{2}$$

Note that the normalization chosen for $\hat{\delta}$ is irrelevant. In addition, the minimization problem concerns a ratio of quadratic forms. The exact solution is given by

$$\hat{s} = T\hat{\lambda}, \tag{3}$$

where $\hat{\lambda}$ is the smallest root solving the eigenvalue problem

$$(Y'Y)^{-1}\, Y' P_z Y \hat{\delta} = \hat{\delta}\hat{\lambda}. \tag{4}$$

See, for example, Rao (1973, p. 74, problem 22). $\hat{\lambda}$ is known as a characteristic root of $Y'P_zY$ with respect to $Y'Y$, and $\hat{\delta}$ is the associated eigenvector. This shows that the Engle–Kozicki statistic can be calculated directly without nonlinear optimization, if desired.

As pointed out in their article, an alternative approach is to use reduced-rank regression. This is the term for maximum likelihood estimation of (1) under the assumption that ξ_t is iid normal. It is known that the maximum likelihood estimator (MLE) $\tilde{\Phi}'$ is given by the first $n - 1$ eigenvectors of $Z'Y\,(Y'Y)^{-1}\,Y'Z$ with respect to $Z'Z$ (where the eigenvectors have been ordered by the declining eigenvalues). Furthermore, $\tilde{\Lambda} = Y'Z\tilde{\Phi}'$. The likelihood ratio (LR) statistic for testing the common-features (reduced-rank) hypothesis against the unrestricted model is given by

$$\mathrm{LR} = T\ln(1 - \tilde{\lambda}) \simeq T\tilde{\lambda}, \qquad (5)$$

where $\tilde{\lambda}$ is the smallest eigenvalue of $Z'Y(Y'Y)^{-1}Y'Z$ with respect to $Z'Z$. This is the eigenvalue problem

$$(Z'Z)^{-1}\,Z'Y(Y'Y)^{-1}\,Y'\tilde{\delta} = \tilde{\delta}\tilde{\lambda}. \qquad (6)$$

But the eigenvalues of $(Z'Z)^{-1}\,Z'Y(Y'Y)^{-1}\,Y'$ are the same as those of $(Y'Y)^{-1}\,Y'Z(Z'Z)^{-1}\,Z'Y = (Y'Y)^{-1}Y'P_zY$, which is the eigenvalue problem (4). Thus these two problems are identical, which implies that $\hat{\lambda} = \tilde{\lambda}$. We find that the only difference between the LR statistic and the Engle–Kozicki statistic is the logarithmic transformation in (5)! The maximum likelihood estimate of the common-features vector δ is given by the eigenvector of (6) associated with the smallest characteristic root $\tilde{\lambda}$.

These results can be shown to carry through for a more general model, which also has additional regressors x_t as in Equation (19) of Engle and Kozicki's article. Although the reduced-rank regression methods are easily generalized to the case in which the rank of Π is $n - r$ for $r > 1$ [so that Λ is an $n \times (n - r)$ matrix], I am unsure how to generalize the analytic solution given in (2)–(4). Thus it is not clear if the equivalence $\tilde{\lambda} = \hat{\lambda}$ carries over to the general case.

In any case, the likelihood approach has a natural advantage over the Engle–Kozicki approach in that the former produces estimates of all of the parameters of the model (Λ, Φ, and δ), as well as a test statistic of the validity of the common-features restriction. The Engle–Kozicki method only produces an estimate of δ and the test. If applied economists find that the common-features restriction is valid, they will often be interested in estimates of the full model, not just the common-features vector.

Both methods are subject to the criticism that the chi-squared distribution approximation requires that the error ξ_t be homoscedastic. An interesting advantage of the Engle–Kozicki approach is that it is easy to construct a robust test statistic. Set

$$r(\delta) = \sum_{t=1}^{T} z_t' u_t(\delta) \left(\sum_{t=1}^{T} z_t z_t' u_t(\delta)^2 \right)^{-1} \sum_{t=1}^{T} z_t u_t(\delta). \qquad (7)$$

Minimization of $r(\delta)$ over δ produces a robust test of the common-features restriction. Nonlinear optimization methods may be necessary to evaluate (7). Alternative tests could also be constructed from the MLE using robust covariance matrix estimates.

In summary, Engle and Kozicki have introduced a new concept that will help to organize communication and research in applied econometrics. They have also introduced a simple method to test this concept, which has fairly general applicability. In many modeling contexts, however, better estimation methods may be available from the literature on reduced-rank regression.

ACKNOWLEDGMENT

This research was supported by the National Science Foundation.

ADDITIONAL REFERENCE

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications* (2nd Ed.) New York: John Wiley.