GENERALISED WALD TYPE TESTS OF NONLINEAR RESTRICTIONS

Zaka Ratsimalahelo

University of Franche-Comté, U.F.R. Science Economique, 45D, av. de l'Observatoire,
25 030 Besançon - France

Abstract: This paper proposes a generalised Wald type tests to test the hypothesis of nonlinear restrictions. We circumvent the problem of singularity of the covariance matrix associated with the usual Wald test by proposing a generalised inverse procedure, and an alternative simple procedure which can be approximated by a suitable chi-square distribution. New threshold values are derived to estimate the rank of the covariance matrix. We also propose an operational procedure to test the hypothesis of the Granger non-causality in cointegrated systems.

Key words: Wald tests; Matrix perturbation theory; Nonlinear restrictions; Spectral decomposition; Covariance matrix degenerate; Granger causality.

1. INTRODUCTION

Analyses of economic data often entail the testing of hypotheses that imply complex nonlinear restrictions on subsets of parameters, and it is thus desirable to employ econometric methods that are flexible both in terms of their applications and implementation for typical data sets. In this paper we are interested in testing the null hypothesis H_0 in the following form

$$H_0 : g(\theta) = 0 \qquad (1)$$

where the vector parameter $\theta \in \Theta$ a k-dimensional compact subspace of $\mathbb{R}^k$ and $g(\theta)$ is a mapping from $\mathbb{R}^k$ to $\mathbb{R}^l$ continuously differentiable functions of θ .

In econometrics literature the Wald tests are commonly used to test this null hypothesis since the generality of its formulation affords the testing of several interesting economic hypotheses which might present formidable difficulties for other procedures (Sargan 1980; Gregory and Veall, 1985, 1986).
Moreover, it is well known that in a linear regression model with normally distributed errors, the Wald statistic for a set of linear restrictions is a monotonic transformation of the likelihood ratio (LR) test statistics.

Under regularity conditions, the null asymptotic distribution is chi-squared distribution, and this is the distribution that one usually uses to carry out hypothesis tests. On the other hand, under the nonregularity, Wald tests fail to have limiting chi-squared distribution in general (Andrew 1987). For example, in vector autoregressive (VAR) processes, if the process is stationary, the multivariate least squares (LS) estimator of the coefficients has a non-singular asymptotic distribution whereas the distribution becomes singular if some variables are integrated or cointegrated.
Another example is the Granger non-causality test. As shown in Toda and Phillips (1993) when there is a cointegrating relationship, in general the Wald statistic of the Granger non-causality test has a non-

standard limiting distribution, depending on nuisance parameters.

To overcome this problem, Lutkepohl and Burda (1997) used a generalized inverse of the asymptotic covariance matrix and they showed that the Wald test has an asymptotically standard distribution. However, the serious problem is how detect the degeneracy or the rank of asymptotic covariance matrix.
It is well known that the rank of matrix is equal to the number of nonzero eigenvalues. Lukepohl and Burda (1997) used an ad-hoc threshold to determine if the eigenvalues are zero or not. Unfortunately their choice of a threshold is more approximate. The values of the threshold do not appear to be based on any explicit analytical expressions. An obvious way of circumventing this difficulty would be to determine the distribution of the eigenvalues.

The purpose of this paper is twofold. First, under nonregularity conditions, we modify the usual Wald test to ensure that the null asymptotic distribution is a chi-squared distribution. Secondly we propose new threshold values to test if the eigenvalues are significantly different from zero or not. This permits to determine the rank of the covariance matrix.

The paper is organised as follows. Section 2 develops the general expression for the tests statistics. Section 3 proposes new threshold values for the eigenvalues. The tests of Granger non-causality are proposed in section 4. Section 5 gives concluding remarks.

2. WALD TESTS

In this section, we will develop general expressions for Wald test statistics for nonlinear restrictions.

Let $\widehat{\theta}$ be an estimator of θ based on a sample of size n. We make the following assumption regarding the estimator $\widehat{\theta}$.

Assumption (i) $\widehat{\theta} = \theta + O_P(n^{-1/2})$;

(ii) $nVar(\widehat{\theta}) = \Omega_n$;

(iii) $\sqrt{n}(\widehat{\theta} - \theta) \to^d N_l(0,\Omega)$, where $\Omega = \lim_{n\to\infty} \Omega_n$ is finite non zero but possibly singular; and the symbol $\to^d$ represents weak convergence of the associated probability measure.

(iv) a consistent estimator $\widehat{\Omega}$ of Ω is available.

The asymptotic normality of $\widehat{\theta}$ implies that under the null hypothesis

$$\sqrt{\mathrm{n}}g(\widehat{\theta}) \to^d N(0,\Sigma) \tag{2}$$

where $\Sigma = G(\theta)\Omega G(\theta)'$
with

$G(\theta) = \dfrac{\partial g(\theta)}{\partial \theta'}$ is the matrix of first – order partial derivatives of the restrictions.

To establish the limiting distribution of the standard Wald test, we assume the following two regularity conditions.

C1. The matrix of first-order partial derivatives of the restrictions $G(\theta)$ is of full rank for all θ in the parameter space.

C2. The asymptotic covariance matrix Ω is non singular.

Let $\widehat{\Sigma}$ be a consistent estimator of Σ, obtained by replacing $G(\theta)$ and Ω by their consistent estimator $G(\widehat{\theta})$ and $\widehat{\Omega}$ respectively. Hence under the two regularity conditions C1 and C2, the Wald statistic for testing (1) is given by

$$W = ng(\widehat{\theta})'\widehat{\Sigma}^{-1}g(\widehat{\theta}) \tag{3}$$

and W is distributed as chi-square with k degrees of freedom on H_0.

Now, when the covariance matrix Σ is singular, i.e. at least one of these regularity conditions (C1 and C2) is not satisfied, the Wald statistic may not have an asymptotic chi-square distribution. The singularity of Σ comes from three possible situations. First the matrix of first-order partial derivatives of the restrictions, $G(\theta)$ has reduced rank. Secondly the asymptotic covariance Ω is a matrix degenerate and finally both matrices $G(\theta)$ and Ω are singular.

If the function $g(\theta)$ involves products of the elements of θ then the matrix of first-order partial derivatives $G(\theta)$ is likely to have a reduced rank over part of the parameter space. Such functions are relatively common in time series analysis. For instance, impulse responses and related quantities of interest in a vector autoregressive (VAR) analysis involve products of the VAR coefficients. Similarly, such functions come up in analyzing multi-step causality in VAR models (Lutkepohl and Burda, 1997); or Granger causality in VARMA models (Lutkepohl, 1991) and in testing of restrictions on the levels of parameters of a cointegrated VAR process.

The singularity of the asymptotic covariance matrix of the estimator parameters occurs for instance in VAR processes. If the process is nonstationary, i.e. some variables are integrated or cointegrated, the multivariate least squares (LS) estimator of the coefficients has a singular asymptotic distribution. This singularity problem has also been noted or discussed in the context of the long-run impact matrix, see (Johansen, 1995).

In that case, it is usual practice to resort to a generalized inverse procedure when we have inverted a singular matrix. That is, we have under the null hypothesis,

$$W^{+} = ng(\widehat{\theta})'\widehat{\Sigma}^{+}g(\widehat{\theta}) \qquad (4)$$

where $\widehat{\Sigma}^{+}$ is the Moore-Penrose generalized inverse of a matrix $\widehat{\Sigma}$. The Wald statistics still have an asymptotic chi-square distribution under an additional condition, that is the rank of $\widehat{\Sigma}$ is equal to the rank of Σ, see (Andrews, 1987). Unfortunately, this latter condition is not always easy to verify in practice.

To overcome this difficulty, the Moore-Penrose generalized inverse can be obtained by using the spectral decomposition of $\widehat{\Sigma}$. Diagonalising $\widehat{\Sigma}$ as $\widehat{Q}_1'\widehat{\Lambda}\widehat{Q}_1$, where $\widehat{\Lambda}$ (r×r) is the diagonal matrix of the positive eigenvalues: $\widehat{\Lambda} = diag(\widehat{\lambda}_1,...,\widehat{\lambda}_r)$ with $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq ... \geq \widehat{\lambda}_r \succ 0$ and the columns of $\widehat{Q}_1$, $\widehat{\eta}_i$ $(\mathrm{i} = 1,...,\mathrm{r})$ are the eigenvectors corresponding to the positive eigenvalues. Hence the generalised Wald statistics

$$W^{+} = ng(\widehat{\theta})'\widehat{Q}_1'\widehat{\Lambda}^{-1}\widehat{Q}_1 g(\widehat{\theta}) \qquad (5)$$

is asymptotically chi-squared with r degrees of freedom on H_0, and $r = rank(\Sigma)$.

In some instances we will need to consider the case where the null hypothesis H_0 holds only approximately. The alternative hypothesis will then be considered

$$H_1 : g(\theta) = \gamma / \sqrt{n}\,. \qquad (6)$$

When H_1 holds, The Wald statistic W^{+} satisfies

$$W^{+} \to^{d} \chi^2(r,\zeta) \qquad (7)$$

where $\chi^2(r,\zeta)$ is the noncentral chi-square distribution with r degrees of freedom, and

$$r = rank(\Sigma)$$

and noncentrality parameter

$$\zeta = \gamma' Q_1' \Lambda^{-1} Q_1 \gamma\,. \qquad (8)$$

This test statistic has fewer degrees of freedom than the usual Wald test since $r = rank(\Sigma) \leq k$. Hence it has improved power, in particular if superfluous restrictions are removed. Gallant (1977) finds in a small sample Monte Carlo investigation that power advantages may in fact result from taking into account fewer restrictions.
The generalised Wald statistic is compared with a critical value draw from a central chi-square distribution with r degrees of freedom. Indeed the determination of the degrees of freedom is a serious problem. The next section proposes a solution to this problem.

3. NEW THRESHOLD VALUES FOR THE EIGENVALUES

However the problem associated with this test statistic is the determination of the number of nonzero eigenvalues of the matrix Σ, that is the rank of Σ. We suggest a sequence pretest procedure. We develop the test procedures for

$$H'_0 : \lambda_i = 0 \qquad (9)$$

where the λ_i are the eigenvalues of the matrix Σ. To decide if the eigenvalues are nonzero or not, Lutkepohl and Burda (1997) used a threshold c which depends on different factors such as the sample size n and the value of the largest eigenvalue $\widehat{\lambda}_1$ of $\widehat{\Sigma}$. Then λ_i is nonzero if the eigenvalue $\widehat{\lambda}_i$ of the estimator $\widehat{\Sigma}$ of Σ is greater than c. In other words, for our purposes one rejects the null hypothesis if $\widehat{\lambda}_i > c$. Unfortunately their choice of a threshold c is more approximate, the values of the threshold do not appear to be based on any explicit analytical expressions but are selected on an ad hoc basis. Consequently, it will not be sure to take all restrictions which correspond to nonzero eigenvalues of the covariance matrix Σ. Thus the rank estimated may be larger than the true rank. We shall propose an appropriate threshold level criterion.
The approach consists to derive the statistical properties of $\widehat{\lambda}_i$. To this end, we use some results of the matrix perturbation theory.

Let us consider the matrix perturbation theory

$$\widehat{\Sigma} = \Sigma + \varepsilon B \tag{10}$$

where εB, with small ε, is the error matrix $\widehat{\Sigma} - \Sigma$ or the matrix perturbation. These results of the matrix perturbation theory indicate how much the eigenvalues and eigenvectors of the matrix $\widehat{\Sigma}$ can differ from those of Σ for small ε.

Assume that the estimator $\widehat{\Sigma}$ of Σ is $root - n$ consistent and has a limiting normal distribution, then the perturbation term εB is of order $O_P(n^{-1/2})$ and can be seen as a zero mean Gaussian random matrix. Hence the eigenvalues $\widehat{\lambda}_i$ of $\widehat{\Sigma}$ are asymptotically normal, that is $\sqrt{n}(\widehat{\lambda}_i - \lambda_i) \to^d N(0, \sigma_i^2)$, see (Ratsimalahelo, 2000, 2003)

Having fully defined the statistical properties of $\widehat{\lambda}_i$, we shall propose two alternative test statistics to decide whether eigenvalue should be declared zero or not.

(i) First, let $\widehat{\sigma}_i^2$ be a consistent estimator of σ_i^2, according to the asymptotic distribution of $\widehat{\lambda}_i$, we reject the null hypothesis H'_0 if

$$\widehat{\lambda}_i > z_\alpha \widehat{\sigma}_i / \sqrt{n} \tag{11}$$

where z_α represents the percentile of the standard normal distribution.

(ii) Secondly, under H'_0, the test statistic

$$q = \frac{n\widehat{\lambda}_i^2}{\widehat{\sigma}_i^2} \tag{12}$$

is a chi-squared random variable with one degree of freedom. We reject H'_0 if $q > \chi_\alpha^2$ where χ_α^2 is the critical value of the chi-square distribution.

We have defined two values of the threshold adequately. To assess the test statistics, approximate $p-$values are computed with reference to standard normal distribution and chi-square distribution respectively.
In the next section we will give an example where testing problem occurs and we will show how it can be solved with the method proposed in the previous sections.

4 GRANGER NON-CAUSALITY TEST

In general, Wald tests for Granger non-causality in vector autoregressive (VAR) process are known to have non-standard asymptotic properties for cointegrated systems. The test based upon the non-standard distribution is very difficult, if not impossible to use in practice.

In this section, we propose a procedure for conducting Granger non-causality tests that are based on methodology developing in sections 2 and 3. We will show that the Granger non-causality tests have an asymptotically standard distribution.

Consider m-vector process $\{y_t\}$ generated by vector autoregressive (VAR) model of order p,

$$\mathrm{y}_t = \mu + \mathrm{A(L)y}_{t-1} + \Phi \mathrm{D}_t + \varepsilon_t \qquad t = 1, 2, \ldots, T \tag{13}$$

where $y_t = \{y_{1t}, \ldots, y_{mt}\}'$, μ is the constant vector, $A(L) = \sum_{i=1}^{p} A_i L^{i-1}$, L is the backward shift operator, Φ is the $m \times k$ coefficient matrix, and $\varepsilon_t = \{\varepsilon_{1t}, \ldots, \varepsilon_{mt}\}'$ is a zero mean independent white noise process with covariance matrix Ψ and, for $\mathrm{i} = 1, \ldots, \mathrm{r}$ $E \mid \varepsilon_{it} \mid^{2+\tau} < \infty$ for some $\tau > 0$. The deterministic terms D_t can contain a linear time, seasonal dummies, intervention dummies, or other regressors that we consider fixed and non-stochastic.

It will be convenient to write the process (13) in a vector error correction (VEC) form :

$$\Delta y_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \Phi D_t + \varepsilon_t \tag{14}$$

where $\Pi = \sum_{i=1}^{p} A_i - I_m$ and $\Gamma_i = -\sum_{j=i+1}^{p} A_i$ for $i = 1, \ldots, p-1$.

By portioning y_t in (h_1) and $(h_2 = m - h_1)$ with $h_2 \geq 1$ dimensional subvectors y_t^1 and y_t^2 and A_i matrices portioned accordingly, then y_t^2 does not Granger-cause the y_t^1 if the following hypothesis

$$H_0 : A_{12i} = 0 \quad \text{for } i = 1,\ldots,r \qquad (15)$$

is true.

Let use define $y_t^1 = L_{h1}y$ with $L_{h1} = [I_{h1}, 0]$ and $y_t^2 = L_{h2}y$ with $L_{h2} = [0, I_{h2}]$ and I_h is the identity matrix of rank h. Let $A = [A_1, A_2, \ldots, A_p]$ and $\alpha = vec(A)$ where $vec(.)$ denotes the vectorization operator that stacks the columns of argument matrix. The null hypothesis that y_t^2 does not Granger-cause the y_t^1 can be written as

$$H_0 : L_{h1}A_iL_{h2}' = 0 \quad (i = 1,\ldots,p) \qquad (16)$$

or equivalently

$$H_0 : L_{h1}AL' = 0 \quad \text{or} \quad R\alpha = 0 \qquad (17)$$

where $R = L_{h1} \otimes L$ and $L = I_p \otimes L_{h2}$.

Let $\widehat{\alpha} = vec(\widehat{A})$ where $\widehat{A}$ is the least squares (LS) estimator of A in equation (13) then following Theorem 2.3 of Phillips (1998), we have

$$T^{1/2}(\widehat{\alpha} - \alpha) \to^d N(0, \Omega_\alpha) \qquad (18)$$

where the asymptotic covariance matrix Ω_α is finite non zero but possibly singular.
Consequently, we have under the null hypothesis of non-causality in equation (17)

$$T^{1/2}R\widehat{\alpha} \to^d N(0, R\Omega_\alpha R') \qquad (19)$$

If $R\Omega_\alpha R'$ is of full rank then a standard Wald test statistic, for H_0 in levels VAR's format, has a chi-square distribution with degrees of freedom equal to the number of restrictions as T grows:

$$W = T(R\widehat{\alpha})'(R\widehat{\Omega}_\alpha R')^{-1}(R\widehat{\alpha}) \to^d \chi_v^2 \qquad (20)$$

However, $R\Omega_\alpha R'$ can be degenerate, as is well known, because Ω_α is singular. When $R\Omega_\alpha R'$ is degenerate, the Wald statistic has an asymptotically non-standard distribution, and cannot be easily tested (Toda and Phillips, 1993, 1994). Thus, the levels VAR's model has not been used in practice in possibly cointegrated systems.

To overcome this problem, we modify the Wald statistic in the following. The generalised Wald test of the null hypothesis of non-causality in equation (17):

$$W^+ = T(R\widehat{\alpha})'(R\widehat{\Omega}_\alpha R')^+(R\widehat{\alpha}) \qquad (21)$$

has an asymptotic chi-squared distribution with r degrees of freedom, and $r = rank(R\Omega_\alpha R')$. Here $(R\Omega_\alpha R')^+$ is the Moore-Penrose generalized inverse of a matrix $R\Omega_\alpha R'$.

The testing procedure depends upon how we detect the rank of $R\Omega_\alpha R'$. In order to detect degeneracy or the rank of its matrix, we use the spectral decomposition of $R\widehat{\Omega}_\alpha R'$ and apply the testing procedure developed in section 3.

5. CONCLUDING REMARKS

In this paper, we have proposed a generalised Wald type tests which guarantees the asymptotic chi-square distribution in the case where the standard Wald statistic fails to have its limiting chi-square distribution due to nonregularity conditions.

Since the limiting distribution of the Wald statistic for testing nonlinear restrictions depends substantially on the singularity or not of the asymptotic covariance matrix, the determination of its rank is of great importance in this context. It is well known that the rank of the matrix is equal to the number of the eigenvalues nonzero. The present paper provides new threshold values for the eigenvalues which permit to decide if the eigenvalues are nonzero or not.

We have also proposed an operational procedure to test the hypothesis of the Granger non-causality in cointegrated systems. It circumvents the problem of possible degeneracy of the variance-covariance matrix associated with the usual Wald test statistic.

The modification of the Wald tests considered here provides general solutions to the nonsingularity problem. It can be used in statistical signal processing.

REFERENCES

Andrews, D.W.K. (1987). "Asymptotic results for generalized Wald tests". *Econometric Theory,* **3**, 348-358.

Dolado, J., J; and Lutkepohl, H. (1996) Making Wald Tests work for cointegrated VAR systems, *Econometrics Reviews,* **15,** 369-386.

Gallant, A.R. (1977). "Testing a nonlinear regression specification: A nonregular case". *Journal of American Statistical Association*, **72**, 523-530.

Gregory, A.W. and M. R. Veall (1985). "Formulating Wald tests of nonlinear restrictions", *Econometrica*, **53.** 1465-1468.

Gregory, A.W. and M. R. Veall, (1986). "Wald tests of Common Factor Restrictions". *Economics Letters*, **22**, 203-208.

Johansen, S. (1995). *Likelihood-Based Inference in cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Lutkepohl, H. (1991). *Introduction to multiple time series analysis*. Springer, Berlin.

Lutkepohl, H. and Burda (1997). "Modified Wald tests under nonregular conditions". *Journal of Econometrics,* **78**. 315-332.

Paruolo, P. (1997) "Asymptotic Inference on Moving Average Impact Matrix in Cointegrated I(1) VAR Systems," *Econometric Theory,* **13,** 79-118.

Phillips, P. C. B. (1998): "Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs," *Journal of Econometrics,* **83**, 21-56.

Ratsimalahelo, Z. (2000). Rank Test Based on Matrix Perturbation Theory. Working paper, University of Franche-Comté, U.F.R Science-Economique. A second version (2002).

Ratsimalahelo, Z. (2003): *Specification of VARMA Echelon Form Models*, in Neck R. (eds) on Modelling and Control of Economic Systems, pp 129-137, Elsevier, Pergamon Press.

Sargan, J. D. (1980). "Some Tests of Dynamic Specification for a single Equation" *Econometrica*, **48**, 879-898.

Toda H. Y. and P.C. B. Phillips (1993): "Vector Autoregression and Causality," *Econometrica*, **59**, 229-255.

Toda H. Y. and P.C. B. Phillips (1994): "Vector Autoregression and Causality: A Theoretical Overview and Simulation Study," *Econometric Reviews*, **13**, 259-285.