COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 761

ASYMPTOTIC RESULTS FOR ESTIMATORS OF MOORE-PENROSE INVERTED MATRICES AND FOR GENERALIZED WALD TESTS[1]

by

Donald W. K. Andrews

September 1985

ASYMPTOTIC RESULTS FOR ESTIMATORS OF MOORE-PENROSE INVERTED MATRICES AND FOR GENERALIZED WALD TESTS[1]

Donald W.K. Andrews[†]
Cowles Foundation for Research in Economics
Yale University

ABSTRACT

This note presents (i) necessary and sufficient conditions for the consistency of estimators of Moore-Penrose inverted matrices, and (ii) sufficient conditions for convergence to a chi-square distribution of quadratic forms based on g-inverted weighting matrices. The latter results are needed to establish asymptotic significance levels and local power properties of generalized Wald tests (i.e., Wald tests with singular covariance matrices). Included in this class of tests are Hausman specification tests and various goodness of fit tests, among others. The results are relevant to procedures currently in the literature, since they illustrate that some results stated in the literature hold only under more restrictive assumptions than those given.

February 1985

Revised: September 1985

† The Cowles Foundation for Research in Economics
Yale University
Box 2125, Yale Station
New Haven, Connecticut 06520

1. Introduction

This note is concerned with the consistency of estimators of Moore-Penrose inverted matrices and with the convergence in distribution of quadratic forms based on g-inverted weighting matrices. Such results are important, since g-inverses, and in particular the Moore-Penrose inverse, are being used increasingly in econometrics and statistics. Examples include inferential procedures for the rank deficient and/or singular covariance matrix linear model (e.g., see Mitra (1980)), Hausman specification tests (see Hausman and Taylor (1981), Holly (1982), and Duncan (1983), goodness of fit tests (see Andrews (1985a,b), Heckman (1984), and Moore (1977)), generalized method of moments specification tests (see Newey (1984)), and more generally, all Wald tests based on statistics that have singular asymptotic normal distributions, i.e., generalized Wald tests (see Moore (1977)).

The purpose of this note is to provide several general asymptotic results involving g-inverses that can be applied to numerous existing and prospective inferential procedures. The results are relevant to procedures currently in the literature, since they illustrate that some results stated in the literature only hold under more restrictive assumptions than are presented.

The g-inverse of an estimator A_n of a singular covariance matrix A often is used to replace in a quadratic form some g-inverse A^- of the covariance matrix A. Commonly, the claim then is made that if A_n is consistent for A, the use of A_n^- rather than A^- will not affect the asymptotic distribution of the quadratic form in question. This claim is important because it is used to determine the critical region of various test procedures (such as those listed above). In some cases, it is justified by the second claim that if A_n is consistent for A, then A_n^+ is consistent for A^+ (where $(\cdot)^+$ denotes the Moose-Penrose inverse). In this note we show that neither claim is true in general, but that both are true under additional conditions, including a rank condition on the matrices A_n.

2. Consistency of Estimators of Moore-Penrose Inverted Matrices

It is easy to see that $A_n \to A$ does not necessarily imply $A_n^+ \to A^+$. Consider the scalar case with $A_n = 1/n$, for $n = 1,2,\ldots,$ and $A = 0$. Then, $A_n^{-1} = n \to \infty \neq A^+ = 0$. The result $A_n^+ \to A^+$ fails in this case, and in others, because the Moore-Penrose inverse is not a continuous function. The following Lemma establishes the conditions under which this discontinuity does, and does not, affect the convergence of A_n^+ to A^+.

Let A_n, $n = 1,2,\ldots,$ and A be nonrandom complex $r \times s$ matrices. Let $\mathrm{rk}[A_n]$ denote the rank of A_n, let $\|\cdot\|$ denote the Euclidean norm, and let ev., i.o., and iff abbreviate "eventually" (i.e., for all but a finite number of n, or equivently, for all n sufficiently large), "infinitely often" (i.e., for infinitely many n), and "if and only if," respectively.

<u>Lemma 1</u>. Suppose $A_n \xrightarrow{n\to\infty} A$. Then, $A_n^+ \xrightarrow{n\to\infty} A^+$ iff $\mathrm{rk}[A_n] = \mathrm{rk}[A]$ ev. iff $\limsup_{n\to\infty}\|A_n^+\| < \infty$.

<u>Proof of Lemma 1</u>. Without loss of generality, assume $\mathrm{rk}[A_n] = g$, $\forall n$, for some constant g. Then, either $g = \mathrm{rk}[A]$ or $g > \mathrm{rk}[A]$.

The following is well known: $A_n^+ \xrightarrow{n\to\infty} A^+$ iff every subsequence of $\{A_n^+\}$ has a sub-subsequence that converges to A^+. Let $\{n_k : k=1, 2, \ldots\}$ be a subsequence of $\{n\}$. By the singular value decomposition (e.g., see Rao (1973, pp. 42-3), we can write

$$A_{n_k} = B_{n_k} D_{n_k} C_{n_k}^* , \qquad \text{for } D_{n_k} = \begin{bmatrix} \Delta_{n_k} & 0 \\ 0 & 0 \end{bmatrix} , \tag{1}$$

where $(\cdot)^*$ denotes the conjugate transpose of a matrix, B_{n_k} and C_{n_k} are unitary complex matrices (i.e., $B_{n_k}^* B_{n_k} = B_{n_k} B_{n_k}^* = I_r$), and Δ_{n_k} is a $g \times g$ diagonal nonsingular matrix with diagonal elements given by the eigenvalues of $A^*_{n_k}A_{n_k}$. The elements of B_{n_k}, C_{n_k}, and D_{n_k} are bounded above, since (i) B_{n_k} and C_{n_k} are unitary, and (ii) $A^*_{n_k}A_{n_k} \xrightarrow{k\to\infty} A^*A$ implies that the set of eigenvalues of $A^*_{n_k}A_{n_k}$ converges to that of A^*A. Hence, there is a subsequence $\{n_m\}$ of $\{n_k\}$ such that B_{n_m}, C_{n_m}, and Δ_{n_m} converge to some matrices B, C, and Δ as $m \to \infty$.

Form the $r \times s$ matrix $D = \begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix}$. It is easy to show that $A = BDC^*$, B and C are unitary, Δ is nonsingular if $g = \mathrm{rk}[A]$ (provided $\mathrm{rk}[A] > 0$), and Δ is singular if $g > \mathrm{rk}[A]$.

If $g = \mathrm{rk}[A] > 0$, then

$$D^+_{n_m} = \begin{bmatrix} \Delta^{-1}_{n_m} & 0 \\ 0 & 0 \end{bmatrix} \xrightarrow{m\to\infty} \begin{bmatrix} \Delta^{-1} & 0 \\ 0 & 0 \end{bmatrix} = D^+, \text{ and}$$

$$\lim_{m\to\infty} A^+_{n_m} = \lim_{m\to\infty} C_{n_m} D^+_{n_m} B^*_{n_m} = CD^+B^* = (BDC^*)^+ = A^+, \qquad (2)$$

as desired. (The same result is trivial when $g = \mathrm{rk}[A] = 0$.) Alternatively, if $g > \mathrm{rk}[A]$, then for some $j \le g$, $[\Delta_{n_m}]_{jj} \xrightarrow{m\to\infty} 0$, $\|\Delta^{-1}_{n_m}\| \xrightarrow{m\to\infty} \infty$, and

$$\lim_{m\to\infty}\|A^+_{n_m}\| = \lim_{m\to\infty}\|C_{n_m} D^+_{n_m} B^*_{n_m}\| = \infty. \qquad \square$$

Lemma 1 can be used to establish necessary and sufficient conditions for weak and strong consistency of estimators of Moore-Penrose inverted matrices. Suppose $\{A_n : n = 1,2,\ldots\}$ are random matrices, but otherwise are as above. Let a.s. abbreviate "almost surely" (i.e., with probability one), and let "$\xrightarrow{p}$" denote convergence in probability.

<u>Theorem 1.</u> (a) Suppose $A_n \xrightarrow{n\to\infty} A$ a.s. Then, $A^+_n \xrightarrow{n\to\infty} A^+$ a.s. iff $\mathrm{rk}[A_n] = \mathrm{rk}[A]$ ev. a.s. iff $\limsup_{n\to\infty}\|A^+_n\| < \infty$ a.s.

(b) Suppose $A_n \xrightarrow{p} A$ as $n \to \infty$. Then, $A^+_n \xrightarrow{p} A^+$ as $n \to \infty$ iff $P(\mathrm{rk}[A_n] = \mathrm{rk}[A]) \xrightarrow{n\to\infty} 1$ iff $\{A^+_n : n = 1,2,\ldots\}$ are stochastically bounded (i.e., $\exists M < \infty$ such that $P(\|A^+_n\| < M) \xrightarrow{n\to\infty} 1$).

<u>Comment</u>: The Theorem shows that if the rank condition $rk[A_n] = rk[A]$ ev. a.s. (or $P(rk[A_n] = rk[A]) \xrightarrow{n\to\infty} 1$) does not hold, then not only is A_n^+ inconsistent for A^+, but the error of estimation is unboundedly large as $n \to \infty$.

<u>Proof of Theorem 1</u>. Part (a) follows immediately from Lemma 1 by a sample path by sample path argument.

The first equivalence of part (b) is obtained as follows: $A_n^+ \xrightarrow{p} A$ iff every subsequence of $\{n\}$ has a sub-subsequence $\{n_k\}$ such that $A_{n_k} \xrightarrow{k\to\infty} A$ a.s. (e.g., see Lukacs (1968, Theorem 2.4.4)) iff every subsequence of $\{n\}$ has a sub-subsequence $\{n_k\}$ such that $1(rk[A_{n_k}] = rk[A]) \xrightarrow{k\to\infty} 1$ a.s. (where $1(\cdot)$ denotes the indicator function) iff $P(rk[A_n] = rk[A]) \xrightarrow{n\to\infty} 1$.

To show the second equivalence of part (b), first suppose $A_n^+ \xrightarrow{p} A^+$. For $M > \|A\|$, $P(\|A_n\| \le M) \ge P(\|A_n - A\| + \|A\| \le M) \xrightarrow{n\to\infty} 1$. For the converse, suppose $P(\|A_n\| < M) \xrightarrow{n\to\infty} 1$, for some $M < \infty$. Then, every subsequence of $\{n\}$ has a sub-subsequence on which $1(\|A_n\| \le M) \xrightarrow{n\to\infty} 1$ a.s., and hence, on which $\lim_{n\to\infty} \sup\|A_n^+\| < \infty$ a.s., and by Lemma 1, on which $A_n^+ \xrightarrow{n\to\infty} A^+$ a.s. The latter implies $A_n^+ \xrightarrow{p} A$. □

3. Asumptotic Distributions of g-Inverse Quadratic Forms

In this section we consider the asymptotic distribution of g-inverse quadratic forms given by $X_n'A_n^-X_n$, where X_n is an asymptotically normal real random vector with singular covariance matrix A, and A_n is a conformable random matrix that converges in probability to A. Quadratic forms of this type are used to construct generalized Wald tests, including Hausman specification tests and chi-square goodness of fit tests, among others. The rank condition and results of Section 2 above are used to determine sufficient conditions for convergence of the g-inverse quadratic form to a chi-square random variable.

We introduce the following conditions:

C1 $X_n \xrightarrow{d} X \sim N(\mu,A)$ as $n \to \infty$,

C2 $A_n \xrightarrow{p} A$ as $n \to \infty$,

C3 $P(rk[A_n] = rk[A]) \to 1$ as $n \to \infty$,

C4 $P(X_n \in \mathcal{M}(A_n)) \to 1$ as $n \to \infty$,

C5 $P(A_n$ is symmetric$) \to 1$ as $n \to \infty$,

where "$\xrightarrow{d}$" denotes convergence in distribution and $\mathcal{M}(A_n)$ denotes the column space of A_n. Also, let $\chi^2(rk[A],\delta)$ denote a non-central chi-square random variable (or distribution) with $rk[A]$ degrees of freedom and non-centrality parameter δ.

The main result is the following:

Theorem 2. (a) Suppose C1-C3 hold. Then,

$$X_n'A_n^+X_n \xrightarrow{d} \chi^2(rk[A],\delta), \quad \text{where} \quad \delta = \mu'A^+\mu.$$

(b) Suppose C1-C5 hold. Then,

$$X_n'A_n^-X_n \xrightarrow{d} \chi^2(rk[A],\delta),$$

for all sequences of g-inverses $\{A_n^-\}$ of $\{A_n\}$.

(c) If A is non-singular, none of the conditions in parts (a) and (b) is redundant.

Comments: 1. Part (c) does not assert that each condition used in parts (a) and (b) is necessary. Rather, it makes the assertion that the if-then statement given in each part does not hold when any of the conditions is dropped.

2. Part (c) shows that some results in the literature are not completely accurate. It is not sufficient to establish C1 and C2, or C1, C2, C4, and C5, in order to assert that a g-inverted quadratic form has a chi-square asymptotic distribution.

3. The proof of part (b) actually shows that $X_n'A_n^-X_n$ is numerically identical for all choices of g-inverse of A_n with probability that converges to one as $n \to \infty$.

The proof of Theorem 2 uses the following lemma. The result of the Lemma is fairly well-known, e.g., see Rao and Mitra (1972, p. 615). For completeness, however, we include the Lemma and its brief proof.

Lemma 2. Given any fixed vector X and any fixed, conformable, symmetric matrix H, if $X \in M(H)$, then $X'H^-X$ is numerically identical for all choices of g-inverse.

Proof of Lemma 2. Since $X \in M(H)$, there exists some vector Z such that $X = HZ$. Hence, $X'H^-X = Z'H'H^-HZ = Z'HZ$, using the symmetry of H. The right-hand-side is independent of the choice of H^-. □

Proof of Theorem 2. Under the assumptions, Theorem 1 and the continuous mapping theorem give $X_n'A_n^+X_n \xrightarrow{d} X'A^+X$ as $n \to \infty$. The result of part (a) follows, since Theorem 9.2.3 of Rao and Mitra (1971) implies that $X'A^{+}X \sim \chi^2(rk[A],\delta)$.

By Lemma 2 and assumptions C4 and C5, $X_n'A_n^{-}X_n - X_n'A_n^{+}X_n \xrightarrow{p} 0$ as $n \to \infty$, for all g-inverses A_n^{-} of A_n. Hence, part (b) is proved.

Next we prove part (c). It is clear that neither C1 nor C2 is redundant in part (a) or part (b). To show that C3 is not redundant in part (a) or part (b), let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $A_n = \begin{bmatrix} 1 & 1/n \\ 1/n & 0 \end{bmatrix}$, $\forall n$, and $X_n = X + (0,1/\sqrt{n})'$, where $X = (X_1,0)' \sim N(0,A)$. The conditions C1, C2, C4, and C5 all hold. But, $A_n^{-1} = \begin{bmatrix} 0 & n \\ n & -n^2 \end{bmatrix}$ and $X_n'A_n^{-1}X_n = 2\sqrt{n}X_1 - n \stackrel{d}{\nrightarrow} \chi^2(rk[A],0)$ as $n \to \infty$.

To show that C4 is not redundant in part (b), take A and X_n as above, and let $A_n = A$, for all n. Conditions C1-C3 and C5 are satisfied. Consider the g-inverse $A_n^{-} = \begin{bmatrix} 1 & 0 \\ 0 & n^2 \end{bmatrix}$ of A_n. For this g-inverse, $X_n'A_n^{-}X_n = X_1^2 + n \stackrel{d}{\nrightarrow} \chi^2(rk[A],0)$ as $n \to \infty$.

Lastly, to show that C5 is not redundant in part (b), take A and X as above, and let $X_n = X$ and $A_n = \begin{bmatrix} 1 & 1/n \\ 0 & 0 \end{bmatrix}$, for all n. Conditions C1-C4 are satisfied. For the g-inverse $A_n^{-} = \begin{bmatrix} n & 0 \\ n(1-n) & 0 \end{bmatrix}$, $X_n'A_n^{-}X_n = nX_1^2 \stackrel{d}{\nrightarrow} \chi^2(rk[A],0)$ as $n \to \infty$. □

4. Conclusion

The use of g-inverse quadratic forms is becoming frequent in econometrics. This note provides results concerning the asymptotic distributions of such statistics. In addition, it considers estimation of Moore-Penrose inverted matrices. Consistent estimation of Moore-Penrose inverted matrices is shown to require the fulfillment of a rank condition. If this condition is not fulfilled, not only does inconsistency result, but the error of estimation is unboundedly large as the sample size increases. The rank condition also is shown to be important in deriving the asymptotic distributions of g-inverse quadratic forms. If the rank condition holds, then g-inverse quadratic forms have χ^2 asymptotic distributions under weak additional conditions. These distributional results are useful for determining asymptotic significance levels and local power properties of generalized Wald tests.

FOOTNOTE

[1] I am greatly indebted to a referee for pointing out several errors in an earlier draft of this paper. I also would like to thank Steven N. Durlauf, Whitney K. Newey, Peter C. B. Phillips, and two referees for their helpful comments and suggestions. The research support of the National Science Foundation, through grant no. SES-8419789, is gratefully acknowledged.

References

Andrews, D.W.K., 1985a, Random cell chi-square diagnostic tests for econometric models: I. Introduction and applications, Cowles Foundation Discussion Paper No. 762, Yale University, New Haven, CT.

Andrews, D.W.K., 1985b, Random cell chi-square diagnostic tests for econometric models: II. Theory, Cowles Foundation Discussion Paper No. 763, Yale University, New Haven, CT.

Duncan, G.M., 1983, A generalized Durbin-Hausman test, unpublished manuscript, Department of Economics, Washington State University, Pullman, Washington.

Hausman, J.A. and W.E. Taylor, 1981, A generalized specification test, Economic Letters 8, 239-245.

Heckman, J.J., 1984, The χ^2 goodness of fit statistic for models with parameters estimated from microdata, Econometrica 52, 1543-1547.

Holly, A., 1982, A remark on Hausman's specification test, Econometrica 50, 749-759.

Lukacs, E., 1968, Stochastic convergence (D.C. Heath, Lexington, MA).

Mitra, S.K., 1980, Generalized inverse of matrices and their application to linear models, Handbook of Statistics 1, 471-512.

Moore, D.S., 1977, Generalized inverses, Wald's method, and the construction of chi-squared tests of fit, Journal of the American Statistical Association 72, 131-137.

Newey, W.K., 1984, Generalized method of moments specification testing, Discussion Paper, Department of Economics, Princeton University.

Rao, C.R., 1973, Linear statistical inference and its applications, second edition (J. Wiley, New York).

Rao, C.R. and S.K. Mitra, 1971, Generalized inverse of matrices and its applications (J. Wiley, New York).

Rao, C.R. and S.K. Mitra, 1972, Generalized inverse of a matrix and its applications, Sixth Berkeley Symposium on Mathematical Statistics and Probability 1, 601-620.