APPLICATIONES MATHEMATICAE
32,1 (2005), pp. 51–55

G. Forchini (Clayton)

A NOTE ON THE CONTINUITY OF PROJECTION MATRICES WITH APPLICATION TO THE ASYMPTOTIC DISTRIBUTION OF QUADRATIC FORMS

Abstract. This paper investigates the continuity of projection matrices and illustrates an important application of this property to the derivation of the asymptotic distribution of quadratic forms. We give a new proof and an extension of a result of Stewart (1977).

An important result in statistics concerning the distribution of quadratic forms is the following: if X is a $k \times 1$ vector having a multivariate normal distribution with mean vector μ and identity covariance matrix, and if P_B is an idempotent matrix of rank p then $X'P_BX$ is $\chi^2(p,\delta)$, where $\delta = \mu' P_B \mu$ (see for instance Muirhead (1982), Theorem 1.4.5).

Sometimes we are interested in an asymptotic version of this result (examples are given below): (i) the $k \times 1$ random vector X_T indexed by T converges in distribution to a multivariate normal random variable with unknown mean μ and identity covariance matrix I_k; (ii) the $k \times n$ $(n \le k)$ matrix $\widehat{B}_T$ converges in probability to the $k \times n$ matrix B (and we write $\operatorname{plim}_{T\to\infty} \widehat{B}_T = B$). In this application $\widehat{B}_T$ has rank n with probability 1 for all but a finite number of T. Let $P_{\widehat{B}_T} = \widehat{B}_T(\widehat{B}_T'\widehat{B}_T)^{-1}\widehat{B}_T'$. We want to find the asymptotic distribution of $X_T' P_{\widehat{B}_T} X_T$ as $T \to \infty$. If the mapping $\widehat{B}_T \to P_{\widehat{B}_T}$ is continuous we can conclude that $X_T' P_{\widehat{B}_T} X_T$ has an asymptotically noncentral chi-square distribution with n degrees of freedom and noncentrality parameter $\mu' P_B \mu$ (Muirhead (1982), Theorem 1.4.5). However, if $n > p = \operatorname{rank}(B)$ for all but a finite number of T, then $X_T' P_{\widehat{B}_T} X_T$ often has an asymptotically noncentral chi-square distribution with (again) n degrees of freedom (rather than p) and noncentrality parameter $\mu'(\operatorname{plim}_{T\to\infty} P_{\widehat{B}_T})\mu$. Two examples where such situation arises are given below.

2000 *Mathematics Subject Classification*: 54C05, 15A03, 62E20.
Key words and phrases: projection matrix, distribution of quadratic forms.

EXAMPLE 1. Consider a single linear structural equation

$$y_1 = Y_2\beta + Z_1\gamma + u, \tag{1}$$

where y_1 and Y_2 are a $T \times 1$ vector and a $T \times n$ matrix of endogenous variables, Z_1 is a $T \times k_1$ matrix of exogenous variables, and β and γ are $n \times 1$ and $k_1 \times 1$ vectors of parameters. The reduced form equation for Y_2 is

$$Y_2 = Z_1\varPhi_2 + Z_2\varPi_2 + V_2, \tag{2}$$

where Z_2 is a $T \times k_2$ matrix of exogenous variables not included in the structural equation, and $\varPhi_2$ and $\varPi_2$ are matrices of parameters of dimension $k_1 \times n$ and $k_2 \times n$ respectively. We assume throughout that $k_2 > n$. Inserting the reduced form (2) into the structural equation (1) gives

$$y_1 = Z_1\phi_1 + Z_2\pi_1 + v_1, \tag{3}$$

where the parameter (ϕ_1, π_1) and the error term v_1 satisfy the *compatibility conditions*:

$$\pi_1 = \varPi_2\beta, \tag{4}$$

$$\phi_1 = \gamma + \varPhi_2\beta, \tag{5}$$

$$v_1 = u + V_2\beta. \tag{6}$$

Together, equations (2) and (3) form a multivariate linear model (MLM)

$$[y_1, Y_2] = Z_1[\phi_1, \varPhi_2] + Z_2[\pi_1, \varPi_2] + [v_1, V_2] \tag{7}$$

with restrictions on its coefficients and its error components. Assume that the following limits hold jointly:

(a) $\widehat{Q}_T = T^{-1}Z_2'M_{Z_1}Z_2$ and $\operatorname{plim}_{T\to\infty}\widehat{Q}_T = Q$, where Q is a fixed, finite, positive definite $k_2 \times k_2$ matrix, and $M_{Z_1} = I_T - Z_1(Z_1'Z_1)^{-1}Z_1'$;

(b) $\widehat{\varOmega}_T = T^{-1}S$ and $\operatorname{plim}_{T\to\infty}\widehat{\varOmega}_T = \varOmega$, where $S = Y'M_ZY$, $Y = [y_1, Y_2]$, $Z = [Z_1, Z_2]$ and $M_Z = I_T - Z(Z'Z)^{-1}Z'$;

(c) $T^{\frac{1}{2}}(\widehat{\varPi} - \varPi) \xrightarrow{L} N(0, Q^{-1} \otimes \varOmega)$, where $\varPi = [\pi_1, \varPi_2]$ and $\widehat{\varPi}$ is the OLS estimator of $\varPi$ in (7);

and suppose one is interested in the vector of coefficients β, and wants to test whether equation (4) is satisfied. In a generalized method of moments framework, one can base such a test on the statistic

$$\mathcal{T} = \widetilde{\pi}_1'(I_{k_2} - P_{\widetilde{\varPi}_2})\widetilde{\pi}_1 = \widehat{u}'Z(Z'Z)^{-1}Z'\widehat{u},$$

where $[\widetilde{\pi}_1, \widetilde{\varPi}_2] = \widehat{Q}_T^{1/2}[\widehat{\pi}_1, \widehat{\varPi}_2]$ and $[\widehat{\pi}_1, \widehat{\varPi}_2]$ is the ordinary least squares estimator of $[\pi_1, \varPi_2]$ in (7). Forchini (2003) shows that if (4) holds then

$$\frac{T\mathcal{T}}{\omega_{11.2}(1 + (\beta_1^*)'\beta_1^*)} \xrightarrow{L} \chi^2(k_2 - n)$$

independently of the rank of Π_2. In the expression above we have used

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{21}' \\ \omega_{21} & \Omega_{22} \end{pmatrix},$$
$$\omega_{11.2} = \omega_{11} - \omega_{12}\Omega_{22}^{-1}\omega_{21},$$
$$\beta^* = (\Omega_{22}^{1/2}\beta - \Omega_{22}^{-1/2}\omega_{21})/\omega_{11.2}^{1/2},$$

and $\beta^* = (\beta_1^{*\prime}, \beta_2^{*\prime})'$ is partitioned conformably to

$$\pi_1 = (\Pi_{21}, \Pi_{22}) \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$

where $\operatorname{rank}(\Pi_2) = \operatorname{rank}(\Pi_{21}) = r \leq n$.

EXAMPLE 2. A similar situation occurs with asymptotically uncooperative regressors (Schmidt (1976), pp. 85–88): consider the case

$$Y_T \sim W_T\beta + U_T, \qquad U_T \sim N(0, \sigma^2 I_T),$$

where the ith row of W_T is $(1, \lambda^i)$ with $|\lambda| < 1$ and σ^2 known. When testing whether $\beta = 0$ one considers the statistic

$$\sigma^{-2}U_T' P_{W_T} U_T = X_T' X_T,$$

where $X_T = \sigma^{-1}(W_T'W_T)^{-1/2}W_T'U_T \sim N(0, \sigma^2 I_2)$. This quantity has a chi-square distribution with 2 ($= \operatorname{rank}(P_{X_T})$) degrees of freedom. Note however that $T^{-1}W_T'W_T \to \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ so that $\operatorname{rank}(P_{\lim_{T\to\infty} X_T}) = 1$.

These results are surprising: the limiting distribution of $X_T' P_{\widehat{B}_T} X_T$ does not depend on p. To understand why, we need to look at the properties of the mapping $\widehat{B}_T \mapsto P_{\widehat{B}_T}$. Let B_T be a sequence of $k \times n$ ($n \leq k$) matrices converging to the $k \times n$ matrix B, and let $P_{B_T} = B_T B_T^\dagger$ and $P_B = BB^\dagger$ be the projections on the spaces spanned by the columns of B_T and B respectively. The matrices $B_T^\dagger$ and $B^\dagger$ are the Moore–Penrose generalised inverses of B_T and B.

Corollary 3.5 of Stewart (1977) shows that a necessary and sufficient condition for $\lim_{T\to\infty} B_T^\dagger = B^\dagger$ is that $\operatorname{rank}(B_T) = \operatorname{rank}(B)$ for all but a finite number of T. Given the uniqueness of the Moore–Penrose generalised inverse, it follows that a necessary and sufficient condition for $\lim_{T\to\infty} P_{B_T} = P_B$ is that $\operatorname{rank}(B_T) = \operatorname{rank}(B)$ for all but a finite number of T. In this note we give a new direct proof of this result, and give an indication of what happens when $\operatorname{rank}(B_T) > \operatorname{rank}(B)$ for all but a finite number of T.

Precisely, we have:

PROPOSITION 1. *Let B_T be a sequence of matrices converging to the matrix B.*

(i) $\lim_{T\to\infty} P_{B_T} = P_B$ *if and only if* $\operatorname{rank}(B_T) = \operatorname{rank}(B)$ *for all but a finite number of* T.

(ii) *If* $\operatorname{rank}(B_T) > \operatorname{rank}(B)$ *for all but a finite number of* T, *then the space spanned by the columns of* B *is contained in the space spanned by* $\lim_{T\to\infty} P_{B_T}$.

Thus, the space spanned by the columns of $\operatorname{plim}_{T\to\infty} P_{\widehat{B}_T}$ is larger than the space spanned by the columns of $P_B = P_{\operatorname{plim}_{T\to\infty}\widehat{B}_T}$. The degrees of freedom of the resulting chi-square distribution are not affected by the rank of B because the space spanned by the columns of $P_{\widehat{B}_T}$ does not collapse to that spanned by the columns of B.

Proof of Proposition 1. Consider a sequence $B_T \to B$. Without loss of generality we can assume that $\operatorname{rank}(B_T) = g \geq p = \operatorname{rank}(B)$. Then write

$$B_T B_T' = H_T \begin{pmatrix} \Lambda_T & 0 \\ 0 & 0 \end{pmatrix} H_T',$$

where H_T is a $k \times k$ orthogonal matrix and Λ_T is a $g \times g$ matrix which contains the g nonzero eigenvalues of $B_T B_T'$ in descending order. Partition H_T as $H_T = (H_{1T}, H_{2T})$, where H_{1T} is $k \times g$ and H_{2T} is $k \times (k-g)$. The space spanned by the columns of B_T is spanned by the eigenvectors corresponding to $B_T B_T'$, i.e. the columns of H_{1T}, so we can write

$$P_{B_T} = H_{1T} H_{1T}'.$$

Now, $P_{B_T} \to P_B$ if and only if every subsequence of P_{B_T} has a subsequence which converges to P_B.

Now consider an arbitrary subsequence $P_{B_{T_q}}$. The elements of H_{T_q} and Λ_{T_q} are bounded above uniformly in T, so there is a subsequence such that $H_{T_{qs}}$ and $\Lambda_{T_{qs}}$ converge to some matrices $H = (H_1, H_2)$ and Λ, and $B = H\left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right)H'$, where H is orthogonal and Λ is diagonal with g nonzero diagonal elements.

Consider the case where $g = p$. Then $H_{T_{qs}} \to H$ and $H_{1T_{qs}} \to H_1$, $\Lambda_{T_{qs}} \to \Lambda$ $(p \times p)$ so that as $q \to \infty$,

$$P_{B_{T_{qs}}} = H_{1T_{qs}} H_{1T_{qs}}' \to H_1 H_1' = P_B.$$

If $g > p$, it is still true that $H_{T_{qs}} \to H$ and $H_{1T_{qs}} \to H_1$. Moreover $\Lambda_{1T_{qs}} \to \Lambda = \left(\begin{smallmatrix} \Lambda_* & 0 \\ 0 & 0 \end{smallmatrix}\right)$ so that

$$P_{B_{T_{qs}}} = H_{1T_{qs}} H_{1T_{qs}}' \to H_1 H_1'.$$

But P_B is the projection on the space spanned by the eigenvector associated to the nonzero eigenvalues of BB', so partitioning $H_1 = (H_{11}, H_{12})$, where H_{11} is $k \times p$ and H_{12} is $k \times (k-p)$, we have

$$P_B = H_{11} H_{11}'$$

so that

$$\lim_{T\to\infty}(P_{B_T} - P_B) = H_{12}H'_{12} \neq 0.$$

Note that $H_{12}H'_{12}$ is itself a projection into the space spanned by the columns of B_T and orthogonal to the space spanned by the columns of B. ∎

References

G. Forchini (2003), *Testing the relevance of instrumental variables in structural estimation*, preprint.

R. J. Muirhead (1982), *Aspects of Multivariate Statistical Theory*, Wiley, New York.

P. Schmidt (1976), *Econometrics*, Dekker, New York.

G. W. Stewart (1977), *On the perturbation of pseudo-inverses, projections and linear least squares problems*, SIAM Rev. 19, 634–662.

Department of Econometrics and Business Statistics
Faculty of Business and Economics
Monash University
Clayton, VIC 3800, Australia

Received on 8.9.2003;
revised version on 7.10.2004 (1706)