An inequality between perpendicular least-squares and ordinary least-squares

Boswijk, P.; Neudecker, H.

DOI
10.1017/S0266466600008537

Publication date
1994

Document Version
Final published version

Published in
Econometric Theory

94.2.4. *An Inequality Between Perpendicular Least-Squares and Ordinary Least-Squares*, proposed by H. Peter Boswijk and Heinz Neudecker. Let $Z = [Z_1 : Z_2]$ be an $n \times p$ matrix with $n > p$, and with Z_i matrices of order $n \times p_i$, $i = 1,2$. Define $M_2 := I_n - Z_2 Z_2^+$, where Z_2^+ is the Moore–Penrose inverse of Z_2; if $\operatorname{rank}(Z_2) = p_2$, then $M_2 = I_n - Z_2(Z_2'Z_2)^{-1}Z_2'$. Moreover, let $\lambda_1 \le \cdots \le \lambda_p$ denote the eigenvalues of $Z'Z$.

1. Prove that

$$\sum_{i=1}^{p_1} \lambda_i \le \operatorname{tr}(Z_1'M_2Z_1), \tag{1}$$

where $\operatorname{rank}(Z_2) = l \le p_2$.

2. Let S denote a $p \times p$ random matrix with a standard central Wishart distribution with n degrees of freedom, i.e., $S \sim W(I_p, n)$, let $\mu_1 \le \cdots \le \mu_p$ denote its eigenvalues, and define

$$s_1 := \sum_{i=1}^{p_1} \mu_i. \tag{2}$$

Finally, let s_2 denote a random variable with a χ^2 distribution with $p_1(n - p_2)$ degrees of freedom. Use (1) to prove that s_1 is stochastically dominated by s_2, i.e., for any $c > 0$,

$$P\{s_1 > c\} \le P\{s_2 > c\}. \tag{3}$$

Remarks.

(i) The right-hand side of (1) is equal to the trace of the residual sum of squares from a multivariate ordinary least-squares (OLS) regression of Z_1 on Z_2. The left-hand side is the corresponding quantity for a multivariate version of perpendicular least-squares (PLS). Hence, (1) entails that PLS yields a smaller residual sum of squares than OLS.

(ii) An application of (3) can be found in reduced rank regression, cf. Anderson [1]. Consider the multivariate normal linear regression $Y = XB + U$, with $Y, U: n \times g$, $X: n \times k$, and $B: k \times g$, and with $n > g \ge k$. The likelihood ratio (LR) statistic for the hypothesis $\operatorname{rank}(B) \le r$, $r < k$, has an asymptotic χ^2 distribution with $(g - r)(k - r)$ degrees of freedom under the null hypothesis, provided that $\operatorname{rank}(B) = r$. On the other hand, if $\operatorname{rank}(B) = q < r$, then the LR statistic has the same asymptotic distribution as the sum of the $(k - r)$ smallest eigenvalues of a $W(I_{k-q}, g - q)$ matrix. Result (3) shows that the latter distribution is more concentrated towards the origin, so that if χ^2_α is the $100\alpha\%$ critical value of the $\chi^2((g - r)(k - r))$ distribution, then under the null hypothesis and as $n \to \infty$, $P\{LR > \chi^2_\alpha\} \to \alpha_q \le \alpha$. Thus the size of the test is controllable (cf. Cragg and Donald [2, Theorem 2] for a related result).

REFERENCES

1. Anderson, T.W. Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics* 22 (1951): 327–351.
2. Cragg, J.G. & S.G. Donald. Testing identifiability and specification in instrumental variable models. *Econometric Theory* 9 (1993): 223–240.