Biometrika (1988), **75**, 3, *pp.* 611–5
Printed in Great Britain

Duality in testing multivariate hypotheses

BY FRANK A. WOLAK
Department of Economics, Stanford University, Stanford, California 94305, U.S.A.

SUMMARY

This paper derives a duality result for a general class of hypothesis testing problems in multivariate analysis utilizing the relationship between convex cones and their polar cones together with the properties of minimum norm problems between points and cones in Euclidian space. Special cases of this result yield generalizations of a well-known duality relation in multivariate equality constraints testing. For example, any multivariate inequality constraints test on the parameters of a multivariate normal random vector has an equivalent multivariate one-sided test in terms of the vector of dual variables associated with the constraints. Also, any combination multivariate inequality and equality constraints test has an equivalent combination multivariate one-sided and two-sided test in terms of the vector of dual variables associated with both sets of constraints.

Some key words: Convex cone; Duality in multivariate hypothesis testing; Multivariate inequality constraints test; Multivariate one-sided test; Polar cone.

1. INTRODUCTION

This paper extends the classical duality result in testing multivariate equality constraints to a general class of testing problems studied by Perlman (1969): H: $\mu \in P_1$ versus K: $\mu \in P_2$, where P_1 and P_2 are positively homogeneous sets with $P_1 \subset P_2$ and μ is the mean of a multivariate normal random vector.

This paper proceeds as follows. In this section we state a familiar special case of our duality result for tests of multivariate equality constraints to provide the appropriate background for subsequent sections and to motivate our work from the viewpoint of empirical multivariate hypothesis testing. In § 2 we state and prove our general duality result. In § 3 we present applications of this result to several testing problems in empirical multivariate analysis.

To state the duality relation in equality constraints testing, let β be a $k \times 1$ vector and $\hat{b} = \beta + \eta$, where η is $N(0, V)$ and V is a known $k \times k$ positive-definite matrix. Let R be a $p \times k$ matrix of full row rank where $p \leq k$ and r a $p \times 1$ vector. Define a quadratic form function $Q(x, M) = \frac{1}{2}x'M^{-1}x$, for $x \in R^n$ and M a positive-definite $n \times n$ matrix.

The linear equality constraints testing problem in primal form is H: $R\beta = r$ versus K: $R\beta \neq r$, with likelihood ratio statistic

$$T_{\text{LR}} = 2 \inf_b [Q\{(\hat{b} - b), V\} \text{ subject to } Rb = r]. \tag{1·1}$$

The optimization problem defining this statistic is a simple quadratic program. The general procedure for solving quadratic programs entails constructing the Lagrangian $Z = Q\{(\hat{b} - b), V\} + L'(r - Rb)$ and computing the saddlepoint of this function in b and L. The saddlepoint is defined as the minimum of Z with respect to b and maximum with respect to L. Bazaraa & Shetty (1979) provide a thorough discussion of the theory of quadratic programming. From the first-order conditions for (1·1) we have $\bar{L} = S(r - R\hat{b})$, where $\bar{L}$ is the optimal L and $S = (RVR')^{-1}$. Define λ as the expectation of $\bar{L}$. This implies the following relation between λ and β:

$$\lambda \equiv E(\bar{L}) = E\{S(r - R\hat{b})\} = S(r - R\beta). \tag{1·2}$$

Given this definition of λ, the equivalent dual test to H: $R\beta = r$ versus K: $R\beta \neq r$ in terms of β is H: $\lambda = 0$ versus K: $\lambda \neq 0$, based on $\bar{L}$, with likelihood ratio statistic $T_{\mathrm{LM}} = 2Q(\bar{L}, S)$. This duality relation follows because $\bar{L}$ is $N(0, S)$ under the two equivalent null hypotheses $R\beta = r$ and $\lambda = 0$, and $T_{\mathrm{LR}} = T_{\mathrm{LM}}$ by definition of $\bar{L}$ in terms of $\hat{b}$. Although the dual test is in terms of λ and $\bar{L}$ and the primal test is in terms of β and $\hat{b}$, both are equivalent tests of H: $R\beta = r$ versus K: $R\beta \neq r$.

2. A DUALITY RELATION IN MULTIVARIATE HYPOTHESIS TESTING

We first introduce some results from convex analysis. If C is a closed convex cone in R^k then its polar cone C^0 is defined as

$$\{w^* \in R^k \mid \langle w, w^* \rangle \leq 0, \text{ for all } w \in C\},$$

where $\langle w, w^* \rangle$ is the inner product of two vectors in R^k. For most applications of our duality result, for example, those in § 3, C is a polyhedral cone in R^k, which is closed and convex by construction and has the general form $C = \{w \in R^k \mid Bw \geq 0\}$, where B is an $n \times k$ matrix and n is any positive integer. The polar cone to C is

$$C^0 = \{y \mid y = -B'z,\ z \in R^n,\ z_i \geq 0,\ i = 1, \ldots, n\}.$$

We now state and prove our general duality result.

THEOREM 1. *Let C be a closed, convex cone in R^k, with polar cone C^0. Let Ω be a known $k \times k$ positive-definite matrix and ξ and γ vectors in R^k. Consider the likelihood ratio statistics for the hypothesis tests*: (P), *H: $\xi \in C$ versus K: $\xi \in R^k$ based on one observation from x which is $N(\xi, \Omega)$, and* (D), *H: $\gamma = 0$ versus K: $\gamma \in C^0$ based on one observation from y which is $N(\gamma, \Omega^{-1})$. The following results hold*: (i) *if $y = \Omega^{-1}x$ then the two likelihood ratio statistics are equal, and* (ii) *if $\gamma = \Omega^{-1}\xi$ then the two test statistics have the same distribution.*

Note that (i) is a statement about the relation between x and y as vectors in R^k which leads to an equality between two test statistics; (ii) specifies a relation between x and y as random vectors in R^k and yields an equality between two distributions.

Proof of (i). For any $x \in R^k$, let $f(z \mid x) = Q\{(x - z), \Omega\}$, $z \in R^k$. The conjugate function to $f(z \mid x)$ is

$$f^*(z^* \mid x) = [Q\{(\Omega^{-1}x + z^*), \Omega^{-1}\} - Q(x, \Omega)], \quad z^* \in R^k.$$

By Theorem 31.4 of Rockafellar (1970),

$$\inf\{f(z \mid x) \mid z \in K\} = -\inf\{f^*(z^* \mid x) \mid z^* \in K^*\},$$

for any nonempty closed convex cone K and its negative polar K^*, where $K^* = \{-z^* \mid z^* \in K^0\}$. The likelihood ratio statistic for (P) is, following the logic of Perlman (1969), $T_{\mathrm{P}} = 2 \inf[Q\{(x - e), \Omega\} \mid e \in C]$. Our conjugate function notation implies $T_{\mathrm{P}} = 2 \inf\{f(e \mid x) \mid e \in C\}$. Using Perlman's approach, the likelihood ratio statistic for (D) is

$$T_{\mathrm{D}} = 2 \sup[Q(y, \Omega^{-1}) - Q\{(y - g), \Omega^{-1}\} \mid g \in C^0].$$

This can be rewritten using the definition of C^* as

$$T_{\mathrm{D}} = -2 \inf[Q\{(y + g), \Omega^{-1}\} - Q(y, \Omega^{-1}) \mid g \in C^*].$$

If the two vectors satisfy $y = \Omega^{-1}x$, our new notation yields $T_{\mathrm{D}} = -2 \inf\{f^*(g \mid x) \mid g \in C^*\}$, so that $T_{\mathrm{P}} = T_{\mathrm{D}}$ by the results of Rockafellar (1970) stated above.

The proof of (ii) follows by rewriting the two statistics as:

$$T_{\mathrm{P}} = 2 \inf[Q\{(\Omega^{-\frac{1}{2}}x - \Omega^{-\frac{1}{2}}e), I\} \mid e \in C], \quad T_{\mathrm{D}} = 2 \sup[Q(\Omega^{\frac{1}{2}}y, I) - Q\{(\Omega^{\frac{1}{2}}y - \Omega^{\frac{1}{2}}g), I\} \mid g \in C^0],$$

where I is a $k \times k$ identity matrix. If $\gamma = \Omega^{-1}\xi$ then both $\Omega^{-\frac{1}{2}}x$ and $\Omega^{\frac{1}{2}}y$ are $N(\Omega^{-\frac{1}{2}}\xi, I)$ random vectors. Therefore T_{P} and T_{D} have the same distribution because they are equivalent quadratic forms based on the same multivariate normal random vector. □

3. Applications of duality result

In this section we apply the duality result to several multivariate hypothesis tests. Consider the test of

$$H\colon R\beta \geqslant r \text{ versus } K\colon \beta \in R^k \tag{3·1}$$

in the notation and under the distributional assumptions for $\hat{b}$ given in § 1. Wolak (1987) discusses several applications of (3·1) for the linear regression model. There the likelihood ratio statistic for (3·1) is given by

$$T_{\text{IU}} = 2 \inf_b [Q\{(\hat{b}-b), V\} \text{ subject to } Rb \geqslant r], \tag{3·2}$$

where IU is shorthand for an inequality constraints null hypothesis versus an unrestricted alternative. The solution of quadratic program (3·2) gives rise to a vector of Kuhn–Tucker multipliers, $\tilde{L} \geqslant 0$, associated with the inequality constraints $Rb \geqslant r$. Using the definition of $\bar{L}$ in terms of $\hat{b}$, the dual of quadratic program (3·2) can be written as

$$T_{\text{EI}} = 2 \sup_L [Q\{\bar{L}, S\} - Q\{(\bar{L}-L), S\} \text{ subject to } L \geqslant 0]. \tag{3·3}$$

The optimal value of the objective function from (3·3) is equal to the likelihood ratio statistic for the test $H\colon \lambda = 0$ versus $K\colon \lambda \geqslant 0$, for the λ defined in (1·2). In this case the subscript EI stands for an equality constraints null hypothesis versus an inequality constrained alternative. The solution to (3·3) is $\tilde{L}$ defined above. Quadratic programs (3·2) and (3·3) satisfy the regularity conditions of Gill, Murray & Wright (1981, p. 76) for the validity of duality theory, so that $T_{\text{IU}} = T_{\text{EI}}$. Wolak (1987) shows that the value of β which maximizes $\text{pr}(T_{\text{IU}} \geqslant c)$ under the null hypothesis for any $c > 0$ is any β satisfying $R\beta = r$. This is the least favourable β used to compute critical values for any size test of (3·1). This β implies the λ defined by (1·2) is zero, which suggests that the equivalent dual test to (3·1) in terms of β and $\hat{b}$ is a multivariate one-sided test in terms of λ and $\bar{L}$. After stating the null distributions for this multivariate one-sided test and hypothesis test (3·1), we show the equivalence of these two tests is a special case of Theorem 1.

One benefit of this duality relation is that the null distribution for any size test of (3·1) can be obtained from the results of the multivariate one-sided test derived by Kudô (1963), Nuesch (1966) and Perlman (1969) applied to hypothesis test (3·1) in terms of λ. These results yield:

$$\text{pr}_{0,S}(T_{\text{EI}} \geqslant c) = \sum_{j=0}^{p} \text{pr}\,(\chi_j^2 \geqslant c)\, w(p, j, S), \tag{3·4}$$

where the notation $\text{pr}_{0,S}\,(T_{\text{EI}} \geqslant c)$ is the probability of $(T_{\text{EI}} \geqslant c)$ given that $\bar{L}$ is $N(\lambda, S)$ with $\lambda = 0$, and $\text{pr}\,(\chi_k^2 \geqslant c)$ is the probability a chi-squared random variable with k degrees of freedom is greater than or equal to c. Wolak (1987) derives the least favourable null distribution for any size test of (3·1) by the primal approach as

$$\sup_\beta \{\text{pr}_{\beta,V}\,(T_{\text{IU}} \geqslant c) \mid R\beta \geqslant r\} = \text{pr}_{\beta^*,V}\,(T_{\text{IU}} \geqslant c)$$

$$= \sum_{j=0}^{p} \text{pr}\,(\chi_j^2 \geqslant c)\, w(p, p-j, S^{-1}), \tag{3·5}$$

where β^* is any β such that $R\beta^* = r$. The weight function, $w(p, j, \Omega)$, used in (3·4) and (3·5) is the probability that $\tilde{\mu}$ has exactly j positive elements, where $\tilde{\mu}$ is the solution to

$$\inf_\mu [Q\{(\hat{\mu}-\mu), \Omega\} \text{ subject to } \mu \geqslant 0]$$

and $\hat{\mu} \in R^n$ is $N(0, \Omega)$. Wolak (1987) states the following relation for the weight function: $w(p, j, \Omega) = w(p, p-j, \omega\Omega^{-1})$ $(j = 0, \ldots, p)$ for all p, positive-definite Ω, and $\omega > 0$. Therefore the two weighted sums of chi-squared distributions given in (3·4) and (3·5) are identical. This leads to the following corollary which is a special case of Theorem 1.

COROLLARY 1. *The test of H: $R\beta \geq r$ versus K: $\beta \in R^k$ based on $\hat{b}$, which is $N(\beta, V)$, is equivalent to, for any size test, H: $\lambda = 0$ versus K: $\lambda \geq 0$ based on $\bar{L}$, which is $N(\lambda, S)$, for λ defined by* (1·2).

Proof. First transform the inequality constraints test in terms of β to one in terms of a convex polyhedral cone and α. There exists an invertible affine transformation from the model $\hat{b} = \beta + \eta$, η is $N(0, V)$, and $R\beta \geq r$ to the model $\hat{a} = \alpha + w$, w is $N(0, I)$ and $A\alpha \geq 0$, where A is a $p \times k$ lower triangular matrix with ones along the diagonal. The reduction of A to lower triangular form and the normalization of the diagonal elements of A to equal one provides the maximum parametric reduction possible for the problem and an illustration of a property of the weights function, $w(p, j, \Omega)$, discussed later. The constraint set $C = \{\alpha \mid A\alpha \geq 0, \alpha \in R^k\}$ defines a polyhedral cone in R^k. In terms of α, the inequality constraints test (3·1) becomes H: $\alpha \in C$ versus K: $\alpha \in R^k$. The polar cone to C is

$$C^0 = \{y \mid y = -A'z, z \in R^p, z_i \geq 0, i = 1, \ldots, p\}.$$

By Theorem 1, the equivalent dual test to H: $\alpha \in C$ versus K: $\alpha \in R^k$ is H: $\alpha = 0$ versus K: $\alpha \in C^0$, because the covariance matrix of $\hat{a}$ is the identity matrix. If we specialize to $\alpha = (\alpha_1, \alpha_2)'$, $\alpha \in R^2$ and A a 2×2 matrix, the graphical representation of C and C^0 given in Fig. 1 is possible. By the definition of A, the two equations defining C are $\alpha_1 \geq 0$ and $\alpha_2 - a\alpha_1 \geq 0$, where a is the (2, 1) element of A. If we assume $a > 0$, C takes the form given in Fig. 1.

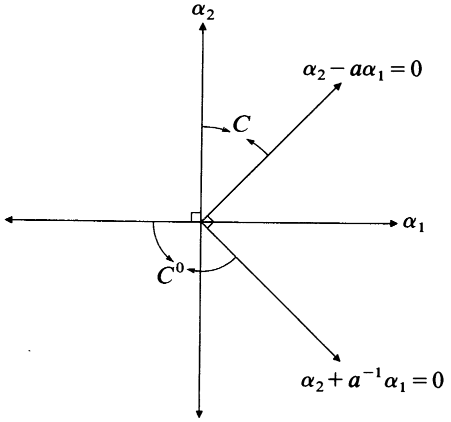

Fig. 1. Graphical representation of cone $C = \{\alpha \mid A\alpha \geq 0, \alpha \in R^2\}$ and polar cone $C^0 = \{\alpha \mid \alpha = -A'z, z \in R^2, z \geq 0\}$.

We now state the invertible affine transformation between testing problem (3·1) in terms of β and this same problem in terms of α. We are guaranteed the existence of $k-p$ rows of length k, denoted by R_*, which make the matrix $T = (R'R'_*)'$ of full rank k. Choose a $(k-p) \times 1$ vector of constants r_* and define a $k \times 1$ vector $t = [r'r'_*]'$. Define $\alpha \equiv h(\beta) \equiv L(T\beta - t)$, where L is a lower triangular matrix such that $L(TVT')L' = I$. The constraint matrix A is defined as follows. By Proposition 5.1 of Eaton (1983), L^{-1} is lower triangular. Define $L^{-1} \equiv (M'M'_*)'$, where M is a $p \times k$ lower triangular matrix and M_* is a $(k-p) \times k$ matrix of the same form. All the diagonal elements of L^{-1} are positive (Dhrymes, 1978, p. 487). Let D be the diagonal matrix with $d_{ii} = l_{ii}$ $(i = 1, \ldots, k)$, where l_{ii} is the ith diagonal element of L. By Proposition 5.1 of Eaton (1983), $l_{ii}^{-1} = l^{ii}$, where l^{ii} is the ith diagonal element of L^{-1}. Define $DL^{-1} = [A'A'_*]'$. This yields the matrix A, a $p \times k$ lower triangular matrix with ones along the diagonal and A_*, a $(k-p) \times k$ matrix of this same form. For α and β related via the invertible function $h(\beta)$, $\{\beta \mid R\beta \geq r, \beta \in R^k\}$ corresponds to C in terms of α and $\{\beta \mid S^{-1}\lambda = R\beta - r, \lambda \geq 0, \lambda \in R^p\}$ corresponds to C^0 in terms of α. Hence for λ defined by (1·2), by Theorem 1 the duality relation between the multivariate inequality constraints test in terms of β and the multivariate one-sided test in terms of λ holds for all $\beta \in R^k$ and corresponding $\lambda \in R^p$. □

By the results of Wolak (1987), the null distribution of our inequality constraints test statistic can be derived in terms of this geometric notation. For the least favourable value of $\alpha = 0$, this distribution is

$$\text{pr}_{0,I}\,(T_{\text{ALR}} \geqslant c) = \sum_{j=0}^{p} \text{pr}\,(\chi_j^2 \geqslant c)\,w(p, p-j, AA'), \tag{3·6}$$

where T_{ALR} is the likelihood ratio statistic from the primal or dual test based on $\hat{a}$, and $w(p, p-j, AA')$ is the weights function associated with $A\tilde{a}$ and $A\hat{a}$, where $\tilde{a}$ is the estimate of α constrained to lie in C. By definition of A, we have $AA' = D_1MM'D_1$, where D_1 is the $(p \times p)$ square submatrix composed of the first intersection of the p rows and columns of D. By definition of M we have $MM' = RVR'$. Therefore AA' simplifies to $D_1RVR'D_1$. By definition of D, $D_1RVR'D_1$ and RVR' have the same correlation matrix. Kudô (1963) shows that $w(p, j, \Omega)$ depends on Ω only through the correlation matrix of Ω. Therefore the null distribution in (3·6) is equal to those given in (3·4) and (3·5).

The results of our theorem can be used to derive other duality relations. Let R, r and A be partitioned as follows:

$$R = (R_1'R_2')', \quad r = (r_1'r_2')', \quad A = (A_1'A_2')',$$

where R_1 is $q \times k$, R_2 is $(p-q) \times k$, r_1 is $q \times 1$, r_2 is $(p-q) \times 1$, A_1 is $q \times k$ and A_2 is $(p-q) \times k$ $(q \leqslant p \leqslant k)$. Let $C_* = \{\alpha \mid A_1\alpha \geqslant 0, A_2\alpha = 0, \alpha \in R^k\}$ and C_*^0 be the polar cone of C_*. For λ defined by (1·2), let $\lambda = (\lambda_1', \lambda_2')'$, with $\lambda_1 \in R^q$ corresponding to $R_1\beta \geqslant r_1$ and $\lambda_2 \in R^{p-q}$ corresponding to $R_2\beta = r_2$. We state the following duality relation without proof. The testing problem H: $R_1\beta \geqslant r_1$ and $R_2\beta = r_2$ versus K: $\beta \in R^k$ is equivalent to, for any size test, H: $\lambda = 0$ versus K: $\lambda_1 \geqslant 0$ and $\lambda_2 \neq 0$. This follows directly from the transformation $h(\beta)$ and Theorem 1 which implies H: $\alpha \in C_*$ versus K: $\alpha \in R^K$ and H: $\alpha = 0$ versus K: $\alpha \in C_*^0$ are equivalent testing problems.

As claimed in § 1, the classical duality result for multivariate equality constraints testing is also a special case of Theorem 1 because $\{\alpha \mid A\alpha = 0, \alpha \in R^k\}$, the set of α corresponding to the set of β such that $R\beta = r$, is a polyhedral cone.

Acknowledgements

I would like to thank the referee and editor for very useful comments and the referee, in particular, for suggesting the present form for Theorem 1.

References

Bazaraa, M. S. & Shetty, C. M. (1979). *Nonlinear Programming.* New York: Wiley.
Dhrymes, P. J. (1978). *Introductory Econometrics.* New York: Springer-Verlag.
Eaton, M. L. (1983). *Multivariate Statistics: A Vector Space Approach.* New York: Wiley.
Gill, P. E., Murray, W. & Wright, M. H. (1981). *Practical Optimization.* New York: Academic Press.
Kudô, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* **50**, 403-18.
Nuesch, P. E. (1966). On the problem of testing location in multivariate problems for restricted alternatives. *Ann. Math. Statist.* **37**, 113-9.
Perlman, M. D. (1969). One-sided problems in multivariate analysis. *Ann. Math. Statist.* **40**, 549-67.
Rockafellar, R. T. (1970). *Convex Analysis.* Princeton University Press.
Wolak, F. A. (1987). An exact test for multiple inequality and equality constraints in the linear regression model. *J. Am. Statist. Assoc.* **82**, 782-93.

[*Received January* 1987. *Revised February* 1988]