The Annals of Statistics
1982, Vol. 10, No. 1. 313–315

COVARIANCE STABILIZING TRANSFORMATIONS AND A CONJECTURE OF HOLLAND

BY C. C. SONG

State University of New York at Albany

This paper gives a proof of a conjecture of P. W. Holland concerning the non-existence of certain covariance stabilizing transformations.

1. Introduction. Suppose the random vector nX has multinomial distribution with parameter vector $(p_1, \cdots, p_{r+1})$ where $X = (x_1, \cdots, x_{r+1})$, and assume that the parameter space is $D = \{(p_1, \cdots, p_r) \mid p_i > 0 \text{ and } \Sigma\, p_i < 1\}$. Let $D^r \subset R^r$ be given by $D^r = \{(x_1, \cdots, x_r) \mid x_i \geq 0 \text{ and } \Sigma\, x_i \leq 1\}$. By a convariance stabilizing transformation, we mean, as in Holland (1973, Section 2), a vector function $F = (f^1, \cdots, f^r): D^r \to R^r$ such that the covariance matrix of the random vector $(f^1(x_1, \cdots, x_r), \cdots, f^r(x_1, \cdots, x_r))$ is nonsingular and independent of the parameter vector $(p_1, \cdots, p_r)$. This may be expressed in terms of matrices as follows.

Let $J_F = (f^i_j)$ be the $r \times r$ Jacobian matrix of F where $f^i_j = \partial f^i/\partial x_j$, and let Σ be the covariance matrix of $\sqrt{n}\,(x_1, \cdots, x_r)$, with (i, j) entry $p_i(\delta_{ij} - p_j)$. Then F will be a covariance stabilizing transformation provided that J_F is invertible and that $\text{Cov}(f^1(x_1, \cdots, x_r), \cdots, f^r(x_1, \cdots, x_r)) = J_F \cdot \Sigma \cdot J_F^T$ is constant. If the matrix $J_F \cdot \Sigma \cdot J_F^T$ is constant then, since it is positive definite, it is equivalent to the identity matrix via a constant orthogonal rotation. Consequently, the question of whether a covariance stabilizing transformation exists can be compactly stated as follows:

Does there exist a transformation $F: D^r \to R^r$ such that J_F is nonsingular and $J_F \cdot \Sigma \cdot J_F^T = I$?

Holland (1973) noted that such transformations exist when $r = 1$ and went on to show that, in the trinomial case ($r = 2$), such transformations do not exist. Holland's technique for attacking the existence problem applies to a number of two-dimensional problems. The technique is based on an elegant reduction of the problem which goes as follows.

The equation $J_F \cdot \Sigma \cdot J_F^T = I$ is equivalent to $J_F^T \cdot J_F = \Sigma^{-1}$. Since Σ^{-1} is symmetric and positive definite, we can rewrite this as:

$$J_F^T \cdot J_F = (\Sigma^{-1/2})^T \cdot \Sigma^{-1/2}, \tag{1.1}$$

where $\Sigma^{-1/2}$ is any square root of Σ^{-1}. Then the existence of a solution to (1.1) is equivalent to the existence of an orthogonal matrix Γ such that

$$J_F = \Gamma \cdot \Sigma^{-1/2} \tag{1.2}$$

is a Jacobian. Holland then derives a general criterion for solving this problem in the case of 2×2 matrices, and applies this criterion to obtain his results on existence or non-existence for several distributions.

The criterion obtained in studying (1.2) relies heavily on the fact that 2×2 orthogonal matrices have simple representations in terms of sines and cosines. Attempts to generalize Holland's construction, even to the 3×3 case, lead immediately to computations of immense and probably unresolvable complexity. Holland noted this difficulty and, despite it, conjectured that covariance stabilization is impossible for the general multinomial distribution. In this paper, Holland's conjecture is proved by attacking the original equation directly.

Received August 1980; revised September 1981.
AMS 1970 *subject classification.* Primary 62H99, secondary 60E05.
Key words and phrases. Jacobian matrix, multinomial distribution.

THEOREM. *There is no C^3 function $F:D^r \to R^r$ such that J_F is nonsingular and $J_F^T \cdot J_F = \Sigma^{-1}$.*

2. Proof of the theorem. The proof is by contradiction. We assume there is a C^3 function $F = (f^1, \cdots, f^r)$ with the properties of the theorem and, by combining various identities derived from $J_F^T \cdot J_F = \Sigma^{-1}$, we find expressions for the third partials f^i_{stu}. Equating these expressions yields the fact that, for each i, the first partials f^i_j are all equal. Thus the rows of J_F are constant and J_F is singular.

NOTATION.
$$\underset{\sim}{1} = (1, \cdots, 1)^T,$$
$$\alpha_i = p_i^{-1}, \quad i = 1, \cdots, r, r+1,$$
$$\mathbf{f}_i = (f_i^1, \cdots, f_i^r)^T, \quad i = 1, \cdots, r,$$
$$\mathbf{f}_{ij} = (f_{ij}^1, \cdots, f_{ij}^r)^T, \quad i, j = 1, \cdots, r,$$
$$e^{ij} = \mathbf{f}_i^T \cdot \mathbf{f}_j, \quad i, j = 1, \cdots, r,$$
$$e_k^{ij} = \mathbf{f}_{ik}^T \cdot \mathbf{f}_j + \mathbf{f}_i^T \cdot \mathbf{f}_{jk}, \quad i, j, k = 1, \cdots, r.$$

First, equate components in the matrix equation $J_F^T \cdot J_F = \Sigma^{-1}$. This gives

$$e^{ij} = \mathbf{f}_i^T \cdot \mathbf{f}_j = \alpha_i \delta_{ij} + \alpha_{r+1}; \qquad i, j = 1, \cdots, r \tag{2.1}$$

where δ_{ij} is the Kronecker delta. Next, computing the partial of each equation in (2.1) with respect to p_k, we have

$$e_k^{ij} = \mathbf{f}_{ik}^T \cdot \mathbf{f}_j + \mathbf{f}_i^T \cdot \mathbf{f}_{jk} = -\alpha_i^2 \delta_{ij} \delta_{ik} + \alpha_{r+1}^2; \qquad i, j, k = 1, \cdots, r. \tag{2.2}$$

Using (2.1) and (2.2), we find that

$$\mathbf{f}_i^T \cdot \mathbf{f}_{jk} = \tfrac{1}{2} \alpha_{r+1}^2, \quad \text{for} \quad j \neq k, \tag{2.3}$$

which gives the following matrix equation:

$$J_F^T \cdot \mathbf{f}_{jk} = \tfrac{1}{2} \alpha_{r+1}^2 \cdot \underset{\sim}{1}$$

or

$$\mathbf{f}_{jk} = \tfrac{1}{2} \alpha_{r+1}^2 \cdot (J_F^T)^{-1} \cdot \underset{\sim}{1}, \quad \text{for} \quad j \neq k. \tag{2.4}$$

Thus, in particular each f^i has the property that all second order partials with distinct indices are equal.

The remainder of the proof involves the development of some useful matrix identities which are related to (2.4).It can be shown that

$$(J_F^T) \cdot \left(\frac{1}{2} \mathbf{f}_i + \textstyle\sum_{j=1}^r \frac{1}{\alpha_j} \mathbf{f}_{ij} \right) = \frac{1}{2} \alpha_{r+1}^2 \cdot \underset{\sim}{1}$$

or

$$\frac{1}{2} \mathbf{f}_i + \textstyle\sum_{j=1}^r \frac{1}{\alpha_j} \mathbf{f}_{ij} = \frac{1}{2} \alpha_{r+1}^2 \cdot (J_F^T)^{-1} \cdot \underset{\sim}{1} \tag{2.5}$$

for each index $i = 1, \cdots, r$. Similarly,

$$\textstyle\sum_{j=1}^r \frac{1}{\alpha_j} \mathbf{f}_i = \alpha_{r+1} (J_F^T)^{-1} \cdot \underset{\sim}{1} = \frac{2}{\alpha_{r+1}} \cdot \frac{1}{2} \alpha_{r+1}^2 \cdot (J_F^T)^{-1} \cdot \underset{\sim}{1}. \tag{2.6}$$

From (2.4) and (2.6) we have

$$f_{st}^i = \frac{\alpha_{r+1}}{2} \textstyle\sum_{j=1}^r \frac{1}{\alpha_j} f_j^i; \quad \text{for} \quad s \neq t, \quad \text{and} \quad i = 1, \cdots, r.$$

Now fix i and differentiate the preceding identity with respect to $1/\alpha_u$ where s, t, u are distinct. This gives the following expression for f^i_{stu}:

$$f^i_{stu} = \frac{\alpha^2_{r+1}}{2}\left[\alpha_{r+1}\left(\sum_{j=1}^{r}\frac{1}{\alpha_j}f^i_{ju} + f^i_u\right) + \sum_{j=1}^{r}\frac{1}{\alpha_j}f^i_j\right].$$

If we compute the third partials f^i_{stu} and f^i_{tus} and equate these expressions (using the assumption that our transformation was C^3), we find that

$$\sum_{j=1}^{r}\frac{1}{\alpha_j}f^i_{ju} + f^i_u = \sum_{j=1}^{r}\frac{1}{\alpha_j}f^i_{js} + f^i_s \tag{2.7}$$

for all u, s. Finally, recall that (2.5) holds for all i,

$$\sum_{j=1}^{r}\frac{1}{\alpha_j}f^i_{uj} + \frac{1}{2}f^i_u = \sum_{j=1}^{r}\frac{1}{\alpha_j}f^i_{sj} + \frac{1}{2}f^i_s \tag{2.8}$$

for all u, s. Subtracting (2.8) from (2.7) yields the result:

$$\frac{f^i_u}{2} = \frac{f^i_s}{2}.$$

Since this holds for all i, u and s, the rows of J_F are identical and J_F is singular after all. This is the desired contradiction.

Acknowledgments. This paper forms part of the author's Ph.D. dissertation. The author is indebted to Professors C. Chen, J. Sedransk and E. S. Thomas for reviewing the manuscript and making several helpful suggestions.

REFERENCE

HOLLAND, P. W. (1973). Covariance stabilizing transformation. *Ann. Statist.* 1 84–92.

DEPARTMENT OF MATHEMATICS AND STATISTICS
STATE UNIVERSITY OF NEW YORK AT ALBANY
1400 WASHINGTON AVENUE
ALBANY, NY 12222