Statistics & Probability Letters 1 (1983) 213–215
North-Holland June 1983

A Note on Spectral Decomposition and Maximum Likelihood Estimation in ANOVA Models with Balanced Data *

Tom Wansbeek
Netherlands Central Bureau of Statistics, Voorburg, The Netherlands

Arie Kapteyn
Department of Econometrics, Tilburg University, Tilburg, The Netherlands

Received July 1982

Abstract. A simple derivation of the spectral decomposition of the covariance matrix for a general multi-way variance components model is presented. So-called balanced data are assumed to be available. Spectral decomposition is exploited to derive the information matrix and the first-order conditions for the maximum likelihood estimation of the variance components parameters.

Keywords. ANOVA, spectral decomposition.

1. Introduction

In this note we present a simple derivation of the spectral decomposition of the covariance matrix for a general multi-way variance components model. Balanced data (to be defined below) are assumed to be available. This problem has been discussed before by Searle and Henderson (1979), who present a longer derivation. To illustrate the importance of the spectral decomposition, we exploit it to derive the information matrix and the first-order conditions for the maximum likelihood estimation of the variance components parameters.

* The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the Netherlands Central Bureau of Statistics.

2. The model

Let there be a random vector u which is normally distributed with zero expectation and covariance matrix V of the following structure

$$V=\sum_{d}\lambda_d N_d, \tag{2.1}$$

where d is a k-vector of zeros and ones. The summation runs over all such vectors (i.e. over 2^k elements). The λ_d are nonnegative parameters and N_d is a matrix consisting of a Kronecker product of k separate matrices of order n_i $(i=1,\dots,k)$, each of which is either a unit matrix (I_{n_i}) or a square matrix of ones (J_{n_i}). Furthermore, let d_i be the ith element of d. A unit matrix occurs in the ith position when $d_i=1$ and a matrix of ones when $d_i=0$. For example, if $k=3$, $d=(0, 1, 1)$,

0167-7152/83/$03.00 © 1983 Elsevier Science Publishers B.V. (North-Holland)

then

$$N_{011} = J_{n_1} \otimes I_{n_2} \otimes I_{n_3}, \tag{2.2}$$

with matrices of order n_1, n_2 and n_3, respectively. Whenever a covariance matrix has format (2.1), we have *balanced data*.

3. The spectral decomposition

Assume it is possible to rewrite the covariance matrix V in the format

$$V = \sum_d \phi_d M_d, \tag{3.1}$$

with the ϕ_d scalar parameters, and M_d a Kronecker product of k separate matrices; but here the ith matrix in M_d equals $E_{n_i} \equiv I_{n_i} - J_{n_i}/n_i$ if $d_i = 1$ and it equals $\bar{J}_{n_i} = J_{n_i}/n_i$ if $d_i = 0$. For example

$$M_{011} = \bar{J}_{n_1} \otimes E_{n_2} \otimes E_{n_3}. \tag{3.2}$$

Whenever V can be written as in (2.1), it can also be writen as in (3.1) (and vice versa). An example may suffice to show why

$$N_{011} = n_1(M_{000} + M_{010} + M_{001} + M_{011}), \tag{3.3}$$

as is easily verified. In general, the relation between N_d and M_d can be written as

$$N_d = n^{\bar{d}} \sum_{e \leqslant d} M_e, \tag{3.4}$$

where the following shorthand notation is used: $\bar{d} = \iota - d$ (ι a k-vector of ones) and

$$n^{\bar{d}} = n^{\iota - d} \equiv \sum_{i=1}^{k} n_i^{1-d_i}; \tag{3.5}$$

the notation $e \leqslant d$ means that the elements of e are not greater than the corresponding elements of d. So,

$$V = \sum_d \lambda_d N_d = \sum_d \lambda_d n^{\bar{d}} \sum_{e \leqslant d} M_e$$

$$= \sum_d \Big(\sum_{e \geqslant d} \lambda_e n^{\bar{e}} \Big) M_d \equiv \sum_d \phi_d M_d, \tag{3.6}$$

where the third equality sign can be checked by writing out the third member. Now the M_d can be verified to have the following properties:

- they add up (over all d) to the unit matrix;
- they are idempotent;
- $M_d M_e = 0$ for $d \neq e$.

Thus (3.6) is exactly the spectral decomposition of V. There are at most 2^k different eigenvalues ϕ_d ($\phi_d = \Sigma_{e \geqslant d} \lambda_e n^{\bar{e}}$), with multiplicity equal to the rank of the corresponding matrix M_d, which equals n^d.

The spectral decomposition of V greatly facilitates the computation of its powers since

$$V^\alpha = \sum_d \phi_d^\alpha M_d \quad \text{for any scalar } \alpha. \tag{3.7}$$

4. Maximum likelihood estimation

Let $V_f = \partial v / \partial \lambda_f$, then the first-order condition for maximum likelihood estimation of λ_f is

$$\operatorname{tr}\big(V_f V^{-1}\big) = u'\big(V^{-1} V_f V^{-1}\big) u, \tag{4.1}$$

where u was introduced at the beginning of Section 2.

A typical element of the information matrix (apart from a factor $-\frac{1}{2}$) is (e.g. Searle (1970))

$$\operatorname{tr}\big(V_f V^{-1} V_g V^{-1}\big). \tag{4.2}$$

Given the spectral decomposition it is easy to calculate these quantities:

$$v_f = \sum_d \frac{\partial \lambda_d}{\partial \lambda_f} N_d \equiv N_f = n^{\bar{f}} \sum_{e \leqslant f} M_e; \tag{4.3}$$

so,

$$\operatorname{tr}\big(V_f V^{-1}\big) = n^{\bar{f}} \operatorname{tr}\Big(\sum_d \frac{1}{\phi_d} M_d \Big)\Big(\sum_{e \leqslant f} M_e \Big)$$

$$= n^{\bar{f}} \operatorname{tr}\Big(\sum_{e \leqslant f} \frac{1}{\phi_e} M_e \Big)$$

$$= n^{\bar{f}} \sum_{e \leqslant f} \frac{n^e}{\phi_e} = n^{\bar{f}} \sum_{e \leqslant f} \frac{n^e}{\sum_{d \geqslant e} \lambda_d n^{\bar{d}}}$$

$$= \sum_{e \leqslant f} \frac{1}{\sum_{d \geqslant e} \lambda_d n^{\bar{d} - \bar{f} - e}}; \tag{4.4}$$

$$u'\big(V^{-1} V_f V^{-1}\big) u = n^{\bar{f}} \sum_{e \leqslant f} \frac{1}{\phi_e^2} u' M_e u; \tag{4.5}$$

$$\operatorname{tr}\left(V_f V^{-1} V_g V^{-1}\right) = n^{\bar{f}+\bar{g}} \sum_{e \leqslant f} \frac{1}{\phi_e^2} \operatorname{tr}(M_e)$$
$$= n^{\bar{f}+\bar{g}} \sum_{e \leqslant f} \frac{n^e}{\phi_e^2} \quad (f \leqslant g). \qquad (4.6)$$

The last two expressions can be worked out further by substituting for ϕ_e, just as in (4.4).

5. Conclusion

The spectral decomposition of the covariance matrix of balanced designs can be obtained in a simple way. The spectral decomposition, in turn, greatly facilitates the derivation of the information matrix and first-order conditions for maximum likelihood estimation of the variance components parameters.

References

Searle, S.R. (1970), Large sample variances of maximum likelihood estimators of variance components, *Biometrics* **26**, 505–524.

Searle, S.R. and H.V. Henderson (1979), Dispersion matrices for variance components models, *J. Amer. Statist. Assoc.* **74**, 465–470.