Asymptotic Covariances For The Parameters Of Biadditive Models

Jean-Baptiste Denis

Laboratoire de Biometrie
I.N.R.A.
route de Saint-Cyr
F-78026 VERSAILLES, FRANCE

and

John C. Gower

Department of Statistics
The Open University
Milton Keynes
MK7 6AA U.K.

ABSTRACT. The statistical model for the two-way table with expectation

$$\mathbf{Y} = m\mathbf{1}\mathbf{1}^T + \mathbf{a}\mathbf{1}^T + \mathbf{1}\mathbf{b}^T + \sum_{u=1}^{r} \mathbf{c}_u \mathbf{d}_u^T$$

is increasingly being used to represent the interaction between genotype and environment in plant-breeding experiments. Proper interpretation of analyses based on this model requires estimates of the standard errors of estimated parameter values. This paper provides the asymptotic variances and covariances of the parameters for this model and for similar models in which one or both of the additive parameters are exluded. It is especially concerned with the technical details of their derivations. The results are given for two equivalent ways of parameterizing the models.

Utilitas Mathematica 46(1994), pp. 193-205

1 Introduction

Here we consider models with expectation of the form:

$$\mathbf{Y} = m\mathbf{1}\mathbf{1}^T + \mathbf{a}\mathbf{1}^T + \mathbf{1}\mathbf{b}^T + \sum_{u=1}^{r} \mathbf{c}_u\mathbf{d}_u^T \tag{1}$$

where $\mathbf{Y}$ represents an $I \times J$ table of data, $\mathbf{a}$ and $\mathbf{c}_u$ $(u = 1, 2, \ldots, r)$ are I-vectors, $\mathbf{b}$ and $\mathbf{d}_u$ $(u = 1, 2, \ldots, r)$ are J-vectors of parameters which require estimation, and $\mathbf{1}$ represents a vector of units whose length is determined from context. The study of such models with multiplicative interaction terms was initiated by Fisher and Mackenzie [5] who considered the simple variant:

$$\mathbf{Y} = \mathbf{c}\mathbf{d}^T. \tag{2}$$

Other variants have been considered in recent years and have found applications, particularly in plant-breeding studies of genotype-environment interaction. Denis and Gower [2] review this family of which (1) is the most general form and discuss computational aspects. A variety of acronyms have been suggested for these models but we favor the neutral mathematical term *Biadditive Models* in acknowledgement that (1) is a special bilinear model which is additive in the row parameters when the column parameters are fixed, and vice versa.

Some constraints are required to identify (1) completely. We write $\mathbf{c}_u^T\mathbf{c}_u = C_u$, $\mathbf{d}_u^T\mathbf{d}_u = D_u$ $(u = 1, 2, \ldots, r)$. Two ways of expressing these constraints are considered in the following.

Parameterization 1: $\mathbf{1}^T\mathbf{a} = \mathbf{1}^T\mathbf{b} = \mathbf{1}^T\mathbf{c}_u = \mathbf{1}^T\mathbf{d}_u = 0$, $C_u = D_u (u = 1, 2, \ldots, r)$ and $\mathbf{c}_u^T\mathbf{c}_v = \mathbf{d}_u^T\mathbf{d}_v = 0$ for $u \neq v$.

Parameterization 2: $\mathbf{1}^T\mathbf{a} = \mathbf{1}^T\mathbf{b} = \mathbf{1}^T\mathbf{c}_u = \mathbf{1}^T\mathbf{d}_u = 0$ $(u = 1, 2, \ldots, r)$, the uth multiplicative term is rewritten $\sigma_u\mathbf{c}_u\mathbf{d}_u^T$ with $C_u = D_u = 1$, $(u = 1, 2, \ldots, r)$ and $\mathbf{c}_u^T\mathbf{c}_v = \mathbf{d}_u^T\mathbf{d}_v = 0$ for $u \neq v$.

These two forms of constraint merely offer different parameterizations of the same model. When estimation is by least-squares, it is clear that the multiplicative terms are given by the singular value decomposition of the matrix of residuals after removing the additive terms (Eckart and Young, [3]) so that in parametrization 2 the σ_u are singular values whereas in parametrization 1 the singular values are absorbed into the lengths of the singular vectors. It is to be understood that if the parameters $\mathbf{a}$ and/or $\mathbf{b}$ are absent from the model then not only are the constraints on these parameters irrelevant but also one or both of the constraints $\mathbf{1}^T\mathbf{c}_u = \mathbf{1}^T\mathbf{d}_u = 0$ $(u = 1, 2, \ldots, r)$; this causes no loss of generality as explained by Denis and Gower [2].

We use the notation $B(m, \mathbf{a}, \mathbf{b}, \pi_r)$ to describe (1) and $B(^*, ^*, ^*, \pi)$ to describe (2). These are examples of an obvious notation for describing the purely additive terms; the suffix r refers to the number of multiplicative terms in the model and if dropped implies that $r = 1$. This paper is concerned with deriving algebraic expressions for the asymptotic variances and covariances of the estimated parameters of these models and considers the increasing complexity of $B(m, ^*, ^*, \pi)$, $B(m, ^*, ^*, \pi_2)$, $B(m, ^*, ^*, \pi_r)$ and $B(m, \mathbf{a}, \mathbf{b}, \pi_r)$, as well as $B(m, \mathbf{a}, ^*, \pi_r)$. From the asymptotic formulae, it is easy to derive asymptotic confidence regions and standard errors of linear combinations of the parameters. Equivalent expressions for the small changes required by other models of the family are indicated.

The usual approach is used of deriving the log-likelihood for $\mathbf{Y}$, and forming the information matrix by taking expectations of the second differentials with respect to the parameters. After changing its sign, the matrix is inverted; the following is mainly concerned with the technical details of this inversion. To keep the algebra symmetric we have imposed the constraints through Lagrangian terms like $\lambda(C - D)$ in the log-likelihood. The consequences of doing this have been studied by Silvey [11] (see also Silvey, [12]) and can be summarized as justifying treating the Lagrange multiplier λ as if it were another parameter for estimation. In the inverse of the information matrix, the rows and columns pertaining to the Lagrange multipliers are ignored.

2 The Tukey Models $B(m, ^*, ^*, \pi_r)$

In this section we derive results using parameterization 1. Because $C_u = D_u$ $(u = 1, 2, \ldots, r)$, we express the results in terms of C_u only or, when $r = 1$, in terms of C. We also use the notation $\mathbf{1}^T\mathbf{c}_u = \gamma_u$, $\mathbf{1}^T\mathbf{d}_u = \delta_u$, $(u = 1, 2, \ldots, r)$; recall that because there are no additive terms in these models the identification constraints $\mathbf{1}^T\mathbf{c}_u = \mathbf{1}^T\mathbf{d}_u = 0$ are inoperative. The model $B(m, ^*, ^*, \pi)$ is a parameterization of the so-called Tukey model for one degree of freedom for non-additivity (implicit in Tukey, [13]); the generalizations $B(m, ^*, ^*, \pi_r)$ have been recently studied by Seyedsadr and Cornelius [10].

2.1 The Simple Tukey Model $B(m, ^*, ^*, \pi)$

In this case there is only a single Lagrange multiplier λ. Assuming independent and identically distributed normal errors with common variance the log-likelihood $\mathcal{L}$ is:

$$\sigma^2\mathcal{L} = -\frac{1}{2}\|\mathbf{Y} - (m\mathbf{1}\mathbf{1}^T + \mathbf{c}\mathbf{d}^T)\|^2 - \frac{1}{2}\lambda(\mathbf{c}^T\mathbf{c} - \mathbf{d}^T\mathbf{d}) + \text{ constant}$$

from which we have:

$$
\begin{aligned}
\sigma^2(\partial\mathcal{L}/\partial m) &= \mathbf{1}^T[\mathbf{Y} - (m\mathbf{1}\mathbf{1}^T + \mathbf{c}\mathbf{d}^T)]\mathbf{1} \\
\sigma^2(\partial\mathcal{L}/\partial \mathbf{c}) &= [\mathbf{Y} - (m\mathbf{1}\mathbf{1}^T + \mathbf{c}\mathbf{d}^T)]\mathbf{d} - \lambda\mathbf{c} \\
\sigma^2(\partial\mathcal{L}/\partial \mathbf{d}) &= [\mathbf{Y} - (m\mathbf{1}\mathbf{1}^T + \mathbf{c}\mathbf{d}^T)]^T\mathbf{c} + \lambda\mathbf{d}
\end{aligned} \qquad (3)
$$

Equating these to zero gives the equations satisfied by the maximum likelihood estimates. Note that $\mathbf{c}^T\mathbf{Y}\mathbf{d} - \mathbf{d}^T\mathbf{Y}^T\mathbf{c} = \lambda(C + D) = 0$, so that $\lambda = 0$. Taking second differentials, expectations and changing sign yields the information matrix:

$$
-\sigma^2\mathcal{E}(\partial^2\mathcal{L}/\partial\Theta^2) = \begin{pmatrix} IJ & \delta\mathbf{1}^T & \gamma\mathbf{1}^T & 0 \\ \delta\mathbf{1} & D\mathbf{I} & \mathbf{c}\mathbf{d}^T & \mathbf{c} \\ \gamma\mathbf{1} & \mathbf{d}\mathbf{c}^T & C\mathbf{I} & -\mathbf{d} \\ 0 & \mathbf{c}^T & -\mathbf{d}^T & 0 \end{pmatrix} \begin{matrix} m \\ \mathbf{c} \\ \mathbf{d} \\ \lambda \end{matrix}
$$

This matrix has to be inverted to obtain the required asymptotic dispersions. It is easy to see that all powers of this matrix have the form:

$$
\begin{pmatrix} p & l_{11}\mathbf{1}^T + m_{11}\mathbf{c}^T & l_{13}\mathbf{1}^T + m_{12}\mathbf{d}^T & q \\ l_{11}\mathbf{1} + m_{11}\mathbf{c} & \rho\mathbf{I} + \theta\mathbf{1}\mathbf{1}^T + \lambda\mathbf{c}\mathbf{c}^T + \alpha(\mathbf{1}\mathbf{c}^T + \mathbf{c}\mathbf{1}^T) & \Psi\mathbf{1}\mathbf{1}^T + \nu\mathbf{c}\mathbf{d}^T + \xi\mathbf{c}\mathbf{1}^T + \zeta\mathbf{1}\mathbf{d}^T & l_{21}\mathbf{1} + m_{21}\mathbf{c} \\ l_{12}\mathbf{1} + m_{12}\mathbf{d} & \Psi\mathbf{1}\mathbf{1}^T + \nu\mathbf{d}\mathbf{c}^T + \xi\mathbf{1}\mathbf{c}^T + \zeta\mathbf{d}\mathbf{1}^T & \sigma\mathbf{I} + \phi\mathbf{1}\mathbf{1}^T + \mu\mathbf{d}\mathbf{d}^T + \beta(\mathbf{1}\mathbf{d}^T + \mathbf{d}\mathbf{1}^T) & l_{22}\mathbf{1} + m_{22}\mathbf{d} \\ q & l_{21}\mathbf{1}^T + m_{21}\mathbf{c}^T & l_{22}\mathbf{1}^T + m_{22}\mathbf{d}^T & r \end{pmatrix}
$$

Following arguments given by Gower [7] and Gower and Groenen [8] it follows that the inverse must have the same form. To determine the unknown parameters, one needs only to multiply the two matrices together and set the result equal to the unit matrix. This is a simple, but tedious, process the results of which can be presented in a simple form by defining

$$
\mathbf{u} = \frac{\delta}{C}\left(1 - \frac{\gamma\mathbf{c}}{2C}\right) \text{ and } \mathbf{v} = \frac{\gamma}{C}\left(1 - \frac{\delta\mathbf{d}}{2C}\right).
$$

The inverse then becomes:

$$
\sigma^2 \begin{pmatrix} p & -p\mathbf{u}^T & -p\mathbf{v}^T & 0 \\ -p\mathbf{u} & \frac{1}{C}\mathbf{1} + p\mathbf{u}\mathbf{u}^T - \frac{3}{4C^2}\mathbf{c}\mathbf{c}^T & p\mathbf{u}\mathbf{v}^T + \frac{1}{4C^2}\mathbf{c}\mathbf{d}^T & \frac{\mathbf{c}}{2C} \\ -p\mathbf{v} & p\mathbf{v}\mathbf{u}^T + \frac{1}{4C^2}\mathbf{d}\mathbf{c}^T & \frac{1}{C}\mathbf{I} + p\mathbf{v}\mathbf{v}^T - \frac{3}{4C^2}\mathbf{d}\mathbf{d}^T & -\frac{\mathbf{d}}{2C} \\ 0 & \frac{\mathbf{c}^T}{2C} & -\frac{\mathbf{d}^T}{2C} & 0 \end{pmatrix} \begin{matrix} m \\ \mathbf{c} \\ \mathbf{d} \\ \lambda \end{matrix} \qquad (4)
$$

where $p = (I - \gamma^2/C)^{-1}(J - \delta^2/C)^{-1}$. Recall that the final row and column arise from the Lagrange multiplier and are irrelevant so far as the asymptotic variances are concerned.

2.2 Tukey Model with Two Multiplicative Terms: $B(m, ^*, ^*, \pi_2)$

The log-likelihood is:

$$
\begin{aligned}
\sigma^2\mathcal{L} = &-\frac{1}{2}\|\mathbf{Y} - (m\mathbf{1}\mathbf{1}^T + \mathbf{c}_1\mathbf{d}_1^T + \mathbf{c}_2\mathbf{d}_2^T)\|^2 - \frac{1}{2}\lambda(\mathbf{c}_1^T\mathbf{c}_1 - \mathbf{d}_1^T\mathbf{d}_1) \\
&- \frac{1}{2}\mu(\mathbf{c}_2^T\mathbf{c}_2 - \mathbf{d}_2^T\mathbf{d}_2) - \nu\mathbf{c}_1^T\mathbf{c}_2 - \tau\mathbf{d}_1^T\mathbf{d}_2 + \text{ constant,}
\end{aligned}
$$

where λ, μ, ν and τ are Lagrange multipliers, the last two of which arise from the orthogonality constraints. This leads to:

$$
\mathbf{L} = -\sigma^2\Delta^{-1}\mathcal{E}(\partial^2\mathcal{L}/\partial\Theta^2)\Delta^{-1} =
$$

$$
\begin{pmatrix}
\mathbf{I} & \mathbf{c}_1\mathbf{d}_1^T & 0 & \mathbf{c}_2\mathbf{d}_1^T & \delta_1\mathbf{1} & \mathbf{c}_1 & 0 & \rho\mathbf{c}_2 & 0 \\
\mathbf{d}_1\mathbf{c}_1^T & \mathbf{I} & \mathbf{d}_2\mathbf{c}_1^T & 0 & \gamma_1\mathbf{1} & -\mathbf{d}_1 & 0 & 0 & \rho\mathbf{d}_2 \\
0 & \mathbf{c}_1\mathbf{d}_2^T & \mathbf{I} & \mathbf{c}_2\mathbf{d}_2^T & \delta_2\mathbf{1} & 0 & \mathbf{c}_2 & \rho^{-1}\mathbf{c}_1 & 0 \\
\mathbf{d}_1\mathbf{c}_2^T & 0 & \mathbf{d}_2\mathbf{c}_2^T & \mathbf{I} & \gamma_2\mathbf{1} & 0 & -\mathbf{d}_2 & 0 & \rho^{-1}\mathbf{d}_1 \\
\delta_1\mathbf{1}^T & \gamma_1\mathbf{1}^T & \delta_2\mathbf{1}^T & \gamma_2\mathbf{1}^T & IJ & 0 & 0 & 0 & 0 \\
\mathbf{c}_1^T & -\mathbf{d}_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{c}_2^T & -\mathbf{d}_2^T & 0 & 0 & 0 & 0 & 0 \\
\rho\mathbf{c}_2^T & 0 & \rho^{-1}\mathbf{c}_1^T & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \rho\mathbf{d}_2^T & 0 & \rho^{-1}\mathbf{d}_1^T & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{matrix} \mathbf{c}_1 \\ \mathbf{d}_1 \\ \mathbf{c}_2 \\ \mathbf{d}_2 \\ m \\ \lambda \\ \mu \\ \nu \\ \tau \end{matrix}
$$

where $\rho = (C_2/C_1)^{1/2}$, $\Delta = diag(C_1^{1/2}\mathbf{1}^T, C_1^{1/2}\mathbf{1}^T, C_2^{1/2}\mathbf{1}^T, C_2^{1/2}\mathbf{1}^T, 1, 1, 1, 1, 1)$ and $\mathbf{c}_1, \mathbf{d}_1, \mathbf{c}_2, \mathbf{d}_2$ are new scaled versions of the previous vectors defined by: $\text{new}(\mathbf{c}_1, \mathbf{d}_1, \mathbf{c}_2, \mathbf{d}_2) = \text{old}(C_1^{-1/2}\mathbf{c}_1, C_1^{-1/2}\mathbf{d}_1, C_2^{-1/2}\mathbf{c}_2, C_2^{-1/2}\mathbf{d}_2)$ and $\mathbf{1}^T\text{new}(\mathbf{c}_1, \mathbf{d}_1, \mathbf{c}_2, \mathbf{d}_2) = \text{new}(\gamma_1, \delta_1, \gamma_2, \delta_2)$, so now $\text{new}(\mathbf{c}_1^T\mathbf{c}_1) = \text{new}(\mathbf{d}_1^T\mathbf{d}_1) = \text{new}(\mathbf{c}_2^T\mathbf{c}_2) = \text{new}(\mathbf{d}_2^T\mathbf{d}_2) = 1$.

Some care should be taken with these changes of notation but they much simplify the algebraic details. For convenience (see below), the row and column arising from m are put at the end of the block matrix of parameters. The inverse of the matrix is obtained as follows. First we write $\mathbf{L}$ in partitioned form:

$$
\begin{pmatrix} \mathbf{A} & \mathbf{C}^* \\ \mathbf{C}^{*T} & \mathbf{B} \end{pmatrix} \text{ with its corresponding inverse, } \begin{pmatrix} \mathbf{P} & \mathbf{R}^* \\ \mathbf{R}^{*T} & \mathbf{Q} \end{pmatrix}
$$

where $\mathbf{A}$ represents the 4×4 top left hand blocks and $\mathbf{B}$ is zero except for the single term IJ; thus this partitioning of $\mathbf{L}$ differs slightly from the natural one into blocks pertaining to the model parameters and to Lagrangian terms, respectively. Repeated powering of the left hand side, but without regard to evaluating detailed values of the coefficients, shows that every column of $\binom{\mathbf{R}^*}{\mathbf{Q}}$ has the form:

$$\begin{pmatrix} l_1\mathbf{1}+m_1\mathbf{c}_1+n_1\mathbf{c}_2 \\ l_2\mathbf{1}+m_2\mathbf{d}_1+n_2\mathbf{d}_2 \\ l_3\mathbf{1}+m_3\mathbf{c}_1+n_3\mathbf{c}_2 \\ l_4\mathbf{1}+m_4\mathbf{d}_1+n_4\mathbf{d}_2 \\ p \\ q \\ r \\ s \\ t \end{pmatrix}.$$

Pre-multiplying this vector by the 9×9 row-blocks of the left hand side and expressing the fact that the result must be a column of the unit matrix gives a series of equations which allow all the elements of $\mathbf{R}^*$ and $\mathbf{Q}$ to be determined. This gives:

$$\mathbf{Q}_{11}=p=(\mathbf{I}-\gamma_1^2-\gamma_2^2)^{-1}(J-\delta_1^2-\delta_2^2)^{-1},$$

all other elements of $\mathbf{Q}$ being zero, and:

$$\mathbf{R}^*=\begin{pmatrix} -p\mathbf{u}_1 & \frac{1}{2}\mathbf{c}_1 & 0 & k\rho\mathbf{c}_2 & k\rho^{-1}\mathbf{c}_2 \\ -p\mathbf{v}_1 & -\frac{1}{2}\mathbf{d}_1 & 0 & k\rho^{-1}\mathbf{d}_2 & k\rho\mathbf{d}_2 \\ -p\mathbf{u}_2 & 0 & \frac{1}{2}\mathbf{c}_2 & -k\rho^{-1}\mathbf{c}_1 & -k\rho\mathbf{c}_1 \\ -p\mathbf{v}_2 & 0 & -\frac{1}{2}\mathbf{d}_2 & -k\rho\mathbf{d}_1 & -k\rho^{-1}\mathbf{d}_1 \end{pmatrix}$$

where $k=(\rho^2-\rho^{-2})^{-1}$ and

$$\begin{aligned} \mathbf{u}_1&=\delta_1\mathbf{1}-\frac{1}{2}\gamma_1\delta_1\mathbf{c}_1-k\rho(\rho\delta_1\gamma_2+\rho^{-1}\gamma_1\delta_2)\mathbf{c}_2 \\ \mathbf{v}_1&=\gamma_1\mathbf{1}-\frac{1}{2}\gamma_1\delta_1\mathbf{d}_1-k\rho(\rho\gamma_1\delta_2+\rho^{-1}\delta_1\gamma_2)\mathbf{d}_2 \\ \mathbf{u}_2&=\delta_2\mathbf{1}-\frac{1}{2}\gamma_2\delta_2\mathbf{c}_2+k\rho^{-1}(\rho\delta_1\gamma_2+\rho^{-1}\gamma_1\delta_2)\mathbf{c}_1 \\ \mathbf{v}_2&=\gamma_2\mathbf{1}-\frac{1}{2}\gamma_2\delta_2\mathbf{d}_2+k\rho^{-1}(\rho\gamma_1\delta_2+\rho^{-1}\delta_1\gamma_2)\mathbf{d}_1 \end{aligned} \qquad (5)$$

It remains to determine $\mathbf{P}$, the most important part of the inverse matrix. We first repartition $\mathbf{L}$ and its inverse to isolate the row/column corresponding to the constant term m, This involves setting $\mathbf{C}^*=(\omega\ \mathbf{C})$ and $\mathbf{R}^*=(-p\mathbf{w}\ \mathbf{R})$ to give:

$$\begin{pmatrix} \mathbf{A} & \omega & \mathbf{C} \\ \omega^T & IJ & 0 \\ \mathbf{C}^T & 0 & 0 \end{pmatrix}\times\begin{pmatrix} \mathbf{P} & -p\mathbf{w} & \mathbf{R} \\ -p\mathbf{w}^T & p & 0 \\ \mathbf{R}^T & 0 & 0 \end{pmatrix}=$$

$$\begin{pmatrix} \mathbf{AP}-p\omega\mathbf{w}^T+\mathbf{CR}^T & -p\mathbf{Aw}+p\omega & \mathbf{AR} \\ \omega^T\mathbf{P}-pIJ\mathbf{w}^T & -p\omega^T\mathbf{w}+pIJ & \omega^T\mathbf{R} \\ \mathbf{C}^T\mathbf{P} & -p\mathbf{C}^T\mathbf{w} & \mathbf{C}^T\mathbf{R} \end{pmatrix} \qquad (6)$$

where $\omega^T=(\delta_1\mathbf{1}^T,\gamma_1\mathbf{1}^T,\delta_2\mathbf{1}^T,\gamma_2\mathbf{1}^T)$ and $\mathbf{w}^T=(\mathbf{u}_1^T,\mathbf{v}_1^T,\mathbf{u}_2^T,\mathbf{v}_2^T)$.

The above choices of $\mathbf{R}^*$ and $\mathbf{Q}$ have already ensured that the last two columns of blocks are part of a unit matrix, so it remains only to choose $\mathbf{P}$ to ensure that the first column is also of the correct form. From the middle column of (6) we have the identities:

$$\begin{aligned} \mathbf{Aw}&=\omega \\ p(IJ-\omega^T\mathbf{w})&=1 \\ \mathbf{C}^T\mathbf{w}&=0. \end{aligned} \qquad (7)$$

In analogy with the single-term Tukey model discussed in §2.1, we conjecture that $\mathbf{P}=\mathbf{I}+p\mathbf{w}\mathbf{w}^T+\mathbf{E}$, where $\mathbf{E}$ is "simple". It follows that the first column of the right hand side of (6) may be written:

$$\begin{aligned} \mathbf{I}=\mathbf{AP}-p\omega\mathbf{w}^T+\mathbf{CR}^T&=\mathbf{A}+\mathbf{AE}+\mathbf{CR}^T \quad \text{(i)} \\ 0=\omega^T\mathbf{P}-pIJ\mathbf{w}^T&=\omega^T-\mathbf{w}^T+\omega^T\mathbf{E} \quad \text{(ii)} \\ 0=\mathbf{C}^T\mathbf{P}&=\mathbf{C}^T+\mathbf{C}^T\mathbf{E} \quad \text{(iii)}. \end{aligned} \qquad (8)$$

Pre-multiplying (i) by $\mathbf{w}^T$ and using the identities, yields $\omega^T\mathbf{E}+\omega^T=\mathbf{w}^T$, which is (ii). Thus only (i) and (iii) have to be solved for $\mathbf{E}$. Evaluating $\mathbf{CR}^T$ gives a matrix whose terms are compatible with those of A-I and which maintains that form when pre-multiplied by $\mathbf{A}$. We therefore expect $\mathbf{E}$ (from (i)) to have this form. Recalling that by its definition $\mathbf{E}$ must be symmetric, we set:

$$\mathbf{E}=\begin{pmatrix} \lambda_{11}\mathbf{c}_1\mathbf{c}_1^T+\lambda_{12}\mathbf{c}_2\mathbf{c}_2^T & \nu_{11}\mathbf{c}_1\mathbf{d}_1^T+\nu_{12}\mathbf{c}_2\mathbf{d}_2^T & \alpha_{12}\mathbf{c}_2\mathbf{c}_1^T & \beta_{12}\mathbf{c}_2\mathbf{d}_1^T \\ \nu_{11}\mathbf{d}_1\mathbf{c}_1^T+\nu_{12}\mathbf{d}_2\mathbf{c}_2^T & \mu_{11}\mathbf{d}_1\mathbf{d}_1^T+\mu_{12}\mathbf{d}_2\mathbf{d}_2^T & \beta_{21}\mathbf{d}_2\mathbf{c}_1^T & \gamma_{12}\mathbf{d}_2\mathbf{d}_1^T \\ \alpha_{12}\mathbf{c}_1\mathbf{c}_2^T & \beta_{21}\mathbf{c}_1\mathbf{d}_2^T & \lambda_{21}\mathbf{c}_1\mathbf{c}_1^T+\lambda_{22}\mathbf{c}_2\mathbf{c}_2^T & \nu_{21}\mathbf{c}_1\mathbf{d}_1^T+\nu_{22}\mathbf{c}_2\mathbf{d}_2^T \\ \beta_{12}\mathbf{d}_1\mathbf{c}_2^T & \gamma_{12}\mathbf{d}_1\mathbf{d}_2^T & \nu_{21}\mathbf{d}_1\mathbf{c}_1^T+\nu_{22}\mathbf{d}_2\mathbf{c}_2^T & \mu_{21}\mathbf{d}_1\mathbf{d}_1^T+\mu_{22}\mathbf{d}_2\mathbf{d}_2^T \end{pmatrix}$$

Substituting $\mathbf{E}$ into (i) yields twelve equations with 16 unknown coefficients and it may be verified that if these twelve equations are satisfied then so is (ii). Substituting into (iii) yields twelve more equations. There is a solution for the 16 unknowns that is consistent with all 24 equations, verifying that the conjectured form of $\mathbf{P}$ must be correct. This solution is:

$$\begin{aligned}
\lambda_{11} &= \mu_{11} = -3/4 \\
\lambda_{12} &= \mu_{12} = (3-\rho^4)/(\rho^2-\rho^{-2})^2 \\
\lambda_{21} &= \mu_{21} = (3-\rho^{-4})/(\rho^2-\rho^{-2})^2 \\
\lambda_{22} &= \mu_{22} = -3/4 \\
\alpha_{12} &= \gamma_{12} = -(\rho^2+\rho^{-2})/(\rho^2-\rho^{-2})^2 \\
\beta_{12} &= \beta_{21} = -2/(\rho^2-\rho^{-2})^2 \\
\nu_{11} &= \nu_{22} = 1/4 \\
\nu_{12} &= 2\rho^{-2}/(\rho^2-\rho^{-2})^2 \\
\nu_{21} &= 2\rho^2/(\rho^2-\rho^{-2})^2
\end{aligned} \tag{9}$$

These coefficients determine $\mathbf{E}$ and then $\mathbf{P} = \mathbf{I} + p\mathbf{w}\mathbf{w}^T + \mathbf{E}$ is fully known, completing the inverse of $\mathbf{L}$. Recall that for practical use these results have to be transformed back into the original definitions of $\mathbf{c}_1$, $\mathbf{d}_1$, $\mathbf{c}_2$, and $\mathbf{d}_2$ and the matrix $\mathbf{P}$ pre- and post-multiplied by Δ^{-1} as well as by an estimate of the variance σ^2.

2.3 The General Tukey Model: $B(m,^*,^*,\pi_r)$

The model is now:

$$\mathbf{Y} = m\mathbf{1}\mathbf{1}^T + \sum_{u=1}^{r} \mathbf{c}_u \mathbf{d}_u^T \tag{10}$$

We may proceed as in the two-term case but now there are r equality constraints $C_u = D_u$ $(u = 1, 2, \ldots, r)$ and $r(r-1)$ orthogonality constraints The corresponding inverse matrix $\mathbf{P}$ is of dimension $r(I+J)$ and, as before, may be written $\mathbf{P} = \mathbf{I} + p\mathbf{w}\mathbf{w}^T + \mathbf{E}$.

We define $\rho_{ij} = (C_j/C_i)^{1/2}$ and write the symmetric matrix $\mathbf{E}$ in the partitioned form corresponding to the r multiplicative terms:

$$\mathbf{E} = \begin{pmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} & \cdots & \cdots & \mathbf{E}_{1r} \\ \mathbf{E}_{21} & \mathbf{E}_{22} & \cdots & \cdots & \mathbf{E}_{2r} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{E}_{r1} & \mathbf{E}_{r2} & \vdots & \cdots & \mathbf{E}_{rr} \end{pmatrix}$$

where $\mathbf{E}_{uv}$ corresponds to the parameters $(\mathbf{c}_u, \mathbf{d}_u)$ for the rows and $(\mathbf{c}_v^T, \mathbf{d}_v^T)$

for the columns. The detailed structures of the $\mathbf{E}_{uv}$ matrices is given by:

$$\mathbf{E}_{uu} = \sum_{q=1}^{r} \begin{pmatrix} \lambda_{uq}\mathbf{c}_q\mathbf{c}_q^T & \nu_{uq}\mathbf{c}_q\mathbf{d}_q^T \\ \nu_{uq}\mathbf{d}_q\mathbf{c}_q^T & \mu_{uq}\mathbf{d}_q\mathbf{d}_q^T \end{pmatrix}$$

$$\mathbf{E}_{uv} = \begin{pmatrix} \alpha_{uv}\mathbf{c}_v\mathbf{c}_u^T & \beta_{uv}\mathbf{c}_v\mathbf{d}_u^T \\ \beta_{vu}\mathbf{d}_v\mathbf{c}_u^T & \gamma_{uv}\mathbf{d}_v\mathbf{d}_u^T \end{pmatrix} \text{ for } u \neq v$$

In this representation the coefficients α_{uv} and γ_{uv} must be symmetric. Indeed, it turns out that the coefficients β_{uv} are also symmetric and, as in the two-term case of §2.2, $\alpha_{uv} = \gamma_{uv}$. The full solution is:

$$\begin{aligned}
p &= (I - \sum_{u=1}^{r} \gamma_u^2)^{-1}(J - \sum_{u=1}^{r} \delta_u^2)^{-1} \\
\mathbf{w}^T &= (\mathbf{u}_1^T, \mathbf{v}_1^T, \ldots, \mathbf{u}_r^T, \mathbf{v}_r^T)
\end{aligned} \tag{11}$$

The coefficients of $\mathbf{E}$ are given by:

$$\begin{aligned}
\alpha_{uv} = \gamma_{uv} &= \frac{-(\rho_{uv}^2 + \rho_{vu}^2)}{(\rho_{uv}^2 - \rho_{vu}^2)^2} & \beta_{uv} &= \frac{-2}{(\rho_{uv}^2 - \rho_{vu}^2)^2} \\
\lambda_{uv} = \mu_{uv} &= \frac{3 - \rho_{uv}^4}{(\rho_{uv}^2 - \rho_{vu}^2)^2} & \nu_{uv} &= \frac{2\rho_{vu}^2}{(\rho_{uv}^2 - \rho_{vu}^2)^2} \quad u \neq v \\
\lambda_{uu} = \mu_{uu} &= -3/4 & \nu_{uu} &= 1/4
\end{aligned} \tag{12}$$

These results are consistent with those of (9) but appear simpler because the inverse property $\rho_{vu} = \rho_{uv}^{-1}$ avoids inverses in the definition of the parameters. Also, the symmetric natures of α_{uv}, β_{uv} and γ_{uv} are now self-evident.

3 The General Biadditive Model: $B(m, \mathbf{a}, \mathbf{b}, \pi_r)$

There are now $I + J$ further parameters and $2(r+1)$ further Lagrange multipliers corresponding to the identification constraints $\mathbf{1}^T\mathbf{a} = \mathbf{1}^T\mathbf{b} = \gamma_u = \delta_u = 0$ $(u = 1, 2, \ldots, r)$. Very similar algebra to that of §2 may be used to obtain the asymptotic variances and covariances of the parameters of (1). However, (6) is replaced by:

$$\begin{pmatrix} \mathbf{A} & \omega & \mathbf{C} \\ \omega^T & IJ & \mathbf{0} \\ \mathbf{C}^T & \mathbf{0} & \mathbf{0} \end{pmatrix} \times \begin{pmatrix} \mathbf{P} & \mathbf{0} & \mathbf{R} \\ \mathbf{0} & p & \mathbf{g}^T \\ \mathbf{R}^T & \mathbf{g} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{AP} + \mathbf{CR}^T & p\omega + \mathbf{Cg} & \mathbf{AR} + \omega\mathbf{g}^T \\ \omega^T\mathbf{P} & pIJ & \omega^T\mathbf{R} + IJ\mathbf{g}^T \\ \mathbf{C}^T\mathbf{P} & 0 & \mathbf{C}^T\mathbf{R} \end{pmatrix} \tag{13}$$

where expressing the first two columns of blocks as part of a unit matrix, in a similar manner to that used above for the general Tukey model, yields the non-null vector $\mathbf{g}$ (previously null) and that the vector corresponding to the previously non-null $\mathbf{w}$ is now null. Specifically $\omega^T = (0,\ldots,0,I\mathbf{1}^T,J\mathbf{1}^T)$ and $\mathbf{g} = (-I^{-1},-J^{-1},0,\ldots,0)$. The matrix $\mathbf{R}$ now has $2r+2$ columns additional to those of the Tukey model, needed to incorporate Lagrange multipliers for the new identification constraints $\mathbf{1}^T\mathbf{a} = \mathbf{1}^T\mathbf{b} = 0$, $\mathbf{1}^T\mathbf{c}_u = \mathbf{1}^T\mathbf{d}_u = 0$ $(u = 1,2,\ldots,r)$. Also $\mathbf{A}$ now has two extra row- and column-blocks for the additive constants $\mathbf{a}$ and $\mathbf{b}$. The second column of (13) yields:

$$p = (IJ)^{-1} \quad (14)$$

$$p\omega + \mathbf{C}\mathbf{g} = 0. \quad (15)$$

Expressing the first column-block to be part of a unit matrix gives:

$$\begin{aligned} \mathbf{AP} + \mathbf{CR}^T &= \mathbf{I} \quad \text{(i)} \\ \omega^T\mathbf{P} &= 0 \quad \text{(ii)} \\ \mathbf{C}^T\mathbf{P} &= 0 \quad \text{(iii)} \end{aligned} \quad (16)$$

which may be compared with (8). Pre-multiplying (iii) by $\mathbf{g}^T$ and using (15) gives (ii), so again only (i) and (iii) have to be satisfied. Evaluating $\mathbf{CR}^T$ yields a matrix the form of which suggests the general form of $\mathbf{P}$, which on using (i) and (iii), and ignoring rows and columns pertaining to the Lagrange multipliers, gives the variance matrix;

$$\begin{pmatrix} (IJ)^{-1} & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & J^{-1}\mathbf{P}_I & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & I^{-1}\mathbf{P}_J & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \mathbf{F}+\mathbf{E}_{11} & \mathbf{E}_{12} & \ldots & \mathbf{E}_{1r} \\ 0 & 0 & 0 & \mathbf{E}_{21} & \mathbf{F}+\mathbf{E}_{22} & \ldots & \mathbf{E}_{2r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{E}_{r1} & \mathbf{E}_{r2} & \ldots & \mathbf{F}+\mathbf{E}_{rr} \end{pmatrix} \begin{matrix} m \\ \mathbf{a} \\ \mathbf{b} \\ \mathbf{c}_1,\mathbf{d}_1 \\ \mathbf{c}_2,\mathbf{d}_2 \\ \\ \mathbf{c}_r,\mathbf{d}_r \end{matrix} \quad (17)$$

where $\mathbf{P}_N = (\mathbf{I}_N - \frac{1}{N}\mathbf{1}\mathbf{1}^T)$, $\mathbf{F} = \begin{pmatrix} \mathbf{P}_I & 0 \\ 0 & \mathbf{P}_J \end{pmatrix}$, and the $\mathbf{E}_{uu}$ and $\mathbf{E}_{uv}$ are defined as above in the general case of the Tukey model and have precisely the same form for the coefficients of the parameters. However, in the two cases the values of ρ_{uv} are ratios of singular values of different matrices, so the coefficients have different numerical values. As with the formulae of §2.2, the working values of $\mathbf{c}_u$ and $\mathbf{d}_u$ must be re-expressed in terms of their original definitions

The first three diagonal values give the variances of the additive parameters m, $\mathbf{a}$ and $\mathbf{b}$. As noted by Chadoeuf and Denis [1], the estimates of the two sets of additive parameters are independent of each other and of all multiplicative parameters.

4 The General Row Regression Model: $B(m,\mathbf{a},^*,\pi_r)$

The terminology is due to Mandel [9] who first studied this member of the biadditive family but the model, albeit not in an algebraically explicit form, has a fuller history (see e.g. Yates, and Cochran [14] and Finlay and Wilkinson [4]) in the applied literature. Similar arguments to those used above, but with the appropriate variants to formulae (11) and (13), give the asymptotic dispersion matrix (18), for the parameters of (17).

$$\mathbf{P} = \begin{pmatrix} (IJ)^{-1} & 0 & 0 & 0 & \ldots & 0 \\ 0 & J^{-1}\mathbf{P}_I & 0 & 0 & \ldots & 0 \\ 0 & 0 & \mathbf{F}+\mathbf{E}_{11} & \mathbf{E}_{12} & \ldots & \mathbf{E}_{1r} \\ 0 & 0 & \mathbf{E}_{21} & \mathbf{F}+\mathbf{E}_{22} & \ldots & \mathbf{E}_{2r} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \mathbf{E}_{r1} & \mathbf{E}_{r2} & \ldots & \mathbf{F}+\mathbf{E}_{rr} \end{pmatrix} \begin{matrix} m \\ \mathbf{a} \\ \mathbf{c}_1,\mathbf{d}_1 \\ \mathbf{c}_2,\mathbf{d}_2 \\ \\ \mathbf{c}_r,\mathbf{d}_r \end{matrix} \quad (18)$$

where the $\mathbf{E}_{uv}$ are defined as before and now $\mathbf{F} = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{P}_J \end{pmatrix}$.

It is clear that the result for the generalized Fisher model $B(^*,^*,^*,\pi_r)$ is as in (18), omitting the rows and columns corresponding to m and $\mathbf{a}$, with $\mathbf{F} = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix}$. Thus, this gives the asymptotic dispersions for the r-dimensional approximation to $\mathbf{Y}$ as given by the theorem of Eckart and Young (1936).

5 Conclusion

All the results given above refer to Parameterization 1 as described in §1. The corresponding results for Parameterization 2 are easily obtained by using the Jacobean $\mathbf{J}$ of the transformation from the first parameterization to the second. We do not give the details here but the results may be summarized as follows:

(i) The asymptotic variances of the singular-values σ_u $(u = 1,2,\ldots,r)$ are all equal to the data-variance σ^2. Also Covariance $(\sigma_u,\sigma_v) = 0$ for $u \neq v$.

(ii) The coefficients of (12) remain unaltered except that now $\lambda_{uu} = \mu_{uu} = -1$ and $\nu_{uu} = 0$ (previously - 3/4 and 1/4, respectively).

In these formulae it is now convenient to define $\rho_{uv} = (\sigma_v/\sigma_u)^{1/2}$. To return to the original scaling divide each E_{uu} by σ_u^2 and E_{uv} by $(\sigma_u\sigma_v)^{3/2}$ as well as multiply by σ^2.

The differences in the formulae for the two parameterizations are a consequence of lengths of the vectors in Parameterization 1 being scaled to equal the corresponding singular value, while in Parameterization 2 they are of unit length.

Finally it has been checked that these asymptotic variances and covariances are consistent with the results already obtained for $B(m, \mathrm{a}, \mathrm{b}, \pi_r)$ by Chadoeuf and Denis [1] for the case $r = 1$ and by Goodman and Haberman [6] for general r but for the variances only. The methods used in this paper differ completely from those use in the two papers just cited.

References

[1] J. Chadoeuf and J.B. Denis, Asymptotic variances for the multiplicative interaction model. *Journal of Applied Statistics* 18(1991), 331–353.

[2] J.B. Denis and J.C. Gower, Biadditive models. Laboratoire de Biometrie INRA, Versailles, France, 1992, 33pp.

[3] C. Eckart and G. Young, The approximation of one matrix by another of lower rank. *Psychometrika* 1(1936), 211–218.

[4] K.W. Finlay and G.N. Wilkinson, The analysis of adaptation in a plant breeding programme. *Australian Journal of Agricultural Research* 14(1963), 742–754.

[5] R.A. Fisher and W.A. Mackenzie, Studies in crop variation II. The manurial response of different potato varieties. *Journal of Agricultural Science* XIII(1923), 311–320.

[6] L.A. Goodman and S.J. Haberman, The analysis of nonadditivity in two-way analysis of variance. *Journal of the American Statistical Association* 85(1990), 139–145.

[7] J.C. Gower, A modified Leverrier-Faddeev algorithm for matrices with multiple eigenvalues. *Linear Algebra and its Applications* **31** (1980), 61–70.

[8] J.C. Gower and P.J.F. Groenen, Applications of the modified Leverrier-Faddeev algorithm for the construction of explicit matrix spectral decompositions and inverses. *Utlilitas Mathematica* 40(1991), 51–64.

[9] J. Mandel, Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **56**(1961), 878–888.

[10] M. Seyedsadr and P.L. Cornelius, Shifted multiplicative models for non-additive two-way tables. *Communications in Statistical Simulation* **21**(3)(1992), 807–832.

[11] S.D. Silvey, The Lagrange multiplier test. *Annals of Mathematical Statistics* 30(1959), 389–407.

[12] S.D. Silvey, "Statistical inference". Chapman and Hall: London, 1975, 191pp.

[13] J.W. Tukey, One degree of freedom for non-additivity. *Biometrics* 5(1949), 232–242.

[14] F. Yates and W.G. Cochran, The analysis of groups of experiments. *Journal of Agricultural Science Cambridge* **28** (1938), 556–580.