Austral. J. Statist., 30(2), 1988, 228–234

CHOLESKY DECOMPOSITION OF A VARIANCE MATRIX IN REPEATED MEASURES ANALYSIS

S. LIANTO

Institut Teknologi Sepuluh Nopember Surabaya, Surabaya, Indonesia

AND

C.A. McGILCHRIST

Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

Summary

The Cholesky decomposition is given for the inverse of a variance matrix occurring in repeated measures problems where observations have a correlation structure both within and between experimental units. The use of this decomposition is outlined for ML and REML estimation procedures.

1. Introduction

In repeated measures studies, observations within the same experimental unit are usually correlated. Sometimes the correlation is considered to be the same for all pairs of observations on the same unit as in split-plot-in-time analyses but more often some stationary error structure is assumed. The units themselves are often considered to be independent but sometimes are considered as being sampled from some hierarchical structure.

Two examples in recent literature are mentioned. Sallas and Harville (1981) consider repeated measures of total lactations of cows. These lactations are taken to have a stationary error structure within each cow and the cows themselves are sampled from offspring of several sires. Other effects are taken as 'fixed effects'. Pantula and Pollock (1985) consider two examples of total leaf area of plants and bobcat movements in which observations on each unit viz. plant/bobcat are serially correlated. In their case no correlation is considered among the experimental units. Sallas and Harville advance a Kalman filter type analysis for their data while Pantula and Pollock use a direct technique related to unbiased quadratic estimation.

Such experiments may be analysed using maximum likelihood or 'residual maximum likelihood (REML)' techniques. The general method for maximum likelihood is given in McGilchrist *et al.* (1981). The REML technique was first given by Patterson and Thompson (1971) and its current use is implicit in Harville (1977).

Section 2 gives the model taken for the observations and presents an outline of the implementation of the ML or REML techniques. This method depends on obtaining a Cholesky factorisation of the inverse of the variance matrix. Such a factorisation is obtained in Section 3.

2. Model and Method of Estimation

Let Y_{ijk} be the kth repeated measure on an experimental unit indexed by (i,j). The model for Y_{ijk} allows fixed effects summarised by a regression vector giving

$$Y_{ijk} = x'_{ijk}\beta + E_{ijk}\,, \qquad i = 1,2,\ldots,I;\; j = 1,2,\ldots,n(i);\; k = 1,2,\ldots,n(i,j)\,.$$

If y is the accumulation of observations Y_{ijk} into a vector of length N, in the order of subscripts given above, and X is the corresponding matrix of regression variables then the whole model may be written as $y = X\beta + e$ with the error vector e distributed as $N[0,\sigma^2\Sigma]$. The covariance matrix Σ is taken to be a function of a parametrising vector θ.

To be specific the covariance matrix Σ is taken to be that which applies when experimental units (i,j) have the nested error structure appropriate for Sallas and Harville. Let 1_{ijk} be a vector of length N with a 1 at position (i,j,k) and zeros elsewhere. A dot replacing a subscript in this vector indicates summation with respect to that subscript. Let

$$A_{100} = \sum_i 1_{i\ldots}1'_{i\ldots}\,, \quad A_{110} = \sum_{i,j} 1_{ij\cdot}1'_{ij\cdot}\,.$$

$$\Gamma = \begin{bmatrix} \gamma_{11} & & & & & & \\ & \gamma_{12} & & & & & \\ & & \ddots & & & & \\ & & & \gamma_{1,n(1)} & & & \\ & & & & \gamma_{21} & & \\ & & & & & \ddots & \\ & & & & & & \gamma_{I,n(I)} \end{bmatrix}$$

$$\gamma_{ij} = \begin{bmatrix} \sigma_0 & \sigma_1 & \sigma_2 & \dots & \\ \sigma_1 & \sigma_0 & \sigma_1 & \dots & \\ \sigma_2 & \sigma_1 & \sigma_0 & \dots & \\ \dots & \dots & \sigma_1 & \sigma_0 & \\ \dots & \dots & & & \\ \dots & \dots & & & \sigma_1 \sigma_0 \end{bmatrix} .$$

Matrix γ_{ij} is symmetric Toeplitz of dimension $n(i,j)$ and elements $\sigma_0, \sigma_1, \dots$ are taken to be functions of $(\theta_4, \theta_5, \dots, \theta_p)$. The whole variance matrix is

$$\operatorname{var} y = \sigma^2 \Sigma = \sigma^2 (I_N + \theta_1 A_{100} + \theta_2 A_{110} + \theta_3 \Gamma)$$

and has parameters σ^2, θ.

The likelihood function L of the observation vector y distributed according to the above model is

$$l = \ln L = -\tfrac{1}{2}[N \ln 2\pi\sigma^2 + \ln|\Sigma| + \sigma^{-2}(y - X\beta)'\sigma^{-1}(y - X\beta)]$$

and the corresponding residual likelihood may be obtained from Harville. Both likelihoods require the evaluation of the determinant and inverse of Σ. To avoid doing this directly, the Cholesky factorisation of Σ^{-1}, is obtained as follows. Let

$$\Sigma_{n+1} = \begin{bmatrix} \Sigma_n & \sigma_n \\ \sigma_n' & \sigma_{n+1,n+1} \end{bmatrix}, \quad \begin{aligned} g_n &= \Sigma_n^{-1} \sigma_n \\ c_{n+1} &= \sigma_{n+1,n+1} - \sigma_n' g_n \end{aligned}$$

then

$$\begin{aligned} \Sigma_{n+1}^{-1} &= \begin{bmatrix} \Sigma_n^{-1} & 0_n \\ 0_n' & 0 \end{bmatrix} + c_{n+1}^{-1} \begin{bmatrix} -g_n' \\ 1 \end{bmatrix} [-g_n', \quad 1] \qquad (2.1) \\ |\Sigma_{n+1}| &= c_{n+1} |\Sigma_n| . \end{aligned}$$

If

$$G = \begin{bmatrix} 1 & -g_{11} & -g_{12} & \dots & -g_{1,N-1} \\ & 1 & -g_{22} & \dots & -g_{2,N-1} \\ & & 1 & \dots & -g_{3,N-1} \\ & & & & \\ & & & & 1 \end{bmatrix}, \quad C = \begin{bmatrix} c_1 & & \\ & c_2 & \\ & & \\ & & c_N \end{bmatrix}$$

where g_{in} is the ith element of g_n then, $\Sigma^{-1} = GC^{-1}G'$, $|\Sigma| = |C|$ and letting $r = G'y$, $z = G'X$ the logarithm of the likelihood becomes

$$l = -\tfrac{1}{2}\left[N \ln 2\pi\sigma^2 + \ln|C| + \sigma^{-2}(r - Z\beta)'C^{-1}(r - Z\beta)\right] .$$

A residual likelihood follows similarly. A detailed recursive estimation procedure is given in McGilchrist (1986). The method depends on being able to compute suitable expressions for g_n, c_n and this is the topic of section 3.

3. Cholesky Factorisation

The recurrence relations which may be used to generate g_n, c_n, from initial values of g_0 = null and $c_1 = \sigma_0$, vary with the different ways in which Σ_n is extended to Σ_{n+1}. Three cases are considered.

$$\begin{aligned} &\text{Case 1:}\quad n \to n+1 \text{ is } (i,j,k) \to (i,j,k+1)\\ &\text{Case 2:}\quad n \to n+1 \text{ is } (i,j,n(i,j)) \to (i,j+1,1)\\ &\text{Case 3:}\quad n \to n+1 \text{ is } (i,n(i),n(i,j)) \to (i+1,1,1)\end{aligned}$$

These cases correspond to staying within an experimental unit, changing to a new unit within the same block, changing to a unit in a new block. The following results prove useful in the sequel.

Lemma 3.1. *Let m be any integer less than n such that $m = n - p$ and*

$$\Sigma_n = \begin{bmatrix} \Sigma_m & \sigma_{mp} \\ \sigma'_{mp} & \Sigma_{mp} \end{bmatrix} .$$

It may be shown by direct multiplication that

$$\Sigma_n^{-1} = \begin{bmatrix} \Sigma_m^{-1} & 0_{mp} \\ 0'_{mp} & 0_{pp} \end{bmatrix} + \begin{bmatrix} -\Sigma_m^{-1}\sigma_{mp} \\ I_p \end{bmatrix} C_{mp}^{-1}[-\sigma'_{mp}\Sigma_m^{-1}\, ,\, I_p]$$

where $C_{mp} = \Sigma_{mp} - \sigma'_{mp}\Sigma_m^{-1}\sigma_{mp}$.

Corollary: *For $p = 1$ the relationship becomes equation 2.1.*

Lemma 3.2. *If $\sigma_{mp} = \sigma_m 1'_p$ the expression in Lemma 3.1 becomes*

$$\Sigma_n^{-1} = \begin{bmatrix} \Sigma_m^{-1} & 0_{mp} \\ 0'_{mp} & 0_{pp} \end{bmatrix} + \begin{bmatrix} -g_m 1'_p \\ I_p \end{bmatrix} C_{mp}^{-1}[-1_p g'_m\, ,\, I_p]$$

and $C_{mp} = \Sigma_{mp} - \sigma'_m\Sigma_m^{-1}\sigma_m 1_p 1'_p = \Sigma_{mp} - (c_{m+1} - \sigma_{mm})1_p 1'_p$.

3.1 Case 1. If $n \to n+1$ remains within an experimental unit consider two partitions

$$\Sigma_{n+1} = \begin{bmatrix} \Sigma_m & \sigma_{m,p+1} \\ \sigma'_{m,p+1} & \Sigma_{m,p+1} \end{bmatrix}, \quad \Sigma_n = \begin{bmatrix} \Sigma_m & \sigma_{mp} \\ \sigma'_{mp} & \Sigma_{mp} \end{bmatrix}$$

where p is the number of observations in the current experimental unit (i,j) contained in y_n. Thus

$$\begin{aligned} \sigma_{mp} &= \sigma_m 1'_p , \quad \sigma_{m,p+1} = \sigma_m 1'_{p+1} \\ \Sigma_{mp} &= I_p + (\theta_1 + \theta_2) 1_p 1'_p + \theta_3 \gamma_{p \text{ rows, columns}} \end{aligned}$$

where $\gamma_{p\text{, rows,columns}}$ is the extraction of the first p rows and columns of γ_{ij}. Similarly

$$\Sigma_{m,p+1} = I_{p+1} + (\theta_1 + \theta_2) 1_{p+1} 1'_{p+1} + \theta_3 \gamma_{p+1 \text{ rows,columns}} \cdot$$

Using Lemma 3.2

$$\begin{aligned} \Sigma^{-1}_{n+1} &= \begin{bmatrix} \Sigma^{-1}_m & 0_{m,p+1} \\ 0'_{m,p+1} & 0_{p+1,p+1} \end{bmatrix} + \begin{bmatrix} -g_m \, 1'_{p+1} \\ I_{p+1} \end{bmatrix} C^{-1}_{m,p+1} [-1_{p+1} g'_m \, , I_{p+1}] \\ C_{m,p+1} &= \Sigma_{m,p+1} - (c_{m+1} - \sigma_{mm}) 1_{p+1} 1'_{p+1} \, . \end{aligned}$$

Equating the expression for Σ^{-1}_{n+1} to that given in equation 2.1 leads to recurrences for c_n, g_n. Before doing so, note that $C_{m,p+1}$ is symmetric Toeplitz so that there exists a_{mp}, V_{mp} such that

$$C^{-1}_{m,p+1} = \begin{bmatrix} C^{-1}_{mp} & 0_p \\ 0'_p & 0 \end{bmatrix} + V^{-1}_{mp} \begin{bmatrix} -a_{mp} \\ 1 \end{bmatrix} [-a'_{mp} \, , 1]$$

and then we have

$$\begin{aligned} c_{n+1} &= V_{mp} \\ g'_n &= [(1 - 1'_p a_{mp}) g'_m \, , a'_{mp}] \, . \end{aligned}$$

Recursive expressions for a_{mp}, V_{mp} for a Toeplitz matrix may be found in McGilchrist and Sandland (1979) or Trench (1967).

3.2. Case 2. If $n \to n+1$ goes from one experimental unit to the next but remains within the same block of units then we choose p equal to the number of observations on the just completed experimental unit. The condition $\sigma_{mp} = \sigma_m 1'_p$ of Lemma 3.2 is satisfied and additionally

$$\sigma'_n = [\sigma'_m\,,\ \theta_1 1'_p]$$

so that

$$\begin{aligned} g_n = \Sigma_n^{-1}\sigma_n &= \left\{\begin{bmatrix}\Sigma_m^{-1} & 0_{mp}\\ 0'_{mp} & 0_{pp}\end{bmatrix} + \begin{bmatrix}-g_m\, 1'_p\\ I_p\end{bmatrix} C_{mp}^{-1}[-1_p g'_m\,,\ I_p]\right\}\begin{bmatrix}\sigma_m\\ \theta_1 1_p\end{bmatrix}\\ &= \begin{bmatrix}g_m\\ 0_p\end{bmatrix} + (\theta_1 + c_{m+1} - \sigma_{mm})\begin{bmatrix}-g_m\, 1'_p\\ I_p\end{bmatrix} C_{mp}^{-1}1_p\\ &= \begin{bmatrix}(1-\kappa_{mp}) & g_m\\ \kappa_{mp} & r_{mp}\end{bmatrix}\end{aligned}$$

where $\kappa_{mp} = (\theta_1 + c_{m+1} - \sigma_{mm})1'_p C_{mp}^{-1}1_p$, $r_{mp} = (C_{mp}^{-1}1_p)(1'_p C_{mp}^{-1}1_p)^{-1}$. Using the results of Case 1, the vector $C_{mp}^{-1}1_p$ may be built up recursively by

$$C_{m,p+1}^{-1}1_{p+1} = \begin{bmatrix}C_{mp}^{-1}\, 1_p\\ 0\end{bmatrix} + V_{mp}^{-1}(1 - a'_{mp}1_p)\begin{bmatrix}-a_{mp}\\ 1\end{bmatrix}\,.$$

On substituting into κ_{mp}, r_{mp} expressions we have

$$\begin{aligned} \kappa_{m,p+1} &= \kappa_{mp} + (c_{m+1} + \theta_1 - \sigma_{mm})V_{mp}^{-1}(1 - a'_{mp}1_p)^2\\ r_{m,p+1} &= \kappa_{m,p+1}^{-1}\left\{\kappa_{mp}\begin{bmatrix}r_{mp}\\ 0\end{bmatrix}\right.\\ &\qquad\qquad \left. +[c_{m+1} + \theta_1 - \sigma_{mm}]^{-1}V_{mp}^{-1}(1 - a'_{mp}\, 1_p]\begin{bmatrix}-a_{mp}\\ 1\end{bmatrix}\right)\,.\end{aligned}$$

We also find

$$c_{n+1} = [1 - \kappa_{mp}]c_{m+1} + \kappa_{mp}[\sigma_{mm} - \theta_1]\,.$$

3.3 Case 3. If $n \to n+1$ goes to the first observation on an experimental unit in a new block then $\sigma_n = 0_n$ and

$$g_n = 0_n\,,\quad c_{n+1} = 1 + \theta_1 + \theta_2 + \theta_3\sigma_0\,.$$

3.4 Discussion. If experimental units are independent then case 2 is deleted. The results would then apply to an experiment such as those of

Pantula and Pollock. All three types of recurrences are required when the experiments have a blocked structure as in Sallas and Harville.

References

HARVILLE, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *J. Amer. Statist. Assoc.* **72**, 320–338.

McGILCHRIST, C.A. (1986). Regression analysis of dependent error models. *Bull. Austral. Math. Soc.* **34**, 199–210.

McGILCHRIST, C.A. & SANDLAND, R.L. (1979). Recursive estimation of the general linear model with dependent errors. *J. Roy. Statist. Soc. Ser. B* **41**, 65–68.

McGILCHRIST, C.A., SANDLAND, R.L. & HILLS, L.J. (1981). Estimation in regression models with stationary dependent errors. *Comm. Statist. A* **10**, 2563–2580.

PANTULA, S.G. & POLLOCK, K.H. (1985). Nested analysis of variance with autocorrelated errors. *Biometrics* **41**, 909–920.

PATTERSON, H.D. & THOMPSON, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554.

SALLAS, W.M. & HARVILLE, D.A. (1981). Best linear recursive estimation for mixed linear models. *J. Amer. Statist. Assoc.* **76**, 860–869.

TRENCH, W. (1967). Weighting coefficients for the prediction of stationary time series from the finite past. *SIAM J. Appl. Math.* **15**, 1502–1510.

Received June 1987; revised February 1988