MAXIMUM LIKELIHOOD ESTIMATION FOR THE GROWTH CURVE MODEL WITH UNEQUAL DISPERSION MATRICES

By

S. R. Chakravorti

Department of Biostatistics
University of North Carolina
and
University of Calcutta, India

Institute of Statistics Mimeo Series No. 893

OCTOBER 1973

Maximum-likelihood Estimation for the Growth Curve Model with Unequal Dispersion Matrices*

S. R. CHAKRAVORTI+

Department of Biostatistics, University of North Carolina, Chapel Hill, N. C. 27514

Abstract

The paper considers the maximum-likelihood estimation problem for the growth curve model — treated as MANOCOVA model — with unequal dispersion matrices. Optimality properties of the estimates have been studied and procedure for testing equality of several growth curves has been indicated.

*Work sponsored by the Aerospace Research Laboratories, U. S. Air Force Systems Command, Contract F 33615-71-C-1927. Reproduction in whole or in part permitted for any purpose of the U. S. Government.

+On leave of absence from the University of Calcutta, India.

1. INTRODUCTION

Consider the growth curve model as multivariate analysis of covariance (MANOCOVA) with stochastic predictors (Rao [9]). The problem of estimation of the parameters of this model with unequal dispersion matrices does not follow the usual procedure of either maximum likelihood method or least squares method when dispersion matrices are equal.

In the univariate situation of linear model under heteroscedastic assumption the methods of estimation have been proposed by C. R. Rao [10] and Hartley and Jaytillake [6]. The later method follows the procedure of Hartley and J. N. K. Rao [5] and is free from the defects of the MINQUE method proposed by C. R. Rao. Here the method of Hartley and Jaytillake has been generalized for the MANOCOVA model stated earlier. The proposed method of estimation yields estimates which are the solutions of maximum likelihood equations by the steepest descent method. The solutions are, in fact, the asymptotic limit to the solution of a system of first order differential equations. The asymptotic optimality properties of these estimates have been studied. Also the procedure of testing equality of several growth curves has been discussed.

2. THE MODEL

The usual growth curve model (viz., Potthoff and Roy [8]) written as MANOCOVA model, under Behrens-Fisher situation, is given by

$$\underset{\sim}{Y}_{\alpha}^{(t)} = \underset{\sim}{\eta}^{(t)} + \underset{\sim}{X}_{\alpha}^{(t)}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}_{\alpha}^{(t)} \tag{2.1}$$

where $\underset{\sim}{Y}_{\alpha}^{(t)}(1\times p)$ is the α-th observation vector in t-th sample ($\alpha=1,2,\ldots,n_t$; $t=1,\ldots,m$), $\underset{\sim}{\eta}^{(t)}(1\times p)$ is the vector of unknown constants in t-th group, which is a p-th degree polynomial in time in growth curve model, $\underset{\sim}{X}_{\alpha}^{(t)}(1\times s)$ is the α-th observation vector of concomitant variables in t-th sample (where $s = q-p$, $q \geq p$),

$\underset{\sim}{\beta}(s\times p)$, the common matrix of regression coefficients of $\underset{\sim}{Y}_\alpha^{(t)}$ on $\underset{\sim}{X}_\alpha^{(t)}$, $\underset{\sim}{\varepsilon}_\alpha^{(t)}$, the error component in t-th sample, distributed as $N_p(\underset{\sim}{0}, \underset{\sim}{\Sigma}_t)$, where $\underset{\sim}{\Sigma}_t(p\times p)$ is the conditional dispersion matrix in t-th group.

The problem is to estimate the parameters $\underset{\sim}{\eta}^{(t)}$, $\underset{\sim}{\Sigma}_t$, $\underset{\sim}{\beta}$, for $t=1,2,\ldots,m$, by the maximum likelihood method.

3. MAXIMUM LIKELIHOOD ESTIMATES OF THE PARAMETERS OF THE MODEL

The log-likelihood of the model (2.1) is given by

$$\log L = \text{const} + \tfrac{1}{2}\sum_{t=1}^{m} n_t|\underset{\sim}{\Sigma}_t^{-1}| - \tfrac{1}{2}\sum_t\sum_\alpha \underset{\sim}{\varepsilon}_\alpha^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{\varepsilon}_\alpha^{(t)'} \tag{3.1}$$

where, from (2.1), $\underset{\sim}{\varepsilon}_\alpha^{(t)} = \underset{\sim}{Y}_\alpha^{(t)} - \underset{\sim}{\eta}^{(t)} - \underset{\sim}{X}_\alpha^{(t)}\underset{\sim}{\beta}$. Then following the method of matrix derivatives (Dwyer and Mephail [2] and Dwyer [3]) we have

$$\frac{\partial \log L}{\partial \underset{\sim}{\eta}^{(t)}} = \sum_{\alpha=1}^{n_t}(\underset{\sim}{Y}_\alpha^{(t)} - \underset{\sim}{\eta}^{(t)} - \underset{\sim}{X}_\alpha^{(t)}\underset{\sim}{\beta})\underset{\sim}{\Sigma}_t^{-1},\ t=1,2,\ldots,m, \tag{3.2}$$

$$\frac{\partial \log L}{\partial \underset{\sim}{\Sigma}_t} = -\tfrac{1}{2}\, n_t\underset{\sim}{\Sigma}_t^{-1} + \tfrac{1}{2}\underset{\sim}{\Sigma}_t^{-1}\Big(\sum_{\alpha=1}^{n_t} \underset{\sim}{\varepsilon}_\alpha^{(t)'}\underset{\sim}{\varepsilon}_\alpha^{(t)}\Big)\underset{\sim}{\Sigma}_t^{-1},\ t=1,\ldots,m, \tag{3.3}$$

$$\frac{\partial \log L}{\partial \underset{\sim}{\beta}} = \sum_t\sum_\alpha \underset{\sim}{X}_\alpha^{(t)'}(\underset{\sim}{Y}_\alpha^{(t)} - \underset{\sim}{\eta}^{(t)} - \underset{\sim}{X}_\alpha^{(t)}\underset{\sim}{\beta})\underset{\sim}{\Sigma}_t^{-1} \tag{3.4}$$

For given $\underset{\sim}{\beta}$, we have the estimates of $\underset{\sim}{\eta}^{(t)}$ and $\underset{\sim}{\Sigma}_t$, by equating (3.2) and (3.3) to zero,

$$\underset{\sim}{\eta}^{(t)}(\underset{\sim}{\beta}) = \bar{\underset{\sim}{Y}}^{(t)} - \bar{\underset{\sim}{X}}^{(t)}\underset{\sim}{\beta} \tag{3.5}$$

$$\underset{\sim}{\Sigma}_t(\underset{\sim}{\beta}) = \frac{1}{n_t}\sum_\alpha \hat{\underset{\sim}{\varepsilon}}_\alpha^{(t)'}\hat{\underset{\sim}{\varepsilon}}_\alpha^{(t)} = \underset{\sim}{S}_t \text{ (say)} \tag{3.6}$$

When $\underset{\sim}{\beta}$ is unknown, to estimate $\underset{\sim}{\beta}$ we have by substituting (3.5) and (3.6) in (3.1)

$$F(\underset{\sim}{\beta}) = \text{const} - \tfrac{1}{2}\sum_{t=1}^{m} n_t \log|\underset{\sim}{S}_t| = \tfrac{1}{2}\sum_{t=1}^{m} \varphi_t(\underset{\sim}{\beta}) \tag{3.7}$$

where $\varphi_t(\underset{\sim}{\beta})$ is some function of $\underset{\sim}{\beta}$. Then the estimate of $\underset{\sim}{\beta}=((\beta_{mn}))$ is obtained by solving

$$\frac{\partial F}{\partial \underset{\sim}{\beta}} = \sum_t \sum_\alpha (\underset{\sim}{X}_\alpha^{(t)} - \bar{\underset{\sim}{X}}^{(t)})'(\underset{\sim}{Y}_\alpha^{(t)} - \bar{\underset{\sim}{Y}}^{(t)} - (\underset{\sim}{X}_\alpha^{(t)} - \bar{\underset{\sim}{X}}^{(t)})\underset{\sim}{\beta})\underset{\sim}{\Sigma}_t^{-1} = 0 \tag{3.8}$$

For this we generate an asymptotically convergent sequence from the steepest descent differential equations

$$\frac{\partial \beta_{mn}}{\partial \theta} = -\frac{\partial F(\underset{\sim}{\beta})}{\partial \beta_{mn}}, \text{ for } m=1,\ldots,s;\ n=1,\ldots,p, \tag{3.9}$$

where θ is the parameter in the parametric representation of the path of descent $\beta_{mn}(\theta)$. Then following Hartley and Jaytillake [6], we observe that for $\theta\to\infty$, the path coordinates $\beta_{mn}(\theta)$ will tend to a limit $\tilde{\beta}_{mn}$ such that

$$\left.\frac{\partial F}{\partial \beta_{mn}}\right|_{\underset{\sim}{\beta}=\tilde{\underset{\sim}{\beta}}} = 0, \text{ for all } m=1,\ldots,s;\ n=1,\ldots,p. \tag{3.10}$$

By following Runge-Kutta procedure (Henrici [7]) a local minimum of $F(\underset{\sim}{\beta})$ is attained in this case which depends on the initial values $\beta_{mn}(\theta)$ selected for the system of simultaneous equations (3.9). The estimates of $\underset{\sim}{\beta}$, $\underset{\sim}{\Sigma}_t$'s and $\underset{\sim}{\eta}^{(t)}$'s are obtainable by solving simultaneously

$$\sum_{\alpha=1}^{n_t} (\underset{\sim}{X}_\alpha^{(t)} - \bar{\underset{\sim}{X}}^{(t)})'(\underset{\sim}{Y}_\alpha^{(t)} - \bar{\underset{\sim}{Y}}^{(t)} - (\underset{\sim}{X}_\alpha^{(t)} - \bar{\underset{\sim}{X}}^{(t)})\underset{\sim}{\beta})\underset{\sim}{S}_t^{-1} = 0 \tag{3.11}$$

and (3.5) and (3.6) by feed back principle. The m.l.e.'s thus obtained are denoted by $\hat{\underset{\sim}{\eta}}^{(t)}$, $\hat{\underset{\sim}{\Sigma}}_t$ and $\hat{\underset{\sim}{\beta}}$ respectively for $\eta^{(t)}$, $\underset{\sim}{\Sigma}_t$ and $\underset{\sim}{\beta}$.

4. ASYMPTOTIC PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATORS

We have altogether 2m+1 (matrix) parameters denoted by $\theta=(\eta^{(1)},\ldots,\eta^{(m)}, \Sigma_1,\ldots,\Sigma_m, \beta)$. To prove the consistency and asymptotic efficiency of θ we assume,

(i) m = the number of groups is fixed, n_t and $n = \sum_{t=1}^{m} n_t$ are large and $r_t = n_t/n$ (for t=1,...,m) are bounded away from 0 and 1.

(ii) The stochastic concomitant vector variable $X_\alpha^{(t)}$ is distributed as $N_s(0, \Sigma_t^*)$, so that

$$E\,\bar{X}^{(t)} = 0,\quad E\sum_{\alpha=1}^{n_t} X_\alpha^{(t)'} X_\alpha^{(t)} = n_t\Sigma_t^*, \tag{4.1}$$

4.1. Consistency

The maximum likelihood estimate $\hat{\theta}=(\hat{\eta}^{(1)},\ldots,\hat{\eta}^{(m)},\hat{\Sigma}_1,\ldots,\hat{\Sigma}_m,\hat{\beta})$ of θ is consistent if

$$\lim_{n\to\infty} \text{Prob}[\hat{\theta} = \theta_0] = 1 \tag{4.2}$$

where $\theta_0=(\eta_0^{(1)},\ldots,\eta_0^{(m)},\Sigma_1^0,\ldots,\Sigma_m^0,\beta_0)$ is the true value of θ.

To prove this, let us first prove that for $\theta\neq\theta_0$

$$V_{\theta_0}[\frac{1}{n}\log L(Y|\theta)] = 0(\frac{1}{n}) \tag{4.3}$$

Let $\delta\eta^{(t)} = \eta^{(t)}-\eta_0^{(t)}$, $\delta\beta = \beta-\beta_0$, where $\eta_0^{(t)}$, β_0 and Σ_t^0 are the true values of $\eta^{(t)}$, β and Σ_t respectively. Then from (3.1)

$$\begin{aligned}\log L &= \text{const} - \tfrac{1}{2}\sum_t\sum_\alpha(Y_\alpha^{(t)}-\eta_0^{(t)}-X_\alpha^{(t)}\beta_0-\delta\gamma_\alpha^{(t)})\Sigma_t^{-1}(Y_\alpha^{(t)}-\eta_0^{(t)}-X_\alpha^{(t)}\beta_0-\delta\gamma_\alpha^{(t)})' \\ &= \text{const} - \tfrac{1}{2}[\sum_t\sum_\alpha(Y_\alpha^{(t)}-\eta_0^{(t)}-X_\alpha^{(t)}\beta_0)\Sigma_t^{-1}(Y_\alpha^{(t)}-\eta_0^{(t)}-X_\alpha^{(t)}\beta_0)' \\ &\qquad -2\sum_t\sum_\alpha(Y_\alpha^{(t)}-\eta_0^{(t)}-X_\alpha^{(t)}\beta_0)\Sigma_t^{-1}\delta\gamma_\alpha^{(t)'} + \sum_t\sum_\alpha\delta\gamma_\alpha^{(t)}\Sigma_t^{-1}\delta\gamma_\alpha^{(t)'}] \\ &= \text{const} + L_{(1)} + L_{(2)} + L_{(3)} \quad \text{(say)}\end{aligned} \tag{4.4}$$

where $\delta\underset{\sim}{\gamma}_\alpha^{(t)} = \delta\underset{\sim}{\eta}^{(t)} + \underset{\sim}{X}_\alpha^{(t)}\delta\underset{\sim}{\beta}$. Let us choose a non-singular matrix $\underset{\sim}{A}_t(p\times p)$ such that $\underset{\sim}{A}_t\underset{\sim}{\Sigma}_t^{o-1}\underset{\sim}{A}_t' = \underset{\sim}{I}$, so that $\underset{\sim}{\Sigma}_t^0 = \underset{\sim}{A}_t'\underset{\sim}{A}_t$. Also let $\underset{\sim}{Z}_\alpha^{(t)}=(\underset{\sim}{Y}_\alpha^{(t)}-\underset{\sim}{\eta}_0^{(t)}-\underset{\sim}{X}_\alpha^{(t)}\underset{\sim}{\beta}_0)\underset{\sim}{A}_t^{-1}$. Then $\underset{\sim}{Z}_\alpha^{(t)}$ is distributed as $N_p(\underset{\sim}{0}, \underset{\sim}{I}_p)$. Under this set of transformations we have from (4.4)

$$L_{(1)} = \sum_t \sum_\alpha \underset{\sim}{Z}_\alpha^{(t)}\underset{\sim}{A}_t\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{A}_t'\underset{\sim}{Z}_\alpha^{(t)'} = \sum_t \sum_\alpha \sum_{i=1}^{p} \lambda_i^{(t)} Z_{\alpha i}^{(t)^2},$$

where $\underset{\sim}{A}_t\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{A}_t' = \underset{\sim}{D}_t = \text{diag}(\lambda_1^{(t)},\ldots,\lambda_p^{(t)})$ and $\lambda_i^{(t)}$'s are the characteristic roots of $\underset{\sim}{\Sigma}_t^{-1}$ and hence finite, so that $V(L_{(1)}) = 2\sum_t n_t \sum_i \lambda_i^{(t)^2}$. Hence from assumption (i)

$$V(\frac{1}{n} L_{(1)}) = \frac{2}{n}\sum_t \sum_i \gamma_t(\lambda_i^{(t)})^2 = O(\frac{1}{n}) \tag{4.5}$$

Since for fixed $\underset{\sim}{X}_\alpha^{(t)}$, $E(L_{(2)})=0$, we have

$$\begin{aligned}
V(L_{(2)}) &= E\sum_t \sum_\alpha \underset{\sim}{Z}_\alpha^{(t)}\underset{\sim}{A}_t\underset{\sim}{\Sigma}_t^{-1}\delta\underset{\sim}{\gamma}_\alpha^{(t)'}\delta\underset{\sim}{\gamma}_\alpha^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{A}_t'\underset{\sim}{Z}_\alpha^{(t)'} \\
&= \text{tr} \sum_t \sum_\alpha \underset{\sim}{A}_t\underset{\sim}{\Sigma}_t^{-1}\delta\underset{\sim}{\gamma}_\alpha^{(t)'}\delta\underset{\sim}{\gamma}_\alpha^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{A}_t'\cdot I \\
&= \sum_t \sum_\alpha \delta\underset{\sim}{\gamma}_\alpha^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{\Sigma}_t^{0}\underset{\sim}{\Sigma}_t^{-1}\delta\underset{\sim}{\gamma}_\alpha^{(t)'} \\
&= \sum_t n_t\delta\underset{\sim}{\eta}^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{\Sigma}_t^{0}\underset{\sim}{\Sigma}_t^{-1}\delta\eta^{(t)'} + 2\sum_t n_t\delta\underset{\sim}{\eta}^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{\Sigma}_t^{0}\underset{\sim}{\Sigma}_t^{-1}\delta\underset{\sim}{\beta}'\bar{\underset{\sim}{X}}^{(t)'} \\
&\quad + \sum_t \sum_\alpha \underset{\sim}{X}_\alpha^{(t)}\delta\underset{\sim}{\beta}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{\Sigma}_t^{0}\underset{\sim}{\Sigma}_t^{-1}\delta\underset{\sim}{\beta}'\underset{\sim}{X}_\alpha^{(t)'}
\end{aligned}$$

This is conditional variance and since $\delta\underset{\sim}{\eta}^{(t)}$ and $\delta\underset{\sim}{\beta}$ are fixed and $\underset{\sim}{\Sigma}_t$ or $\underset{\sim}{\Sigma}_t^0$ are positive definite symmetric matrices, we have on taking expectation over $\underset{\sim}{X}_\alpha^{(t)}$ and assumptions (i) and (ii)

$$V(\frac{1}{n} L_{(2)}) = O(\frac{1}{n}). \tag{4.6}$$

Since (4.4) is conditional likelihood function, $L_{(3)}$ can be treated as a constant and hence its conditional variance is zero and since $E\,\underset{\sim}{X}_\alpha^{(t)}=\underset{\sim}{0}$, its variance is zero also unconditionally. Again the three covariance terms are zero due to the fact that $\underset{\sim}{Z}_{\alpha}^{(t)}$'s are distributed as $N_p(\underset{\sim}{O}, \underset{\sim}{I})$. Hence the result (4.3) follows from (4.5) and (4.6).

Now from (4.3) and Chebychev's inequality we have

$$\lim_{n\to\infty} \text{Prob}[\frac{1}{n} \log L\ (\underset{\sim}{Y}|\underset{\sim}{\theta}) = E_0(\frac{1}{n} \log L\ (\underset{\sim}{Y}|\underset{\sim}{\theta}))] = 1, \tag{4.7}$$

where E_0 is the expectation when true parameter $\underset{\sim}{\theta}_0$ holds. Also for any $\underset{\sim}{\theta}\neq\underset{\sim}{\theta}_0$, we have from Lemma 1 of Wald [11]

$$E_0[\frac{1}{n} \log L(\underset{\sim}{Y}|\underset{\sim}{\theta})] < E_0[\frac{1}{n} \log L(\underset{\sim}{Y}|\underset{\sim}{\theta}_0)] \tag{4.8}$$

If the maximum likelihood estimate $\hat{\underset{\sim}{\theta}}$ provides global maximum of the likelihood, then with probability one,

$$\log L(\underset{\sim}{Y}|\hat{\underset{\sim}{\theta}}) \geq \log L(\underset{\sim}{Y}|\underset{\sim}{\theta}_0), \tag{4.9}$$

which satisfies the conditions of theorems 2 of Wald to hold. Hence using (4.7) and (4.8) we have the result (4.2) from the theorem 2 of Wald[11]. This establishes the consistency of the estimate $\hat{\underset{\sim}{\theta}}$.

4.2. Asymptotic Efficiency

To establish the asymptotic efficiency of the estimates $\hat{\underset{\sim}{\theta}}=(\hat{\underset{\sim}{\eta}}^{(1)},\ldots,\hat{\underset{\sim}{\eta}}^{(m)},\hat{\underset{\sim}{\Sigma}}_1,\ldots,\hat{\underset{\sim}{\Sigma}}_m,\hat{\underset{\sim}{\beta}})$ we are to prove the following.

<u>Theorem 1</u>. The derivative of the log-likelihood, $\partial \log L/\partial\underset{\sim}{\theta}$ is asymptotically normally distributed with a null-matrix as mean and variance-covariance matrix as the information matrix $\underset{\sim}{\mathcal{J}}(\underset{\sim}{\theta})$

This theorem will then imply that $\hat{\underset{\sim}{\theta}}$ is asymptotically normally distributed with mean $\underset{\sim}{\theta}$ and dispersion matrix $\underset{\sim}{\mathcal{J}}^{-1}$.

<u>Proof</u>. The elements of the information matrix $\underset{\sim}{\mathcal{I}}$ are obtained by considering second derivatives of log-likelihood with respect to parameters. Let us first of all show that the off-diagonal submatrices of the information matrix are zero and diagonal submatrices are the inverse of the dispersion matrices corresponding to the estimates $\hat{\eta}^{(t)}$, $\hat{\Sigma}_t$ and $\hat{\beta}$, t=1,2,...,m.

To show this let us denote (3.2), (3.3) and (3.4) by $U_1^{(t)}$, $U_2^{(t)}$ and U_3 respectively. Then applying matrix derivatives method we have

$$\frac{\partial U_1^{(t)}}{\partial \eta^{(t)'}} = \frac{\partial^2 \log L}{\partial \eta^{(t)'} \partial \eta^{(t)}} = - n_t \Sigma_t^{-1} \tag{4.10}$$

$$\frac{\partial U_1^{(t)'}}{\partial \beta} = - \begin{pmatrix} n_t \Sigma_t^{-1} J_{11}' \bar{X}^{(t)'} & \cdots & n_t \Sigma_t^{-1} J_{1p}' \bar{X}^{(t)'} \\ \vdots & & \\ n_t \Sigma_t^{-1} J_{s1}' \bar{X}^{(t)'} & \cdots & n_t \Sigma_t^{-1} J_{sp}' \bar{X}^{(t)'} \end{pmatrix}, \tag{4.11}$$

where $J_{r\ell}$ is a (s×p) matrix with (r,ℓ)th element unity and rest are zero, r=1,...,s; ℓ=1,...,p and each element of (4.11) is a vector of order (p×1).

$$\frac{\partial U_1^{(t)'}}{\partial \Sigma_t} = - \begin{pmatrix} n_t \Sigma_t^{-1} J_{11}' \Sigma_t^{-1} \bar{\varepsilon}^{(t)'} & \cdots & n_t \Sigma_t^{-1} J_{1p}' \Sigma_t^{-1} \bar{\varepsilon}^{(t)'} \\ \vdots & & \\ n_t \Sigma_t^{-1} J_{p1}' \Sigma_t^{-1} \bar{\varepsilon}^{(t)'} & \cdots & n_t \Sigma_t^{-1} J_{pp}' \Sigma_t^{-1} \bar{\varepsilon}^{(t)'} \end{pmatrix} \tag{4.12}$$

where $J_{r\ell}$ is a (p×p) matrix with (r,ℓ)th and (ℓ,r)th elements unity for $r \neq \ell$ and only (r,ℓ)th element unity for $r=\ell$ for r,ℓ=1,...,p and rest are zero.

$$\frac{\partial \underset{\sim}{U}_2^{(t)}}{\partial \underset{\sim}{\Sigma}_t} = \tfrac{1}{2}n_t\left\{\begin{pmatrix} \Sigma_t^{-1}J_{11}\Sigma_t^{-1} & \cdots & \Sigma_t^{-1}J_{p1}\Sigma_t^{-1} \\ \vdots & & \\ \Sigma_t^{-1}J_{1p}\Sigma_t^{-1} & \cdots & \Sigma_t^{-1}J_{pp}\Sigma_t^{-1} \end{pmatrix} - \begin{pmatrix} \Sigma_t^{-1}J_{11}\Sigma_t^{-1}S_t\Sigma_t^{-1} & \cdots & \Sigma_t^{-1}J_{p1}\Sigma_t^{-1}S_t\Sigma_t^{-1} \\ \vdots & & \\ \Sigma_t^{-1}J_{1p}\Sigma_t^{-1}S_t\Sigma_t^{-1} & \cdots & \Sigma_t^{-1}J_{pp}\Sigma_t^{-1}S_t\Sigma_t^{-1} \end{pmatrix}\right\} \tag{4.13}$$

where $\underset{\sim}{S}_t = \sum_{\alpha=1}^{n_t} \underset{\sim}{\varepsilon}_\alpha^{(t)'} \underset{\sim}{\varepsilon}_\alpha^{(t)}$.

$$\frac{\partial \underset{\sim}{U}_2^{(t)}}{\partial \beta_{r\ell}} = -\tfrac{1}{2}[\Sigma_t^{-1}\sum_\alpha \varepsilon_\alpha^{(t)'} X_\alpha^{(t)} J_{r\ell}\Sigma_t^{-1} + \Sigma_t^{-1}\sum_\alpha J_{r\ell}' X_\alpha^{(t)'} \varepsilon_\alpha^{(t)}\Sigma_t^{-1}] \tag{4.14}$$

for $r=1,\ldots,s$; $\ell=1,\ldots,p$.

$$\frac{\partial \underset{\sim}{U}_3}{\partial \underset{\sim}{\beta}'} = \frac{\partial^2 \log L}{\partial \underset{\sim}{\beta}' \partial \underset{\sim}{\beta}} = -\begin{pmatrix} \sum_\alpha X_\alpha^{(t)'} X_\alpha^{(t)} J_{11}\Sigma_t^{-1} & \cdots & \sum_\alpha X_\alpha^{(t)'} X_\alpha^{(t)} J_{s1}\Sigma_t^{-1} \\ \vdots & & \\ \sum_\alpha X_\alpha^{(t)'} X_\alpha^{(t)} J_{1p}\Sigma_t^{-1} & \cdots & \sum_\alpha X_\alpha^{(t)'} X_\alpha^{(t)} J_{sp}\Sigma_t^{-1} \end{pmatrix} \tag{4.15}$$

Now from the assumptions on the model (2.1) and assumption (ii) of Section 4, it follows that on taking expectations, expressions (4.11), (4.12) and (4.14) are zero, which proves that off-diagonal submatrices of the information matrix are zero. Again on taking expectations over (4.10), (4.13) and (4.15) and after some simplifications, the inverse of the dispersion matrices of the estimates $\hat{\underset{\sim}{\eta}}^{(t)}$, $\hat{\underset{\sim}{\Sigma}}_t$ and $\underset{\sim}{\beta}$, which are the diagonal submatrices of information matrix are obtained as follows.

$$\underset{\sim}{V}_{\hat{\eta}^{(t)}}^{-1} = n_t\underset{\sim}{\Sigma}_t^{-1}, \quad t=1,2,\ldots,m, \tag{4.16}$$

$$\underset{\sim}{V}_{\hat{\Sigma}_t}^{-1} = \tfrac{1}{2}\, n_t(\underset{\sim}{\Sigma}_t^{-1} \boxdot \underset{\sim}{\Sigma}_t^{-1})\ , \ t=1,2,\cdots m, \tag{4.17}$$

$$\underset{\sim}{V}_{\hat{\beta}}^{-1} = \sum_{t=1}^{m} n_t\underset{\sim}{\Sigma}_t^{-1} \otimes \underset{\sim}{\Sigma}_t^{*} \tag{4.18}$$

where $A \boxdot B$ is the product notation for $((a_{ij}b_{ij}))$ when A and B are of some order (Rao [10]) and $P \otimes Q$ is the Kronecker's product notation.

Now to prove the theorem we are only to prove that, unconditionally, any linear function of $(\underset{\sim}{U}_1^{(t)}, \underset{\sim}{U}_2^{(t)}$ and $\underset{\sim}{U}_3$, $t=1,2,\ldots,m)$ is asymptotically normally distributed. Let us, therefore, consider the linear function

$$T = \sum_{t=1}^{m} \underset{\sim}{U}_1^{(t)} \underset{\sim}{D}_1^{(t)'} + \mathrm{tr} \sum_{t=1}^{m} \underset{\sim}{U}_2^{(t)} \underset{\sim}{D}_2^{(t)'} + \mathrm{tr}\, \underset{\sim}{U}_3 \underset{\sim}{D}_3' \tag{4.19}$$

where $\underset{\sim}{D}_1^{(t)}(1\times p)$, $\underset{\sim}{D}_2^{(t)}(p\times p)$, $\underset{\sim}{D}_3(s\times p)$ are matrices of real elements. Then from (3.2), (3.3) and (3.4) we can rewrite T as follows.

$$T = \sum_{t=1}^{m} \Big[\sum_{\alpha=1}^{n_t} \{ \underset{\sim}{\varepsilon}_\alpha^{(t)} \underset{\sim}{F}_1^{(t)} + \mathrm{tr}(-\underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_2^{(t)'} + \underset{\sim}{\varepsilon}_\alpha^{(t)'} \underset{\sim}{\varepsilon}_\alpha^{(t)} \underset{\sim}{F}_2^{(t)}) + \underset{\sim}{\varepsilon}_\alpha^{(t)} \underset{\sim}{F}_{3\alpha}^{(t)} \} \Big] \tag{4.20}$$

where $\underset{\sim}{F}_1^{(t)} = \underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_1^{(t)'}$, $\underset{\sim}{F}_2^{(t)} = \underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_2^{(t)'} \underset{\sim}{\Sigma}_t^{-1}$ and $\underset{\sim}{F}_{3\alpha}^{(t)} = \underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_3' \underset{\sim}{X}_\alpha^{(t)'}$.

Writing (4.20) as $T = \sum_t \sum_\alpha T_\alpha^{(t)}$, where $T_\alpha^{(t)} = T_{1\alpha}^{(t)} + T_{2\alpha}^{(t)} + T_{3\alpha}^{(t)}$, we note that the undonditional second and fourth central moments of $T_\alpha^{(t)}$, are

$$\mu_2(T_\alpha^{(t)}) = E(T_{1\alpha}^{(t)^2} + T_{2\alpha}^{(t)^2} + T_{3\alpha}^{(t)^2}) \tag{4.21}$$

$$\begin{aligned} \mu_4(T_\alpha^{(t)}) = E(T_{1\alpha}^{(t)^4} + T_{2\alpha}^{(t)^4} + T_{3\alpha}^{(t)^4}) + 6E(T_{1\alpha}^{(t)^2} T_{2\alpha}^{(t)^2} \\ + T_{1\alpha}^{(t)^2} T_{3\alpha}^{(t)^2} + T_{2\alpha}^{(t)^2} T_{3\alpha}^{(t)^2}) \end{aligned} \tag{4.22}$$

(which follow from the assumptions in (2.1) and assumption (ii) of Section 4). Now $\varepsilon_\alpha^{(t)}$ being distributed as $N_p(\underset{\sim}{0}, \underset{\sim}{\Sigma}_t)$, $T_{1\alpha}^{(t)} = \underset{\sim}{\varepsilon}_\alpha^{(t)} \underset{\sim}{F}_1^{(t)}$, a linear function of the elements of $\underset{\sim}{\varepsilon}_\alpha^{(t)}$, is distributed as $N(0, \underset{\sim}{F}_1^{(t)'} \underset{\sim}{\Sigma}_t \underset{\sim}{F}_1^{(t)})$. Hence

$$\left.\begin{aligned} E(T_{1\alpha}^{(t)^2}) &= \underset{\sim}{F}_1^{(t)'}\underset{\sim}{\Sigma}_t\underset{\sim}{F}_1^{(t)} = \underset{\sim}{D}_1^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{D}_1^{(t)'} \\ E(T_{1\alpha}^{(t)^4}) &= 3(\underset{\sim}{D}_1^{(t)}\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{D}_1^{(t)'})^2 \end{aligned}\right\} \tag{4.23}$$

From (4.20), $T_{2\alpha}^{(t)} = tr(\underset{\sim}{\varepsilon}_\alpha^{(t)'}\underset{\sim}{\varepsilon}_\alpha^{(t)}\underset{\sim}{F}_2^{(t)} - \underset{\sim}{\Sigma}_t\underset{\sim}{F}_2^{(t)})$. Since $\underset{\sim}{\varepsilon}_\alpha^{(t)'}\underset{\sim}{\varepsilon}_\alpha^{(t)} \sim W_p(1, \underset{\sim}{\Sigma}_t)$, the characteristic function of $T_{2\alpha}^{(t)}$ is obtained as (Anderson [1])

$$\begin{aligned} \varphi(\theta) = Ee^{i\theta T_{2\alpha}^{(t)}} &= e^{-i\theta tr(\underset{\sim}{\Sigma}_t\underset{\sim}{F}_2^{(t)})} \cdot Ee^{i\theta tr\underset{\sim}{\varepsilon}_\alpha^{(t)'}\underset{\sim}{\varepsilon}_\alpha^{(t)}\underset{\sim}{F}_2^{(t)}} \\ &= e^{-i\theta tr(\underset{\sim}{\Sigma}_t\underset{\sim}{F}_2^{(t)})} \cdot |I-2i\theta\underset{\sim}{\Sigma}_t\underset{\sim}{F}_2^{(t)}|^{-\frac{1}{2}} \end{aligned} \tag{4.24}$$

On taking logarithm of both sides of (4.24) and expanding r.h.s., we have by collecting coefficients of $(i\theta)^2/2!$ and $(i\theta)^4/4!$,

$$\left.\begin{aligned} \mu_2(T_{2\alpha}^{(t)}) &= 2\ tr(\underset{\sim}{\Sigma}_t\underset{\sim}{F}_2^{(t)})^2 = 2\ tr(\underset{\sim}{D}_2^{(t)'}\underset{\sim}{\Sigma}_t^{-1})^2 \\ \mu_4(T_{2\alpha}^{(t)}) &= 12[4\ tr(\underset{\sim}{D}_2^{(t)'}\underset{\sim}{\Sigma}_t^{-1})^4 + (tr(\underset{\sim}{D}_2^{(t)'}\underset{\sim}{\Sigma}_t^{-1})^2)^2] \end{aligned}\right\} \tag{4.25}$$

From (4.20) it is clear that for fixed $\underset{\sim}{X}_\alpha^{(t)}$, $T_{3\alpha}^{(t)}$ is a linear function of the elements of $\underset{\sim}{\varepsilon}_\alpha^{(t)}$ and hence distributed as $N(0, \underset{\sim}{X}_\alpha^{(t)}\underset{\sim}{R}_t\underset{\sim}{X}_\alpha^{(t)'})$, where $\underset{\sim}{R}_t = \underset{\sim}{D}_3\underset{\sim}{\Sigma}_t^{-1}\underset{\sim}{D}_3'$. Therefore, the c.f. of $T_{3\alpha}^{(t)}$ for fixed $\underset{\sim}{X}_\alpha^{(t)}$ is

$$\varphi_x(\theta) = e^{\frac{1}{2}(i\theta)^2\underset{\sim}{X}_\alpha^{(t)}\underset{\sim}{R}_t\underset{\sim}{X}_\alpha^{(t)'}} \tag{4.26}$$

Then from assumption (ii), since $\underset{\sim}{X}_\alpha^{(t)} \sim N(\underset{\sim}{0}, \underset{\sim}{\Sigma}_t^*)$ we have the unconditional c.f. of $T_{3\alpha}^{(t)}$ by integrating (4.26) over $\underset{\sim}{X}_\alpha^{(t)}$,

$$\begin{aligned} \varphi(\theta) &= |\underset{\sim}{\Sigma}_t^{*-1}-(i\theta)^2\underset{\sim}{R}_t|^{-\frac{1}{2}}/|\underset{\sim}{\Sigma}_t^*|^{\frac{1}{2}} \\ &= |I_s - (i\theta)^2\underset{\sim}{\Sigma}_t^*\underset{\sim}{R}_t|^{-\frac{1}{2}} \end{aligned} \tag{4.27}$$

Hence on taking logarithm of both sides of (4.27) and expanding r.h.s. we have

$$\left.\begin{aligned} \mu_2(T_{3\alpha}^{(t)}) &= \operatorname{tr}(\underset{\sim}{\Sigma}_t^* \underset{\sim}{D}_3 \underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_3') \\ \mu_4(T_{3\alpha}^{(t)}) &= 3[4\, \operatorname{tr}(\underset{\sim}{\Sigma}_t^* \underset{\sim}{D}_3 \underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_3')^2 + (\operatorname{tr}(\underset{\sim}{\Sigma}_t^* \underset{\sim}{D}_3 \underset{\sim}{\Sigma}_t^{-1} \underset{\sim}{D}_3'))^2] \end{aligned}\right\} \quad (4.28)$$

Thus substituting results from (4.23), (4.25) and (4.28) in (4.21) and (4.22) and remembering that $T_{1\alpha}^{(t)}$, $T_{2\alpha}^{(t)}$ and $T_{3\alpha}^{(t)}$ are independently distributed we obtain the unconditional moments $\mu_2(T_\alpha^{(t)})$ and $\mu_4(T_\alpha^{(t)})$ which are finite and same for all α. Let $\nu_3(T_\alpha^{(t)})$ be the third absolute central moment of $T_\alpha^{(t)}$. Then defining

$$B_{n_t} = \Big(\sum_{\alpha=1}^{n_t} \nu_3(T_\alpha^{(t)})\Big)^{1/3}, \quad C_{n_t} = \Big(\sum_{\alpha=1}^{n_t} \mu_2(T_\alpha^{(t)})\Big)^{1/2}$$

we have

$$\frac{B_{n_t}}{C_{n_t}} = \frac{(\sum_\alpha \nu_3(T_\alpha^{(t)}))^{1/3}}{(\sum_\alpha \mu_2(T_\alpha^{(t)}))^{1/2}} < \frac{(\sum_\alpha \nu_4^{3/4}(T_\alpha^{(t)}))^{1/3}}{(\sum_\alpha \mu_2(T_\alpha^{(t)}))^{1/2}} \quad (4.29)$$

Since $\nu_4(T_\alpha^{(t)}) = \mu_4(T_\alpha^{(t)})$ is finite and constant for all α, it follows that $B_{n_t}/C_{n_t} \to 0$ as $n_t \to \infty$.

Thus for the sequence of independent random variables $\{T_\alpha^{(t)}\}$, all the conditions for Liapounoff's central limit theorem [4] are satisfied and hence $T^{(t)} = \sum_{\alpha=1}^{n_t} T_\alpha^{(t)}$ is asymptotically normally distributed. Since $T^{(t)}$ for $t=1,\ldots,m$ are independently distributed it follows that T defined by (4.19) is asymptotically normally distributed. This T being linear function of $(\underset{\sim}{U}_1^{(t)}, \underset{\sim}{U}_2^{(t)}, \underset{\sim}{U}_3,\ t=1,\ldots,m)$ it follows that $\partial \operatorname{Log} L/\partial\underset{\sim}{\theta}$ asymptotically follows multinormal law with a null-matrix as mean and dispersion matrix $\underset{\sim}{\mathcal{J}}$, whose diagonal elements are given by

(4.16), (4.17) and (4.18). Hence the theorem.

4.3. Unbiasedness

The small sample property of unbiasedness of the estimates follow in the same line as proved by Hartley and Jaytillake [6] since from the assumptions in model (2.1) the condition $P(\underset{\sim}{\varepsilon}_{\alpha}^{(t)}) = P(-\underset{\sim}{\varepsilon}_{\alpha}^{(t)})$ is satisfied.

5. LIKELIHOOD RATIO TEST FOR THE HYPOTHESIS OF EQUALITY OF SEVERAL GROWTH CURVES

From the model (2.1) it is clear that the desired hypothesis is

$$H_o[\underset{\sim}{\eta}^{(1)} = \ldots = \underset{\sim}{\eta}^{(m)}] \tag{5.1}$$

We have seen that the parameter matrix of the model (2.1) is $\underset{\sim}{\theta} = [\underset{\sim}{\eta}^{(1)},\ldots,\underset{\sim}{\eta}^{(m)},\underset{\sim}{\Sigma}_1,\ldots,\underset{\sim}{\Sigma}_m,\underset{\sim}{\beta}]$. Let the unconditional maximum of the likelihood function, obtained in Section 3, be denoted by $L(\underset{\sim}{Y}|\hat{\underset{\sim}{\theta}})$. Now under H_o, the parameter matrix contains m+2 parameters $\underset{\sim}{\theta}^o = (\underset{\sim}{\eta},\underset{\sim}{\Sigma}_1,\ldots,\underset{\sim}{\Sigma}_m,\underset{\sim}{\beta})$ and the model (2.1) reduces to the MANOCOVA model of the same kind. So that by the same procedure as in Section 3 we obtain the maximum of the likelihood function, given by $L(\underset{\sim}{Y}|\hat{\underset{\sim}{\theta}}^o)$. Hence the likelihood ratio test is given by

$$\lambda = L(\underset{\sim}{Y}|\hat{\underset{\sim}{\theta}}^o)/L(\underset{\sim}{Y}|\hat{\underset{\sim}{\theta}}) \tag{5.2}$$

Since all our maximum likelihood estimates are asymptotically normally distributed and efficient, $-2 \log_e \lambda$ is asymptotically distributed, under H_o, as a central chi-square with p(m-1) d.f.

ACKNOWLEDGMENT

The author is grateful to Professor P. K. Sen for his help and guidance throughout this investigation.

REFERENCES

[1] Anderson, T. W., Introduction to Multivariate Statistical Analysis, (1958) John Wiley & Sons Inc., New York.

[2] Dwyer, P. S. and Macphail, M. S., Symbolic matrix derivatives, Annals of Mathematical Statistics, 19, (1948), 517-34.

[3] Dwyer, P. S., Some applications of matrix derivatives in multivariate analysis, Journal of American Statistical Association, 62, (June 1967), 607-25.

[4] Gnedenko, B. W. and Kolmogorov, A. N., Limit Distributions for Sums of Independent Random Variables, (1954), Addison-Wesley, Cambridge.

[5] Hartley, H. O. and Rao, J. N. K., Maximum-likelihood estimation for mixed analysis of variance model, Biometrika, 54, (June 1967), 93-108.

[6] Hartley, H. O. and Jaytillake, K. S. E., Estimation of linear models with unequal variances, Journal of American Statistical Association, 68, (March 1973), 189-92.

[7] Henrici, P., Discrete Variable Methods in Ordinary Differential Equations, (1962), John Wiley & Sons Inc., New York.

[8] Potthoff, R. F. and Roy, S. N., A generalized multivariate analysis of variance model useful specially for growth curve problems, Biometrika, 51, (June 1964), 313-26.

[9] Rao, C. R., The theory of least squares when parameters are stochastic and its applications to the analysis of growth curve, Biometrika, 52, (December 1965), 447-58.

[10] Rao, C. R., Estimation of heteroscedastic variances in linear models, Journal of American Statistical Association, 65 (March 1970), 161-72.

[11] Wald, A., Note on the consistency of the maximum likelihood estimate, Annals of Mathematical Statistics, 20 (1949), 595-601.