Revista Colombiana de Estadística
N$^{\underline{o}}$s 19-20, junio-diciembre 1989

CALCULO DE LA MATRIZ INVERSA DE COMPONENTES DE VARIANZA EN ESTRUCTURAS BALANCEADAS Y DESBALANCEADAS

Luis Alberto López Pérez

Profesor Asistente
Universidad Nacional

Resumen. De la matriz V de componentes de varianza asociada a un modelo superparametrizado con n-vías de clasificación, se desarrolla un procedimiento para obtener su inversa cuando la estructura de datos es balanceada y jerárquica desbalanceada, haciendo uso de productos de Kronecker de matrices idénticas y cuadradas de orden uno.

Introducción.

Los modelos de efectos mixtos, los cuales dentro del contexto de modelos lineales caracterizan las estructuras de componentes de varianza, tienen gran relevancia por sus múltiples aplicaciones especialmente en la determinación de parámetros genéticos e índices de selección de especies animales y/o vegetales.

En este tipo de modelos es muy importante la determinación de los mejores estimadores lineales no viciados o insesgados (MELI) y mejores predictores lineales no viciados (BLUP), ya que el conocimiento de éstos nos permiten hacer inferencia y de esta forma tomar decisiones apropiadas.

Sin embargo como puede verse en SEARLE (1987) tanto la estructura de los MELI como de los BLUP dependen del conocimiento de la matriz de componentes de varianza y consecuentemente de su inversa la cual tiene una estructura compleja dependiente de la naturaleza del modelo.

Por ello algunos trabajos han contribuído a buscar soluciones generales al problema de la inversión de la matriz de componentes de varianza, particularmente en SEARLE y HENDERSON (1979) se encuentra una solución general para el caso de estructuras balenceados.

En este trabajo se presenta una metodología a partir de la cual se obtiene en forma rápida la inversa de la matriz de componentes en estructura balanceadas y no balanceadas dentro de un orden jerárquico.

2. Caracterización general de los modelos mixtos.

En esta sección se van a identificar los modelos mixtos y sus componentes de varianza considerando los supuestos necesarios que permita identificar la matriz asociada a esta estructura de modelos mixtos.

2.1. El modelo: siguiendo la notación de SEARLE (1988) se tiene que:

$$y = X\beta + Zu + e..... \qquad (2.1)$$

En donde:

$y_{N\times 1}$: Vector de observaciones

$\beta_{k\times 1}$: Vector de parámetros asociados a los efectos fijos

$X_{N\times k}$: Matriz conocida asociada a los efectos fijos

$u'_{1\times q}$: Vector no observable asociado a los efectos aleatorios

$$q = \sum_{i=1}^{s} q_i \quad ; \qquad u = (u_1';u_2';\ldots;u_s').$$

$Z_{N\times q}$: Usualmente es una matriz de incidente asociada a los efectos aleatorios y en general observable.

$$Z = (z_1;z_2;\ldots;z_s)$$

$e_{N\times 1}$: Vector no observable o vector de términos de error.

2.2. Supuestos generales.

Partiendo del modelo (2.1) y bajo el supuesto que $u_o = e$, entonces la matriz U es de la forma

$$U = (u_o', u_1', \ldots, u_s') \qquad (2.2)$$

i) $E\{U_i\} = 0 \quad \forall_i = 0,1,\ldots,s$.

ii) $\operatorname{cov}(u_i, u_i') = \begin{cases} \sigma_i^2\, I\, q_i \;, & i = i' \\ 0 & i \neq i' \end{cases}$

iii) $$\begin{aligned} \text{var}(Y) &= \text{var}(X\beta + Zu + e) \\ &= Z\text{var}(U)Z' + V(e) + Z\text{cov}(u,e) \qquad (2.3) \\ &= ZDZ' + R = ZDZ' + \sigma^2 e\, I_N \,. \end{aligned}$$

En donde

$$D = \bigoplus_{i=1}^{s} \sigma_i^2 I_{q_i} \qquad (2.4)$$

$$R = V(e).$$

En este caso $\oplus$: representa la suma directa de matrices.

Al sustituir (2.4) en (2.3), se sigue que:

$$\text{var}(Y) = Z(\bigoplus_{i=1}^{s} \sigma_i^2 I q_i)Z' + \sigma^2 e\, I_N = \sum_{i=0}^{s} Z_i Z_i' \sigma^2 i = V \qquad (2.5)$$

iv) $$\begin{aligned} \text{cov}(Y;U') &= E\{(y-E(y)\}\{u-E(u)^T\} \\ &= \sum_{i=1}^{s} Z_i \sigma_i^2 = ZD = C. \end{aligned} \qquad (2.6)$$

Los resultados obtenidos en (2.3) a (2.6) son resumidos en la siguiente estructura general

$$\text{var}\begin{bmatrix} Y \\ U \\ e \end{bmatrix} = \begin{bmatrix} V & C & R \\ C & D & 0 \\ R & 0 & R \end{bmatrix} \,. \qquad (2.7)$$

3. Cálculo de la matriz inversa de V con estructuras de datos balanceada.

En esta sección se va a presentar una metodología con la cual se puede implementar en forma inmediata la matriz inversa de la estructura presentada en (2.5), aunque esta estructura de pende de las siguientes características generales:

i) Si el modelo es balanceado.

ii) Si el modelo es desbalanceado.

iii) Si R es de naturaleza compleja.

Los resultados obtenidos en este trabajo satisfacen las características (i) y (ii) pero teniendo en cuenta estructuras jerárquicas y en caso de (iii) se restringe al hecho que

$$R = \sigma^2 I_N$$

De la estructura de V presentada en (2.5), notemos que cada una de las matrices $Z_i Z_i'$ puede ser escrita en términos de productos Kronecker, $[\otimes]$ de matrices I_s y J_s los cuales satisfacen:

a) I_s: Matriz identidad de orden s.

b) J_s: Matriz cuadrada de orden s con elementos todos iguales a uno.

c) $I_{s_1} \otimes I_{s_2} \otimes I_{s_3} = I_{s_1 s_2 s_3}$

d) $J_{s_1} \otimes J_{s_2} \otimes J_{s_3} = J_{s_1 s_2 s_3}$

c) $(J_s)^o = I_s$

Si en el modelo (2.1) describimos una respuesta de la forma:

$$y_{ijk\ldots t}\,;\ i = 1,\ldots,I,\quad j = 1,\ldots,J,\ldots t = 1,\ldots,T$$

Sea p el número de subíndices asociados a la variable respuesta, los cuales dependen de los efectos e interacciones fijos y aleatorios del modelo; se va a tener entonces 2^p particiones de productos de Kronecker con matrices I_δ, J_δ, las cuales son resumidas en la expresión

$$A = (I_I \otimes I_J \otimes .. \otimes I_T; I_I \otimes I_J \otimes .. \otimes J_T; .. ; J_I \otimes J_J \otimes J_T) \qquad (3.2)$$

Definiendo ahora la función indicadora

$$\delta = \begin{cases} 1 \text{ si aparece } J_\delta \\ 0 \text{ si aparece } I_\delta . \end{cases}$$

Se nota entonces que (3.2) puede ser reescrita de la forma

$$\Delta^o = (00...0;\ 00...1;\ 11...1) \qquad (3.3)$$

Caracterizado el vector Δ^o, entonces se construye el vector asociado con los componentes de la varianza, digamos $\theta_{1\times 2^p}$ de componentes $((\theta_\delta))$.

Según el siguiente criterio

$$\theta_\delta = \begin{cases} \sigma^2 & \text{si existe una relación directa entre el subíndice del efecto y la posición del cero en (3.3).} \\ 0 & \text{en caso contrario.} \end{cases}$$

Con estos resultados la matriz (2.5) es reescrita de la forma:

$$V = \Delta^* \, \theta'^* \qquad (3.4)$$

En donde Δ^* representa la matriz final de permutaciones de tal manera que exista una relación directa entre los subíndices del último vector columna en la matriz T_p y la estructura de Δ^* siendo T_p el producto Kronecher de submatrices.

$$T_p = \begin{bmatrix} 1 & 0 \\ 1 & I \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 1 & J \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} 1 & 0 \\ 1 & T \end{bmatrix} \tag{3.5}$$

En forma similar θ^* es una permutación final de θ .

Al hacer

$$\lambda p = T_p\, \theta^* \tag{3.6}$$

Se sigue de (3.4), (3.5) y (3.6) que

$$V^{-1} = T_p^{-1}\ \lambda_p^{-1} \Delta^* \ . \tag{3.7}$$

Ejemplo: sea el modelo

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \tag{3.8}$$

Notamos entonces que $p = 3$, y considerando solo el efecto de la media como no aleatorio de (3.2) se sigue que:

$$\Delta = (I_I \otimes I_J \otimes I_K;\ I_I \otimes I_J \otimes I_K; \ldots ; J_I \otimes J_J \otimes J_K) \tag{3.9}$$

Entonces

$$\Delta^* = [J_I \otimes J_J \otimes J_K;\ J_I \otimes J_J \otimes I_K; \ldots ; I_I \otimes I_J \otimes I_K]$$

$$[J_{IJK};\ J_{IJ} \otimes I_K; \ldots ; I_{IJK}]$$

$$\Delta' = [\overset{\varepsilon_{ijk}}{000},\ \overset{\gamma_{ij}}{001}, 010,\ 100,\ 110,\ \overset{\beta_j}{101},\ \overset{\alpha_i}{011}, 111]$$

Así:

$$\theta = [\sigma_\varepsilon^2 \; ; \; \sigma_\gamma^2 \; ,0,00,\sigma_\varepsilon^2 \; ; \; \sigma_\alpha^2 \; ;0]$$

y

$$\theta^* = [0,0,\sigma_\beta^2 \; ; \; 0;\sigma_\alpha^2 \; ,0,\sigma_\gamma^2 \; ; \; \sigma_\varepsilon^2 \;]$$

El vector de componentes de varianza asociado al modelo (3.8) es θ^*; y, entonces la matriz (2.5) es de la forma

$$V = \Delta'\theta^* = I_{IJK}\,\sigma_\varepsilon^2 + I_{IJ}\otimes J_K\,\sigma_\gamma^2 + I_I\otimes J_{JK}\sigma_\alpha^2 + J_I\otimes I_J\otimes J_K\sigma_\beta^2$$

Y de (3.6) se sigue que:

$$\lambda' = [\sigma_\varepsilon^2;\; \sigma_\varepsilon^2 + K\sigma_\gamma^2;\; \sigma_\varepsilon^2;\; \sigma_\varepsilon^2 + K\sigma_\gamma^2 + JK\sigma_\alpha^2;\; \sigma_\varepsilon^2\,;\sigma_\varepsilon^2 + K\sigma_\gamma^2 + IK\sigma_\beta^2\;;\sigma_\varepsilon^2\;;\sigma_\varepsilon^2 + K\sigma_\gamma^2 + JK\sigma_\alpha^2 + IK\sigma_\beta^2\;]\;.$$

Obteniendo finalmente de (3.7)

$$V^{-1} = \frac{1}{\sigma_\varepsilon^2} I_{IJK} - \frac{\sigma_\gamma^2}{\sigma_\varepsilon^2\,(\sigma_\varepsilon^2 + K\sigma_\gamma^2\,)} I_{IJ}\otimes J_K$$

$$- \frac{\sigma_\alpha^2}{(\sigma_\varepsilon^2 + K\sigma_\gamma^2\,)\,(\sigma_\varepsilon^2 + K\sigma_\gamma^2 + JK\sigma_\alpha^2\,)} I_I\otimes J_{jK}$$

$$- \frac{\sigma_\beta^2}{(\sigma_\varepsilon^2 + K\sigma_\gamma^2\,)(\sigma_\varepsilon^2 + K\sigma_\gamma^2 + IK\sigma_\beta^2\,)} J_I\otimes I_J\otimes J_K$$

$$+ \frac{K\sigma_\alpha^2\,\sigma_\beta^2\,[(\sigma_\varepsilon^2+K\sigma_\gamma^2) + (\sigma_\varepsilon^2+K\sigma_\gamma^2) + JK\sigma_\alpha^2 + IK\sigma_\beta^2)]}{(\sigma_\varepsilon^2+K\sigma_\alpha^2+IK\sigma_\beta^2)\,(\sigma_\varepsilon^2+K\sigma_\gamma^2)\,(\sigma_\varepsilon^2+K\sigma_\alpha^2+JK\sigma_\alpha^2)\,(\sigma_\varepsilon^2+K\sigma_\gamma^2+JK\sigma_\alpha^2+IK\sigma_\beta^2)} J_{IJK}.$$

Obsérvese en el resultado anterior que:

Si $K = 1$ y $\gamma = 0$, se identifica en forma inmediata el modelo de dos vías sin interacción donde sólo el efecto de la media es considerado como fijo.

4. Cálculo de la matriz inversa de componentes de varianza con estructuras jerárquicas desbalanceadas.

En esta sección, se va a presentar un algoritmo el cual permite obtener la estructura de la matriz inversa cuando la información se ajusta a un modelo jerárquico desbalanceado, por facilidad de comprensión se hace uso de dos modelos simples y encima de ellos se presenta el resultado propuesto:

Sea el modelo

$$Y_{ij} = \mu + \alpha i + \varepsilon_{IJK} \qquad i = 1\ldots a; \quad j = 1\ldots ni$$

de tal manera que

$$N = \sum_{i=1}^{a} ni$$

de (2.5) y (3.4) se sigue que

$$V = \sigma_\varepsilon^2\, IN + \sigma_\alpha^2\, I_a \otimes J_{ni} = D(\sigma_\varepsilon^2\, I_{ni} + \sigma_\alpha^2\, J_{ni}) \tag{4.1}$$

De donde $D(t)$: Matriz diagonal con elementos en la diagonal principal iguales a t.

Se puede observar también que la forma presentada en (4.1) puede ser reescrita en la forma de sumas directas como:

$$V = \bigoplus_{i=1}^{a} [\sigma_\varepsilon^2\, I_{ni} + \sigma_\alpha^2\; 1_{ni\times 1} I(1)\; 1_{1\times ni}] \tag{4.2}$$

donde $I(1) = 1$.

$$= \bigoplus_{i=1}^{a} V_i \; . \tag{4.3}$$

De Herrnderson y Searle (1981) y Searle (1982) se obtiene que:

$$(A + U B \mathring{V})^{-1} = A^{-1} - A^{-1} U (I + B \mathring{V} A^{-1} U)^{-1} B \mathring{V} A^{-1} \tag{4.4}$$

Que al ser sustituído en (4.2) obtenemos la estructura de la matriz inversa como

$$V^{-1} = \bigoplus_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2 (\sigma_\varepsilon^2 + n i \sigma_\alpha^2)^2} [\sigma_\varepsilon^2 I_{ni} + D(n_i \sigma_\alpha^2) - \sigma_\alpha^2 J_{ni}] \tag{4.5}$$

Si ahora tenemos el modelo

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk} \qquad i = 1 \ldots a; \quad j = 1 \ldots, b$$

$$k = 1, \ldots, n_{ij}$$

De (2.5) y (3.4) se tiene que

$$V = \sigma_\varepsilon^2 I_{ni} + \sigma_\alpha^2 I_a \otimes J_{ni} + \sigma_\beta^2 D(J_{nij}) \tag{4.6}$$

Al escribir

$$J_{nij} = 1_{nij \times 1} I(1) 1_{1 \times nij}$$

y

$$V = \bigoplus_{i=1}^{a} V_i^* \; .$$

Tenemos entonces que

$$V = \bigoplus_{i=1}^{a} [\sigma_\varepsilon^2 I_{ni} + \sigma_\alpha^2 I_a \otimes J_{ni} + \sigma_\beta^2 (1_{nij \times 1} I(1) 1_{1 \times nij}] \tag{4.7}$$

$$= \bigoplus_{i=1}^{a} [V_i + \sigma_\beta^2 \, (1_{nij\times 1} I(1) 1_{1\times nij}] \qquad (4.8)$$

donde

$$V_i = \sigma_\varepsilon^2 \; 1_{ni} + \sigma_\alpha^2 \; 1_a \otimes J_{ni} \, .$$

Al aplicar nuevamente el resultado obtenido en (4.4) en (4.8) se obtiene la inversa de la V asociada al modelo jerárquico de dos factores, sin pérdida de generalidad podemos afirmar que este procedimiento es aplicado en el caso de estructuras multijerárquicas entre tanto cuando el modelo es cruzado y desbalanceado no es posible presentar un algoritmo de inversión de la expresión (2.5) o no, se encontró en la literatura un procedimiento general para invertir V en este tipo de información.

Una aplicación de este procedimiento de inversión se desarrolla en el trabajo de Novoa/López (1991).

Conclusión.

Dada la relación entre los subíndices del modelo superparametrizado y las matrices I_s y J_s; facilitó la implementación del procedimiento de inversión de la matriz V, haciendo el método más operacional y simple frente a otros procedimientos como el propuesto por Henderson y Searle (1979) el cual está basado en los autovalores de V.

*

BIBLIOGRAFIA

Henderson, H.V. y Searle. S.R. (1981), *On deriving the inverse of a some of matrizes*, SIAM Revew 53-60.

Novoa, E., López, L.A. (1991). *Aspectos computacionales de modelos mixtos desbalanceados*. Tesis de grado. Depto. de Matemáticas y Estadística. Universidad Nacional. Santa Fe de Bogotá.

Searle y Henderson, H.V. (1979). *Dispersion Matrizes for variance comnonents models*, JASA 34, 465-470.

Searle, S.R. (1982), *Matrix Algebra useful for statisties*, John Wiley, N.Y.

Searle, S.R. (1988), *Mixed models and un balanced data where from. where at and where to*: Commun Statist theory Meth. 17 (4), 935-968.

Searle, S.R. (1987), *Linear Models for um balanced data*, John Wiley N.Y.

*