Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
C. R. Acad. Sci. Paris, Ser. I 337 (2003) 219–222
Comptes Rendus Mathematique

Statistics/Probability Theory

Constrained covariance matrix estimation in road accident modelling with Schur complements

Assi N'Guessan

École polytechnique universitaire de Lille et laboratoire de mathématiques appliquées: CNRS, FRE 2222, Université de Lille 1, 59655 Villeneuve d'Ascq cedex, France

Received 31 March 2003; accepted after revision 17 June 2003

Presented by Paul Deheuvels

Abstract

We consider a collection $(Y_{11}, Y_{21}), \dots, (Y_{1s}, Y_{2s})$ of s independent couples of $2r$-dimensional random vector with $r > 1$. We assume that each couple has a multinomial distribution linked to an unknown vector parameter and an extra set data. We deal with a formal inversion of a Fisher's information matrix connected to those couples using Schur complements approach. ***To cite this article: A. N'Guessan, C. R. Acad. Sci. Paris, Ser. I 337 (2003).***

Résumé

Estimation sous contraintes d'une matrice de covariances dans la modélisation d'accidents de la route par complements de Schur. On considère $(Y_{11}, Y_{21}), \dots, (Y_{1s}, Y_{2s})$ une collection de s couples indépendants de vecteurs aléatoires de dimension $2r$ avec $(r > 1)$. On suppose que chaque couple est distribué selon une loi multinomiale dépendant à la fois d'un vecteur paramètre inconnu et d'un ensemble de données supplémentaire. On étudie l'inversion formelle de l'information de Fisher issue de ces s couples par le biais de la technique des compléments de Schur. ***Pour citer cet article : A. N'Guessan, C. R. Acad. Sci. Paris, Ser. I 337 (2003).***

1. Introduction and motivation

N'Guessan, Essai and Langrand [6] deal with the multidimensional estimation of the effect of road safety measure applied to s $(s > 0)$ entities (sites), each one including r $(r > 1)$ mutually exclusive different accident types and linked to a specific control area. They consider a collection $(Y_{11}, Y_{21}), (Y_{12}, Y_{22}), \dots, (Y_{1s}, Y_{2s})$ of s independent couples of $2r$-dimensional random vector and assume that, for fixed k, each couple (Y_{1k}, Y_{2k}) has multinomial distribution denoted by $\mathcal{M}(n_k; \Pi_{1k}(\alpha), \Pi_{2k}(\alpha))$, where

$$\Pi_{tk}(\alpha) = \left(\pi_{t1k}(\alpha), \pi_{t2k}(\alpha), \dots, \pi_{trk}(\alpha)\right),$$

E-mail address: Assi.Nguessan@polytech-lille.fr (A. N'Guessan).

1631-073X/$ – see front matter
doi:10.1016/S1631-073X(03)00309-1

n_k ($n_k > 0$) is a given integer, $\alpha = (\beta_0, \beta^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{1+sr}$ a unknown vector parameter, β_0 ($\beta_0 > 0$) a scalar, $\beta = (\beta_1^{\mathrm{T}}, \ldots, \beta_s^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{sr}$ with $\beta_k = (\beta_{1k}, \beta_{2k}, \ldots, \beta_{rk})^{\mathrm{T}} \in \mathbb{R}^r$ and $\beta_{jk} > 0$. The components $\pi_{trk}(\alpha)$ of the $r \times 1$ vector probability $\Pi_{tk}(\alpha)$ are given by a continuous differentiable functional link $g : \mathbb{R}^{1+sr} \mapsto]0,1[$ so that $\pi_{trk}(\alpha) = g(z_k, \alpha)$ and

$$g(z_k,\alpha) = \begin{cases} \dfrac{\beta_{jk}}{1+\beta_0\langle z_k,\beta_k\rangle}, & t=1;\ j=1,2,\ldots,r,\\[2mm] \dfrac{z_{jk}\beta_0\beta_{jk}}{1+\beta_0\langle z_k,\beta_k\rangle}, & t=2;\ j=1,2,\ldots,r,\end{cases} \tag{1}$$

where $\langle , \rangle$ is the usual inner product, $z_k = (z_{1k}, z_{2k}, \ldots, z_{rk})^{\mathrm{T}}$ a given extra set data connected to the control area of each couple k. The latter authors are interested in the m.l.e. $\hat{\alpha}$ of α and its numerical consistency under the restraints $h_k(\alpha) = 0$ where

$$h_k(\alpha) = \langle \mathbf{1}_r, \beta_k\rangle - 1; \quad k = 1,2,\ldots,s, \tag{2}$$

with $\mathbf{1}_r = (1,\ldots,1)^{\mathrm{T}} \in \mathbb{R}^r$, the vector of unity. The main motivation of this Note is to make their paper complete by investigating the asymptotic precision of $\hat{\alpha}$, i.e., the asymptotic variances–covariances matrix.

2. Technical intermediate results

In the following we note $H_\alpha = (H_1, H_2)$ the constraints matrix with $H_1 = \frac{\partial h}{\partial \beta_0}$ a $s \times 1$ matrix, $H_2 = \frac{\partial h}{\partial \beta}$ a $s \times (sr)$ matrix, $h = (h_1, h_2, \ldots, h_s)$, $L(\alpha)$ the log-likelihood function based on the s couples, $J_\alpha = E(-\frac{\partial^2 L}{\partial\alpha\partial\alpha})$, the $(1+sr) \times (1+sr)$ information matrix evaluated at different points on the line segment joining α^0 (the true vector parameter) and $\hat{\alpha}$. Statistical aspects of asymptotic results of restraints m.l.e. $\hat{\alpha}$ and related problems have been extensively discussed in Aitchison and Silvey [1], Crowder [2,3], Harville [4], Lee [5], and Van Eeden [9]. It can been shown under regularity conditions, that $\hat{\alpha}$ is asymptotically normally distributed and its covariance matrix strongly depends on the fact that both J_α and $H_\alpha J_\alpha^{-1} H_\alpha^{\mathrm{T}}$ are nonsingular. Technical lemmas below elaborate these points. The technical demonstrations and further straightforward matrix manipulations are displayed in N'Guessan [7].

Lemma 2.1. *Let us note* $V_{\alpha,k} = \frac{\beta_0}{1+\beta_0\langle z_k,\beta_k\rangle}(z_{1k}, z_{2k}, \ldots, z_{rk})^{\mathrm{T}}$ *a* $r \times 1$ *vector and* $\Omega_{\alpha,k} = \mathrm{diag}(\frac{1+\beta_0 z_{1k}}{\beta_{1k}}, \frac{1+\beta_0 z_{2k}}{\beta_{2k}}, \ldots, \frac{1+\beta_0 z_{rk}}{\beta_{rk}})$ *a* $r \times r$ *matrix. Then for* $k = 1,2,\ldots,s$: (i) $\|z_k\|^2_{\Omega^{-1}_{\alpha,k}} < \frac{\langle z_k,\beta_k\rangle}{\beta_0}$; (ii) $\|V_k\|^2_{\Omega^{-1}_{\alpha,k}} < \frac{\beta_0\langle z_k,\beta_k\rangle}{(1+\beta_0\langle z_k,\beta_k\rangle)^2} < 1$.

Lemma 2.2. *Let us note* $B_{\alpha,k} = \gamma_k(\Omega_{\alpha,k} - V_{\alpha,k}V_{\alpha,k}^{\mathrm{T}})$ *with* $\gamma_k = \frac{n_k}{1+\beta_0\langle z_k,\beta_k\rangle}$. *Then* (i) *for* $k = 1,2,\ldots,s$; $B^{-1}_{\alpha,k} = \gamma_k^{-1}(\Omega^{-1}_{\alpha,k} + t_k\Omega^{-1}_{\alpha,k}V_{\alpha,k}V^{\mathrm{T}}_{\alpha,k}\Omega^{-1}_{\alpha,k})$; $t_k = (1 - \|V_k\|^2_{\Omega^{-1}_{\alpha,k}})^{-1}$; (ii) $\|U_\alpha\|^2_{B_\alpha^{-1}} < \tau_\alpha$; *where* $B_\alpha = \mathrm{Bloc}-\mathrm{diag}(B_{\alpha,1}, \ldots, B_{\alpha,s})$, $U_\alpha = (U^{\mathrm{T}}_{\alpha,1}, \ldots, U^{\mathrm{T}}_{\alpha,s})^{\mathrm{T}}$ *a* $(sr) \times 1$ *vector with* $U_{\alpha,k} = \frac{\gamma_k}{\beta_0}V_{\alpha,k}$ *and* $\tau_\alpha = \sum_{k=1}^{s}\frac{\gamma_k\langle z_k,\beta_k\rangle}{\beta_0(1+\beta_0\langle z_k,\beta_k\rangle)}$.

Using model (1) above and after some straightforward derivations and matrix manipulations one obtains:

$$J_\alpha = \begin{pmatrix} \tau_\alpha & U_\alpha^{\mathrm{T}} \\ U_\alpha & B_\alpha \end{pmatrix}.$$

Following Ouellette [8] we note (J_α/B_α) (resp. (J_α/τ_α)) the Schur complement of B_α (resp. τ_α) in J_α and defined by $(J_\alpha/B_\alpha) = \tau_\alpha - U_\alpha^{\mathrm{T}}B_\alpha^{-1}U_\alpha$ (resp. $(J_\alpha/\tau_\alpha) = B_\alpha - U_\alpha\tau_\alpha^{-1}U_\alpha^{\mathrm{T}}$).

Lemma 2.3. (i) $(J_\alpha/B_\alpha) > 0$; (ii) (J_α/τ_α) *is nonsingular and*

$$(J_\alpha/\tau_\alpha)^{-1} = B_\alpha^{-1} + (J_\alpha/B_\alpha)^{-1}B_\alpha^{-1}U_\alpha U_\alpha^{\mathrm{T}}B_\alpha^{-1}. \tag{3}$$

Lemma 2.4. (i) J_α^{-1} *exits and*

$$J_\alpha^{-1} = \begin{pmatrix} (J_\alpha/B_\alpha)^{-1} & -(J_\alpha/B_\alpha)^{-1}U_\alpha^{\mathrm{T}}B_\alpha^{-1} \\ -B_\alpha^{-1}U_\alpha(J_\alpha/B_\alpha)^{-1} & (J_\alpha/\tau_\alpha)^{-1} \end{pmatrix}; \tag{4}$$

(ii)

$$\big(H_\alpha J_\alpha^{-1} H_\alpha^{\mathrm{T}}\big)^{-1} = \Lambda_\alpha^{-1} - \big((J_\alpha/B_\alpha) + \|\xi_\alpha\|^2_{\Lambda_\alpha^{-1}}\big)^{-1}\Lambda_\alpha^{-1}\xi_\alpha\xi_\alpha^{\mathrm{T}}\Lambda_\alpha^{-1}, \tag{5}$$

where $\xi_\alpha = H_2 B_\alpha^{-1} U_\alpha$ *a* $s \times 1$ *vector and* $\Lambda_\alpha = \mathrm{diag}(\|\mathbf{1}_r\|^2_{B^{-1}_{\alpha,1}}, \|\mathbf{1}_r\|^2_{B^{-1}_{\alpha,2}}, \ldots, \|\mathbf{1}_r\|^2_{B^{-1}_{\alpha,s}})$ *a* $s \times s$ *matrix.*

Remark 1. Lemmas 2.3 and 2.4 provide from general applications of well-known results of Schur complement about the inverse of partitioned matrix (see Ouellette [8], pp. 201–204) and combine model (1) and constraints (2).

3. Main result

The restraints m.l.e. of α_0, provided it exists, is the vector $\hat\alpha$ which satisfies

$$\left(\frac{\partial L}{\partial \alpha}\right)_{\hat\alpha} + H_{\hat\alpha}^{\mathrm{T}}\hat\lambda = 0, \quad h(\hat\alpha) = 0, \tag{6}$$

where λ is a vector of s Lagrange multipliers.

Let $M_\alpha = J_\alpha^{-1}(J_\alpha + (\frac{\partial^2 L}{\partial\alpha\partial\alpha}))$ a $(1+sr) \times (1+sr)$ matrix, $R_\alpha = H_2(J_\alpha/\tau_\alpha)^{-1}H_2^{\mathrm{T}}$ a $s \times s$ matrix, $\Phi_\alpha = J_\alpha^{-1}H_\alpha^{\mathrm{T}}R_\alpha^{-1}$ a $(1+sr) \times s$ matrix.

Theorem 3.1. *Suppose that*

(A1) $(\hat\alpha, \hat\lambda)$ *is a consistent solution of* (6).
(A2) $M_{\hat\alpha} \overset{\text{Proba}}{\longrightarrow} 0$, $\Phi_{\hat\alpha} - \Phi_{\alpha^0} \overset{\text{Proba}}{\longrightarrow} 0$, $|\Phi_{\alpha^0}| < \infty$.
(A3) $(\frac{\partial L}{\partial\alpha})_{\alpha^0} \overset{\text{Distr.}}{\sim} \mathcal{N}(0, J_{\alpha^0})$.

Then the asymptotic components of the covariance matrix of $\hat\alpha$ *are*:

$$\sigma^2(\hat\beta_0) = \big((J_\alpha/B_\alpha) + \|\xi_\alpha\|^2_{\Lambda_\alpha^{-1}}\big)^{-1},$$
$$\mathrm{Var}(\hat\beta) = (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1}H_2^{\mathrm{T}}R_\alpha^{-1}H_2(J_\alpha/\tau_\alpha)^{-1},$$
$$\mathrm{COV}(\hat\beta, \hat\beta_0) = -(J_\alpha/B_\alpha)^{-1}\big(B_\alpha^{-1}U_\alpha - (J_\alpha/\tau_\alpha)^{-1}H_2^{\mathrm{T}}R_\alpha^{-1}\xi_\alpha\big).$$

Proof (outline). Application of the mean value theorem at α^0 to (6) and Lemma 2.4 give

$$\begin{aligned} &\left(\frac{\partial L}{\partial \alpha}\right)_{\alpha^0} + J_\alpha(M_\alpha - I_{1+sr})\big(\hat\alpha - \alpha^0\big) + H_\alpha^{\mathrm{T}}\hat\lambda = 0, \\ &H_\alpha\big(\hat\alpha - \alpha^0\big) = 0. \end{aligned} \tag{7}$$

Under constraints (2) and (A1) to (A3) and using a version of a theorem of Crowder ([3], Section 2), then

$$\begin{pmatrix} \hat\alpha - \alpha^0 \\ \eta_{\hat\alpha} \end{pmatrix} \overset{\text{Distr.}}{\sim} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} W_{\alpha^0} & 0 \\ 0 & R_{\alpha^0} \end{pmatrix}\right), \tag{8}$$

where $W_\alpha = J_\alpha^{-1} - J_\alpha^{-1}H_\alpha^{\mathrm{T}}R_\alpha^{-1}H_\alpha J_\alpha^{-1}$; $\eta_{\hat\alpha} = R_\alpha\hat\lambda$. Now using formula (4) of Lemma 2.4 and after some matrix manipulation (N'Guessan [7], Appendix A.2) the components of W_{α^0} above fellow. We will make capital out of those results at subsequent papers in order to bring forward asymptotic confidence intervals and correlation coefficients (N'Guessan [7], Section 5). □

Acknowledgements

I am especially grateful to Prof. Claude Langrand for his helpful comments and suggestions, which have led to an improvement in this article and my Ph.D. thesis in road safety measure analysis. I wish also thank Prof. Claude Brezinsky who drew my attention to the paper about Schur Complements. I thank the referees and Prof. Paul Deheveuls for constructive suggestions. This research was partially supported by *France–Maroc action intégrée programme* $N^0 MA01/02$.

References

[1] J. Aitchison, S.D. Silvey, Maximum likelihood estimation of parameters subject to restraints, Ann. Math. Statist. 29 (1958) 813–829.
[2] M. Crowder, Maximum likelihood estimation for dependent observations, J. Roy. Statist. Soc. 38 (1975) 45–53.
[3] M. Crowder, On the constrained maximum likelihood estimation with non i.i.d. observations, Ann. Inst. Statist. Math. Part A 36 (1984) 239–249.
[4] D.A. Harville, Maximum likelihood approaches to variance component estimation and to related problems, J. Amer. Statist. Assoc. 72 (358) (1977) 320–340.
[5] S.-Y. Lee, Constrained estimation in covariance structure analysis, Biometrika 66 (3) (1979) 539–545.
[6] A. N'Guessan, A. Essai, C. Langrand, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière, Rev. Statist. Appl. XLIX (2) (2001) 83–100.
[7] A. N'Guessan, On a use of Schur complements for an inverse of a constrained accident data information matrix, Pub. IRMA Lille 60 (VI) (2003).
[8] D.V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981) 187–295.
[9] C. Van Eeden, Estimation in restricted parameter spaces – some history and some recent developments, CWI Quarterly 9 (1 and 2) (1996) 69–76.