ELSEVIER

Statistics & Probability Letters 55 (2001) 53–61

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

On the efficiency of extended generalized estimating equation approaches

Brajendra C. Sutradhar[a,*], Pranesh Kumar[b]

[a]*Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NF, Canada A1C 5S7*
[b]*Department of Mathematics and Computer Science, University of Northern British Columbia, Prince George, BC, Canada V2N 4Z9*

Received December 2000; received in revised form March 2001

Abstract

On top of the generalized estimating equation (GEE) approach, there exist two extended generalized estimating equation (EGEE) approaches where two sets of estimating equations are simultaneously solved for the estimation of the regression and the so-called 'working' correlation parameters. The loss of efficiency of the GEE approach based regression estimators was recently studied by Sutradhar and Das (Biometrika 86 (1999) 459). In this paper, we study the efficiency loss problem for the two EGEE approaches by utilizing the approach of Sutradhar and Das. © 2001 Elsevier Science B.V. All rights reserved

Keywords: Consistency; Efficiency; Generalized linear model; Misspecification of correlation structure; Normal based higher-order moments; Regression estimator; Working correlation

1. Introduction

Suppose that a scalar response y_{it} and a p-dimensional vector of covariates x_{it} are observed for clusters $i=1,\dots,K$, at a time point t $(t=1,\dots,n_i)$. For the ith cluster, let $y_i=(y_{i1},\dots,y_{it},\dots,y_{in_i})^{\mathrm{T}}$ be the response vector and $X_i=(x_{i1},\dots,x_{it},\dots,x_{in_i})^{\mathrm{T}}$ be the $n_i \times p$ matrix of covariates. Further suppose that the marginal density of the responses y_{it} is of the exponential family form

$$f(y_{it})=\exp[\{y_{it}\theta_{it}-a(\theta_{it})\}\phi+b(y_{it},\phi)], \tag{1.1}$$

(cf. Liang and Zeger, 1986) where $\theta_{it}=h(\eta_{it})$ with $\eta_{it}=x_{it}^{\mathrm{T}}\beta$, $a(\cdot)$, $b(\cdot)$, and $h(\cdot)$ are of known functional form, ϕ is a possibly unknown scale parameter, and β is the $p\times 1$ vector of parameters of interest. It then follows that $E(Y_{it})=a'(\theta_{it})$ and $\mathrm{var}(Y_{it})=\phi^{-1}a''(\theta_{it})$, where $a'(\theta_{it})$ and $a''(\theta_{it})$ are, respectively, the first and second derivatives of $a(\theta_{it})$ with respect to θ_{it}.

* Corresponding author.
E-mail addresses: bsutradh@math.mun.ca (B.C. Sutradhar), kumarp@unbc.ca (P. Kumar).

PII: S0167-7152(01)00127-4

In the cluster regression set-up, the components of the response vector y_i are correlated. Let $C_i(\rho)$ be the $n_i \times n_i$ true correlation matrix of Y_i $(i=1,\dots,K)$, which is usually unknown in practice. Here, ρ is an $s_1 \times 1$ vector of correlation parameters which fully characterizes $C_i(\rho)$. For $A_i = \text{diag}[a''(\theta_{it})]$, and for known $C_i(\rho)$, the quasilikelihood estimator $\tilde{\beta}$ of β under (1.1), is the solution of the score equation

$$\sum_{i=1}^{K} X_i^{\mathrm{T}} A_i \Sigma_i^{-1}(\rho)(y_i - \mu_i) = 0, \tag{1.2}$$

where $\mu_i = (a'(\theta_{i1}),\dots,a'(\theta_{it}),\dots,a'(\theta_{in_i}))^{\mathrm{T}}$ and $\Sigma_i(\rho) = \phi^{-1} A_i^{1/2} C_i(\rho) A_i^{1/2}$ is the true covariance matrix of Y_i. In many important situations, for example for binary and Poisson data, one may use $\phi = 1$. In what follows, we therefore consider the case $\phi = 1$, for simplicity. Under mild regularity conditions, $K^{1/2}(\tilde{\beta} - \beta)$ is asymptotically multivariate normal with zero mean vector and covariance matrix V_{T} given by

$$V_{\mathrm{T}} = \lim_{K\to\infty} K \left\{ \sum_{i=1}^{K} X_i^{\mathrm{T}} A_i^{1/2} C_i^{-1}(\rho) A_i^{1/2} X_i \right\}^{-1}. \tag{1.3}$$

In practice, $C_i(\rho)$ is unknown. This makes it impossible to estimate β by solving the estimating Eq. (1.1). To overcome the problem of unknown $C_i(\rho)$, Liang and Zeger (1986) have used a 'working' correlation matrix $R_i(\alpha)$ for $C_i(\rho)$ and solved the estimating equations

$$\sum_{i=1}^{K} X_i^{\mathrm{T}} A_i^{1/2} R_i^{-1}(\hat{\alpha}) A_i^{-1/2}(y_i - \mu_i) = 0, \tag{1.4}$$

where α, an $s_2 \times 1$ vector of correlation parameters fully characterizes $R(a)$ matrix, and $\hat{\alpha}$ is a consistent estimate of α. Let $\hat{\beta}_{\mathrm{G}}$ be the solution for β based on the generalized estimating equation (GEE) (1.4).

In the context of Liang–Zeger (1986) model, Fitzmaurice et al. (1993, Eqs. (2) and (4), pp. 286–87) use two sets of generalized estimating equations to estimate β and the 'working' correlation parameter α. Their estimating equation for α is quite similar to that of Prentice and Zhao (1991) but unlike Fitzmaurice et al. (1993), Prentice and Zhao (1991) estimate the true correlation parameters ρ. Thus, Fitzmaurice et al. (1993) used an extended GEE (EGEE) approach for the estimation of the parameters as in the original paper of Liang and Zeger. We refer to this technique as the EGEE2 approach. Note that similar to Prentice and Zhao (1991), this EGEE2 approach of Fitzmaurice et al. (1993) requires the computations of the third- and the fourth-order moments of the responses, in order to construct the generalized estimating equation for α. The exact computation of these higher-order moments is, however, not possible. One standard approach for the estimation of these moments is to use the covariance matrix of the responses by pretending it as the covariance matrix of the normally distributed responses (Prentice and Zhao, 1991). To avoid the complexity of computing higher-order moments, recently Hall and Severini (1998) use an extended GEE approach that uses only up to second-order moments. We refer to this approach as the EGEE1 approach. In this EGEE1 approach, Hall and Severini (1993) like Fitzmaurice et al. (1993), jointly estimate the regression and the 'working' correlation parameters α. Similar to Liang and Zeger (1986), by using the $R(\alpha)$ matrix in the covariance matrices of the estimates of the regression and the 'working' correlation parameters, Hall and Severini (1998) have also examined different types of efficiency losses due to misspecification of the correlation structure of the responses. We refer to their Tables 3–5 (Hall and Severini, 1998, pp. 1370–1371) in particular to make this point clear. But as suggested by Sutradhar and Das (1999), the computations of the efficiencies should be based on the $R(\alpha_0^*)$ matrix, where α_0^* is the convergent value of the estimate of α, irrespective of the approaches whether β and α are jointly or separately estimated. The main objective of the paper is to deal with this important issue of computing the efficiencies correctly to understand the merits of the proposed EGEE1 and EGEE2 approaches for the estimation of the parameters, mainly the regression parameters.

2. EGEE1 and EGEE2 based estimators for working correlations and their convergence

2.1. EGEE1 based estimator and convergence

Let $\hat{\alpha}_{\mathrm{EG}}$ denote the EGEE1 based estimator of α due to Hall and Severini (1998). Note, however, that as the α parameter involved in the $R_i(\alpha)$ (1.4) working correlation matrix is subject to uncertainty of definition (cf. Crowder, 1995), it is essential to see the limiting value of $\hat{\alpha}_{\mathrm{EG}}$ obtained based on EGEE1 approach. Suppose that $\hat{\alpha}_{\mathrm{EG}}$ converges to a quantity $\tilde{\alpha}$, say. We then compute the efficiency of the EGEE1 based estimator of β, say $\hat{\beta}_{\mathrm{EG}}$, by computing the covariance matrix of $\hat{\beta}_{\mathrm{EG}}$ with $\tilde{\alpha}$ for α (cf. Sutradhar and Das, 1999), whereas Hall and Severini computed the covariance matrix of $\hat{\beta}_{\mathrm{EG}}$ by using $R_i(\alpha)$ matrix itself (cf. Tables 3–5 in Hall and Severini, 1998). We now examine the convergence of $\hat{\alpha}_{\mathrm{EG}}$ due to misspecification of the correlation structure.

Following the notation in (1.4), let $u_i=[u_{i1}^{\mathrm{T}},u_{i2}^{\mathrm{T}}]^{\mathrm{T}}$, where $u_{i1}=[(y_{i1}-\mu_{i1})^2,\ldots,(y_{it}-\mu_{it})^2,\ldots,(y_{in_i}-\mu_{in_i})^2]^{\mathrm{T}}$, and $u_{i2}=[(y_{i1}-\mu_{i1})(y_{i2}-\mu_{i2}),\ldots,(y_{it}-\mu_{it})(y_{it'}-\mu_{it'}),\ldots,(y_{i(n_i-1)}-\mu_{i(n_i-1)})(y_{in_i}-\mu_{in_i})]^{\mathrm{T}}$ are, respectively, the n_i and $n_i(n_i-1)/2$ dimensional vectors of corrected squares and distinct cross-products of the observations $y_{i1},\ldots,y_{it},\ldots,y_{in_i}$ for all $i=1,\ldots,K$. Under the assumption of working correlation structure $R_i(\alpha)$ or the working covariance matrix $V_i(\beta,\alpha)=A_i^{1/2}R_i(\alpha)A_i^{1/2}=(v_{itt'})$, let $\tilde{v}_{i1}=(v_{i11},\ldots,v_{itt},\ldots,v_{in_in_i})^{\mathrm{T}}$ and $\tilde{v}_{i2}=(v_{i212},\ldots,v_{i2tt'},\ldots,v_{i2(n_i-1)n_i})^{\mathrm{T}}$, respectively, be the vectors of diagonal and distinct off-diagonal elements of the $V_i(\beta,\alpha)$ matrix. In the EGEE1 approach, one then solves the estimating equations (Hall and Severini, 1998)

$$K^{-1}\sum_{i=1}^{K}\frac{\partial\mu_i^{\mathrm{T}}}{\partial\beta}V_i^{-1}(\beta,\alpha)(y_i-\mu_i)=0 \tag{2.1}$$

and

$$K^{-1}\sum_{i=1}^{K}[W_{id}^{\mathrm{T}},W_{i\bar{d}}^{\mathrm{T}}][(u_{i1}-\tilde{v}_{i1})^{\mathrm{T}},(u_{i2}-\tilde{v}_{i2})^{\mathrm{T}}]^{\mathrm{T}}=0 \tag{2.2}$$

for β and α jointly, where W_{id} and $W_{i\bar{d}}$ are $n_i\times 1$ and $n_i(n_i-1)/2\times 1$ vectors consisting of the diagonal and distinct off-diagonal elements of the $W_i(\alpha)=\partial V_i^{-1}(\beta,\alpha)/\partial\alpha$ matrix, respectively. In

$$W_i(\alpha)=-A_i^{-1/2}R_i^{-1}(\alpha)\frac{\partial R_i(\alpha)}{\partial\alpha}R_i^{-1}(\alpha)A_i^{-1/2}, \tag{2.3}$$

the specific form of $\partial R_i(\alpha)/\partial\alpha$ will depend on the structure of the $R_i(\alpha)$ matrix, chosen as the working correlation matrix.

Note that to obtain the best possible estimate of α, one assumes that β is known. Thus, one solves Eq. (2.2) only to obtain $\hat{\alpha}_{\mathrm{EG}}$. Now by similar arguments as in Sutradhar and Das (1999), $\hat{\alpha}_{\mathrm{EG}}$ from (2.2) would converge to a quantity determined by the form chosen for the working correlation matrix $R_i(\alpha)$ and the form of the true correlation matrix $C_i(\rho)$ for the data y_i. This is because $W_i(\alpha)$, $\tilde{v}_{i1}$ and $\tilde{v}_{i2}$ are functions of α, whereas $E(u_{i1})$ and $E(u_{i2})$ are naturally the functions of the elements of the $C_i(\rho)$ matrix, $C_i(\rho)$ being the true correlation structure of the data. In order to examine the loss or gain of efficiency of the EGEE1 based regression estimator, one first needs to compute $\tilde{\alpha}$, the convergent quantity of $\hat{\alpha}_{\mathrm{EG}}$, due to misspecification of the correlation structure. We do this by solving the equation

$$K^{-1}\sum_{i=1}^{K}[W_{id}^{\mathrm{T}},W_{i\bar{d}}^{\mathrm{T}}][(\tilde{\sigma}_{i1}-\tilde{v}_{i1})^{\mathrm{T}},(\tilde{\sigma}_{i2}-\tilde{v}_{i2})^{\mathrm{T}}]^{\mathrm{T}}=0 \tag{2.4}$$

for α as a function of ρ. Note that Eq. (2.4) is obtained from (2.2) by taking expectations of u_{i1} and u_{i2} under the true correlation structure. Consequently, in (2.4), $\tilde{\sigma}_{i1}$ and $\tilde{\sigma}_{i2}$ are the $n_i\times 1$ and $n_i(n_i-1)/2\times 1$ vectors consisting of the diagonal and distinct off-diagonal elements of the $\Sigma_i(\rho)=A_i^{1/2}C_i(\rho)A_i^{1/2}$ matrix, respectively.

Next, for three commonly encountered correlation processes, as in Liang and Zeger (1986), namely for equi-correlation (EQC), MA(1) and AR(1) processes, we simplify Eq. (2.4) to compute $\hat{\alpha}_{\mathrm{EG}}$ and obtain its convergent value $\tilde{\alpha}$ under misspecification of correlation structures, as follows. For the true MA(1) versus working EQC, it can be shown after some algebras that $\hat{\alpha}_{\mathrm{EG}}$ obtained from (2.4) converges to $\tilde{\alpha}=2\rho/n$. For the true MA(1) versus working AR(1) structure, $\hat{\alpha}_{\mathrm{EG}}$ appears to converge to $\tilde{\alpha}=\rho$, where $-0.5<\rho<0.5$.

For the true AR(1) versus working equicorrelation process, $\hat{\alpha}_{\mathrm{EG}}$ converges to $\tilde{\alpha}$ satisfying

$$\tilde{\alpha}=[2\rho/\{(1-\rho)n(n-1)\}]\left[n-\frac{1-\rho^n}{1-\rho}\right],$$

whereas for the true AR(1) versus working MA(1) structure, $\hat{\alpha}_{\mathrm{EG}}$ appears to converge to

$$\tilde{\alpha}=\sum_{i=1}^{K}\sum_{j=1}^{n-1}\left\{\rho^j\sum_{k=1}^{n-j}r_2^{k,k+j}\right\}\Big/\sum_{i=1}^{K}\sum_{k=1}^{n-1}r_2^{k,k+1}, \tag{2.5}$$

where $r_2^{j,k}=\sum_{t=1}^{n}(r^{j,t-1}r^{t,k}+r^{j,t+1}r^{t,k})$, with $r^{j,k}$ as the (j,k)th element of the inverse matrix of the working correlation matrix $R_{i,M}(\alpha)$, say, for the moving average process of order 1. The formulas for these elements $r^{j,k}$ are available in Shaman (1969, Eq. (8)) (see also Tanaka and Satchell (1989)), for example. Note that in this true AR(1) versus working MA(1) structure, $\tilde{\alpha}$ appears to satisfy a complicated polynomial relationship with ρ. For a given ρ, $\tilde{\alpha}$ may be solved by using a trial and error (or search method) technique so that Eq. (2.5) is satisfied. Next for the true EQC versus working AR(1) case, $\hat{\alpha}_{\mathrm{EG}}$ appears to converge to $\tilde{\alpha}=\rho$, whereas for the true EQC versus working MA(1) structure, $\hat{\alpha}_{\mathrm{EG}}$ converges to

$$\tilde{\alpha}=\frac{\rho\sum_{j<k=1}^{n-1}\sum_{t=1}^{n}(r^{j,t-1}r^{t,k}+r^{j,t+1}r^{t,k})}{\sum_{j=1}^{n-1}\sum_{t=1}^{n}(r^{j,t-1}r^{t,t+1}+r^{j,t+1}r^{t,i+1})}, \tag{2.6}$$

where $r^{j,k}$, the (j,k)th element of the inverse matrix of $R_{i,M}(\alpha)$ is computed as in (2.5). The values of $\tilde{\alpha}$ to the corresponding values of ρ for all three correlation structures are shown in Table 1.

2.2. EGEE2 based estimator and convergence

Recall from (1.4) that in the GEE approach, Liang and Zeger (1986) estimate the regression vector β by solving the estimating equation

$$\sum_{i=1}^{K}\frac{\partial\mu_i^{\mathrm{T}}}{\partial\beta}V_i^{-1}(\beta,\hat{\alpha})(y_i-\mu_i)=0,$$

where the 'working' correlation estimate $\hat{\alpha}$ is computed separately by using the method of moments. In this section, unlike the Hall and Severini (1998) joint estimation approach EGEE1 for β and α, we consider a third and fourth moments based direct generalization of the Liang–Zeger approach for such joint estimation (cf. Fitzmaurice et al., 1993), which has been referred to as the EGEE2 approach.

Define $f_i=[(y_i-\mu_i)^{\mathrm{T}},(u_{i1}-\tilde{v}_{i1})^{\mathrm{T}},(u_{i2}-\tilde{v}_{i2})^{\mathrm{T}}]^{\mathrm{T}}$, and

$$D_i=\begin{bmatrix}\dfrac{\partial\mu_i}{\partial\beta^{\mathrm{T}}} & 0\\[2ex] \dfrac{\partial\tilde{v}_i}{\partial\beta^{\mathrm{T}}} & \dfrac{\partial\tilde{v}_i}{\partial\alpha}\end{bmatrix},\quad \Omega_i=\begin{bmatrix}\mathrm{var}(y_i) & \mathrm{cov}(y_i,u_i)\\ & \mathrm{var}(u_i)\end{bmatrix},$$

where $u_i=[u_{i1}^{\mathrm{T}},u_{i2}^{\mathrm{T}}]^{\mathrm{T}}$ with u_{i1} and u_{i2} as the vectors of corrected squares and distinct cross-products of the n_i observations $y_{i1},\ldots,y_{in_i}$ under the ith cluster, and $\tilde{v}_i=(\tilde{v}_{i1}^{\mathrm{T}},\tilde{v}_{i2}^{\mathrm{T}})^{\mathrm{T}}$ with $E(u_{i1})=\tilde{v}_{i1}$ and $E(u_{i2})=\tilde{v}_{i2}$ under the

Table 1
Percentage relative efficiency of $R(\alpha)$ working correlation based $\hat{\beta}_{\mathrm{I}} = (\hat{\beta}_{0\mathrm{I}}, \hat{\beta}_{1\mathrm{I}})^{\mathrm{T}}$, $\hat{\beta}_{\mathrm{G}} = (\hat{\beta}_{0(\mathrm{G})}, \hat{\beta}_{1(\mathrm{G})})^{\mathrm{T}}$ and $\hat{\beta}_{\mathrm{EG}} = (\hat{\beta}_{0(\mathrm{EG})}, \hat{\beta}_{1(\mathrm{EG})})^{\mathrm{T}}$ to the generalized estimator $\tilde{\beta} = (\tilde{\beta}_0, \tilde{\beta}_1)^{\mathrm{T}}$ with true correlation matrix $C(\rho)$, for $\log\{\mu_{it}\} = \beta_0 + \beta_1 t/n$ with $\beta_0 = \beta_1 = 1$ for $n = 5$

		Estimation approach										
		Working independence		EGEE2 with working normal based higher moments			EGEE2 with working independence based higher moments			EGEE1		
$R(a)\|C(\rho)$	ρ	$\hat{\beta}_{0\mathrm{I}}$	$\hat{\beta}_{1\mathrm{I}}$	α_1^*	$\hat{\beta}_{0(\mathrm{G})}$	$\hat{\beta}_{1(\mathrm{G})}$	α_2^*	$\hat{\beta}_{0(\mathrm{G})}$	$\hat{\beta}_{1(\mathrm{G})}$	$\tilde{\alpha}$	$\hat{\beta}_{0(\mathrm{EG})}$	$\hat{\beta}_{1(\mathrm{EG})}$
EQC|MA(1)	−0.3	97	97	−0.13	97	97	−0.13	97	97	−0.12	97	97
	0.1	100	100	0.04	100	100	0.04	100	100	0.04	100	100
	0.2	99	99	0.08	99	99	0.08	99	99	0.08	99	99
	0.4	95	94	0.17	95	94	0.17	95	94	0.16	95	94
AR(1)|MA(1)	−0.3	97	97	−0.29	100	100	—	—	—	−0.3	100	100
	0.1	100	100	0.10	100	100	—	—	—	0.10	100	100
	0.2	99	99	0.20	100	100	—	—	—	0.20	100	100
	0.4	95	94	0.37	99	99	—	—	—	0.40	100	99
EQC|AR(1)	−0.3	96	96	−0.10	96	96	−0.10	96	96	−0.10	96	96
	0.3	98	98	0.16	98	98	0.16	98	98	0.15	98	98
	0.7	97	95	0.53	97	95	0.53	97	95	0.52	97	95
MA(1)|AR(1)	−0.7	88	87	−0.36	98	98	−0.49	99	99	−0.02	89	88
	−0.3	96	96	−0.26	100	100	−0.30	100	100	−0.03	97	96
	0.3	98	98	0.27	100	100	0.30	100	100	0.21	100	100
	0.7	97	96	0.39	100	100	0.49	99	99	0.05	98	97
AR(1)|EQC	0.3	100	100	0.33	98	96	0.05	100	100	0.3	98	97
	0.7	100	100	−0.53	98	94	0.05	100	100	0.7	96	87
	0.9	100	100	−0.52	99	96	0.05	100	100	0.9	98	88
MA(1)|EQC	0.1	100	100	0.09	100	100	0.10	100	100	0.12	100	100
	0.3	100	100	0.22	99	98	0.30	97	95	0.12	100	99
	0.4	100	100	0.27	98	96	0.40	95	90	0.06	100	100
	0.49	100	100	0.31	98	95	0.49	91	82	0.03	100	100

'working' correlation model. Then, in EGEE2 approach, one obtains the estimates of β and α simultaneously by solving the joint estimating equations

$$\sum_{i=1}^{K} D_i^{\mathrm{T}} \Omega_{iw}^{-1} f_i = 0 \tag{2.7}$$

for β and α, where Ω_{iw}^{-1} represents a working covariance matrix of f_i. Two different choices of Ω_{iw} matrix will be considered. First, the Ω_{iw} matrix will be constructed by pretending $\operatorname{var}(y_i) = V_i(\beta,\alpha) \equiv V_i(\beta,\alpha)$ as though it is the covariance matrix of the normal vector y_i. Second, Ω_{iw} matrix will be constructed by pretending that the components of y_i vector follow the true exponential family model but they are independent.

We now concentrate on the estimation of the working correlation parameter α as in the following. Irrespective of the choice for the working covariance matrix, we now estimate α by assuming that β is known. It then follows from (2.7) that the estimating equation for α is given by

$$\sum_{i=1}^{K} \frac{\partial \tilde{v}_i^{\mathrm{T}}}{\partial \alpha} V_{iuu}^{-1}(u_i - \tilde{v}_i) = 0, \tag{2.8}$$

where V_{iuu} is the covariance matrix of u_i, and $\tilde{v}_i = [\tilde{v}_{i1}^{\mathrm{T}}, \tilde{v}_{i2}^{\mathrm{T}}]^{\mathrm{T}}$. Let $\hat{\alpha}_{\mathrm{G}}$ be the solution of (2.8) for α. There is, however, no guarantee that $\hat{\alpha}_{\mathrm{G}}$ will converge to α, as the working correlation parameter α is subject to an uncertainty of definition (cf. Crowder, 1995). In order to examine this, we now simplify the estimating equation (2.8) for two different types of third- and fourth-order moment matrix V_{iuu}, analogously to Hall and Severini (1998). First, we consider a working structure for V_{iuu} based on the assumption that the observations have Gaussian distribution with correct mean vector μ_i and covariance matrix $V_i(\beta, \alpha)$. Second we construct a working structure for V_{iuu} under the assumption that the observations are independent following the exponential family density (1.1) so that one may compute the necessary third- and fourth-order moments by exploiting the density (1.1) itself. The two estimators based on these two approaches will be denoted by $\hat{\alpha}_{\mathrm{G1}}$ and $\hat{\alpha}_{\mathrm{G2}}$, respectively.

2.2.1. Normal based EGEE2 estimator of α

Let

$$V_{iuu} = \begin{bmatrix} m_{11}^* & m_{12}^* \\ & m_{22}^* \end{bmatrix}, \tag{2.9}$$

where $m_{11}^* = \mathrm{var}(u_{i1})$, $m_{12}^* = \mathrm{cov}(u_{i1}, u_{i2})$, and $m_{22}^* = \mathrm{var}(u_{i2})$, u_{i1} and u_{i2} being the n_i and $n_i(n_i-1)/2$ dimensional vectors of squares and cross-products, respectively. For $j, k = 1, \ldots, n_i$, the general (j,k)th element of m_{11}^* is given by

$$\begin{aligned} m_{11}^*(j,k) &= E_w[\{(y_{ij} - \mu_{ij})^2 - v_{ijj}(\beta,\alpha)\}\{(y_{ik} - \mu_{ik})^2 - v_{ikk}(\beta,\alpha)\}] \\ &= 2\alpha_{|j-k|}^2 v_{ijj}(\beta,\alpha) v_{ikk}(\beta,\alpha), \end{aligned} \tag{2.10}$$

as under normality and the working expectation assumption $E_w(y_i - \mu_i)(y_i - \mu_i)^{\mathrm{T}} = V_i(\beta,\alpha) = (v_{ijk}(\beta,\alpha))$, one obtains $E_w[(y_{ij} - \mu_{ij})^2 (y_{ik} - \mu_{ik})^2] = v_{ijj}(\beta,\alpha) v_{ikk}(\beta,\alpha) + 2v_{ijk}^2(\beta,\alpha)$. In (2.10), $\alpha_{|j-k|}$ is the (j,k)th element of the $R_i(\alpha)$ matrix in $V_i(\beta,\alpha)$, with $\alpha_0 = 1$. The elements of the m_{12}^* and m_{22}^* matrices may be computed similarly.

Note that when $R_i(\alpha)$ is assumed to be a working correlation matrix following AR(1), MA(1) or equicorrelation process, s_2 becomes 1 implying that α is a scalar parameter. Since these processes are widely used in practice, in this paper we consider $s_2 = 1$ only. It then follows that $\partial \tilde{v}_i^{\mathrm{T}} / \partial \alpha$ in (2.8) is a $1 \times n_i(n_i+1)/2$ vector given by

$$\begin{aligned} \frac{\partial \tilde{v}_i^{\mathrm{T}}}{\partial \alpha} &= \left[0^{\mathrm{T}}, \frac{\partial}{\partial \alpha}\{\alpha_{|1-2|}(v_{i11}v_{i22})^{1/2}, \ldots, \alpha_{|j-k|}(v_{ijj}v_{ikk})^{1/2}, \ldots, \alpha_{|(n_i-1)-n_i|}(v_{i(n_i-1)(n_i-1)}v_{in_in_i})^{1/2}\}\right] \\ &= [b_{11}, \ldots, b_{n_in_i}, b_{12}, \ldots, b_{jk}, \ldots, b_{(n_i-1)n_i}] \quad \text{(say)} \end{aligned} \tag{2.11}$$

Remark that estimating equation (2.8) is unbiased only with respect to α, the working correlation parameter. Now suppose that $C_i(\rho)$ is the true correlation structure of the data $y_{i1}, \ldots, y_{in_i}$ as mentioned in (1.2). Consequently, $E(u_i)$ in (2.8) is no longer $\tilde{v}_i$, although $E(u_{i1}) = \tilde{v}_{i1}$. This is because, for $C_i(\rho) = (\rho_{|j-k|})$ with $\rho_0 = 1$,

$$\begin{aligned} E(u_{i2}) = [&\{v_{i11}v_{i22}\}^{1/2}\rho_{|1-2|}, \ldots, \{v_{i11}v_{in_in_i}\}^{1/2}\rho_{|1-n_i|}, \{v_{i22}v_{i33}\}^{1/2} \\ &\rho_{|2-3|}, \ldots, \{v_{i(n_i-1)(n_i-1)}v_{in_in_i}\}^{1/2}\rho_{|(n_i-1)-n_i|}]^{\mathrm{T}} \end{aligned}$$

leading to

$$E(u_i - \tilde{v}_i) = [d_{11}, \ldots, d_{n_in_i}, d_{12}, \ldots, d_{jk}, \ldots, d_{(n_i-1)n_i}]^{\mathrm{T}} \quad \text{(say)}, \tag{2.12}$$

where for $j<k$, $d_{jk}=\{v_{ijj}v_{ikk}\}^{1/2}(\rho_{|j-k|}-\alpha_{|j-k|})$. Next, for $j,k=1,\ldots,n_i(n_i+1)/2$, let $V_{iuu}^{-1}=(c_{jk})$. Now, by combining (2.11) and (2.12), one solves

$$\sum_{k=1}^{n_i^*}\sum_{j=1}^{n_i^*} b_{1j}c_{jk}d_{k1}=0, \tag{2.13}$$

following (2.8) for α in terms of ρ, which will be the convergent value of $\hat{\alpha}_{G1}$. In (2.13), $n_i^*=n_i(n_i+1)/2$. Let α_1^* be the convergent value of $\hat{\alpha}_{G1}$. For clarity, take for example the data considered by Hall and Severini (1998), where $v_{itt}=\mu_{itt}=\exp\{\beta_1+\beta_2 t/n_i\}$ with $\beta_1=\beta_2=1$, following a Poisson model. We, however, assume that these cluster data ($n_i=5$ say) really follow an equi-correlation process, whereas one uses MA(1) working correlation structure to estimate α and hence β. For such a case, the elements b_{jk} in (2.11) are: $b_{jk}=\{v_{ijj}v_{ikk}\}^{1/2}$ for $|j-k|=1$, and $b_{jk}=0$ for $|j-k|>1$; the elements d_{jk} in (2.12) are: $d_{jj}=0$ for $j=1,\ldots,n_i$, $d_{jk}=(\rho-\alpha)\{v_{ijj}v_{ikk}\}^{1/2}$ for $|j-k|=1$, and $d_{jk}=\rho\{v_{ijj}v_{ikk}\}^{1/2}$ for $|j-k|>1$; and the elements of the submatrices in (2.9) are computed by putting $\alpha_{|j-k|}=\alpha$ for $|j-k|=1$ and $\alpha_{|j-k|}=0$ for $|j-k|>1$, yielding the appropriate values for c_{jk}. Now for $\rho=0.1$, 0.3, 0.4 and 0.49, the solutions for α by (2.13) are given by $\alpha_1^*=0.09$, 0.22, 0.27 and 0.31, respectively. Note that these α_1^* values should be used in the efficiency computations for the estimates of β, but not the values of α. In the same manner, we also compute the α_1^* values for other model misspecification such as: true equi-correlation versus AR(1) working processes; true AR(1) versus working MA(1) and equi-correlation processes; and true MA(1) versus working AR(1) and equi-correlation processes. The α_1^* values for all these cases are reported in Table 1. The efficiencies for the β estimates are computed in Section 2.2.2.

2.2.2. *Independence (for third and fourth moments) based EGEE2 estimator of α*

In this approach, one computes the third- and fourth-order working covariance matrix V_{iuu} under the assumption that the observations follow the true exponential family model but they are independent. The formulas for b_{jk} and d_{jk} remain the same as in Section 2.2.1 constructed for the normal based GEE2 estimator of α.

We now develop the submatrices m_{11}^*, m_{12}^* and m_{22}^* for the construction of the V_{iuu} matrix in (2.8). By similar calculations as in (2.10), one obtains

$$m_{11}^*(j,j)=m_{ij4}-2m_{ij2}v_{ijj}+v_{ijj}^2 \quad \text{for } j=1,\ldots,n_i \tag{2.14}$$

and for $j\neq k$, $j,\ k=1,\ldots,n_i$, $m_{11}^*(j,k)=m_{ij2}m_{ik2}-v_{ikk}m_{ij2}-v_{ijj}m_{ik2}+v_{ijj}v_{ikk}$, where $m_{ij2}=a''(\theta_{ij})$ and $m_{ij4}=a^{\mathrm{IV}}(\theta_{ij})+3\{a''(\theta_{ij})\}^2$ are the second and fourth moments of the exponential family (1.1) based response variable y_{ij}. For the Poisson case $m_{ij4}=\mu_{ij}+3\mu_{ij}^2$, and $m_{ij2}=\mu_{ij}$ with $\mu_{ij}=v_{ijj}=\sigma_{ijj}$. It then follows that for the Poisson model

$$m_{11}^*(j,j)=\mu_{ij}+2\mu_{ij}^2 \quad \text{for } j=1,\ldots,n_i \tag{2.15}$$

and $m_{11}^*(j,k)=0$, otherwise. The elements of m_{12}^* and m_{22}^* matrices may be computed similarly.

Now, by similar operations as in the case for normal based EGEE2 estimation, we solve (2.13) for α in terms of ρ. Note that c_{jk} are obtained from the inverse matrix of V_{iuu}. The b_{1j} and d_{k1} are the same as in (2.13) for the normal based EGEE2 estimation for α. For the same equi-correlation $\rho=0.1$, 0.3, 0.4, and 0.49 used in the normal based EGEE2 estimation, we now obtain $\alpha_2^*=0.1$, 0.3, 0.4, and 0.49 , as the convergent values of $\hat{\alpha}_{G2}$, which are the same as the values of ρ. Here, α was chosen to characterize the working correlation matrix of the MA(1) process. Similarly to the values of α_1^*, the values of α_2^* are also computed for other model misspecification, and they are reported in the same Table 1.

3. Efficiency under correlation structure misspecification

Recall that for known true correlation structure $C_i(\rho)$, the quasi-likelihood estimator $\tilde{\beta}$ obtained by solving the estimating equation (1.2) has the asymptotic covariance matrix given by (1.3). We identify this covariance matrix as $\mathrm{var}(\tilde{\beta})$. Similarly, for the EGEE1 based estimator $\hat{\beta}_{\mathrm{EG}}$, we first compute the covariance matrix of $(\hat{\beta}_{\mathrm{EG}}^{\mathrm{T}}, \hat{\alpha}_{\mathrm{EG}})^{\mathrm{T}}$, where $\hat{\beta}_{\mathrm{EG}}^{\mathrm{T}}$ and $\hat{\alpha}_{\mathrm{EG}}$ are obtained from (2.1) and (2.2) jointly, and then use the first p diagonal elements of this covariance matrix as the variances of the elements of $\hat{\beta}_{\mathrm{EG}}$. Let $\mathrm{var}(\hat{\beta}_{\mathrm{EG}})$ denote the $p \times p$ covariance matrix of $\hat{\beta}_{\mathrm{EG}}$. Next, we obtain the covariance of EGEE2 based estimator $\hat{\beta}_{\mathrm{G}}$ from the covariance matrix of $(\hat{\beta}_{\mathrm{G}}^{\mathrm{T}}, \hat{\alpha}_{\mathrm{G}})^{\mathrm{T}}$ computed based on (2.7). We denote this covariance matrix by $\mathrm{var}(\hat{\beta}_{\mathrm{G}})$.

For the Poisson model considered by Hall and Severini (1998) (see also Section 2), we now apply the proposed approach to deal with misspecification of the correlation structure and compute the efficiency of $\hat{\beta}_{\mathrm{EG}}$ and $\hat{\beta}_{\mathrm{G}}$ by using $\mathrm{var}(\tilde{\beta})/\mathrm{var}(\hat{\beta}_{\mathrm{EG}})$ and $\mathrm{var}(\tilde{\beta})/\mathrm{var}(\hat{\beta}_{\mathrm{G}})$, respectively. For efficiency comparison, we also include the working independence estimator $\hat{\beta}_{\mathrm{I}}$, which is the solution of the independence estimating equation (IEE)

$$\sum_{i=1}^{K} X_i^{\mathrm{T}}(y_i - \mu_i(\beta)) = 0 \tag{3.1}$$

and which has the asymptotic covariance V_{I} given by

$$V_{\mathrm{I}} = \lim_{K\to\infty} \left(\sum_{i=1}^{K} X_i^{\mathrm{T}} A_i X_i\right)^{-1} \left(\sum_{i=1}^{K} X_i^{\mathrm{T}} A_i^{1/2} C_i(\rho) A_i^{1/2} X_i\right) \left(\sum_{i=1}^{K} X_i^{\mathrm{T}} A_i X_i\right)^{-1}. \tag{3.2}$$

The efficiencies of $\hat{\beta}_{\mathrm{I}}$, $\hat{\beta}_{\mathrm{EG}}$ and $\hat{\beta}_{\mathrm{G}}$ with respect to $\tilde{\beta}$ are reported in Table 1.

It is clear from Table 1 that in four out of six cases, namely for the cases with true MA(1) versus working EQC, true AR(1) versus working EQC, true EQC versus working AR(1) and MA(1) correlation structures, the working independence estimator $\hat{\beta}_{\mathrm{I}}$ is either equally or more efficient as compared to the EGEE1 and EGEE2 estimators of β. In other two cases, $\hat{\beta}_{\mathrm{I}}$ appears to trail to both EGEE1 and EGEE2 estimators. The EGEE1 and Gaussian EGEE2 estimators appear to be equally efficient when MA(1) is a true correlation structure but one uses AR(1) as the working correlation structure. In the last case, when AR(1) is a true correlation structure but one uses MA(1) as the working correlation structure, Gaussian and independence EGEE2 estimators appear to be more efficient than the EGEE1 based estimators. Thus, in general, EGEE1 and EGEE2 approaches do not appear to perform well as compared to the working independence estimator $\hat{\beta}_{\mathrm{I}}$. Note that similar conclusion was reached by Sutradhar and Das (1999) regarding the performance of the Liang–Zeger (1986) GEE estimator as compared to the working independence estimating equations (IEE) based estimator of β. Since, EGEE1 and EGEE2 are more complex to compute the estimate of β, the findings of the present paper clearly reveal that in a situation where it is not possible to specify true correlation structure of the responses, it is much better to use the simple IEE approach as opposed to the EGEE1 and EGEE2 approaches in estimating β.

4. Concluding remarks

It was shown by Sutradhar and Das (1999) that in the longitudinal regression set-up, the generalized regression estimators (Liang and Zeger, 1986) may be less efficient as compared to the regression estimators obtained by using the independence estimating equations approach. In the present paper, it is shown that the extended generalized estimators of the regression coefficients may also be less efficient than the regression estimators obtained based on the independence estimating equations approach. This findings clearly suggests that in situations where it is not possible to specify the correct correlation structure, it is much better to use the independence based regression estimators which are easy to compute.

Acknowledgements

This research was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Crowder, M., 1995. On the use of a working correlation matrix in using generalized linear models for repeated measures. Biometrika 82, 407–410.

Fitzmaurice, G.M., Laird, N.M., Rotnitzky, A.G., 1993. Regression models for discrete longitudinal responses (with discussion). Statist. Sci. 8, 284–309.

Hall, D.B., Severini, T.A., 1998. Extended generalized estimating equations for clustered data. J. Amer. Statist. Assoc. 93, 1365–1375.

Liang, K.Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 78, 13–22.

Prentice, R.L., Zhao, L.P., 1991. Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. Biometrics 47, 825–839.

Shaman, P., 1969. On the inverse of the covariance matrix of a first order moving average. Biometrika 56, 595–600.

Sutradhar, B.C., Das, K., 1999. On the efficiency of regression estimators in generalized linear models for longitudinal data. Biometrika 86, 459–465.

Tanaka, K., Satchell, S.E., 1989. Asymptotic properties of the maximum-likelihood and nonlinear least-squares estimators for noninvertible moving average models. Econometric Theory 5, 333–353.