A Computationally Efficient Algorithm to Solve Generalized Method of Moments Estimating Equations based on Secant-Vector Divisions Procedure

N. Mamode Khan n.mamodekhan@uom.ac.mu
Faculty of Social Studies and Humanities
/Department of Economics and Statistics
University of Mauritius
Reduit, Mauritius

M. Heenaye m.mamodekhan@uom.ac.mu
Faculty of Engineering
/Department of Computer Science and Engineering
University of Mauritius
Reduit, Mauritus

Abstract

Generalized method of moment estimating function enables one to estimate regression parameters consistently and efficiently. However, it involves one major computational problem: in complex data settings, solving generalized method of moments estimating function via Newton-Raphson technique gives rise often to non-invertible Jacobian matrices. Thus, parameter estimation becomes unreliable and computationally difficult. To overcome this problem, we propose to use secant method based on vector divisions instead of the usual Newton-Raphson technique to estimate the regression parameters. This new method of estimation demonstrates a decrease in the number of non-convergence iterations as compared to the Newton-Raphson technique and provides reliable estimates. We compare these two estimation approaches through a simulation study.

Keywords: Quadratic Inference Function, Newton-Raphson, Jacobian, Secant Method, Vector Divisions.

1. INTRODUCTION

GENERALIZED method of moments (GMM) is a popular tool developed by Hansen (1982) to estimate regression parameters especially in settings where the number of equations exceeds the number of unknown parameters. Recently, Qu et al. (2000) and Qu and Lindsay (2003) have formulated a GMM function known as the quadratic inference function (QIF) to analyze the effects of explanatory variables on repeated responses. Since, in general, the correlation structure of repeated measures are unknown, Qu et al.(2000) assume a working structure which can be decomposed into several basis matrices. These basis matrices are then combined to form a score vector whose dimension is quite large. The objective is to use this score vector to estimate the vector of regression parameters. The authors proposed to construct a generalized moment estimating function based on this score vector and thereafter use calculus to optimize the function and obtain estimates of the regression parameters. The optimized function is non-linear and thus, the Newton-Raphson algorithm is implemented to solve iteratively the equation. However, we often remark in simulation studies that the Jacobian matrix of the Newton-Raphson iterative equation is close to singularity. This may lead to unreliable parameter estimates or a blockage of the iterative process. Our objective in this paper is to apply an alternative iterative method that omits the computation of the inverse Jacobian matrix. Yixun (2008) and Mamode Khan (2011) considered the secant method of estimation which is based on vector divisions. In this paper, we use this approach to estimate the regression parameters based on the GMM objective function and compare the Newton-Raphson estimation approach with the Secant method based vector

divisions. The comparison between these two techniques is made through simulating AR(1) correlated Poisson counts with different covariate designs.
The organization of the paper is as follows: In section 2, we provide the estimating equations of the Generalized method of moments and its estimation procedures. In section 3, we introduce the method of secant iterative scheme based on vector divisions following Mamode Khan (2011). In the next section, we present a simulation study whereby we generate AR(1) correlated Poisson counts and use GMM and Secant based on vector divisions to estimate the regression parameters. In the last section, we provide the conclusions and recommendations based on comparisons of these two techniques.

2. GENERALIZED METHOD OF MOMENTS

Qu et al. (2000) introduced a GMM of the form of a quadratic objective function that combines an extended score function with its covariance matrix.
The construction of the extended score function is based on the generalized estimating equation, that is

$$g(\beta)=\sum_{i=1}^{I}\left(\frac{\partial\mu_i}{\partial\beta^T}\right)^T V_i^{-1}(y_i-\mu_i)=0 \tag{1}$$

where

$$V_i = A_i^{\frac{1}{2}}R(\alpha)A_i^{\frac{1}{2}} \tag{2}$$

and $R(\alpha)$ is the working correlation structure. Their method of GMM models the inverse of the correlation structure $R(\alpha)^{-1}$ by a class of matrices

$$R(\alpha)^{-1}=\sum_{i=1}^{m}a_i M_i \tag{3}$$

where $M_1, M_2,....,M_m$ are known basis matrices and $a_1,a_2,...,a_M$ are constants. Equation (3) can accommodate the popular correlation structures. Equation (1) can then be written as

$$g(\beta)=\sum_{i=1}^{I}\left(\frac{\partial\mu_i}{\partial\beta^T}\right)^T A_i^{-\frac{1}{2}}(a_1M_1+a_2M_2+.....+a_mM_m)A_i^{-\frac{1}{2}}(y_i-\mu_i)=0 \tag{4}$$

Based on this representation, Qu et al. (2000) defined an extended score

$$g^*(\beta)=\frac{1}{I}\sum_{i=1}^{I}g_i(\beta)=\frac{1}{I}\begin{pmatrix}\sum_{i=1}^{I}\left(\frac{\partial\mu_i}{\partial\beta^T}\right)^T A_i^{-\frac{1}{2}}M_1A_i^{-\frac{1}{2}}(y_i-\mu_i)\\ \sum_{i=1}^{I}\left(\frac{\partial\mu_i}{\partial\beta^T}\right)^T A_i^{-\frac{1}{2}}M_2A_i^{-\frac{1}{2}}(y_i-\mu_i)\\ \vdots\\ \sum_{i=1}^{I}\left(\frac{\partial\mu_i}{\partial\beta^T}\right)^T A_i^{-\frac{1}{2}}M_mA_i^{-\frac{1}{2}}(y_i-\mu_i)\end{pmatrix} \tag{5}$$

In principle, the vector $g^*(\beta)$ contains more equations than parameters but they can be combined optimally following GMM to form a quadratic objective function of the form

$$S(\beta)=[g^*(\beta)]^T W^{-1}[g^*(\beta)] \tag{6}$$

Where

$$W=\frac{1}{I^2}\sum_{i=1}^{I}g_i(\beta)g_i(\beta)^T \tag{7}$$

The idea is to minimize $S(\beta)$. Qu et al. (2000) showed that asymptotically,

$$\dot{S}(\beta) = 2\left[\frac{\partial g^*}{\partial \beta^T}\right]^T W^{-1} g^* \tag{8}$$

$$\ddot{S}(\beta) = 2\left[\frac{\partial g^*}{\partial \beta^T}\right]^T W^{-1} \left[\frac{\partial g^*}{\partial \beta^T}\right] \tag{9}$$

Then the vector of regression parameters β is estimated iteratively using the Newton-Raphson technique

$$\hat{\beta}_{(j+1)} = \hat{\beta}_{(j)} - [\ddot{S}(\hat{\beta}_j)]^{-1}\, \dot{S}(\hat{\beta}_j) \tag{10}$$

Qu et al. (2000) showed that asymptotically $\hat{\beta}$ is consistent and its variance reaches the Cramer-Rao type lower bound. The algorithm works as follows: After assuming a working structure for the basis matrices, we construct $\dot{S}(\beta)$ and $\ddot{S}(\beta)$ for an initial of vector regression parameters $\hat{\beta}_0$. We replace in equation (10) to obtain an updated $\hat{\beta}_1$. We then use $\hat{\beta}_1$ to obtain $\ddot{S}(\hat{\beta}_1)$ and $\dot{S}(\hat{\beta}_1)$. However, this iterative equation may not be successful in estimating parameters for every type of setting. In fact, we carried out an experiment that involves the simulation of correlated Poisson counts and noted that while estimating the regression parameters, the Jacobian matrix $\ddot{S}(\hat{\beta})$ often turns out to be singular and ill-conditioned. This ultimately blocks the computation process. To remedy the situation, we propose an alternative approach known as the secant method based on vector divisions to estimate the parameters. In the next section, we introduce the secant method and show its iterative scheme.

3. SECANT METHOD

The traditional secant iterative formula to estimate a scalar parameter β is given by

$$\hat{\beta}_{r+1} = \hat{\beta}_r - \frac{F(\hat{\beta}_r)}{F[\hat{\beta}_r, \hat{\beta}_{r-1}]} \tag{11}$$

with

$$F[\hat{\beta}_r, \hat{\beta}_{r-1}] = \frac{F(\hat{\beta}_r) - F(\hat{\beta}_{r-1})}{\hat{\beta}_r - \hat{\beta}_{r-1}} \tag{12}$$

where $F(\beta) = 0$

However, this iterative formula cannot be applied directly to obtain the vector of regression parameters β in equation (8) since β is here multi-dimensional. To overcome this issue, Yixun Shi (2008) developed an iterative multi-dimensional secant formula using vector divisions. We adapt his procedures to solve equation (8). By letting $F(\beta) = \dot{S}(\beta)$, we estimate iteratively using

$$\hat{\beta}_{j+1} = \hat{\beta}_j - \frac{(\hat{\beta}_j - \hat{\beta}_{j-1})^T (\hat{\beta}_j - \hat{\beta}_{j-1})}{(\hat{\beta}_j - \hat{\beta}_{j-1})^T (F(\hat{\beta}_j) - F(\hat{\beta}_{j-1}))} F(\hat{\beta}_j) \tag{13}$$

or

$$\hat{\beta}_{j+1} = \hat{\beta}_j - \frac{(F(\hat{\beta}_j) - F(\hat{\beta}_{j-1}))^T F(\hat{\beta}_j)}{(F(\hat{\beta}_j) - F(\hat{\beta}_{j-1}))^T (F(\hat{\beta}_j) - F(\hat{\beta}_{j-1}))} (\hat{\beta}_j - \hat{\beta}_{j-1}) \tag{14}$$

The iterative process works as follows: For initial values of $\hat{\beta}_0, \hat{\beta}_1$, we calculate $\hat{\beta}_2$ using equation (13) or (14). Then using $\hat{\beta}_1$ and $\hat{\beta}_2$, we calculate $\hat{\beta}_3$. The iterative process continues until convergence, i.e, $\| \hat{\beta}_{t+1} - \hat{\beta}_t \| < 10^{-5}$. However, to ensure convergence, we can use a steepest direction coefficient following Mamode Khan (2011) and Yixun (2008).

4 SIMULATION STUDY

In this section, we generate AR(1) correlated Poisson counts following McKenzie (1986) with true mean parameters $\beta_0 = 1, \beta_1 = 1$. Note that in GMM, the bases matrices M_0, M_1 and M_2 are : M_0: The identity matrix, M_1 has one on the two main off-diagonals and M_2 has 1 on the corners $(1,1)$ and $(n.n)$. We consider different covariates designs:

Design 1: $$x_{it1} = 1, (i = 1, \ldots, \frac{I}{4}), x_{it2} = 0, (i = \frac{I}{4} + 1, \ldots, \frac{3I}{4}), x_{it3} = 1, (i = \frac{3I}{4} + 1, \ldots, I) \quad (15)$$

Design 2: $$x_{it1} = rbin(3,0.5), (i = 1, \ldots, \frac{I}{4}), x_{it2} = 0, (i = \frac{I}{4} + 1, \ldots, \frac{3I}{4}), x_{it3} = rbin(3,0.5), (i = \frac{3I}{4} + 1, \ldots, I) \quad (16)$$

Design 3: $$x_{it1} = rpois(3), (i = 1, \ldots, \frac{I}{4}), x_{it2} = 0, (i = \frac{I}{4} + 1, \ldots, \frac{3I}{4}), x_{it3} = rpois(0.7), (i = \frac{3I}{4} + 1, \ldots, I) \quad (17)$$

and for the second covariate x_{it2}, we generate I standard normal values.

For each design, we run 5,000 simulations for $I = 20, 60, 100$ and 500. The following tables provide the simulated mean of the estimates and the number of non-convergent simulations under both techniques.

Size	$\hat{\beta}_{1,NR}$	$\hat{\beta}_{2\,NR}$	$\hat{\beta}_{1\sec}$	$\hat{\beta}_{2\sec}$	Number of non-convergent simulations in the Newton-Raphson	Number of non-convergent simulations in the Secant approach
20	0.9621	0.9872	0.9632	0.9881	2341	1550
60	0.9991	0.9999	0.9992	1.0054	1201	910
100	0.9990	1.0024	1.0014	0.9996	825	534
400	1.0010	0.9998	0.9999	0.9999	230	100
600	0.9999	0.9999	0.9999	0.9999	98	30
1000	1.0000	1.0000	1.0000	1.0000	54	10

TABLE 1: Estimates of GMM regression parameters under Newton-Raphson and Secant method: Design 1

Size	$\hat{\beta}_{1,NR}$	$\hat{\beta}_{2\,NR}$	$\hat{\beta}_{1\sec}$	$\hat{\beta}_{2\sec}$	Number of non-convergent simulations in the Newton-Raphson	Number of non-convergent simulations in the Secant approach
20	0.9943	0.9899	0.9942	0.9892	2562	1899
60	0.9993	0.9997	0.9992	0.9998	1666	1032
100	0.9990	1.0003	0.9999	0.9999	1321	998
400	1.0001	0.9998	0.9999	0.9999	344	223
600	0.9999	0.9999	0.9999	0.9999	142	97
1000	1.0000	1.0000	1.0000	1.0000	76	55

TABLE 2: Estimates of GMM regression parameters under Newton-Raphson and Secant method: Design 2

Size	$\hat{\beta}_{1,NR}$	$\hat{\beta}_{2\,NR}$	$\hat{\beta}_{1\sec}$	$\hat{\beta}_{2\sec}$	Number of non-convergent simulations in the Newton-Raphson	Number of non-convergent simulations in the Secant approach
20	0.9897	0.9899	0.9901	0.9900	1040	999
60	0.9997	0.9999	0.9998	1.0001	889	762
100	0.9999	1.0001	1.0000	0.9996	762	566
400	1.0001	0.9999	0.9999	0.9999	444	320
600	0.9999	0.9999	0.9999	0.9999	102	87
1000	1.0000	1.0000	1.0000	1.0000	65	34

TABLE 3: Estimates of GMM regression parameters under Newton-Raphson and Secant method: Design 3

For each design, we assume small initial values of the mean parameters to run the simulations. As noted, there is no huge discrepancy between the estimated parameters and the true value of the regression parameters. As the cluster size increases, the discrepancies become lesser under both estimation techniques in all of the designs. This is in accordance with the consistency properties of the estimators under the GMM approach. As regards to the number of non-convergent simulations, the Newton-Raphson technique reports a comparatively higher number of non-convergent simulations than the Secant method as the Jacobian matrix becomes close to singularity. This problem was noted in almost all cluster sizes. However, as the cluster size increases, the number of non-convergent simulations decreases significantly under both approaches. The non-convergence problem also occurs because of the choice of the steepest descent coefficient as reported by Mamode Khan (2011). Under some simulations, these coefficients were modified to yield convergence and to speed convergence. Based on the simulation results, we may conclude that GMM based on the secant method using vector divisions is a computationally fast and efficient approach. Also, its computational complexities compared with the Newton-Raphson method will be lesser. In the same context, Mamode Khan (2011) showed through simulation studies that the secant method is an efficient estimation approach from a computational perspective as it reduces the number of non-convergent simulations and provides equally consistent estimates.

5 : CONCLUSION

Generalized method of moments is an efficient estimation approach that yields consistent and reliable estimates of regression parameters particularly in an over-determined system of non-linear equations but its estimation procedures often give rise to singular Jacobian matrices. This makes computation quite difficult. In this paper, we propose an alternative to Newton-Raphson known as the Secant method based on vector divisions. This approach omits the computation of the Jacobian matrix and provides equally consistent and reliable estimates than GMM under Newton-Raphson approach. Another advantage of this method is the computational complexities are lesser than Newton-Raphson as the inverse of a matrix requires quite a number of flop counts. Based on simulation results, we note that both Newton-Raphson and Secant method based on vector divisions yield consistent estimates but the secant method yields fewer non-convergent simulations than Newton-Raphson. However, care must be taken when choosing initial values of the parameters. Otherwise, the resulting estimates may be unreliable. Thus, we may conclude that GMM based Secant method is a more optimal estimation methodology.

6 : REFERENCES

[1] Hansen, L. (1982) ‘Large sample properties of generalized method of moments estimators’. Econometrika 50, 1029-54.

[2] Mamode Khan, N. (2011), ‘ Use of vector equations in solving quasi-likelihood equations for Poisson model’, Journal of Statistical computation and simulation, iFirst,1-6, ISSN 0094-9655

[3] McKenzie, E. (1986), ‘Autoregressive moving-average processes with negative binomial and geometric marginal distributions. Advanced Applied Probability, 18, 679-705.

[4] Qu, A and Lindsay, B. (2003), ‘ Building adaptive estimating equations when inverse of covariance estimation is difficult’. Journal of Royal Statistical Society 65, 127-142.

[5] Qu, A, Lindsay, B and Li, B. (2000), ‘ Improving generalized estimating equations and quadratic inference functions’. Biometrika 91(2), 447-59.

[6] Yixun,S. (2008) ‘Using vector divisions in solving non-linear equations’. International Journal of Contemporary Math Sciences 3(16): 753-759.