Computational Economics **10:** 231–250, 1997.

Computing 3SLS Solutions of Simultaneous Equation Models with a Possible Singular Variance–Covariance Matrix

ERRICOS J. KONTOGHIORGHES[1*] and ELIAS DINENIS[2]
[1]*Institut d'Informatique, Université de Neuchâtel, Rue Emile-Argand 11, CH-2007 Neuchâtel, Switzerland. e-mail: kontoghiorghes@info.unine.ch*
[2]*Centre for Mathematical Trading and Finance, City University Business School, Frobisher Crescent, Barbican Centre, London EC2Y 8HB, UK*
([*] *Corresponding author*)

Abstract. Algorithms for computing the three-stage least squares (3SLS) estimator usually require the disturbance covariance matrix to be non-singular. However, the solution of a reformulated simultaneous equation model (SEM) results into the redundancy of this condition. Having as a basic tool the QR decomposition, the 3SLS estimator, its dispersion matrix and methods for estimating the singular disturbance covariance matrix are derived. Expressions revealing linear combinations between the observations which become redundant have also been presented. Algorithms for computing the 3SLS estimator after the SEM has been modified by deleting or adding new observations or variables are found not to be very efficient, due to the necessity of removing the endogeneity of the new data or by re-estimating the disturbance covariance matrix. Three methods have been described for solving SEMs subject to *separable* linear equalities constraints. The first method considers the constraints as additional precise observations while the other two methods reparameterized the constraints to solve reduced unconstrained SEMs. Methods for computing the main matrix factorizations illustrate the basic principles to be adopted for solving SEMs on serial or parallel computers.

Key words: simultaneous equation models, 3SLS, QR decomposition, parallel algorithms

1. Introduction

It is not always possible for the disturbance covariance matrix of a simultaneous equation model (SEM) to be non-singular. In allocation models, for example, or models with precise observations that imply linear constraints on the parameters, or models in which the number of structural equations exceeds the number of observations, the disturbance covariance matrix is singular (Judge et al., 1985; Theil, 1971). Such models can be estimated using generalized inverses which are expensive and can lead to loss of accuracy. Here we provide a computational strategy for solving an alternative formulation of the 3SLS problem in which the disturbance covariance matrix can be singular (Court, 1974; Narayanan, 1969; Pollock, 1979; Srivastava and Tiwari, 1978; Zellner and Theil, 1962).

First, recent methods for solving SURE (seemingly unrelated regression equation) (Kontoghiorghes, 1993b; Kontoghiorghes and Clarke, 1995a; Kontoghiorghes and Dinenis, 1996b) models are extended to 3SLS estimation of SEMs. Efficient, stable approaches for computing the coefficient estimates and their dispersion matrix are presented. Then means for updating SEMs and their solutions with linear equality constraints are presented. Finally, strategies for computing relevant matrix operations are suggested.

2. 3SLS Estimation of SEMs

The ith structural equation of the SEM can be written as

$$y_i = X_i\beta_i + Y_i\gamma_i + \varepsilon_i, \quad i = 1, \ldots, G, \tag{1}$$

where, for the ith structural equation, $y_i \in \Re^T$ is the dependent vector, X_i is the $T \times k_i$ matrix of full column rank of exogenous variables, Y_i is the $T \times g_i$ matrix of other included endogenous variables, β_i and γ_i are the structural parameters, and $\varepsilon_i \in \Re^T$ are the disturbance terms. For $W_i \equiv (X_i \; Y_i)$ and $\delta_i^T \equiv (\beta_i^T \; \gamma_i^T)$, the stacked system of the structural equations can be written as

$$\mathsf{vec}(Y) = (I_G \otimes W)S\delta + \varepsilon, \tag{2}$$

where $W \equiv (X \; Y), X$ is a $T \times K$ matrix of all predetermined variables, $Y \equiv (y_1 \ldots y_G), S \equiv \mathsf{diag}(S_1, \ldots, S_G)$, where S_i is a selector matrix such that $WS_i = W_i \; (i = 1 \ldots G), \delta^T \equiv (\delta_1^T \ldots \delta_G^T)$, and $\varepsilon^T \equiv (\varepsilon_1^T \ldots \varepsilon_G^T)$. The disturbance vector ε satisfies $\mathrm{E}(\varepsilon) = 0$ and $\mathrm{E}(\varepsilon\varepsilon^T) = \Sigma \otimes I_T$, where Σ is a $G \times G$ non-negative definite matrix. It is assumed that all structural equations are identifiable, that is, $e_i = k_i + g_i \leqslant K$.

The 2SLS and Generalised LS (GLS) estimators of (2) are defined, respectively, from the application of OLS and GLS to the transformed SEM (hereafter TSEM)

$$\mathsf{vec}(Q_1^T Y) = (I_G \otimes Q_1^T W)S\delta + (I_G \otimes Q_1^T)\varepsilon,$$

or

$$\mathsf{vec}(R^{(2)}) = (I_G \otimes R)S\delta + \tilde{\varepsilon}, \tag{3}$$

where $\tilde{\varepsilon} = (I_G \otimes Q_1^T)\varepsilon$, and the Qs and Rs come from the *incomplete* QR decomposition (QRD) of the augmented matrix $W = (X \; Y)$ given by

$$Q^T W = \begin{array}{cc} \begin{array}{cc} K & G \end{array} & \\ \begin{pmatrix} R^{(1)} & R^{(2)} \\ 0 & R^{(3)} \end{pmatrix} & \begin{array}{l} {\scriptstyle K} \\ {\scriptstyle T-K} \end{array} \end{array}, \quad \text{with} \quad Q = \begin{array}{cc} K & T-K \\ (Q_1 & Q_2), \end{array} \tag{4}$$

$Q \in \Re^{T \times T}$ orthogonal, $R^{(1)}$ upper triangular and non-singular, and

$$R = (R^{(1)} \;\; R^{(2)}) = Q_1^T W. \tag{5}$$

Note that $E(\widetilde{\varepsilon}) = 0$ and $E(\widetilde{\varepsilon}, \widetilde{\varepsilon}^T) = \Sigma \otimes I_K$.

The 3SLS estimator, denoted by $\widehat{\delta}_{\text{3SLS}}$, is the GLS estimator with Σ replaced by its consistent estimator $\widehat{\Sigma}$ based on the 2SLS residuals. Computing the Cholesky decomposition

$$\widehat{\Sigma} = CC^T, \tag{6}$$

the $\widehat{\delta}_{\text{3SLS}}$ estimator derives from the solution of the normal equations

$$(\widetilde{W}^T \widetilde{W})\widehat{\delta}_{\text{3SLS}} = \widetilde{W}^T \widetilde{y}, \tag{7}$$

where C is a $K \times K$ non-singular upper triangular matrix, $\widetilde{W} = (C^{-1} \otimes R)S$, and $\widetilde{y} = (C^{-1} \otimes I_K)\mathsf{vec}(R^{(2)}) = \mathsf{vec}(R^{(2)}C^{-T})$ (Belsley, 1992; Dent, 1976; Jennings, 1980).

For $\widehat{\Sigma}$ singular or badly ill-conditioned the above estimation procedure fails, since C^{-1} either does not exist or computes badly. However, this can be overcomed by rewriting the TSEM (3) in the equivalent form

$$\mathsf{vec}(R^{(2)}) = (I_G \otimes R)S\delta + (\widehat{C} \otimes I_K)V, \tag{8}$$

where the rank of $\widehat{\Sigma} = \widehat{C}\widehat{C}^T$ is $g \leqslant G$, $\widehat{C} \in \Re^{G \times g}$ has a full column rank, and V is a random gK element vector with zero mean and variance–covariance matrix I_{gK}, defined as $(\widehat{C} \otimes I_K)V = \widetilde{\varepsilon}$. Under this formulation, the 3SLS estimator of δ comes from the solution to the generalized linear least squares problem

$$\underset{\delta}{\operatorname{argmin}}\, V^T V \;\; \text{subject to} \;\; \mathsf{vec}(R^{(2)}) = (I_G \otimes R)S\delta + (\widehat{C} \otimes I_K)V, \tag{9}$$

which does not require that the variance–covariance matrix be non-singular (Kontoghiorghes, 1993b; Kontoghiorghes and Clarke, 1995a; Kontoghiorghes and Dinenis, 1996b; Kourouklis and Paige, 1981; Paige, 1979a; Paige, 1979b).

For the solution of (9) consider the following QRDs involving $(I_G \otimes R)S$ and $(\widehat{C} \otimes I_K)$:

$$\widetilde{Q}^T((I_G \otimes R)S \quad \mathsf{vec}(R^{(2)})) = \begin{array}{cl} \begin{array}{cc} E & 1 \end{array} & \\ \begin{pmatrix} \widetilde{R} & \widetilde{y}^{(1)} \\ 0 & \widetilde{y}^{(2)} \\ 0 & \widetilde{y}^{(3)} \end{pmatrix} & \begin{array}{l} E \\ q \\ GK-E-q \end{array} \end{array} \tag{10}$$

and

$$\widetilde{Q}^T(\widehat{C} \otimes I_K)P = \begin{array}{cl} \begin{array}{cc} gK-q & q \end{array} & \\ \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \\ 0 & 0 \end{pmatrix} & \begin{array}{l} E \\ q \\ GK-E-q \end{array} \end{array}. \tag{11}$$

Here $\widetilde{R} \equiv \mathsf{diag}(\widetilde{R}^{(1)}, \ldots, \widetilde{R}^{(G)})$, and the $\widetilde{R}^{(i)} \in \Re^{e_i \times e_i}$ and L_{22} are upper triangular non-singular matrices; $\widetilde{Q}$ and P are $GK \times GK$ and $gK \times gK$ orthogonal matrices, respectively; $E = \Sigma_{i=1}^{G} e_i$, and $E + q$ is the column rank of $((I_G \otimes R)S\ (\widehat{C} \otimes I_K))$ (Anderson et al., 1992; Golub and Loan, 1983; Paige, 1990; De Moor and Van Dooren, 1992). The orthogonal matrix $\widetilde{Q}$ is defined as

$$\widetilde{Q}^T = \begin{pmatrix} I_E & 0 \\ 0 & \widetilde{Q}_C^T \end{pmatrix} \begin{pmatrix} \widetilde{Q}_A^T \\ \widetilde{Q}_B^T \end{pmatrix} = \begin{array}{l} \begin{pmatrix} \widetilde{Q}_A^T \\ \widetilde{Q}_C^T \widetilde{Q}_B^T \end{pmatrix} \begin{array}{l} E \\ GK-E \end{array} \end{array}, \tag{12}$$

where the QRD of RS_i $(i = 1, \ldots, G)$ and the *complete* QRD of $\widetilde{Q}_B^T(\widehat{C} \otimes I_K)$ (Golub and Loan, 1983; Lawson and Hanson, 1974) are given, respectively, by

$$\widetilde{Q}_i^T(RS_i) = \begin{pmatrix} \widetilde{R}^{(i)} \\ 0 \end{pmatrix}, \quad \text{with } \widetilde{Q}_i = \begin{array}{cc} e_i & K - e_i \\ (\widetilde{Q}_{Ai} & \widetilde{Q}_{Bi}) \end{array} \tag{13}$$

and

$$\begin{aligned} &\widetilde{Q}_C^T(\widetilde{Q}_B^T(\widehat{C} \otimes I_K))P = \begin{pmatrix} 0 & L_{22} \\ 0 & 0 \end{pmatrix}, \\ &\widetilde{Q}_A \equiv \mathsf{diag}(\widetilde{Q}_{A1}, \ldots, \widetilde{Q}_{AG}), \qquad \widetilde{Q}_B \equiv \mathsf{diag}(\widetilde{Q}_{B1}, \ldots, \widetilde{Q}_{BG}), \end{aligned} \tag{14}$$

$$\begin{pmatrix} \widetilde{Q}_{Ai}^T \\ \widetilde{Q}_{Bi}^T \end{pmatrix} R_{\cdot i}^{(2)} = \begin{pmatrix} \widetilde{y}_i^{(1)} \\ \widehat{y}_i \end{pmatrix} \begin{array}{l} e_i \\ K-e_i \end{array}, \tag{15}$$

$R_{\cdot i}^{(2)}$ is the ith column of $R^{(2)}, \widetilde{y}^{(1)} \equiv \mathsf{vec}(\widetilde{y}_1^{(1)} \ldots \widetilde{y}_G^{(1)})$ and

$$\widetilde{Q}_C^T\ \mathsf{vec}(\widehat{y}_1 \ldots \widehat{y}_G) = \begin{pmatrix} \widetilde{y}^{(2)} \\ \widetilde{y}^{(3)} \end{pmatrix}. \tag{16}$$

Conformally partitioning $\widetilde{V}^T = V^T P$ as $(\widetilde{V}_1^T\ \widetilde{V}_2^T)$, it follows that the SEM is consistent iff

$$\widetilde{y}^{(3)} = 0, \tag{17}$$

where $\widetilde{V}_2$ derives from the solution of the triangular system $L_{22}\widetilde{V}_2 = \widetilde{y}^{(2)}$ and the arbitrary vector $\widetilde{V}_1$ is set to zero. The 3SLS estimator is the solution to the block triangular system $\widetilde{R}\widehat{\delta}_{3\text{SLS}} = \widetilde{y}^{(1)} - L_{12}\widetilde{V}_2$, which can be equivalently written as

$$\widetilde{R}^{(i)}\widehat{\delta}_{3\text{SLS}}^{(i)} = \widetilde{y}_i^{(1)} - h_i, \quad (i = 1, \ldots, G), \tag{18}$$

where $\widehat{\delta}_{3\text{SLS}}^{(i)} \in \Re^{e_i}$ corresponds to the 3SLS estimator of δ_i, and $\widetilde{V}_2^T L_{12}^T = (h_1^T \ldots h_G^T)$. Elementary algebraic manipulations produce

$$\widehat{\delta}_{3\text{SLS}}^{(i)} = \delta_i + (\widetilde{R}^{(i)})^{-1}\Lambda_i\widetilde{V}_1,$$

implying that $\mathrm{E}(\widehat{\delta}^{(i)}_{\text{3SLS}}) = \delta_i$ and that the covariance matrix between $\widehat{\delta}^{(i)}_{\text{3SLS}}$ and $\widehat{\delta}^{(j)}_{\text{3SLS}}$ is given by

$$\mathsf{Cov}(\widehat{\delta}^{(i)}_{\text{3SLS}}, \widehat{\delta}^{(j)}_{\text{3SLS}}) = \Gamma_i \Gamma_j^T, \quad (i, j = 1, \ldots, G), \tag{19}$$

where Γ_p $(p = i, j)$ is computed from the solution of the triangular system $\widetilde{R}^{(p)}\Gamma_p = \Lambda_p, L_{11}^T = (\Lambda_1^T \ldots \Lambda_G^T)$ and $\Lambda_p^T \in \Re^{e_p \times (gk-q)}$ (Kontoghiorghes and Clarke, 1995a).

2.1. Computing $\widehat{\Sigma}$

A consistent estimator of Σ is computed by $\widehat{\Sigma} = U^T U/T$, where $U = (u_1, \ldots, u_G)$ denotes the residuals of the structural equations. Initially, U is formed from the residuals of the 2SLS estimators

$$\widehat{\delta}^{(i)}_{\text{2SLS}} = (\widetilde{R}^{(i)})^{-1} \widetilde{y}_i^{(1)}, \quad (i = 1, \ldots, G),$$

that is,

$$\mathsf{vec}(U) = \mathsf{vec}(Y) - (I_G \otimes W) S \widehat{\delta}_{\text{2SLS}}.$$

Since $U^T U = U^T Q Q^T U$, the premultiplication of both sides of the latter by $I_G \otimes Q^T$ gives

$$\mathsf{vec}(Q^T U) = \mathsf{vec}(\overline{R}^{(2)}) - (I_G \otimes \overline{R}) S \widehat{\delta}_{\text{2SLS}}, \tag{20}$$

where in (4) $Q^T W = \overline{R} = (\overline{R}^{(1)}\ \overline{R}^{(2)})$. Then, residuals iteratively based on 3SLS estimators are used to recompute $Q^T U$, until convergence has been achieved.

If $\widehat{\Sigma}$ is computed explicitly, then $\widehat{C}$ in (9) could be obtained by removing the $G - g$ zero columns of the Cholesky factor C in (6) (Lawson and Hanson, 1974). An alternative numerically stable method is to compute the factorization

$$Q_u^T U \Pi_u = \begin{array}{cc} \begin{array}{cc} g & G-g \end{array} & \\ \begin{pmatrix} \overline{C}_{11} & \overline{C}_{12} \\ 0 & 0 \end{pmatrix} & \begin{array}{l} g \\ T-g \end{array} \end{array} \quad \text{or} \quad U = \widetilde{Q}_u (\widetilde{C}_{11}\ \widetilde{C}_{12}) \Pi_u^T, \tag{21}$$

where $\widetilde{Q}_u$ comprises the first g columns of the orthogonal matrix $Q_u \in \Re^{T \times T}$, $\widetilde{C}_{11}$ is upper-triangular and non-singular, and Π_u is a $G \times G$ permutation matrix. In this case, the matrix $\widehat{C}$ is defined as

$$\widehat{C} = \frac{1}{\sqrt{T}} \Pi_u \begin{pmatrix} \widetilde{C}_{11}^T \\ \widetilde{C}_{12}^T \end{pmatrix}.$$

The factorization (21) could be extended to compute the *complete* QRD of U as

$$U = \widetilde{Q}_u \widetilde{C} \widetilde{P}_u^T \Pi_u^T, \tag{22}$$

where $(\widetilde{C}_{11}\ \ \widetilde{C}_{12})P_u = (\widetilde{C}\ \ 0)$, $\widetilde{C} \in \Re^{g\times g}$ is a non- singular upper-triangular matrix and $\widetilde{P}_u$ comprises the first g columns of the orthogonal matrix $P_u \in \Re^{G\times G}$. From the latter it follows that $\widehat{C} = \Pi_u \widetilde{P}_u \widetilde{C}^T/\sqrt{T}$. Note from (21) that, if the number of structural equations exceeds the number of observations in each variable – the so-called undersized sample problem –, then $\widehat{\Sigma}$ will be singular with rank $g \leqslant T < G$.

2.2. REDUNDANCES

Under the assumption that the consistency condition (17) is satisfied, factorizations (10) and (11) show that $GK - E - q$ rows of the TSEM (8) become redundant due to linear combinations (Hammarling et al., 1983). Let $\widetilde{Q}_{\widetilde{C}}$ comprises the last $GK - E - q$ columns of $\widetilde{Q}_C$ and $N = \widetilde{Q}_{\widetilde{C}}^T \widetilde{Q}_B^T = (\widetilde{N}^{(1)} \ldots \widetilde{N}^{(G)})$, where $\widetilde{N}^{(i)}$ is an $(GK - E - q) \times K$ matrix $(i = 1, \ldots, G)$. The elements of the pth row of N, denoted by $N_{p\cdot}$, can reveal a linear dependency among the equations of the TSEM $(p = 1, \ldots, GK - E - q)$. Premultiplication of the TSEM by $N_{p\cdot}$ gives

$$\sum_{i=1}^{G} \widetilde{N}_{p\cdot}^{(i)} R_{\cdot i}^{(2)} = \sum_{i=1}^{G} \widetilde{N}_{p\cdot}^{(i)} R^{(i)} \delta_i + \sum_{i=1}^{G} \widetilde{N}_{p\cdot}^{(i)} \sum_{j=1}^{g} \widehat{C}_{ij} V_{\cdot j} = 0$$

or

$$\sum_{i=1}^{G} \sum_{t=1}^{K} \widetilde{N}_{pt}^{(i)} R_{ti}^{(2)} = \sum_{i=1}^{G} \sum_{t=1}^{K} \widetilde{N}_{pt}^{(i)} \widetilde{R}_{t\cdot}^{(i)} \delta_i + \sum_{i=1}^{G} \sum_{t=1}^{K} \widetilde{N}_{pt}^{(i)} \widetilde{V}_{ti} = 0, \tag{23}$$

where $V^T = (V_{\cdot 1}^T \ldots V_{\cdot g}^T)$, $V_{\cdot j} \in \Re^K$ $(i = 1, \ldots, g)$ and

$$\widetilde{V}_{ti} = \sum_{j=1}^{g} \widehat{C}_{ij} V_{tj} \quad (i = 1, \ldots, G).$$

Assume that the μth equation of the λth transformed structural equation

$$R_{\mu\lambda}^{(2)} = \widetilde{R}_{\mu\cdot}^{(\lambda)} \delta_\lambda + \widetilde{V}_{\mu\lambda} \tag{24}$$

occurs in the linear dependency (23), that is, $\widetilde{N}_{p\mu}^{(\lambda)} \neq 0$. Writing (23) as

$$\begin{aligned}
&\sum_{\substack{i=1\\ i\neq\lambda}}^{G} \sum_{\substack{t=1\\ t\neq\mu}}^{K} \widetilde{N}_{pt}^{(i)} R_{ti}^{(2)} + \widetilde{N}_{p\mu}^{(\lambda)} R_{\mu\lambda}^{(2)} \\
&\quad = \sum_{\substack{i=1\\ i\neq\lambda}}^{G} \sum_{\substack{t=1\\ t\neq\mu}}^{K} \widetilde{N}_{pt}^{(i)} (\widetilde{R}_{t\cdot}^{(i)} \delta_i + \widetilde{V}_{ti}) + \widetilde{N}_{p\mu}^{(\lambda)} (\widetilde{R}_{\mu\cdot}^{(\lambda)} \delta_\lambda + \widetilde{V}_{\mu\lambda}),
\end{aligned}$$

it follows that

$$\frac{1}{\widetilde{N}_{p\mu}^{(\lambda)}} \sum_{\substack{i=1 \\ i\neq\lambda}}^{G} \sum_{\substack{t=1 \\ t\neq\mu}}^{K} \widetilde{N}_{pt}^{(i)} R_{ti}^{(2)} = \frac{1}{\widetilde{N}_{p\mu}^{(\lambda)}} \sum_{\substack{i=1 \\ i\neq\lambda}}^{G} \sum_{\substack{t=1 \\ t\neq\mu}}^{K} \widetilde{N}_{pt}^{(i)} (\widetilde{R}_{t\cdot}^{(i)} \delta_i + \widetilde{V}_{ti})$$

is an equivalent form of (24). Observe that, if $\widetilde{Q}_{\widetilde{C}}^T = (\widetilde{Q}_{\widetilde{C}}^{(1)} \dots \widetilde{Q}_{\widetilde{C}}^{(G)})$, then $\widetilde{N}^{(i)} = \widetilde{Q}_{\widetilde{C}}^{(i)} \widetilde{Q}_{Bi}^T$. Furthermore, if $\widehat{q}_{pi}^T$ and $\widetilde{q}_{ip}$ denote the pth row and column of $\widetilde{Q}_{\widetilde{C}}^{(i)}$ and $\widetilde{Q}_{Bi}^T$, respectively, then $\widetilde{N}_{pt}^{(i)} / \widetilde{N}_{p\mu}^{(\lambda)} = \widehat{q}_{pi}^T \widetilde{q}_{it} / \widehat{q}_{p\lambda}^T \widetilde{q}_{\lambda\mu}$.

3. Modifying the SEM

It is often desirable to modify the SEM by adding or deleting observations or variables. This might be necessary if new data become available, *old* or incorrect data are deleted from the SEM, or variables are added or deleted from structural equations. First consider the case of updating the SEM with new data. Let the additional sample information be denoted by

$$\mathsf{vec}(\check{Y}) = (I_G \otimes \check{W}) S\delta + \check{\varepsilon}, \tag{25}$$

where $\check{W}_i = \check{W} S_i = (\check{X}_i \; \check{Y}_i) \in \Re^{\check{T}\times e_i}$; $\check{X} \in \Re^{\check{T}\times K}$ is the matrix of all predetermined variables in (25); $\mathrm{E}(\check{\varepsilon}) = 0$ and $\mathrm{E}(\check{\varepsilon}\check{\varepsilon}^T) = \Sigma \otimes I_{\check{T}}$. Computing the *updated incomplete* QRD

$$\check{Q}^T \begin{pmatrix} \check{X} & \check{Y} \\ R^{(1)} & R^{(2)} \end{pmatrix} = \begin{array}{c} \begin{array}{cc} K & G \end{array} \\ \begin{pmatrix} \check{R}^{(1)} & \check{R}^{(2)} \\ 0 & \check{R}^{(3)} \end{pmatrix} \end{array} \begin{array}{l} K \\ \check{T} \end{array}, \tag{26}$$

the 3SLS estimator of the updated SEM solves

$$\operatorname*{argmin}_{\delta} V^T V \quad \text{subject to} \quad \mathsf{vec}(\check{R}^{(2)}) = (I_G \otimes \check{R}) S\delta + (\check{C} \otimes I_K) V, \tag{27}$$

where $\check{R}^{(1)}$ is upper triangular, $\check{R} = (\check{R}^{(1)} \; \check{R}^{(2)})$, and $\check{\Sigma} = \check{C}\check{C}^T$ is a new estimator of Σ. The only computational advantage of not solving the updated SEM afresh is the use of the already computed matrices $R^{(1)}$ and $R^{(2)}$ to construct the updated TSEM. The solution of (9) cannot be used to reduce the computational burden for solving (27).

Similarly, the downdating problem can be described as solving the SEM (2) after the sample information denoted by (25) has been deleted. If the original matrix W is available, then the downdated SEM can be solved afresh or the matrix that corresponds to R in (5) can be derived from downdating the *incomplete* QRD of W (Elden and Park, 1994; Gill et al., 1974; Golub and Loan, 1983; Kontoghiorghes and Clarke, 1993b; Olszanskyj et al., 1994; Paige, 1978). However, as in the updating problem, the solution of the downdated TSEM will need to be recomputed from scratch.

Assume that the additional variables denoted by $W\widehat{S}_i \in \Re^{T\times\widehat{e}_i}$ have been introduced to the ith structural equation. After computing the QRD

$$\widehat{Q}_i^T(\widetilde{Q}_{Bi}^T R\widehat{S}_i) = \begin{pmatrix} \widehat{R}^{(i)} \\ 0 \end{pmatrix} \begin{matrix} \scriptstyle \widehat{e}_i \\ \scriptstyle K-e_i-\widehat{e}_i \end{matrix}, \quad \text{with } \widehat{Q}_i = \begin{array}{cc} \scriptstyle \widehat{e}_i & \scriptstyle K-e_i-\widehat{e}_i \\ (\widehat{Q}_{Ai} & \widehat{Q}_{Bi}), \end{array}$$

the matrix computations corresponding to (13) and (15) are given respectively by

$$\check{Q}_i^T(R(S_i\ \widehat{S}_i)) = \begin{pmatrix} \widetilde{R}^{(i)} & \widetilde{Q}_{Ai}^T R\widehat{S}_i \\ 0 & \widehat{R}^{(i)} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \check{R}^{(i)} \\ 0 \end{pmatrix} \begin{matrix} \scriptstyle e_i+\widehat{e}_i \\ \scriptstyle K-e_i-\widehat{e}_i \end{matrix}$$

and

$$\begin{pmatrix} \check{Q}_{Ai}^T \\ \check{Q}_{Bi}^T \end{pmatrix} R_{\cdot i}^{(2)} = \begin{pmatrix} \widetilde{y}_i^{(1)} \\ \widehat{Q}_{Ai}^T \widehat{y}_i \\ \hline \widehat{Q}_{Bi}^T \widehat{y}_i \end{pmatrix} = \begin{pmatrix} \check{y}_i^{(1)} \\ \widehat{y}_i^* \end{pmatrix} \begin{matrix} \scriptstyle e_i+\widehat{e}_i \\ \scriptstyle K-e_i-\widehat{e}_i \end{matrix},$$

where $\check{Q}_{Bi} = \widetilde{Q}_{Bi}\widehat{Q}_{Bi}$ and $\check{Q}_{Ai} = (\widetilde{Q}_{Ai}\ \widetilde{Q}_{Bi}\widehat{Q}_{Ai})$. Computing the complete QRD of $\check{Q}_B^T(\check{C}\otimes I_K)$ as in (14) and the equivalent of (16), the 3SLS solution of the modified SEM can be found using (18), where $\widehat{Q}_B \equiv \mathsf{diag}(\widehat{Q}_{B1},\ldots,\widehat{Q}_{BG})$, $\check{Q}_B^T = \widehat{Q}_B^T\widetilde{Q}_B^T$, and, as in the updating problem, $\check{C}\check{C}^T$ is a new estimator of Σ.

Deleting the $W\widehat{S}_i$ data matrix from the ith structural equation is equivalent to re-triangularizing $\widetilde{R}^{(i)}$ by orthogonal transformations after deleting the columns $\widetilde{R}^{(i)}\widehat{S}_i$ $(i = 1,\ldots,G)$. Thus, if the new selector matrix of the ith equation is denoted by S_i^*, and the QRD of $\widetilde{R}^{(i)}S_i^*$ is given by

$$\widehat{Q}_i^T(\widetilde{R}^{(i)}S_i^*) = \begin{pmatrix} \widehat{R}^{(i)} \\ 0 \end{pmatrix} \begin{matrix} \scriptstyle e_i-\widehat{e}_i \\ \scriptstyle \widehat{e}_i \end{matrix}, \quad \text{with } \widehat{Q}_i = \begin{array}{cc} \scriptstyle e_i-\widehat{e}_i & \scriptstyle \widehat{e}_i \\ (\widehat{Q}_{Ai} & \widehat{Q}_{Bi}) \end{array} \tag{28}$$

then (14)–(16) need to be recomputed with $\widetilde{Q}_{Ai}$ and $\widetilde{Q}_{Bi}$ replaced by $\widetilde{Q}_{Ai}\widehat{Q}_{Ai}$ and $(\widetilde{Q}_{Ai}\widehat{Q}_{Bi}\ \ \widetilde{Q}_{Bi})$, respectively.

Now consider the case where *new* predetermined variables, denoted by the $T\times\widehat{K}$ matrix $\widehat{X}$, are added to the SEM. The modified SEM can be written as

$$\mathsf{vec}(Y) = (I_G\otimes\check{W})\check{S}\check{\delta} + \varepsilon,$$

where $\check{W} \equiv (X \;\; \widehat{X} \;\; Y)$, $\check{S} \equiv \mathsf{diag}(\check{S}_1, \ldots, \check{S}_G)$, $\check{S}_i$ is a $(K + G + \widehat{K}) \times (e_i + \widehat{k}_i)$ selector matrix defined as $\check{S}_i \equiv \mathsf{diag}(S_i^{(1)}, S_i^{(2)}, S_i^{(3)})$, and $S_i \equiv \mathsf{diag}(S_i^{(1)}, S_i^{(3)})$. Computing the incomplete QRD

$$\widehat{Q}^T (Q_2^T \widehat{X} \;\; R^{(3)}) = \begin{pmatrix} \widehat{R}^{(1)} & \widehat{Q}_1^T R^{(3)} \\ 0 & \widehat{Q}_2^T R^{(3)} \end{pmatrix} \begin{matrix} \scriptstyle \widehat{K} \\ \scriptstyle T-K-\widehat{K} \end{matrix},$$

$$\widehat{K} \qquad T - K - \widehat{K}$$

with $\widehat{Q} = (\widehat{Q}_1 \qquad \widehat{Q}_2)$, it follows that the modified TSEM can be written in the form

$$\mathsf{vec}(\check{R}^{(2)}) = (I_G \otimes \check{R})\check{S}\check{\delta} + (\check{C} \otimes I_{K+\widehat{K}})\check{V}, \tag{29}$$

where now $\check{V}$ and $\check{\delta}$ are $(K + \widehat{K})$ and $(E + \Sigma_{i=1}^G \widehat{k}_i)$-element vectors, respectively, and

$$\check{R}^{(1)} = \begin{pmatrix} R^{(1)} & Q_1^T \widehat{X} \\ 0 & \widehat{R}^{(1)} \end{pmatrix}, \check{R}^{(2)} = \begin{pmatrix} R^{(2)} \\ \widehat{Q}_1^T R^{(3)} \end{pmatrix} \quad \text{and} \quad \check{R} = (\check{R}^{(1)} \;\; \check{R}^{(2)}).$$

The solution of (29) can be obtained as in the original case. However, computations for forming the QRDs of $\check{R}\check{S}_i$ $(i = 1, \ldots, G)$ can be reduced significantly if both sides of (29) are premultiplied by the orthogonal matrix $(\check{Q}_A \;\; \check{Q}_B)^T$, where $\check{Q}_A \equiv \mathsf{diag}(\check{Q}_{A1}, \ldots, \check{Q}_{AG})$, $\check{Q}_B \equiv \mathsf{diag}(\check{Q}_{B1}, \ldots, \check{Q}_{BG})$, $\check{Q}_{Ai} \equiv \left(\begin{smallmatrix} \widetilde{Q}_{Ai} & 0 \\ 0 & I_{\widehat{k}_i} \end{smallmatrix}\right)$ and $\check{Q}_{Bi} \equiv (\widetilde{Q}_{Bi} \;\; 0)$ for $i = 1, \ldots, G$. In this case the upper triangular factor in the QRD of $\check{R}S_i^{(1)}$ is given by the already computed $RS_i^{(1)}$.

4. 3SLS of SEMs Subject to Linear Equality Constraints

Consider the solution of the SEM (2) with the *separable* constraints

$$H\delta = \xi, \tag{30}$$

where $H \equiv \mathsf{diag}(H_1, \ldots, H_G)$, $\xi^T \equiv (\xi_1^T \ldots \xi_G^T)$, $H_i \in \Re^{d_i \times e_i}$ has full row rank, $\xi_i \in \Re^{d_i}$, $d \equiv \Sigma_{i=1}^G d_i$, and $d_i < e_i (i = 1, \ldots, G)$. The constrained 3SLS estimator can be found from the solution of

$$\underset{\delta}{\text{argmin}}\, V^T V \;\; \text{s.t.} \;\; \begin{cases} \mathsf{vec}(R^{(2)}) = (I_G \otimes R)S\delta + (\widehat{C} \otimes I_K)V, \\ \xi = H\delta, \end{cases} \tag{31}$$

which, under the assumption that the consistency rule (17) is satisfied, can be written as

$$\underset{\delta}{\text{argmin}}\, V^T V \;\; \text{s.t.} \begin{pmatrix} \widetilde{y}^{(1)} \\ \xi \\ \widetilde{y}^{(2)} \end{pmatrix} = \begin{pmatrix} \widetilde{R} \\ H \\ 0 \end{pmatrix} \delta + \begin{pmatrix} L_{11} & L_{12} \\ 0 & 0 \\ 0 & L_{22} \end{pmatrix} \begin{pmatrix} \widetilde{V}_1 \\ \widetilde{V}_2 \end{pmatrix}. \tag{32}$$

Computing the QRD

$$\widehat{Q}_i^T \begin{pmatrix} \widetilde{R}^{(i)} \\ H_i \end{pmatrix} = \begin{pmatrix} \widehat{R}^{(i)} \\ 0 \end{pmatrix} \begin{matrix} e_i \\ d_i \end{matrix}, \quad \text{with } \widehat{Q}_i^T = \begin{array}{c} \begin{matrix} e_i & & d_i \end{matrix} \\ \begin{pmatrix} \widehat{Q}_{11}^{(i)} & \widehat{Q}_{12}^{(i)} \\ \widehat{Q}_{21}^{(i)} & \widehat{Q}_{22}^{(i)} \end{pmatrix} \end{array} \begin{matrix} e_i \\ d_i \end{matrix}, \tag{33}$$

let

$$\widehat{Q}_{11}^T = \mathsf{diag}(\widehat{Q}_{11}^{(1)}, \ldots, \widehat{Q}_{11}^{(G)}), \qquad \widehat{Q}_{12}^T = \mathsf{diag}(\widehat{Q}_{12}^{(1)}, \ldots, \widehat{Q}_{12}^{(G)}),$$
$$\widehat{Q}_{21}^T = \mathsf{diag}(\widehat{Q}_{21}^{(1)}, \ldots, \widehat{Q}_{21}^{(G)}), \qquad \widehat{Q}_{22}^T = \mathsf{diag}(\widehat{Q}_{22}^{(1)}, \ldots, \widehat{Q}_{22}^{(G)}),$$
$$\widehat{Q}^T = \begin{pmatrix} \widehat{Q}_{11}^T & \widehat{Q}_{12}^T \\ \widehat{Q}_{21}^T & \widehat{Q}_{22}^T \end{pmatrix} \quad \text{and}$$
$$\widehat{Q}^T \begin{pmatrix} \widetilde{y}^{(1)} & L_{11} & L_{12} \\ \xi & 0 & 0 \end{pmatrix} = \begin{pmatrix} \widehat{y}^{(1)} & \widetilde{L}_{11} & \widetilde{L}_{12} \\ y_2^* & \widetilde{L}_{21} & \widetilde{L}_{22} \end{pmatrix} \begin{matrix} E \\ d \end{matrix}.$$

The constrained 3SLS solution can be derived analogously to the original problem after computing the complete QRD

$$\widehat{Q}_i^T \begin{pmatrix} \widetilde{L}_{21} & \widetilde{L}_{22} \\ 0 & L_{22} \end{pmatrix} \widehat{P} = \begin{array}{c} \begin{matrix} gK - \widehat{q} & \quad \widehat{q} \end{matrix} \\ \begin{pmatrix} 0 & \widehat{L}_{22} \\ 0 & 0 \end{pmatrix} \end{array} \begin{matrix} \widehat{q} \\ d+q-\widehat{q} \end{matrix} \quad \text{and}$$
$$\widehat{Q}_C^T = \begin{pmatrix} y_2^* \\ \widehat{y}^{(2)} \end{pmatrix} = \begin{pmatrix} \widehat{y}^{(2)} \\ \widehat{y}^{(3)} \end{pmatrix}.$$

In the case where (30) allows cross-section constraints, the solution of (31) is derived as above except that, in the QRD $\widehat{Q}^T(\widetilde{R}^T \quad H^T)^T = (\widehat{R}^T \quad 0^T)^T$, the triangular matrix $\widehat{R}$ is not block diagonal.

4.1. Annihilation and direct elimination methods

The annihilation (basis-of-the-null-space) method and the direct elimination method are alternative means for solving the constrained 3SLS problem. Both methods reparameterize the constraints and solve a reduced unconstrained SEM of $E - d$ parameters (Barlow and Handy, 1988; Bjorck, 1984; Kontoghiorghes and Clarke, 1994; Lawson and Hanson, 1974; Sargan, 1988; Schittkowski and Stoer, 1979; Stoer, 1971). Consider the case of separable constraints. In the annihilation method, the coefficient vector δ_i is expressed as

$$\delta_i = \overline{Q}_{Ai}\vartheta_i + \overline{Q}_{Bi}\zeta_i, \tag{34}$$

where the QRD of H_i^T is given by

$$H_i\overline{Q}_i = \begin{array}{cc} d_i & e_i - d_i \\ (L_i & 0) \end{array}, \quad \text{with } \overline{Q}_i = \begin{array}{cc} d_i & e_i - d_i \\ (\overline{Q}_{Ai} & \overline{Q}_{Bi}) \end{array}, \tag{35}$$

$L_i \in \Re^{d_i \times d_i}$ is non-singular lower-triangular matrix, $L_i\vartheta_i = \xi_i$ (from (30)), and ζ_i is an unknown non-zero $(e_i - d_i)$-element vector. Substituting (34) into (8) gives the reduced problem

$$\mathsf{vec}(R^{(2)}) - (I_G \otimes R)S\overline{Q}_A\theta = (I_G \otimes R)S\overline{Q}_B\zeta + (\widehat{C} \otimes I_K)V, \tag{36}$$

where $\theta^T \equiv (\vartheta_1^T, \ldots, \vartheta_G^T), \zeta^T \equiv (\zeta_1^T, \ldots, \zeta_G^T), \overline{Q}_A \equiv \mathsf{diag}(\overline{Q}_{A1}, \ldots, \overline{Q}_{AG})$ and $\overline{Q}_B \equiv \mathsf{diag}(\overline{Q}_{B1}, \ldots, \overline{Q}_{BG})$. Premultiplying both sides of (36) by $\widetilde{Q}^T$ gives

$$\begin{pmatrix} \check{y} \\ \tilde{y}^{(2)} \end{pmatrix} = \begin{pmatrix} \check{R} \\ 0 \end{pmatrix} \zeta + \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix} \begin{pmatrix} \widetilde{V}_1 \\ \widetilde{V}_2 \end{pmatrix}, \tag{37}$$

where it is assumed that (17) holds, $V^TP = (\widetilde{V}_1^T \;\; \widetilde{V}_2^T), \check{y} \equiv \mathsf{vec}(\check{y}_1 \ldots \check{y}_G), \check{y}_i = \tilde{y}_i^{(1)} - \widetilde{R}^{(i)}\overline{Q}_{Ai}\vartheta_i (i = 1, \ldots, G)$, and $\check{R} \equiv \mathsf{diag}(\widetilde{R}^{(1)}\overline{Q}_{B1}, \ldots, \widetilde{R}^{(G)}\overline{Q}_{BG})$. Once the estimator of ζ_i, say $\widehat{\zeta}_i$, is derived from the solution of (37), then the constrained 3SLS estimator of δ_i can be found from (34) with ζ_i replaced by $\widehat{\zeta}_i$.

In the direct elimination method, one computes the QRD

$$\overline{Q}_i^T H_i \Pi_i = (\widehat{L}_i \;\; \widetilde{L}_i), \tag{38}$$

where Π_i is a permutation matrix and $\widetilde{L}_i \in \Re^{d_i \times d_i}$ is non-singular lower-triangular matrix. If

$$\Pi_i^T \delta_i = \begin{pmatrix} \widehat{\delta}_i \\ \widetilde{\delta}_i \end{pmatrix} \begin{array}{l} e_i - d_i \\ d_i \end{array}, \tag{39}$$

$\widetilde{L}_i\vartheta_i = \overline{Q}_i^T\xi_i$ and $\widetilde{L}_i\check{L}_i = \widehat{L}_i$, then $\widetilde{\delta}_i$ can be written as

$$\widetilde{\delta}_i = \vartheta_i - \check{L}_i\widehat{\delta}_i. \tag{40}$$

Furthermore, if $S_i\Pi_i = (\widehat{S}_i \;\; \widetilde{S}_i)$, where $\widehat{S} \equiv \mathsf{diag}(\widehat{S}_1, \ldots, \widehat{S}_G)$ and $\widetilde{S} \equiv \mathsf{diag}(\widetilde{S}_1, \ldots, \widetilde{S}_G)$, then (8) can be written equivalently as

$$\mathsf{vec}(R^{(2)}) = (I_G \otimes R)(\widehat{S} \;\; \widetilde{S}) \begin{pmatrix} \widehat{\delta} \\ \theta - \check{L}\widehat{\delta} \end{pmatrix} + (\widehat{C} \otimes I_K)V$$

or

$$\mathsf{vec}(R^{(2)}) - (I_G \otimes R)\widetilde{S}\theta = (I_G \otimes R)(\widehat{S} - \widetilde{S}\check{L})\widehat{\delta} + (\widehat{C} \otimes I_K)V,$$

where $\widehat{\delta}^T \equiv (\widehat{\delta}_1^T, \ldots, \widehat{\delta}_G^T)$ and $\check{L} \equiv \mathsf{diag}(\check{L}_1, \ldots, \check{L}_G)$. As in the annihilation method, the premultiplication of the latter by $\widetilde{Q}^T$ gives

$$\begin{pmatrix} \check{y} \\ \widetilde{y}^{(2)} \end{pmatrix} = \begin{pmatrix} \check{R} \\ 0 \end{pmatrix} \widehat{\delta} + \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix} \begin{pmatrix} \widetilde{V}_1 \\ \widetilde{V}_2 \end{pmatrix}, \tag{41}$$

where $\check{y}_i = \widetilde{y}_i^{(1)} - \check{R}_2^{(i)} \vartheta_i, \widetilde{R}^{(i)}(\widehat{S}_i \;\; \widetilde{S}_i) = (\check{R}_1^{(i)} \;\; \check{R}_2^{(i)}), \check{y} \equiv \mathsf{vec}(\widetilde{y}_1 \ldots \check{y}_G), \check{R}_i = \check{R}_1^{(i)} - \check{R}_2^{(i)} \check{L}_i$, and $\check{R} \equiv \mathsf{diag}(\check{R}_1, \ldots, \check{R}_G)$. The solution to (41) will give an estimator for $\widehat{\delta}_i$, which is then used in (40) to compute $\widetilde{\delta}_i$ $(i = 1, \ldots, G)$. Finally, the constrained 3SLS estimator of δ_i is computed by $\Pi_i(\widehat{\delta}_i^T \;\; \widetilde{\delta}_i^T)^T$. The above methods are trivially extended for the case of cross-section constraints.

5. Computational Strategies

The QRD and its modification are the main components of the approach to computing the 3SLS estimator given here. We now examine strategies for computing the factorizations (13) and (14) when $\widehat{\Sigma}$ is non-singular. Extensions of the well-known Householder transformations, Givens rotations, or (classical or modified) Gram–Schmidt orthogonalization procedure can be employed to compute the remaining factorizations. The Gram–Schmidt method will be particularly efficient if it is necessary to compute the orthogonal basis of the matrices, while Givens rotations are suitable when matrices are sparse or have a special banded structure. However, Householder transformations are the fastest means for deriving the upper-triangular factor in the QRD of large-scale dense matrices (Clint et al., 1966; Golub and Loan, 1983; Kontoghiorghes, 1993a; Lawson and Hanson, 1974).

5.1. Computing the QRDs of the transformed data matrices

The QRDs of RS_i $(i = 1, \ldots, G)$ in (13) are mutually independent and can be computed simultaneously. In a *Multiple Instruction stream–Multiple Data stream* (MIMD) multiprocessor system, each processing unit will compute one or more factorizations (Kontoghiorghes and Dinenis, 1996a). Furthermore, the particular structure of RS_i can be exploited to reduce the computational burden of the QRD (13). If $\widehat{S}_i$ comprises the first k_i columns of S_i, then the QRD of $R\widehat{S}_i$ is a re-triangularization of an upper-triangular matrix after deleting columns by orthogonal transformations. Let $S_i \equiv (\mathbf{e}_{\lambda_{i,1}} \ldots \mathbf{e}_{\lambda_{i,e_i}})$ and define the e_i-element integer vector $\sigma_i = (\lambda_{i,1} \ldots \lambda_{i,k_i} K \ldots K)$, where $\mathbf{e}_{\lambda_{i,j}}$ $(j = 1, \ldots, e_i)$ is the λ_jth column of I_{K+G} and $\lambda_{i,1} < \cdots < \lambda_{i,e_i}$. Figure 1 shows a Givens annihilation scheme for computing the QRD (13), where $\sigma_i = (3, 6, 7, 9, 12, 15, 15, 15)$ and $g_i = 3$. A number j $(1, \ldots, 46)$, a blank and a $\bullet$ denote the element annihilated by the jth rotation, a zero element and a non-zero element, respectively. An element of RS_i at position (l, q) is zeroed by rotating the lth and $(l-1)$th adjacent rows. Generally, the total number of rotations applied to compute the QRD (13) for $i = 1, \ldots, G$ is given by

$$T_1(\sigma_i, k_i, g_i, G) = \sum_{i=1}^{G} \sum_{j=1}^{g_i+k_i} (\sigma_{ij} - j)$$

$$= \sum_{i=1}^{G} \left(\left(\sum_{j=1}^{k_i} \sigma_{ij} - k_i(k_i+1)/2 \right) + g_i(2K - 2k_i - g_i - 1)/2 \right). \qquad (42)$$

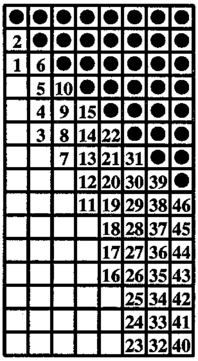

Figure 1. Serial Givens annihilation scheme.

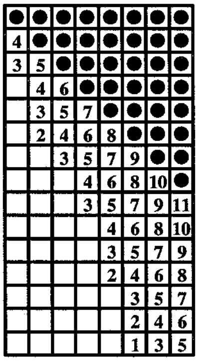

Figure 2. Parallel Givens annihilation scheme.

Notice that (42) gives the maximum number of rotations for computing the QRDs (13). This number can be possibly reduced by exploiting the structure of the matrices RS_i $(i = 1, \ldots, G)$, which depends on the specific characteristics of the SEM. In order to illustrate this, consider the case where $RS_i = (R_1^{(i)} \;\; R_2^{(i)})$ and $RS_j = (R_1^{(i)} \;\; R_2^{(j)})$ for some $i \neq j$. Conformally partitioning $\widetilde{R}^{(i)}$ as

$$\widetilde{R}^{(i)} = (\widetilde{R}_1^{(i)} \;\; \widetilde{R}_2^{(i)}) = \begin{pmatrix} \widetilde{R}_{11}^{(i)} & \widetilde{R}_{21}^{(i)} \\ 0 & \widetilde{R}_{22}^{(i)} \end{pmatrix},$$

it follows that $\widetilde{R}^{(j)}$ can be derived from the QRD

$$\begin{pmatrix} I & 0 \\ 0 & \widehat{Q}_j^T \end{pmatrix} \widetilde{Q}_i^T RS_j$$

$$= \begin{pmatrix} I & 0 \\ 0 & \widehat{Q}_j^T \end{pmatrix} \begin{pmatrix} \widetilde{R}_{11}^{(i)} & \widehat{R}_{21}^{(j)} \\ 0 & \widehat{R}_{22}^{(j)} \end{pmatrix} = \begin{pmatrix} \widetilde{R}_{11}^{(i)} & \widehat{R}_{21}^{(j)} \\ 0 & \widehat{Q}_j^T \widehat{R}_{22}^{(j)} \end{pmatrix}$$

$$= \begin{pmatrix} \widetilde{R}_{11}^{(i)} & \widehat{R}_{21}^{(j)} \\ 0 & \widetilde{R}_{22}^{(j)} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \widetilde{R}^{(j)} \\ 0 \end{pmatrix},$$

where $((\widehat{R}_{21}^{(j)})^T \ (\widehat{R}_{22}^{(j)})^T)^T = \widetilde{Q}_i^T R_2^{(j)}$. Thus, the number of rotations for computing the QRD of RS_j is determined by the triangularization of the smaller submatrix $\widehat{R}_{22}^{(j)}$.

A parallelisation of the above strategy is shown in Figure 2. In this case, a number j $(1, \ldots, 11)$ denotes the elements annihilated by the jth compound disjoint Givens rotation (CDGR). A CDGR is a product of Givens rotations that can be applied simultaneously (Kontoghiorghes, 1995; Kontoghiorghes and Clarke, 1993b; Kontoghiorghes and Clarke, 1995b). Notice that elements in the sub-diagonals are annihilated by successive CDGRs.

Using this annihilation scheme, the total number of CDGRs applied to compute the QRDs of RS_i $(i = 1, \ldots, G)$ simultaneously, is given by

$$T_3(\sigma_i, e_i, G) = \max(T_2(\sigma_i, e_i)), \quad (i = 1, \ldots, G), \tag{43}$$

where, for $\lambda_{i,j} \neq j$,

$$T_2(\sigma_i, e_i) = e_i - \min(2j - \sigma_{ij}), \quad (j = 1, \ldots, e_i), \tag{44}$$

is the total number of CDGRs required to compute the QRD of RS_i. A detailed description of a similar parallel Givens sequence is found in Kontoghiorghes and Clarke (1993a) and a *bitonic* algorithm which could possibly be adapted to compute the QRD (13) using fewer CDGRs is discussed in Kontoghiorghes (1995).

5.2. Computing the QRD (14) when $\widehat{\Sigma}$ is non-singular

Consider the computation of the orthogonal factorization (14) when $\widehat{C}$ is non-singular; that is, $\widetilde{Q}_C = I_{GK-E}, g = G, q = GK - E$, and

$$P^T((\widehat{C}^T \otimes I_K)\widetilde{Q}_B) = \begin{pmatrix} 0 \\ L_{22}^T \end{pmatrix}\begin{matrix} {\scriptstyle E} \\ {\scriptstyle GK-E} \end{matrix}. \tag{45}$$

Let P be defined as $P = (\widetilde{Q}_A \ \ \widetilde{Q}_B)\widetilde{P}$ such that

$$\begin{pmatrix}\widetilde{Q}_A^T\\ \widetilde{Q}_B^T\end{pmatrix}(\widehat{C}^T\otimes I_K)\widetilde{Q}_B$$

$$= \begin{array}{cc} \begin{matrix} K-e_1 & K-e_2 & K-e_3 & & K-e_G \end{matrix} & \\ \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ \widetilde{A}_{21}^{(1)} & 0 & 0 & \dots & 0 \\ \widetilde{A}_{31}^{(1)} & \widetilde{A}_{32}^{(1)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{A}_{G1}^{(1)} & \widetilde{A}_{G2}^{(1)} & \widetilde{A}_{G3}^{(1)} & \dots & 0 \\ \widehat{L}_1^{(1)} & 0 & 0 & \dots & 0 \\ \widehat{A}_{21}^{(1)} & \widehat{L}_2^{(1)} & 0 & \dots & 0 \\ \widehat{A}_{31}^{(1)} & \widehat{A}_{32}^{(1)} & \widehat{L}_3^{(1)} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{A}_{G1}^{(1)} & \widehat{A}_{G2}^{(1)} & \widehat{A}_{G3}^{(1)} & \dots & \widehat{L}_G^{(1)} \end{pmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \\ \\ e_G \\ K-e_1 \\ K-e_2 \\ K-e_3 \\ \\ K-e_G \end{matrix} \end{array}, \qquad (46)$$

where $\widetilde{A}_{ij}^{(1)} = \widehat{C}_{ji}\widetilde{Q}_{Ai}^T\widetilde{Q}_{Bj}$ and $\widehat{A}_{ij}^{(1)} = \widehat{C}_{ji}\widetilde{Q}_{Bi}^T\widetilde{Q}_{Bj}$ for $i > j$, and $\widehat{L}_i^{(1)} = \widehat{C}_{ii}I_{k-e_i}$ $(i = 1,\dots,G)$. The matrix $\widetilde{P}$ is the product of orthogonal matrices that reduce (46) to lower-triangular form. Let $\widetilde{P}_{ij}^T$ $(i = 1,\dots,G-1$ and $j = 1,\dots,G-i)$ be orthogonal such that

$$\widetilde{P}_{ij}^T\begin{pmatrix}\widetilde{A}_{i+j1}^{(i)} & \dots & \widetilde{A}_{i+jj}^{(i)}\\ \widehat{A}_{j1}^{(i)} & \dots & \widehat{L}_j^{(i)}\end{pmatrix} = \begin{array}{cc}\begin{pmatrix}\widetilde{A}_{i+j1}^{(i+1)} & \dots & 0\\ \widehat{A}_{j1}^{(i+1)} & \dots & \widehat{L}_j^{(i+1)}\end{pmatrix} & \begin{matrix} e_{i+j} \\ K-e_j \end{matrix}\end{array}, \qquad (47)$$

where the $\widehat{L}_j^{(i+1)}$ matrix is lower-triangular. From (47) it follows that the triangular matrix L_{22}^T in (45) is given by

$$L_{22}^T = \begin{pmatrix}\widehat{L}_1^{(G)} & 0 & \dots & 0\\ \widehat{A}_{21}^{(G-1)} & \widehat{L}_2^{(G-1)} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ \widehat{A}_{G1}^{(1)} & \widehat{A}_{G2}^{(1)} & \dots & \widehat{L}_G^{(1)}\end{pmatrix}.$$

The orthogonal matrix $\widetilde{P}^T$ is defined as

$$\widetilde{P}^T = \prod_{i=1}^{G-1}\prod_{j=1}^{G-i}\widehat{P}_{ij}^T = \widehat{P}_{G-11}^T\widehat{P}_{G-22}^T\widehat{P}_{G-21}^T\dots\widehat{P}_{1G-1}^T\dots\widehat{P}_{12}^T\widehat{P}_{11}^T,$$

where the orthogonal matrices $\widehat{P}_{i1}^T, \ldots, \widehat{P}_{iG-i}$ $(i = 1, \ldots, G-1)$ are disjoint and can be applied simultaneously. Partitioning $\widetilde{P}_{ij}^T$ as

$$\widetilde{P}_{ij}^T = \begin{array}{cc} \begin{array}{cc} e_{i+j} & K-e_j \end{array} & \\ \begin{pmatrix} \widetilde{P}_{11}^{(ij)} & \widetilde{P}_{12}^{(ij)} \\ \widetilde{P}_{21}^{(ij)} & \widetilde{P}_{22}^{(ij)} \end{pmatrix} & \begin{array}{l} e_{i+j} \\ K-e_j \end{array} \end{array}, \tag{48}$$

the $GK \times GK$ matrix $\widehat{P}_{ij}^T$ and the product of this disjoint orthogonal matrices (PDOM) $\widetilde{P}^{(i)} = \Pi_{j=1}^{G-i} \widehat{P}_{ij}^T$ are given, respectively, by

$$\widehat{P}_{ij}^T = \begin{pmatrix} I & 0 & 0 & 0 & 0 \\ 0 & \widetilde{P}_{11}^{(ij)} & 0 & \widetilde{P}_{12}^{(ij)} & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & \widetilde{P}_{21}^{(ij)} & 0 & \widetilde{P}_{22}^{(ij)} & 0 \\ 0 & 0 & 0 & 0 & I \end{pmatrix} \tag{49}$$

and

$$\widetilde{P}^{(i)} = \left(\begin{array}{ccccc|ccccc} I_{\widetilde{E}_i} & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \widetilde{P}_{11}^{(i1)} & 0 & \ldots & 0 & \widetilde{P}_{12}^{(i1)} & 0 & \ldots & 0 & 0 \\ 0 & 0 & \widetilde{P}_{11}^{(i2)} & \ldots & 0 & 0 & \widetilde{P}_{12}^{(i2)} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \widetilde{P}_{11}^{(iG-i)} & 0 & 0 & \ldots & \widetilde{P}_{12}^{(iG-i)} & 0 \\ \hline 0 & \widetilde{P}_{21}^{(i1)} & 0 & \ldots & 0 & \widetilde{P}_{22}^{(i1)} & 0 & \ldots & 0 & 0 \\ 0 & 0 & \widetilde{P}_{21}^{(i2)} & \ldots & 0 & 0 & \widetilde{P}_{22}^{(i2)} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \widetilde{P}_{21}^{(iG-i)} & 0 & 0 & \ldots & \widetilde{P}_{22}^{(iG-1)} & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & I_{\widehat{E}_i} \end{array}\right), \tag{50}$$

where $\widetilde{E}_i = \Sigma_{p=1}^i e_p$, and $\widehat{E}_i = iK - \Sigma_{p=1}^i e_{G+1-p}$. Notice that (48) and (50) are block generalizations of a single Givens rotation and a CDGR, respectively. That is, if $\forall i \quad K - e_i = e_i = 1$, then (48) is identical to the single Givens rotation $\left(\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right)$ (Kontoghiorghes, 1993a; Kontoghiorghes, 1995; Kontoghiorghes and Clarke, 1993b; Kontoghiorghes and Clarke, 1995b).

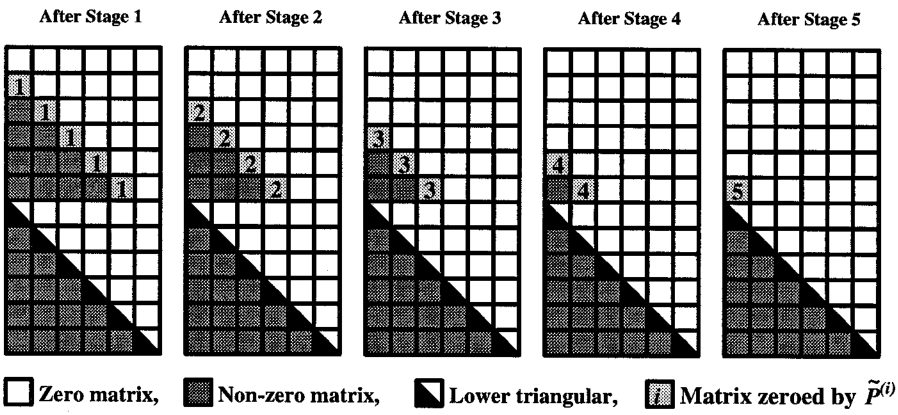

Figure 3. Parallel reduction of (46) into lower triangular form.

Figure 3 illustrates the process of reducing (46) to triangular form, using the block-parallel strategy described above. Now a number i $(1, \ldots, 5)$ denotes the blocks annihilated simultaneously from the application of $\widetilde{P}^{(i)}$. The total number of steps required to compute the QRD (46) is given by

$$T_5(G, K, e) = \sum_{i=1}^{G-1} \max(T_4(K, e, i, j)), \quad (j = 1, \ldots, G-i), \tag{51}$$

where $T_4(K, e, i, j)$ and $\max(T_4(K, e, i, j))$ are the number of steps needed to compute the orthogonal factorization (47) and apply the PODM (50), respectively, and $e = (e_1, \ldots, e_G)$.

6. Conclusions and Future Work

Our purpose has been to present computational methods for solving the SEM without requiring the non-singularity of the disturbance covariance matrix. This is achieved by considering the transformed SEM as a generalized linear least-squares problem. This formulation facilitates the use of orthogonal factorisations for deriving the 3SLS estimator and its dispersion matrix. Computational strategies for solving the main two factorisations (13) and (14) when $\widehat{\Sigma}$ is non-singular have been presented. These are by no means the most efficient strategies, but they show the general principles to be adopted for solving the SEM on a serial or parallel computer. Actual implementation depends on the number of processing units available and the problem's dimensions. For example, the block-parallel algorithm described in Figure 3 will be inefficient if there is a $p > 1$ such that $e_p \ggg e_j$ $(j = 1, \ldots, G-1)$. This is because the steps of the first $p-1$ stages will be given by $T_4(K, e, l, p-l)$, which by assumption is of higher order than that of $T_4(K, e, l, j)$, where $l = 1, \ldots, p-1,\ j = 1, \ldots, G-l$ and $j \neq (p-l)$.

The algorithms for solving the modified SEM fail to exploit the calculations used to solve (9) fully, because of the need to remove the endogeneity of the new data or to re-estimate $\widehat{\Sigma} = \widehat{C}\widehat{C}^T$. In contrast, the algorithms for solving the SEM with separable linear equality constraints can be seen as a continuation of the unconstrained problem with minor differences. Further research is needed to investigate the computational merits of these algorithms and propose efficient strategies for solving the SEM with combinations of separable and cross-section constraints.

Expressions that reveal linear combinations between the observations that become redundant are derived. These combinations can provide further insights into the properties of the particular SEM. However, inconsistencies in the SEM have not been considered. Unlike the case of inconsistent SURE models, the setting of $\widetilde{y}^{(3)}$ to zero will result in incompatibilities in the specification of the SEM (Hammarling et al., 1983; Kontoghiorghes and Dinenis, 1996b). In order to illustrate this, assume for simplicity that $\widetilde{y}^{(3)} = N\,\mathsf{vec}(R^{(2)})$ is uniformly non-zero, where $N = \widetilde{Q}^T_{\widetilde{C}}\widetilde{Q}^T_B$, using the notation of subsection 2.2. Premultiplying the TSEM (8) by the idempotent matrix $D = (I_{GK-E-q} - N^TN)$ gives

$$\mathsf{vec}(Q_1^TY) - N^T\widetilde{y}^{(3)} = (I_G \otimes R)S\delta + (\widehat{C} \otimes I_K)V,$$

from which it can be observed that

$$\widetilde{Q}^T(\mathsf{vec}(Q_1^TY) - N^T\widetilde{y}^{(3)}) = \begin{pmatrix} \widetilde{y}^{(1)} \\ \widetilde{y}^{(2)} \\ \widetilde{y}^{(3)} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ \widetilde{y}^{(3)} \end{pmatrix} = \begin{pmatrix} \widetilde{y}^{(1)} \\ \widetilde{y}^{(2)} \\ 0 \end{pmatrix}.$$

If $\mathsf{vec}(Q_1^T\widehat{Y})$ denotes the modified vector $\mathsf{vec}(Q_1^TY)$ in the TSEM such that $\mathsf{vec}(Q_1^T\widehat{Y}) = \mathsf{vec}(Q_1^TY) - N^T\widetilde{y}^{(3)}$, then

$$\mathsf{vec}(\widehat{Y}) = \widetilde{D}\,\mathsf{vec}(Y) + \mathsf{vec}(Q_2\Gamma),$$

where Γ is a random $(T-K)\times G$ matrix and $\widetilde{D} = (I_G \otimes Q_1)D(I_G \otimes Q_1^T)$. Therefore, the premultiplication of (2) by $\widetilde{D}$ will give the consistent modified SEM

$$\widetilde{D}\,\mathsf{vec}(Y) = \widetilde{D}(I_G \otimes W)S\delta + \widetilde{D}\varepsilon \qquad (52)$$

or, equivalently,

$$\mathsf{vec}(Q_1Q_1^T\widehat{Y}) = (I_G \otimes \widehat{W})S\delta + \widetilde{D}\varepsilon, \qquad (53)$$

where $\widehat{W} = (X \;\; Y - Q_2R^{(3)})$. Thus, the above modified model is incompatible with the specification of the original SEM, since the replacement of Y by $Q_1Q_1^T\widehat{Y}$

contradicts the replacement of Y by $Y - Q_2R^{(3)}$ in W. Further research is needed to define the modification in the endogenous matrix Y that yields a consistent and correctly specified model.

Currently under investigation are the solutions of a SEM with linear inequality constraints and block-recursive SEMs, and the development of optimum parallel strategies for computing (13) and (14).

Acknowledgements

The authors are grateful to David Belsley for helpful and constructive comments.

References

Anderson, E., Bai, Z. and Dongarra, J. (1992). Generalized QR factorization and its applications, *Linear Algebra and its Applications*, **162**, 243–271.

Barlow, J. and Handy, S. (1988). The direct solution of weighted and equality constrained least-squares problems, *SIAM Journal of Scientific and Statistical Computing*, **9**(4), 704–716.

Belsley, D. (1992). Paring 3SLS calculations down to manageable proportions, *Computer Science in Economics and Management*, **5**, 157–169.

Bjorck, A. (1984). A general updating algorithm for constrained linear least squares problems, *SIAM Journal of Scientific and Statistical Computing*, **5**(2), 394–402.

Clint, M., Kontoghiorghes, E. and Weston, J. (1996). Parallel Gram–Schmidt orthogonalisation and QR factorisation on an array processor, *Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM)*, **76**(S1), 377–378.

Court, R. (1974). Three stage least squares and some extensions where the structural disturbance covariance matrix may be singular, *Econometrica*, **42**(3), 547–558.

Dent, W. (1976). Information and computation in simultaneous equation estimation, *Journal of Econometrics*, **4**, 89–95.

Elden, L. and Park, H. (1994). Block downdating of least squares solutions, *SIAM Journal of Matrix Analysis and Applications*, **15**(3), 1018–1034.

Gill, P., Golub, G., Murray, W. and Saunders, M. (1974). Methods for modifying matrix factorizations, *Mathematics of Computation*, **28**(126), 505–535.

Golub, G. and Loan, C.V. (1983). *Matrix Computations*. North Oxford Academic.

Hammarling, S., Long, E. and Martin, P. (1983). *A Generalized Linear Least Squares Algorithm for Correlated Observations, with Special Reference to Degenerate Data*. DITC 33/83, National Physical Laboratory.

Jennings, L. (1980). Simultaneous equations estimation (computational aspects), *Journal of Econometrics*, **12**, 23–39.

Judge, G., Griffiths, W., Hill, R., Lütkepohl, H. and Lee, T. (1985). *The Theory and Practise of Econometrics*. Wiley series in probability and mathematical statistics. John Wiley and Sons, second edition.

Kontoghiorghes, E. (1993a). *Algorithms for Linear Model Estimation on Massively Parallel Systems*. PhD Thesis, University of London. (Also Technical report TR-655, Department of Computer Science, Queen Mary and Westfield College, University of London).

Kontoghiorghes, E. (1993b). *Solving Seemingly Unrelated Regression Equations Models using Orthogonal Decompositions*. Technical Report TR-631, Department of Computer Science, Queen Mary and Westfield College, University of London.

Kontoghiorghes, E. (1995). New parallel strategies for block updating the QR decomposition, *Parallel Algorithms and Applications*, **5**$(1 + 2)$, 229–239.

Kontoghiorghes, E. and Clarke, M. (1993a). Parallel reorthogonalization of the QR decomposition after deleting columns, *Parallel Computing*, **19**(6), 703–707.

Kontoghiorghes, E. and Clarke, M. (1993b). Solving the updated and downdated ordinary linear model on massively parallel SIMD systems, *Parallel Algorithms and Applications*, **1**(2), 243–252.

Kontoghiorghes, E. and Clarke, M. (1994). A parallel algorithm for repeated processing estimation of linear models with equality constraints. In: G. Joubert, D. Trystram, and F. Peters (eds): *Parallel Computing: Trends and Applications*, pp. 525–528.

Kontoghiorghes, E. and Clarke, M. (1995a). An alternative approach for the numerical solution of seemingly unrelated regression equation models, *Computational Statistics & Data Analysis*, **19**(4), 369–377.

Kontoghiorghes, E. and Clarke, M. (1995b). Solving the general linear model on a SIMD array processor, *Computers and Artificial Intelligence*, **14**(4), 353–370.

Kontoghiorghes, E. and Dinenis, E. (1996a). Data parallel QR decompositions of a set of equal size matrices used in SURE model estimation, *Journal of Mathematical Modelling and Scientific Computing*, **6**. (In press).

Kontoghiorghes, E. and Dinenis, E. (1996b). Solving triangular seemingly unrelated regression equation models on massively parallel systems. In: M. Gilli (ed.): *Computational Economic Systems: Models, Methods & Econometrics*, Vol. 5 of *Advances in Computational Economics*. pp. 191–201.

Kourouklis, S. and Paige, C. (1981). A constrained least squares approach to the general Gauss–Markov linear model, *Journal of the American Statistical Association*, **76**(375), 620–625.

Lawson, C. and Hanson, R. (1974). *Solving Least Squares Problems.* Prentice-Hall Englewood Cliffs.

De Moor, B. and Van Dooren, P. (1992). Generalizations of the singular value and QR decompositions, *SIAM Journal of Matrix Analysis and Applications*, **13**(4), 993–1014.

Narayanan, R. (1969). Computation of Zellner–Theil's three stage least squares estimates, *Econometrica*, **37**(2), 298–306.

Olszanskyj, S., Lebak, J. and Bojanczyk, A. (1994). Rank-k modification methods for recursive least squares problems, *Numerical Algorithms*, **7**, 325–354.

Paige, C. (1978). Numerically stable computations for general univariate linear models, *Communications on Statistical and Simulation Computation*, **7**(5), 437–453.

Paige, C. (1979a). Computer solution and perturbation analysis of generalized linear least squares problems, *Mathematics of Computation*, **33**(145), 171–183.

Paige, C. (1979b). Fast numerically stable computations for generalized linear least squares problems, *SIAM Journal of Numerical Analysis*, **16**(1), 165–171.

Paige, C. (1990). Some aspects of generalized QR factorization. In: M. Cox and S. Hammarling (eds): *Reliable Numerical Computations*.

Pollock, D. (1979). *The Algebra of Econometrics*, Wiley series in probability and mathematical statistics. John Wiley and Sons.

Sargan, D. (1988). *Lectures on Advanced Econometric Theory.* Basil Blackwell Inc.

Schittkowski, K. and Stoer, J. (1979). A factorization method for the solution of constrained linear least squares problems allowing subsequent data changes, *Numerische Mathematik*, **31**, 431–463.

Srivastava, V. and Tiwari, R. (1978). Efficiency of two-stage and three-stage least squares estimators, *Econometrica*, **46**(6), 1495–1498.

Stoer, J. (1971). On the numerical solution of constrained least-squares problems, *SIAM Journal of Numerical Analysis*, **8**(2), 382–411.

Theil, H. (1971). *Principles of Econometrics*. John Wiley & Sons, Inc.

Zellner, A. and Theil, H. (1962). Three-stage least squares: Simultaneous estimation of simultaneous equations, *Econometrica*, **30**(1), 54–78.