subfaculteit der econometrie

RESEARCH MEMORANDUM

TILBURG UNIVERSITY
DEPARTMENT OF ECONOMICS
Postbus 90153 - 5000 LE Tilburg
Netherlands

FEW
154

Statistically and Computationally Efficient Estimation of the Gravity Model

Tom Wansbeek*)

and

Arie Kapteyn**)

*) Groningen University, Econometric Institute, P.O. Box 800, 9700 AV Groningen, The Netherlands.

**) Tilburg University, Department of Econometrics, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

Abstract

Least squares estimation of the gravity model in log-linear form is, under certain assumptions, statistically efficient. Computational efficiency (essential when the number of nodes distinguished is large) is obtained by estimation via the so-called covariance transformation. This transformation can only be applied when flows are observed between all nodes. It therefore breaks down in the generic case where there are no flows observed from node i to itself. In the paper, an alternative transformation is presented that allows for efficient computation of the estimates in the generic case.

1. Introduction

In this paper we present a computationally efficient calibration method[1)] for the widely used gravity model. In log-linear form the model can be written as follows:

$$(1.1) \qquad y_{ij} = \alpha_i + \gamma_j + \beta\, d_{ij} + \varepsilon_{ij}\,, \qquad i,j = 1,\ldots,R,$$

where y_{ij} is the logarithm of the traffic flow from node i to node j; α_i and γ_j are generation and attraction factors respectively, specific to nodes i and j; d_{ij} is some measure of the distance between i and j, ε_{ij} is a random disturbance term, which we assume to be normally distributed with mean zero and constant variance, σ^2, for all i and j. The parameters α_i, γ_j and β are unknown and have to be estimated.

Depending upon the particular context to which the model is applied, a different operationalisation of d_{ij} will be used. A fairly general specification would be to replace $\beta\, d_{ij}$ by $\Sigma_{k=1}^{K}\beta_k x_{ijk}$, where the β_k are parameters and x_{ijk} are different measures of the distance between i and j. This transforms model (1.1) into

$$(1.2) \qquad y_{ij} = \alpha_i + \gamma_j + \sum_{k=1}^{K} \beta_k x_{ijk} + \varepsilon_{ij}\,, \qquad i,j = 1,\ldots,R$$

If we would have observations on y_{ij} and x_{ijk} for all i, j and k, estimation of (1.2) by means of ordinary least squares would be statistically efficient, i.e. satisfy the Gauss-Markov theorem. Although the number of parameters may be large in practice, namely 2R + K, use of the well-known covariance transformation reduces the computational burden of calibration to manageable proportions. For later reference, we give these computationally efficient formulas for the estimates of α_i, γ_j and β_k in a Lemma.

Define:

1) We will use the terms "calibration" and "estimation" interchangeably, but the latter more frequently.

$$(1.3) \qquad x_{i.k} \equiv \frac{1}{R}\sum_j x_{ijk},\; x_{.jk} \equiv \frac{1}{R}\sum_i x_{ijk},\; x_{..k} \equiv \frac{1}{R^2}\sum_i\sum_j x_{ijk}\;, \qquad k = 1,\ldots,K$$

and $y_{i.}$, $y_{.j}$, $y_{..}$ analogously. Next define

$$(1.4) \qquad \bar{x}_{ijk} \equiv x_{ijk} - x_{i.k} - x_{.jk} + x_{k..}\;, \qquad k = 1,\ldots,K\;;$$

$$\bar{y}_{ij} \equiv y_{ij} - y_{i.} - y_{.j} + y_{..}\;.$$

Furthermore, let b_k be the OLS-estimate of β_k, $k = 1,\ldots,K$, in the transformed model

$$(1.5) \qquad \bar{y}_{ij} = \sum_{k=1}^{K} \beta_k\, \bar{x}_{ijk} + \bar{\varepsilon}_{ij}\;,$$

where $\bar{\varepsilon}_{ij} \equiv \varepsilon_{ij} - \varepsilon_{i.} - \varepsilon_{.j} + \varepsilon_{..}$, in obvious notation.

Lemma. Least squares estimates of β_k, α_i and γ_j in model (1.2) are b_k, a_i and c_j, respectively, where

$$(1.6) \qquad a_i \equiv (y_{i.}-y_{..}) - \sum_k (x_{i.k}-x_{..k})b_k$$

$$(1.7) \qquad c_j \equiv (y_{.j}-y_{..}) - \sum_k (x_{.jk}-x_{..k})b_k + y_{..} - \sum_k x_{..k}b_k\;.$$

The estimates of β_k are unique, those of α_i and γ_j are unique up to an additive constant.

Proof: e.g. Scheffé (1959), Cesario (1975), Judge et al. (1980, Ch. 8).

The non-uniqueness of the estimates of α_i and γ_j is already clear from model (1.2) where we can add a constant to all α_i and subtract the same constant from all γ_j without affecting the y_{ij}. This lack of uniqueness can be used to obtain a slightly more elegant expression for c_j. Notice that the a_i sum to zero. We can introduce a constant term β_0 in (1.2) and impose the restriction that both the a_i and the c_j

add up to zero. In that case the last two terms on the right hand side of (1.7) drop out, giving (1.6) and (1.7) a similar structure. The estimate of β_0 is then

$$(1.8) \qquad b_0 = y_{..} - \sum_k x_{..k} b_k \, .$$

The obvious advantage of the least squares solution here is its computational simplicity. Rather than having to invert[1] a matrix of order (2R+K) × (2R+K), when applying OLS to (1.2) directly, we can carry out a simple transformation of the data, first estimate the β_k which requires inversion of a K×K-matrix only and then compute the estimates of the α_i and γ_j straightforwardly.

A typical characteristic of traffic flow data is that there will be no observations for i = j, i.e. one does not observe a flow from node i to itself. As a result, the lemma cannot be applied. Although one could still try to apply least squares to (1.2) directly, this will turn out to be non-feasible for many realistic problems where R may large (we return to this in section 3). Hence, we require results similar to those presented in the lemma, for the model where observations are missing whenever i = j. These results are the main objective of this paper and are collected in the theorem below.

Define:

$$(1.9) \qquad \tilde{x}_{i.k} \equiv \frac{1}{R-1} \sum_j x_{ijk}, \quad \tilde{x}_{.jk} \equiv \frac{1}{R-1} \sum_i x_{ijk}, \quad \tilde{x}_{..k} \equiv \frac{1}{R(R-1)} \sum_i \sum_j x_{ijk} \, , \qquad k = 1,\ldots,K,$$

and $\tilde{y}_{i.}$, $\tilde{y}_{.j}$, $\tilde{y}_{..}$, $\tilde{\varepsilon}_{i.}$, $\tilde{\varepsilon}_{.j}$, $\tilde{\varepsilon}_{..}$ analogously. To appreciate the definitions, notice that the total number of observations is now R(R-1) instead of R^2 and that for given k and i (and for given k and j), there are R-1 observations x_{ijk}. Next define

$$(1.10) \qquad \tilde{x}_{ijk} \equiv x_{ijk} - \frac{(R-1)^2}{R(R-2)} \tilde{x}_{i.k} - \frac{(R-1)^2}{R(R-2)} \tilde{x}_{.jk} - \frac{R-1}{R(R-2)} \tilde{x}_{j.k} -$$

1) Since this matrix has at most rank 2R + K-1, a generalized inverse has to he taken, of course.

$$- \frac{R-1}{R(R-2)} \tilde{x}_{.ik} + \frac{R}{R-2} \tilde{x}_{..k} \quad , \quad k = 1,\ldots,K$$

and $\tilde{y}_{ij}$ and $\tilde{\varepsilon}_{ij}$ analogously. Note that for large R, $\tilde{x}_{ijk}$ will be close to $\bar{x}_{ijk}$ defined by (1.4). Furthermore, let $\hat{b}_k$ be the OLS-estimate of β_k, k = 1,...,K, in the transformed model

$$(1.11) \qquad \tilde{y}_{ij} = \sum_{k=1}^{K} \beta_k \tilde{x}_{ijk} + \tilde{\varepsilon}_{ij} ,$$

then we define $\tilde{\Omega}$ as $\tilde{\Omega} = \sigma^2(\tilde{X}'\tilde{X})^{-1}$, where $\tilde{X}$ is the R(R-1) × K matrix with typical element $\tilde{x}_{ijk}$ and denote a typical element of $\tilde{\Omega}$ by $\tilde{\omega}_{k\ell}$. The quantity $\tilde{\sigma}^2$ is defined as the "usual" best quadratic unbiased estimator of $\mathrm{var}(\tilde{\varepsilon}_{ij})$ in this model, and

$$(1.12) \qquad r_{ij} \equiv y_{ij} - \sum_k \hat{b}_k x_{ijk}$$

$$(1.13) \qquad r_{i.} \equiv \frac{1}{R-1} \sum_j r_{ij}, \quad r \equiv \frac{1}{R-1} \sum_i r_{ij}, \quad r_{..} \equiv \frac{1}{R(R-1)} \sum_i \sum_j r_{ij} .$$

In addition we define:

$$(1.14) \qquad q_{ik} \equiv \frac{(R-1)^2}{R(R-2)} x_{i.k} + \frac{R-1}{R(R-2)} x_{.ik} - \frac{1}{R-2} x_{..k}$$

$$(1.15) \qquad s_{jk} \equiv \frac{R-1}{R(R-2)} x_{j.k} + \frac{(R-1)^2}{R(R-2)} x_{.jk} - \frac{R-1}{R-2} x_{..k} .$$

<u>Theorem</u>. <u>Let in model</u> (1.2) <u>observations be missing if and only if</u> i = j. <u>Least squares estimates of</u> β_k, α_i <u>and</u> γ_j <u>are</u> $\hat{b}_k$, $\hat{a}_i$ and $\hat{c}_j$, <u>where</u>

$$(1.16) \qquad \hat{a}_i \equiv \frac{(R-1)^2}{R(R-2)} r_{i.} + \frac{R-1}{R(R-2)} r_{.i} - \frac{1}{R-2} r_{..}$$

$$(1.17) \qquad \hat{c}_j \equiv \frac{R-1}{R(R-2)} r_{j.} + \frac{(R-1)^2}{R(R-2)} r_{.j} - \frac{R-1}{R-2} r_{..} .$$

<u>The estimates of</u> β_k <u>are unique, those of</u> α_i <u>and</u> γ_j <u>are unique up to an additive constant. The variance covariance matrix of the estimates</u> $\hat{b}_k$ <u>is</u> $\tilde{\Omega}$. <u>The best quadratic unbiased estimator of</u> σ^2 <u>is</u>

$$(1.18) \qquad \hat{\sigma}^2 \equiv \tilde{\sigma}^2 \frac{R(R-1)-K}{R(R-1)-(2R+K-1)} \;.$$

The coefficient of determination is

$$(1.19) \qquad \tilde{R}^2 \equiv 1 - \frac{\{R(R-1)-(2R+K-1)\}\hat{\sigma}^2}{\sum_i \sum_j (y_{ij}-y_{..})^2}$$

The variances of $\hat{a}_i$ and $\hat{c}_j$ are

$$(1.20) \qquad \mathrm{var}(\hat{a}_i) = \hat{\sigma}^2 \frac{2R^2-3R-1}{R(R-1)(R-2)} + \sum_k \sum_\ell q_{ik} q_{i\ell} \tilde{\omega}_{k\ell}, \qquad i = 1,\ldots,R \;,$$

$$(1.21) \qquad \mathrm{var}(\hat{c}_j) = \hat{\sigma}^2 \frac{(R-1)^2}{R^2(R-2)} + \sum_k \sum_\ell s_{jk} s_{j\ell} \tilde{\omega}_{k\ell} \;, \qquad j = 1,\ldots,R \;.$$

Section 2 will be devoted to a proof of this theorem and section 3 to a discussion of its practical importance.

2. Proof of the Theorem

Rewrite model (1.2) in matrix form as follows,

$$(2.1) \qquad y = Z\delta + X\beta + \varepsilon,$$

where

$$(2.2) \qquad y \equiv (y_{12}, y_{13}, \ldots, y_{1R}, y_{21}, y_{23}, \ldots, y_{2R}, \ldots, y_{R1}, \ldots, y_{R,R-1})'$$

$$(2.3) \qquad \varepsilon \equiv (\varepsilon_{12}, \varepsilon_{13}, \ldots, \varepsilon_{1R}, \varepsilon_{21}, \varepsilon_{23}, \ldots, \varepsilon_{2R}, \ldots, \varepsilon_{R1}, \ldots, \varepsilon_{R,R-1})'$$

X is a $R(R-1) \times K$ matrix of which the k-th column x_k is

$$(2.4) \qquad x_k \equiv (x_{12k}, x_{13k}, \ldots, x_{1Rk}, x_{21k}, x_{23k}, \ldots, x_{2Rk}, \ldots, x_{R1k}, \ldots, x_{R,R-1,k})'$$

$$(2.5) \qquad Z \equiv (Z_1, Z_2) \;, \quad R(R-1) \times 2R,$$

with

$$(2.6) \qquad Z_1 \equiv \left[\begin{array}{cccc} 1 & & & \\ 1 & & & \\ \vdots & & & \\ 1 & & & \\ \hline & 1 & & \\ & 1 & & \\ & \vdots & & \\ & 1 & & \\ \hline & & \ddots & \\ \hline & & & 1 \\ & & & 1 \\ & & & \vdots \\ & & & 1 \end{array}\right] \begin{array}{l} \\ (R-1) \\ \\ \\ \\ (R-1) \\ \\ \\ \\ \\ (R-1) \\ \\ \\ \end{array} \qquad Z_2 \equiv \left[\begin{array}{ccccc} 0 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ \hline 1 & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ \hline & & \vdots & & \\ \hline 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \end{array}\right] \begin{array}{l} \\ (R-1) \\ \\ \\ \\ (R-1) \\ \\ \\ \\ \\ (R-1) \\ \\ \\ \end{array}$$

$$\qquad\qquad\qquad R \qquad\qquad\qquad\qquad\qquad\qquad R$$

$$(2.7) \qquad \beta \equiv (\beta_1, \beta_2, \ldots, \beta_k)' \ , \ \delta \equiv (\alpha', \gamma')' \ ,$$

with

$$(2.8) \qquad \alpha \equiv (\alpha_1, \alpha_2, \ldots, \alpha_R)', \ \gamma \equiv (\gamma_1, \gamma_2, \ldots, \gamma_R)'$$

In a slightly more compact form, (2.1) can be written as

$$(2.9) \qquad y = W\rho + \varepsilon,$$

with

$$(2.10) \qquad W \equiv (Z, X), \quad \rho \equiv (\beta', \delta')'$$

The matrix Z defined by (2.5) has rank 2R-1 and hence W will generally be of rank 2R-1+K. Consequently, there is no unique least squares solution for the (2R+K)-vector ρ. It is well-known however, (cf., e.g., Searle, 1971, Ch. 5), that all least squares solutions for ρ are generated by

$$(2.11)\qquad \hat{\rho} = (W'W)^{-}\ W'y,$$

where $(W'W)^{-}$ is a g-inverse of $W'W$. By varying over the set of all possible g-inverses of $W'W$ we generate all possible solutions $\hat{\rho}$. Furthermore,

$$(2.12)\qquad \operatorname{var}(\hat{\rho}) \doteq \sigma^{2}(W'W)^{-}\ W'W(W'W)^{-}$$

and the best quadratic unbiased estimator of σ^2 is

$$(2.13)\qquad \hat{\sigma}^{2} = \frac{y'[I-W(W'W)^{-}W']y}{R(R-1)-\operatorname{rank}(W)}\ .$$

The coefficient of determination is

$$(2.14)\qquad \tilde{R}^{2} = 1 - \frac{y'[I-W(W'W)^{-}W']y}{y'y-R(R-1)y_{..}^{2}}\ .$$

The proof of the theorem consists of an elaboration of these formulas, with respect to the model at hand. We will start with (2.11).

$$(2.15)\qquad W'W = \begin{bmatrix} Z'Z & Z'X \\ X'Z & X'X \end{bmatrix},$$

so that a g-inverse of $W'W$ is (cf. Rohde, 1965):

$$(2.16)\ (W'W)^{-} = \begin{bmatrix} (Z'Z)^{-}+(Z'Z)^{-}Z'X(X'PX)^{-1}X'Z(Z'Z)^{-} & -(Z'Z)^{-}Z'X(X'PX)^{-1} \\ -(X'PX)^{-1}X'Z(Z'Z)^{-} & (X'PX)^{-1} \end{bmatrix},$$

where $(Z'Z)^{-}$ is a g-inverse of $Z'Z$ and $P \equiv I-Z(Z'Z)^{-}Z$; P is idempotent and invariant under the choice of $(Z'Z)^{-}$. Using (2.16) to elaborate (2.11) yields

$$(2.17)\qquad \hat{\rho} \equiv \begin{pmatrix} \hat{d} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} (Z'Z)^{-}Z'\{y-(X'PX)^{-1}X'Py\} \\ (X'PX)^{-1}X'Py \end{pmatrix}$$

Notice that P and hence $\hat{b}$ is invariant under the choice of a g-inverse of Z'Z, i.e. $\hat{b}$ is unique, as claimed in the theorem. The non-uniqueness of $\hat{d}$ can be investigated formally by employing the theory of estimable functions. For the sake of brevity, this part of the proof is omitted here.

The next step is to elaborate the matrix P, to obtain a formula for $\hat{b}$ that is computationally efficient. First, notice that

$$(2.18) \qquad Z'Z = \begin{bmatrix} (R-1)I_R & J_R-I_R \\ J_R-I_R & (R-1)I_R \end{bmatrix},$$

where I_R is the identity matrix of order R and J_R is a square R×R-matrix of ones. It can be verified directly that the following matrix is a g-inverse of Z'Z:

$$(2.19) \qquad (Z'Z)^- = \begin{bmatrix} \frac{1}{R-1} I_R & 0 \\ 0 & 0 \end{bmatrix} + \frac{R-1}{R(R-2)} \begin{bmatrix} \frac{1}{(R-1)^2} E_R & \frac{1}{R-1} E_R \\ \frac{1}{R-1} E_R & E_R \end{bmatrix},$$

with $E_R \equiv I_R - \frac{1}{R} J_R$. Next, it is a matter of straightforward manipulation to show that

$$(2.20) \quad Z(Z'Z)^-Z' = \frac{1}{R(R-2)}\{(R-1)Z_1Z_1' + Z_1Z_2' + Z_2Z_1' + (R-1)Z_2Z_2' - \frac{R}{R-1} J_N\},$$

where J_N is an N×N-matrix of ones, $N \equiv R(R-1)$. For P we thus obtain

$$(2.21) \qquad P = I-Z(Z'Z)^-Z' =$$

$$= I_N - \frac{1}{R-2} (Z_1Z_1'+Z_2Z_2') + \frac{1}{R(R-2)} (Z_1-Z_2)(Z_1-Z_2)' + \frac{1}{(R-1)(R-2)} J_N$$

To see how this works out for $\hat{b}$, consider as an example the expression for Py. Define $\eta_1 \equiv Z_1'y$, $\eta_2 \equiv Z_2'y$, $\eta_3 \equiv \iota_N'y$, with ι_N an N-vector of ones. This yields

$$(2.22) \quad Py = y - \frac{1}{R-2} (Z_1\eta_1+Z_2\eta_2) + \frac{1}{R(R-2)} (Z_1-Z_2)(\eta_1-\eta_2) + \frac{\eta}{(R-1)(R-2)} \iota_N$$

Thus the (i,j)-element of the vector Py is

$$(2.23)\quad (Py)_{ij} = y_{ij} - \frac{1}{R-2}(\eta_{1i}+\eta_{2j}) + \frac{1}{R(R-2)}(\eta_{1i}-\eta_{1j}-\eta_{2i}+\eta_{2j}) + \frac{\eta}{(R-1)(R-2)}$$

or

$$(2.24)\quad (Py)_{ij} = \tilde{y}_{ij} \,,$$

as defined on the previous section (cf. (1.10) and below). Premultiplication of X by P transforms the columns of X analogously, so that x_{ijk} is replaced by $\tilde{x}_{ijk}$, cf. (1.10). As a result we recognize the formula for $\hat{b}$ as the least squares estimate of β in model (1.11), as claimed in the theorem.

Now consider $\hat{d} \equiv (\hat{a}',\hat{c}')'$, where $\hat{a} = (\hat{a}_1,\hat{a}_2,\ldots,\hat{a}_R)'$, $\hat{c} = (\hat{c}_1,\hat{c}_2,\ldots,\hat{c}_R)'$. The expression for $\hat{d}$ in (2.17) can be rewritten as

$$(2.25)\quad \hat{d} = (Z'Z)^{-}Z'r \,,$$

where $r \equiv y - X\hat{b}$. Define

$$(2.26)\quad p_1 \equiv Z_1'r, \quad p_2 \equiv Z_2'r, \quad p \equiv \iota_N'r = \iota_R' p_1 = \iota_R' p_2$$

Using (2.19), we can write

$$(2.27)\quad \hat{d} = \begin{pmatrix} \frac{1}{R-1} p_1 \\ 0 \end{pmatrix} + \frac{R-1}{R(R-2)} \begin{bmatrix} \frac{1}{R-1}\left(\frac{1}{R-1} p_1 + p_2 - \frac{p}{R-1}\iota_R\right) \\ \frac{1}{R-1} p_1 + p_2 - \frac{p}{R-1}\iota_R \end{bmatrix}$$

$$= \frac{1}{R(R-2)} \begin{bmatrix} (R-1)\, p_1 + p_2 - \frac{p}{R-1}\iota_R \\ p_1 + (R-1)\, p_2 - p\cdot\iota_R \end{bmatrix} = \begin{pmatrix} \hat{a} \\ \hat{c} \end{pmatrix}$$

Inserting the expressions for p_1, p_2 and p in (2.27) directly yields (1.16) and (1.17) of the Theorem.

Now consider the variances of the estimators. It is easy to verify that $(Z'Z)^{-}$ given by (2.19) is a reflexive g-inverse of Z'Z and that $(W'W)^-$ given by (2.16) is a reflexive g-inverse of W'W. Hence the variance covariance matrix of $\hat{\rho}$ is $\sigma^2(W'W)^-$. Consequently $\sigma^2(X'PX)^{-1}$ is the variance covariance matrix of $\hat{b}$ and this is nothing else than $\tilde{\Omega}$ defined below (1.11), as claimed in the Theorem.

From (2.16), we have that the variance covariance matrix of $\hat{d}$ is equal to

$$\text{(2.28)} \qquad \text{var}(\hat{d}) = \text{var}\begin{pmatrix} \hat{a} \\ \hat{c} \end{pmatrix} = \sigma^2\{(Z'Z)^- + (Z'Z)^-Z'X(X'PX)^{-1}X'Z(Z'Z)^-\}$$

To prove (1.20) and (1.21) in the Theorem, we will show that these expressions correspond to diagonal elements of the matrix on the right hand side of (2.28). First consider $(Z'Z)^-$. From (2.19) it is clear that its diagonal elements in the upper left (R×R)-block are equal to $(2R^2-3R-1)/R(R-1)(R-2)$. The diagonal elements in the lower right (R×R)-block are equal to $(R-1)^2/R^2(R-2)$. Next, consider $(Z'Z)^-Z'X(X'PX)^{-1}.$ $X'Z(Z'Z)^-$. To evaluate this expression, we note that the k-th column of $(Z'Z)^-Z'X$ has the same structure as $\hat{d}$, given in (2.25). By analogy with (2.27) and (1.16) and (1.17) it is clear that

$$\text{(2.29)} \qquad (Z'Z)^-Z'X = \begin{pmatrix} Q \\ S \end{pmatrix} ,$$

where the (R×K)-matrices Q en S have typical elements q_{ik} and s_{jk}, defined in (1.14) and (1.15). Denoting the (k,ℓ)-element of $\sigma^2(X'PX)^{-1}$ by $\tilde{\omega}_{k\ell}$, as in section 1, we obtain as typical diagonal elements of $(Z'Z)^-Z'X(X'PX)^{-1}X'Z(Z'Z)^-$, $\Sigma_k\Sigma_\ell q_{ik}q_{i\ell}\tilde{\omega}_{k\ell}$, $i = 1,\dots,R$, in the upper left (R×R)-block and $\Sigma_k\Sigma_\ell s_{jk}s_{j\ell}\tilde{\omega}_{k\ell}$, $j = 1,\dots,R$, in the lower right (R×R)-block. Combining this with the results for $(Z'Z)^-$ yields (1.20) and (1.21) of the Theorem.

To arrive at the best quadratic estimate for σ^2 we note that in (2.13) $y'[I-W(W'W)^-W']y$ is equal to $y'P(I-PX(X'PX)^{-1}X'P)Py$, which is simply the residual sum of squares for model (1.11). However, when applying least squares to (1.11), we would divide the residual sum of squares by R(R-1)-K to arrive at the best quadratic estimator of σ^2 in that model, whereas in (2.13) we divide by R(R-1)-(2R+K-1), which explains the "degrees of freedom corrections" in (1.18) of the Theorem.

Finally, (2.14) immediately implies (1.19).

3. Discussion

As with the standard model, we can use the indeterminateness of the estimates of α_i and γ_j to introduce an intercept β_0, and to restrict $\hat{a}_i$ and $\hat{c}_j$ to sum to zero. It is easily seen from (1.17) that the $\hat{c}_j$ already add up to zero, so that we only have to adapt the $r_{..}$-term in (1.16); (1.16) is hence replaced by

$$(1.16)' \quad \hat{a}_i = \frac{(R-1)^2}{R(R-2)} r_{i.} + \frac{R-1}{R(R-2)} r_{.i} - \frac{R-1}{R-2} r_{..}$$

and the estimate of β_0 is then

$$(3.1) \qquad \hat{b}_0 \equiv r_{..} ,$$

which is analogous to (1.8).

A comparison of the formulas in the Theorem with the corresponding ones in the Lemma, shows that for $R \to \infty$ they become pairwise identical. This is what one might have expected, because for increasing R the missing observations make up a smaller proportion of the total number of observations.

The formulas in the Theorem look a bit more complicated than those in the Lemma, but their computational complexity is the same. Given the estimate for β, $\hat{b}$, the computation of all other quantities ($\hat{a}_i$, $\hat{c}_i$, etc.) requires at most $O(N)$ time ($N \equiv R(R-1)$). The computation of $\hat{b}$ itself requires simple manipulations of the data to arrive at model (1.11). These manipulations and the computation of $\hat{b}$ also require $O(N)$ time. In contrast, if least squares would be applied to (1.2) directly, one would have to invert a $(2R+K) \times (2R+K)$-matrix, which requires $O((2R+K)^3)$ time. Thus, application of the Theorem reduces the computing time required approximately by a factor of R. In practice, where R may be in the hundreds, this may very well be the difference between feasibility and non-feasibility of the estimation task.

References

Cesario, F.J. (1975), "Least Squares Estimation of Trip Distribution Parameters", Transportation Research, 9, pp. 13-18.

Judge, G.G., W.E. Griffiths, R.C. Hill, T.C. Lee (1980), The Theory and Practice of Econometrics, Wiley, New York.

Rohde, C.A. (1966), "Some Results on Generalized Inverse ", SIAM Review, 8, pp. 201-205.

Scheffé, H. (1959), The Analysis of Variance, Wiley, New York.

Searle, S.R. (1971), Linear Models, Wiley, New York.

IN 1983 REEDS VERSCHENEN

126 H.H. Tigelaar
Identification of noisy linear systems with multiple arma inputs.

127 J.P.C. Kleijnen
Statistical Analysis of Steady-State Simulations: Survey of Recent Progress.

128 A.J. de Zeeuw
Two notes on Nash and Information.

129 H.L. Theuns en A.M.L. Passier-Grootjans
Toeristische ontwikkeling - voorwaarden en systematiek; een selectief literatuuroverzicht.

130 J. Plasmans en V. Somers
A Maximum Likelihood Estimation Method of a Three Market Disequilibrium Model.

131 R. van Montfort, R. Schippers, R. Heuts
Johnson S_U-transformations for parameter estimation in arma-models when data are non-gaussian.

132 J. Glombowski en M. Krüger
On the Rôle of Distribution in Different Theories of Cyclical Growth.

133 J.W.A. Vingerhoets en H.J.A. Coppens
Internationale Grondstoffenovereenkomsten.
Effecten, kosten en oligopolisten.

134 W.J. Oomens
The economic interpretation of the advertising effect of Lydia Pinkham.

135 J.P.C. Kleijnen
Regression analysis: assumptions, alternatives, applications.

136 J.P.C. Kleijnen
On the interpretation of variables.

137 G. van der Laan en A.J.J. Talman
Simplicial approximation of solutions to the nonlinear complementarity problem with lower and upper bounds.

IN 1984 REEDS VERSCHENEN

138 G.J. Cuypers, J.P.C. Kleijnen en J.W.M. van Rooyen
Testing the Mean of an Asymetric Population:
Four Procedures Evaluated

139 T. Wansbeek en A. Kapteyn
Estimation in a linear model with serially correlated errors when observations are missing

140 A. Kapteyn, S. van de Geer, H. van de Stadt, T. Wansbeek
Interdependent preferences: an econometric analysis

141 W.J.H. van Groenendaal
Discrete and continuous univariate modelling

142 J.P.C. Kleijnen, P. Cremers, F. van Belle
The power of weighted and ordinary least squares with estimated unequal variances in experimental design

143 J.P.C. Kleijnen
Superefficient estimation of power functions in simulation experiments

144 P.A. Bekker, D.S.G. Pollock
Identification of linear stochastic models with covariance restrictions.

145 Max D. Merbis, Aart J. de Zeeuw
From structural form to state-space form

146 T.M. Doup and A.J.J. Talman
A new variable dimension simplicial algorithm to find equilibria on the product space of unit simplices.

147 G. van der Laan, A.J.J. Talman and L. Van der Heyden
Variable dimension algorithms for unproper labellings.

148 G.J.C.Th. van Schijndel
Dynamic firm behaviour and financial leverage clienteles

149 M. Plattel, J. Peil
The ethico-political and theoretical reconstruction of contemporary economic doctrines

150 F.J.A.M. Hoes, C.W. Vroom
Japanese Business Policy: The Cash Flow Triangle
an exercise in sociological demystification

151 T.M. Doup, G. van der Laan and A.J.J. Talman
The $(2^{n+1}-2)$-ray algorithm: a new simplicial algorithm to compute economic equilibria

IN 1984 REEDS VERSCHENEN (vervolg)

152 A.L. Hempenius, P.G.H. Mulder
Total Mortality Analysis of the Rotterdam Sample of the Kaunas-Rotterdam Intervention Study (KRIS)

153 A. Kapteyn, P. Kooreman
A disaggregated analysis of the allocation of time within the household.