Deriving Bounds on the Structural Vector when the Measurement Errors are Correlated: An elaboration of the Frisch/Reiersøl Approach

by

Yngve Willassen

Department of Economics
University of Oslo
P.O. Box 1095, Blindern
N-0317 Oslo, Norway

1. Introduction

In the general linear errors-in-variables model the main results have been derived under the assumption that the measurement errors are uncorrelated. However, as recognized by Bekker, Kapteyn and Wansbeek (BKW) (1987) and Lach (1993) this is often a problematic assumption to maintain in empirical applications since quite trivial variable transformations will often create correlation between the errors. BKW (op.cit.) derived parameter bounds without assuming a diagonal covariance-matrix of the errors. Instead they derived their results by supposing that the econometricians are able to impose an a priori upper bound on the covariance-matrix of the measurements errors. Lach (op.cit.) examined various implications of correlated errors introduced by one particular variable transformation.

However, the more succinct study on the bounds of the structural parameters when the errors are correlated is given by an interesting study by Erickson (1993). Erickson deduces his result by studying the covariance equations of the observable variables. Although, he doesn't give a complete solution to the problem he poses, his analysis turns out to be extremely difficult. Indeed, Erickson shows a remarkable amount of skill and ingeniuity by being able to clarify parts of this problem by this approach. However, reading his paper one gets a strong feeling that one should face this problem by quite a different approach. An approach which can also be generalized to more complicated models. This is the purpose of the present paper.

By an apporach initiated by Frisch (1934) and elaborated an extended by Reiersøl (1941) and (1945) we shall show the complete solution to Erickson's model. By appealing to the basic structure of the model the necessary analysis proves to be quite simple and transparent.

The plan of the paper is the following: In section 2 we specify the model to be studied and list results from matrix theory which we will use repeatedly. In section 3 we prove the mathmatical results which prepare the ground for deducing the bounds on the structural vector. As regards these bounds it turns out that there are four separate cases to consider. These cases are studied in sections 4–7. Finally, in Appendix A we illustrate important concepts in this problem.

2. Specification of the model together with definitions and useful results from matrix theory

The observable variables are denoted by a $(n \times 1)$ column vector X, which we conveniently write in partitioned form:

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \tag{2.1}$$

where $X^{(1)}$ denotes the (2×1) vector of error distorted variables, while $X^{(2)}$ denotes the $((n-2) \times 1)$ vector of exactly observable variables.

Let $\xi^{(1)}$ denote the (2×1) vector of unobservable systematic variables which are related to their observable counterparts by the equation:

$$X^{(1)} = \xi^{(1)} + \epsilon^{(1)} \tag{2.2}$$

where $\epsilon^{(1)}$ denotes the (2×1) vector of measurement errors.

As regardes $\epsilon^{(1)}$ we assume:

$$\text{The random vectors } \xi^{(1)} \text{ and } \epsilon^{(1)} \text{ are uncorrelated.} \tag{2.3}$$

The entries of $\epsilon^{(1)}$ have zero means and the covariance-matrix L^{11} is given by: (2.4)

$$E(\epsilon^{(1)}\epsilon^{(1)'}) = L^{11} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \cdot & \sigma_2^2 \end{pmatrix} \tag{2.5}$$

where σ_1^2, σ_2^2 denote the variances of ϵ_1 and ϵ_2 respectively, and ρ denotes the correlation coefficient between ϵ_1 and ϵ_2.

Since $X^{(2)} = \xi^2$ the general form of the covariance-matrix of the errors are given by:

$$L = \begin{pmatrix} L^{11} & O \\ O & O \end{pmatrix} \tag{2.6}$$

where the O's denote the zero submatrices of appropriate dimensions.

The covariance-matrix of the observable variables is supposed to be positive definite, and is denoted by:

$$M = \begin{pmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1n} \\ \cdot & \mu_{22} & \dots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & \mu_{nn} \end{pmatrix} \tag{2.7}$$

Below we shall by the notation $L = L(\sigma_1, \sigma_2, \rho, O)$ repeatedly refer to the matrix function given by (2.6). The determinant and cofactors of M are denoted by $|M|$ and M_{ij} $(i, j = 1, 2, \dots, n)$.

Let $(\xi^{(1)})' = (\xi_1, \xi_2)$ then we suppose that the structural variables $(\xi_1, \xi_2, X_3, \dots, X_n)$ satisfy the linear relation:

$$\gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_3 + \dots + \gamma_n\xi_n + \gamma_0 = 0 \tag{2.8}$$

From the assumption (2.3)–(2.5) and the specification (2.8) it follows immediately:

$$(M - L)\gamma = 0 \tag{2.9}$$

where $(M - L)$ is the covariance-matrix of the structural variables $(\xi_1, \xi_2, X_3, \dots, X_n)$, $\gamma' = (\gamma_1, \gamma_2, \dots, \gamma_n)$ denotes the vector of structural parameters, and, finally, 0 denotes a $(n \times 1)$ vector of zeros.

Some useful definitions.

DEFINITION 2.1 For a given symmetric, positive definite matrix M, $\mathscr{L}$ will denote the set of covariance-matrices L such that $L \in \mathscr{L}$ implies:

$$(M - L) \text{ is non-negative definite} \tag{2.10}$$

$$\text{The determinant of } (M - L) \text{ denoted } |M - L| \text{ is zero} \tag{2.11}$$

DEFINITION 2.2 (i) A diagonal matrix T is called a sign-matrix if any of its elements are $+1$ og -1. (ii) A matrix A is said to have compatible signs if there exists a sign-matrix T such that (TAT) has positive elements only.

DEFINITION 2.3 Let $(X^{(1)})' = (X_1, X_2)$ then r will denote the partial correlation coefficient between X_1 and X_2. From Cramér ((1946), ch. 23.4) we know:

$$r = \frac{-M_{12}}{\sqrt{M_{11}M_{22}}} \tag{2.12}$$

In the present study we assume:

$$M_{12} < 0 \Longrightarrow r > 0. \tag{2.13}$$

We shall often use the following results from matrix theory.

RESULT 2.1 Let A be a $(n \times n)$ matrix with rank $n-1$. Then the rank of adj A (the adjoint of A) is 1.

RESULT 2.2 Let A be a $(n \times n)$ matrix with rank $n-1$, and consider the set of linear homogeneous equations $A\gamma = 0$. Then the solution vector γ can be written:

$$\gamma' = t(A_{r1}, A_{r2}, \ldots, A_{rn}) \tag{2.14}$$

where t is any real parameter and $(A_{r1}, A_{r2}, \ldots, A_{rn})$ denotes any non-zero vector of adj A.

RESULT 2.3 Let A be a square matrix with cofactors A_{ij}. Then from Cramér ((1946), ch. 11) we quote the identities:

$$A_{jj}A_{kk} - A_{jk}^2 = |A|A_{jj\cdot kk} \tag{2.15}$$

$$A_{11}A_{ik} - A_{i1}A_{1k} = |A|A_{11\cdot ik} \tag{2.16}$$

where $A_{11\cdot ik}$ denotes the cofactor attained by deleting the first and the i'th row and the first and the k'th column of A.

Finally, the basic idea of our approach is simple. In general terms it can be described as follows. For a given, fixed covariance-matrix M, the set of equations given by (2.9) will determine a mapping from $\mathscr{L}$ into the coefficient space Γ. Hence, when the covariance-matrix L of the errors varies over the set $\mathscr{L}$, the structural vector γ will vary over a subset of Γ determined by (2.9). Since $L \in \mathscr{L}$ the rank of $(M - L)$ is by contruction less than n which implies that γ has a non-trivial solution. However, the rank of $(M-L)$ and the entries of adj$(M - L)$ are sensitive to the value of the correlation coefficient ρ. Subsequently, this will be reflected in the solution set for the structural vector γ.

Below we shall analyse the different cases which emerge. In all cases the rank of $(M - L)$ and the zero/non-zero elements of adj$(M - L)$ proved to be important. Therefore, we shall first consider these questions.

3. *The rank of* $(M - L)$ *and the structure of* adj$(M - L)$ *when* $L \in \mathscr{L}$

In the following we shall for simplicity use the notation: $x = \sigma_1$, $y = \sigma_2$. The following function turns out to be important:

$$f(x, y, \rho) = |M - L| = M_{11\cdot 22}(1-\rho^2)x^2y^2 - M_{11}x^2 - M_{22}y^2 - 2M_{12}\rho xy + |M| \tag{3.1}$$

where the principal minors $M_{11\cdot 22}$, M_{11}, M_{22} and the determinant $|M|$ and finally the cofactor M_{12} are defined above. In particular, $M_{11\cdot 22}$ is the principal minor obtained from M by deleting the first two rows and columns.

Since by assumption M is positive definite $M_{11\cdot 22} > 0$. Also, since x and y are standard errors of ϵ_1 and ϵ_2, we are only interested in the range of f when $x \geq 0$, $y \geq 0$.

PROPOSITION 3.1 Let the partial correlation coefficient $r > 0$ be given by (2.12). Then for any feasible value of ρ:

(i) the function f (3.1) has a saddle point at:

$$x_0 = \sqrt{\frac{M_{22}(1-\rho r)}{M_{11\cdot 22}(1-\rho^2)}} \tag{3.2}$$

$$y_0 = \sqrt{\frac{M_{11}(1-\rho r)}{M_{11\cdot 22}(1-\rho^2)}} \tag{3.3}$$

(ii) When $\rho = r$ we attain $f(x_0, y_0, \rho) = 0$.

(iii) When $\rho \neq r$ we attain $f(x_0, y_0, \rho) < 0$.

PROOF (i) For a given value of ρ the stationary point (x_0, y_0) of f is determined by eqs.

$$\frac{\partial f}{\partial x} = 0 \tag{3.4}$$

$$\frac{\partial f}{\partial y} = 0 \tag{3.5}$$

The Hessian corresponding to this point is given by:

$$-16(1-\rho r)M_{11}M_{22} < 0 \tag{3.6}$$

since $(1-\rho r)M_{11}M_{22}$ is always positive. ∎

(ii) When $\rho = r$, x_0 and y_0 become: $x_0 = \sqrt{M_{22}}/\sqrt{M_{11\cdot 22}}$, $y_0 = \sqrt{M_{11}}/\sqrt{M_{11\cdot 22}}$. By direct calculation we attain:

$$f = -\frac{M_{11}M_{22}(1-\rho^2)}{M_{11\cdot 22}} + |M| \tag{3.7}$$

Since $M_{11\cdot 22}|M| = (M_{11}M_{22} - M_{12}^2)$ by (2.14) and since $\rho^2 = \dfrac{M_{12}^2}{M_{11}M_{22}}$ when $\rho = r$ we attain:

$$f(x_0, y_0, \rho) = 0$$

∎

(iii) By direct calculation we attain:

$$\frac{\partial f(x_0, y_0, \rho)}{\partial \rho} = \frac{2M_{11}M_{22}(1-\rho r)(r-\rho)}{1-\rho^2} \tag{3.9}$$

Hence, $(\partial f/\partial \rho) > 0$ when $r > \rho$ and $(\partial f/\partial \rho) < 0$ when $r < \rho$. Since $f(x_0, y_0, \rho) = 0$ when $r = \rho$, the conclusion follows. ∎

REMARK 3.1 The saddle-point (3.2)–(3.3) has an interesting implication for the covariance-matrix $(M - L)$. Since $(M - L)$ is non-negative definite, all its principal minors must be non-negative. In particular, $(M-L)_{11} \geq 0$ and $(M-L)_{22} \geq 0$ imply that $x^2 \leq M_{22}/M_{11\cdot 22}$ and $y^2 \leq M_{11}/M_{11\cdot 22}$. When $r = \rho$ the saddle-point (x_0, y_0) corresponds exactly to these values of x and y. Hence, when $\rho = r$ the matrix $(M - L)$ evaluated at the saddle-point (x_0, y_0) will imply $|M-L|_{11} = |M-L|_{22} = 0$. The fact that $f(x_0, y_0, \rho) = |M-L| = 0$ makes sure that $L \in \mathscr{L}$. This proposition enables us to characterize the rank properties of $(M - L)$ when $L \in \mathscr{L}$.

PROPOSITION 3.2 Let again the covariance-matrix of the observable variables M be given, and assume as always that the partial correlation coefficient r is positive. If we vary the covariance matrix L of the errors over $\mathscr{L}$, then it follows:

(i) If $\rho = r$, then the matrix L^* denoted by $L^* = L\left(\sqrt{M_{22}/M_{11\cdot 22}}, \sqrt{M_{11}/M_{11\cdot 22}}, \rho, O\right)$ (see (2.6)) is contained in $\mathscr{L}$ and the rank of $(M - L^*)$ is not larger than $(n-2)$.

(ii) If $\rho < r$, then for all $L = L(x, y, \rho, O) \in \mathscr{L}$ the rank of $(M - L)$ is $(n-1)$ and all elements of the first two rows/columns of $(M - L)$ are non-zero.

(iii) If $\rho > r$, then there exist matrices $\tilde{L}$ and $\underset{\sim}{L}$ where respectively the second row/column, and the first row/column of $\text{adj}(M - L)$ consist of zeros only.

PROOF (i) Since M is non-singular we have

$$(M - L) = (I - LM^{-1})M \tag{3.10}$$

Then we partition M og M^{-1} in accordance with the partition (2.6) of L. That is

$$M = \begin{pmatrix} M(1.1) & M(1.2) \\ M(2.1) & M(2.2) \end{pmatrix} \tag{3.11}$$

where $M(1.1) \sim (2 \times 2)$, $M(1.2) \sim (2 \times (n-2))$, $M(2.1) \sim ((n-2) \times 2)$, $M(2.2) \sim ((n-2) \times (n-2))$.

Furthermore:

$$M^{-1} = \begin{pmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{pmatrix} \tag{3.12}$$

where, in particular,

$$M^{11} = (M(1.1) - M(1.2)M(2.2)^{-1}M(2.1))^{-1} \tag{3.13}$$

(see f.ex. Anderson ((1984), A.3)).

Then we have:

$$(I - LM^{-1}) = \begin{pmatrix} I_2 - L^{11}M^{11} & -L^{11}M^{12} \\ O & I_{n-2} \end{pmatrix} \tag{3.14}$$

where $I_2 \sim (2 \times 2)$ and $I_{n-2} \sim ((n-2) \times (n-2))$ are identity matrices.

Using the definition (2.12) of r and the fact that $\rho = r$ we have by evaluating L^{11} at L^*:

$$L^{11} = \begin{pmatrix} \dfrac{M_{22}}{M_{11\cdot 22}} & -\dfrac{M_{12}}{M_{11\cdot 22}} \\ . & \dfrac{M_{11}}{M_{11\cdot 22}} \end{pmatrix} \tag{3.15}$$

Similarly, from (3.12) we have:

$$M^{11} = \begin{pmatrix} \dfrac{M_{11}}{|M|} & \dfrac{M_{21}}{|M|} \\ . & \dfrac{M_{22}}{|M|} \end{pmatrix} \tag{3.16}$$

We have $M_{21} = M_{12}$ since M is symmetric and by identity (2.15) $M_{11}M_{22} - M_{12}^2 = |M|M_{11\cdot 22}$. Using these facts we attain from (3.15)–(3.16) that:

$$L^{11}M^{11} = I_2 \tag{3.17}$$

Hence, when $L = L^*$ we attain:

$$(I - L^*M^{-1}) = \begin{pmatrix} O & -L^{11}M^{12} \\ O & I_{n-2} \end{pmatrix} \tag{3.18}$$

Since the first two columns of $(I - L^*M^{-1})$ consist of zero elements only, we conclude that the rank of $(I - L^*M^{-1})$, to be denoted $\text{rank}(I - L^*M^{-1})$, is $n - 2$.

Finally, since:

$$\text{rank}((I - LM^{-1})M) \leq \min(\text{rank}(I - LM^{-1}), \text{rank } M^{-1}) \tag{3.19}$$

it follows from (3.10) and (3.19) that

$$\text{rank}(M - L) \leq n - 2 \tag{3.20}$$

proving (i). ∎

(ii) $\rho < r$. For a given ρ we solve the eq. (3.1), i.e. the equation $f(x, y, \rho) = |M - L| = 0$ w.r.t. y, attaining:

$$y = \frac{M_{12}\rho x \pm \sqrt{M_{12}^2\rho^2x^2 - (M_{11\cdot 22}(1-\rho^2)x^2 - M_{22})(|M| - M_{11}x^2)}}{(M_{11\cdot 22}(1-\rho^2)x^2 - M_{22})} \tag{3.21}$$

Since x and y are standard errors only real, non-negative values of x, y are feasible. Since the denominator of (3.21) is negative we shall obviously use the negative “root” for small values of x. However, by increasing x from $x = 0$ we observe that the expression in the square root of (3.21) will sooner or later become negative. Simple calculations show that the discriminant of (3.21) will be zero for:

$$(x^2)_1 = \frac{M_{11\cdot 22}|M|}{M_{11\cdot 22}(1-\rho^2)M_{11}} = \frac{(1-r^2)M_{11}M_{22}}{M_{11\cdot 22}(1-\rho^2)M_{11}} = \frac{(1-r^2)M_{22}}{(1-\rho^2)M_{11\cdot 22}} \tag{3.22}$$

$$(x^2)_2 = \frac{M_{22}}{M_{11\cdot 22}} \tag{3.23}$$

In deducing (3.22) we have used the fact that $M_{11\cdot 22}|M| = (M_{11}M_{22} - M_{12}^2)$ (identity (2.15)), and the expression (2.12) for r. Since $\rho < r$ implies $((1-r^2)/(1-\rho^2)) < 1$ we observe by comparing (3.22) and (3.23) that

$$x^2 = (x^2)_1 < (x^2)_2 \tag{3.24}$$

when $\rho < r$. Hence when $\rho < r$, $(x^2)_1$ is the largest value of x^2 which is consistent with the requirement that the corresponding covariance-matrix L is contained in $\mathscr{L}$. The value of y or y^2 corresponding to $\tilde{x}^2$ is calculated from (3.21) giving:

$$y^2 = \frac{(1-r^2)\rho^2 M_{11}}{(1-\rho^2)r^2 M_{11\cdot 22}} < \frac{M_{11}}{M_{11\cdot 22}} \tag{3.25}$$

when $\rho < r$.

Hence, when $\rho < r$ we have $x^2 < M_{22}/M_{11\cdot 22}$ and $y^2 < M_{11}/M_{11\cdot 22}$ which implies (Remark 3.1) that the principal minors $|M-L|_{11}$ and $|M-L|_{22}$ are always positive in this case. Similar reasoning implies that also the remaining principal minors of order $n-1$ are positive when $\rho < r$, $L \in \mathscr{L}$. Hence, in this case the rank of $(M-L)$ is $n-1$, and since $|M-L| = 0$ it follows from identity (2.15) that all cofactors $|M-L|_{ik}$ $(i, k = 1, 2, \dots, n)$ will be different from zero. This proves (ii). ∎

(iii) $(\rho > r)$. By repeating the arguments above we shall now choose the root (3.23), i.e.

$$\tilde{x} = \sqrt{\frac{M_{22}}{M_{11\cdot 22}}} \tag{3.26}$$

The corresponding value of y is again obtained from (3.21), giving:

$$\tilde{y} = \frac{r}{\rho}\sqrt{\frac{M_{11}}{M_{11\cdot 22}}} \tag{3.27}$$

Again, by Remark 3.1 the principal minor $|M-L|_{22}$ is zero for this value of x (3.26). Since $|M-L| = 0$, it follows again from identity (2.15) that all the cofactors $|M-L|_{12}, |M-L|_{23}, \dots, |M-L|_{2n}$ will also be zero. Hence, the second row/column of $\text{adj}(M-L)$ will consist of zeros only for $L = \tilde{L}$ where $\tilde{L} = L(\tilde{x}, \tilde{y}, \rho, O) \in \mathscr{L}$. By symmetry, it is evident that the matrix corresponding to $L = \underset{\sim}{L}$, where $\underset{\sim}{L} = L(\underset{\sim}{x}, \underset{\sim}{y}, \rho, O)$ and $\underset{\sim}{x}, \underset{\sim}{y}$ are given by:

$$\underset{\sim}{x} = \frac{r}{\rho}\sqrt{\frac{M_{22}}{M_{11\cdot 22}}} \tag{3.28}$$

$$\underset{\sim}{y} = \sqrt{\frac{M_{11}}{M_{11\cdot 22}}} \tag{3.29}$$

will be contained in $\mathscr{L}$. Repeating the above arguments now imply that the first row/column of $\text{adj}(M-L)$ consist of zeros only. This proves (iii). ∎

Remark 3.2 Proposition 3.1 enables us to give instructive pictures of the set $\mathscr{L}$. The proposition indicates that we shall separate four different cases: (i) $\rho \leq 0$, (ii) $0 < \rho < r$, (iii) $\rho = r$, (iv) $\rho > r$. It is also evident from this proposition that the level curves determined by $f(x, y, \rho) = |M-L| = 0$ have two branches. These branches are disjoint when $\rho \neq r$, when $\rho = r$ they have a common point at the saddle-point (x_0, y_0). At this point, as we know from proposition 3.2, the rank of $(M-L)$ drops from $(n-1)$ to a value not greater than $(n-2)$. In order to show what $\mathscr{L}$ looks like, we have carried out the necessary calculation for a particular M in appendix A.

Propositions 3.1–3.2 prepare the ground for deducing the bounds on the structural parameters. Below we shall study the four cases in succession. We always suppose $M_{12} < 0$.

4. Bounds on γ when $\rho \leq 0$

We start with the equations (2.9) or:

$$(M - L)\gamma = 0, \qquad L \in \mathscr{L} \tag{4.1}$$

Suppose that $\gamma_1 \neq 0$ and let us define:

$$\gamma_1^{-1}(\gamma) = \begin{pmatrix} 1 \\ g \end{pmatrix} \tag{4.2}$$

Then, since M is non-singular we can write (4.1):

$$(I - LM^{-1})M \begin{pmatrix} 1 \\ g \end{pmatrix} = 0 \tag{4.3}$$

Let us define:

$$\psi = \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \end{pmatrix} = M \begin{pmatrix} 1 \\ g \end{pmatrix} \tag{4.4}$$

where $(\psi^{(1)})' = (\psi_1, \psi_2)$ and $(\psi^{(2)})' = (\psi_3, \ldots, \psi_n)$. Then, using the partition (3.14), (4.3) can be written:

$$\begin{pmatrix} I_2 - L^{11}M^{11} & -L^{11}M^{12} \\ 0 & I_{n-2} \end{pmatrix} \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \end{pmatrix} = 0 \tag{4.5}$$

The following observations are immediate from (4.3) and (4.5). Firstly, it follows from (4.5) that:

$$\psi^{(2)} = 0 \tag{4.6}$$

so that (4.5) reduces to:

$$(I_2 - L^{11}M^{11})\psi^{(1)} = 0 \tag{4.7}$$

Secondly, we observe that $|M - L| = 0$ implies:

$$|I_2 - L^{11}M^{11}| = 0 \tag{4.8}$$

so that with the present specification of the model, LM^{-1} and $L^{11}M^{11}$ have the same eigenvalues. Since $L \in \mathscr{L}$ it can be proved (Klette and Willassen (1996), th. 2.2) that 1 is the largest eigenvalues of LM^{-1}. Since (LM^{-1}) and $(L^{11}M^{11})$ have the same eigenvalues, 1 is also the largest eigenvalue of $(L^{11}M^{11})$.

When $\rho < 0$ and $M_{12} < 0$ it follows from the definition (2.5) of L^{11} and (3.12) that L^{11} and M^{11} have the same sign-pattern:

$$\begin{pmatrix} + & - \\ - & + \end{pmatrix} \tag{4.9}$$

which implies that $L^{11}M^{11}$ has compatible signs. That is, if we define the sign-matrix T by:

$$T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.10}$$

then the matrix $(T(L^{11}M^{11})T)$ has positive entries only. Since $TT = I_2$ it is evident that $L^{11}M^{11}$ and $(TL^{11}M^{11}T)$ have the same eigenvalues. Since $(TL^{11}M^{11}T)$ has positive elements only and its largest eigenvalue is 1, it follows by applying Frobenious' matrix theory that we can choose the eigenvector $(T\psi^{(1)})$ in (4.7) so that it has only positive elements.

Combining (3.12), (4.4), (4.6) we observe:

$$\begin{pmatrix} 1 \\ g \end{pmatrix} = M^{-1} \begin{pmatrix} \psi^{(1)} \\ 0 \end{pmatrix} = \begin{pmatrix} M^{11} \\ M^{21} \end{pmatrix} \psi^{(1)} \tag{4.11}$$

where 0 denotes the $((n-2) \times 1)$ zero-vector.

Let us define the sign-corrected $(n \times 2)$ matrix

$$\Omega^{-1} = \begin{pmatrix} M^{11} \\ M^{21} \end{pmatrix} T \tag{4.12}$$

where T is the sign-matrix (4.10).

Hence, the first column of Ω^{-1} is identical to the first column of (M^{-1}), while the second one is equal to second column of M^{-1} multiplied by -1. Since, by assumption $M_{12} < 0$, it follows that the first row of Ω^{-1} consists of two positive entries. Then let us define:

$$\ell' = (1, 1) \tag{4.13}$$

$$D = \text{the diagonal matrix with elements consisting of the first row of } \Omega^{-1} \tag{4.14}$$

$$\begin{pmatrix} \ell' \\ P \end{pmatrix} = (\Omega^{-1}D^{-1}) \tag{4.15}$$

where P consists of the two columns

$$P_j' = (M_{2j}/M_{1j}, M_{3j}/M_{1j}, \ldots, M_{nj}/M_{1j}), \qquad j = 1, 2 \tag{4.16}$$

$$\mathscr{P} = \text{the convex combination of } P_1 \text{ and } P_2 \tag{4.17}$$

Then we summarize all the details for this case in the following proposition.

Proposition 4.1 Suppose the model is specified by eqs. (2.1)–(2.9). Suppose also that $M_{12} < 0$ and $\rho \leq 0$. Then the eqs.

$$(M - L)\begin{pmatrix} 1 \\ g \end{pmatrix} = 0, \qquad L \in \mathscr{L} \tag{4.18}$$

determines a mapping from $\mathscr{L}$ to the convex set $\mathscr{P}$ (4.17), such that for every $L \in \mathscr{L}$ there corresponds one and only one $g \in \mathscr{P}$.

PROOF Most of the details have been given above, so it is sufficient to start with (4.11). Let T be the sign-matrix (4.10). Since $TT = I_2$ we have:

$$\begin{pmatrix} 1 \\ g \end{pmatrix} = \begin{pmatrix} M^{11} \\ M^{21} \end{pmatrix} T(T\psi^{(1)}) = \Omega^{-1}(T\psi^{(1)}) \tag{4.19}$$

where $(T\psi^{(1)})$ is the positive eigenvector satisfying:

$$(I_2 - TL^{11}M^{11}T)(T\psi^{(1)}) = 0 \tag{4.20}$$

(see (4.7)). Then by (4.14) we obtain:

$$\begin{pmatrix} 1 \\ g \end{pmatrix} = (\Omega^{-1}D^{-1})D(T\psi^{(1)}) = \begin{pmatrix} \ell' \\ P \end{pmatrix} w \tag{4.21}$$

where:

$$w = D(T\psi^{(1)}) \tag{4.22}$$

The elements of the diagonal matrix D and the eigenvector $(T\psi^{(i)})$ are all positive. Therefore, the elements of the column-vector w will also be positive. Then it follows from (4.21) that:

$$1 = \ell' w = (w_1 + w_2) \tag{4.23}$$

$$g = Pw = w_1 P_1 + w_2 P_2 \tag{4.24}$$

Eqs. (4.23)–(4.24) show that the structural vector g is a convex combination of P_1 and P_2. ∎

REMARK 4.1 We note that P_1 is the regression vector obtained by taking ordinary least square regression of X_1 on $X_2, \ldots, X_n$, while P_2 is the regression vector obtained by taking the regression of X_2 on $X_1, X_3, \ldots, X_n$ and then solve wrt. X_1.

5. *Bounds on γ when $0 < \rho < r$*

In this case the submatrix L^{11} has the sign-pattern:

$$\begin{pmatrix} + & + \\ + & + \end{pmatrix} \tag{5.1}$$

while M^{11} still has the pattern:

$$\begin{pmatrix} + & - \\ - & + \end{pmatrix} \tag{5.2}$$

Hence, in this case the matrix $L^{11}M^{11}$ doesn't have compatible signs. Therefore, proposition 4.1 is not applicable, and we conclude that the structural vector γ is not restricted to the convex combination $\mathscr{P}$.

According to proposition 3.2 (ii) we can in this case always suppose $\gamma_1 \neq 0$. Then we can write (2.9):

$$(M - L)\begin{pmatrix} 1 \\ g \end{pmatrix} = 0 \tag{5.3}$$

$$\begin{pmatrix} 1 \\ g \end{pmatrix} = \gamma_1^{-1}(\gamma) \tag{5.4}$$

By proposition 3.2 (ii) and matrix Result 2.2 we can take the first row of $\mathrm{adj}(M - L)$ as the solution of the structural vector γ. Hence, the structural equation becomes:

$$\xi_1 + g_2\xi_2 + g_3X_3 + \cdots + g_nX_n + g_0 = 0 \tag{5.5}$$

where g_0 denotes an inessential constant and the structural parameters are given by:

$$g_j = \frac{(M - L)_{1j}}{(M - L)_{11}}, \qquad j = 2, 3, \ldots, n \tag{5.6}$$

In order to study the range of the structural parameters when L varies over $\mathscr{L}$, it is enough to consider one of them in detail, f.ex. g_2, since the remaining g's can be handled in exactly the same way.

For a given value of $\rho \in (0, r)$ we know from proposition 3.2 (ii) that the largest value of σ_1^2 $(= x^2)$ consistent with $L \in L$ is given by:

$$x^2 = \frac{(1 - r^2)M_{22}}{(1 - \rho^2)M_{11\cdot 22}} \tag{5.7}$$

The value of σ_2^2 $(= y^2)$ is then given by:

$$y^2 = \frac{(1 - r^2)\rho^2 M_{11}}{(1 - \rho^2)r^2 M_{11\cdot 22}} \tag{5.8}$$

By the symmetry of this problem we observe that the largest value of σ_2^2 and the corresponding value of σ_1^2 are given by:

$$y^2 = \frac{(1 - r^2)M_{11}}{(1 - \rho^2)M_{11\cdot 22}} \tag{5.9}$$

$$x^2 = \frac{(1 - r^2)\rho^2 M_{22}}{(1 - \rho^2)r^2 M_{11\cdot 22}} \tag{5.10}$$

Hence, for a given $\rho \in (0, r)$ (3.21) becomes

$$y(x) = \frac{M_{12}\rho x - h(x)}{M_{11\cdot 22}(1 - \rho^2)x^2 - M_{22}} \qquad \text{for} \quad x \in \left[0, \sqrt{\frac{(1 - r^2)M_{22}}{(1 - \rho^2)M_{11\cdot 22}}}\,\right] \tag{5.11}$$

$$y(x) = \frac{M_{12}\rho x + h(x)}{M_{11\cdot 22}(1 - \rho^2)x^2 - M_{22}} \qquad \text{for} \quad x \in \left[\sqrt{\frac{|M|}{M_{11}}}, \sqrt{\frac{(1 - r^2)M_{22}}{(1 - \rho^2)M_{11\cdot 22}}}\,\right] \tag{5.12}$$

where:

$$h(x) = \sqrt{M_{12}^2\rho^2x^2 - (M_{11\cdot22}(1-\rho^2)x^2 - M_{22})(|M| - M_{11}x^2)} \tag{5.13}$$

For $\rho \in (0, r)$ the values of y and x given by (5.11)–(5.12) constitute the nonzero values of $L \in \mathscr{L}$. These are the feasible values of x and y for the present problem.

By (5.6) we know that:

$$g_2 = \frac{M_{21} + M_{11\cdot22}\rho xy}{M_{11} - M_{11\cdot22}y^2} \tag{5.14}$$

By proposition 3.2 (ii) the numerator and denominator of (5.14) are non-zero for all feasible values of x and y. Since $M_{21} < 0$ if follows from the continuity that the numerator is negative. Similarly, it follows that the denominator is positive for all feasible x, y.

The set of values given by (5.11)–(5.12) is closed and bounded and hence compact. Since g_2 is continuous it follows from a well-known mathematical result that g_2 will attain both its supremum and infimum on $\mathscr{L}$ when $\rho \in (0, r)$. Using the quantity:

$$R = \frac{\left(r - \sqrt{(r^2-\rho^2)(1-\rho^2)}\right)}{\rho^2} \tag{5.15}$$

we summarize the details in the following proposition.

PROPOSITION 5.1 Fix $\rho \in (0, r)$ and let (x, y) vary over the feasible region given by (5.11)–(5.12). Then we attain:

$$\max g_2 = -\sqrt{\frac{M_{22}}{M_{11}}}\left(R - \sqrt{R^2-1}\right) \tag{5.16}$$

$$\min g_2 = -\sqrt{\frac{M_{22}}{M_{11}}}\left(R + \sqrt{R^2-1}\right) \tag{5.17}$$

PROOF Direct evaluation. (Omitted). ■

The bounds (5.16) and (5.17) agree with those given by Erickson (op.cit. p. 961) in his theorem 1 (b).

6. Bounds on γ when $\rho = r$

Consider the covariance-matrix L of the errors given by:

$$L = L^* = \begin{pmatrix} L^{*11} & O \\ O & O \end{pmatrix} \tag{6.1}$$

where:

$$L^{*11} = \begin{pmatrix} \dfrac{M_{22}}{M_{11\cdot 22}} & -\dfrac{M_{12}}{M_{11\cdot 22}} \\ -\dfrac{M_{12}}{M_{11\cdot 22}} & \dfrac{M_{11}}{M_{11\cdot 22}} \end{pmatrix} \tag{6.2}$$

When $\rho = r = -M_{12}/\sqrt{M_{11}M_{22}}$ we know from proposition 3.2 (i) that $L^* \in \mathscr{L}$. We also know from this proposition that the rank of $(M - L^*)$ is not larger than $n - 2$.

PROPOSITION 6.1 Suppose that $\gamma_1 \neq 0$ and let L^* be given by (6.1)–(6.2). Then the solution of

$$(M - L^*)\begin{pmatrix} 1 \\ g \end{pmatrix} = 0 \tag{6.3}$$

has one degree of freedom at least.

PROOF Since the conclusion of proposition 5.1 follows directly from proposition 3.2 (i), we omit the proof. ■

Prop. 5.1 tells us that the possible range of each of the structural parameters $g_2, g_3, \ldots, g_n$ are from $-\infty$ to $+\infty$ by proper choice of the free parameter(s). Therefore, we can make g outside any bounded set.

7. Bounds on γ when $\rho > r$

Repeating the analysis of section 5 we know that the structural equation can be written:

$$\xi_1 + g_2\xi_2 + g_3X_3 + \cdots + g_nX_n + g_0 = 0 \tag{7.1}$$

where

$$g_j = \frac{(M-L)_{1j}}{(M-L)_{11}}, \qquad j = 2, 3, \ldots, n \tag{7.2}$$

Again it is enough to study one of the g's in detail, f.ex. g_2, since the remaining g's can be handled in exactly the same way. By (7.2) we know that:

$$g_2 = \frac{M_{21} + M_{11\cdot 22}\rho xy}{M_{11} - M_{11\cdot 22}y^2} \tag{7.3}$$

Where:

$$L = \underset{\sim}{L} = \begin{pmatrix} \underset{\sim}{L}^{11} & O \\ O & O \end{pmatrix} \tag{7.4}$$

where:

$$\underset{\sim}{L}^{11} = \begin{pmatrix} \dfrac{r^2 M_{22}}{\rho^2 M_{11\cdot 22}} & \dfrac{r\sqrt{M_{11}M_{22}}}{M_{11\cdot 22}} \\ \dfrac{r\sqrt{M_{11}M_{22}}}{M_{11\cdot 22}} & \dfrac{M_{11}}{M_{11\cdot 22}} \end{pmatrix} \tag{7.5}$$

we know from prop. 3.2 (iii) that $\underset{\sim}{L} \in \mathscr{L}$ and that the principal minor $|M - \underset{\sim}{L}|_{11}$ and the cofactors $|M - \underset{\sim}{L}|_{1j}$ $(j = 2, 3, \ldots, n)$ are all zero.

In the analysis to come we also need the derivative of $y(x)$ (3.21). For a fixed value of ρ we attain:

$$y'(x) = \frac{-((M_{11} - M_{11\cdot 22}(1-\rho^2)y^2)x + M_{12}\rho y)}{h(x)} \tag{7.6}$$

It is evident from our proof of prop. 3.2 (iii) and easily verified directly that $y'(x) = 0$ when

$$\underset{\sim}{x} = \frac{r}{\rho}\sqrt{\frac{M_{22}}{M_{11\cdot 22}}} \tag{7.7}$$

$$\underset{\sim}{y} = \sqrt{\frac{M_{11}}{M_{11\cdot 22}}} \tag{7.8}$$

PROPOSITION 7.1 (i) At the two edges of $\mathscr{L}$, i.e. at $\left(0, \sqrt{|M|/M_{22}}\right)$ and $\left(\sqrt{|M|/M_{11}}, 0\right)$ g attains the values $g_2(0) = M_{22}/M_{21}$ and $g_2\left(\sqrt{|M|/M_{11}}\right) = M_{21}/M_{11}$. (ii) g_2 doesn't have a limit when $x \to \underset{\sim}{x}$ (7.7) and $y \to \underset{\sim}{y}$ (7.8). (iii) $g_2 \to -\infty$ when $x \to \underset{\sim}{x}$ from the left and $g_2 \to +\infty$ when $x \to \underset{\sim}{x}$ from the right.

PROOF (i) The two values of g_2 are attained by direct calculations using (7.3). Both values are negative since $M_{21} = M_{12}$ by the symmetry of M and since $M_{12} < 0$ by assumption. (ii) By prop. 3.2 (iii) the minors $|M - L|_{1j}$ $(j = 1, 2, \ldots, n)$. are all zero at $(x, y) = (\underset{\sim}{x}, \underset{\sim}{y})$. This implies that $g_2 = 0/0$ at $L = \underset{\sim}{L}$. By applying L'Hôspital's rule to g_2 at this point we observe that the derivative of the numerator tends to $M_{11\cdot 22}\rho y > 0$, and the derivative of the denominator tends to $-2M_{11\cdot 22}yy'(x) = 0$ at $x = \underset{\sim}{x}$. This proves (ii). Finally, since $y'(x)$ is positive to the left and negative to the right of $x = \underset{\sim}{x}$ statement (iii) also follows from L'Hôspital's rule. ■

REMARK 7.1 Prop. 7.1 has an interesting implication. First we note that the identity:

$$M_{11}M_{22} - M_{12}^2 = |M|M_{11\cdot 22} > 0 \tag{7.9}$$

implies:

$$g_2\left(\sqrt{|M|/M_{11}}\right) > g_2(0) \tag{7.10}$$

We also note that the derivative of g_2 (7.3) as a function of x is negative to the left and positive to the right of $x = \underset{\sim}{x}$. This fact together with prop. 7.1 (iii) imply:

$$g_2 \in [g_2(0), -\infty) \qquad \text{for} \quad x \in [0, \underset{\sim}{x}) \tag{7.11}$$

$$g_2 \in \left[g_2\left(\sqrt{|M|/M_{11}}\right), +\infty\right) \qquad \text{for} \quad x \in \left(\underset{\sim}{x}, \sqrt{\frac{M_{22}}{M_{11\cdot 22}}}\,\right] \tag{7.12}$$

Hence, except for the open interval $\left(g_2(0), g_2\left(\sqrt{|M|/M_{11}}\right)\right)$ g_2 can take any real number when $\rho > r$.

Finally, it is evident that we can apply the same analysis to the remainig structural parameters $g_3, g_4, \ldots, g_n$ and the corresponding conclusion will follow.

8. Conclusion

In the bivariate case with uncorrelated errors the bounds on the slope parameter is given by Gini (1921). For the generalization of this result to the multivariate case Erickson (op.cit.) refers to Kalman (1982) and Klepper and Leamer (1984). These references appear to be standard in econometric literature. However, this generalization is certainly contained in Koopmans (1937) and Reiersøl (1941). Regarding this theorem we quote from Reiersøl (op.cit. p. 4): "The general formulation of the theorem and the proof of it is given by Koopmans (op.cit. p. 101)." The theorem is also explicitly stated and given an elegant proof in Reiersøl (op.cit. Theorem 1, p. 3).

Above we have given the complete solution to the multivariate version of the model considered by Erickson by an elaboration of Reiersøl's approach. This approach enables us to consider all structural parameters simultaneously and the results emerge quite easily.

Our results deduced in sections 4–7 demonstrate that applying regression theory to this type of models when the structural parameters are not identifiable can be very hazardous.

Finally, there can be no doubt that the present approach is much more powerfull and general that the procedure based on solving covariance equation which is applied by Erickson (op.cit.) and many others, for instance Moran (1971). Hence, the old scholars should not be overlooked.

Appendix A

The four figures below illustrate the set $\mathscr{L}$ defined by (2.10)–(2.11). For a given correlation coefficient (ρ) $\mathscr{L}$ is the branch of the level curve determined by $f(x, y, \rho) = |M - L| = 0$ which is lying nearest to the origin in the x, y plane. The branch of the level curve lying further away from the origin is part of the indefinite region of $(M - L)$.

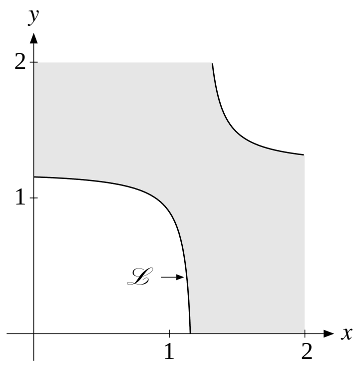
Our figures, which have been drawn by use of proper software, illustrate the four cases: (i) $\rho \leq 0$, (ii) $0 < \rho < r$, (iii) $\rho = r$ and (iv) $\rho > r$ for the particular covariance-matrix:

$$M = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

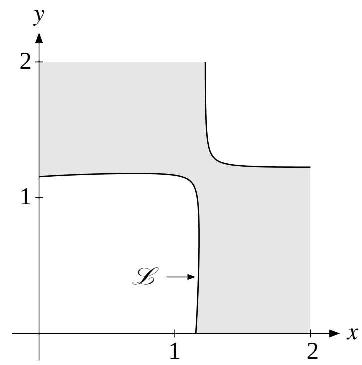
For this matrix the partial correlation coefficient $r = \frac{1}{3}$ (2.12), and the function f (3.1) defined in the xy-plane becomes:

$$f(x, y) = 2(1 - \rho^2)x^2y^2 - 3x^2 - 3y^2 + 2\rho xy + 4 \qquad (A.1)$$

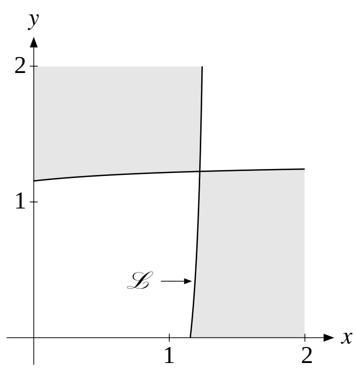
The white and grey regions in the xy-plane indicate where f is positive and negative respectively. Figure (iii) illustrates the case $\rho = r$, and indicates that the two branches of the level curve of $f(x, y, \frac{1}{3}) = 0$ have a common point at the saddle-point $\left(\sqrt{3/2}, \sqrt{3/2}\right)$.



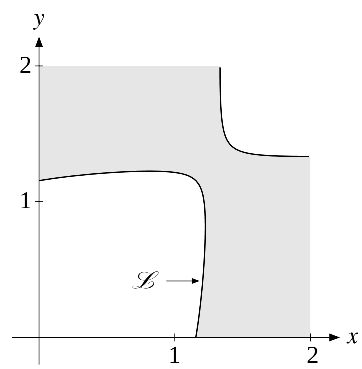
$\rho = -0.1 < r = 1/3$



$\rho = 0.2 < r = 1/3$



$\rho = r = 1/3$



$\rho = 0.5 > r = 1/3$

References

Anderson, T. W. (1984): *An Introduction to Multivariate Analysis*, Wiley Publishing Co., New York.

Bekker, P. , A. Kapteyn and T. Wansbeek (1987): "Consistent Sets of Estimates for Regressions with Correlated or Uncorrelated Measurement Errors in Arbitrary Subsets of All Variables," Econometrica, 55, 1223–1230.

Cramér, H. (1946): *Mathematical Methods of Statistics*, Princeton University Press, Princeton N.J.

Erickson, T. (1993) "Restricting Regression Slopes in the Errors-in-Variables Model by Bounding the Error Correlation," Econometrica, 61, 959–969.

Frisch, R. (1934): *Statistical Confluence Analysis by Means of Complete Regression Systems*, Universitetets Økonomiske Institutt, Publikasjon nr. 5, Oslo.

Kalman, R. (1982): "System Identification from Noisy Data," in Dynamical Systems II, ed. by A. Bednarek and L. Cesari., Academic Press, New York.

Klepper, S. and E. Leamer (1984): "Consistent Sets of Estimates for Regressions with Errors in All Variables," Econometrica, 52, 163–183.

Klette, T. J. and Y. Willassen (1996): "Correlated Errors-in-Variables, too Few Instrumental Variables, and Bounds on Parameters." Memorandum nr. 30, Department of Economics, University of Oslo.

Koopmans. T. C. (1937): *Linear Regression Analysis of Economic Time Series*, Netherlands Economic Institute, Publication nr. 20.

Lach, S. (1993): "Decomposition of Variables and Correlated Measurement Errors," International Economic Review 34, 715–725.

Moran, P. A. P. (1971): "Estimating Structural and Functional Relationships," Journal of Multivariate Analysis, 1, 232–255.

Reiersøl, O. (1941): "Confluence Analysis by Means of Lag Moments and other Methods of Confluence Analysis," Econometrica 9, 1–24.

Reiersøl, O. (1945): *Confluence Analysis by Means of Instrumental Sets of Variables*, Arkiv för Matematik, Astronomi och Fysik, vol. 32 A, nr. 4.