Università Politecnica delle Marche

DIPARTIMENTO DI ECONOMIA

Identification of covariance structures

Riccardo Lucchetti

QUADERNI DI RICERCA n. 214

Luglio 2004

Abstract

The issue of identification of covariance structures, which arises in a number of different contexts, has been so far linked to conditions on the true parameters to be estimated. In this paper, this limitation is removed.

As done by Johansen (1995) in the context of linear models, the present paper provides necessary and sufficient conditions for the identification of a covariance structure that depend only on the constraints, and can therefore be checked independently of estimated parameters.

A sufficient condition is developed, which only depends on the structure of the constraints. It is shown that this *structure condition*, if coupled with the familiar order condition, provides a sufficient condition for identification. In practice, since the structure condition holds if and only if a certain matrix, constructed from the constraint matrices, is invertible, automatic software checking for identification is feasible even for large-scale systems.

JEL Class.: C13, C30

Indirizzo: Dipartimento di Economia, Università di Ancona. E-mail: jack@dea.unian.it

Identification of covariance structures*

Riccardo Lucchetti

1 Introduction

The aim of this paper is to shed some light on a problem that arises in models which impose some sort of structure on covariance matrices. This is the case of several popular models commonly employed in econometrics and multivariate statistics.

In all these models, it is assumed that a vector of n observable variables y_t can be thought of as some function of a vector of n unobservable variables u_t (called the *structural shocks*), which can be assumed independent, or at least uncorrelated, and to have unit variance.

Most models of this kind can be written as

$$A(\theta)y_t = B(\theta)u_t, \tag{1}$$

where A and B are square nonsingular matrices of order n, which are a function of the parameters to be estimated θ. We may define a *covariance structure* as a parametric model featuring two matrices A and B such that

$$A(\theta)\Sigma A(\theta)' = B(\theta)B(\theta)'. \tag{2}$$

where $\Sigma = V(y_t|\mathscr{F}_t)$, ie the covariance matrix of the observable variables y_t (possibly conditional on some information set $\mathscr{F}_t$). In order to avoid cumbersome notation, the reference to θ will be henceforth dropped, and A and B will be implicitly assumed to be functions of θ.

Examples include:

Structural VARs The methodology of structural VARs (or SVARs for short) was pioneered by Sims (1980) and then made popular by a countless number of applications, some of which highly influential (Blanchard

I would like to thank all the participants to the meeting held in Pavia on the 11th of June 2004 in honor of Carlo Giannini for their comments; it goes without saying that Carlo himself provided not only acute observations on the day, but also the main inspiration for this piece of work. The usual disclaimer obviously applies.

and Quah (1989) come to mind). The issues involved in the identification and estimation of such models were thoroughly investigated in Giannini (1992), and the present paper builds heavily on that work[1].

Multivariate GARCH models Factor models have been put forward rather often in order to overcome problems arising from the incredible number of parameters necessary for medium-scale models. One of the latest examples in this vein is Vrontos, Dellaportas, and Politis (2003).

LISREL models These models are widely used in sociology. For a survey, see Mueller (1996).

As equation (2) obviously implies

$$\Sigma = A^{-1}BB'(A')^{-1},$$

it is possible to define the vector function

$$\sigma(\theta) = \text{vech}\left[A^{-1}BB'(A')^{-1}\right]; \tag{3}$$

the difficulty in establishing whether such a model is identified arises because in most cases $\sigma(\theta)$ is a complex, highly nonlinear (albeit smooth) function of the vector θ.

In this paper, we will consider the cases when estimation of the parameters θ is carried out by optmising some objective function: typically, this is the log-likelihood (possibly concentrated), but not necessarily. The objective function which can be written as

$$\mathscr{L}(\mathbf{y}, \sigma(\theta)). \tag{4}$$

At the optimum of the objective function the following equation is satisfied:

$$\hat{A}\hat{\Sigma}\hat{A}' = \hat{B}\hat{B}', \tag{5}$$

where $\hat{\Sigma}$ is implicitly defined by $\hat{A}$ and $\hat{B}$ being optima. The identification question arises because there is no guarantee that $\hat{A}$ and $\hat{B}$ should be unique: in fact, equation (5) is satisfied by any pair of matrices A_1 and B_1 such that

$$A_1 = Q\hat{A} \tag{6}$$
$$B_1 = Q\hat{B}H \tag{7}$$

where Q is invertible and H is orthogonal.

[1] In 1997, a second edition of the same book was issued, co-authored by Gianni Amisano. Regrettably, it is marred by too many misprints to be cited as a reference.

It is clear that the matrices A_1 and B_1 can be considered an equivalent reparametrisation of the original model. The simplest case that can be taken as an example is when H is a permutation matrix. In this case, B_1 is simply $\hat{B}$ with its columns reordered, that is to say, the ordering of the structural shocks u_t in (1) is changed. The case when H or Q are arbitrary is nevertheless uninteresting, because the practically relevant issues arise when considering whether such matrices Q and H can exist in an arbirarily small neighbourood of the parameter space. In the terminology of Rothenberg (1971), therefore, we are dealing with *local*, rather than *global* identification.

In order to achieve identification, some constraints on A and B have therefore to be imposed. In this paper, we will follow a long-standing tradition of imposing a system of $p_a + p_b = p$ linear constraints[2]:

$$R_a \text{vec} A = R_a a = d_a \tag{8}$$
$$R_b \text{vec} B = R_b b = d_b; \tag{9}$$

(lowercase symbols will be used throughout to indicate vectorisations of matrices, ie $a \equiv \text{vec}A$). Alternatively, the constraints can be written in explicit form as follows

$$a = S_a \theta + s_a \tag{10}$$
$$b = S_b \theta + s_b \tag{11}$$

where θ are the unconstrained parameters, whose number equals $m = 2n^2 \cdot p$ (the number of elements in A and B minus the number of constraints). The columns of the matrix S_a and S_b form bases for the null space of the rows of R_a and R_b respectively, so that $R_a S_a = R_b S_b = 0$ and the matrices

$$[S_a \,|\, R_a'] \qquad \text{and} \qquad [S_b \,|\, R_b']$$

are square and invertible (the symbol "$|$" will be used troughout the paper to indicate horizontal stacking of matrices).

If a pair of matrices A and B satisfies the system of constraints (8)–(9), we call them *admissible*. The aim of this paper is to establish conditions under which an optimum of the objective function (4) corresponds to the only admissible pair within a given neighbourhood.

If the objective function (4) possesses derivatives up to the second order, identification can be investigated through definiteness of the Hessian matrix

[2]This excludes from the present setting multivariate techniques such as principal components analysis, since in those cases identification is typically achieved via nonlinear constraints. However, linear constraints are much more common, and possibly interesting, because they lend themselves more naturally to the representation of some economic theory: a typical example are zero restrictions on some parameters.

at the optimum. As a rule, however, the Hessian matrix is itself a function of the parameters to be estimated, so it cannot be computed before estimation is carried out, which renders this type of check non-operational[3]. As done by Johansen(1995) in the context of linear models, the present paper provides necessary and sufficient conditions for the identification of a covariance structure that depend only on the constraints, and can therefore be checked independently of estimated parameters.

The plan of the paper is as follows: section 2 provides a brief reminder of some mathematical concepts that will be used in the rest of the paper. Section 3 explores a special case in which the A matrix is fully restricted, in order to establish some concepts and exemplify them more clearly, while section 4 deals with to the general case; some examples are given in section 5.

2 Definitions and notation

In this section, we define some concepts and terms that will be used in the rest of the paper.

2.1 Decomposition of square matrices

As is well known (see for example Lütkepohl (1996)), any $n \times n$ matrix X can be written as

$$X = X_+ + X_-,$$

where X_+ is symmetric and X_- is hemisymmetric, and are defined as

$$\begin{aligned} X_+ &= \frac{1}{2}(X + X') \\ X_- &= \frac{1}{2}(X - X'). \end{aligned}$$

Let us now consider the space $\Omega = \mathbb{R}^{n^2}$: any element of this space can be considered the vectorisation of an $(n \times n)$ matrix. In a parallel fashion, the space Ω can be subdivided into two orthogonal subspaces Ω^+ and Ω^-: any vector $x \in \Omega$ can be written as

$$x = x_+ + x_-,$$

[3]A procedure is suggested in Giannini (1992), where identification is checked by computing the information matrix at a random point in the parameter space. It is argued that under-identification occurs on an area with zero measure, and thus the probability of making an incorrect decision is 0. This insight will be made more precise in section 3.3.

where $x_+ = \text{vec}(X_+) \in \Omega^+$ and $x_- = \text{vec}(X_-) \in \Omega^-$. It can be shown that Ω^+ has dimension $\frac{n(n+1)}{2}$ and Ω^- has dimension $\frac{n(n-1)}{2}$. We call D_n and $\widetilde{D}_n$ any matrices whose columns form bases for Ω^+ and Ω^- respectively, so that any symmetric matrix X has a vectorised form which satisfies $x = D_n\xi$ for some vector ξ and any hemisymmetric matrix X has a vectorised form which satisfies $x = \widetilde{D}_n\varphi$ for some vector φ.

One useful operator[4] in this context is the $n \times n$ matrix K_{nn}, which is defined by the property

$$K_{nn}\text{vec}(A) = \text{vec}(A').$$

It can be shown (see for instance Magnus and Neudecker (1988, pag. 46)) that K_{nn} is symmetric and orthogonal, ie

$$K_{nn} = K'_{nn} \qquad K_{nn}K'_{nn} = I.$$

2.2 The infinitesimal rotation operator

A matrix operator that will be useful in the rest of the paper is the so-called *infinitesimal rotation operator*[5]. Consider a vector x_0: we are interested in an infinitesimal displacement of x_0 which preserves its norm, ie a vector $x_1 = x_0 + \mathrm{d}x$ such that $\mathrm{d}\|x\| = 0$. Then, it is possible to define a matrix H as the matrix such that $x_1 = (I + H)x_0$. Two properties of H will be of interest:

1. $(I+H)$ must be orthogonal because $x_1'x_1 = x_0'(I+H)'(I+H)x_0 = x_0'x_0$ must hold for any x_0, and therefore $(I + H)'(I + H) = I$;

2. H is hemisymmetric because $\mathrm{d}(x'x) = x_0'\mathrm{d}x + (\mathrm{d}x')x_0 = 0$ and therefore $H = -H'$.

In this case, we say that the transformation is an infinitesimal rotation and H is the corresponding infinitesimal rotation operator.

2.3 Identification — generalities

We assume estimation of the parameters θ is carried out by optimising the following objective function over a parameter space Θ (typically, Θ coincides with $\mathbb{R}^k$):

$$\mathscr{L}(\mathbf{y}, \sigma(\theta));$$

[4]An alternate notation for K_{nn}, which is sometimes used, eg in Pollock (1979), is Ⓣ.

[5]For the properties of the infinitesimal rotation operator, see Weisstein (2004)

often the empirical covariance matrix of y_t, namely

$$\widehat{V} \equiv \frac{1}{T} \sum_{t=1}^{T} (y_t - \bar{y})(y_t - \bar{y})',$$

will serve as a sufficient statistic for θ, so that the objective function can be written as

$$\mathscr{L}(\widehat{v}, \sigma(\theta)),$$

where $\widehat{v} \equiv \text{vech}\widehat{V}$.

It will be assumed that the objective function possesses derivatives up to the second order and has at least one optimum $\hat{\theta}$. The covariance matrix of y_t, evaluated at $\hat{\theta}$ obeys

$$\hat{A}\hat{\Sigma}\hat{A}' = \hat{B}\hat{B}'.$$

where $\hat{A}$ and $\hat{B}$ are admissible[6].

Obvious examples include:

- Gaussian likelihood: $\mathscr{L}(y_t, \theta) = \text{const} - 1/2 \ln |\hat{\Sigma}| - 1/2\text{tr}(\hat{\Sigma}^{-1}\widehat{V})$
- GMM: $\mathscr{L}(y_t, \theta) = \delta'\Omega\delta$ where $\delta = \text{vech}(\hat{\Sigma} - \widehat{V})$.

In a locally identified model, if we treat $\mathscr{L}(\cdot)$ as a function of Σ, the optimiser $\hat{\Sigma}$ is unique in a neighbourhood of θ. Therefore, the only way for θ to move and still be optimal is when $\hat{\Sigma}$ is unchanged. In other words, the only way to go better (higher/lower) from $\hat{\Sigma}$ is to end up in a non-admissible point.

The first-order conditions for an optimum can be written in terms of Lagrange multipliers as

$$\frac{\partial \mathscr{L}}{\partial \sigma} + \lambda' \frac{\partial G}{\partial \sigma} = 0$$

where $G(\sigma)$ is the system of constraints induced on σ by the restrictions (8) and (9). Another equivalent way of expressing these conditions is to say that

$$\text{d}\mathscr{L} = \frac{\partial \mathscr{L}}{\partial \sigma} \frac{\partial \sigma}{\partial \theta} \text{d}\theta = J_{\sigma}^{\mathscr{L}} J_{\theta}^{\sigma} \text{d}\theta = 0$$

if the Jacobian matrix J_{θ}^{σ} has full column rank, then we have identification at the optimum. This, in turn, implies that the system is identified if, inside an ϵ-neighbourhood of $\hat{\theta}$, no other vector θ_1 (and a corresponding pair of admissible matrices A_1 and B_1) exists such that

$$A_1 \hat{\Sigma} A_1' = B_1 B_1'.$$

[6] In many cases, $\hat{\Sigma}$ coincides with $\widehat{V}$; this need not be the case, however. One may think of $\hat{\Sigma}$ and $\widehat{V}$ as structural- and reduced-form estimates respectively.

Obviously, identification imposes an order condition: since the number of elements in σ is $\frac{n(n+1)}{2}$, the column rank of the Jacobian matrix J^{σ}_{θ} cannot be full if θ has more than $\frac{n(n+1)}{2}$ elements. As is well known, however, this is only a necessary condition. Another necessary condition, called the *structure condition*, which covers some cases of interest not covered by the order condition and is based only on the constraint matrices will be developed in the next subsections. It will be argued that a sufficient condition for identification can be obtained by requiring that both order and structure conditions hold.

3 Identification in the C-model

We begin by examining the special case in which $A = I$. In the terminology introduced in Giannini (1992), which will be adopted from here on, this corresponds to the so-called C-model. In this section, therefore, we analyse the special case where

$$\begin{aligned} R_a &= I \\ s_a &= \text{vec} I \end{aligned}$$

and S_a does not exist, since R_a is full rank and there are no elements of A which depend on unknown parameters. All the m free parameters are contained in the matrix B. To avoid clutter, we might as well forget about the matrix A altogether and assume, in this section, that $n^2 = p + m$; consequently, the index $_b$ will be dropped from all the constraint matrices.

The structure of interest here is

$$\hat{\Sigma} = B_0 B_0'.$$

Consider now another admissible matrix, $B_1 = B_0 + \mathrm{d}B$, where $\mathrm{d}B$ is infinitesimal. The model is under-identified if B_1 is observationally equivalent to B_0, ie if

$$B_1 B_1' = \hat{\Sigma} = B_0 B_0'. \tag{12}$$

If (12) holds, it is possible to define an infinitesimal rotation matrix H as the matrix such that $B_1 = B_0(I + H)$. From its definition, the matrix B_1 can be written in vectorised form as follows

$$B_1 = B_0(I + H) \Longrightarrow b_1 = b_0 + (H' \otimes I) b_0$$

If B_1 is to be admissible, then we must have

$$R' b_1 = d;$$

however, this implies

$$R'b_0 + R'(H' \otimes I)b_0 = d$$

and therefore

$$R'(H' \otimes I)b_0 = R'(H' \otimes I)(S\theta_0 + s) = 0; \tag{13}$$

The above reasoning could have been equivalently, and certainly more compactly, put by requiring that, for $B_0 + \mathrm{d}B$ to be admissible, the condition

$$R'\mathrm{d}b = 0 \tag{14}$$

has to be satisfied, where $\mathrm{d}B = BH$ for some hemisymmetric H, though possibly this would have come to the expense of clarity.

Consider now the i-th row of $R'(H' \otimes I)$. If we call e_i the i-th column of the identity matrix, we have

$$e_i' R(H' \otimes I) = h'(I \otimes R_i');$$

where $h = \mathrm{vec}H$ and R_i is an $n \times n$ matrix whose vectorisation is the i-th row of R. Moreover, in order to be an infinitesimal rotation matrix, H must be hemisymmetric. Therefore, it is possible to write h as

$$h = \widetilde{D}_n \varphi,$$

where $\widetilde{D}_n$ is any basis for Ω^-. Therefore, condition (13) can be rephrased as

$$\varphi' \widetilde{D}_n'(I \otimes R_i')[S\theta + s] = \varphi' T_i \theta + \varphi' t_i = 0 \quad \text{for}\, i = 1 \ldots p \tag{15}$$

where

$$\begin{aligned} T_i &\equiv \widetilde{D}_n'(I \otimes R_i')S \\ t_i &\equiv \widetilde{D}_n'(I \otimes R_i')s \end{aligned}$$

and p is the number of constraints (the number of rows in R). In other words, the system is unidentified at θ if some $\varphi \neq 0$ exists which satisfies the above equations for every i. It is important to note that the existence of non-zero solutions to the system (15) (and therefore under-identification) depends only on the matrices T_i and the vectors t_i, *which are functions of the constraints alone*.

Giannini (1992), in his analysis of the C-model in a structural VAR context, indicates (page 28) that identification holds if and only if the the matrix

$$R(I \otimes B)\widetilde{D}_n$$

has full column rank[7]. This condition was obtained by considering maximum likelihood estimation, since it ensures that the information matrix is positive definite. It is easy to see that this condition is exactly equivalent to requiring that equation (13) — and hence system (15) — has no nontrivial solutions.

In general, if a solution to the system (15) $\varphi \neq 0$ exists, it may depend on θ. However, some special cases can be considered.

3.1 The order condition

First, the order condition can be established: for any θ, equation (15) can be written as

$$\varphi' \left[\; T_1\theta + t_1 \;\middle|\; T_2\theta + t_2 \;\middle|\; \cdots \;\middle|\; T_p\theta + t_p \; \right] = [0]; \tag{16}$$

since the matrix in square brackets is $(\frac{n(n-1)}{2} \times p)$, its rows span a space of dimension $\min(p, \frac{n(n-1)}{2})$ at most, so if $p < \frac{n(n-1)}{2}$ its right null space has (at least) dimension $\left(p - \frac{n(n-1)}{2}\right)$: therefore, there is always a nonzero φ which satisfies equation (15). This implies that the model is certainly under-identified if the number of constraints p is less than $\frac{n(n-1)}{2}$, which means that the number of free parameters m cannot exceed $\frac{n(n+1)}{2}$. This is, again, the familiar order condition.

3.2 The structure condition

Additionally, there might be cases when (13) holds for any value of $\theta \in \Theta$ even though the order condition is met. We will call *structurally under-identified* such a model: in these cases there exists at least one H that satisfies equations (12) and (13) whatever the choice of B_0. If identification of the structure does not depend on the particular value of B_0, then equation (13) must be satisfied for any θ_0; therefore, it must also be true that

$$R'(H' \otimes I)[S \mid s] = 0 \tag{17}$$

for some infinitesimal rotation H.

By employing the properties of the Kronecker product and of the vectorization operator in a manner similar to the one used in the derivation of eq. (15), it is possible to write the i-th column of $(H' \otimes I)S$ as

$$(H' \otimes I)Se_i = (I \otimes S_i)h;$$

[7]The original notation was slightly altered to match ours.

where S_i is an $n \times n$ matrix whose vectorisation is the i-th column of S. Similarly, we can transform $(H' \otimes I)s$ into

$$(H' \otimes I)s = (I \otimes \bar{S})h,$$

where $\text{vec}(\bar{S}) = s$.

The problem can now be stated as follows: given the matrix

$$T = \begin{bmatrix} R'(I \otimes S_1) \\ R'(I \otimes S_2) \\ \vdots \\ R'(I \otimes S_m) \\ R'(I \otimes \bar{S}) \end{bmatrix} \tag{18}$$

if there is an infinitesimal rotation H such that $Th = 0$, then the system is structurally unidentified[8]; in turn, this implies the existence of a non-null vector φ such that $T\widetilde{D}_n\varphi = 0$ As a consequence, *a C-model is structurally identified if and only if the matrix* $T\widetilde{D}_n$ *has full column rank* $\frac{n(n-1)}{2}$. This condition will be henceforth called *structure condition.*

The structure condition can be more easily checked via the matrix

$$\mathcal{M} = \widetilde{D}_n' T' T \widetilde{D}_n = \sum_{i=1}^{m} \left[\widetilde{D}_n'(I \otimes S_i')RR'(I \otimes S_i)\widetilde{D}_n \right] + \widetilde{D}_n'(I \otimes \bar{S}')RR'(I \otimes \bar{S})\widetilde{D}_n$$

The system is structurally identified if and only if $\mathcal{M}$ is invertible. If not, the matrix H has a vectorisation which lies in the right null space of $T\widetilde{D}_n$.

3.3 Sufficiency

Both the order and structure condition conditions are, by themselves, necessary but not sufficient. However, the order condition and the structural identification condition form a quasi-sufficient condition when taken together: *if the equations (15) are satisfied for some* $\varphi \neq 0$ *beyond the trivial case of* $p < \frac{n(n-1)}{2}$*, then the values of* θ *that satisfy (15) define a set of measure 0, and therefore the model is identified almost everywhere in* Θ*.*

This assertion can be proven by a line of reasoning similar to that in Johansen (1995, p. 130): consider the matrix between square brackets in

[8]Notice that each row of the matrix T can be written as $\text{vec}(S_i'R_j)'$ for all possible combinations of the columns of R and $[S\,|\,s]$. This may help computationally. Another point that may help in a practical software implementation is that, as a rule, a large number of rows in T are zero, and clearly can be left out without influencing the rank of T.

(16) as a function of θ. For (16) to have non-zero solutions, this matrix must have rank less than $\frac{n(n-1)}{2}$, and therefore the determinant of the following matrix:

$$\begin{aligned} D &= \left[\ T_1\theta + t_1 \mid T_2\theta + t_2 \mid \cdots \mid T_p\theta + t_p\ \right] \begin{bmatrix} \theta' T_1' + t_1' \\ \theta' T_2' + t_2' \\ \vdots \\ \theta' T_p + t_p' \end{bmatrix} = \\ &= \sum_{j=1}^{p} (T_1\theta + t_1)(T_1\theta + t_1)' \end{aligned}$$

must be 0. Now the equation $|D| = 0$ is a polynomial in θ, and therefore is either identically 0, or has a set of solutions which forms a closed set in Θ which has zero Lebesgue measure. The former can happen in two cases: either the number of columns is less than $\frac{n(n-1)}{2}$ (which takes us back to the order condition) or equation (16) is satisfied for any θ. But if this is the case, then the model fails to meet the structure condition; as a consequence, if both conditions hold (16) has no solutions in Θ but for a set of measure 0. Moreover, this set is closed, and the set of points in Θ where the model is identified if and only if both the order and the structure conditions hold is open and dense in Θ.

3.4 An example

Let us analyse a C-model where $n = 2$ and B is lower-triangular. As is well known, this case is identified, as B is simply the Cholesky decomposition of Σ. In this case, the number of free parameters is 3 and the parameter space Θ coincides with $\mathbb{R}^3$. The matrix B has the form

$$B = \begin{bmatrix} \theta_1 & 0 \\ \theta_2 & \theta_3 \end{bmatrix}$$

so that the constraint matrices can be written as:

$$\begin{aligned} R &= \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \\ d &= 0 \end{aligned}$$

and the corresponding S matrix is

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(the vector s equals 0); as for $\widetilde{D}_n$, the only possible choice (up to a scalar) when $n = 2$ is the vector

$$\widetilde{D}_n = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}.$$

The order condition is evidently satisfied, as the number of free parameters equals $\frac{n(n+1)}{2} = 3$. The structure condition can be checked via

$$\mathcal{M} = \widetilde{D}_n' \left[\sum_{i=1}^{3} (I \otimes S_i) R' R (I \otimes S_i') \right] \widetilde{D}_n = 1,$$

and therefore the model is identified almost everywhere in $\mathbb{R}^3$.

In this case, the dimension of the matrices is small enough to make it feasible to check identification by developing the argument analytically, which is rather instructive. The structure of constraints shows clearly that B is admissible if and only if it is lower-triangular. The identification problem then reduces to establishing whether postmultiplying B by an arbitrary infinitesimal rotation may result in another lower triangular matrix. In formulae:

$$\begin{bmatrix} \theta_1 & 0 \\ \theta_2 & \theta_3 \end{bmatrix} \begin{bmatrix} 1 & -\varphi \\ \varphi & 1 \end{bmatrix} = \begin{bmatrix} \theta_1 & -\theta_1 \varphi \\ \theta_2 + \theta_3 \varphi & \theta_3 - \theta_2 \varphi \end{bmatrix}$$

The resulting matrix is clearly not admissible (ie lower triangular) unless $\theta_1 \varphi = 0$. If $\varphi \neq 0$, then θ_1 must be 0.

The same result can be obtained by considering directly condition (15); since we have only one constraint, then we have only one T_i matrix and one t_i vector, and these equal:

$$\begin{aligned} T_1 &= \widetilde{D}_n' (I \otimes R_1') S \\ &= \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix} \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \end{aligned}$$

whereas $t_1 = 0$.

The model is under-identified as long as there exist $\varphi \neq 0$ and $\theta \neq 0$ that verify

$$\varphi' \widetilde{D}_n' (I \otimes R_1')(S\theta + s) = \varphi \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \varphi \cdot \theta_1 = 0.$$

If φ has to be non-zero, the previous equation has the only solution $\theta_1 = 0$. This implies that the only region in Θ where the model is under-identified is the plane $\theta_1 = 0$, which has zero Lebesgue measure in $\mathbb{R}^3$. As a consequence, the model is identified almost everywhere[9] in Θ.

4 Identification in the AB-model

This case is the most general: contrary to the previous section, we analyse the situation where neither A nor B are fully restricted, so the identification question revolves around establishing the existence of two matrices $(A + \mathrm{d}A)$ and $(B + \mathrm{d}B)$ which are still admissible.

The constraints to be put on $A + \mathrm{d}A$ and $B + \mathrm{d}B$ to remain admissible can be specified in a way akin to equation (14) by writing

$$\begin{aligned} A + \mathrm{d}A &= (I + Q)A \\ B + \mathrm{d}B &= (I + Q)B(I + H), \end{aligned}$$

where $(I + Q)$ is nonsingular and $(I + H)$ is orthogonal, so that

$$\mathrm{d}a = (I \otimes Q)(S_a\theta + s_a) \quad (19)$$

$$\mathrm{d}b = \left[(I \otimes Q) + (H' \otimes I)\right](S_b\theta + s_b) \quad (20)$$

(in equation (20), there would be a term $(H' \otimes Q)$ which, however, disappears because both matrices are infinitesimal).

Like in the previous section, H has to be an infinitesimal rotation for $(I + H)$ to be orthogonal. The only additional complication with respect to section 3 is that it is now necessary to take into consideration the matrix Q as well. Moreover, Q is not necessarily an infinitesimal rotation operator, since $I + Q$ need not be orthogonal, although it is required that it is invertible. However, if we only consider what happens in a neighbourhood of the optimum, it suffices to say that $|I + Q| \neq 0$ for any Q, as long as Q is infinitesimal, by the continuity of the determinant function. Therefore, if we define $q = \mathrm{vec}Q$, it is sufficient to consider the condition $q \neq 0$ without any further qualifications.

The equivalent of equation (14) is the following system:

$$R_a'\mathrm{d}a = R_a'(I \otimes Q)(S_a\theta + s_a) = 0 \quad (21)$$

$$R_b'\mathrm{d}b = R_b'\left[(I \otimes Q) + (H' \otimes I)\right](S_b\theta + s_b) = 0; \quad (22)$$

[9] It is interesting to note that, for $\theta_1 = 0$, the B matrix is singular by construction.

we may re-express some of the matrix products as

$$\begin{aligned} e_i'R_a'(I \otimes Q) &= q'K_{nn}(R_{a,i} \otimes I) \\ e_i'R_b'(I \otimes Q) &= q'K_{nn}(R_{b,i} \otimes I) \\ e_i'R(H' \otimes I) &= \varphi'\widetilde{D}_n(I \otimes R_i') \end{aligned}$$

because $K_{nn}q = \text{vec}(Q')$ (see subsection 2.1) and $h = \widetilde{D}_n\varphi$. Such rearrangments lead us to examine the following system of equations:

$$\begin{aligned} q'U_i^a\theta + q'u_i^a &= 0 \quad \text{for } i = 1 \dots p_a \\ q'U_j^b\theta + \varphi'T_j^b\theta + q'u_j^b + \varphi't_j^b &= 0 \quad \text{for } j = 1 \dots p_b \end{aligned}$$

where

$$\begin{aligned} U_i^a &\equiv K_{nn}(R_{a,i}' \otimes I)S_a & u_i^a &\equiv K_{nn}(R_{a,i}' \otimes I)s_a \\ U_i^b &\equiv K_{nn}(R_{b,i}' \otimes I)S_b & u_i^b &\equiv K_{nn}(R_{b,i}' \otimes I)s_b \\ T_i^b &\equiv \widetilde{D}_n'(I \otimes R_{b,i}')S_b & t_i^b &\equiv \widetilde{D}_n'(I \otimes R_{b,i}')s_b \end{aligned}$$

which admits non-zero solutions if, and only if, the model is under-identified.

4.1 The order and structure conditions

The order and structure conditions can be stated in a way very similar to section 3.

Again, the order condition requires that the number of free parameters does not exceed $\frac{n(n+1)}{2}$, or equivalently, that the number of restrictions put on A and B is at least $n^2 + \frac{n(n-1)}{2}$. If this requirement were not met, then the following equation

$$\left[\begin{array}{c|c} q' & \varphi' \end{array}\right] \left[\begin{array}{c|c|c|c|c|c} U_1^a\theta + u_1^a & \cdots & U_{p_a}^a\theta + u_{p_a}^a & U_1^b\theta + u_1^b & \cdots & U_{p_b}^b\theta + u_{p_b}^a \\ \hline 0 & \cdots & 0 & T_1^b\theta + t_1^b & \cdots & T_{p_b}^b\theta + t_{p_b}^a \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \tag{23}$$

would always have solutions for some nonzero q and/or φ.

The structure condition requires that there are no infinitesimal matrices Q and H (with H orthogonal) that satisfy

$$R_a'(I \otimes Q)[S_a|s_a] = 0 \tag{24}$$

$$R_b'\left[(I \otimes Q) + (H' \otimes I)\right][S_b|s_b] = 0. \tag{25}$$

By considering the columns of S_a and S_b one at a time, we get

$$R_a'(I \otimes Q)[S_a|s_a]e_i = R_a'(S_{ai}' \otimes I)q = 0 \quad \forall i$$

and

$$R'_b\left[(I \otimes Q) + (H' \otimes I)\right][S_b|s_b]e_i = R'_b(S'_{bi} \otimes I)q + R'_b(I \otimes S_{bi})\widetilde{D}_n\varphi = 0 \quad \forall i$$

so that the question reduces to the existence of non-trivial solutions to this system, which evidently parallels the problem of finding solutions to equation (17) in subsection 3.2.

The matrix T equivalent to the one in equation (18) becomes

$$T = \left[\begin{array}{c|c} R'_a(S'_{a1} \otimes I) & 0 \\ R'_a(S'_{a2} \otimes I) & 0 \\ \vdots & \vdots \\ R'_a(S'_{ap} \otimes I) & 0 \\ R'_a(\bar{S}'_a \otimes I) & 0 \\ \hline R'_b(S'_{b1} \otimes I) & R'_b(I \otimes S_{b1})\widetilde{D}_n \\ R'_b(S'_{b2} \otimes I) & R'_b(I \otimes S_{b2})\widetilde{D}_n \\ \vdots & \vdots \\ R'_b(S'_{bq} \otimes I) & R'_b(I \otimes S_{bq})\widetilde{D}_n \\ R'_a(\bar{S}'_b \otimes I) & R'_b(I \otimes \bar{S}_b)\widetilde{D}_n \end{array}\right] = \left[\begin{array}{cc} U_a & 0 \\ U_b & T_b\widetilde{D}_n \end{array}\right] \tag{26}$$

and the system is structurally identified provided there are no trivial solutions to

$$\left[\begin{array}{cc} U_a & 0 \\ U_b & T_b\widetilde{D}_n \end{array}\right]\left[\begin{array}{c} q \\ \varphi \end{array}\right] = 0. \tag{27}$$

As a consequence, an operational procedure for checking the structure condition could simply amount to verifying whether the matrix

$$\mathcal{M} = \left[\begin{array}{cc} U'_a & U'_b \\ 0 & \widetilde{D}'_n T'_b \end{array}\right]\left[\begin{array}{cc} U_a & 0 \\ U_b & T_b\widetilde{D}_n \end{array}\right] = \left[\begin{array}{cc} U'_aU_a + U'_bU_b & U'_bT_b\widetilde{D}_n \\ \widetilde{D}'_nT'_bU'_b & \widetilde{D}'_nT'_bT_b\widetilde{D}_n \end{array}\right] \tag{28}$$

is singular[10].

The quasi-sufficiency property of the order and structure conditions combined can be assessed by means of an argument similar to that developed in subsection 3.3 by noting that the rank of the matrix in equation (23) is less than $n^2 + \frac{n(n-1)}{2}$ either identically, or for a set of measure 0 in Θ.

[10] An ox class for checking the structure condition will be available shortly at http://www.econ.univpm.it/lucchetti/.

4.2 Special cases

In certain cases, checking the structure condition involves simpler matrices: the C-model, for example, clearly emerges as a special case: if we impose the constraint $A = I$, then we have

$$\begin{aligned} R_a &= I \\ [S_a | s_a] &= \text{vec} I \end{aligned}$$

and therefore the matrix T reduces to

$$T = \left[\begin{array}{c|c} I & 0 \\ \hline U_b & T_b \widetilde{D}_n \end{array} \right],$$

which clearly has full column rank if and only if its south-east corner has. This, in turn, is precisely the matrix T in equation (18).

Another notable special case arises when it is assumed that $A\Sigma A' = I$. This case is referred to as the K-model in Giannini (1992), and, since $A\Sigma A' = I$ implies $A'A = \Sigma^{-1}$, it could be conjectured that the requisites for structural identification ought to be similar to those in the C-model. This is indeed the case, as the restrictions put on B parallel those previously put on A:

$$\begin{aligned} R_b &= I \\ [S_b | s_b] &= \text{vec} I; \end{aligned}$$

as a consequence, the matrix T becomes

$$T = \left[\begin{array}{c|c} U_a & 0 \\ \hline I & \widetilde{D}_n \end{array} \right],$$

and equation (27) becomes

$$\left[\begin{array}{cc} U_a & 0 \\ I & \widetilde{D}_n \end{array} \right] \left[\begin{array}{c} q \\ \varphi \end{array} \right] = 0;$$

which implies $q = -\widetilde{D}_n \varphi$; therefore, if a non-zero solution exists, the equality

$$U_a \widetilde{D}_n \varphi = 0 \tag{29}$$

must hold. This happens only if the rank of $U_a \widetilde{D}_n$ is less than $\frac{n(n-1)}{2}$.

5 A few examples

5.1 An unidentified 2 × 2 case

This is an example of a model reported in Giannini (1992), where it is shown that the order condition is not sufficient, by itself, to ensure identification. In the context of the present paper, it provides a simple yet enlightening example of failure of the structure condition. We have $n = 2$, and A has the following structure:

$$A = \begin{bmatrix} \theta_1 & \theta_2 \\ -\theta_2 & \theta_1 \end{bmatrix};$$

in other words, any (2×2) admissible matrix has the same elements on the diagonal, whereas the off-diagonal elements have opposite sign.

The fact that $B = I$ by hypothesis allows us to classify this model as a K-model, and thus focus on (29) for checking the structure condition, rather than on the more complex equation (28).

The constraint matrices can be written as

$$\begin{aligned} R_a &= \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \\ S_a &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

and both d and s are vectors of zeros.

The order condition is obviously met, since the number of constraints (two), is greater than $\frac{n(n-1)}{2} = 1$. As far as the structure condition goes, we can use equation (29); using the same $\widetilde{D}_n$ as in subsection 3.4, we have

$$\mathcal{M} = \widetilde{D}'_n \left[\sum_{i=1}^{2} (S_i \otimes I) R' R (S'_i \otimes I) \right] \widetilde{D}_n = 0,$$

which has rank 0, and therefore the model is structurally unidentified.

In this case, the matrices are small enough to analyze symbolically the effect of an infinitesimal rotation on the matrix A in some detail. Since any hemisymmetric matrix of order 2 can be written as

$$H = \begin{bmatrix} 0 & \lambda \\ -\lambda & 0 \end{bmatrix},$$

we have that the product $A_1 = A(I + H)$ produces

$$A(I+H) = \begin{bmatrix} \theta_1 & \theta_2 \\ -\theta_2 & \theta_1 \end{bmatrix} \begin{bmatrix} 1 & \lambda \\ -\lambda & 1 \end{bmatrix} = \begin{bmatrix} \theta_1 - \lambda\theta_2 & \theta_2 + \lambda\theta_1 \\ -(\theta_2 + \lambda\theta_1) & \theta_1 - \lambda\theta_2 \end{bmatrix}$$

which is clearly admissible. Therefore, there is an infinity of admissible matrices A_1 which satisfy $A_1' A_1 = A'A = \hat{\Sigma}^{-1}$.

5.2 The "standard" AB model

The model analysed here corresponds to the most common setup in structural VARs: the one where A is lower-triangular with ones on the diagonal and B is diagonal.

$$A = \begin{bmatrix} 1 & 0 \\ \theta_1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} \theta_2 & 0 \\ 0 & \theta_3 \end{bmatrix}.$$

This example is admittedly rather contrived, as this model could be easily reparametrised into a K-model, but here it just serves the purpose of providing a simple example of constraints put on both A and B.

The constraint matrices are:

$$R_a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad [S_a \mid s_a] = \left[\begin{array}{c|c} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{array}\right]$$

$$R_b = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad [S_b \mid s_b] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right]$$

The matrix $\mathcal{M}$, computed via equation (28), equals

$$\mathcal{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

Since $\mathcal{M}$ is invertible, the system is structurally identified. Therefore, both order and structure conditions hold, and identification is attained.

6 Conclusions and directions for future research

The main object of this paper is to put forth a method for assessing identification of a covariance structure. As such, the ideas presented here are readily applicable to a number of different contexts, such as the specification and estimation of structural VAR or multivariate GARCH models.

However, some points may deserve further investigation. First, it might be interesting to analyse the relationship between the set of points in the parameter space which give rise to singular A and/or B matrices and the set where identification fails, despite the fact that both the order and structure conditions are satisfied. We do know that both are sets with 0 Lebesgue measure. Furthermore, the example presented in section 3.4 shows a case where the latter is a strict subset of the former. It is natural to ask if anything more stringent can be said in general.

Another extension of the present results that may be useful in applied contexts would be incorporating into the present framework the interesting approach to identification of linear systems through heteroskedasticity recently proposed by Rigobon (2003). Research by the author is currently under way in this direction.

Finally, we like to mention the possibility of generalising the present approach to nonlinear contraints. Since all the arguments we presented applied to an arbitrarily small neighbourhood of the parameter space, perhaps some kind of linearisation of the constraints might be attempted.

References

Blanchard, O., and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Aggregate Supply Shocks," *American Economic Review*, 79(4), 655–73.

Giannini, C. (1992): *Topics in structural VAR econometrics*. Springer-Verlag, 1st edn.

Johansen, S. (1995): "Identifying restrictions of linear equations with applications to simultaneous equations and cointegration," *Journal of Econometrics*, 69(1), 112–132.

Lütkepohl, H. (1996): *Handbook of Matrices*. John Wiley & Sons.

Magnus, J., and H. Neudecker (1988): *Matrix Differential Calculus.* John Wiley & Sons.

Mueller, R. O. (1996): *Basic Principles of Structural Equation Modelling: An Introduction to LISREL and EQS.* Springer Verlag.

Pollock, D. (1979): *The Algebra of Econometrics.* John Wiley & Sons.

Rigobon, R. (2003): "Identification Through Heteroskedasticity," *Review of Economics and Statistics*, 85(4), 777–92.

Rothenberg, T. (1971): "Identification in Parametric Models," *Econometrica*, 39, 577–591.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

Vrontos, I., P. Dellaportas, and D. Politis (2003): "A full-factor multivariate GARCH model," *Econometrics Journal*, 6, 312–334.

Weisstein, E. W. (2004): "Infinitesimal Rotation," MathWorld – A Wolfram Web Resource. http://mathworld.wolfram.com/InfinitesimalRotation.html.