Economics Letters 144 (2016) 107–111

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

A topological view on the identification of structural vector autoregressions

Klaus Neusser [1]

Department of Economics, University of Bern, Schanzeneckstrasse 1, P.O. Box 8573, CH-3001 Berne, Switzerland

ARTICLE INFO

Article history:
Received 15 March 2016
Accepted 4 May 2016
Available online 12 May 2016

JEL classification:
C1
C18
C32

Keywords:
SVAR
Identification
Group action
Haar measure
Perturbation

ABSTRACT

The notion of the group of orthogonal matrices acting on the set of all feasible identification schemes is used to characterize the identification problem arising in structural vector autoregressions. This approach presents several conceptual advantages. First, it provides a fundamental justification for the use of the normalized Haar measure as the natural uninformative prior. Second, it allows to derive the joint distribution of blocks of parameters defining an identification scheme. Finally, it provides a coherent way for studying perturbations of identification schemes which becomes relevant, among other things, for the specification of vector autoregressions with time-varying covariance matrices.

1. Introduction

Structural vector autoregressive (SVAR) models have established themselves as an indispensable tool in empirical macroeconomics. While these models capture reasonably well the dynamic properties of the data, their economic interpretation in terms of structural shocks is discussed controversially because these models suffer from a fundamental identification problem. This problem is addressed by imposing some restrictions (short-run, long-run, sign restrictions, etc.) which are more or less funded in a priori economic reasoning. The econometric aspects of the identification problem has been analyzed by Rubio-Ramírez et al. (2010) and Waggoner and Zha (2003) in the spirit of Rothenberg (1971).

We view the identification problem as an invariance property of the group of orthogonal matrices on the set of observationally equivalent identification schemes. While already anticipated in the previously mentioned papers, following this route rigorously presents several advantages. First, the identification problem is given a precise mathematical framework. In this framework, the invariance principle naturally leads to the use of the normalized Haar measure as an uninformative prior (Jaynes, 1968). Second, it allows the derivation of the *joint* distribution of the impact effects and not just of a single coefficient as in Baumeister and Hamilton (2015, Section 3). Third, the action of the group allows to conceive a kind of perturbation analysis of the identification scheme. This is not only interesting in itself, but can be used to formulate time-varying covariance matrices in a coherent way.

2. Structural vector autoregressive models

Consider a vector autoregressive (VAR) processes $\{X_t\}$ with observations in the state space $\mathbb{R}^n$ and defined as the stationary solution of the stochastic difference equations of order p with constant coefficients $\Phi_1, \ldots, \Phi_p$:

$$X_t = \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + Z_t, \quad Z_t \sim \mathrm{WN}(0, \Sigma), \tag{2.1}$$

where Σ is symmetric and positive definite. The reduced form shocks Z_t are obtained from the structural shocks $\{V_t\}$ by a linear weighting scheme

$$Z_t = B' V_t, \quad V_t \sim \mathrm{WN}(0, I_n), \tag{2.2}$$

where the $n \times n$ matrix B is left unrestricted. The uncorrelatedness assumption of the structural shocks is very much accepted in the literature. Otherwise, there would remain some unexplained relationship between them. The assumption that the structural shocks have a covariance matrix equal to the identity is just a convenient normalization.

E-mail address: klaus.neusser@vwi.unibe.ch.

[1] I thank Katharina Neusser for her perceptive comments.

http://dx.doi.org/10.1016/j.econlet.2016.05.003
0165-1765/

Although this is not necessary for the discussion, it will be assumed that $\{X_t\}$ admits a causal representation with respect to $\{Z_t\}$. Thus, there exists a sequence of matrices $\{\Psi_j\}$, $j = 0, 1, 2, \ldots$, with $\Psi_0 = I_n$ and $\sum_{j=0}^{\infty} \|\Psi_j\| < \infty$ such that

$$X_t = \Psi_0 Z_t + \Psi_1 Z_{t-1} + \Psi_2 Z_{t-2} + \cdots = \sum_{j=0}^{\infty} \Psi_j Z_{t-j} = \Psi(\mathrm{L}) Z_t \quad (2.3)$$

$$= \Psi_0 B' V_t + \Psi_1 B' V_{t-1} + \Psi_2 B' V_{t-2} + \cdots$$

$$= \sum_{j=0}^{\infty} \Psi_j B' V_{t-j} = \Psi(\mathrm{L}) B' V_t. \quad (2.4)$$

Such a causal representation of X_t in terms of current and past Z_t's exists if and only if $\det(I_n - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p) = \det \Phi(z) \neq 0$ for all $z \in \mathbb{C}$ with $|z| \leq 1$.

While the VAR, usually, gives a good summary of the data, at least up to the second moments, it is just a preliminary first step in the analysis. The second and more controversial step aims at identifying the structural shocks $\{V_t\}$ and their effects on X_{t+j}, $j = 0, 1, 2, \ldots$. These effects are propagated over time and captured by the sequence $\{\Psi_j B'\}$, $j = 0, 1, 2, \ldots$, known as the impulse response function. The shocks and their propagation are usually given an economic interpretation and are at the core of the SVAR approach.

Relying on second moments only or assuming a Gaussian framework, it is easy to see that the simultaneous equation system (2.2) is not identified, i.e. it is impossible to extract B just from the knowledge of Z_t alone.[2] Indeed, taking the symmetry of covariance matrices into account, the nonlinear equation system

$$\Sigma = \mathbb{E}(Z_t Z_t') = \mathbb{E}(B' V_t V_t' B) = B' B \quad (2.5)$$

delivers only $n(n+1)/2$ independent equations for n^2 unknown coefficients in B. Thus, there is a need of $n^2 - n(n+1)/2 = n(n-1)/2$ additional equations. A customary solution to the underidentification problem is to place enough restrictions on the matrix B so that the equation system (2.5) admits a unique solution. We call these identifying restrictions an *identification scheme*.

One popular form of restrictions is to set some coefficients a priori to zero. These, so-called, short-run restrictions have to come from either additional, usually theoretical, reasoning or other a priori reasoning and are subject to controversy. Another common way to place restrictions on B is to assume that the cumulated effects of some particular shocks on some variable equals zero. Thus, these so-called long-run restrictions impose zeros on $\Psi(1)B$. Obviously, short- and long-run restriction do not exclude each other, but can complement each other. As the gap between equations and unknowns grows quadratically, it becomes more and more difficult to incorporate reasonable restrictions as the dimension n of the VAR increases.[3]

3. An algebraic interpretation of the identification problem

3.1. The group action of orthogonal matrices

One aim of this paper is to provide a deeper conceptual framework which in the end should allow a better understanding of the fundamental identification problem and of the solution techniques proposed in this context.

Before presenting some results it is necessary to introduce some algebraic and topological notions. Let $\mathbb{M}_n$ be the vector space of $n \times n$ matrices with real entries. It is clear that to any matrix $A = A_{ij} \in \mathbb{M}_n$ we can associate a point in $\mathbb{R}^{n^2}$ and hence identify the vector space $\mathbb{M}_n$ with $\mathbb{R}^{n^2}$. In this way, $\mathbb{M}_n$ can be equipped with the Euclidian metric of $\mathbb{R}^{n^2}$. With respect to this metric, the usual matrix operations are continuous and even smooth.[4] Since $\det A$ is a continuous function from $\mathbb{M}_n$ to $\mathbb{R}$, the set of invertible matrices is an open subset of $\mathbb{M}_n$ which forms a group with respect to the matrix multiplication. This group is called the general linear group and denoted by $\mathbb{GL}_n$.

In the following, the subgroup of orthogonal matrices $\mathbb{O}_n$, i.e. matrices Q with the property $Q'Q = I_n$, will be of special interest. It can be shown that $\mathbb{O}_n$ is a compact (closed and bounded) subgroup of $\mathbb{GL}_n$.[5] This implies that $\mathbb{O}_n$ has a finite Haar measure (see Diestel and Spalsbury, 2014, Chapter 5). This measure can be normalized to make it a probability distribution.[6]

This distribution can be efficiently implemented numerically by applying the QR-decomposition to a random matrix A with law $\mathcal{L}(A) = \mathrm{N}(0, I_n \otimes I_n)$, i.e. the elements of A are i.i.d. $N(0, 1)$ random variables (see Birkhoff and Gulati, 1979; Stewart, 1980; Edelman and Raj Rao, 2005, for details).

In Section 3.2, we derive an analytic expression for the density of subblocks of the normalized Haar measure on $\mathbb{O}_n$. This result will then be used to derive a corresponding result for the identification schemes. For this purpose, we define the set of conceivable identification schemes, called the set of structural factorizations, and an action of $\mathbb{O}_n$ on this set.

Definition 1. For any given positive definite symmetric matrix Σ, the set

$$\mathbb{B}(\Sigma) = \{B \in \mathbb{GL}_n : \Sigma = B'B\}$$

is called the set of *feasible structural factorizations* of Σ.

This set is nonempty because every positive definite symmetric matrix admits a unique Cholesky factor R such that $\Sigma = R'R$ with R being an upper-triangular matrix with positive diagonal entries (see, for example, Meyer, 2000, 154–155). Clearly, any $B_1, B_2 \in \mathbb{B}(\Sigma), B_1 \neq B_2$, are observationally equivalent with respect to $\{Z_t\}$.[7]

Proposition 1. *$\mathbb{B}(\Sigma)$ is compact in $\mathbb{M}_n$.*

Proof. Consider the function $F(B) = B'B - \Sigma$. Because the usual matrix operations are continuous and the set consisting just of the zero matrix is closed, $F^{-1}(\{0\}) = \mathbb{B}(\Sigma)$ is closed. Moreover,

$$\sum_{j=1}^{n} B_{ij} B_{ij} = \Sigma_{ii} > 0$$

implies that $\|B\| := \sqrt{\sum_{i,j}^{n} B_{ij}^2} = \sqrt{\sum_{i=1}^{n} \Sigma_{ii}}$. Thus, the set $\mathbb{B}(\Sigma)$ is bounded. □

Consider the following map:

$$\mathbb{O}_n \times \mathbb{B}(\Sigma) \to \mathbb{B}(\Sigma) : (Q, B) = QB.$$

Note it is well-defined because $QB \in \mathbb{B}(\Sigma)$ as $B'Q'QB = B'B = \Sigma$ and continuous, in fact even smooth. Moreover, the map satisfies:

[2] In a non-Gaussian framework, it is conceivable to rely on higher moments.

[3] An early example of how difficult this can be, is given by the five-dimensional VAR analyzed by Blanchard (1989).

[4] A function is called smooth if it is infinitely differentiable.

[5] This is proven in Proposition 1 if Σ is set to I_n.

[6] The normalized Haar measure is unique and is the analogue to the uniform distribution on the real line. Denote the normalized Haar measure by μ, then in the case of orthogonal matrices we must have $\mu(\mathbb{O}_n) = 1$, and $\mu(Q\Omega) = \mu(\Omega Q) = \mu(\Omega)$ for every measurable set $\Omega \subseteq \mathbb{O}_n$ and every $Q \in \mathbb{O}_n$.

[7] In the terminology of Dufour and Hsiao (2008) $\mathbb{B}(\Sigma)$ is called a model and its elements structures.