

Neusser, Klaus

Working Paper

A topological view on the identification of structural vector autoregressions

Discussion Papers, No. 16-04

Provided in Cooperation with:

Department of Economics, University of Bern

Suggested Citation: Neusser, Klaus (2016) : A topological view on the identification of structural vector autoregressions, Discussion Papers, No. 16-04, University of Bern, Department of Economics, Bern

This Version is available at:

<https://hdl.handle.net/10419/149117>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



^b
**UNIVERSITÄT
BERN**

Faculty of Business, Economics
and Social Sciences

Department of Economics

**A Topological View on the Identification
of Structural Vector Autoregressions**

Klaus Neusser

16-04

March 2016

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
<http://www.vwi.unibe.ch>

A Topological View on the Identification of Structural Vector Autoregressions

Klaus Neusser *

March 16, 2016

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 becomes relevant, among other things, for the specification of vector autoregressions with time-varying covariance matrices.

JEL CLASSIFICATION: C1, C18, C32

KEYWORDS: SVAR, identification, group action, Haar measure, perturbation

*I thank Katharina Neusser for her perceptive comments.
Department of Economics, University of Bern, Schanzeneckstrasse 1, P.O. Box 8573, CH-3001 Berne, Switzerland. email: klaus.neusser@vwi.unibe.ch

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 founded in a priori economic reasoning. The econometric aspects of the identification problem has been analyzed by RUBIO-RAMÍREZ, WAGGONER, and ZHA (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 Vectorautoregressive 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 Φ_1, \dots, Φ_p :

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

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

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

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.

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 $\{\Psi_j\}$, $j = 0, 1, 2, \dots$, with $\Psi_0 = I_n$ and $\sum_{j=0}^{\infty} \|\Psi_j\| < \infty$ such that

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

$$= \Psi_0 B' V_t + \Psi_1 B' V_{t-1} + \Psi_2 B' V_{t-2} + \dots = \sum_{j=0}^{\infty} \Psi_j B' V_{t-j} = \Psi(L)B' V_t \quad (2.4)$$

Such a causal representation of X_t in terms of current and past Z_t 's exists if and only if $\det(I_n - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p) = \det \Phi(z) \neq 0$ 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} , $j = 0, 1, 2, \dots$. These effects are propagated over time and captured by the sequence $\{\Psi_j B'\}$, $j = 0, 1, 2, \dots$, 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.¹ 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

¹In a non-Gaussian framework, it is conceivable to rely on higher moments.

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.²

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.³ 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 .⁴ 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.⁵

This distribution can be efficiently implemented numerically by applying

²An early example of how difficult this can be, is given by the five-dimensional VAR analyzed by BLANCHARD (1989).

³A function is called smooth if it is infinitely differentiable.

⁴This is proven in Proposition 1 if Σ is set to I_n .

⁵The 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 $Q \in \mathbb{O}_n$.

the QR -decomposition to a random matrix A with law $\mathcal{L}(A) = N(0, I_n \otimes I_n)$, i.e. the elements of A are iid $N(0,1)$ random variables (see BIRKHOFF and GULATI, 1979; STEWART, 1980; EDELMAN and RAO, 2005, for details).

In Section 3.2, we derive an analytic expression for the density of sub-blocks 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 $\{Z_t\}$.⁶

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

Consider the following map:

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

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:

- (i) $I_n \in \mathbb{O}_n$ and $I_n B = B$ for all $B \in \mathbb{B}(\Sigma)$.

⁶In the terminology of DUFOUR and HSIAO (2008) $\mathbb{B}(\Sigma)$ is called a model and its elements structures.

- (ii) Because the matrix multiplication is associative, $(Q_1Q_2, B) = (Q_1, Q_2B)$ for all $Q_1, Q_2 \in \mathbb{O}_n$ and all $B \in \mathbb{B}(\Sigma)$.

Thus, the map defines a continuous (even smooth) *group action* of \mathbb{O}_n on $\mathbb{B}(\Sigma)$.

In addition, the group action is *transitive* because for any $B_1, B_2 \in \mathbb{B}(\Sigma)$ we can find $Q \in \mathbb{O}_n$ such that $(Q, B_1) = B_2$. To see this take $Q = B_2(B_1' B_1)^{-1} B_1' = B_2 \Sigma^{-1} B_1'$. This property means that we can move matrices in $\mathbb{B}(\Sigma)$ around via homeomorphism of $\mathbb{B}(\Sigma)$ onto itself. For *any* given element $B \in \mathbb{B}(\Sigma)$ the isotropy subgroup or stabilizer \mathbb{H}_B is defined as

$$\mathbb{H}_B = \{Q \in \mathbb{O}_n : QB = B\}.$$

Given the invertibility of B , this subgroup is the trivial subgroup which consists just of the identity element I_n . Thus, the group action is not only transitive, but also free. The quotient group \mathbb{O}_n/\mathbb{H} is therefore just \mathbb{O}_n itself. Given these preliminaries, we can apply two classic theorems of Weil (see DIESTEL and SPALSBURY, 2014, chapter 6) to draw the following two conclusions:

- (i) $\mathbb{O}_n/\mathbb{H}_B = \mathbb{O}_n$ and $\mathbb{B}(\Sigma)$ are homeomorphic. Thus, given any fixed matrix $B_0 \in \mathbb{B}(\Sigma)$, there corresponds to every identification scheme $B \in \mathbb{B}(\Sigma)$ a unique orthogonal matrix Q with $B = QB_0$ and vice-versa. A convenient choice of B_0 is given by the Cholesky factor of Σ .
- (ii) There is a unique \mathbb{O}_n -invariant probability measure on $\mathbb{B}(\Sigma)$.

This invariance property provides an important justification to state our ignorance about the correct identification scheme in $\mathbb{B}(\Sigma)$ in terms of the normalized Haar measure on \mathbb{O}_n (see JAYNES, 1968, sections 7 and 8). This view stands in contrast to the arguments put forward by BAUMEISTER and HAMILTON (2015).

3.2 The distribution of structural factorizations

The next step consist in the characterization of the Haar probability measure on \mathbb{O}_n which will then allows us to derive the corresponding probability measure on $\mathbb{B}(\Sigma)$. For this purpose, we closely follow the exposition in EATON (1989, chapter 7). Let $\mathbb{F}_{q,n}$, $q \leq n$, denote the set of $n \times q$ real matrices F such that $F'F = I_q$. These are matrices such that the q columns of F belong to some orthogonal matrix in $\mathbb{O}(n)$. For $q \leq n$, let A denote a $n \times q$ random matrix with probability law $\mathcal{L}(A) = \text{N}(0, I_n \otimes I_q)$, then $F = A(A'A)^{-1/2}$ is

a well-defined element of $\mathbb{F}_{q,n}$. According to EATON (1989, proposition 7.1), F has the uniform distribution on $\mathbb{F}_{q,n}$.

Partition the matrix A as $A = (A'_1, A'_2)'$ with A_1 and A_2 being $n_1 \times q$, respectively $(n - n_1) \times q$ matrices, then F can be written as

$$F = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} (A'_1 A_1 + A'_2 A_2)^{-1/2}.$$

Denote the upper $n_1 \times q$ block of F by F_1 . Thus,

$$F_1 = A_1 (A'_1 A_1 + A'_2 A_2)^{-1/2}.$$

F_1 can be thought of as the $n_1 \times q$ upper left block of some random orthogonal matrix.

With these preliminaries, we can quote two theorems of EATON (1989, Chapter 7, propositions 7.2 and 7.3).⁷

Theorem 1. *When $n_1 \geq q$ and $n_1 + q \leq n$, the random matrix $F'_1 F_1$ has a multivariate beta distribution $\mathbf{B}(n_1, n - n_1; I_q)$. Thus, $F'_1 F_1$ has density*

$$f_0(X) = C_0 (\det X)^{(n_1 - q - 1)/2} (\det(I_q - X))^{(n - n_1 - q - 1)/2} \mathbf{1}(X) \quad (3.1)$$

where $\mathbf{1}$ denotes the indicator function of the $q \times q$ symmetric matrices all of whose eigenvalues are in $(0, 1)$ and the constant C_0 is

$$C_0 = \omega(n_1, q) \omega(n - n_1, q) / \omega(n, q)$$

where ω is the Wishart constant, i.e.

$$\omega(r, q) = \left[\pi^{q(q-1)/4} 2^{rq/2} \prod_{j=1}^q \Gamma\left(\frac{r-j+1}{2}\right) \right]^{-1}.$$

Using the same notation as in Theorem 1, it is possible to derive an explicit formula for the density of F_1 .

Theorem 2. *For $n_1 \geq q$ and $n_1 + q \leq n$, the density of F_1 is given by*

$$f_1(X) = C_1 (\det(I_q - X'X))^{(n - n_1 - q - 1)/2} \mathbf{1}(X'X) \quad (3.2)$$

where the constant C_1 is given by

$$C_1 = \left(\sqrt{2\pi}\right)^{-n_1 q} \omega(n - n_1, q) / \omega(n, q)$$

⁷Further background material can be found in EATON (1983) and ANDERSON (1984).

Remark 1. *The case $n_1 \leq q$ is treated by taking transposes.*

It is easy to see that for the special case $n_1 = q = 1$, equation (2) reduces to

$$f_1(x) = \frac{\Gamma(n/2)}{\Gamma(1/2)\Gamma((n-1)/2)}(1-x^2)^{(n-3)/2} \mathbf{1}(x^2)$$

which is exactly the formula given in BAUMEISTER and HAMILTON (2015, equation (32)).

From the preceding Section 3.1 it is clear that every $B \in \mathbb{B}(\Sigma)$ can be represented as QR where Q is some orthogonal matrix and R is the Cholesky factor of Σ . Assuming as in Theorem 2 that $n_1 \geq q$ and $n_1 + q \leq n$, partition Q and the Cholesky factor R conformably into four submatrices

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$$

such that Q_{11} and R_{11} are $n_1 \times q$, respectively $q \times q$ submatrices. Remember that R_{11} is an invertible upper triangular square matrix. The upper left block $n_1 \times q$ block of $B \in \mathbb{B}(\Sigma)$, B_{11} , is then given by $B_{11} = Q_{11}R_{11}$. From this relation we can deduce the distribution of B_{11}

Corollary 1. *Given that the $n_1 \times q$ matrix Q_{11} , $n_1 \geq q$, has density (3.2), $B_{11} = Q_{11}R_{11}$ has a density given by*

$$\begin{aligned} f_2(X) &= C_1 |\det R_{11}|^{-n_1} \\ &\quad (\det(I_q - R_{11}^{-1'} X' X R_{11}^{-1}))^{(n-n_1-q-1)/2} \mathbf{1}(R_{11}^{-1'} X' X R_{11}^{-1}) \\ &= C_1 (\det \Sigma_{11})^{-(n-q-1)/2} \\ &\quad (\det(\Sigma_{11} - X' X))^{(n-n_1-q-1)/2} \mathbf{1}(R_{11}^{-1'} X' X R_{11}^{-1}) \end{aligned} \quad (3.3)$$

where Σ_{11} is the upper left $q \times q$ block of Σ . Using transposes a similar argument can be made for the case $n_1 \leq q$.

Proof. As Q_{11} is distributed according to the density in equation (3.2), the theory of multidimensional analysis (see, for example MAGNUS and NEUDECKER, 1999) implies that B_{11} has a density given by (3.3). \square

3.3 Perturbation analysis

As \mathcal{O}_n is a smooth submanifold in \mathbb{R}^{n^2} we have the notion of a tangent space at any point $Q \in \mathcal{O}_n$ denoted by $\mathbf{T}_Q \mathcal{O}_n$. We can explicitly realize this tangent space as:

$$\mathbf{T}_Q \mathcal{O}_n = \left\{ \left. \frac{d\gamma(t)}{dt} \right|_{t=0} \in \mathbb{M}_n : \gamma \text{ a smooth curve in } \mathcal{O}_n \text{ with } \gamma(0) = Q \right\}.$$

It can be shown by simple algebra that $\mathbf{T}_Q\mathbb{O}_n$ is a vector space. Because $\gamma(t) \in \mathbb{O}_n$, $\gamma(t)^T\gamma(t) = I_n$ where the superscript T denotes transposition. Differentiating this expression with respect to t and evaluating at $t = 0$ gives $\gamma'(0)^T\gamma(0) + \gamma(0)^T\gamma'(0) = \gamma'(0)^TQ + Q^T\gamma'(0) = 0$. This implies that $Q^T\gamma'(0)$ equals some skew-symmetric matrix S so that $\gamma'(0) = QS$. The tangent space at the identity therefore consists of the skew-symmetric matrices and we have $\mathbf{T}_Q\mathbb{O}_n = Q\mathbf{T}_{I_n}\mathbb{O}_n = Q\mathfrak{o}_n$ where \mathfrak{o}_n is the so-called Lie algebra of \mathbb{O}_n .⁸ The Lie algebra of a Lie group is just the tangent space at the identity together with the Lie bracket $[\cdot, \cdot]$ which for matrix groups is just the commutator $[A, B] = AB - BA$. In our case the Lie algebra \mathfrak{o}_n is the vector space of $n \times n$ skew-symmetric matrices S , i.e. of matrices with the property $S' = -S$. These matrices have trace, $\text{tr}(S)$, equal to zero and purely imaginary eigenvalues. The dimension of the vector space \mathfrak{o}_n is given by the number of free parameters of $n \times n$ skew-symmetric matrices. Thus,

$$\begin{aligned} \frac{n(n-1)}{2} &= \dim \mathfrak{o}_n = \dim \mathbf{T}_{I_n}\mathbb{O}_n = \dim \mathbb{O}_n = \dim \mathbb{B}(\Sigma) \\ &= \text{ "number of missing equation" }. \end{aligned}$$

The advantage of working with the Lie algebra is that it is a vector space where many computations can be accomplished more easily. The mechanism for passing information from the Lie algebra to the matrix Lie group is the matrix exponential map which is defined through the power series expansion:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

It can be shown that the exponential map is defined for every $A \in \mathbb{M}_n$ and that $\exp(A)$ is invertible. Thus, $\exp : \mathbb{M}_n \rightarrow \mathbb{GL}_n$. Note that $\exp(A+B)$ is in general not equal to $\exp(A)\exp(B)$. A sufficient condition for this to hold is that A and B commute, i.e. that $AB = BA$.

For skew-symmetric matrices, the matrix exponential delivers the special orthogonal group \mathbb{SO}_n of rotations which is defined as the set of orthogonal matrices with determinant equal to one. Thus, we have $\exp(\mathfrak{o}_n) = \mathbb{SO}_n \subset \mathbb{O}_n$. Given a fixed identification scheme $B \in \mathbb{B}(\Sigma)$, we can consider (small) random perturbation of this scheme as follows

$$\exp(S)B$$

where the free elements of the skew-symmetric matrix S are independently drawn from a zero mean normal distribution. Other centered distributions

⁸See BAKER (2001) and HALL (2003) for light introductions.

are, of course, reasonably as well. Thus, S is centered at the zero matrix whose exponential equals the identity matrix. Clearly, $\exp(S)B \in \mathbb{B}(\Sigma)$. Note that because \mathbb{SO}_n has determinant equal to one, $\exp(S)$ excludes reflections and simple interchanges of in the ordering of the structural shocks.

Alternatively, we can construct a random walk like time-varying scheme

$$B_t = \exp(S_t)B_{t-1}, \quad t = 1, 2, \dots \quad (3.4)$$

starting with some identification scheme $B_0 \in \mathbb{B}(\Sigma)$. The term random walk is motivated by taking the matrix logarithm of B_t and expressing it using the Baker-Campbell-Hausdorff formula (see HALL, 2003; HIGHAM, 2008):

$$\begin{aligned} \log B_t &= \log(\exp(S_t)B_{t-1}) \\ &= \log(B_{t-1}) + \log(\exp(S_t)) + \frac{1}{2}[\log(\exp(S_t)), \log(B_{t-1})] \\ &\quad + \text{higher order terms.} \end{aligned}$$

The “higher order terms” are given only in terms of Lie brackets of $\log(B_{t-1})$ and $\log(\exp(S_t))$. Ignoring the higher order terms and noting that $\log(\exp(S_t)) = S_t$, we can approximate $\log B_t$ as

$$\log B_t \approx \log(B_{t-1}) + S_t + \frac{1}{2}[S_t, \log(B_{t-1})].$$

Because $\|\log B_t - \log(B_{t-1}) - S_t\| = \|\frac{1}{2}[S_t, \log(B_{t-1})]\| \leq \frac{\sqrt{2}}{2}\|S_t\|\|\log(B_{t-1})\|$ (BÖTTCHER and WENZEL, 2008), we can make $\log B_t$ following approximately a random walk by making the perturbation S_t small. It is, however, not wise to use the logarithmic formulation because the matrix logarithm is not unique and the principal matrix logarithm not always defined (see HIGHAM, 2008, for more details). Thus, it is recommended to work directly with the propagation mechanism (3.4) and take the logarithmic specification only for motivation.

As $B_t \in \mathbb{B}(\Sigma)$ the resulting identification schemes will hold the covariance matrix constant. To relax this assumption, one can follow a suggestion of PRIMICERI (2005) and factorize the time-varying covariance matrix Σ_t as $\Sigma_t = L_t' \Omega_t L_t$ where L_t is an upper triangular matrix with ones on the diagonal and Ω_t is a diagonal matrix with strictly positive diagonal elements. Like the Cholesky decomposition, this factorization is also unique. The evolution of Σ_t can then be specified as $L_t = \exp(T_t)L_{t-1}$ where T_t is an upper triangular matrix with zeros on the diagonal and normally distributed mean zero elements above the main diagonal and $\Omega_t = \Omega_{t-1} \exp(D_t)$ and D_t is

a diagonal matrix with mean zero normally distributed diagonal elements.⁹ The scheme is initialized with L_0 being an upper triangular matrix with one on the diagonal and Ω_0 being a diagonal matrix with strictly positive diagonal elements. As $\exp(T_t)$ results in an upper triangular matrix with ones on the diagonal, the above scheme is well-defined and produces symmetric and positive definite matrices Σ_t .

4 Conclusion

Although the flavor of this note is on the theoretical side, it should be clear that the suggestions derived from viewing the identification problem in terms of a group action are awaiting to be implemented in empirical applications. This should affirm their usefulness and open the way for further insights along this route.

References

- ANDERSON, THEODORE W. (1984), *An Introduction to Multivariate Statistical Analysis*, 2nd edn., New York: John Wiley & Sons.
- BAKER, ANDREW (2001), *Matrix Groups – An Introduction to Lie Group Theory*, London: Springer-Verlag.
- BAUMEISTER, CHRISTIANE and JAMES D. HAMILTON (2015), “Sign restrictions, structural vector autoregressions, and useful prior information”, *Econometrica*, 83, 1963–1999.
- BIRKHOFF, GARRETT and SURENDER GULATI (1979), “Isotropic distributions of test matrices”, *Journal of Applied Mathematics and Physics*, 30, 148–158.
- BLANCHARD, OLIVIER J. (1989), “A traditional interpretation of macroeconomic fluctuations”, *American Economic Review*, 79, 1146–1164.
- BÖTTCHER, ALBRECHT and DAVID WENZEL (2008), “The Frobenius norm and the commutator”, *Linear Algebra and its Applications*, 429, 1864–1885.
- DIESTEL, JOE and ANGELA SPALSBURY (2014), *The Joys of Haar Measure*, vol. 150 of *Graduate Studies in Mathematics*, Providence, Rhode Island: American Mathematical Society.

⁹In the case of diagonal matrices the matrix logarithm can be applied without reservations so that one can write alternatively $\log(\Omega_t) = \log(\Omega_{t-1}) + D_t$.

- DUFOUR, JEAN-MARIE and CHENG HSIAO (2008), “Identification”, in *The New Palgrave Dictionary of Economics*, Steven N. Durlauf and Lawrence E. Blume, eds., 2nd edn., Palgrave Macmillan.
- EATON, MORRIS L. (1983), *Multivariate Statistics – A Vector Space Approach*, New York: John Wiley & Sons.
- EATON, MORRIS L. (1989), *Group Invariance Applications in Statistics*, vol. 1 of *Regional Conference Series in Probability and Statistics*, Alexandria, Virginia: American Statistical Association.
- EDELMAN, ALAN and N. RAJ RAO (2005), “Random matrix theory”, *Acta Numerica*, 14, 1–65.
- HALL, BRIAN C. (2003), *Lie Groups, Lie Algebras, and Representations – An Elementary Introduction*, vol. 222 of *Graduate Texts in Mathematics*, New York: Springer-Verlag.
- HIGHAM, NICHOLAS J. (2008), *Functions of Matrices – Theory and Computation*, Philadelphia: SIAM Society for Industrial and Applied Mathematics.
- JAYNES, EDWIN T. (1968), “Prior probabilities”, *IEEE Transactions On Systems Sciences and Cybernetics*, 4, 227–241.
- MAGNUS, JAN R. and HEINZ NEUDECKER (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, revised edn., Chichester: John Wiley and Sons.
- MEYER, CARL D. (2000), *Matrix Analysis and Applied Linear Algebra*, Philadelphia: Society for Industrial and Applied Mathematics.
- PRIMICERI, GIORGIO E. (2005), “Time varying structural vector autoregressions and monetary policy”, *Review of Economic Studies*, 72, 821–852.
- ROTHENBERG, THOMAS J. (1971), “Identification of parametric models”, *Econometrica*, 39, 577–591.
- RUBIO-RAMÍREZ, JUAN F., DANIEL F. WAGGONER, and TAO ZHA (2010), “Structural vector autoregressions: Theory of identification and algorithms for inference”, *Review of Economic Studies*, 77, 665–696.
- STEWART, GILBERT W. (1980), “The efficient generation of random orthogonal matrices with an application to condition estimation”, *SIAM Journal of Numerical Analysis*, 17, 403–409.

WAGGONER, DANIEL F. and TAO ZHA (2003), “Likelihood preserving normalization in multiple equation models”, *Journal of Econometrics*, 114, 329–347.