BAYESIAN IDENTIFICATION OF STRUCTURAL VARS

NIKOLAY AREFIEV AND RAMIS KHABIBULLIN

1. Introduction

Macroeconometricians use Structural Vector Autoregressions (SVARs) to estimate the response of aggregate activity to policy shocks, such as a restrictive monetary policy shock. One of steps of estimation of SVAR model is identification. However, there is no solution to the identification problem that would satisfy most of macroeconomic researchers. The approach to the identification that uses theoretically justified identifying restrictions does not satisfy many macroeconomic researchers, because it uses non-convincing identifying assumptions. The "sign-restrictions" approach is theoretically more appealing, but at our best knowledge, it has never produced an application with informative confidence intervals. The heteroskedasity-based approach uses assumption that only the variance of the structural shocks change over time, but the other structural parameters remain constant, which may be not reliable, and produces vague confidence intervals.

We propose a new method for identification. Our method relies on the following three assumptions:

Assumption 1 (True model is sparse)**.** *The matrix of parameters of the true structural model is sparse.*

This assumption is similar to the assumption used in the approach that uses identifying restrictions, but it has two important distinctions. First, we do not assume that we know in advance which parameters are restricted to zero, and which are not restricted. Second, if this assumption does not hold, it is asymptotically almost surely rejected by the proposed procedure.

Assumption 2 (True model is identified)**.** *The true model has enough restrictions at least for the just identification*

Without this assumption, the measure of the correctly identified model in the space of all possible models decreases proportionally to the square root of the number of observations, and exponentially with the number of variables included into the structural model. Therefore, without Assumption 2 the true model is not asymptotically revealed, and this assumption drastically increases the measure of correctly identified models in the space of all possible models. With this assumption, if the sufficient condition formulated in

Arefiev (2014) are satisfied, the approach asymptotically imposes high posterior weight onto the true model and uniform low weight on the other models given that the true model is sufficiently sparse.

Assumption 3 (Minnesota prior)**.** *The prior matrices of lagged effects are close to diagonal.*

Using Assumptions 1 and 2, the true model can be revealed only up to a permutation. The third assumption picks up the correct permutation matrix.

Numerical examples considered in the qualifying research paper by Khabibullin (2015) suggest that the proposed in this conference application has a great potential. At this moment we developed the theory required for the Gibbs sampler procedure of model search, see the appendices to this application. We expect to propose an working application example at the XVIII HSE April conference.

Appendix A. A few Lemma

The objective of this section is to take integral $\Theta_i(\alpha,\beta) \equiv \int_{-\infty}^{\infty} |\alpha x - \beta|^i \exp\left(-x^2\right) dx$ that we use in the following appendix section.

Lemma 1. *Let i be a non-negative integer and α be a parameter. I have:*

$$\int_{-\infty}^{\infty} (\alpha x - \beta)^i \exp\left(-x^2\right) dx = \Phi_i(\alpha,\beta), \tag{1}$$

where the sequence of functions $\Phi_i(\alpha,\beta)$ is defined recursively by: $\Phi_0(\alpha,\beta) = \sqrt{\pi}$, $\Phi_1(\alpha,\beta) = -\beta\sqrt{\pi}$, and for $i \geq 2$:

$$\Phi_i(\alpha,\beta) = \frac{i-1}{2}\alpha^2 \Phi_{i-2}(\alpha,\beta) - \beta\Phi_{i-1}(\alpha,\beta)$$

Proof. $\Phi_0(\alpha,\beta)$ is the Gaussian integral: $\Phi_0(\alpha) = \int_{-\infty}^{\infty} \exp\left(-x^2\right) dx = \sqrt{\pi}$. Since function $x\exp(-x^2)$ is odd, $\Phi_1(\alpha,\beta)$ is:

$$\Phi_1(\alpha) = \int_{-\infty}^{\infty} (\alpha x - \beta) \exp\left(-x^2\right) dx \quad = \quad \alpha \int_{-\infty}^{\infty} x \exp\left(-x^2\right) dx - \beta \int_{-\infty}^{\infty} \exp\left(-x^2\right) dx = -\beta\sqrt{\pi}$$

For $i \geq 2$ I have:

$$\begin{aligned}
\Phi_i(\alpha) &= \int_{-\infty}^{\infty} (\alpha x - \beta)^i \exp\left(-x^2\right) dx = \int_{-\infty}^{\infty} (\alpha x - \beta)(\alpha x - \beta)^{i-1} \exp\left(-x^2\right) dx \\
&= \alpha \int_{-\infty}^{\infty} x(\alpha x - \beta)^{i-1} \exp\left(-x^2\right) dx - \beta \int_{-\infty}^{\infty} (\alpha x - \beta)^{i-1} \exp\left(-x^2\right) dx \\
&= -\frac{\alpha}{2} \int_{-\infty}^{\infty} (\alpha x - \beta)^{i-1} d \exp\left(-x^2\right) - \beta \Phi_{i-1}(\alpha) \\
&= -\frac{\alpha}{2} \left((\alpha x - \beta)^{i-1} \exp\left(-x^2\right) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \exp\left(-x^2\right) d(\alpha x - \beta)^{i-1} \right) - \beta \Phi_{i-1}(\alpha) \\
&= \frac{\alpha^2}{2} (i-1) \int_{-\infty}^{\infty} (\alpha x - \beta)^{i-2} \exp(-x^2) dx - \beta \Phi_{i-1}(\alpha) \\
&= \frac{i-1}{2} \alpha^2 \Phi_{i-2}(\alpha) - \beta \Phi_{i-1}(\alpha)
\end{aligned}$$

□

Lemma 2. *Let i be a non-negative integer, α be a real parameter, and $\mathcal{N}(\cdot)$ be the normal CDF. I have:*

$$\int_{-\infty}^{\gamma} (\alpha x - \beta)^i \exp\left(-x^2\right) dx = \Psi_i(\alpha, \beta, \gamma), \tag{2}$$

where the sequence of functions $\Psi_i(\alpha, \beta, \gamma)$ is defined recursively by: $\Psi_0(\alpha, \beta, \gamma) = \sqrt{\pi}\mathcal{N}\left(\gamma\sqrt{2}\right)$, $\Psi_1(\alpha, \beta, \gamma) = -\frac{\alpha}{2} \exp\left(-\gamma^2\right) - \beta\Psi_0$, and for $i \geq 2$:

$$\Psi_i(\alpha) = \frac{i-1}{2} \alpha^2 \Psi_{i-2}(\alpha, \beta, \gamma) - \beta \Psi_{i-1}(\alpha, \beta, \gamma) - \frac{\alpha}{2} (\alpha\gamma - \beta)^{i-1} \exp(-\gamma^2)$$

Proof. Let $\sigma = \frac{1}{\sqrt{2}}$

$$\begin{aligned}
\Psi_0(\alpha, \beta, \gamma) &= \int_{-\infty}^{\gamma} \exp(-x^2) dx = \sqrt{\pi} \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{1}{2\sigma^2} x^2) dx = \sqrt{\pi}\mathcal{N}(\gamma\sqrt{2}) \\
\Psi_1(\alpha, \beta, \gamma) &= \int_{-\infty}^{\gamma} (\alpha x - \beta) \exp\left(-x^2\right) dx = \alpha \int_{-\infty}^{\gamma} x \exp\left(-x^2\right) dx - \beta \int_{-\infty}^{\gamma} \exp\left(-x^2\right) dx \\
&= -\frac{\alpha}{2} \int_{-\infty}^{\gamma} d \exp\left(-x^2\right) - \beta \Psi_0 = -\frac{\alpha}{2} \exp\left(-\gamma^2\right) - \beta \Psi_0
\end{aligned}$$

$$\begin{aligned}
i \geq 2 : \Psi_i(\alpha,\beta,\gamma) &= \int_{-\infty}^{\gamma} (\alpha x-\beta)^i \exp\left(-x^2\right) dx \\
&= \alpha \int_{-\infty}^{\gamma} x(\alpha x-\beta)^{i-1} \exp\left(-x^2\right) dx - \beta \int_{-\infty}^{\gamma} (\alpha x-\beta)^{i-1} \exp\left(-x^2\right) dx \\
&= -\frac{\alpha}{2} \int_{-\infty}^{\gamma} (\alpha x-\beta)^{i-1} d \exp\left(-x^2\right) - \beta \Psi_{i-1}(\alpha,\beta,\gamma) \\
&= -\frac{\alpha}{2} \left((\alpha x-\beta)^{i-1} \exp(-x^2)|_{-\infty}^{\gamma} \right) + \frac{\alpha^2}{2}(i-1) \int_{-\infty}^{\gamma} (\alpha x-\beta)^{i-2} \exp(-x^2) dx - \beta \Psi_{i-1}(\alpha,\beta,\gamma) \\
&= \frac{\alpha^2}{2}(i-1) \int_{-\infty}^{\gamma} (\alpha x-\beta)^{i-2} \exp(-x^2) dx - \beta \Psi_{i-1}(\alpha,\beta,\gamma) - \frac{\alpha}{2}(\alpha\gamma-\beta)^{i-1}\exp(-\gamma^2) \\
&= \frac{i-1}{2}\alpha^2 \Psi_{i-2}(\alpha,\beta,\gamma) - \beta \Psi_{i-1}(\alpha,\beta,\gamma) - \frac{\alpha}{2}(\alpha\gamma-\beta)^{i-1}\exp(-\gamma^2)
\end{aligned}$$

□

Summarizing Lemmas 1 and 2, I get:

$$\text{(3)} \quad \Theta_i(\alpha,\beta) \equiv \int_{-\infty}^{\infty} |\alpha x-\beta|^i \exp\left(-x^2\right) dx = \begin{cases} \Phi_i(\alpha,\beta) & \text{if } i \text{ is even} \\ \operatorname{sign}(\alpha)\left(\Phi_i(\alpha,\beta) - 2\Psi_i\left(\alpha,\beta,\frac{\beta}{\alpha}\right)\right) & \text{if } i \text{ is odd and } \alpha \neq 0 \\ |\beta|^T \sqrt(\pi) & \text{if } i \text{ is odd and } \alpha = 0 \end{cases}$$

Appendix B. Parameter distribution function required for Gibbs Sampling

The structural model is:

$$XA = \varepsilon \tag{4}$$

where $X = [YZ]$ is $T \times m$ matrix of centralized endogenous and exogenous variables, A is $m \times n$ matrix of parameters to be estimated, and ε is the matrix of structural shocks, $\mathbb{E}(\varepsilon'\varepsilon) = \mathbb{I}$ is the identity matrix. Let A_0 be the top $n \times n$ block of A.

The likelihood function is:

$$\mathcal{L} = (2\pi)^{-\frac{n}{2}} |\det A_0|^T \exp\left(-\frac{1}{2}\operatorname{tr}\left(XAA'X'\right)\right) \tag{5}$$

(6)

Proposition 1 (Likelihood as a function of $a_{ij} \in A_0$)**.** *Consider $a_{ij} \in A_0$ as a free parameter, taking the other parameters as given. The likelihood function as a function of a_{ij} is proportional to:*

$$\mathcal{L}(a_{ij}) \propto L(a_{ij}) = |a_{ij}C_{ij} + R_{ij}|^T \exp\left(-0.5\omega_{ii}\left(a_{ij}+\rho_{ij}\right)^2\right)$$

where ω_{ij} is the entry of matrix $\Omega = X'X$, $\rho_{ij} = \sum_{k\neq i} a_{kj}\frac{\omega_{ki}}{\omega_{ii}}$, C_{ij} is the cofactor of a_{ij} in the determinant, and R_{ij} is the remainder of the determinant, which does not depend on a_{ij}, but which can be calculated using shortcut: $R_{ij} = \det A_0 - a_{ij}C_{ij}$.

Proof. Rewrite the likelihood function as:

$$\mathcal{L} = (2\pi)^{-\frac{nT}{2}} \left|\det A_0\right|^T \exp\left(-\frac{1}{2}\mathrm{tr}\left(\Omega A A'\right)\right),$$

where $\Omega = X'X$. Drop the terms that do not depend on A:

$$\mathcal{L} \propto \left|\det A_0\right|^T \exp\left(-\frac{1}{2}\mathrm{tr}\left(\Omega A A'\right)\right)$$

Write $\Omega A A'$ in the extensive form and drop the terms that do not depend on a_{ij}:

$$\mathcal{L} \propto \left|\det A_0\right|^T \exp\left(-\frac{1}{2}\left(a_{ij}^2\omega_{ii} + 2a_{ij}\sum_{k\neq i} a_{kj}\omega_{ki}\right)\right)$$

Complete to the trace part to the sum of squares and drop the term that do not depend on a_{ij}:

$$\begin{aligned}\mathcal{L} &\propto \left|\det A_0\right|^T \exp\left(-\frac{\omega_{ii}}{2}\left(a_{ij} + \sum_{k\neq i} a_{kj}\frac{\omega_{ki}}{\omega_{ii}}\right)^2\right)\\ &= \left|\det A_0\right|^T \exp\left(-\frac{\omega_{ii}}{2}\left(a_{ij}+\rho_{ij}\right)^2\right),\end{aligned}$$

where ρ_{ij} is defined to be $\sum_{k\neq i} a_{kj}\frac{\omega_{ki}}{\omega_{ii}}$.

Finally, rite the determinant as $\det A_0 = a_{ij}C_{ij} + R_{ij}$, where C_{ij} is the cofactor of a_ij, and R_{ij} is the remainder. By construction, C_{ij} and R_{ij} do not depend on a_{ij}. I get:

$$\mathcal{L} \propto \left|a_{ij}C_{ij} + R_{ij}\right|^T \exp\left(-\frac{\omega_{ii}}{2}\left(a_{ij}+\rho_{ij}\right)^2\right)$$

□

Distribution of $a_{ij} \in A_0$ by Lemma 1

Proposition 2 (Odds ratio)**.** *The posterior odds that $a_{ij} \in A_0$ is unrestricted equals to* PrOR $\cdot$ LOR, *where* PrOR *is the prior odds ratio, and the likelihood odds ratio is given by:*

$$\mathrm{LOR} = \frac{\exp\left(.5\omega_{ii}\rho_{ij}^2\right)}{\sqrt{2}\left|R_{ij}\right|^T}\frac{\Theta_{2T}(\alpha_{ij},\beta_{ij})}{\Theta_T\left(\sqrt{2}\alpha_{ij},\beta_{ij}\right)}, \tag{7}$$

where $\alpha_{ij} = C_{ij}\omega_{ii}^{-1/2}$ *and* $\beta_{ij} = C_{ij}\rho_{ij} - R_{ij}$

Proof. The likelihood odds ratio is:

$$\text{LOR} = \frac{\mathbb{E}_{a_{ij}}\mathcal{L}(a_{ij})}{\mathcal{L}(0)} = \frac{\int_{-\infty}^{\infty} L(a_{ij})^2 da_{ij}}{L(0)\int_{-\infty}^{\infty} L(a_{ij}) da_{ij}}$$

$$\int_{-\infty}^{\infty} L(a_{ij})^2 da_{ij} = |a_{ij}C_{ij} + R_{ij}|^T \exp\left(-\omega_{ii}\left(a_{ij} + \rho_{ij}\right)^2\right) da_{ij} =: \left| \begin{array}{c} \sqrt{\omega_{ii}}\left(a_{ij} + \rho_{ij}\right) = x \\ da_{ij} = \frac{dx}{\sqrt{\omega_{ii}}} \\ a_{ij} = \frac{x}{\sqrt{\omega_{ii}}} - \rho_{ij} \end{array} \right|$$

$$:= \sqrt{\omega_{ii}^{-1}} \int_{-\infty}^{\infty} \left(\frac{xC_{ij}}{\sqrt{\omega_{ii}}} - \rho_{ij}C_{ij} + R_{ij}\right)^{2T} \exp\left(-x^2\right) dx = \sqrt{\omega_{ii}^{-1}}\Theta_{2T}(C_{ij}\omega_{ii}^{-1/2}, C_{ij}\rho_{ij} - R_{ij})$$

$$\int_{-\infty}^{\infty} L(a_{ij}) da_{ij} = \int_{-\infty}^{\infty} |a_{ij}C_{ij} + R_{ij}|^T \exp\left(-.5\omega_{ii}\left(a_{ij} + \rho_{ij}\right)^2\right) =: \left| \begin{array}{c} \sqrt{.5\omega_{ii}}\left(a_{ij} + \rho_{ij}\right) = x \\ a_{ij} = \frac{x}{\sqrt{.5\omega_{ii}}} - \rho_{ij} \\ da_{ij} = \frac{dx}{\sqrt{.5\omega_{ii}}} \end{array} \right|$$

$$:= \frac{1}{\sqrt{.5\omega_{ii}}} \int_{-\infty}^{\infty} \left|\frac{xC_{ij}}{\sqrt{.5\omega_{ii}}} - \rho_{ij}C_{ij} + R_{ij}\right|^T \exp\left(-x^2\right) dx = \sqrt{2\omega_{ii}^{-1}}\Theta_T\left(C_{ij}\sqrt{2\omega_{ii}^{-1}}, C_{ij}\rho_{ij} - R_{ij}\right)$$

$$\begin{aligned} \text{LOR} &= \frac{\sqrt{\omega_{ii}^{-1}}\Theta_{2T}(C_{ij}\omega_{ii}^{-1/2}, C_{ij}\rho_{ij} - R_{ij})}{\sqrt{2\omega_{ii}^{-1}}\Theta_T\left(C_{ij}\sqrt{2\omega_{ii}^{-1}}, C_{ij}\rho_{ij} - R_{ij}\right)|R_{ij}|^T \exp\left(-.5\omega_{ii}\rho_{ij}^2\right)} \\ &= \frac{\exp\left(.5\omega_{ii}\rho_{ij}^2\right)}{\sqrt{2}\,|R_{ij}|^T} \frac{\Theta_{2T}(C_{ij}\omega_{ii}^{-1/2}, C_{ij}\rho_{ij} - R_{ij})}{\Theta_T\left(C_{ij}\sqrt{2\omega_{ii}^{-1}}, C_{ij}\rho_{ij} - R_{ij}\right)} \\ &= \frac{\exp\left(.5\omega_{ii}\rho_{ij}^2\right)}{\sqrt{2}\,|R_{ij}|^T} \frac{\Theta_{2T}(\alpha_{ij}, \beta_{ij})}{\Theta_T\left(\sqrt{2}\alpha_{ij}, \beta_{ij}\right)} \end{aligned}$$

□

Proposition 3 (Posterior distribution of $a_{ij} \in A_0$)**.** *Under uninformative prior, the posterior distribution of* a_{ij} *given the other parameters and data is:*

$$a_{ij} \sim \begin{cases} \text{constrained to 0} & \text{prob } \frac{1}{1+\text{LOR}} \\ \frac{|a_{ij}C_{ij} + R_{ij}|^T}{\sqrt{2\omega_{ii}^{-1}}\Theta_T\left(\sqrt{2}\alpha_{ij}, \beta_{ij}\right)} \exp\left(-\frac{\omega_{ii}}{2}\left(a_{ij} + \rho_{ij}\right)^2\right) & \text{prob } \frac{\text{LOR}}{1+\text{LOR}} \end{cases}$$

The CDF for a_{ij} in case $i \leq n$ given that a_{ij} is not constrained can be written explicitly using functions Φ and Ψ defined in Appendix A.

Example: Assume the parameters of A_0 can be written as:

$$A_0 = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Then a_{11}, a_{22} and a_{33} follow the ±square root of the inverse chi squared distribution with T degrees of freedom, and parameters a_{21}, a_{31}, and a_{32} follow a normal distribution.

Proposition 4 (Likelihood as a function of $a_{ij} \notin A_0$)**.** *Consider $a_{ij} \notin A_0$ as a free parameter, taking the other parameters as given. The likelihood function as a function of a_{ij} is proportional to:*

$$\mathcal{L}(a_{ij}) \propto L(a_{ij}) = \exp\left(-0.5\omega_{ii}\left(a_{ij} + \rho_{ij}\right)^2\right)$$

where ω_{ij} is the entry of matrix $\Omega = X'X$, $\rho_{ij} = \sum_{k \neq i} a_{kj} \frac{\omega_{ki}}{\omega_{ii}}$.

Proposition 5 (Posterior of $a_{ij} \notin A_0$)**.**

$$\text{LOR} = \sqrt{.5}\exp(.5\omega_{ii}\rho_{ij}^2)$$

$$a_{ij} \sim \begin{cases} \text{constrained to 0} & \text{prob } \frac{1}{1+\text{LOR}} \\ \mathcal{N}\left(-\rho_{ij}, \sqrt{\omega_{ii}^{-1}}\right) & \text{prob } \frac{\text{LOR}}{1+\text{LOR}} \end{cases}$$