1 Modeling SVARs in RISE

1.1 The structural model

$$\bar{A}_0 y_t + \bar{A}_{-1} y_{t-1} + ... + \bar{A}_{-p} y_{t-p} + \bar{C} + \varepsilon_t = 0 \tag{1}$$

- $\varepsilon_t \sim N(0, I_n)$ is an $n \times 1$ vector of structural shocks
- y_t is an $n \times 1$ vector of endogenous variables
- $\bar{C}$ is an $n \times 1$ vector of constants
- $1 \leq t \leq t_{\max}$
- $\bar{A}_0$, $\bar{A}_{-1}, ..., \bar{A}_{-p}$ are $n \times n$ matrices of parameters
- p is the lag length
- The initial conditions $y_{t-1}...y_{t-p}$ are taken as given
- $\left\{\bar{A}_0, \bar{A}_{-1}, ..., \bar{A}_{-p}, \bar{C}\right\}$ are the structural parameters of the system

1.2 The solution

$$\begin{array}{rcl} y_t & = & \underbrace{\left(-\bar{A}_0^{-1}\bar{A}_{-1}\right)}_{\bar{T}_{-1}} y_{t-1} + ... + \underbrace{\left(-\bar{A}_0^{-1}\bar{A}_{-p}\right)}_{\bar{T}_{-p}} y_{t-p} + \underbrace{\left(-\bar{A}_0^{-1}\bar{C}\right)}_{K} + \underbrace{\left(-\bar{A}_0^{-1}\right)}_{\bar{R}} \varepsilon_t \\ & = & \bar{T}_{-1} y_{t-1} + ... + \bar{T}_{-p} y_{t-p} + K + \bar{R}\varepsilon_t \\ & = & \bar{T}_{-1} y_{t-1} + ... + \bar{T}_{-p} y_{t-p} + K + u_t \end{array} \tag{2}$$

with

$$u_t \equiv \bar{R}\varepsilon_t$$

so that the short-run variance of y_t is given by

$$\begin{array}{rcl} V\left(y_t | y_{t-1}...y_{t-p}\right) & = & V\left(u_t\right) \\ & = & \bar{R} V\left(\varepsilon_t\right) \bar{R}' \\ & = & \bar{R}\bar{R}' \\ & \equiv & \Sigma_u \end{array} \tag{3}$$

We can also compute the long-run variance. Define $\bar{T}(L) \equiv I_n - \bar{T}_{-1}L - ... - \bar{T}_{-p}L^p$. The solution can also be written in a moving average representation form

$$y_t = \bar{T}(L)^{-1} K + \bar{T}(L)^{-1} u_t \tag{4}$$

We compute the long-run variance by letting $L=1$ in the expression above, so that

$$\begin{aligned} V\left(y_t\right) &= \bar{T}\left(1\right)^{-1}V\left(u_t\right)\left[\bar{T}\left(1\right)^{-1}\right]' \\ &= \bar{T}\left(1\right)^{-1}\Sigma_u\left[\bar{T}\left(1\right)^{-1}\right]' \\ &\equiv \Omega_u \\ &= \bar{T}\left(1\right)^{-1}\bar{R}\bar{R}'\left[\bar{T}\left(1\right)^{-1}\right]' \\ &= \left[\bar{T}\left(1\right)^{-1}\bar{R}\right]\left[\bar{T}\left(1\right)^{-1}\bar{R}\right]' \\ &\equiv FF' \end{aligned} \tag{5}$$

where $F\equiv\bar{T}\left(1\right)^{-1}\bar{R}$ or $\bar{R}=\bar{T}\left(1\right)F$

- Estimates of $\left\{\bar{T}_{-1},...,\bar{T}_{-p},K,\Sigma_u,\Omega_u\right\}$ are readily available from the OLS/MLE. Those parameters constitute the reduced-form parameters in the VAR language.
- In the special case where there are no lags in the model, $\bar{T}\left(1\right)=I_n$, and as we could have expected, we have :
  - $\Omega_u=\Sigma_u$ i.e. the short-run variance is equal to the long run
  - $F=\bar{R}$
- In order to compute interpretable IRFs, Variance decompositions, historical decompositions, etc. we need the structural parameters, which, conditional on the reduced-form parameters, all depend on $\bar{R}$. So in a sense, identifying the structural parameters is synonymous to identifying the structural shocks.
- Note for any orthogonal matrix Π such that $\Pi^{-1}=\Pi'$, we have $\left(\bar{R}\Pi\right)\left(\bar{R}\Pi\right)'=\bar{R}\Pi\Pi'\bar{R}'=\bar{R}\bar{R}'=\Sigma_u$, hence $\bar{R}$ is not unique. By the same argument, F is not unique
- It follows that the likelihood is independent of the specific values of R or F.

$$\log Lik=-\frac{1}{2}\sum_{t=1}^{t_{\max}}\left(n\log\left(2\pi\right)+\log\left(\det\left(\Sigma_u\right)\right)+trace\left(e_t'\Sigma_u^{-1}e_t\right)\right) \tag{6}$$

with

$$e_t\equiv y_t-\bar{T}_{-1}y_{t-1}+...+\bar{T}_{-p}y_{t-p}+K$$

1.3 Identification strategies

Consider the expressions

$$\Sigma_u=\bar{R}\bar{R}'$$

and

$$\Omega_u = \left[\bar{T}(1)^{-1}\bar{R}\right]\left[\bar{T}(1)^{-1}\bar{R}\right]'$$

In each of those, the left-hand side has n variances on the diagonal and $\frac{n(n-1)}{2}$ distinct elements off diagonal. The total is then $\frac{n(n+1)}{2}$ distinct elements. In order to identify the right-hand side, we need $\frac{n(n+1)}{2}$ conditions. From the beginning of the problem, we had already imposed n conditions with the assumption that $\varepsilon_t \sim N(0, I_n)$, i.e. all structural variances are 1. In order to fully identify $\bar{R}$ we need at least another $\frac{n(n-1)}{2}$ restrictions.

If we have more than $\frac{n(n-1)}{2}$ additional restrictions, the system is over-identified. If we have fewer, the system is under-identified and if we have exactly $\frac{n(n-1)}{2}$ additional restrictions , the system is just-identified.

1.3.1 Short-run identification

Short run identification imposes restrictions on the entries of $\bar{R}$. Once the restrictions are imposed, the entries of $\bar{R}$ are estimated through maximizing the concentrated likelihood

$$\log Lik = -\frac{1}{2}\sum_{t=1}^{t_{\max}}\left(n\log(2\pi) + \log\left(\det\left(\bar{R}\bar{R}'\right)\right) + \hat{e}_t'\left(\bar{R}\bar{R}'\right)^{-1}\hat{e}_t\right) \tag{7}$$

with

$$\hat{e}_t \equiv y_t - \widehat{\bar{T}_{-1}}y_{t-1} + ... + \widehat{\bar{T}_{-p}}y_{t-p} + \widehat{K}$$

1.3.2 Long-run identification

Long run identification imposes restrictions on the entries of F and recovers $\bar{R}$ as $\bar{R} = \bar{T}(1)F$. Once the restrictions are imposed, the entries of F are estimated through maximizing the concentrated likelihood (7) with $\bar{R} = \bar{T}(1)F$.

1.3.3 Mixing short and long-run identification

It is easy to mix short and long-run restrictions. RISE starts out by imposing restrictions on $\bar{R}$, then conditional on $\bar{R}$, it computes $F = \bar{T}(1)^{-1}\bar{R}$ and minimizes the distance between F_{ij} and F_{ij}^*, where F_{ij}^* embodies the long-run restrictions.

1.3.4 Cholesky identification

This simple identification strategy imposes a recursive structure on the $\bar{R}$ matrix. Without loss of generality, we could order the variables in the y vector as $y^1 \longrightarrow y^2 \longrightarrow y^3...y^{n-1} \longrightarrow y^n$, where shocks affect the variables as shown by the arrows. More explicity, shocks that affect y^1 also affect the rest of the system but shocks that directly hit y^2 do not affect y^1, but affect the other variables.

- Although this strategy is often used for short-run identification, it can also be used for long-run identification. After ordering the variables, we estimate the system (the other way around would be correct as well) and get Σ_u and Ω_u.
  - Short-run identification: take Cholesky decomposition of Σ_u so that $\Sigma_u = \bar{R}\bar{R}'$
  - Short-run identification: take Cholesky decomposition of Ω_u so that $\Sigma_u = FF'$ and recover $\bar{R}$ as $\bar{R} = \bar{T}(1)F$
- The Cholesky identification can also be computed through optimization and the results are identical... at least in theory[1].

1.3.5 Sign restrictions identification

I hate sign restrictions and have not implemented them yet. I will do so when my nausea for them is passed.

2 The RISE representation

RISE represents both the structural and solution forms in companion form. In particular, (1) becomes

$$A_{+}\begin{bmatrix} y_{t+1} \\ y_t \\ \vdots \\ y_{t-p} \end{bmatrix} + A_0 \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} + A_{-}\begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + C + B\varepsilon_t = 0$$

where $A_{+} \equiv \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}$, $A_0 \equiv \begin{bmatrix} \bar{A}_0 & 0 & \cdots & 0 \\ 0 & I & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & I \end{bmatrix}$, $A_{-} \equiv \begin{bmatrix} \bar{A}_{-1} & \bar{A}_{-2} & \cdots & \bar{A}_{-p} \\ -I & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & -I & 0 \end{bmatrix}$,
$C \equiv \begin{bmatrix} \bar{C} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $B \equiv \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ and (2) becomes

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \bar{T}_{-1} & \bar{T}_{-2} & \cdots & \bar{T}_{-p} \\ I & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} K \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \bar{R} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varepsilon_t = 0$$

[1] The optimization algorithm may fail to converge to the global maximum.

3 Numerical issues

- The advantage of the Cholesky identification strategy is that it does not require optimization.
- The disadvantage of identification procedures that require optimization is that sometimes, the optimization may not converge to the maximum.
- RISE is armed with debugging tools that can be used to check that the global peak is found.
- In addition to other restrictions, RISE imposes some normalization restrictions restricting to the diagonal of the $\bar{R}$ to be positive. This turns out to assuage the estimation burden.