structural vector autoregressions

Structural vector autoregressions (SVARs) are a multivariate, linear representation of a vector of observables on its own lags and (possibly) other variables as a trend or a constant. SVARs make explicit identifying assumptions to isolate estimates of policy and/or private agents’ behaviour and its effects on the economy while keeping the model free of the many additional restrictive assumptions needed to give every parameter a behavioural interpretation. Introduced by Sims (1980), SVARs have been used to document the effects of money on output (Sims and Zha, 2006a), the relative importance of supply and demand shocks on business cycles (Blanchard and Quah, 1989), the effects of fiscal policy (Blanchard and Perotti, 2002), or the relation between technology shocks and worked hours (Galí, 1999), among many other applications.

Economic theory and the SVAR representation

Dynamic economic models can be viewed as restrictions on stochastic processes. From this perspective, an economic theory is a mapping between a vector of $k$ economic shocks $w_t$ and a vector of $n$ observables $y_t$ of the form $y_t = D(w')$, where $w'$ represents the whole history of shocks $w_t$ up to period $t$. The economic shocks are those shocks to the fundamental elements of the theory: preferences, technology, informational sets, government policy, measurement errors, and so on. The observables are all variables that the researcher has access to. Often, $y_t$ includes a constant to capture the mean of the process. The mapping $D(\cdot)$ is the product of the equilibrium behaviour of the agents in the model, implied by their optimal decision rules and consistency conditions like resource constraints and market clearing. The construction of the mapping $D(\cdot)$ is the sense in which economic theory tightly relates shocks and observables. Also, the mapping $D(\cdot)$ can be interpreted as the impulse response of the model to an economic shock.

Often, we restrict our attention to linear mappings of the form $y_t = D(L)w_t$, where $L$ is the lag operator. For simplicity of exposition, $w_t$ will be i.i.d. random variables and normally distributed, $w_t \sim N(0, \Sigma)$. More involved structures – for example, allowing for autocorrelation between the shocks – can be accommodated with additional notation.

We pick the neoclassical growth model, the workhorse of dynamic macroeconomics, to illustrate the previous paragraphs. In its basic version, the model maps productivity shocks, the $w_t$ of the theory, into observables, $y_t$, like output or investment. The mapping comes from the optimal investment and labour supply decisions of the households, the resource constraint of the economy, and the law of motion for productivity. If the productivity shocks are normally distributed and we solve the model by linearizing its equilibrium conditions, we obtain a mapping of the form $y_t = D(L)w_t$ described above.
If $k = n$, that is, we have as many economic shocks as observables, and $|D(L)|$ has all its roots outside the unit circle, we can invert the mapping $D(L)$ (see Fernández-Villaverde, Rubio-Ramírez and Sargent, 2005) and obtain

$$A(L)y_t = w_t$$

where $A(L) = A_0 - \sum_{k=1}^{\infty} A_k L^k$ is a one-sided matrix lag polynomial that embodies all the (usually nonlinear) cross-equation restrictions derived by the equilibrium solution of the model. In general, $A(L)$ is of infinite order. This representation is known as the SVAR representation. The name comes from realizing that $A(L)y_t = w_t$ is a vector autoregression (VAR) generated by an economic model (a ‘structure’).

**Reduced form representation, normalization, and identification**

Consider now the case where a researcher does not have access to the SVAR representation. Instead, she has access to the VAR representation of $y_t$:

$$y_t = B_1 y_{t-1} + B_2 y_{t-2} + \cdots + a_t,$$

where $Ey_{t-j} a_t = 0$ for all $j$ and $Ea_t a_t' = \Omega$. This representation is known as the reduced-form representation. Can the researcher recover the SVAR representation using the reduced-form representation? Fernández-Villaverde, Rubio-Ramírez, and Sargent (2005) show that, given a strictly invertible economic model, that is, $|D(L)|$ has all its roots strictly outside the unit circle, there is one and only one identification scheme to recover the SVAR from the reduced form. In addition, they show that the mapping between $a_t$ and $w_t$ is $a_t = A_0^{-1} w_t$. Hence, if the researcher knew $A_0^{-1}$, she could recover the SVAR representation from the reduced form, noting that $A_j = A_0 B_j$ for all $j$ and $w_t = A_0 a_t$.

Hence, the recovery of $w_t$ from $y_T$ requires the knowledge of the dynamic economic model. Can we avoid this step? Unfortunately, the answer is, in general, ‘no’, because knowledge of the reduced-form matrices $B_i$ and $\Omega$ does not imply, by itself, knowledge of the $A_i$ and $\Sigma$, for two reasons.

The first is normalization. Reversing the signs of two rows or columns of the $A_i$ does not matter for the $B_i$. Thus, without the correct normalization restrictions, statistical inference about the $A_i$’s is essentially meaningless. Waggoner and Zha (2003) provide a general normalization rule that maintains coherent economic interpretations.

The second is identification. If we knew $A_0$, each equation $B_i = A_0^{-1} A_i$ would determine $A_i$ given some $B_i$. But the only restrictions that the reduced-form representation imposes on the matrix $A_0$ comes from $\Omega = A_0 \Sigma A_0'$. In this relationship, we have $n(3n + 1)/2$ unknowns (the $n^2$ distinct elements of $A_0$ and the $n(n+1)/2$ distinct elements of $\Sigma$) for $n^2$ knowns (the $n(n+1)/2$ distinct elements of $\Omega$). Thus, we require $n^2$ identification restrictions. Since we can set the diagonal elements of $A_0$ equal to 1 by scaling, we are left with the need of $n(n-1)$ additional identification restrictions (alternatively, we could scale the shocks such that the diagonal of $\Sigma$ is