Vector Autoregressions VARs.
Text pp. 312-331

Consider two stationary time series, $y_{1,t}$ and $y_{2,t}$ and the dynamic regressions (suppress the constant)

$$\begin{aligned} y_{1,t} &= a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + \epsilon_{1,t} \\ y_{2,t} &= a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + \epsilon_{2,t} \end{aligned}$$

Rewrite this system in vector (matrix) form.

$$\underbrace{\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix}}_{y_{t-1}} + \underbrace{\begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}}_{\epsilon_t}$$

$$y_t = Ay_{t-1} + \epsilon_t$$

One lag of y_t :VAR(1). A VAR(2) would look like,

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \epsilon_t$$

Here, we'd be regressing $y_{1,t}$ on 2 lags of itself and 2 lags of y_{2t}, and similarly for the y_{2t} equation.

Estimation: Do least squares on each equation individually. We assume

$$\begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}}_{\Sigma} \right)$$

Grab the regression residuals to estimate Σ

$$\begin{aligned} \hat{\sigma}_{11} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_{1,t}^2 \\ \hat{\sigma}_{12} = \hat{\sigma}_{21} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_{1,t}\hat{\epsilon}_{2,t} \\ \hat{\sigma}_{22} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_{2,t}^2 \end{aligned}$$

How to choose lag length? Use information criteria (AIC, BIC, etc). Let k be the number of regression coefficients in the system (the $a_{ij}'s$).

$$\begin{aligned} AIC &= 2\ln|\Sigma| + \frac{2k}{T} \\ BIC &= 2\ln|\Sigma| + \frac{k\ln(T)}{T} \end{aligned}$$

where $|\Sigma|$ is the determinant of the covariance matrix.

Impulse response analysis. Remember when we did the AR(1) model, let y_t be a scalar time series. We can get the MA(∞) representation

$$\begin{aligned}
y_t &= \rho y_{t-1} + \epsilon_t \\
&= \rho\left(\rho y_{t-2} + \epsilon_{t-1}\right) + \epsilon_t = \epsilon_t + \rho\epsilon_{t-1} + \rho^2 y_{t-2} = \epsilon_t + \rho\epsilon_{t-1} + \rho^2\left(\rho y_{t-3} + \epsilon_{t-2}\right) \\
&= \epsilon_t + \rho\epsilon_{t-1} + \rho^2\epsilon_{t-2} + \rho^3\epsilon_{t-3} + \cdots
\end{aligned}$$

Mimic this with the VAR. Now y_t is a vector

$$\begin{aligned}
y_t &= Ay_{t-1} + \epsilon_t \\
&= A\left(Ay_{t-2} + \epsilon_{t-1}\right) + \epsilon_t \\
&= A^2 y_{t-2} + \epsilon_t + A\epsilon_{t-1} \\
y_t &= \epsilon_t + A\epsilon_{t-1} + A^2\epsilon_{t-2} + A^3\epsilon_{t-3} + \cdots
\end{aligned}$$

A^2 means AA. Not raising each element in A to the second power.

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{12}a_{21} + a_{11}^2 & a_{11}a_{12} + a_{12}a_{22} \\ a_{11}a_{21} + a_{21}a_{22} & a_{12}a_{21} + a_{22}^2 \end{pmatrix}$$

We use this to do the impulse response analysis.

$$y_t = \epsilon_t + A\epsilon_{t-1} + A^2\epsilon_{t-2} + A^3\epsilon_{t-3} + \cdots$$

Two issues: (1) How big should the shock be? Usually, people want the shock to be one standard deviation. (2) How to unambigiously attribute the shocks if the shocks are correlated? We need to make the shocks uncorrelated. This is called orthogonalizing the shocks. Let ϵ_1 and ϵ_2 be correlated, and let z_1 and z_2 be independent standard normals $N(0,1)$. We can build up ϵ_1 and ϵ_2 as functions of z_1 and z_2. Let

$$\begin{aligned}
\epsilon_1 &= \sqrt{\sigma_{11}}z_1 \\
\epsilon_2 &= \sqrt{\sigma_{22}}\left(\rho z_1 + z_2\sqrt{(1-\rho^2)}\right)
\end{aligned}$$

where $\rho = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$. Verify

$$\begin{aligned}
E\left(\epsilon_1^2\right) &= \sigma_{11}E\left(z_1^2\right) = \sigma_{11} \\
E\left(\epsilon_2^2\right) &= E\left\{\sigma_{22}\left(\rho^2 z_1^2 + \left(1-\rho^2\right)z_2^2 + 2\rho\sqrt{(1-\rho^2)}z_1 z_2\right)\right\} \\
&= \sigma_{22}\left(\rho^2 + 1 - \rho^2\right) = \sigma_{22} \\
E\left(\epsilon_1\epsilon_2\right) &= E\left\{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\left(\rho z_1^2 + z_1 z_2\sqrt{(1-\rho^2)}\right)\right\} = \sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\rho = \sigma_{12}
\end{aligned}$$

Now get the $z's$ as a function of the $\epsilon's$. We call the $z's$ structural shocks.

$$\begin{aligned}
z_1 &= \frac{1}{\sqrt{\sigma_{11}}}\epsilon_1 \\
z_2 &= \frac{\sqrt{\sigma_{11}}\epsilon_2 - \rho\sqrt{\sigma_{22}}\epsilon_1}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\sqrt{(1-\rho^2)}}
\end{aligned}$$

This is known as the Choleski decomposition of the error covariance matrix. Let's write the shocks in matrix form

$$\underbrace{\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}}_{\epsilon_t} = \underbrace{\begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ \rho\sqrt{\sigma_{22}} & \sqrt{\sigma_{11}}\sqrt{(1-\rho^2)} \end{pmatrix}}_{\Lambda} \underbrace{\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix}}_{z_t}$$

and substitute this back into the MA(∞) representation.

$$y_t = \Lambda z_t + A\Lambda z_{t-1} + A^2\Lambda z_{t-2} + + \cdots$$

So now I can shock y_{1t} with z_{1t} and not worry about how z_{2t} moves, because it doesn't.