Autoregressions: Problems

Suppose that you estimate the following VAR(1) process:

$$y_t = \alpha + \beta y_{t-1} + e_t \tag{1}$$

1. Often true. The regression coefficients are estimated through OLS. Generally, a VAR allows for correlation of the error terms across equations to allow for more efficient estimation. If the set of regressors is the same for all equations, however, then this process is also the same as OLS. Most, but not all, VARs have the same set of regressors.

2. False. As with OLS, the order does not affect the coefficients. The order matters only in how we impose the identifying restrictions in order to obtain orthogonalized impulse response functions.

3. They use lags. Usually, lags will be correlated with the independent (and possibly endogenous) variable but uncorrelated with the time t error term.

For #4-8, suppose that you are interested in estimating an economic system using a VAR with three variables: rainfall, crop yields, and agriculture profits.

4. False. Non-stationarity is a source of potentially major misspecification as with an autoregression.

5. The most common approach is to run the VAR for different lag lengths and then choose th best length using an information criteria.

6. We want to order our variables from the most exogenous to the most endogenous. We can try to try different orders and hope that results are robust. Or we can use theory and intuition to try to convince the reader that a given order is best. Here, rainfall is surely highly exogenous (it probably should just be an exogenous variable rather than part of the system). I would argue that crop yields are probably the next most exogenous. Although farmers may respond to changes in profits by planting different amounts, I expect that that this effect may not occur right away.

7. Our vector of endogenous variables, x_t includes rainfall, yields, and profits. Assume they are differenced as needed to ensure stationary. We now run our VAR. Lets say it is a VAR(2):

$$x_t = \alpha + \beta_1 x_{t-1} + \beta_2 x_{t-2} + u_t \tag{2}$$

We assume that the true data generating process takes the following form that allows for contemporaneous effects:

$$\theta x_t = \kappa + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + e_t \tag{3}$$

We now assume that the off-diagonal elements of θ (and $theta^{-1}$) are zero. This is why order matters. We are assuming that shocks to yield and profits cannot affect rainfall contemporaneously. For once, this seems like a perfectly safe assumption.

Now, as in class, we set Σ, our estimated covariance matrix equal to $\theta^{-1}(\theta^{-1})'$ to solve for θ^{-1}.

We then start the model at its steady state. We then set $e_t' = [1, 0, 0]$ for a shock to rainfall. The impact IRFs are then:

$$x_t = \bar{x} + \theta^{-1} e_t \tag{4}$$

Where t is now the period where the impulse occurs. For $t+1$ we then use (2):

$$x_{t+1} = \alpha + \beta_1 x_t + \beta_2 \bar{x} \tag{5}$$

where the two period lag still reflects the steady state value. Confidence intervals may be computed based on the standard errors of the regression coefficients. The shocks are now set to zero for all periods after t. For $t+2$:

$$x_{t+2} = \alpha + \beta_1 x_{t+1} + \beta_2 x_t \tag{6}$$

and we can then continue for as many periods as we wish.

8. Cointergration refers to cases where a linear combination of non-stationary variables is stationary. A standard VAR (as opposed to an error correction model) omits this information and it is reasonable to interpret this as a case of omitted variable bias.