

Reduced-form VARs

Tyler Sotomayor

2024

Worked calculations for a lag-major design, recursive forecasts, and the parameter penalties in multivariate information criteria.

Based on the 2024 BSE course taught by Luca Gambetti and Konstantin Boss. I have revised the classroom code and added the accompanying explanations. The notes identify the papers, data, and assumptions used in each application. These are my course notes, not an official BSE publication. A classroom replication should not be read as a replication of every result in its underlying paper.

1 The first design row

The dependent dates are 3, 4, and 5. Intercept-first, lag-major ordering gives

$$X = \begin{pmatrix} 1 & 3 & 1 & 1 & 2 \\ 1 & 2 & 4 & 3 & 1 \\ 1 & 5 & 3 & 2 & 4 \end{pmatrix}, \quad Y = \begin{pmatrix} 2 & 4 \\ 5 & 3 \\ 4 & 6 \end{pmatrix}.$$

There are only three usable observations but $k = 1 + 2 \cdot 2 = 5$ coefficients per equation. X cannot have full column rank; computing a pseudoinverse would not create identification or positive residual degrees of freedom. The production estimator correctly rejects this specification.

After removing date 3, date 4 lacks its first lag and date 5 lacks its second lag. Dates 1 and 2 do not have the required two-period history. No dependent date is usable. Shifting the remaining rows would incorrectly treat date 2 as date 3's replacement. This is why the lag function matches calendar labels.

2 Two forecasts

The first forecast uses the observed y_T :

$$\hat{y}_{T+1|T} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} .5 & .2 \\ 0 & .4 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

The second step substitutes that forecast, not the original state:

$$\hat{y}_{T+2|T} = \begin{pmatrix} 1 + .5(3) + .2(2) \\ .4(2) \end{pmatrix} = \begin{pmatrix} 2.9 \\ .8 \end{pmatrix}.$$

The triangular matrix has eigenvalues .5 and .4, both inside the unit circle. Solve $(I - A)\mu = c$. The second equation gives $.6\mu_2 = 0$, so $\mu_2 = 0$; the first gives $.5\mu_1 - .2\mu_2 = 1$, so $\mu_1 = 2$. As a check, forecasts converge toward $(2, 0)'$.

An innovation $b = (0, 1)'$ instead has impact response b and one-step response $Ab = (.2, .4)'$. The intercept cancels when subtracting the no-shock path from the shock path. Thus the forecast $(3, 2)'$ and the response $(.2, .4)'$ answer different questions even though both use the same matrix A .

3 Comparing information criteria

At one lag, each of two equations has three coefficients, hence six system coefficients. At two lags, each has five, hence ten. The criteria are

$$AIC(1) = -2 + 2(6)/100 = -1.88,$$

$$AIC(2) = -2.1 + 2(10)/100 = -1.90,$$

$$BIC(1) = -2 + \log(100)(6)/100 \simeq -1.72369,$$

$$BIC(2) = -2.1 + \log(100)(10)/100 \simeq -1.63948.$$

AIC selects two lags; BIC selects one. The second lag lowers the fit term by .1. That exceeds AIC's extra penalty of .08 but not BIC's extra penalty of about .18421. This arithmetic is more informative than saying one criterion is always better.

If the larger model loses ten additional observations, its fit term refers to a different set of outcomes. A better score may reflect which observations were removed as well as a better lag specification. Re-estimate both candidates on their common dependent dates before applying the comparison. The data can still be used in full after selecting the order, provided that refit is recorded separately.