Lecture note 3

General companion form, and the ADL derived from the VAR.

The companion form of a VAR

In Lecture 5 we saw that the VAR in two variables X_t and Y_t, and with first order dynamics, an VAR(1), is already on so called **companion form**.

We now demonstrate that also a general VAR in n variables and dynamics of order p can be expressed as a VAR(1) with a companion matrix that contain all the parameters of the system.

We define $\mathbf{y}_t$ as the $n \times 1$ vector

$$\mathbf{y}_t = [Y_{1t}, Y_{2t}, \ldots, Y_{nt}]'$$

and write the system with dynamics of degree p as:

$$\mathbf{y}_t = \boldsymbol{\phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\phi}_2 \mathbf{y}_{t-2} + ... + \boldsymbol{\phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \tag{1}$$

where $\boldsymbol{\phi}_i$ is a $n \times n$ matrix with parameters and $\boldsymbol{\epsilon}_t$ is a vector of random variables (it can be jointly normally distributed (Gaussian) or at least stationary). The *companion form* of this general system is:

$$\underbrace{\begin{pmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{pmatrix}}_{\mathbf{z}_t} {}_{np\times 1} = \underbrace{\begin{pmatrix} \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_{p-1} & \boldsymbol{\phi}_p \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{pmatrix}}_{\mathbf{F}_{np\times np}} \underbrace{\begin{pmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-p} \end{pmatrix}}_{\substack{np\times 1 \\ \mathbf{z}_{t-1}}} + \underbrace{\begin{pmatrix} \boldsymbol{\epsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}}_{\boldsymbol{\epsilon}_t} {}_{np\times 1} \tag{2}$$

or, written compactly, with the suggested notation:

$$\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \tag{3}$$

The vector variable $\mathbf{z}_t$ has a global asymptotic stable solution and is stationary if all eigenvalues of $\mathbf{F}$ from

$$|\mathbf{F} - \lambda \mathbf{I}| = 0 \tag{4}$$

are less than one in magnitude. The means all real roots must be between -1 and 1, and all complex roots must have modulus less than one.

The ADL model derived from a Gaussian VAR

The following system is an example of a first order Gaussian VAR in the two time series X_t and Y_t:

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \pi_{10} \\ \pi_{20} \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix} \tag{5}$$

$$\begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \mid Y_{t-1}, X_{t-1}\right). \tag{6}$$

The conditioning in (6) means that the normal distribution of $\boldsymbol{\epsilon}_t = (\varepsilon_{yt},\ \varepsilon_{xt})'$ in the VAR is conditional on the history of the system up to period $t-1$. In Lecture note 4, we generalize the argument by starting from the so called Haavelmo distribution, but as said in the lectures and computer class, it is enough to study the first order case carefully to get the right understanding.

Start by writing (5) as

$$Y_t = \mu_{y,t-1} + \varepsilon_{yt} \tag{7}$$
$$X_t = \mu_{x,t-1} + \varepsilon_{xt} \tag{8}$$

where the conditional expectations $\mu_{y,t-1} \equiv E(Y_t \mid Y_{t-1}, X_{t-1})$ and $\mu_{x,t-1} \equiv E(X_t \mid Y_{t-1}, X_{t-1})$ are given by

$$\mu_{y,t-1} = \pi_{10} + \pi_{11}Y_{t-1} + \pi_{12}X_{t-1} \tag{9}$$
$$\mu_{x,t-1} = \pi_{20} + \pi_{21}Y_{t-1} + \pi_{22}X_{t-1}. \tag{10}$$

Y_t and X_t given by (7), (8) and (6) have a joint normal distribution which is conditional on $X_{t-1}$and Y_{t-1}. It follows from the properties of the normal distribution that the conditional distribution of Y_t given X_t is also normal, with expectation:

$$\begin{aligned}
E(Y_t \mid X_t, X_{t-1}, Y_{t-1}) &= \mu_{y,t-1} - \rho_{xy}\frac{\sigma_y}{\sigma_x}\mu_{x,t-1} + \rho_{xy}\frac{\sigma_y}{\sigma_x}X_t \\
&= \pi_{10} - \frac{\omega_{xy}}{\sigma_x^2}\pi_{20} + \frac{\omega_{xy}}{\sigma_x^2}X_t + (\pi_{12} - \frac{\omega_{xy}}{\sigma_x^2}\pi_{22})X_{t-1} \\
&\quad + (\pi_{11} - \frac{\omega_{xy}}{\sigma_x^2}\pi_{21})Y_{t-1}
\end{aligned}$$

where ρ_{xy} is the correlation coefficient between ε_{xt} and ε_{yt}:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2}\sqrt{\sigma_y^2}}. \tag{11}$$

We can now define parameters:

$$\phi_0 = \pi_{10} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{20} \tag{12}$$
$$\phi_1 = \pi_{11} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{21} \tag{13}$$
$$\beta_0 = \frac{\sigma_{xy}}{\sigma_x^2} \tag{14}$$
$$\beta_1 = \pi_{12} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{22} \tag{15}$$

and write the conditional expectation as

$$E(Y_t \mid X_t, X_{t-1}, Y_{t-1}) = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1}. \tag{16}$$

Finally, define the disturbance ε_t as

$$\varepsilon_t = Y_t - E(Y_t \mid X_t, X_{t-1}, Y_{t-1}) \tag{17}$$

and write the conditional model for Y_t as

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \tag{18}$$

which is an ADL model. (The same as equation (13.58) in DM, but with the obvious change in notation (and with $p = q = 1$)).

Note that, again with reference to the normal distribution, the ADL-disturbance is:

$$\varepsilon_t = \varepsilon_{yt} - \frac{\sigma_{xy}}{\sigma_x^2}\varepsilon_{xt} \tag{19}$$

with variance:

$$\sigma^2 = \sigma_y^2(1 - \rho_{xy}^2). \tag{20}$$

If you guess that an ADL model with p-lags can be derived form a $VAR(p)$ that condition on X_{t-p-1} and Y_{t-p-1} you are right! We loose nothing in generality about the status of the ADL as a conditional model by just looking at the simplest, $VAR(1)$, case. For those interested, we refer to *Lecture note 4*.

Exogeneity and pre-determinedness: Remember that the starting point is the joint distribution of ε_{yt} and ε_{xt} conditional on X_{t-1} and Y_{t-1}, cf (6). Therefore:

$$E(\varepsilon_t \mid X_{t-1}, Y_{t-1}) = 0$$

so the disturbance of the ADL model is uncorrelated with the conditioning variables of the VAR. But we have also

$$E(\varepsilon_t \mid \varepsilon_{xt}) = 0 \tag{21}$$

and since $\varepsilon_{xt} = [X_t - E[X_t \mid X_{t-1}, Y_{t-1}]$ by definition, (21) can alternatively be expressed as:

$$E(\varepsilon_t \mid X_t, X_{t-1}, Y_{t-1}) = 0 \tag{22}$$

showing that ε_t is uncorrelated with **all** the explanatory variables of the model, just like in the static regression models we have seen before. In sum: X_t, X_{t-1},are exogenous variables, exactly in the sense given by (22).

However

$$E(\varepsilon_{t-j} \mid X_t, X_{t-1}, Y_{t-1}) \neq 0 \text{ for } j = 1, 2, \ldots$$

since Y_{t-1} must depend on ε_{t-1} and earlier disturbances via the solution for Y_t obtained by repeated substitution of lagged Y_t in (18):

$$\begin{aligned} Y_t &= \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \\ &= \phi_0(1 + \phi_1) + X\text{-terms} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + ... + \phi_1^2 Y_{t-2} \end{aligned}$$

and so on. Hence Y_{t-1} is a pre-determined variable in (18), not an exogenous variable.

Role of Gaussian VAR assumption: In all important respects, the above remains valid if (6) is replaced by an IID assumption for the VAR disturbance. The only expectation is the equations that maps from the parameters of the normal distribution to the parameter of the ADL. But the parameters of the ADL will still be parameters in a conditional expectation (again, just as in the static/ordinary regression model case).