Lecture 25
Forecasting Using ARMAV Models
(aka Forecasting Using Leading Indicators)

Goal: Improve prediction of one time-series by adding information from another, related time-series. This "related" time-series is often referred to as the "leading indicator".

Essentially, it's prediction using ARMAV models

Will show it using the ARV(1) model (simplest dynamics)

$$X_{2t} = \phi_{1_{21}} X_{1_{t-1}} + \phi_{1_{22}} X_{2_{t-1}} + a_{2t}$$

Taking conditional expectation of $X_{2_{t+1}}$ at time t, gives

$$E[X_{2t+1} | t] = \hat{X}_{2t}(1) = \phi_{1_{21}} X_{1t} + \phi_{1_{22}} X_{2t}$$

and for $t+2$, we have

$$E[X_{2t+2} | t] = \hat{X}_{2t}(2) = \phi_{1_{21}} \hat{X}_{1t}(1) + \phi_{1_{22}} \hat{X}_{2t}(1)$$

For $t+3$ we have

$$\hat{X}_{2t}(3) = \phi_{1_{21}} \hat{X}_{1t}(2) + \phi_{1_{22}} \hat{X}_{2t}(2)$$

Basically, I intend to use predictions of X_{1t} to make predictions in X_{2t}. If there is any structure in X_{1t}, I can use it!

Let's assume X_{1t} also follows ARV(1) model

$$X_{1t} = \phi_{1_{11}} X_{1t-1} + \phi_{1_{12}} X_{2t-1} + a_{1t}$$

$$\hat{X}_{1t}(1) = \phi_{1_{11}} X_{1t} + \phi_{1_{12}} X_{2t}$$

$$\hat{X}_{1t}(2) = \phi_{1_{11}} \hat{X}_{1t}(1) + \phi_{1_{12}} \hat{X}_{2t}(1)$$

$$\hat{X}_{1t}(2) = \phi_{1_{11}} \hat{X}_{1t}(1) + \phi_{1_{12}} (\phi_{1_{21}} X_{1t} + \phi_{1_{22}} X_{2t}) =$$

$$= \phi_{1_{11}} (\phi_{1_{11}} X_{1t} + \phi_{1_{12}} X_{2t}) + \phi_{1_{12}} (\phi_{1_{21}} X_{1t} + \phi_{1_{22}} X_{2t})$$

$$\hat{X}_{1t}(3) = \phi_{1_{11}} \hat{X}_{1t}(2) + \phi_{1_{12}} \hat{X}_{2t}(2) =$$

$$\phi_{1_{11}}[\phi_{1_{11}} (\phi_{1_{11}} X_{1t} + \phi_{1_{12}} X_{2t}) + \phi_{1_{12}} (\phi_{1_{21}} X_{1t} + \phi_{1_{22}} X_{2t})] + \phi_{1_{12}} \hat{X}_{2t}(2)$$

Essentially, we're using one t-s to predict another one.

Good for forecasting when no. of samples is small.

Sometimes it may not work. Whethear we can do prediction or not depends on whether there is some internal dynamic structure in the t-s or not (whether it came from some other t-s or from some dynamic mechanism). If such internal structure does not exist, one can not do prediction with or without ARMAV.

Rule of thumb: If a single TS model is predicting the TS well, then ARMAV will do even better. If single ARMA model does not work, then even ARMAV probably won't work.

General Case of Leading Indicator Prediction

$$X_{1t} = T_{11}(B) X_{1t} + T_{12}(B) X_{2t} + T_{a_1}(B) a_{1t} \qquad (1)$$

$$X_{2t} = T_{21}(B) X_{1t} + T_{22}(B) X_{2t} + T_{a_2}(B) a_{2t} \qquad (2)$$

$$(1) \Rightarrow \quad X_{1t} = \frac{T_{12}(B)}{1 - T_{11}(B)} X_{2t} + \frac{T_{a_1}(B)}{1 - T_{11}(B)} a_{1t}$$

substitute this into (2) (X_{1t} is the leading indicators

$$(1 - T_{22}(B)) X_{2t} = \frac{T_{21}(B) T_{12}(B)}{1 - T_{11}(B)} X_{2t} + \frac{T_{a_1}(B) T_{21}(B)}{1 - T_{11}(B)} a_{1t}$$

$$+ T_{a_2}(B) a_{2t}$$

$$\left(1 - \frac{T_{21}(B) T_{12}(B)}{(1 - T_{11}(B))(1 - T_{22}(B))}\right) X_{2t} =$$

$$= \frac{T_{a_1}(B) T_{21}(B)}{(1 - T_{11}(B))(1 - T_{22}(B))} a_{1t} + \frac{T_{a_2}(B)}{1 - T_{22}(B)} a_{2t}$$

$$X_{2t} = \frac{T_{a_1}(B) T_{21}(B)}{(1 - T_{11}(B))(1 - T_{22}(B)) - T_{21}(B) T_{12}(B)} a_{1t} +$$

$$+ \frac{T_{a_2}(B) (1 - T_{11}(B))}{(1 - T_{11}(B))(1 - T_{22}(B)) - T_{21}(B) T_{12}(B)} a_{2t}$$

Now, I can Break X_{2t} into a series of operators and get

$$X_{2t} = (G_{10} + G_{11} B + G_{12} B^2 + \dots + G_{1\ell} B^\ell + \dots) a_{1t} +$$

$$+ (G_{20} + G_{21} B + G_{22} B^2 + \dots + G_{2\ell} B^\ell + \dots) a_{2t}$$

Which is analogous to the Green's fcn decomposition of simple ARMA models

$$\hat{X}_{2t}(\ell) = E[X_{2t+\ell} | t] =$$

$$= (G_{1\ell} B^\ell + G_{1\ell+1} B^{\ell+1} + \dots) a_{1t+\ell} + (G_{2\ell} B^\ell + G_{2\ell+1} B^{\ell+1} + \dots) a_{2t+\ell}$$

$$Var(\hat{e}_{2t}(\ell)) = Var[(G_{10} + G_{11} B + \dots + G_{1\ell-1} B^{\ell-1}) a_{1t+\ell}] +$$

$$+ Var[(G_{20} + G_{21} B + \dots + G_{2\ell-1} B^{\ell-1}) a_{2t+\ell}] =$$

$$= \sigma_{a_1}^2 (G_{10}^2 + G_{11}^2 + \dots + G_{1\ell-1}^2) + \sigma_{a_2}^2 (G_{20}^2 + G_{21}^2 + \dots + G_{2\ell-1}^2)$$

Example:

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} -0.0148 & 0.035 \\ 0.06584 & 0.098 \end{bmatrix} \begin{bmatrix} X_{1\,t-1} \\ X_{2\,t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

$$\gamma_a = \begin{bmatrix} 0.3653 & 0 \\ 0 & 0.6016 \end{bmatrix} \qquad \begin{matrix} \sigma_{a_{22}} = 0.6016 \\ \gamma_{a_{11}} = 0.3653 \end{matrix}$$

Just modeling X_{2t} alone gives

$$X_{2t} = 0.0273\, X_{2\,t-1} + a_{2t} \qquad \sigma_{a_2} = 0.7604$$

$$X_{1t}' = \frac{0.035B}{1+0.0148B} X_{2t} + \frac{1}{1+0.0148B} a_{1t}$$

$$X_{2t} = \frac{0.06584B}{1-0.098B} X_{1t} + \frac{1}{1-0.098B} a_{2t} =$$

$$= \frac{0.0023B^2}{1-0.0832B-0.0015B^2} X_{2t} + \frac{0.06584B}{1-0.0832B-0.0015B^2} a_{1t}$$

$$+ \frac{1}{1-0.098B} a_{2t}$$

$$\left(1 - \frac{0.0023\, B^2}{1 - 0.0832\, B - 0.0015 B^2}\right) X_{2t} = \frac{0.06584\, B}{1 - 0.0832\, B - 0.0015\, B^2}\, a_{1t}$$

$$+ \frac{1}{1 - 0.098 B}\, a_{2t}$$

$$X_{2t} = \frac{0.06584\, B}{1 - 0.0832\, B - 0.0038 B^2}\, a_{1t} + \frac{1 - 0.0148\, B}{1 - 0.0832 B - 0.0038 B^2}\, a_{2t}$$

Breaking the expression above into PFE terms (partial fraction expansion) will lead to a Green's fn (orthogonal) decomposition in terms of a_{1t} and a_{2t}.

ARMAV Based Regulation

Let us now look at the case when we can control one time-series in order to make another time-series do what we want it to do!

<u>Goal</u>: keep the system output ("the other t-s") at a desired level, with minimal (in the mean-least-squares sense) variations!

<u>Strategy</u> Fiddle around with X_{1t} to make sure $X_{2\,t+L}$ is as close to 0 as possible

L - lag (time-delay) between input action and output reaction to that input.

Method: We will fix X_{1t} to make sure that

$$\hat{X}_{2t}(L) = E[X_{2\,t+L} \mid t] = 0$$

After control, $X_{2\,t+L} = \overset{0}{\hat{X}_{2t}(L)} + \hat{e}_{2t}(L) = \hat{e}_{2t}(L)$

Steps to accomplish this:

i) Obtain the output model before control

ii) Derive a control law based on the minimal mean-squared-error forecast

iii) Obtain the model after control

iv) Evaluate control efficiency

Examples:

X_{1t} - gate opening input $\qquad$ X_{2t} paper weight output

(i) $\rightarrow$ $X_{2t} = 0.25\,\underset{\text{control signal}}{X_{1\,t-1}} + 0.7\,X_{2\,t-1} + a_{2t}$ $\qquad$ Model before control

$$\sigma_{a_{2t}} = 0.0062$$

ii) Deriving the control law

$$X_{2t+1} = 0.25 X_{1t} + 0.7 X_{2t} + a_{2t+1} \quad \Big| \; E[\cdot \mid t]$$

$$\hat{X}_{2t}(1) = 0.25 X_{1t} + 0.7 X_{2t} + 0 = 0$$

we want to set it to 0 to make sure X_{2t+1} is minimized in the mean LS sense

$$\Rightarrow 0.25 X_{1t} + 0.7 X_{2t} = 0 \Rightarrow X_{1t} = -\frac{0.7}{0.25} X_{2t}$$

$$X_{1t} = -2.8 X_{2t}$$ CONTROL LAW

iii) Model after control

$$X_{2t+1} = \hat{X}_{2t}(1) + a_{2t+1} = a_{2t+1}$$

$$Var[X_{2t+1}] = Var[X_{2t}] = Var[a_{2t}] = \sigma_{22} = 0.0062$$

iv) Evaluating control efficiency

$$[Var\ X_{2t}]_{\text{Before control}} = ? \qquad [Var\ X_{2t}]_{\text{with control}} = 0.0062$$