Mathematics Behind VAR

1 Motivations

A Vector Autoregression (VAR) model is a type of autoregressive model, which aim to predict the next value in a sequence by measuring the previous values. They're used in time series data of the like that we're looking in to.

2 Process

For each time step, the model has one equation for each variable within the system, so in our case there will be 8 equations at each time step. Below I give an example of the equations for a VAR model with 2 variables,

$$\begin{bmatrix} y_{t,1} \\ y_{t,2} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{11,1} & \phi_{12,1} \\ \phi_{21,1} & \phi_{22,1} \end{bmatrix} \begin{bmatrix} y_{t-1,1} \\ y_{t-1,2} \end{bmatrix} + \begin{bmatrix} e_{t,1} \\ e_{t,2} \end{bmatrix}$$

or expanded,

$$y_{t,1} = c_1 + \phi_{11,1} y_{t-1,1} + \phi_{12,1} y_{t-1,2} + e_{t,1}$$

$$y_{t,2} = c_2 + \phi_{21,1} y_{t-1,1} + \phi_{22,1} y_{t-1,2} + e_{t,2}$$

where,
$y_{t,i}$ is the tth observation of the ith variable
c_i is a constant,
$\phi_{ii,l}$ captures the influence of the lth lag variable of y_i on itself,
$\phi_{ij,l}$ captures the influence of the lth lag variable of y_j on y_i,
$e_{t,i}$ are white noise processes that maybe be related.
The goal of training is to estimate parameters for each equation so as to minimise,

$$\sum e_{t,i}^2$$

As we can see from the equations above, each target variable depends not only on its lag, but also on the other target variables lag. This is useful as some variables in our dataset are likely dependent on one another, giving us greater predictive power.

2.1 Lag Variable

Lag variables are key to the VAR model. It is a previous observation of the variable in question and allows us to predict our next value dependent on those we have already seen.

3 How and why we're using the VAR model

For our project, we're using the VAR model as a baseline. It's appropriate for our data as it's a model that is used for time series data and it is also a lot simpler than the models we're going to delve into later. This gives us a good starting point for comparison when it comes to how it performs, how long it takes to run and the amount of memory it requires.

4 References

https://online.stat.psu.edu/stat510/lesson/11/11.2#:~:text=VAR%20models%20(vector%20autoregressive%20models,%2C%20and%20x%20t%20%2C%203%20