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 and are used in time series data of the like that we're looking in to.

A VAR model is an extension of this that models multiple variables and looks at the relationships between them.

In a VAR model all variables are modelled as if they influence eachother equally

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,

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

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

where,

$$y_{i,t} \text{ is the } t\text{th observation of the } i\text{th variable}$$
$$c_i \text{ is a constant}$$
$$\phi_{ii,l} \text{ captures the influence of the } l\text{th lag variable of } y_i \text{ on itself}$$
$$\phi_{ij,l} \text{ captures the influence of the } l\text{th lag variable of } y_j \text{ on } y_i$$
$$e_{i,t} \text{ are white noise processes that maybe be related}$$

The goal is to estimate parameters for each equation so as to minimise,

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

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 if there is such a relation between the current values and those we have encountered previously.

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 commonly used for time series 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 long it takes to run and the amount of memory it takes.