COVID-19 Model

April 2020

Let $x_{t,j}$ be the COVID-19 cases at time t in county j, for t=1, ...and T, j=1, ..., R. Let $x_t = [x_{t,1}, \dots, x_{t,R}]$, the R $\times$ 1 vector of COVID-19 cases at time t. We can write the model as

$$\underbrace{x_t}_{1\times R} = \underbrace{u_t}_{1\times RL}\underbrace{B}_{RL\times R} + \underbrace{e_t}_{1\times R}$$

for t=1, ..., T, where $u_t = [x_{t1}, x_{t2}, \dots, x_{tL}]$, is the 1 $\times$ RL vector of concatenated lagged cases; and B is the RL $\times$ R matrix county and time lag associations. For T time points, this is

$$\underbrace{X}_{(T-L)\times R} = \underbrace{U}_{(T-L)\times RL}\underbrace{B}_{(RL)\times R} + \underbrace{E}_{(T-L)\times R}$$

As typical with VAR models we use the vec notation

$$\begin{aligned}\underline{x} &= vec(X)\\ \underline{\beta} &= vec(B)\\ \underline{e} &= vec(E)\end{aligned}$$

where $vec(X(s))$ denotes the columns of X stacked on top of each other. Thus, we write

$$\underbrace{\underline{x}}_{(T-L)R\times 1} = (\underbrace{I}_{R\times R} \otimes \underbrace{U}_{(T-L)\times RL})\underbrace{\underline{\beta}}_{(RL)R\times 1} + \underbrace{\underline{e}}_{(T-L)R\times 1}$$

where $e \sim N(0, \Xi \otimes I)$.

This is more or less a linear regression problem. Estimating the values of $\underline{\beta}$ using a sparsity inducing method will give us our network. Non-zero values in $\underline{\beta}$ will indicate that there is an association between the corresponding county and time lag.