Work Work

Jyotishka & his dear Aytijhya

HELLO

Let the vector $X = (X_1, X_2, ..., X_k)$ and Cov(X) = V which is partitioned as

$$\begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

where $V_{11} = Cov(Y_1), V_{22} = Cov(Y_2)$ and $V_{12} = V_{21} = Cov(Y_1, Y_2)$, with $Y_1 = (X_1, X_2)$, $Y_2 = (X_3, ..., X_k)$. \ Then $V_{11.2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$.\ Consider the individual elements of

$$V_{11.2} = \begin{bmatrix} v_{11.2} & v_{12.2} \\ v_{21.2} & v_{22.2} \end{bmatrix}$$

We would like to compare, between $v_{12.2}/v_{11.2}$ and $v_{12.2}/v_{22.2}$ and keep whichever is bigger among them and discard the other one. This would give regression coefficient between each pair of random variables eliminating the effect of the other variables. This way one can select the coefficients W matrix which would be non-zero, in the model

$$Y = WY + \epsilon$$

and would be a way to see directional dependence in the fixed time period. Then we would move the window. This would give a way to do path analysis.

This process can be compared with other ones, such backward substitution or forward selection, eliminating the multi colinearity effect.

Implementation

For each district, we started with only those districts which we found out to be possible regressors for that district according to the comparison rule stated above, and performed backward regression.We have performed the computation of the estimated W matrix for 14 overlapping timespans, viz.

- Day 41 to Day 100
- Day 71 to Day 130
- Day 100 to Day 160

- …
- …
- …
- Day 400 to Day 460
- Day 430 to Day 490

Estimation of W in the model $Y = WY + \epsilon$

Traditional method

Rigobon's Method of Partitioning

GKB Method

Estimation reduced model $Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + \delta_t$

We estimate $ B_1$, B_2 using Vector Auto-regressive(2) model.

Estimation of the original spatio-temporal model $Y_t = WY_t + A_1 Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$

Method of estimation of W is discussed in section 1.(Why this W and the W of the model in section 1 should be same?)\ For i = 1,2; $A_i = (I - W)B_i$

$$\begin{aligned}
\text{cov}(p, q) &= \text{cov}(\alpha q + \varepsilon, \beta p + \eta) \\
&= \alpha\beta \cdot \text{cov}(p, q) + \alpha \cdot \text{cov}(q, \eta) + \beta \cdot \text{cov}(p, \varepsilon) \\
&= \alpha\beta \cdot \text{cov}(p, q) + \alpha \cdot \text{cov}(q, q - \beta p) + \beta \cdot \text{cov}(p, p - \alpha q) \\
&= \alpha \cdot \text{var}(q) + \beta \cdot \text{var}(p) - \alpha\beta \cdot \text{cov}(p, q)
\end{aligned}$$

Therefore, we have:

$$\text{cov}(p, q) = \frac{1}{1 + \alpha\beta}\Big[\alpha \cdot \text{var}(q) + \beta \cdot \text{var}(p)\Big]$$

Also,we calculated

$$var(q) = \frac{\beta^2 \sigma^2_{\eta,s} + \sigma^2_{\epsilon,s}}{(1 - \alpha\beta)^2}$$

$$var(p) = \frac{\alpha^2 \sigma^2_{\epsilon,s} + \sigma^2_{\eta,s}}{(1 - \alpha\beta)^2}$$

For $s = 1, 2$; we have 6 equations using the sample estimates of the variances and covariances over the two regimes and there are 6 unknowns :

$\alpha, \beta, \sigma^2_{\epsilon,1}, \sigma^2_{\epsilon,2}, \sigma^2_{\eta,1}, \sigma^2_{\eta,2}$

An Approach to estimate W, V_1, and V_2

We have been following the following spatio-temporal model:

$$Y_t = WY_t + A_1 Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$$

where ε_t is the unknown error vector at time t. Note that in this model, W, A_1, and A_2 are unknown $d \times d$ matrices.\

As of now, we shall assume that the errors over time are independently and identically distributed. We can perform the following calculation:

$$(I - W)Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t \implies Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + \delta_t$$

where B_1, B_2, δ_t are obtained by pre-multiplying A_1, A_2, ε_t by $(I - W)^{-1}$. \begin{align*}