

# Work Work

Jyotishka & his dear Aytijhya

## HELLO

Let the vector  $X = (X_1, X_2, \dots, X_k)$  and  $\text{cov}(X) = V$  which can be partitioned as

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

where  $V_{11} = \text{cov}(Y_1)$ ,  $V_{22} = \text{cov}(Y_2)$  and  $V_{12} = V_{21} = \text{cov}(Y_1, Y_2)$ , with  $Y_1 = (X_1, X_2)$ ,  $Y_2 = (X_3, \dots, X_k)$ . \ 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

[https://docs.google.com/spreadsheets/d/1RDhOMN0rF\\_tmNRCtIUM1PpAdJuo0pNU9t/edit?usp=sharing&oid=115423038723750570450&rtpof=true&sd=true](https://docs.google.com/spreadsheets/d/1RDhOMN0rF_tmNRCtIUM1PpAdJuo0pNU9t/edit?usp=sharing&oid=115423038723750570450&rtpof=true&sd=true)

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

### Traditional method

### Rigobon's Method of Partitioning

### GKB Method

## Estimation reduced model $Y_t = B_1Y_{t-1} + B_2Y_{t-2} + \delta_t$

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

## Estimation of the original spatio-temporal model $Y_t = WY_t + A_1Y_{t-1} + A_2Y_{t-2} + \varepsilon_t$

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:

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

Also, we calculated

$$\begin{aligned}\text{var}(q) &= \frac{\beta^2 \sigma_{\eta, s}^2 + \sigma_{\varepsilon, s}^2}{(1 - \alpha\beta)^2} \\ \text{var}(p) &= \frac{\alpha^2 \sigma_{\varepsilon, s}^2 + \sigma_{\eta, s}^2}{(1 - \alpha\beta)^2}\end{aligned}$$

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 :

$$\alpha, \beta, \sigma_{\varepsilon, 1}^2, \sigma_{\varepsilon, 2}^2, \sigma_{\eta, 1}^2, \sigma_{\eta, 2}^2$$

## An Approach to estimate $W$ , $V_1$ and $V_2$

We have been working with the following spatio-temporal model:

$$Y_t = WY_t + A_1Y_{t-1} + A_2Y_{t-2} + \varepsilon_t$$

where  $\varepsilon_t$  is the unknown error vector at time  $t$ . Note that in this model,  $W$ ,  $A_1$ , and  $A_2$  are unknown, time-invariant,  $d \times d$  matrices, i.e. parameters of the spatio-temporal model.

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

$$(I - W)Y_t = A_1Y_{t-1} + A_2Y_{t-2} + \varepsilon_t \implies Y_t = B_1Y_{t-1} + B_2Y_{t-2} + \delta_t$$

where  $B_1$ ,  $B_2$ ,  $\delta_t$  are obtained by pre-multiplying  $A_1$ ,  $A_2$ ,  $\varepsilon_t$  by  $(I - W)^{-1}$ . The equation on the LHS is the actual model, and the transformed one on the RHS is a reduced model.

Note that the reduced model is nothing but an usual VAR(2) model, with parameters  $B_1$  and  $B_2$ . So, we can simply estimate  $B_1$  and  $B_2$  using traditional methodologies. Let us denote their estimates by  $\hat{B}_1$  and  $\hat{B}_2$  respectively.

Now, we can estimate the error vectors in the following manner:

$$(I - W)^{-1}\varepsilon_t = \delta_t = Y_t - (\hat{B}_1Y_{t-1} + \hat{B}_2Y_{t-2})$$

We shall be making an assumption that the covariance matrix of  $\varepsilon_t$ , say  $V$ , is a diagonal matrix.

Taking covariance on both sides, we get that

$$\begin{aligned} (I - W)^{-1}\varepsilon_t &= Y_t - (\hat{B}_1Y_{t-1} + \hat{B}_2Y_{t-2}) \\ \implies \text{cov}((I - W)^{-1}\varepsilon_t) &= \text{cov}(Y_t - (\hat{B}_1Y_{t-1} + \hat{B}_2Y_{t-2})) \\ \implies (I - W)^{-1}V_\varepsilon((I - W)^{-1})^T &= \text{ResCov}_Y \\ \implies (I - W)^TV_\varepsilon^{-1}(I - W) &= \text{ResCov}_Y^{-1} \end{aligned}$$

We shall be writing the last equation as:  $M^TVM = C$ , where we are defining:  $M := I - W$ ,  $V := V_\varepsilon^{-1}$ , and  $C := \text{ResCov}_Y^{-1}$ .

This is to be noted that by our assumption, the following structures have been imposed on the matrices that we are supposed to estimate:

- The principal diagonal of  $W$  should contain only 0's. Naturally, the principal diagonal of  $M$  should contain only 1's.
- $V_\varepsilon$  is, by assumption, a diagonal matrix with strictly positive entries. Thus, the matrix  $V$  should be a diagonal matrix, with its diagonal entries equal to the reciprocals of the diagonal entries of  $V_\varepsilon$ .

- Being a covariance matrix,  $\text{ResCov}_Y$  is a known, symmetric, positive definite matrix (That is to be estimated from data). Since the inverse of a positive definite matrix is necessarily positive definite,  $C$  is a known, symmetric, positive definite matrix.
- All the matrices mentioned above are  $d \times d$  matrices.  $M$  and  $V$  are to be estimated, and  $C$  is a known, symmetric, positive definite matrix.

### Estimation for two regimes

We can partition our time series dataset into multiple regimes, and estimate  $M$  and  $V$  accordingly. For simplicity, we may consider two regimes. In general, for the  $s$ -th regime, we shall be denoting the the corresponding matrices by  $M$ ,  $V_s$  and  $C_s$ . Note that  $M$  is independent of  $s$ , because we are assuming that the spatial matrix  $W$  remains invariant over time.

In the case of  $s = 2$ , i.e. two regimes, we are required to solve for  $M$ ,  $V_1$  and  $V_2$  the following system of equations:

$$M^T V_1 M = C_1 \quad \text{and} \quad M^T V_2 M = C_2$$

where the structures of all the matrices have been discussed earlier.

We can start off by writing down a spectral decomposition of  $C_2$ , let's say  $C_2 = D_2^T \Sigma_2 D_2$  where  $\Sigma_2$  is a diagonal matrix containing eigenvalues of  $C_2$ , and the rows of  $D_2$  contain the corresponding eigenvectors.