Theoretical framework VAR model

Michael Aichoun

April 2023

Formalisation mathématiques du modèle VAR

Le modèle VAR (Vector Autoregression) est un modèle statistique qui permet de modéliser la dynamique des relations entre plusieurs séries chronologiques. Il permet de capturer les corrélations entre les variables à différents moments dans le temps. En considèrant le jeu de données de la manière suivante :

- $(Z_t)_{t\geq 0}$ la série temporelle stationnaire associée à la production.
- $(Y_t)_{t\geq 0}$ la série temporelle stationnaire associée aux prix du baril de Brent.
- $(X_t)_{t\geq 0}$ la série temporelle stationnaire associée à l'activité économique globale.

Alors les équations réduites du modèle à estimer sont les suivantes pour un VAR d'ordre p :

$$\begin{cases} Z_t = c_1 + a_{111}Z_{t-1} + ... + a_{p11}Z_{t-p} + a_{112}Y_{t-1} + ... + a_{p12}Y_{t-p} + a_{113}X_{t-1} + ... + a_{p13}X_{t-p} + \epsilon_{1t} \\ Y_t = c_2 + a_{121}Z_{t-1} + ... + a_{p21}Z_{t-p} + a_{122}Y_{t-1} + ... + a_{p22}Y_{t-p} + a_{123}X_{t-1} + ... + a_{p23}X_{t-p} + \epsilon_{2t} \\ X_t = c_3 + a_{131}Z_{t-1} + ... + a_{p31}Z_{t-p} + a_{132}Y_{t-1} + ... + a_{p32}Y_{t-p} + a_{133}X_{t-1} + ... + a_{p33}X_{t-p} + \epsilon_{3t} \end{cases}$$

Où :

- c_1, c_2 et c_3 sont des constantes.
- a_{pij} représente le décalage p du coefficient associé à la variable j dans l'équation i.
- ϵ_{it} est un terme d'erreur aléatoire de loi $\mathcal{N}(0, \sigma^2)$

À l'aide de la méthode des Moindres Carrés Ordinaires, on détermine l'ordre optimal p du VAR en se servant des critères AIC et BIC comme indice de sélection par cross-validation. Ainsi une fois déterminé l'ordre p optimal, on transforme notre VAR d'ordre p en un VAR d'ordre 1. Si l'on considère nos données, cela donne :

$$\begin{cases} \mathcal{Z}_t = F_{11}\mathcal{Z}_{t-1} + F_{12}\mathcal{Y}_{t-1} + F_{13}\mathcal{X}_{t-1} + \gamma_{1t} \\ \mathcal{Y}_t = F_{21}\mathcal{Z}_{t-1} + F_{22}\mathcal{Y}_{t-1} + F_{23}\mathcal{X}_{t-1} + \gamma_{2t} \\ \mathcal{X}_t = F_{31}\mathcal{Z}_{t-1} + F_{32}\mathcal{Y}_{t-1} + F_{33}\mathcal{X}_{t-1} + \gamma_{3t} \end{cases}$$

Où :

$$\mathcal{Z}_t = \begin{pmatrix} Z_t \\ \vdots \\ Z_{t-p+1} \end{pmatrix}, \mathcal{Y}_t = \begin{pmatrix} Y_t \\ \vdots \\ Y_{t-p+1} \end{pmatrix}, \mathcal{X}_t = \begin{pmatrix} X_t \\ \vdots \\ X_{t-p+1} \end{pmatrix}, F_{12} = \begin{pmatrix} a_{112} & a_{212} & \cdots & a_{p12} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\text{et } \gamma_{2t} = \begin{pmatrix} \epsilon_{2t} \\ \vdots \\ 0 \end{pmatrix}$$

Modélisation sous forme primitive pour un VAR(1) :

$$\begin{cases} Z_t = d_{10} + d_{11}Y_t + d_{12}X_t + d_{13}Z_{t-1} + d_{14}Y_{t-1} + d_{15}X_{t-1} + \delta_{1t} \\ Y_t = d_{20} + d_{21}Z_t + d_{22}X_t + d_{23}Z_{t-1} + d_{24}Y_{t-1} + d_{25}X_{t-1} + \delta_{2t} \\ X_t = d_{30} + d_{31}Z_t + d_{32}Y_t + d_{33}Z_{t-1} + d_{34}Y_{t-1} + d_{35}X_{t-1} + \delta_{3t} \end{cases}$$

Qui se ré-écrit sous forme matricielle de la manière suivante :

$$\begin{pmatrix} Z_t \\ Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} d_{10} \\ d_{20} \\ d_{30} \end{pmatrix} + \begin{pmatrix} 0 & d_{11} & d_{12} \\ d_{21} & 0 & d_{22} \\ d_{31} & d_{32} & 0 \end{pmatrix} \begin{pmatrix} Z_t \\ Y_t \\ X_t \end{pmatrix} + \begin{pmatrix} d_{13} & d_{14} & d_{15} \\ d_{23} & d_{24} & d_{25} \\ d_{33} & d_{34} & d_{35} \end{pmatrix} \begin{pmatrix} Z_{t-1} \\ Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{pmatrix}$$

Ce qui nous amène à l'équation structurelle :

$$\begin{pmatrix} 1 & -d_{11} & -d_{12} \\ -d_{21} & 1 & -d_{22} \\ -d_{31} & -d_{32} & 1 \end{pmatrix} \begin{pmatrix} Z_t \\ Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} d_{10} \\ d_{20} \\ d_{30} \end{pmatrix} + \begin{pmatrix} d_{13} & d_{14} & d_{15} \\ d_{23} & d_{24} & d_{25} \\ d_{33} & d_{34} & d_{35} \end{pmatrix} \begin{pmatrix} Z_{t-1} \\ Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{pmatrix}$$

Pour simplifier, on prend les notations ci-dessous :

$$\mathrm{A} = \begin{pmatrix} 1 & -d_{11} & -d_{12} \\ -d_{21} & 1 & -d_{22} \\ -d_{31} & -d_{32} & 1 \end{pmatrix}, N_t = \begin{pmatrix} Z_t \\ Y_t \\ X_t \end{pmatrix}, \lambda = \begin{pmatrix} d_{10} \\ d_{20} \\ d_{30} \end{pmatrix}, \mathrm{B} = \begin{pmatrix} d_{13} & d_{14} & d_{15} \\ d_{23} & d_{24} & d_{25} \\ d_{33} & d_{34} & d_{35} \end{pmatrix}$$

$$\text{et } u_t = \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{pmatrix}$$

Ainsi on obtient d'une part :

$$N_t = A^{-1}\lambda + A^{-1}BN_{t-1} + A^{-1}u_t$$

Puis d'autre part à l'aide de la méthode OLS, on est capable empiriquement de récupérer l'équation réduite suivante :

$$N_t = \alpha + CN_{t-1} + e_t$$

Ainsi par identification, on estime les coefficients de notre matrice A et B en résolvant le système ci-dessous :

$$\begin{cases} A^{-1}\lambda = \alpha \\ A^{-1}B = C \\ A^{-1}u_t = e_t \end{cases}$$