

3.3 Analyse structurelle avec modèles VAR

Comme les modèles *VAR* représentent les corrélations entre un ensemble de variables, ils sont souvent utilisés pour analyser certains aspects des relations entre les variables d'intérêt. On donne trois outils qui permettent d'interpréter un modèle *VAR*.

3.3.1 Causalité au sens de Granger

Une des questions posées par Sims (1980) avec les processus *VAR* est de savoir s'il existe une relation de causalité entre les différentes variables du système. Il existe plusieurs définitions de la causalité. On présente la causalité au sens de Granger (1969) qui est la plus fréquemment utilisée. L'idée est que la cause ne peut pas venir après l'effet et donc elle sera utile pour améliorer la prévision.

Définitions

1) Y cause X à la date t si et seulement si :

$$E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}) \neq E(X_t / \underline{X}_{t-1})$$

où \underline{X}_{t-1} est le passé de X jusqu'à l'instant $t - 1$. Le passé de Y apporte une information supplémentaire par rapport à la seule prise en compte du passé de X .

2) Y cause instantanément X à la date t si et seulement si :

$$E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}, Y_t) \neq E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1})$$

La valeur présente de Y apporte une information de plus.

Exemple

Soit les processus suivants :

$$\begin{cases} X_t = \varepsilon_t + 0.5Z_t + 0.9Z_{t-1} \\ Y_t = Z_t \end{cases}$$

où ε_t et Z_t sont 2 BB indépendants, alors

$$\begin{aligned} E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}) &= 0.9Z_{t-1} \\ E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}, Y_t) &= 0.5Z_t + 0.9Z_{t-1} \end{aligned}$$

Donc Y_t cause instantanément X_t .

Caractérisation de la causalité au sens de Granger

Pour déterminer les relations de causalité entre les variables d'un processus Y_t de dimension k , on le partitionne en Z_t de dimension m et X_t de dimension $(k - m)$:

$$Y_t = \begin{pmatrix} Z_t \\ X_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} A_{11,1} & A_{12,1} \\ A_{21,1} & A_{22,1} \end{pmatrix} \begin{pmatrix} Z_{t-1} \\ X_{t-1} \end{pmatrix} + \dots + \begin{pmatrix} A_{11,p} & A_{12,p} \\ A_{21,p} & A_{22,p} \end{pmatrix} \begin{pmatrix} Z_{t-p} \\ X_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \quad (*)$$

Alors on a le résultat suivant :

Corollaire

Si Y_t est un processus $VAR(p)$ stationnaire comme dans (*), alors :

X_t ne cause pas $Z_t \iff A_{12,i} = 0$ pour $i = 1, \dots, p$.

Alternativement :

Z_t ne cause pas $X_t \iff A_{21,i} = 0$ pour $i = 1, \dots, p$.

3.3.2 Analyse ou fonction de réponses impulsionnelles (ARI ou FRI)

L'ARI va permettre de déterminer l'influence d'un choc relié à l'évolution d'une des variables sur les autres variables du système. Par exemple : soit un $VAR(1)$: $X_t = AX_{t-1} + \varepsilon_t$, on veut analyser l'effet d'un choc en $t = 0$, on aura ε_0 qui portera le choc puis la suite des $\varepsilon_i = 0$ pour $i = 1, \dots, t$ d'où

$$\begin{aligned} X_0 &= \varepsilon_0 \\ X_1 &= AX_0 \\ X_2 &= AX_1 = A^2X_0 \\ &\vdots \\ X_t &= A^tX_0. \end{aligned}$$

Si on suppose que le choc est égal à 1 et porte sur la 1ère variable, on définit $\varepsilon'_0 = (1, 0, \dots, 0)$, alors la 1ère colonne de A^t représentera l'effet sur le système d'un choc sur la 1ère variable après t périodes et comme

$$(I - AL)^{-1} = \Psi(L) = \sum_{i=0}^{\infty} A^i L^i$$

Les coefficients de la forme $MA(\infty)$ (formule de Wold) vont donner la suite des RI du système à un choc unitaire sur les innovations du processus avec le multiplicateur dynamique :

$$\frac{\partial X_{t+h}}{\partial \varepsilon'_t} = \psi_h.$$

En faisant le graphe de $\frac{\partial X_{i,t+h}}{\partial \varepsilon_{j,t}} = \psi_{i,j}^{(h)}$, pour $h = 0, 1, 2, \dots$, on obtient la FRI avec $\psi_{i,j}^{(h)}$:

l'effet du choc j sur la variable i , h périodes après le choc.

Problème : On peut avoir des corrélations instantanées entre les erreurs donc une impulsion sur $\varepsilon_{2,t}$ va influencer automatiquement $\varepsilon_{1,t}$.

Solution : Dans le cas où $E(\varepsilon_t \varepsilon'_t) = \Sigma$ est non diagonale alors il faut orthogonaliser les innovations en utilisant la décomposition de Cholesky et utiliser la représentation VAR Structurelle (SVAR) où le terme d'erreur est orthogonal.

Pour avoir le SVAR il suffit de multiplier $X_t = AX_{t-1} + \varepsilon_t$ par P^{-1} avec $\Sigma = PP'$:

$$P^{-1}X_t = P^{-1}AX_{t-1} + P^{-1}\varepsilon_t$$

On pose $u_t = P^{-1}\varepsilon_t$: qui est le choc structurel (orthogonal) car

$$E(u_t u'_t) = E(P^{-1}\varepsilon_t \varepsilon'_t P^{-1'}) = P^{-1}\Sigma P^{-1'} = I$$

on a l'indépendance entre $u_{1,t}$ et $u_{2,t}$. Pour avoir l'ARI on utilise le VMA(∞) structurel avec $\Sigma = PP'$

$$\begin{aligned} X_t &= \mu + \sum_{j=0}^{\infty} \psi_j P P^{-1} \varepsilon_{t-j} \\ &= \mu + \sum_{j=0}^{\infty} C_j u_{t-j} \end{aligned}$$

avec $C_j = \psi_j P$ et $u_t = P^{-1} \varepsilon_t$.

Les colonnes de C_j nous donne l'ARI par rapport à un choc indépendant et normalisé sur l'innovation d'une variable après j périodes.

Suite de l'exemple : Σ est non diagonale alors on applique la décomposition de Cholesky

$$\begin{aligned} \Sigma &= \begin{pmatrix} P_1 & 0 \\ P_2 & P_3 \end{pmatrix} \begin{pmatrix} P_1 & P_2 \\ 0 & P_3 \end{pmatrix} \\ &= \begin{pmatrix} P_1^2 & P_1 P_2 \\ P_1 P_2 & P_1^2 + P_3^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1.49 \end{pmatrix} \end{aligned}$$

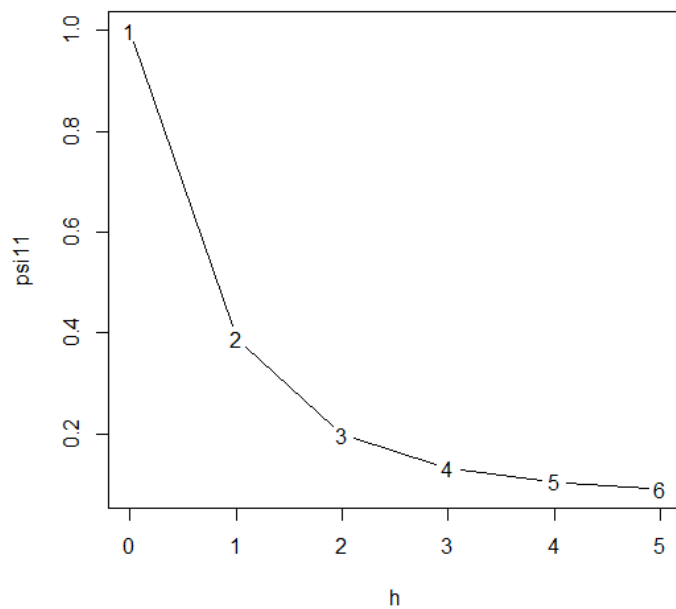
Donc $P = \begin{pmatrix} 1 & 0 \\ 0.7 & 1 \end{pmatrix}$, d'où le $VMA(\infty)$ structurel est :

$$\begin{aligned} X_t &= \sum_{j=0}^{\infty} \begin{pmatrix} 0.6 & -0.3 \\ -0.3 & 0.6 \end{pmatrix}^j P P^{-1} \varepsilon_{t-j} \\ &= \begin{pmatrix} 1 & 0 \\ 0.7 & 1 \end{pmatrix} u_t + \sum_{j=1}^{\infty} \left[\begin{pmatrix} 0.6 & -0.3 \\ -0.3 & 0.6 \end{pmatrix}^j P \right] u_{t-j} \end{aligned}$$

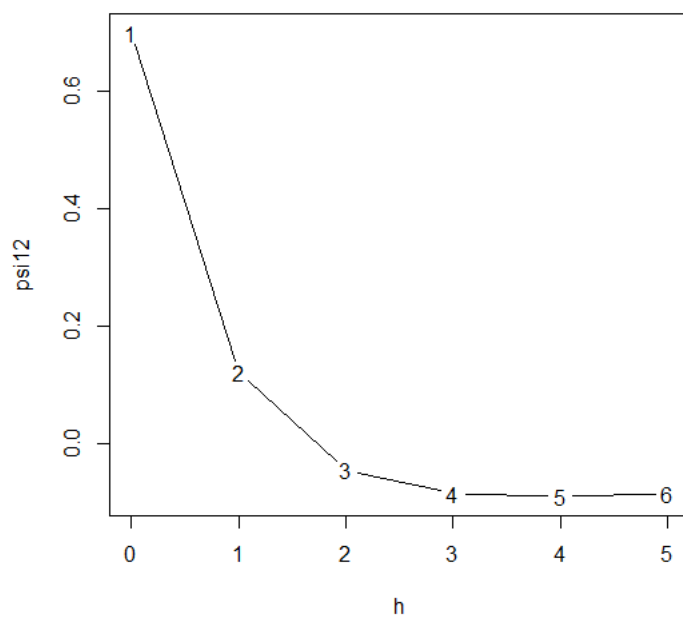
Nous allons maintenant faire le graphe de l'ARI pour la première variable, pour cela nous allons considérer 5 périodes après l'innovation.

$$\begin{aligned} X_t &= \begin{pmatrix} 1 & 0 \\ 0.7 & 1 \end{pmatrix} u_t + \begin{pmatrix} 0.39 & -0.3 \\ 0.12 & 0.6 \end{pmatrix} u_{t-1} + \begin{pmatrix} 0.198 & -0.36 \\ -0.045 & 0.45 \end{pmatrix} u_{t-2} + \begin{pmatrix} 0.1323 & -0.351 \\ -0.0864 & 0.378 \end{pmatrix} u_{t-3} \\ &\quad + \begin{pmatrix} 0.1053 & -0.324 \\ -0.09153 & 0.3321 \end{pmatrix} u_{t-4} + \begin{pmatrix} 0.090 & -0.294 \\ -0.086 & 0.296 \end{pmatrix} u_{t-5} + \dots \end{aligned} \quad (**)$$

alors : $\psi_{1,1}^{(0)} = 1, \psi_{1,1}^{(1)} = 0.39, \psi_{1,1}^{(2)} = 0.198, \psi_{1,1}^{(3)} = 0.1323, \psi_{1,1}^{(4)} = 0.1053, \psi_{1,1}^{(5)} = 0.09$. De la même manière on obtient $\psi_{1,2}^{(h)}$.



Effet d'un choc de la V_1 sur V_1 ($V_1 \rightarrow V_1$)



Effet d'un choc de la V_2 sur V_1 ($V_2 \rightarrow V_1$)

Dans les deux cas on voit que l'effet du choc diminue et tend vers 0 rapidement due à la stabilité du système.

Remarque :

L'orthogonalisation implique que l'ordre dans lequel sont disposées les variables du VAR affecte l'analyse des FRI. Dans le VAR, la variable que l'on suppose économiquement être la variable explicative doit être après la variable dont on veut expliquer les évolutions.

$$VAR = \begin{pmatrix} \text{Variable à expliquer} \\ \text{Variable explicative} \end{pmatrix}.$$

par exemple pour les séries de revenu (R_t) et de consommation (C_t) des ménages on aura :

$$X_t = \begin{pmatrix} C_t \\ R_t \end{pmatrix} \text{ et pour les séries de production } (P_t) \text{ et de ventes } (V_t) \text{ on aura : } X_t = \begin{pmatrix} P_t \\ V_t \end{pmatrix}.$$

De manière générale quand on veut analyser l'impact de la série S_1 sur la série S_2 on utilise

$$\text{le VAR : } X_t = \begin{pmatrix} S_2 \\ S_1 \end{pmatrix}$$

3.3.3 Décomposition de la variance

L'objectif est de calculer la contribution de chacune des innovations orthogonales à la variance de l'erreur ce qui donnera une information sur l'importance relatif de chaque innovation. On commence par calculer la variance totale de l'erreur de prévision à un horizon h puis la variance de l'erreur attribué aux variables et enfin on divise sur la variance totale et on obtient les poids en pourcentages.

Soit le $VAR(p)$ suivant :

$$\Phi(L) X_t = \varepsilon_t \iff X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} : \quad VMA(\infty)$$

L'erreur de prévision à l'horizon h est

$$\begin{aligned} X_{t+h} - \hat{X}_t(h) &= X_{t+h} - E(X_{t+h}/I_t) \\ &= \sum_{i=0}^{h-1} \psi_i \varepsilon_{t+h-i} \end{aligned}$$

Donc

$$\begin{aligned}
 E \left(X_{t+h} - \widehat{X}_t(h) \right) &= 0 \\
 E \left(\left(X_{t+h} - \widehat{X}_t(h) \right) \left(X_{t+h} - \widehat{X}_t(h) \right)' \right) &= \Sigma + \sum_{i=1}^{h-1} \psi_i \Sigma \psi_i' \\
 &= \sum_{i=0}^{h-1} C_i C_i'
 \end{aligned}$$

la diagonale représente la variance totale des variables d'où la proportion de la variance de l'erreur de prévision à l'horizon h de la variable i attribué à l'innovation $\varepsilon_{j,t}$ est

$$\omega_{i,j}^{(h)} = \frac{\text{variance de l'erreur attribué à } \varepsilon_{j,t}}{\text{Variance totale de la variable } X_{i,t}} \times 100\%, i, j = 1, \dots, k.$$

Suite de l'exemple :

Faire la décomposition de la variance de l'erreur prévisionnelle pour la première variable pour $h = 1, \dots, 5$.

Pour $h = 1$, $X_{t+1} - \widehat{X}_t(1) = \begin{pmatrix} 1 & 0 \\ 0.7 & 1 \end{pmatrix} u_{t+1} = \begin{pmatrix} u_{1,t+1} \\ 0.7u_{1,t+1} + u_{2,t+1} \end{pmatrix}$ (trouver à partir de(**)) d'où

$$E \left(\left(X_{t+1} - \widehat{X}_t(1) \right) \left(X_{t+1} - \widehat{X}_t(1) \right)' \right) = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1.49 \end{pmatrix}$$

Donc

$$E \left(X_{1,t+1} - E \left(X_{1,t+1} / I_t \right) \right)^2 = E \left(u_{1,t+1}^2 \right) = 1$$

$$\omega_{1,1}^{(1)} = \frac{1}{1} = 100\%, \omega_{1,2}^{(1)} = \frac{0}{1} = 0\%.$$

Pour $h = 2$, $X_{t+2} - \widehat{X}_t(2) = \begin{pmatrix} 1 & 0 \\ 0.7 & 1 \end{pmatrix} u_{t+2} + \begin{pmatrix} 0.39 & -0.3 \\ 0.12 & 0.6 \end{pmatrix} u_{t+1}$ d'où

$$\begin{aligned}
 E \left(\left(X_{t+2} - \widehat{X}_t(2) \right) \left(X_{t+2} - \widehat{X}_t(2) \right)' \right) &= \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1.49 \end{pmatrix} + \begin{pmatrix} 0.39 & -0.3 \\ 0.12 & 0.6 \end{pmatrix} \times \begin{pmatrix} 0.39 & 0.12 \\ -0.3 & 0.6 \end{pmatrix} \\
 &= \begin{pmatrix} 1.2421 & 0.5668 \\ 0.5668 & 1.8644 \end{pmatrix}
 \end{aligned}$$

Donc

$$\begin{aligned} E(X_{1,t+2} - E(X_{1,t+2}/I_t))^2 &= E(u_{1,t+2} + 0.39u_{1,t+1} - 0.3u_{2,t+1})^2 \\ &= 1.2421 \end{aligned}$$

$$\omega_{1,1}^{(2)} = \frac{1 + 0.39^2}{1.2421} = 92.754\%, \omega_{1,2}^{(2)} = \frac{0.3^2}{1.2421} = 7.2458\%.$$

$$\text{Pour } h = 3, X_{t+3} - \hat{X}_t(3) = \begin{pmatrix} 1 & 0 \\ 0.7 & 1 \end{pmatrix} u_{t+3} + \begin{pmatrix} 0.39 & -0.3 \\ 0.12 & 0.6 \end{pmatrix} u_{t+2} + \begin{pmatrix} 0.198 & -0.36 \\ -0.045 & 0.45 \end{pmatrix} u_{t+1}$$

d'où

$$\begin{aligned} E\left(\left(X_{t+3} - \hat{X}_t(3)\right)\left(X_{t+3} - \hat{X}_t(3)\right)'\right) &= \begin{pmatrix} 1.2421 & 0.5668 \\ 0.5668 & 1.8644 \end{pmatrix} + \begin{pmatrix} 0.198 & -0.36 \\ -0.045 & 0.45 \end{pmatrix} \times \begin{pmatrix} 0.198 & -0.045 \\ -0.36 & 0.45 \end{pmatrix} \\ &= \begin{pmatrix} 1.4109 & 0.39589 \\ 0.39589 & 2.0689 \end{pmatrix} \end{aligned}$$

Donc

$$\begin{aligned} E(X_{1,t+3} - E(X_{1,t+3}/I_t))^2 &= E(u_{1,t+3} + 0.39u_{1,t+2} - 0.3u_{2,t+2} + 0.198u_{1,t+1} - 0.36u_{2,t+1})^2 \\ &= 1.4109 \end{aligned}$$

$$\omega_{1,1}^{(3)} = \frac{1 + 0.39^2 + 0.198^2}{1.4109} = 84.436\%, \omega_{1,2}^{(3)} = \frac{0.3^2 + 0.36^2}{1.4109} = 15.565\%$$

Je vous laisse le plaisir de faire les calculs jusqu'à $h = 6$. Généralement on présente les résultats sous forme de tableau

Erreur de prévision en $X_{1,t}$	Horizon (h)	Décomposition de la variance attribué à l'innovation de	
		$X_{1,t}$	$X_{2,t}$
	1	100%	0%
	2	92.754%	7.2458%
	3	84.436%	15.565%

On remarque qu'on a 100% de la variance de l'erreur prévisionnelle de $X_{1,t}$ pour l'horizon 1 due à ses propres innovations et pour l'horizon 3, on a 84.436% et 15.565% de la variance de l'erreur due aux innovations des variables $X_{1,t}$ et $X_{2,t}$ respectivement.

Bibliographie

- [1] Brockwell, J. P. and Davis, R. A. (2002). Introduction to Time Series and Forecasting. 2nd ed. Springer texts in statistics. Springer-Verlag New York, Inc.
- [2] Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press. Princeton, New Jersey.
- [3] Kirchgässner, G. and Wolters, J. (2007). Introduction to modern time series analysis. ISBN 978-3-540-73290-7 Springer Berlin Heidelberg New York.
- [4] Lütkepohl, H. (2005). New Introduction to Multiple Time Series Analysis. Springer-Verlag Berlin Heidelberg.
- [5] Shumway, R. H. and Stoffer, D. S. (2011). Time Series Analysis and Its Applications With R Examples. Springer Science+Business Media, LLC 2011.