

## Chapitre 4

# Processus VAR et VARMA

Révision février 2010

### 4.1 Introduction. Exemples

On peut rencontrer des séries temporelles multivariées dans beaucoup de situations. Considérons quelques exemples.

- **Appels arrivant à un centre d'appels.** Une compagnie de distribution de l'eau a un centre d'appels pour ses clients pour ce qui concerne la facturation, les incidents de distribution. La compagnie émet des factures en quantité et à des dates décidées par le service Comptabilité et aimerait prévoir le nombre d'appels qu'elle reçoit chaque jour pendant une semaine pour prévoir l'effectif des opérateurs au centre d'appel. Ici on s'intéresse à une série, le nombre d'appels, qu'on sait lié au nombre de factures envoyées les jours précédents et aux conditions climatiques, gel notamment. Il y a une série aléatoire, le nombre d'appels mais le nombre de factures émis chaque jour est purement explicatif et déterministe. Les conditions météorologiques sont un processus aléatoire mais on n'est pas intéressé par la modélisation du gel dans ce problème. En résumé, l'étude du nombre d'appels reçu chaque jour est essentiellement un problème unidimensionnel, les conditions climatiques, les factures émises sont des informations purement exogènes.

Il y a une série  $y_t$  nombre d'appels le jour  $t$  et des variables explicatives  $x_{1t}, x_{2t}, \dots$  qui peuvent être le fait que le jour  $t$  est un lundi, le nombre de factures envoyées le jour  $t - k$ ,  $k$  à choisir... et on s'intéresse à la régression de  $y_t$  sur  $x_{1t}, x_{2t}, \dots$ . Les variables explicatives ne sont pas aléatoires.

$$y_t = \mathbf{x}'_t \beta + u_t \quad (*)$$

Très probablement, l'erreur  $u_t$  est autocorrélée et pour être efficace, l'estimation de  $\beta$  doit tenir compte de cette autocorrélation. C'est le même type de situation que pour le lac Huron.

- **Ventes d'un produit et campagne de publicité.** Des données classiques en marketing (Données Lydia Pinckam) concernent les ventes et la publicité pour une potion. Cette série célèbre a été souvent étudiée et son histoire racontée.

La vente du produit a commencé en 1873. Elle a été immédiatement accompagnée d'une intense campagne de publicité. Le produit est resté essentiellement inchangé tout au long de la période en question. Il n'y avait pas de produit de substitution sur le marché. La compagnie ne s'appuyait que sur la publicité pour augmenter ses ventes.

Ici on s'intéresse à l'influence de la publicité sur les ventes. Si l'effort publicitaire n'a pas été décidé exogènement on peut considérer qu'on dispose ainsi d'une série bidimensionnelle dont les composantes sont des séries aléatoires.

- **Séries macroéconomiques.** Pour une région, l'investissement, le revenu disponible et les dépenses de consommation sont des séries temporelles dépendantes, sans qu'on puisse dire que telle ou telle série est exogène.

- **Série pilote.** On observe parfois qu'une série dépend à la date  $t$  de la valeur d'une autre série à une date antérieure,  $t - k$ . Les deux séries étant modélisables par des modèles classiques (ARMA...). On se trouve dans un cas restreint par rapport à l'exemple des séries macroéconomiques : une série a un rôle explicatif de l'autre série, ceci entraîne des restrictions sur la nature de la dépendance cherchée.

Nous verrons dans l'e-thème consacré aux fonctions de transfert une façon d'identifier un tel modèle. A travers ces exemples on voit que l'objectif est l'analyse de la dépendance de chaque série par rapport aux autres ou d'une série par rapport à certaines autres séries ... et avec des présupposés variés sur la structure des séries.

Avant de mettre en oeuvre une technique il faut donc bien comprendre le phénomène pour ensuite choisir la bonne technique.

Nous commençons par définir les séries temporelles vectorielles stationnaires et donnons quelques exemples. Ensuite nous introduisons les modèles VAR (Vector Auto Regressive) et VMA (Vector Moving Average) pour lesquels nous nous posons les mêmes questions que pour les ARMA univariés : stationnarité, inversibilité. Enfin nous étudions un exemple.

## 4.2 Série vectorielle stationnaire

### Analyse exploratoire

Pour explorer une série bidimensionnelle  $(y_{1t}, y_{2t})$ ,  $t = 1, \dots, T$ , on examine les composantes. On dessine les chronogrammes (temps en abscisse, séries en ordonnée). On peut aussi dessiner le nuage des points  $(y_{1t}, y_{2t})$  et parfois un nuage avec des séries décalées :  $(y_{1,t-k}, y_{2t})$ .

L'examen des corrélations croisées :  $r_{12}(h)$ ,  $h = \dots, -2, -1, 0, 1, 2, 3 \dots$  ne peut donner des résultats significatifs que si une des deux séries est un bruit blanc. Dans l'e-thème consacré aux fonctions de transfert, nous verrons d'où vient ce problème et comment on peut parfois transformer les séries pour que l'une des deux soit un bruit blanc.

### Stationnarité jointe

Etant donné une série bidimensionnelle  $\{X_t, Y_t\}$ , on définit la fonction moyenne :  $\begin{pmatrix} E(X_t) \\ E(Y_t) \end{pmatrix}$  et les fonctions de covariances :  $\gamma_X(r, s)$ ,  $\gamma_Y(r, s)$ ,  $\gamma_{X,Y}(r, s) = \text{cov}(X_r, Y_s)$ .

On dit que  $\{X_t, Y_t\}$  est (faiblement) stationnaire si :  $\{X_t\}$  et  $\{Y_t\}$  sont (faiblement) stationnaires et  $\gamma_{X,Y}(t+h, t)$  ne dépend que de  $h$ . On note alors  $\gamma_{X,Y}(h) = E(X_{t+h} - \mu_X)(Y_t - \mu_Y)$ .

La fonction

$$h \rightsquigarrow \gamma_{X,Y}(h)$$

est la *fonction de corrélation croisée* entre les séries  $X$  et  $Y$ .

Pour une série vectorielle stationnaire on a :

$$\gamma_{X,Y}(h) = \gamma_{Y,X}(-h)$$

mais

$$\gamma_{X,Y}(h) \neq \gamma_{X,Y}(-h). \quad (4.1)$$

La fonction de corrélation croisée est :

$$\rho_{X,Y}(h) = \frac{\gamma_{X,Y}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}} \quad (4.2)$$

Elle permet de mesurer l'intensité de la liaison entre les séries  $\{X_t\}$  et  $\{Y_t\}$ .

**Exemple de stationnarité jointe.** Définissons :  $X_t = Z_t + Z_{t-1}$  et  $Y_t = Z_t - Z_{t-1}$ , où les  $Z_t$  sont i.i.d.  $(0, \sigma_Z^2)$ . On trouve en particulier que  $\rho_{X,Y}(1) = 0.5$  et  $\rho_{X,Y}(-1) = -0.5$

### 4.2.1 Estimation de la fonction de corrélation croisée

Soit  $\{X_t, Y_t\}$  stationnaire, observée sur  $t = 1, \dots, T$ . On estime  $\gamma_{X,Y}(h)$  par :

$$\hat{\gamma}_{X,Y}(h) = \frac{1}{T} \sum_{t=1}^{T-h} (x_{t+h} - \bar{x})(y_t - \bar{y}) \quad h \geq 0$$

puis  $\hat{\gamma}_{X,Y}(-h)$  par  $\hat{\gamma}_{Y,X}(h)$ . Le coefficient de corrélation croisée empirique d'ordre  $h$  est :

$$\hat{\rho}_{X,Y}(h) = \frac{\hat{\gamma}_{X,Y}(h)}{\sqrt{\hat{\gamma}_X(0)\hat{\gamma}_Y(0)}}$$

**Propriété.** Si au moins un des deux processus  $X, Y$  est un BB (pas nécessairement gaussien mais avec des conditions sur les moments jusqu'à l'ordre 4), alors

$$\hat{\rho}_{X,Y}(h) \simeq N(0, \frac{1}{T})$$

**Exemple empirique.** Illustration de (4.1). Les séries notées SALES et INDEX dans la sortie de la proc ARIMA ci-dessous sont considérées, après certaines transformations comme le préblanchiment (*prewhitened*), conjointement stationnaires. Elles sont reprises dans l'étude des modèles à fonction de transfert, chapitre suivant.

---

```

Correlation of SALES and INDEX
Variable INDEX has been differenced.
Period(s) of Differencing = 1.
Both series have been prewhitened.
Variance of transformed series = 3.794675 and 0.078036
Number of observations =      149
NOTE: The first observation was eliminated by differencing.

                                Crosscorrelations
Lag Covariance Correlation -1  9  8  7  6  5  4  3  2  1  0  1  2  3  4  5  6  7  8  9  1
-6 -0.034684 -0.06374 | . * | . |
-5  0.013016  0.02392 | . | . |
-4  0.0012583  0.00231 | . | . |
-3  0.022045  0.04051 | . | * |
-2  0.0054125  0.00995 | . | . |
-1  0.051478  0.09460 | . | ** |
0  0.034232  0.06291 | . | * |
1  0.043060  0.07913 | . | ** |
2  0.010062  0.01849 | . | . |
3  0.367442  0.67523 | . | ***** |
4  0.246112  0.45227 | . | ***** |
5  0.185447  0.34079 | . | ***** |
6  0.140160  0.25757 | . | ***** |
7  0.145861  0.26804 | . | ***** |
8  0.107803  0.19811 | . | ***** |
9  0.094235  0.17317 | . | *** |
10 0.053115  0.09761 | . | ** |
11 0.078822  0.14485 | . | *** |
12 0.038038  0.06990 | . | * |
13 -0.0078178 -0.01437 | . | . |

```

---

Cette sortie SAS donne le graphe de la fonction de corrélation croisée empirique : entre  $\text{SALES}(t)$  et  $\text{INDEX}(t - \text{Lag})$ .

On constate que  $\text{SALES}(t)$  et  $\text{INDEX}(t-3)$  sont fortement corrélés alors que  $\text{SALES}(t)$  et  $\text{INDEX}(t+3)$  ne le sont pratiquement pas. Pour vérifier cette lecture on a dessiné ci-dessous les graphes de  $(\text{SALES}(t), \text{INDEX}(t+3))$ , figure (4.1) à gauche, et de  $(\text{SALES}(t), \text{INDEX}(t-3))$ , figure (4.2) à droite.

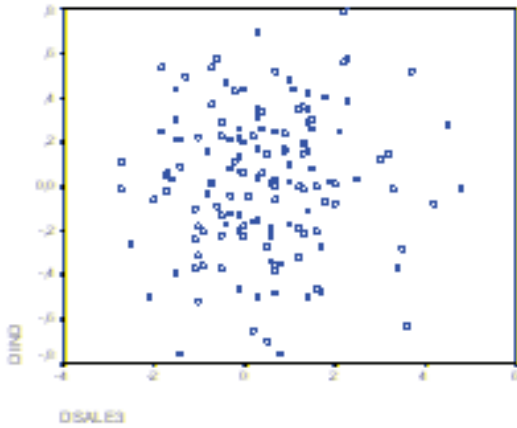


FIGURE 4.1 – INDEX en ordonnée, contre SALES retardé de 3.

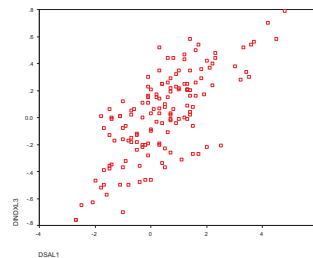


FIGURE 4.2 – INDEX retardé de 3 en ordonnée contre SALES.

*Stationnarité jointe – cas général*

Considérons la série vectorielle des vecteurs aléatoires,  $\mathbf{Y}_t = (Y_{1t}, \dots, Y_{mt})'$  et définissons le vecteur moyenne :

$$\boldsymbol{\mu}_t = \mathbf{E}(\mathbf{Y}_t) = (\mathbf{E}(Y_{1t}), \dots, \mathbf{E}(Y_{mt}))'$$

et la matrice de covariance d'ordre  $h$

$$\boldsymbol{\Gamma}(t+h, t) = \text{cov}(\mathbf{Y}_{t+h}, \mathbf{Y}_t) = \mathbf{E}(\mathbf{Y}_{t+h} - \boldsymbol{\mu}_{t+h})(\mathbf{Y}_t - \boldsymbol{\mu}_t)'$$

d'élément  $(i, j)$  :  $\gamma_{ij}(t+h, t) = \text{cov}(Y_{i,t+h}, Y_{j,t})$ .

La série  $\mathbf{Y}_t$  est stationnaire (faiblement) si  $\boldsymbol{\mu}_t$  et  $\boldsymbol{\Gamma}(t+h, t)$  ne dépendent pas de  $t$  (cf les notes de cours n° 1). Dans ce cas on note :

$$\boldsymbol{\mu} = \mathbf{E}(\mathbf{Y}_t)$$

et

$$\boldsymbol{\Gamma}(h) = \text{cov}(\mathbf{Y}_{t+h}, \mathbf{Y}_t).$$

On a :  $\gamma_{ij}(h) = \gamma_{ji}(-h)$ .

La matrice des autocorrélations d'ordre  $h$  est

$$\mathbf{R}(h) = \begin{pmatrix} \rho_{11}(h) & \cdots & \rho_{1m}(h) \\ \vdots & & \vdots \\ \rho_{m1}(h) & \cdots & \rho_{mm}(h) \end{pmatrix}$$

où

$$\rho_{ij}(h) = \gamma_{ij}(h)(\gamma_{ii}(0)\gamma_{jj}(0))^{-1/2}$$

Ainsi  $\mathbf{R}(h) = \mathbf{D}^{-1}\boldsymbol{\Gamma}(h)\mathbf{D}^{-1}$  où  $\mathbf{D}$  est la matrice diagonale des écart-types des composantes.

### 4.3 Exemples de processus vectoriels stationnaires

#### 4.3.1 Bruit blanc multidimensionnel.

Soit une série vectorielle  $\{\mathbf{Z}_t\}$  à  $m$  composantes. On dit que  $\{\mathbf{Z}_t\}$  est un bruit blanc si et seulement si  $\{\mathbf{Z}_t\}$  est stationnaire de moyenne nulle et de fonction de covariance :

$$\Gamma(h) = \begin{cases} \Sigma & \text{si } h = 0 \\ 0 & \text{si } h \neq 0 \end{cases}$$

On écrit :  $\{\mathbf{Z}_t\} \sim BB(0, \Sigma)$ .

**Test de blancheur d'une série vectorielle** Pour tester  $H_0 : \mathbf{R}(1) = \dots = \mathbf{R}(k) = 0$ , la statistique du portemanteau du cas unidimensionnel devient

$$Q_m(k) = T^2 \sum_{l=1}^k \frac{1}{T-l} \text{tr}(\hat{\Gamma}(l)' \hat{\Gamma}(0)^{-1} \hat{\Gamma}(l) \hat{\Gamma}(0)^{-1}) \quad (4.3)$$

où  $T$  est la taille de l'échantillon,  $m$  la dimension du processus. Sous l'hypothèse nulle et quelques conditions,  $Q_m(k)$  est asymptotiquement  $\chi^2$  à  $m^2 k$  ddl.

Observons que si la série est unidimensionnelle,  $m = 1$ , alors  $Q_m(k)$  revient à la statistique du portemanteau.

Exemple. On a simulé un bruit blanc gaussien bi-dimensionnel, voir en annexe du chapitre (4.6). Par la proc varmax on teste sa blancheur. (Il est nécessaire de régresser le vecteur sur la constante pour que la procédure fasse le test.)

---

```
proc varmax data= st9.simul_bb;
model y1 y2 = un/noint;
run;
...
Covariances croisées des résidus
```

Retard	Variable	y1	y2
0	y1	1.26715	0.26219
	y2	0.26219	2.23424
1	y1	-0.02390	-0.04198
	y2	-0.11906	0.05323
2	y1	0.00486	0.07105
	y2	0.04258	0.04327

```
Corrélations croisées des résidus
```

Retard	Variable	y1	y2
0	y1	1.00000	0.15583
	y2	0.15583	1.00000
1	y1	-0.01886	-0.02495
	y2	-0.07076	0.02382
2	y1	0.00383	0.04223
	y2	0.02531	0.01937

```
Représentation schématique des corrélations croisées des résidus
Variable/Retard  0  1  2  3  4  5  6  7  8  9  10  11  12
y1               +.  ..  ..  ..  -.  ..  ..  ..  ..  ..  ..  ..  ..
y2               .+  ..  ..  ..  ..  ..  ..  ..  ..  ..  ..  ..  ..

+ is > 2*std error, - is < -2*std error, . is between

Test de Portmanteau pour corrélations
```

croisées des résidus			
Jusqu' au retard	DF	Khi 2	Pr > Khi 2
1	4	0.72	0.9491
2	8	0.97	0.9984
3	12	2.48	0.9982
4	16	7.57	0.9607
5	20	11.13	0.9428
6	24	17.34	0.8338
7	28	22.49	0.7582

On n'a fourni qu'un extrait des sorties mais on y voit que, comme on s'y attendait, l'hypothèse nulle de blancheur n'est rejetée pour aucun décalage (ou retard). Notons aussi la façon dont SAS schématise les matrices d'autocorrélation de la série : pour chaque retard on a une matrice  $2 \times 2$  dont les éléments sont symbolisés.

### 4.3.2 Processus VAR

**Processus VAR(1)**  $\{\mathbf{Y}_t\} \in R^m$  est un VAR(1) s'il est stationnaire et vérifie quelque soit  $t$  :

$$\mathbf{Y}_t - \boldsymbol{\mu} = \Phi_1(\mathbf{Y}_{t-1} - \boldsymbol{\mu}) + \mathbf{Z}_t, \quad \mathbf{Z}_t \sim \text{BB}(0, \Sigma)$$

où  $\Phi_1$  est une matrice  $m \times m$  et  $\mathbf{Z}_t$  est  $m$ -dimensionnel.  $\{\mathbf{Y}_t\}$  stationnaire est dit causal si on peut écrire :

$$\mathbf{Y}_t - \boldsymbol{\mu} = \sum_{j=0}^{\infty} \psi_j \mathbf{Z}_{t-j}$$

avec  $\{\psi_j\}$  absolument sommable. C'est la représentation MA( $\infty$ ) d'un VAR(1).

Voyons les conditions sur  $\Phi_1$  qui assurent la stationnarité. Posons  $\Phi(\mathbf{B}) := (I - \Phi_1 \mathbf{B})$ . On veut pouvoir écrire :

$$\mathbf{Y}_t - \boldsymbol{\mu} = \Phi(\mathbf{B})^{-1} \mathbf{Z}_t$$

Or

$$\Phi(\mathbf{B})^{-1} = \frac{1}{\det(\Phi(\mathbf{B}))} \Phi^*(\mathbf{B})$$

où  $\Phi^*(\mathbf{B})$  est la transposée de la comatrice de  $\Phi(\mathbf{B})$  (voir par exemple

<http://www.techno-science.net/?onglet=glossaire&definition=5091>).

Pour que  $\Phi(\mathbf{B})^{-1}$  soit définie, il faut donc pouvoir développer en série  $1/\det(\Phi(\mathbf{B}))$ . Vu ce qu'on a montré pour un AR(p), il faut que les racines de,  $\det(\Phi(z)) = 0$ , soient  $> 1$  en module.

Par exemple, pour  $m = 2$ ,  $\Phi(\mathbf{B}) = \begin{bmatrix} 1 - \phi_{11}\mathbf{B} & -\phi_{12}\mathbf{B} \\ -\phi_{21}\mathbf{B} & 1 - \phi_{22}\mathbf{B} \end{bmatrix}$  et cette co-matrice est :

$$\Phi^*(\mathbf{B}) = \begin{bmatrix} 1 - \phi_{22}\mathbf{B} & \phi_{12}\mathbf{B} \\ \phi_{21}\mathbf{B} & 1 - \phi_{11}\mathbf{B} \end{bmatrix}$$

$$\text{et } \det(\Phi(\mathbf{B})) = (1 - \phi_{11}\mathbf{B})(1 - \phi_{22}\mathbf{B}) - \phi_{12}\phi_{21}\mathbf{B}^2.$$

Exercice. Considérons le VAR(1) :

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1.2 & -0.5 \\ 0.6 & 0.3 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \mathbf{Z}_t \quad \text{où } \mathbf{Z}_t \sim \text{BBN} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1.25 \end{bmatrix} \right). \quad (4.4)$$

Vérifier que ce processus est bien stationnaire.

Ce processus est simulé en annexe (4.6) et on donne ci-dessous le test de blancheur de la série simulée.

```
proc varmax data= st9.simull;
model y1 y2 = un/noint;
run;
```

Covariances croisées des résidus			
Retard	Variable	y1	y2
0	y1	7.66266	4.76739
	y2	4.76739	6.59924
1	y1	6.37094	5.99680
	y2	2.11248	5.09011
2	y1	4.29829	5.82911
	y2	-0.06654	3.21104

Corrélations croisées des résidus			
Retard	Variable	y1	y2
0	y1	1.00000	0.67041
	y2	0.67041	1.00000
1	y1	0.83143	0.84330
	y2	0.29707	0.77132
2	y1	0.56094	0.81972
	y2	-0.00936	0.48658

Représentation schématique des corrélations croisées des résidus													
Variable/ Retard	0	1	2	3	4	5	6	7	8	9	10	11	12
y1	++	++	++	++	.+	.+	-.	-.	..	..	..	..	..
y2	++	++	.+	-+	-.	-.	--	--	..	..	..	..	..

+ is > 2\*std error, - is < -2\*std error, . is between

Test de Portmanteau pour corrélations  
croisées des résidus

Jusqu'  
au  
retard

DF	Khi 2	Pr > Khi 2
1	145.50	<.0001
2	254.48	<.0001
3	335.97	<.0001

...

12	48	485.38	<.0001
----	----	--------	--------

Etant donné qu'on a ajusté la série à un vecteur de 1, les corrélations croisées des résidus (vocabulaire de SAS) sont bien les corrélations croisées de la série. La p-value qui correspond à l'hypothèse nulle de blancheur de la série, est faible à tous les retards. On rejette, comme on s'y attendait, la blancheur de cette série.

**Processus VAR(p)**  $\{\mathbf{Y}_t\} \in R^m$  est un VAR(p) s'il vérifie quelque soit  $t$  :

$$\mathbf{Y}_t - \boldsymbol{\mu} = \sum_{j=1}^p \Phi_j (\mathbf{Y}_{t-j} - \boldsymbol{\mu}) + \mathbf{Z}_t, \quad \mathbf{Z}_t \sim \text{BB}(0, \Sigma)$$

ou  $\Phi(\mathbf{B})(\mathbf{Y}_t - \boldsymbol{\mu}) = \mathbf{Z}_t$  avec :

$$\Phi(\mathbf{B}) = I - \Phi_1 \mathbf{B} - \dots - \Phi_p \mathbf{B}^p$$

ou  $\Phi(\mathbf{B})\mathbf{Y}_t = \boldsymbol{\delta} + \mathbf{Z}_t$ ,  $\boldsymbol{\delta} = (I - \Phi_1 - \dots - \Phi_p)\boldsymbol{\mu}$ .

$\{\mathbf{Y}_t\}$ , VAR(p), est stationnaire si on peut écrire :

$$\mathbf{Y}_t - \boldsymbol{\mu} = \sum_{j=0}^{\infty} \psi_j \mathbf{Z}_{t-j}$$

avec  $\{\psi_j\}$  absolument sommable.

Quelles conditions sur le polynôme de l'autorégression permettent cette écriture ?

Remarque préliminaire : un VAR(p) peut se représenter comme un VAR(1) à  $pm$  composantes. (Voir par exemple Hamilton [13] p. 259).

$$\xi_t = \begin{bmatrix} y_t - \mu \\ y_{t-1} - \mu \\ \vdots \\ y_{t-p+1} - \mu \end{bmatrix}$$

$$F_{mp \times mp} = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \cdots & \Phi_{p-1} & \Phi_p \\ I_m & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_m & 0 \end{bmatrix} \quad (4.5)$$

$$\mathbf{v}_t = \begin{bmatrix} Z_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\xi_t = F\xi_{t-1} + \mathbf{v}_t \quad (4.6)$$

$$y_t - \mu = \begin{bmatrix} 1, 0, \dots, 0 \end{bmatrix}_{1 \times p} \xi_t \quad (4.7)$$

**Propriété.**  $\mathbf{Y}_t$  est stationnaire si les valeurs propres de la matrice  $F$  ci-dessus sont en module  $< 1$  ou de façon équivalente si les racines de  $\det(\Phi(z)) = 0$  sont  $> 1$  en module.

### 4.3.3 Processus VMA (Vector Moving Average)

**VMA(1)**  $\mathbf{Y}_t \in R^m$  est un VMA(1) (processus moyenne mobile vectoriel d'ordre 1) s'il vérifie une équation :

$$\mathbf{Y}_t = \boldsymbol{\mu} + \mathbf{Z}_t - \Theta_1 \mathbf{Z}_{t-1}$$

où  $\mathbf{Z}_t \sim \text{BB}(0, \Sigma)$ . Un VMA(1) est toujours stationnaire.

Fonction de covariance d'un VMA(1).

**Inversibilité** (représentation AR( $\infty$ ) d'un VMA(1)). Le problème est symétrique de la condition de la représentation MA( $\infty$ ) d'un VAR(1). On obtient donc le résultat suivant.

Si  $\det(I - \Theta_1 z) = 0$  a ses racines  $> 1$  en module, alors  $\mathbf{Y}_t$  admet une représentation VAR( $\infty$ ).

Fonction de covariance d'un VMA(1). Elle est nulle à partir du décalage 2

Exercice. Calculer la fonction d'autocovariance de :

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix} - \begin{bmatrix} 0.3 & -.9 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix}$$

où  $\begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix}$  est un BB de matrice des covariances  $\begin{bmatrix} 1 & .3 \\ .3 & 1 \end{bmatrix}$

### 4.3.4 Processus VARMA

**Définition.**  $\mathbf{Y}_t \in R^m$  est un VARMA(p,q) s'il est stationnaire et vérifie une équation :

$$(I - \Phi_1 B - \cdots - \Phi_p B^p) \mathbf{Y}_t = \boldsymbol{\delta} + (I - \Theta_1 B - \cdots - \Theta_q B^q) \mathbf{Z}_t$$

où  $\mathbf{Z}_t \sim \text{BB}(0, \Sigma)$ .



#### 4.4 Représentation $MA(\infty)$ et réponse impulsionnelle

Rappelons d'abord la représentation  $MA(\infty)$  d'un ARMA pour bien situer ce qui est différent dans le cas multidimensionnel. Soit  $Y_t$  ARMA stationnaire obéissant à

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p} = \theta_0 + Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

où  $Z_t \sim BB(0, \sigma_Z^2)$ . Sa représentation par un  $MA(\infty)$ , si elle est permise, est :

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i Z_{t-i}. \quad (4.8)$$

où, on l'a vu :  $\mu = \theta_0 / (1 - \phi_1 - \phi_2 - \cdots - \phi_p)$ . Cette représentation existe si les racines du polynôme  $1 - \phi_1 \mathbf{B} - \phi_2 \mathbf{B}^2 - \cdots - \phi_p \mathbf{B}^p$  sont supérieures à 1 en module.

On voit que  $\psi_k$  peut s'interpréter comme la réponse de  $Y_t$  à une variation de 1 de  $Z_{t-k}$ . Comme  $Z_t$  est une suite de variables non corrélées, on peut imaginer une variation d'un  $Z_t$  sans variation des autres (c'est-à-dire toutes choses égales par ailleurs). Ce ne serait pas le cas si les  $Z_t$  étaient corrélés.

Considérons maintenant un VARMA pour lequel on a la représentation (4.8) :

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{j=0}^{\infty} \begin{bmatrix} \psi_{j,11} & \psi_{j,12} \\ \psi_{j,21} & \psi_{j,22} \end{bmatrix} \begin{bmatrix} Z_{1,t-j} \\ Z_{2,t-j} \end{bmatrix} \quad (4.9)$$

Notons  $\Sigma$  la matrice des covariances de  $\mathbf{Z}_t$ . On veut quantifier la réponse d'une composante de  $\mathbf{Y}_t$ , par exemple  $Y_{1t}$  à une variation de 1 d'une composante de  $\mathbf{Z}_{t-k}$ , par exemple  $Z_{2,t-k}$ . Mais  $Z_{1t}$  et  $Z_{2t}$  peuvent être corrélés. Il faut donc transformer un peu le problème pour avoir un bruit blanc de composantes non corrélées. On procède ainsi.

Soit  $A$  une matrice inversible choisie de façon que dans :  $\mathbf{Y}_t = \sum_{i=0}^{\infty} \Psi_i A^{-1} A \mathbf{Z}_{t-i}$  on ait :  $\text{cov}(A\mathbf{Z}) = A\Sigma A' = I$ .

Définissons :

$$\Psi_i^{\#} = \Psi_i A^{-1}, \text{ et } Z_t^{\#} = A \mathbf{Z}_t. \quad (4.10)$$

Alors

$$\mathbf{Y}_t = \sum_{i=0}^{\infty} \Psi_i^{\#} Z_{t-i}^{\#}$$

et nous avons obtenu une représentation où les composantes du bruit blanc sont non corrélées. Dans cette décomposition, l'élément  $k, l$  de  $\Psi_h^{\#}$  est la réponse de  $Y_{kt}$  à une variation de 1 de  $Z_{l,t-h}^{\#}$ .

Notons que le choix de  $A$  n'est pas unique : la factorisation de Cholesky et la décomposition en valeurs singulières sont deux possibilités.

Remarque. La notion d'identifiabilité pour un VARMA est plus compliquée que pour les modèles ARMA. Nous nous contenterons d'observer le problème sur l'exemple suivant.

Considérons le processus VAR(1) où

$$\Phi_1 = \begin{pmatrix} 0 & 0.5 \\ 0 & 0 \end{pmatrix}$$

Il admet une représentation  $MA(\infty)$  qui se réduit à

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \mathbf{Z}_t + \Phi_1 \mathbf{Z}_{t-1}$$

Ainsi ce VAR(1) peut se représenter comme un VMA(1). D'une façon générale, on ne peut pas identifier un VARMA à partir de sa représentation  $MA(\infty)$ . Nous arrivons maintenant à notre exemple.

4.5 Exemple d’inférence sur un VAR

4.5.1 Données et analyse préliminaire

On considère les séries des logarithmes des ventes annuelles de fourrures de vison (mink) et de rat musqué (muskrat) en baie d’Hudson de 1850 à 1911, fichier `muskrat` . . On commence par dessiner ces séries (fig. 4.3). `logmusk` =, log rat musqué, est la série supérieure. Elle montre une légère tendance mais noyée dans d’amples fluctuations.

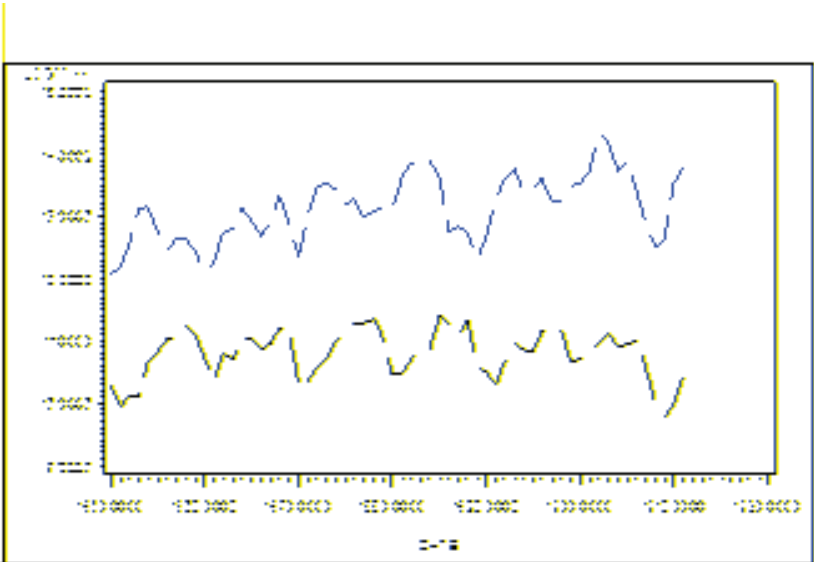


FIGURE 4.3 – Log rat musqué (trait supérieur) et vison.

```
proc arima data=st.muskrat;
i var= logmusk stationarity=(adf= (3,4,5)) ;
run;
i var=logmink stationarity=(adf= (3,4,5)) ;
run;
quit;
```

The ARIMA Procedure

    Name of Variable = LOGMUSK

Mean of Working Series        13.1248 Standard Deviation    0.528868 Number of Observations

62

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	
Pr > F							
Zero Mean	3	0.0243	0.6847	0.13	0.7210		
	4	0.0316	0.6863	0.23	0.7484		
	5	0.0466	0.6897	0.35	0.7831		
Single Mean	3	-24.4908	0.0016	-2.83	0.0597	4.04	0.0892
	4	-15.9986	0.0220	-2.30	0.1760	2.69	0.3945
	5	-18.8175	0.0094	-2.34	0.1624	2.84	0.3562
Trend	3	-549.251	0.0001	-4.62	0.0024	10.70	0.0010
	4	423.8145	0.9999	-3.96	0.0156	7.86	0.0173
	5	58.8701	0.9999	-4.28	0.0066	9.22	0.0010

Name of Variable = LOGMINK							
Mean of Working Series      10.78493   Standard Deviation   0.388156   Number of Observations      62							
Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	3	0.0004	0.6793	0.00	0.6799		
	4	-0.0278	0.6729	-0.18	0.6183		
	5	-0.0204	0.6745	-0.16	0.6235		
Single Mean	3	-374.340	0.0001	-4.81	0.0003	11.56	0.0010
	4	158.7255	0.9999	-4.45	0.0007	9.92	0.0010
	5	78.5820	0.9999	-4.13	0.0018	8.54	0.0010
Trend	3	-224.665	0.0001	-4.69	0.0019	11.87	0.0010
	4	207.3058	0.9999	-4.35	0.0053	10.00	0.0010
	5	92.7428	0.9999	-4.03	0.0130	8.76	0.0010

Les séries initiales sont clairement de moyenne non nulle.

Testons la présence d'une racine unité.

Cas de LOGMUSK. Elle montre une légère tendance, on examine donc le bloc `Trend`. Que ce soit sur `F` ou sur `Tau`, on rejette l'hypothèse nulle de racine unité. Nous concluons que cette série est stationnaire. En procédant comme pour le lac Huron, on pourrait lui ajuster une droite, identifier le résidu, puis estimer simultanément la droite et le modèle de l'erreur et enfin tester la nullité de la pente.

Cas de LOGMINK. Elle ne montre pas de tendance, donc on examine le bloc `Single Mean`. On voit qu'on rejette l'hypothèse de racine unité et qu'il n'est pas pertinent d'envisager pour cette série un modèle avec tendance.

Le corrélogramme empirique (fig. 4.4) décroît rapidement vers 0. On peut par cet examen également, conclure à la stationnarité de la série. Nous examinerons après estimation, la pertinence d'un modèle stationnaire pour ces séries.

## 4.5.2 Identification

On doit identifier un modèle VAR pour cette série, c'est-à-dire trouver un ordre d'autorégression convenable pour modéliser cette série vectorielle.

On fait ce travail à l'aide de la `proc varmax` de SAS. Elle fournit différents critères d'information.

1. Dans une première étape on demande le calcul du critère qu'on a retenu pour un ensemble de modèles et on retiendra le modèle d'AIC minimum.
2. Dans une deuxième étape on estime ce modèle et on examine en détail la qualité de l'ajustement : blancheur du résidu...

La `proc varmax` est ainsi organisée que l'on doit d'abord estimer un modèle particulier pour avoir la valeur d'un critère d'information sur un ensemble de modèles.

- On doit donc donner un ordre de modèle à estimer ; dans la syntaxe ci-dessous on a choisi `p=1`, mais a priori nous ne savons pas si ce modèle nous intéresse.
- C'est l'option `minic` de la commande `model` qui fournit les critères d'information. Par défaut le critère est le AICC, la commande `type=aicc` ne s'impose donc pas. [Les autres critères sont : FPE (Final Prediction Error criterion), HQC (Hanna-Quinn Criterion), (SBC Schwarz Bayesian Criterion)]. Il faut d'abord examiner le résultat de `minic`. On choisit de demander le critère pour tous les modèles jusqu'à un VARMA(5,1). Et on demande l'estimation d'un VAR(1). (L'organisation d'un logiciel est souvent guidée par des considérations plus numériques que logiques.)

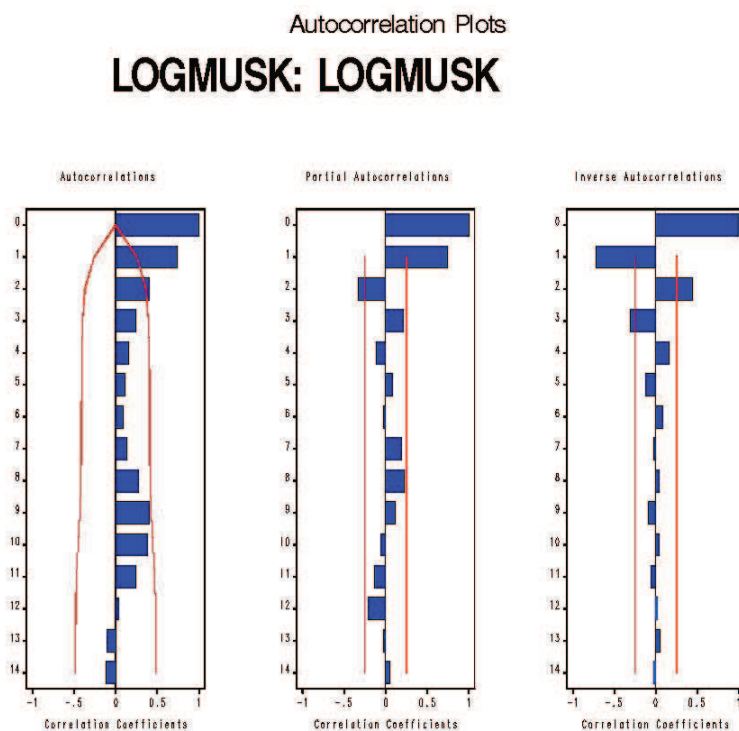


FIGURE 4.4 – ACF et PACF de la série Logmusk.

- Notons enfin que la méthode retenue est ls, moindres carrés, alors que le critère AIC suppose une estimation par maximum de vraisemblance. Les calculs numériques correspondant au maximum de vraisemblance sont très longs et on se contente donc d'un calcul de l'AIC à partir d'estimations par moindres carrés.

---

```
* estimation préliminaire ;
proc varmax data=st6.muskrat;
  id date interval=year;
  model logmink logmusk/ minic=( type=aicc p=5 q=1) p=1 method=ls;
  run;
quit;
```

---

The VARMAX Procedure

Number of Observations                      62    Number of Pairwise Missing                      0

Variable	Type	NoMissN	Mean	StdDev	Min	Max
LOGMINK	DEP	62	10.78493	0.39133	9.79020	11.41670
LOGMUSK	DEP	62	13.12480	0.53319	12.07520	14.31640

Minimum Information Criterion

Lag	MA 0	MA 1
AR 0	-3.503494	-3.760649
AR 1	-5.55847	-5.425568
AR 2	-5.63956	-5.626477
AR 3	-5.634281	-5.559498

```
AR 4      -5.615905      -5.534291
AR 5      -5.478451      -5.468305
```

... et d'autres résultats qui ne nous intéressent pas encore ...

### 4.5.3 Estimation

On note qu'un VAR(2) correspond à la valeur la plus faible du critère. On décide donc de faire cet ajustement et on demande un certain nombre de résultats complémentaires.

Quelques explications sur la syntaxe mise en oeuvre.

L'option

lagmax=5 calcule une estimation des matrices des autocorrélations des erreurs jusqu'au retard 5,  
 impulse= (simple orth) de l'option print=( ) demande la fonction de réponse impulsionnelle ordinaire et la fonction de réponse impulsionnelle correspondant à un BB(0,I),  
 roots donne les valeurs propres de la matrice de la représentation VAR(1).

```
proc varmax data=st.muskrat;
  id date interval=year;
  model logmink logmusk/ p=2 lagmax=5 method=ls
    print= ( impulse= (simple orth ) roots);
  run;
quit;
```

On obtient différentes sorties dont certaines soulèvent des questions.

```
Type of Model                                VAR(2) Estimation Method
Least Squares Estimation
5
    Constant Estimates

Variable          Constant
10
LOGMINK            2.28389
LOGMUSK            5.17706

15          AR Coefficient Estimates

Lag   Variable      LOGMINK      LOGMUSK
20
  1   LOGMINK        0.75943      0.29391
      LOGMUSK       -0.73224      1.16678
  2   LOGMINK       -0.09844     -0.18878
      LOGMUSK        0.42490     -0.30697

25          Model Parameter Estimates

Equation Parameter      Estimate   Std Error   T Ratio Prob>|T|
Variable
30 LOGMI (t)  CONST1      2.28389    1.43871     1.59    0.1181
              AR1_1_1      0.75943    0.13438     5.65    0.0001 LOGMI (t-1)
              AR1_1_2      0.29391    0.13274     2.21    0.0310 LOGMU (t-1)
              AR2_1_1     -0.09844    0.15640    -0.63    0.5317 LOGMI (t-2)
              AR2_1_2     -0.18878    0.12919    -1.46    0.1496 LOGMU (t-2)
35 LOGMU (t)  CONST2      5.17706    1.41342     3.66    0.0006
```

AR1_2_1	-0.73224	0.13202	-5.55	0.0001	LOGMI (t-1)
AR1_2_2	1.16678	0.13041	8.95	0.0001	LOGMU (t-1)
AR2_2_1	0.42490	0.15365	2.77	0.0077	LOGMI (t-2)
AR2_2_2	-0.30697	0.12692	-2.42	0.0189	LOGMU (t-2)

40

On peut observer que certains paramètres ne sont pas significatifs : LOGMI (t-2) semble avoir peu d'influence sur LOGMI (t), de même que LOGMU (t-2) sur LOGMI (t). Une deuxième étape devrait simplifier le modèle en contraignant à 0 certains coefficients, *un par un*.

Covariance Matrix for the Innovation

5 Variable	LOGMINK	LOGMUSK
LOGMINK	0.06505	0.02043
LOGMUSK	0.02043	0.06278

10

Information Criteria

AICC(Corrected AIC)	-5.41904
HQC(Hannan-Quinn Criterion)	-5.3128
15 AIC(Akaike Information Criterion)	-5.44934
SBC(Schwarz Bayesian Criterion)	-5.10028
FPEC(Final Prediction Error Criterion)	0.004302

Residual Cross-Covariance Matrices

20 Lag	Variable	LOGMINK	LOGMUSK
	0 LOGMINK	0.05963	0.01873
	LOGMUSK	0.01873	0.05755
25 1	LOGMINK	0.00019581	0.00158
	LOGMUSK	0.00907	-0.00042401
	2 LOGMINK	0.00280	-0.00686
	LOGMUSK	-0.01377	-0.01000
	3 LOGMINK	-0.00023489	0.00008787
	LOGMUSK	0.00238	0.00788
30 4	LOGMINK	-0.00607	-0.00352
	LOGMUSK	-0.00200	0.00478
	5 LOGMINK	-0.00559	0.00858
	LOGMUSK	0.00069520	-0.00002974

35

Residual Cross-Correlation Matrices

40 Lag	Variable	LOGMINK	LOGMUSK
	0 LOGMINK	1.00000	0.31973
	LOGMUSK	0.31973	1.00000
	1 LOGMINK	0.00328	0.02697
	LOGMUSK	0.15484	-0.00737
45 2	LOGMINK	0.04696	-0.11707
	LOGMUSK	-0.23503	-0.17386
	3 LOGMINK	-0.00394	0.00150
	LOGMUSK	0.04065	0.13700
	4 LOGMINK	-0.10174	-0.06016
	LOGMUSK	-0.03415	0.08302
50 5	LOGMINK	-0.09383	0.14648
	LOGMUSK	0.01187	-0.00051672

55

Schematic Representation of  
Residual Cross Correlations

```

60 Variable/ Lag          0      1      2      3      4      5

    LOGMINK              ++      ..      ..      ..      ..      ..
    LOGMUSK              ++      ..      ..      ..      ..      ..

65      + is > 2*std error, - is <
      -2*std error, . is between
-----
    Portmanteau Test for Residual
      Cross Correlations
70 To      Chi-      Prob>
   Lag      Square    DF      ChiSq

      3         9.20      4      0.0563
      4        10.76      8      0.2155
75      5        14.04     12      0.2984

```

Les lignes (13) et suivantes donnent différents critères d'information. Quand on s'est fixé un critère en vue de choisir un modèle (ordre ou contraintes sur certains paramètres), on retient le modèle donnant la plus faible valeur pour le critère retenu. Mais il est recommandé d'examiner l'ensemble des critères pour tous les modèles essayés.

Si l'ajustement est satisfaisant, les lignes (43) à (52) sont des estimations de matrices nulles. Cette hypothèse est testée par le test du portemanteau ci-dessus, introduit formule (4.3). On peut considérer que le modèle retenu est acceptable, même si le niveau de signification empirique quand on se limite au décalage 3 reste un peu faible (0.0563).

SAS fournit également, voir output ci-dessous, un certain nombre de statistiques sur les séries univariées des résidus de chaque composante.

- Chaque équation de l'ajustement est considérée isolément et le R2 et le F associé sont calculés. La ligne (6) donne le R2 entre logmink et sa prédiction, l'écart-type de l'erreur (déjà fourni plus haut) et la statistique de Fisher pour tester la significativité de la dépendance.
- La statistique de Durbin-Watson est calculée pour chaque résidu.
- Le test de normalité de Jarque-Bera (basé sur la distance du  $\chi^2$  de [aplatissement, asymétrie] observés à [aplatissement, asymétrie] d'une variable normale, est calculé.
- Un test d'hétéroscédasticité est également fourni.

```

Univariate Model Diagnostic Checks

Variable      R-square      StdDev      F Value      Prob>F
5
LOGMINK        0.5737        0.2550        18.50      <.0001
LOGMUSK        0.7740        0.2506        47.10      <.0001

10      Univariate Model Diagnostic Checks

Variable      DW(1)      Normality      Prob>      ARCH1
           ChiSq      ChiSq      F Value      Prob>F

15 LOGMINK        1.99        3.96      0.1382      0.48      0.4905
LOGMUSK        2.00        1.11      0.5739      0.62      0.4344

Univariate Model Diagnostic Checks

20      AR1      AR1-2      AR1-3      AR1-4
Variable  F Value Prob>F F Value Prob>F F Value Prob>F F Value Prob>F

LOGMINK    0.00 0.9803    0.06 0.9402    0.05 0.9864    0.20 0.9379
25 LOGMUSK    0.00 0.9547    0.92 0.4035    1.33 0.2755    0.87 0.4859

```

Examinons maintenant la réponse impulsionnelle. SAS fournit d'abord les matrices  $\Psi$  puis les matrices  $\Psi^\#$ .

Impulse Response Matrices					
Lag	Variable	LOGMINK	LOGMUSK		
1	LOGMINK	0.75943	0.29391		
	LOGMUSK	-0.73224	1.16678		
2	LOGMINK	0.26308	0.37736		
	LOGMUSK	-0.98555	0.83919		
3	LOGMINK	-0.02640	0.28403		
	LOGMUSK	-0.79510	0.46955		
4	LOGMINK	-0.09359	0.15814		
	LOGMUSK	-0.49406	0.24262		
5	LOGMINK	-0.06359	0.07481		
	LOGMUSK	-0.27508	0.14383		
-----					
Orthogonalized Impulse Response Matrices					
Lag	Variable	LOGMINK	LOGMUSK		
0	LOGMINK	0.25504	0		
	LOGMUSK	0.08011	0.23740		
1	LOGMINK	0.21723	0.06978		
	LOGMUSK	-0.09328	0.27700		
2	LOGMINK	0.09733	0.08959		
	LOGMUSK	-0.18413	0.19923		
3	LOGMINK	0.01602	0.06743		
	LOGMUSK	-0.16517	0.11147		
4	LOGMINK	-0.01120	0.03754		
	LOGMUSK	-0.10657	0.05760		
5	LOGMINK	-0.01022	0.01776		
	LOGMUSK	-0.05863	0.03415		
Roots of AR Characteristic Polynomial					
Index	Real	Imaginary	Modulus	ATAN (I/R)	
Degree					
1	0.32458	0.40894	0.5221	0.8999	51.5610
2	0.32458	-0.40894	0.5221	-0.8999	-51.5610
3	0.58752	0	0.5875	0	0
4	0.68953	0	0.6895	0	0

Examen des matrices de réponses impulsionnelles, ligne (5) ci-dessus et suivantes.

Le 1<sup>er</sup> tableau Impulse Response Matrices est l'empilement des  $\Psi_i, i = 1, \dots, 5$ . Le 2<sup>e</sup> tableau est l'empilement des  $\Psi_i^\#, i = 0, \dots, 5$ .

On obtient donc aux lignes (21) et suivantes que  $A^{-1} = \Psi_0^\#$  et on peut vérifier avec les lignes (2) que  $(\Psi_0^\#)^{-1} \Sigma_Z ((\Psi_0^\#)^{-1})' = I$ . Par exemple, ligne (27), on note que quand la composante du bruit blanc normalisé (4.10) relative à logmusk augmente de 1 en  $t - 3$ , logmink augmente de 0.06743 en  $t$ .

Les lignes (39) donnent les valeurs propres de

$$\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 \\ I_2 & 0_{2,2} \end{bmatrix}$$

matrice donnant la représentation VAR(1), voir l'équation 4.6, à 4 composantes du VAR(2) :

$$\begin{bmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \end{bmatrix} = \Phi \begin{bmatrix} \mathbf{Y}_{t-1} \\ \mathbf{Y}_{t-2} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_t \\ 0 \end{bmatrix}.$$



La condition : valeurs propres de cette matrice,  $< 1$  en module, assure la stationnarité.

Pour calculer et stocker les prévisions à l'horizon 5 de ces séries on ajoute une commande `output` :

```
output out=predic lead=5;
```

On obtient :

Forecasts						
Variable	Obs	Time	Forecast	Standard Error	95% Confidence Limits	
LOGMINK	63	2028	10.6982	0.2550	10.1983	11.1980
	64	2029	10.8176	0.3422	10.1469	11.4883
	65	2030	10.8375	0.3669	10.1184	11.5566
	66	2031	10.8293	0.3734	10.0975	11.5611
	67	2032	10.8242	0.3754	10.0884	11.5600
LOGMUSK	63	2028	13.7269	0.2506	13.2358	14.2180
	64	2029	13.5510	0.3850	12.7964	14.3055
	65	2030	13.3989	0.4710	12.4758	14.3219
	66	2031	13.3116	0.5114	12.3093	14.3139
	67	2032	13.2709	0.5255	12.2409	14.3009

Estimation en imposant des restrictions sur certains paramètres – exemple de syntaxe.

```
proc varmax data=st.muskrat;
  id date interval=year;
  model logmink logmusk/ p=2 lagmax=5 method=ls ;
  restrict ar(1,1,2)=0 ar(2,1,2)=0 ;
run;
quit;
```

## 4.6 Simulation

Simulation d'un bruit blanc gaussien bidimensionnel  $N(0, \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 2 \end{pmatrix})$  en SAS avec la `proc iml` et ajout d'une constante à la table des séries simulées.

```
* Simulation d'un BB multidim ;
proc iml;
sig = {1.0 0.5, 0.5 2};
phi = {0 0 , 0 0};
call varmasim(y,phi)
sigma = sig n = 100 seed = 34657; cn = {"y1" "y2"};
create st9.simul_bb from y[colname=cn];
append from y;
quit;

data st9.simul_bb;
set st9.simul_bb;
un = 1;
run;
```

Simulation d'un VAR(1) suivant le modèle (4.4).

```
proc iml;
sig = {1.0 0.5, 0.5 1.25};
```

```

phi = {1.2 -0.5, 0.6 0.3};
call varmasim(y,phi) sigma = sig n = 100 seed = 34657;
cn = {"y1" "y2"};
create st9.simul1 from y[colname=cn]; append from y;
quit;

data st9.simul1;
set st9.simul1;
un = 1;
run;

```

---

Estimation d'un VAR(1) sur la table simul1.

```

proc varmax data=st9.simul1;
    model y1 y2 / p=1 ;
run;

```

---

Calcul de la réponse impulsionnelle pour le modèle théorique (4.4). Les données ne servent à rien, mais il n'y a pas de fonction particulière pour ce travail. On contraint, par `restrict`, les paramètres aux vraies valeurs.

```

proc varmax data=st9.simul1;
    model y1 y2 / noint p=1 print=(impulse(8));
    restrict AR(1,1,1)=1.2, AR(1,1,2)=-.5, AR(1,2,1)=0.6, AR(1,2,2)=0.3;
run;

```

---

Parmi divers résultats, dont des estimations des paramètres, SAS donne la fonction de réponse impulsionnelle jusqu'à l'ordre 8.

Réponse impulsionnelle simple			
Retard	Variable	y1	y2
1	y1	1.20000	-0.50000
	y2	0.60000	0.30000
2	y1	1.14000	-0.75000
	y2	0.90000	-0.21000
...			
8	y1	-0.19368	0.00439
	y2	-0.00527	-0.18577

---

## Chapitre 5

# Modèle à fonction de transfert

### Présentation

Ce chapitre présente des modèles VARMA particuliers, où les séries n'ont pas des rôles symétriques : il y a un input et un output. Pour apprendre ce chapitre, le plus simple est de commencer par la section 2, méthodologie de Box-Jenkins puis faire l'exemple et enfin aborder la première section qui est un peu plus générale.

### 5.1 Représentation MA et fonction de transfert

Parfois, la dépendance entre séries temporelles est modélisée par une fonction de transfert

$$Y_{it} = \sum_{j \neq i} \frac{\omega_{ij}(B)}{\delta_{ij}(B)} Y_{jt} + \frac{\theta_i(B)}{\phi_i(B)} Z_{it}, \quad i = 1, \dots, m$$

Si on multiplie à gauche et à droite par les  $\delta_{ij}(B)$  et les  $\phi_i(B)$  on obtient un modèle VARMA(p,q) dont la structure MA est diagonale.

$$\mathbf{Y}_t - \sum_{j=1}^p \Phi_j \mathbf{Y}_{t-j} = \mathbf{Z}_t - \sum_{j=1}^q \Theta_j \mathbf{Z}_{t-j}$$

où la structure MA est diagonale.

Envisageons le cas où  $\mathbf{Y}_t$  est partitionné en deux sous-vecteurs  $\mathbf{Y}_{1t}$  et  $\mathbf{Y}_{2t}$  de dimensions  $m_1$  et  $m_2$ ,  $m_1 + m_2 = m$  et partitionnons de même les opérateurs AR et MA :

$$\Phi(B) = \begin{bmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{bmatrix} \quad \Theta(B) = \begin{bmatrix} \Theta_{11}(B) & \Theta_{12}(B) \\ \Theta_{21}(B) & \Theta_{22}(B) \end{bmatrix}.$$

Considérons la situation où  $\Phi_{12}(B)$  et  $\Theta_{12}(B)$  sont nulles et supposons pour simplifier que  $\Theta_{21}(B) = 0$ . Alors les équations de ce modèle peuvent se répartir en deux groupes distincts :

$$\begin{cases} \Phi_{11}(B) \mathbf{Y}_{1t} = & \Theta_{11}(B) \mathbf{Z}_{1t} \\ \Phi_{22}(B) \mathbf{Y}_{2t} = & -\Phi_{21}(B) \mathbf{Y}_{1t} + \Theta_{22}(B) \mathbf{Z}_{2t}. \end{cases} \quad (*)$$

Dans la terminologie de la causalité on dit que dans la structure (\*), les variables  $\mathbf{Y}_{1t}$  causent (ou précèdent)  $\mathbf{Y}_{2t}$  au sens de Granger et que les variables  $\mathbf{Y}_{2t}$  ne causent pas  $\mathbf{Y}_{1t}$ , voir la section (5.4). Dans cette structure, les valeurs futures de  $\mathbf{Y}_{1t}$  ne dépendent que de son propre passé alors que les valeurs futures de  $\mathbf{Y}_{2t}$  dépendent des passés de  $\mathbf{Y}_{1t}$  et de  $\mathbf{Y}_{2t}$ .

Remarque. Si on avait  $\Theta_{21}(B) \neq 0$ , on pourrait remplacer  $\Theta_{21}(B) \mathbf{Z}_{1t}$  par  $\Theta_{21}(B) \Theta_{11}(B)^{-1} \Phi_{11}(B) \mathbf{Y}_{1t}$ .

Exemple. Considérons un VMA(1) à valeurs dans  $R^2$ , que nous écrivons sous forme de fonction de transfert.

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1 - \theta_{11}B & -\theta_{12}B \\ -\theta_{21}B & 1 - \theta_{22}B \end{bmatrix} \begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix}$$

$$d(B) = \det(I - \Theta B) = (1 - \theta_{11}B)(1 - \theta_{22}B) - \theta_{12}\theta_{21}B^2$$

Posons

$$\Theta^*(B) = \text{adj}(I - \Theta B) = \begin{bmatrix} 1 - \theta_{22}B & \theta_{12}B \\ \theta_{21}B & 1 - \theta_{11}B \end{bmatrix}$$

On peut écrire :

$$\Theta^*(B)Y_t = d(B)Z_t$$

c'est-à-dire

$$\begin{cases} (1 - \theta_{22}B)Y_{1t} &= -\theta_{12}BY_{2t} + d(B)Z_{1t} \\ (1 - \theta_{11}B)Y_{2t} &= -\theta_{21}BY_{1t} + d(B)Z_{2t} \end{cases}$$

ou :

$$\begin{cases} Y_{1t} &= \frac{-\theta_{12}B}{1 - \theta_{22}B}Y_{2t} + \frac{d(B)}{1 - \theta_{22}B}Z_{1t} \\ Y_{2t} &= \frac{-\theta_{21}B}{1 - \theta_{11}B}Y_{1t} + \frac{d(B)}{1 - \theta_{11}B}Z_{2t} \end{cases}$$

Supposons maintenant que  $\theta_{12} = 0$ . Ces équations se réduisent à une fonction de transfert unidirectionnelle :

$$\begin{cases} Y_{1t} &= (1 - \theta_{11}B)Z_{1t} \\ Y_{2t} &= \frac{-\theta_{21}B}{1 - \theta_{11}B}Y_{1t} + (1 - \theta_{22}B)Z_{2t} \end{cases}$$

Il est possible que  $Z_{1t}$  et  $Z_{2t}$  soient corrélés. Pour traiter ce cas supposons que  $Z_t$  est un BB gaussien. On peut projeter  $Z_{2t}$  sur  $Z_{1t}$  :

$$Z_{2t} = \beta Z_{1t} + a_{2t}, \quad \text{avec : } \beta = \sigma_{12}/\sigma_{11} \text{ et } a_{2t} \perp Z_{1t}$$

et donc :

$$Y_{2t} = \frac{1}{1 - \theta_{11}B}(\beta - (\theta_{21} - \beta\theta_{22})B)Y_{1t} + (1 - \theta_{22}B)a_{2t} \quad \text{où } a_{2t} \perp Y_{1t}$$

c'est un modèle classique de fonction de transfert.

On arrive à un modèle semblable quand on part d'un VAR(1) dont  $\Phi_{12} = 0$ .

## 5.2 Identification d'une fonction de transfert - méthodologie de Box-Jenkins

### 5.2.1 Données et modèle

On dispose de l'observation de deux séries conjointement stationnaires  $\{X_t, Y_t\}$ .  $X_t$  est la série en entrée (exogène) et  $Y_t$  la série en sortie (endogène). Ces séries sont reliées par un filtre linéaire :

$$Y_t = \nu(B)X_t + a_t$$

où  $\nu(B) = \sum_{j=0}^{\infty} \nu_j B^j$ .

Le bruit  $\{a_t\}$  est stationnaire et supposé non corrélé avec  $\{X_t\}$  :  $\text{cov}(a_t, X_s) = 0, \forall t, s$ . On appelle  $\{\nu_j\}_{j=0}^{\infty}$  la fonction de réponse impulsionnelle. On se restreint au cas où on peut représenter ou approcher  $\nu(B)$  par une fraction rationnelle :

$$\nu(B) = \frac{\omega_s(B)B^b}{\delta_r(B)}$$

où

$b$  est un paramètre de retard,

$$\omega_s(\mathbf{B}) = \omega_0 - \omega_1 \mathbf{B} - \cdots - \omega_s \mathbf{B}^s,$$

$$\delta_r(\mathbf{B}) = \delta_0 - \delta_1 \mathbf{B} - \cdots - \delta_r \mathbf{B}^r.$$

Les fonctions de réponse impulsionnelles typiques s'interprètent de la même façon que pour les modèles d'intervention.

### 5.2.2 Relation entre fonction de corrélation croisée et fonction de transfert

L'objectif est d'identifier la fonction de transfert, c'est-à-dire de découvrir des valeurs possibles pour les degrés  $b$ ,  $r$ ,  $s$ .

Posons :  $\gamma_{XY}(k) = \mathbf{E}(X_t - \mu_X)(Y_{t+k} - \mu_Y)$  et  $\rho_{XY}(k) = \frac{\gamma_{XY}(k)}{\sigma_X \sigma_Y}$ . Nous supposons pour simplifier :  $\mu_X = \mu_Y = 0$ . Multipliant l'expression  $Y_{t+k} = \nu_0 X_{t+k} + \nu_1 X_{t+k-1} + \nu_2 X_{t+k-2} + \cdots + a_{t+k}$  des deux côtés par  $X_t$  et prenant l'espérance mathématique on obtient :

$$\gamma_{XY}(k) = \nu_0 \gamma_X(k) + \nu_1 \gamma_X(k-1) + \nu_2 \gamma_X(k-2) + \cdots$$

ou

$$\rho_{XY}(k) = \frac{\sigma_X}{\sigma_Y} [\nu_0 \rho_X(k) + \nu_1 \rho_X(k-1) + \nu_2 \rho_X(k-2) + \cdots]$$

Ainsi, à cause de l'autocorrélation de la série input, on ne peut attacher un  $\nu$  particulier à la valeur  $\rho_{XY}(k)$ . Mais si  $\{X_t\}$  était un bruit blanc on aurait :

$$\nu_k = \frac{\sigma_Y}{\sigma_X} \rho_{XY}(k).$$

La fonction de réponse impulsionnelle est dans ce cas directement proportionnelle à la fonction de corrélation croisée.

### 5.2.3 Préblanchiment de l'input

Supposons que l'input suive un ARMA(p,q) :

$$\Phi_X(\mathbf{B})X_t = \Theta_X(\mathbf{B})Z_t.$$

On appelle série input préblanchie, la série

$$\alpha_t = \frac{\Phi_X(\mathbf{B})}{\Theta_X(\mathbf{B})} X_t$$

Appliquons la même transformation à  $Y_t$  :

$$\beta_t = \frac{\Phi_X(\mathbf{B})}{\Theta_X(\mathbf{B})} Y_t$$

et définissons également :

$$\epsilon_t = \frac{\Phi_X(\mathbf{B})}{\Theta_X(\mathbf{B})} a_t.$$

On a finalement :

$$\beta_t = \nu(\mathbf{B})\alpha_t + \epsilon_t$$

et  $\nu_k = \frac{\sigma_\alpha}{\sigma_\beta} \rho_{\alpha\beta}(k)$ .

On a maintenant les éléments pour identifier un modèle à fonction de transfert.

### 5.2.4 Les étapes

1. Vérifier la stationnarité des deux séries et calculer la fonction de corrélation croisée (CCF) empirique
2. Identification.

(a) Préblanchir l'input :

- \* identifier et estimer  $X_t$ ,
- \* prédire le bruit blanc qui engendre  $X_t : \alpha_t = \Phi_X(B)/\Theta_X(B)X_t$ .

(b) Appliquer le filtre préblanchissant à l'output : calculer

$$\beta_t = \frac{\Phi_X(B)}{\Theta_X(B)} Y_t$$

(c) Calculer la CCF empirique  $\hat{\rho}_{\alpha\beta}(k)$  pour estimer  $\nu_k$  :

$$\hat{\nu}_k = \frac{\hat{\sigma}_\beta}{\hat{\sigma}_\alpha} \hat{\rho}_{\alpha\beta}(k).$$

On sait d'autre part que si  $\nu_k = 0$  alors l'écart-type de  $(\hat{\sigma}_\alpha/\hat{\sigma}_\beta)\hat{\nu}_k$  est  $\simeq 1/\sqrt{n-k}$  où  $n$  est la longueur de la série (Wei par. 13.3, Reinsel par. 4.1.2, Brockwell et Davis p.411).

- (d) Identifier  $b$ ,  $\omega_s(B)$  et  $\delta_r(B)$  d'après des formes typiques.
- (e) Faire une estimation préliminaire de  $\nu(B)$  en supposant que  $a_t$  est un BB.
- (f) Identifier le modèle du bruit  $a_t$  à partir de

$$\hat{a}_t = Y_t - \hat{\nu}(B)X_t = Y_t - \frac{\hat{\omega}_s(B)}{\hat{\delta}_r(B)} B^b X_t.$$

(g) Estimer le modèle complet.

3. Contrôle de qualité.

(a) Contrôle de la corrélation croisée entre  $a_t$  et  $X_t$

- \* Contrôle à un décalage particulier. On doit avoir avec une probabilité d'environ 95% :

$$\hat{\rho}_{x\hat{a}}(k) \in \pm 2/\sqrt{n-k}$$

- \* Test du portemanteau. S'il n'y a pas de corrélation croisée,

$$Q_0 = m(m+2) \sum_{j=0}^K \frac{1}{m-j} \hat{\rho}_{x\hat{a}}^2(j) \simeq \chi^2(K+1-M)$$

où  $m$  est le nombre de résidus calculés,  $M$  le nombre de paramètres  $\delta_i$  et  $\omega_j$  estimés.

(b) Contrôle d'autocorrélation de  $a_t$ .

## 5.3 Exemple

### 5.3.1 Données

On dispose de deux séries : `index` et `sales`, série de ventes. On a observé par le passé, que les variations de `index` précèdent celles de `sales`. On veut donc modéliser la dépendance de `sales` par rapport à `index`.

Le fichier `indicipil` contient les séries : `index`, `sales`, les séries différenciées à l'ordre 1 `dindex1` et `dsales1` et `num` le numéro de l'observation.

Ces données sont tirées du livre de Box-Jenkins et se trouvent aussi comme exemple dans de nombreux logiciels et ouvrages. Les graphes (5.1 et 5.2) de ces séries montrent qu'elles ne sont pas stationnaires.

On les différencie donc et la suite du travail va consister à modéliser par une fonction de transfert, la relation entre `dindex1` et `dsales1`.

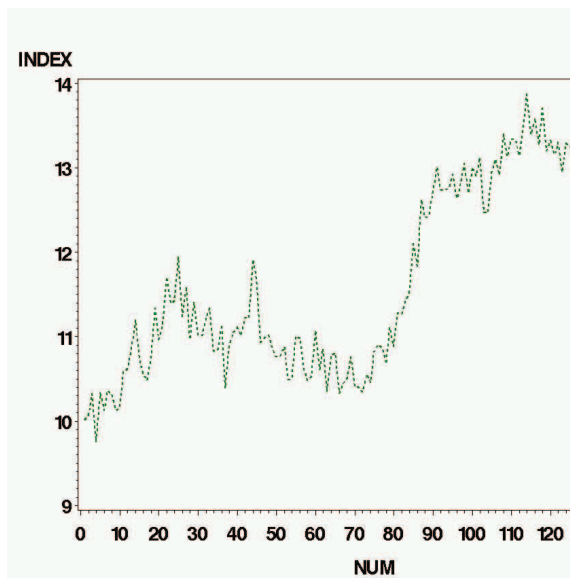


FIGURE 5.1 – Série indicatrice

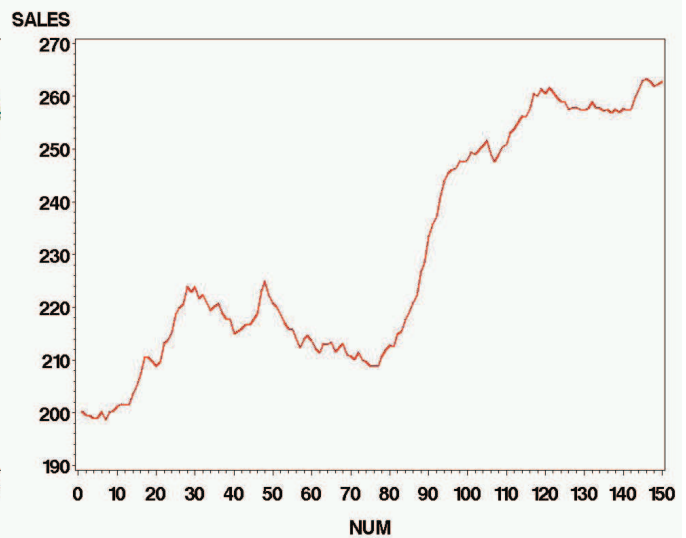


FIGURE 5.2 – Série des ventes

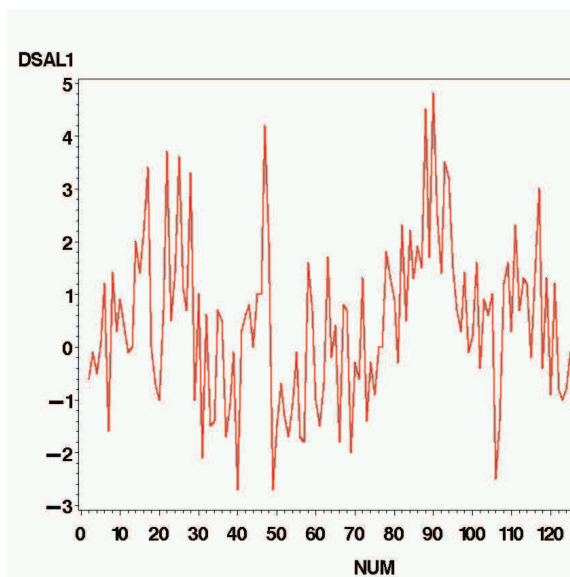


FIGURE 5.3 – Différence première des ventes

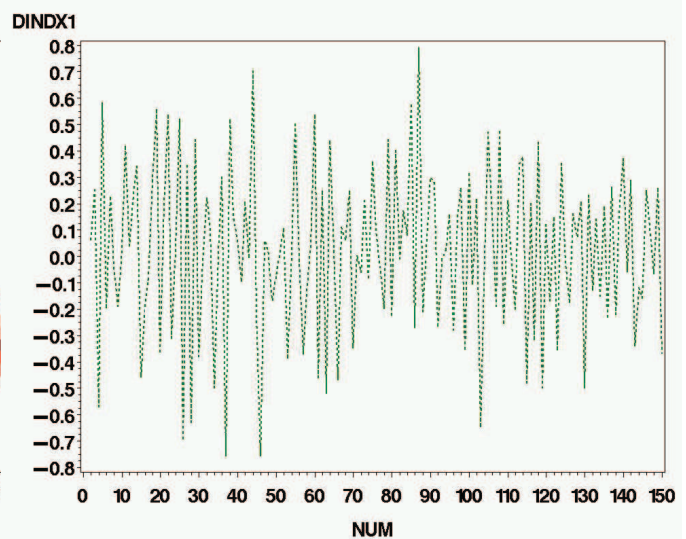


FIGURE 5.4 – Différence première de l'indicateur

### 5.3.2 Modélisation de l'input

On met en œuvre la proc arima.

---

```
proc arima data = aa;
identify var= index(1); run;
```

---

```

Name of variable = INDEX.
Period(s) of Differencing = 1.
Mean of working series = 0.022752
Standard deviation      = 0.315162
Number of observations =      149
NOTE: The first observation was eliminated by differencing.

Autocorrelations

Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1 Std
0 0.099327 1.00000 | | | | | | | | | | | | | | | | | | | | | 0
1 -0.044402 -0.44703 | | | | | | | | | | | | | | | | | | | | | 0.081923
2 0.0084831 0.08541 | | | | | | | | | | | | | | | | | | | | | 0.096921
3 -0.0069778 -0.07025 | | | | | | | | | | | | | | | | | | | | | 0.097425
4 0.012869 0.12956 | | | | | | | | | | | | | | | | | | | | | 0.097764
5 -0.0090300 -0.09091 | | | | | | | | | | | | | | | | | | | | | 0.098910
6 0.0077101 0.07762 | | | | | | | | | | | | | | | | | | | | | 0.099469
7 -0.0077688 -0.07821 | | | | | | | | | | | | | | | | | | | | | 0.099875
8 0.011913 0.11994 | | | | | | | | | | | | | | | | | | | | | 0.100285
9 -0.0051837 -0.05219 | | | | | | | | | | | | | | | | | | | | | 0.101243
10 -0.012395 -0.12479 | | | | | | | | | | | | | | | | | | | | | 0.101424
11 0.018569 0.18695 | | | | | | | | | | | | | | | | | | | | | 0.102449
```

---

Ce graphique suggère un MA(1). Les statistiques descriptives qui précèdent suggèrent une moyenne nulle.

---

```
estimate q=1 noconstant; run;
```

---

```

ARIMA Procedure
Conditional Least Squares Estimation
Approx.
Parameter Estimate Std Error T Ratio Lag
MA1,1 0.44920 0.07347 6.11 1

Variance Estimate = 0.08038194
Std Error Estimate = 0.28351709
AIC = 48.2164141*
SBC = 51.2203604*
Number of Residuals= 149
* Does not include log determinant.

Autocorrelation Check of Residuals

To Chi Autocorrelations
Lag Square DF Prob
6 5.93 5 0.313 -0.055 0.082 0.018 0.146 -0.001 0.083
12 13.82 11 0.243 -0.002 0.116 -0.029 -0.071 0.171 -0.009
18 15.07 17 0.590 0.029 -0.031 -0.023 0.010 0.038 -0.059
24 19.45 23 0.675 0.054 0.141 -0.017 0.028 0.015 -0.033

Model for variable INDEX

No mean term in this model.
Period(s) of Differencing = 1.

Moving Average Factors
Factor 1: 1 - 0.4492 B**(1)
```

---



Le modèle est satisfaisant.

### 5.3.3 Identification du modèle de la fonction de transfert

```
identify var= sales(1) crosscor= index(1);
run;
```

Noter l'usage de la commande `crosscor`. On obtient d'abord la fonction d'autocorrélation de la série  $y$  (= sales différenciée 1 fois) :

```

              ARIMA Procedure
Name of variable = SALES.
Period(s) of Differencing = 1.
Mean of working series = 0.420134
Standard deviation      = 1.439145
Number of observations =    149
NOTE: The first observation was eliminated by differencing.

              Autocorrelations
Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1 Std
0   2.071138    1.00000 |          |*****|          0
1   0.645779    0.31180 |          . |*****|          0.081923
2   0.576179    0.27819 |          . |*****|          0.089534
3   0.468885    0.22639 |          . |*****|          0.095159
4   0.522143    0.25210 |          . |*****|          0.098707
5   0.309832    0.14960 |          . |***. |          0.102938
6   0.276728    0.13361 |          . |***. |          0.104387
7   0.130231    0.06288 |          . |* . |          0.105528
8   0.274203    0.13239 |          . |***. |          0.105780
9  -0.039048   -0.01885 |          . | . |          0.106886
10 -0.0077359  -0.00374 |          . | . |          0.106908

```

On obtient ensuite ci-dessous, la fonction de corrélation croisée des *versions préblanchies* de  $y$ =Sales (1) et de  $x$ =Index (1), préblanchiment défini d'après le modèle qu'on a estimé auparavant pour l'input. Noter la valeur très élevée en 3, c'est-à-dire  $\text{COV}(x_t, y_{t+3})$  est très élevée. On observe ensuite une décroissance géométrique de raison 0.7 environ de cette corrélation croisée. Ceci suggère que la fonction de transfert est en  $1/(1 - \omega B)$ . On décide d'estimer une telle fonction.

```

              Correlation of SALES and INDEX

Variable INDEX has been differenced.
Period(s) of Differencing = 1.
Both series have been prewhitened.
Variance of transformed series = 3.794675 and 0.078036
Number of observations =    149
NOTE: The first observation was eliminated by differencing.

              Crosscorrelations
Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
-6  -0.034684  -0.06374 |          . *| . |
-5   0.013016   0.02392 |          . | . |
-4   0.0012583  0.00231 |          . | . |
-3   0.022045   0.04051 |          . |* . |
-2   0.0054125  0.00995 |          . | . |
-1   0.051478   0.09460 |          . |** . |
0    0.034232   0.06291 |          . |* . |
1    0.043060   0.07913 |          . |** . |
2    0.010062   0.01849 |          . | . |
3    0.367442   0.67523 |          . |*****|
4    0.246112   0.45227 |          . |*****|
5    0.185447   0.34079 |          . |*****|
6    0.140160   0.25757 |          . |*****|

```

```

7  0.145861  0.26804 | . |*****|
8  0.107803  0.19811 | . |****|
9  0.094235  0.17317 | . |***|
10 0.053115  0.09761 | . |**.|
11 0.078822  0.14485 | . |***|
12 0.038038  0.06990 | . |* .|
13 -0.0078178 -0.01437 | . | .|

```

### 5.3.4 Estimation du modèle

On estime un modèle avec un retard de 3 sur l'input et une constante additive. D'abord on suppose que le bruit du modèle est blanc :

```
estimate input=(3 $(0)/(1) index) plot ; run;
```

On dessine la fonction d'autocorrélation des résidus de ce modèle (option plot ci-dessus). On note que ce résidu pourrait être modélisé comme un MA(1).

Autocorrelation Plot of Residuals																								
Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	Std
0	0.070381	1.00000													*****									0
1	-0.022758	-0.32335									*****		.											0.083045
2	0.0048545	0.06897								.		*	.											0.091317
3	0.00016559	0.00235								.		.												0.091675
4	-0.0031744	-0.04510								.	*		.											0.091676
5	0.010729	0.15244								.		***.												0.091829
6	0.0013107	0.01862								.		.												0.093557
7	-0.0059165	-0.08406								.	**		.											0.093583
8	0.0052426	0.07449								.		*	.											0.094102
9	-0.0080949	-0.11502								.	**		.											0.094508
10	0.0087124	0.12379								.		**.												0.095469
11	-0.0010988	-0.01561								.		.												0.096569

On ajuste donc la même fonction de transfert que plus haut mais avec une erreur MA(1).

```
estimate q=1 input=(3 $(0)/(1) index); run; quit;
```

Both variables have been prewhitened by the following filter:

```

Prewhitening Filter
Moving Average Factors
Factor 1: 1 - 0.4492 B**(1)

```

#### Conditional Least Squares Estimation

Parameter	Estimate	Std Error	T Ratio	Lag	Variable	Shift
MU	0.02373	0.01633	1.45	0	SALES	0
MA1,1	0.31200	0.08005	3.90	1	SALES	0
NUM1	4.71265	0.07064	66.71	0	INDEX	3
DEN1,1	0.72191	0.0058449	123.51	1	INDEX	3

Constant Estimate = 0.02373434

```

Variance Estimate = 0.06364271
Std Error Estimate = 0.25227506
AIC = 16.0377405*
SBC = 27.9446755*
Number of Residuals= 145
* Does not include log determinant.

```

#### Correlations of the Estimates

Variable	Parameter	SALES MU	SALES MA1,1	INDEX NUM1	INDEX DEN1,1
SALES	MU	1.000	-0.007	-0.060	-0.345
SALES	MA1,1	-0.007	1.000	0.035	-0.021
INDEX	NUM1	-0.060	0.035	1.000	-0.572
INDEX	DEN1,1	-0.345	-0.021	-0.572	1.000

On note que maintenant, le bruit résiduel est assimilable à un bruit blanc.

Autocorrelation Check of Residuals									
To Lag	Chi Square	DF	Prob	Autocorrelations					
6	5.90	5	0.316	-0.029	0.077	0.026	0.011	0.170	0.050
12	9.41	11	0.584	-0.059	0.038	-0.073	0.102	0.003	-0.040
18	14.10	17	0.660	-0.083	0.005	-0.047	0.039	-0.125	0.047
24	18.74	23	0.716	0.050	-0.033	0.004	-0.098	-0.091	0.072

Il faut vérifier que la série des résidus de l'ajustement est non corrélée avec la série input. Il reste encore un peu de corrélation aux ordres 2 et 3.

Crosscorrelation Check of Residuals with Input INDEX									
To Lag	Chi Square	DF	Prob	Crosscorrelations					
5	10.72	5	0.057	-0.001	0.115	-0.120	0.064	-0.047	0.204
11	12.65	11	0.317	0.015	0.005	0.043	0.062	0.013	0.086
17	13.95	17	0.671	0.009	-0.029	-0.027	-0.043	0.069	0.028
23	20.41	23	0.617	-0.148	0.087	0.013	0.049	-0.062	0.098

### Modèle estimé

Model for variable SALES  
 Estimated Intercept = 0.02373434  
 Period(s) of Differencing = 1.  
  
 Moving Average Factors  
 Factor 1: 1 - 0.312 B\*\*(1)  
 Input Number 1 is INDEX with a shift of 3.  
 Period(s) of Differencing = 1.  
 Overall Regression Factor = 4.712648  
  
 The Denominator Factors are  
 Factor 1: 1 - 0.72191 B\*\*(1)

## 5.4 Test de causalité au sens de Granger

Supposons  $\mathbf{Y}_t$  à  $m$  composantes, VAR(p) stationnaire, partitionné en  $\mathbf{Y}_t = \begin{bmatrix} \mathbf{Y}_{1t} \\ \mathbf{Y}_{2t} \end{bmatrix}$ , de tailles  $m_1$  et  $m_2$ ,  $m_1 + m_2 = m$ . Les coefficients d'autorégression et  $\mathbf{Z}_t$  sont partitionnés de la même façon.

$$\Phi(\mathbf{B})\mathbf{Y}_t = \mathbf{Z}_t$$

$$\begin{bmatrix} \Phi_{11}(\mathbf{B}) & \Phi_{12}(\mathbf{B}) \\ \Phi_{21}(\mathbf{B}) & \Phi_{22}(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{1t} \\ \mathbf{Y}_{2t} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{1,t-1} \\ \mathbf{Z}_{2,t-1} \end{bmatrix}.$$

Si  $\Phi_{12}(\mathbf{B}) = 0$ , on dit que les variables  $\mathbf{Y}_1$  causent les variables  $\mathbf{Y}_2$  et que les variables  $\mathbf{Y}_2$  ne causent pas les variables  $\mathbf{Y}_1$ . Autrement dit,  $\mathbf{Y}_{1t}$  ne dépend pas du passé de  $\mathbf{Y}_2$  :  $\mathbf{Y}_{2,t-1}, \mathbf{Y}_{2,t-2}, \dots$  alors que  $\mathbf{Y}_{2t}$  dépend du passé de  $\mathbf{Y}_1$  :  $\mathbf{Y}_{1,t-1}, \mathbf{Y}_{1,t-2}, \dots$ . Observons que  $\mathbf{Y}_{1t}$  dépend de  $\mathbf{Y}_{2t}$  par  $\mathbf{Z}_t$ .

**Test de causalité de Granger.** Tester une non causalité revient à tester que des coefficients d'une régression linéaire sont nuls. L'hypothèse nulle est de la forme  $H_0 : C\beta = c$ . Dans le modèle VAR,  $\beta$  est la

vectorisation des  $\Phi_1, \dots, \Phi_p, C, pm_1m_2 \times pm^2$ ;  $c, pm_1m_2 \times 1$ , matrice nulle. La statistique du test de causalité de Granger et la statistique de Wald, distance du  $\chi^2$  entre  $C\hat{\beta}$  et  $0_{pm_1m_2 \times 1}$ .

Dans SAS, le test se fait par l'instruction `causal`. On spécifie les deux groupes de variables par l'option `GROUP1=(VARIABLES) GROUP2=(VARIABLES)`. Si l'hypothèse nulle n'est pas rejetée on peut considérer que les variables du GROUP1 sont influencées par elles-mêmes mais pas par celles du GROUP2.

Exemple de syntaxe.

---

```

proc varmax data=tourism;
  id year interval=year;
  model lpvsp lppdi lrcsp / p=1;
  /* Test de causalité */
5   causal group1=(lrcsp) group2=(lpvsp lppdi);
  run;

```

---