REVUE DE STATISTIQUE APPLIQUÉE

ANTOINE DE FALGUEROLLES

Autorégression avec écrêtement

Revue de statistique appliquée, tome 27, n° 1 (1979), p. 29-44

<http://www.numdam.org/item?id=RSA_1979__27_1_29_0>

AUTORÉGRESSION AVEC ÉCRÊTEMENT

Antoine de FALGUEROLLES

Laboratoire de Statistique et Probabilités,
Université Paul Sabatier
E.R.A. C.N.R.S. 591

I – INTRODUCTION

Lorsqu'on étudie une série chronologique scalaire $\{x_1, \ldots, x_T\}$ – brute ou éventuellement transformée – il est courant, notamment depuis les travaux de G. BOX et G. JENKINS (cf. [2]), de considérer que l'on dispose d'une portion de réalisation d'un processus $\{X_t ; t \in Z\}$ stationnaire du second ordre (i.e. tel que $E[X_t] = \mu$ et $E[(X_{t-i} - \mu)(X_{t-j} - \mu)] = \sigma(|i-j|)$). Ce processus est en outre supposé autorégressif-moyenne mobile inversible (en abrégé A.R.M.A. inversible), c'est-à-dire solution stationnaire d'une équation stochastique linéaire à coefficients constants :

$$\sum_{i=0}^{p} \phi_i (X_{t-i} - \mu) = \sum_{j=0}^{q} \chi_j \epsilon_{t-j} \quad (\text{avec } \phi_0 = \chi_0 = 1 \text{ et } \phi_p \chi_q \neq 0) .$$

Dans cette équation, $\{\epsilon_t ; t \in Z\}$ désigne un bruit blanc (i.e. un processus tel que $E[\epsilon_t] = 0$, $E[\epsilon_t^2] = \sigma_\epsilon > 0$ et $E[\epsilon_{t'} \epsilon_{t''}] = 0$ si $t' \neq t''$) et les coefficients ϕ_i et χ_j sont tels que les polynômes $\sum_{i=0}^{p} \phi_i z^{p-i}$ et $\sum_{j=0}^{q} \chi_j z^{q-j}$ de la variable complexe z n'admettent pas de zéro pour $|z| \geqslant 1$.

Si l'on considère l'espace de Hilbert engendré par le processus $\{X_t - \mu ; t \in Z\}$, muni du produit scalaire de la covariance, et son sous-espace engendré par $\{X_{t'} - \mu ; t' < t\}$, $X_t - \mu$ s'écrit comme la somme de sa projection orthogonale sur ce sous-espace et de ϵ_t (innovation au temps t du processus $\{X_t - \mu ; t \in Z\}$). Pour ces modèles de processus cette projection s'écrit sous la forme d'une série :

$$\sum_{j=1}^{+\infty} \psi_j (X_{t-j} - \mu) .$$

Si $q = 0$, il vient $\psi_j = \phi_j$, $j \in \{1, \ldots, p\}$, et $\psi_j = 0$, $j > p$; le processus $\{X_t - \mu ; t \in Z\}$ est alors dit autorégressif d'ordre p (en abrégé A.R. (p)).

G. BOX et G. JENKINS ont montré comment, au vu du corrélogramme et de l'ensemble des coefficients d'autocorrélation partielle d'un processus A.R.M.A., on peut tenter d'en reconnaître les ordres p et q. On rappelle que le corrélogramme

Mots-clés: **Processus, Autorégression, Auto-corrélations partielles, Ridge-regression.**

est l'ensemble $\{\rho(k) ; k \in N^*\}$ avec $\rho(k) = \sigma(k)/\sigma(0)$. L'ensemble des coefficients d'autocorrélation partielle, $\{\alpha_k^k ; k \in N^*\}$ est pratiquement déterminé par autorégression pas à pas c'est-à-dire en projetant $X_t - \mu$ sur le sous-espace engendré par $X_{t-1} - \mu, \ldots, X_{t-k} - \mu$, k parcourant N*. Soit $\widehat{(X_t - \mu)}_{(k)}$ cette projection ; on montre (cf. [2] pp. 64-65 et pp. 82-84) que l'on a :

$$\widehat{(X_t - \mu)}_{(k)} = \alpha_k^1 (X_{t-1} - \mu) + \cdots + \alpha_k^k (X_{t-k} - \mu) .$$

Dans cet article, partant d'un problème concret, on adapte cette démarche à des processus $\{X_t ; t \in \{1, \ldots, T\}\}$ non stationnaires, autorégressifs à coefficients dépendants du temps. L'estimation de leurs paramètres conduit assez naturellement à des procédures de type "ridge-regression".

II – LE PROBLEME GUIDE

On effectue sur chaque "individu" i d'une population I une séquence ordonnée x^i de T mesures : $x^i = (x_1^i, \ldots, x_t^i, \ldots, x_T^i)$.

On note :

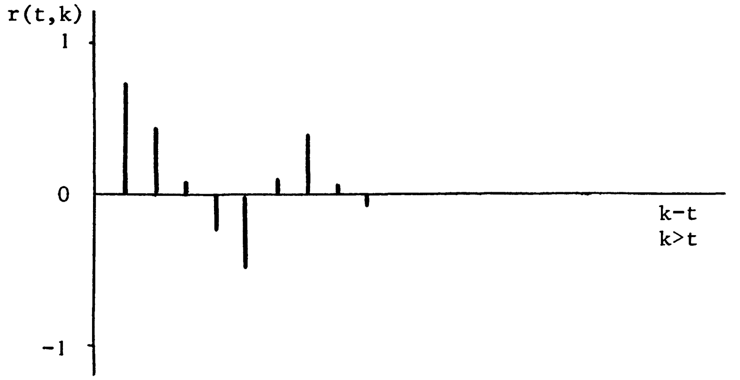
$$m_t = \frac{1}{|I|} \sum_{i \in I} x_t^i, \quad t \in \{1, \ldots, T\} ;$$

$$s(k, \ell) = \frac{1}{|I|} \sum_{i \in I} (x_k^i - m_k)(x_\ell^i - m_\ell), \quad (k, \ell) \in \{1, \ldots, T\}^2 ;$$

$$r(k, \ell) = \frac{s(k, \ell)}{\sqrt{s(k, k)}\sqrt{s(\ell, \ell)}}, \quad (k, \ell) \in \{1, \ldots, T\}^2 ;$$

$$c(t) = (r(t, t+1), r(t, t+2), \ldots, r(t, T)), \quad t \in \{1, \ldots, T-1\} .$$

Il apparaît dans certains cas de morphogénèse végétale (cf. [3]) que les termes des séquences c (t) présentent des variations analogues (sur les portions où on peut les comparer) : l'évolution des coefficients de corrélation $r(t, k), k > t$, en fonction de leur ordre $k - t$ présente une certaine stabilité pour tout t. On observe en outre que si $|r(t, k)|, k > t$, devient petit quand k croît, la séquence des $r(t, k), k > t$, peut présenter des oscillations amorties :



Les séquences x^i étant considérées comme des trajectoires indépendantes d'un processus $\{X_t \,; t \in \{1, \ldots, T\}\}$, on étudie une classe de modèles de processus dont les matrices de corrélation sont susceptibles de restituer certains des traits empiriquement mis en évidence.

III – LE MODELE AUTOREGRESSIF

Les X_t sont des v.a.r. de moyenne μ_t et de variance $\sigma(t, t)$. On désigne par $\sigma(t', t'')$ la covariance de $X_{t'}$ et de $X_{t''}$, par $\rho(t', t'')$ leur coefficient de corrélation. On note Y_t la v.a.r. réduite :

$$\frac{X_t - \mu_t}{\sqrt{\sigma(t, t)}} .$$

On suppose que le processus $\{Y_t \,; t \in \{1, \ldots, T\}\}$ est autorégressif au sens suivant :

$$Y_t = \sum_{j=1}^{p(t)} \alpha_t^j \, Y_{t-j} + \epsilon_t \quad (\text{où } \; p(t) \leqslant t - 1) \quad [E.1]$$

avec $\quad E[\epsilon_t] = 0 \quad$ et $\quad E[\epsilon_{t'} \epsilon_{t''}] = \begin{cases} 0 & \text{si } \; t' \neq t'' \\ \sigma_{\epsilon_t} & \text{si } \; t' = t'' = t \end{cases} .$

Y_t est encore une moyenne mobile d'ordre au plus égal à t :

$$Y_t = \sum_{j=0}^{t-1} \theta_t^j \, \epsilon_{t-j} \quad \text{avec} \quad \theta_t^0 = 1 .$$

D'où par suite :

$$E[Y_{t'} \epsilon_{t''}] = 0 \,, \quad t'' > t' \,, \quad \text{et} \quad E[Y_t \epsilon_t] = \sigma_{\epsilon_t} .$$

IV – LA MATRICE DES AUTOCORRELATIONS

Il vient :

$$\rho(k, t) = \sum_{j=1}^{p(k)} \alpha_k^j \, \rho(k - j, t) \,, \quad k > t \,, \quad [E.2] .$$

Ce sont donc des équations aux différences finies qui déterminent les éléments des séquences $\gamma(t) = (\rho(t, t+1), \rho(t, t+2), \ldots, \rho(t, T))$. Et l'on sait que les solutions de telles équations sont susceptibles d'osciller.

Remarque IV.1.

Il vient encore $\rho(k,t) = \sum_{j=1}^{p(t)} \alpha_t^j \rho(k, t-j) + E[\epsilon_t Y_k]$, $k \leqslant t$.

Exemple 1 :

Cas d'un processus autorégressif d'ordre constant égal à 1.

On note $\alpha_t^1 = \alpha_t$, $t \geqslant 2$. Il vient :

$$\left\{\begin{array}{l} \gamma(1) = (\alpha_2, \alpha_2, \alpha_3, \ldots, (\alpha_2, \ldots, \alpha_T)), \\ \ldots \\ \gamma(T-2) = (\alpha_{T-1}, \alpha_{T-1}\, \alpha_T). \\ \gamma(T-1) = (\alpha_T). \end{array}\right.$$

On note que $|\rho(t,k)| < 1$ et $\sigma_{\epsilon_t} > 0$ imposent $|\alpha_t| < 1$. Il y a décroissance monotone des modules des termes des séquences $\gamma(t)$, $t \in \{1, \ldots, T-1\}$.

Exemple 2 :

Cas d'un processus autorégressif d'ordre constant égal à 2 et à coefficients constants $\alpha_t^1 = \alpha_1$ et $\alpha_t^2 = \alpha_2$.

La même équation aux différences finies détermine les termes des séquences $\gamma(t)$, $t \in \{1, \ldots, T-1\}$:

$$\rho(t+k, t) = \alpha_1 \rho(t+k-1, t) + \alpha_2 \rho(t+k-2, t), \quad k \geqslant 1 .$$

Seules les valeurs initiales de ces séquences diffèrent : 0 et 1 pour $\gamma(1)$, $\rho(2,1)$ et 1 pour $\gamma(2), \ldots, \rho(t, t-1)$ et 1 pour $\gamma(t), \ldots$

Désignant par $P(z) = z^2 - \alpha_1 z - \alpha_2$ le polynôme caractéristique associé, de racines z_1 et z_2, il vient :

a) si $z_1, z_2 \in R$ et $z_1 \neq z_2$, $\rho(t+k, t) = A(t) z_1^k + B(t) z_2^k$,

b) si $z_1, z_2 \in R$ et $z_1 = z_2$, $\rho(t+k, t) = A(t) z_1^k + B(t)\, k\, z_1^k$

c) si $z_1 = \xi e^{i\omega}$ et $z_2 = \xi e^{-i\omega}$, $\rho(t+k, t) = \xi^k A(t) \sin(k\omega + \psi_t)$.

Ainsi si $P(z)$ a ses racines z_1 et z_2 dans le cercle unité, tous les termes des séquences $\gamma(t)$ tendent vers zéro en oscillant généralement.

Ces exemples montrent que, ayant réduit les variables X_t, des restrictions sont imposées aux α_k^j pour que soient vérifiés :

a) $|\rho(k,t)| \leqslant 1$,

b) $1 - \sum_{j=1}^{p(k)} \alpha_k^j \rho(k, k-j) = \sigma_{\epsilon_k} \geqslant 0$.

Remarque IV.2.

Il suffit d'imposer $\sum_{j=1}^{p(k)} |\alpha_k^j|^2 < \frac{1}{p(k)}$ pour que les solutions de [E.2] vérifient a) et b) ci-dessus. En effet :

$$|\rho(k,t)|^2 = \left|\sum_{j=1}^{p(k)} \alpha_k^j \rho(k-j,t)\right|^2 \leqslant \sum_{j=1}^{p(k)} |\alpha_k^j|^2 \sum_{j=1}^{p(k)} |\rho(k-j,t)|^2$$

$$\leqslant p(k) \sum_{j=1}^{p(k)} |\alpha_k^j|^2 .$$

C'est cette condition sur les α_k^j qui motive l'étude de l'extention à l'autorégression d'une procédure de "ridge-regression".

Remarque IV.3.

Si $\sum_{j=1}^{p(k)} |\alpha_k^j|^2 < \frac{1}{p(k)}$, le polynôme $P_k(z) = z^{p(k)} - \alpha_k^1 z^{p(k)-1} \ldots - \alpha_k^{p(k)}$, $z \in C$, a toutes ses racines dans le cercle unité. On retrouve donc une spécification courante des processus stationnaires autorégressifs $\{Y_t\ ;\ t \in Z\}$ (écriture canonique du modèle dont le processus est solution cf. [2] et [5]).

Démonstration

Soit $R_k(z) = z^{p(k)}$ et $Q_k(z) = P_k(z) - R_k(z), z \in C$. On a :

$$|R_k(z)| > |Q_k(z)|, \quad |z| = 1 .$$

Par suite (théorème de Rouché cf. [8] p. 392) $P_k(z) = Q_k(z) + R_k(z)$ et $R_k(z)$ ont même nombre de racines à l'intérieur du cercle unité.

V – IDENTIFICATION DES ORDRES p(t)

On utilise pour chaque valeur de t la propriété considérée par G. BOX et G. JENKINS pour identifier l'ordre d'un processus stationnaire autorégressif $\{Y_t\ ;\ t \in Z\}$ (cf. [2] pp. 54-56). En effet les équations de YULE-WALKER se généralisent aisément à ce cas non stationnaire.

Soit $Y_t = \sum_{j=1}^{p(t)} \alpha_t^j Y_{t-j} + \epsilon_t$. Soit $\mathcal{Y}_t$ l'espace de Hilbert engendré par les v.a.r. réduites $Y_t, \ldots, Y_1$ et muni du produit scalaire de la covariance, $\mathcal{Y}_t^p$ le sous-espace engendré par les v.a.r. $Y_{t-1}, \ldots, Y_{t-p}$ $(1 \leqslant p \leqslant t-1)$.

On désigne par $\hat{Y}_{j,p}$ la projection de Y_j, $Y_j \in \mathcal{Y}_t$, sur $\mathcal{Y}_t^p$.

Il vient :

$$\hat{Y}_{t,p} = a_t^1 Y_{t-1} + \cdots + a_t^p Y_{t-p} .$$

On appelle coefficient d'autocorrélation partielle d'ordre p + 1 au temps t le coefficient de corrélation (noté $\delta_t(p+1)$) des v.a. $Y_t - \hat{Y}_{t,p}$ et $Y_{t-p-1} - \hat{Y}_{t-p-1,p}$. On pose $\delta_t(1) = \rho(t, t-1)$ et $\delta_t(t) = 0$.

Théorème

$$\delta_t(p(t)) \neq 0 \quad \text{et} \quad \delta_t(p) = 0 , \quad p > p(t) .$$

Démonstration

$R_{t,p}$ désigne la matrice des autocorrélations des v.a. $Y_{t-1}, \ldots, Y_{t-p}$; elle est définie-positive par hypothèse. On note :

$$ {}^t r^0_{t,p} = [\rho(t, t-1) \ldots \rho(t, t-p)] , $$

$$ {}^t r^{p+1}_{t,p} = [\rho(t-p-1, t-1) \ldots \rho(t-p-1, t-p)] , $$

$$ {}^t a_{t,p} = [a^1_t \ldots a^p_t] , \quad s^0_{t,p} = \|Y_t - \hat{Y}_{t,p}\|^2 \quad \text{et} $$

$$ s^{p+1}_{t,p} = \|Y_{t-p-1} - \hat{Y}_{t-p-1,p}\|^2 . $$

Par analogie avec le cas stationnaire plus usuel (cf. [2] pp. 54-66) on appelle équations de YULE-WALKER d'ordre p au temps t le système :

$$ R_{t,p}\, a_{t,p} = r^0_{t,p} . $$

Il vient :

$$ a_{t,p} = R^{-1}_{t,p} r^0_{t,p} , \quad s^0_{t,p} = 1 - {}^t r^0_{t,p} R^{-1}_{t,p} r^0_{t,p} > 0 , \quad s^{p+1}_{t,p} = 1 - {}^t r^{p+1}_{t,p} R^{-1}_{t,p} r^{p+1}_{t,p} > 0. $$

On vérifie que

$$ \delta_t(p+1) = \frac{\langle Y_t - \hat{Y}_{t,p} , Y_{t-p-1} - \hat{Y}_{t-p-1,p} \rangle}{\|Y_t - \hat{Y}_{t,p}\| \, \|Y_{t-p-1} - \hat{Y}_{t-p-1,p}\|} = \frac{\rho(t, t-p-1) - {}^t r^{p+1}_{t,p} a_{t,p}}{\sqrt{s^0_{t,p}} \sqrt{s^{p+1}_{t,p}}} $$

L'inverse de $R_{t,p+1}$ a pour expression :

$$ R^{-1}_{t,p+1} = \left[\begin{array}{c|c} R^{-1}_{t,p} + \dfrac{1}{s^{p+1}_{t,p}} R^{-1}_{t,p} r^{p+1}_{t,p} {}^t r^{p+1}_{t,p} R^{-1}_{t,p} & -\dfrac{1}{s^{p+1}_{t,p}} R^{-1}_{t,p} r^{p+1}_{t,p} \\ \hline -\dfrac{1}{s^{p+1}_{t,p}} {}^t r^{p+1}_{t,p} R^{-1}_{t,p} & \dfrac{1}{s^{p+1}_{t,p}} \end{array} \right] $$

Et $\qquad {}^t r^0_{t,p+1} = [{}^t r^0_{t,p} \mid \rho(t, t-p-1)] .$

D'où : $\quad a^{p+1}_{t,p+1} = \dfrac{\rho(t, t-p-1) - {}^t r^{p+1}_{t,p} R^{-1}_{t,p} r^0_{t,p}}{s^{p+1}_{t,p}} = \dfrac{\sqrt{s^0_{t,p}}}{\sqrt{s^{p+1}_{t,p}}} \delta_t(p+1)$

Le modèle étant autorégressif d'ordre p (t), il vérifie :

a) ${}^t[\alpha^1_t \ldots \alpha^{p(t)}_t] = a_{t,p(t)} = R^{-1}_{t,p(t)} r^0_{t,p(t)}$;

b) $\rho(t, t-p(t)-h) = \sum_{j=1}^{p(t)} \alpha^j_t \rho(t-j, t-p(t)-h)$,

$$ h = 1, 2, \ldots, t-p(t)-1 . $$

On démontre alors par récurrence sur h que :

$$ \delta_t(p(t)+h) = a^{p(t)+h}_{t,p(t)+h} = 0 \quad \text{et} \quad a^j_{t,p(t)+h} = a^j_{t,p(t)} \ \forall j \in \{1, \ldots, p(t)\} \ \text{et} \ h \geqslant 1 . $$

Remarque V.1.

Si $\{Y_t \,;\, t \in Z\}$ est un processus stationnaire, on a $a^j_{t,p} = a^j_p$ et

$$\delta_t(p) = \delta(p) = a^p_p \,, \quad \forall t \in Z \,.$$

Remarque V.2.

On montre que $\|Y_t - \hat{Y}_{t,p+1}\|^2 = \|Y_t - \hat{Y}_{t,p}\|^2 \,(1 - \delta^2_t(p+1))$.

Disposant de plusieurs trajectoires indépendantes du processus

$$\{Y_t \,;\, t \in \{1\,, \ldots, T)\}$$

donc d'estimations empiriques $r(t\,,k)$ des $\rho(t\,,k)$, pour chaque valeur de t on résoud successivement les équations de YULE-WALKER d'ordre $1\,,2\,,\ldots,t-1$ et l'on peut tenter d'identifier les $p(t)$.

VI – RIDGE-AUTOREGRESSION

La restriction $\sum_{j=1}^{p(t)} |\alpha^j_t|^2 < 1/p(t)$ suggère l'emploi d'une procédure de "ridge-regression" pour estimer ${}^t\alpha_t = [\alpha^1_t \ldots \alpha^{p(t)}_t]$. On sait (cf. [4]) que l'estimation $\tilde{a}_{t,p(t)}$ de α_t est donnée par :

$$\tilde{a}_{t,p(t)} = (R_{t,p(t)} + \chi I)^{-1}\, r^0_{t,p(t)} \quad [E\text{-}3]$$

avec $\chi = 0$ si $\sum_{j=1}^{p(t)} |a^j_t|^2 \leqslant \dfrac{1}{p(t)}$, sinon $\chi > 0$ tel que $\sum_{j=1}^{p(t)} |\tilde{a}^j_t|^2 = \dfrac{1}{p(t)}$.

Dans la pratique, on applique [E-3] pour diverses valeurs positives de χ et on compare pour chaque valeur de χ la valeur de $\sum_{j=1}^{p(t)} |\tilde{a}^j_t|^2$ à $1/p(t)$. Compte tenu de [E-3] et les éléments diagonaux de $R_{t,p(t)}$ étant ici égaux à 1, la méthode consiste à résoudre les équations de YULE-WALKER : $\tilde{a}_{t,p(t)} = \tilde{R}^{-1}_{t,p(t)}\,\tilde{r}^0_{t,p(t)}$ où $\tilde{R}_{t,p(t)}$ est la matrice obtenue en multipliant tous les éléments extradiagonaux de $R_{t,p(t)}$ par $\dfrac{1}{1+\chi}$ et où $\tilde{r}^0_{t,p(t)}$ est le vecteur $\dfrac{1}{1+\chi} r^0_{t,p(t)}$. Ainsi on substitue aux estimations $r(k\,,e)$, $k \neq e$, de $\rho(k\,,e)$ l'estimation $\tilde{r}(k\,,e) = \dfrac{r(k\,,e)}{1+\chi}$.

Dans la méthode précédente, on atténue uniformément les autocorrélations $r(k\,,e)$ quelque soit leur ordre $|k-e|$. On est "naturellement" conduit à considérer des variantes dans lesquelles on fait :

$$\tilde{r}(k\,,e) = \lambda_{|k-e|}\, r(k\,,e)$$

où les coefficients $\lambda_{|k-e|}$ sont les éléments d'une suite $(\lambda_j)_{j\in N}$ décroissante, à valeurs positives et telle que $\lambda_0 = 1$.

Lemme 1 : (propriété du produit de Schur ou de Hadamard de deux matrices)

Soient $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices semi-definies-positives d'ordre p. Alors $C = (a_{ij}\, b_{ij})$ est semi-définie-positive.

De plus si A est définie-positive et si B est semi-définie-positive telle que $b_{ii} > 0$, $i \in \{1, \ldots, p\}$, alors C est définie-positive. (démonstration cf [1] p. 365).

Application

Soit $A = R_{t,p}$. Soit $\chi > 0$ et $B_p = (b_{ij})$ la matrice carrée d'ordre p telle que : $b_{ii} = 1$ et $b_{ij} = \dfrac{1}{1+\chi}$, $i \neq j$. Alors B_p est définie-positive (elle admet les valeurs propres $1 - \dfrac{1}{1+\chi} > 0$, de multiplicité $p-1$, et $1 + (p+1)\dfrac{1}{1+\chi} > 0$). Par suite $C = \widetilde{R}_{t,p}$ est définie-positive.

Plus généralement, on est amené à considérer des suites réelles, $(\lambda_k)_{k\in N}$, de type strictement positif (c'est-à-dire telles que toute matrice carrée d'ordre p $B_p = (b_{ij}) = (\lambda_{|i-j|})$ est définie-positive).

Lemme 2 :

Soit f une fonction positive, paire et continue sur $[-\pi, \pi]$. La suite $(\lambda_k)_{k\in N}$ de ses coefficients de Fourier est une suite de type positif (cf. [6]).

Exemple :

Les problèmes posés par l'estimation de la densité spectrale d'un processus stationnaire $\{Y_t ; t \in Z\}$ conduisent à recourir à de telles fonctions. La plus connue est celle de Bartlett (cf. [7] pp. 209-237) :

$$f(\phi) = \frac{1}{K}\left(\frac{\sin K\dfrac{\phi}{2}}{\sin\dfrac{\phi}{2}}\right)^2, \quad K \in N\,; \quad \text{d'où} \quad \lambda_j = \max\{1 - \frac{j}{K}\,; 0\}, \quad j \in N\,.$$

Propriétés de $\widetilde{a}_{t,p}$:

Soit $(\lambda_k)_{k\in N}$ une suite de type strictement positif utilisée pour définir $\widetilde{R}_{t,p}$. On désigne par L_p la matrice diagonale d'éléments diagonaux $\lambda_1, \ldots, \lambda_p$.

$\widetilde{a}_{t,p}$ se déduit de $a_{t,p}$ (solution des moindres carrés ordinaires) par une transformation linéaire. En effet :

$$\widetilde{a}_{t,p} = \widetilde{R}_{t,p}^{-1}\, \widetilde{r}_{t,p}^{0} = \widetilde{R}_{t,p}^{-1}\, L_p\, r_{t,p}^{0} = (\widetilde{R}_{t,p}^{-1}\, L_p\, R_{t,p})\, a_{t,p}\,.$$

Remarque :

On constante sur des exemples (voir notamment en annexe) que le choix de la "fenêtre" de Bartlett conduit, comme dans la "ridge" usuelle, à des estimations $\widetilde{a}_{t,p}$ telles que :

$${}^t\widetilde{a}_{t,p}\, \widetilde{a}_{t,p} < {}^t a_{t,p}\, a_{t,p}\,.$$

De plus $\widetilde{s}_{t,p}^{0} > s_{t,p}^{0}$.

Il est vraisemblable que ce sont des propriétés assez générales qu'on se propose d'étudier ultérieurement.

REFERENCES

[1] BALLANTINE C. – *On the Hadamard product.* Mathematische Zeitschrift 105 (1968), 365-366.

[2] BOX G. et JENKINS G. – *Time series analysis, forecasting and control.* Holden-Day 1970.

[3] BUIS R. et BRIERE C. – *Caractéristiques des matrices de corrélation dans les ensembles de variables ordonnées en morphogénése végétale, leur analyse en composante principale.* Séminaire de Statistique du Laboratoire de Statistique de l'Université Paul SABATIER (Toulouse) le 11 mars 1977.

[4] CAZES P. – *Protection de la régression, régression sous contrainte.* Document 74/17 du Laboratoire de Statistique Mathématique de l'Université Paris VI.

[5] DOSSOU-GBETE S. et al. – Contribution à l'étude des processus ARMA. *Publication du Laboratoire de Statistique de l'Université Paul SABATIER* (Toulouse), n° 05, 1976.

[6] GRENANDER V. et SGEGO G. – *Toeplitz forms and their application.* University of California Press, 1958.

[7] JENKINS G et WATTS D. – *Spectral analysis and its applications.* Hoden-Day 1968.

[8] VALIRON G. – *Théorie des fonctions.* Masson, 1955.

EXEMPLE

(données communiquées par R. BUIS et C. BRIERE)

Ayant sélectionné dans un même lot de sarrasin argenté (*Fagopyrum Esculentum* Moench) les tiges présentant à l'âge adulte 12 entre-noeuds, on mesure sur chacune de ces tiges i, $i \in \{1, \ldots, 23\}$, la longueur des 12 feuilles associées, chaque nœud ne portant qu'une feuille. Les corrélations empiriques des variables représentatives des longueurs des feuilles donnent, pour ce petit échantillon ($|I| = 23$) la matrice de corrélation suivante (cf. [3]) :

1.00	.75	.58	.51	.35	.07	–.20	–.18	–.12	.14	.31	.28
.75	1.00	.65	.69	.52	.02	–.35	–.24	–.23	–.01	.16	.27
.58	.65	1.00	.63	.50	.21	–.18	–.37	–.13	–.09	.02	.19
.51	.69	.63	1.00	.78	.45	–.03	–.20	–.27	–.13	.11	.35
.35	.52	.50	.78	1.00	.56	.08	.03	–.17	–.19	.08	.31
.07	.02	.21	.45	.56	1.00	.69	.50	.37	.14	.17	.16
–.20	–.35	–.18	–.03	–.08	.69	1.00	.56	.54	.20	.19	.16
–.18	–.24	–.37	–.20	.03	.50	.56	1.00	.66	.31	.06	–.15
–.12	–.23	–.13	–.27	–.17	.37	.54	.66	1.00	.74	.45	.27
.14	–.01	–.09	–.13	–.19	.14	.20	.31	.74	1.00	.74	.53
.31	.16	.02	.11	.08	.17	.19	.06	.45	.74	1.00	.78
.28	.27	.19	.35	.31	.16	.16	–.15	.27	.53	.78	1.00

On cherche à étudier une relation exprimant que la longueur de la feuille de niveau t, t variant de 2 à 12, dépend des longueurs des feuilles des niveaux inférieurs.

Les tableaux I à VI ci-après concernent la sixième feuille. Le tableau I donne les résultats de l'autorégression ordinaire d'ordre p(6), p(6) variant pas à pas de 1 à 5. On notera par exemple l'évolution des α_6^j que provoque l'introduction de la variable explicative Y_2 (p(6) variant de 3 à 4) faiblement corrélée à Y_6 et fortement corrélée à Y_5, Y_4 et Y_3 (effet classique de multicolinéarité). Les tableaux II à VI permettent de comparer, pour un ordre d'autorégression p(6) fixé, les résultats de l'autorégression sous fenêtre de Bartlett d'ordre K, $K \in \{2, \ldots, 30\}$, aux résultats précédents (moindres carrés ordinaires). On constate sur cet exemple que la méthode, comme la "ridge classique", permet de réduire la somme des carrés des coefficients d'autorégression.

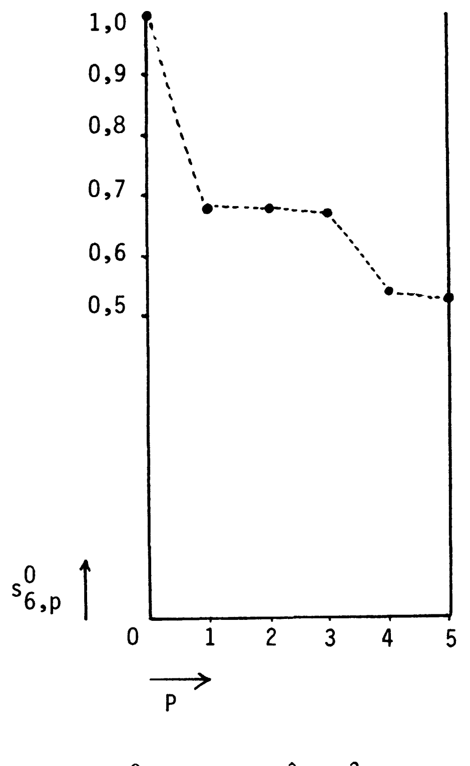
Compte tenu des résultats analogues obtenus pour les autres feuilles il semblerait que l'on puisse retenir pour chaque feuille un modèle autorégressif, d'ordre 2 au maximum, susceptible par suite de restituer des corrélations de signe négatif. L'étroitesse de l'échantillon diminue toutefois la portée de ce résultat.

TABLEAU I

Résultats concernant la sixième feuille

(autorégression ordinaire)

p(6)	α_6^1	α_6^2	α_6^3	α_6^4	α_6^5	$\delta_6(p(6))$	$\sum_{j=1}^{p}(\alpha_6^j)^2$	$s^0_{6,p}$
1	0.564					0.564	0.318	0.682
2	0.550	0.017				0.013	0.303	0.682
3	0.555	0.096	-0.131			- 0.124	0.334	0.672
4	0.521	0.377	0.068	-0.552		- 0.450	0.722	0.535
5	0.534	0.378	0.038	-0.658	0.159	0.140	0.888	0.525



Variation de $s^0_{6,p} = \| Y_6 - \hat{Y}_{6,p} \|^2, p \in \{1, \ldots, 5\}$
(en posant $s^0_{6,0} = \| Y_6 \|^2 = 1$)

TABLEAU II

$p(6) = 1 : Y_{6,1} = \alpha_6^1 Y_5$

		α_6^1	$\delta_6(2)$	$\sum_{j=1}^{p(6)} (\alpha_6^j)^2$	$s_{6,1}^0$
FENETRE DE BARTLETT, K =	2	0.282	-0.124	0.079	0.921
	3	0.376	-0.059	0.141	0.859
	4	0.423	-0.033	0.179	0.821
	5	0.451	-0.019	0.203	0.797
	6	0.470	-0.011	0.221	0.779
	7	0.483	-0.006	0.233	0.767
	8	0.493	-0.003	0.243	0.757
	9	0.501	-0.000	0.251	0.749
	10	0.507	0.002	0.257	0.743
	⋮				
	15	0.526	0.007	0.277	0.723
	⋮				
	20	0.535	0.009	0.287	0.713
	⋮				
	25	0.541	0.010	0.293	0.707
	⋮				
	30	0.545	0.011	0.297	0.703
	⋮				
M.C.O.(*)		0.564	0.013	0.318	0.682

(*) M.C.O. = moindres carrés ordinaires

TABLEAU III

$p(6) = 2 : \hat{Y}_{6,2} = \alpha_6^1 Y_5 + \alpha_6^2 Y_4$

		α_6^1	α_6^2	$\delta_6(3)$	$\sum_{j=1}^{p(6)} (\alpha_6^j)^2$	$s_{6,2}^o$
FENETRE DE BARTLETT , K =	2	0.332	- 0.129	0.045	0.127	0.906
	3	0.409	- 0.064	- 0.050	0.171	0.856
	4	0.444	- 0.036	- 0.053	0.198	0.820
	5	0.465	- 0.022	- 0.059	0.216	0.796
	6	0.478	- 0.013	- 0.066	0.229	0.779
	7	0.488	- 0.007	- 0.071	0.238	0.767
	8	0.495	- 0.003	- 0.076	0.245	0.757
	9	0.501	- 0.000	- 0.080	0.251	0.749
	10	0.506	0.002	- 0.083	0.256	0.743
	. . .					
	15	0.520	0.008	- 0.094	0.270	0.723
	. . .					
	20	0.527	0.011	- 0.101	0.278	0.713
	. . .					
	25	0.532	0.013	- 0.105	0.283	0.707
	. . .					
	30	0.535	0.014	- 0.108	0.286	0.703
	. . .					
M.C.O.(*)		0.550	0.017	-0.124	0.303	0.682

(*) M.C.O. = moindres carrés ordinaires

TABLEAU IV

$$p(6) = 3 : \hat{Y}_{6,3} = \alpha_6^1 Y_5 + \alpha_6^2 Y_4 + \alpha_6^3 Y_3$$

		α_6^1	α_6^2	α_6^3	$\delta_6(4)$	$\sum_{j=1}^{p(6)} (\alpha_6^j)^2$	$s_{6,3}^o$
FENETRE DE BARTLETT, K =	2	0.339	-0.146	0.046	- 0.017	0.138	0.905
	3	0.405	-0.040	-0.051	0.038	0.168	0.854
	4	0.442	-0.010	-0.055	- 0.035	0.198	0.818
	5	0.463	0.010	-0.061	- 0.086	0.219	0.794
	6	0.478	0.023	-0.068	- 0.126	0.233	0.776
	7	0.488	0.032	-0.074	- 0.158	0.245	0.763
	8	0.496	0.040	-0.079	- 0.184	0.254	0.752
	9	0.502	0.045	-0.083	- 0.206	0.261	0.744
	10	0.507	0.050	-0.087	- 0.225	0.267	0.738
	. . .						
	15	0.522	0.065	-0.099	- 0.288	0.287	0.717
	. . .						
	20	0.530	0.072	-0.106	- 0.323	0.297	0.706
	. . .						
	25	0.535	0.077	-0.111	- 0.346	0.304	0.699
	. . .						
	30	0.539	0.080	-0.114	- 0.364	0.310	0.694
	. . .						
M.C.O.(*)		0.555	0.096	-0.131	- 0.450	0.334	0.672

(*) M.C.O. = moindres carrés ordinaires

TABLEAU V

$$p(6) = 4 : \hat{Y}_{6,4} = \alpha_6^1 Y_5 + \alpha_6^2 Y_4 + \alpha_6^3 Y_3 + \alpha_6^4 Y_2$$

FENETRE DE BARTLETT, K =	α_6^1	α_6^2	α_6^3	α_6^4	$\delta_6(5)$	$\sum_{j=1}^{p(6)} (\alpha_6^j)^2$	$s_{6,4}^o$
2	0.340	- 0.148	0.052	- 0.017	0.007	0.140	0.904
3	0.411	- 0.046	- 0.067	0.039	-0.009	0.177	0.852
4	0.439	- 0.002	- 0.039	- 0.037	0.043	0.195	0.817
5	0.457	0.032	- 0.023	- 0.091	0.034	0.218	0.788
6	0.469	0.061	- 0.012	- 0.136	0.044	0.242	0.764
7	0.478	0.086	- 0.003	- 0.173	0.052	0.266	0.744
8	0.484	0.107	0.004	- 0.204	0.059	0.288	0.727
9	0.490	0.125	0.010	- 0.230	0.065	0.308	0.712
10	0.494	0.141	0.015	- 0.253	0.069	0.328	0.700
⋮							
15	0.505	0.198	0.030	- 0.330	0.086	0.404	0.658
⋮							
20	0.510	0.233	0.039	- 0.376	0.097	0.457	0.632
⋮							
25	0.513	0.257	0.044	- 0.406	0.104	0.496	0.616
⋮							
30	0.515	0.276	0.048	- 0.438	0.110	0.530	0.602
⋮							
M.C.O.(*)	0.521	0.377	0.068	-0.552	0.140	0.722	0.535

(*) M.C.O. = moindres carrés ordinaires

TABLEAU VI

$$p(6) = 5 : \hat{Y}_{6,5} = \alpha_6^1 Y_5 + \alpha_6^2 Y_4 + \alpha_6^3 Y_3 + \alpha_6^4 Y_2 + \alpha_6^5 Y_1$$

FENETRE DE BARTLETT, K =	α_6^1	α_6^2	α_6^3	α_6^4	α_6^5	$\delta_6(6)$ (**)	P(6) $\sum_{j=1} (\alpha_6^j)^2$	$s_{6,5}^o$
2	0.340	-0.149	0.053	-0.020	0.007	0	0.141	0.904
3	0.411	-0.047	- 0.066	0.044	0.009	0	0.178	0.852
4	0.441	0.001	- 0.042	-0.063	0.048	0	0.202	0.816
5	0.458	0.034	- 0.026	-0.114	0.038	0	0.226	0.787
6	0.471	0.064	- 0.016	-0.166	0.050	0	0.256	0.762
7	0.480	0.088	- 0.009	-0.209	0.059	0	0.286	0.742
8	0.487	0.109	- 0.004	-0.245	0.067	0	0.314	0.724
9	0.493	0.128	0.001	-0.276	0.074	0	0.341	0.710
10	0.497	0.144	0.004	-0.303	0.079	0	0.366	0.697
. . .								
15	0.511	0.201	0.015	-0.394	0.099	0	0.466	0.653
. . .								
20	0.517	0.236	0.021	-0.448	0.111	0	0.536	0.626
. . .								
25	0.521	0.259	0.024	-0.484	0.119	0	0.587	0.609
. . .								
30	0.524	0.279	0.027	-0.613	0.125	0	0.632	0.595
. .								
(*) M.C.O.	0.534	0.378	0.038	-0.650	0.158	0	0.888	0.525

(*) M.C.O. = moindres carrés ordinaires

(**) $\delta_6(6) = \delta_t(t) = 0$ par convention