Régression PLS sur un processus stochastique

Cristian Preda, Gilbert Saporta

REVUE DE STATISTIQUE APPLIQUÉE

CRISTIAN PREDA
GILBERT SAPORTA

Régression PLS sur un processus stochastique

Revue de statistique appliquée, tome 50, n° 2 (2002), p. 27-45
<http://www.numdam.org/item?id=RSA_2002__50_2_27_0>

Rev. Statistique Appliquée, 2002, **L** (2), 27-45

RÉGRESSION PLS SUR UN PROCESSUS STOCHASTIQUE

Cristian PREDA[(1)] et Gilbert SAPORTA[(2)]

[(1)] *CERIM - Département de Statistique, Faculté de Médecine, Université de Lille 2, 1, place de Verdun, 59045 Lille cedex, e-mail : cpreda@univ-lille2.fr*

[(2)] *CNAM - Paris, Chaire de Statistique Appliquée, 292, rue Saint Martin, 75141 Paris cedex 03, e-mail : saporta@cnam.fr*

RÉSUMÉ

Après avoir montré le principe de la régression PLS dans le cas fini, nous allons développer la régression PLS sur un processus $(X_t)_{t\in[0,T]}$ L_2-continu. On montre l'existence des composantes PLS comme vecteurs propres d'un certain opérateur ainsi que la convergence de l'approximation PLS vers celle donnée par la régression classique. Les résultats d'une application sur des données boursières seront comparés avec ceux fournis par d'autres méthodes.

Mots-clés : *Régression PLS, opérateur d'Escoufier, analyse en composantes principales.*

ABSTRACT

We give an extension of PLS regression to the case where the set of predictor variables forms a L_2-continuous stochastic process and the response is a random vector of finite or infinite dimension. We prove the existence of PLS components as eigenvectors of some operator and also some convergence properties of the PLS approximation. The results of an application to stock-exchange data will be compared with those obtained by others methods.

Keywords : *PLS regression, Escoufier's operator, principal component analysis.*

1. Introduction

Il ne semble pas usuel d'effectuer une régression linéaire sur une infinité de variables explicatives. Cela correspond pourtant à la situation suivante souvent rencontrée (Figure 1) : on observe n courbes (ou trajectoires) en continu sur l'intervalle de temps [0,T] – que nous allons considérer comme réalisations d'un processus stochastique $(X_t)_{t\in[0,T]}$ – et on veut utiliser cette information pour prédire une réponse Y qui peut être X_{T+h} – on parle alors de prédiction à l'horizon h, $h > 0$ – ou une variable aléatoire réelle externe quelconque (par exemple, $(X_t)_{t\in[0,T]}$ peut représenter des courbes de températures observées en n lieux et Y le montant de récoltes). Théoriquement, cela s'exprime par la régression de la variable Y sur le processus $(X_t)_{t\in[0,T]}$.

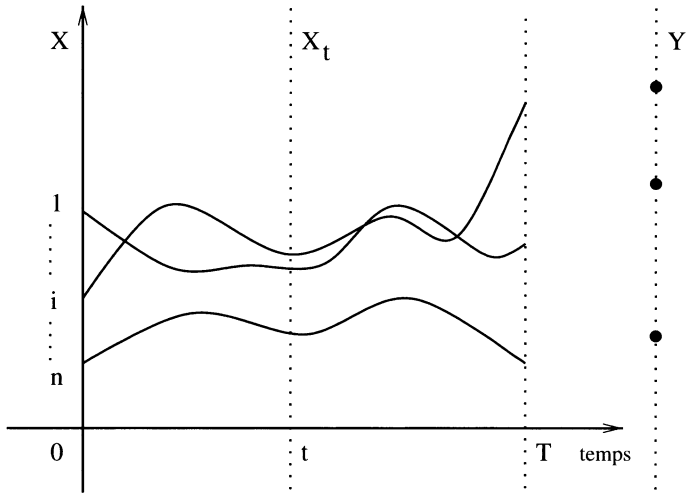

FIGURE 1
Régression sur un processus

Le but de cet article est d'adapter la régression PLS lorsque l'ensemble de variables explicatives est un processus stochastique.

Les problèmes posés par la régression linéaire classique sur un processus – l'indétermination des coefficients de régression (Ramsay *et al.* [10], [11], Saporta [12]) ou encore le choix des composantes principales de $(X_t)_{t\in[0,T]}$ comme variables explicatives (Deville [4], Saporta [12], Aguilera *et al.* [1]) – trouvent dans ce cadre des solutions satisfaisantes dont les principales propriétés découlent de celles de l'opérateur d'Escoufier associé au processus $(X_t)_{t\in[0,T]}$ (Saporta [12]).

L'article est divisé en 6 sections. Nous faisons d'abord (Section 2) un rappel théorique de la régression sur un processus en mettant en évidence les principales difficultés rencontrées et présentons brièvement la régression sur les composantes principales du processus. Le principe de la régression PLS d'une v.a.r. Y sur un vecteur aléatoire de dimension finie $\mathbf{X} = (X_i)_{i=1,\ldots,p}$ est présenté (*cf.* Tenenhaus [14]) dans la Section 3. Dans la Section 4 on développe la régression PLS sur un processus L_2-continu $\mathbf{X} = (X_t)_{t\in[0,T]}$ lorsque l'ensemble de variables à expliquer est formé d'une seule variable (PLS univariée). On montre l'existence des composantes PLS ainsi que quelques propriétés de convergence vers la régression linéaire classique. Le cadre général de la régression PLS sur un processus est développé dans la Section 5 et correspond au cas où $\mathbf{Y}$ est un vecteur aléatoire multidimensionnel. Le cas particulier $\mathbf{Y} = (X_t)_{t\in[T,T+a]}$, $a > 0$ est traité dans la dernière partie de cette section et offre une alternative aux méthodes de prévision proposées par Aguilera *et al.* ([1]) et Ramsay *et al.* ([11]). La Section 6 présente une application de la régression PLS ayant comme but la prévision. Les données auxquelles nous nous intéressons sont un ensemble de 85 actions cotées à la Bourse de Paris. Les résultats [1] de cette analyse

[1] Ces résultats ont été obtenus utilisant le logiciel SIMCA-P (Umetri 1996, [15])

seront comparés avec ceux fournis par la régression sur les composantes principales et l'algorithme NIPALS ([14]).

2. Quelques rappels théoriques

Soit $(X_t)_{t\in[0,T]}$ un processus du second ordre L_2-continu et Y une variable aléatoire réelle définie sur le même espace de probabilité, $(\mathbf{\Omega}, \mathcal{A}, \mathbf{P})$. Notons par $L_2(X)$ le sous-espace de Hilbert de $L_2(\mathbf{\Omega})$ engendré par les combinaisons linéaires finies de variables $\{X_t : t \in [0,T]\}$ et supposons que $E(X_t) = 0$, $\forall t \in [0,T]$ et $E(Y) = 0$.

L'analyse en composantes principales du processus $(X_t)_{t\in[0,T]}$ a été définie par Deville ([3]) et repose essentiellement sur la décomposition de Karhunen-Loève de $(X_t)_{t\in[0,T]}$ sous la forme d'une somme de processus quasi-déterministes ([12]) :

$$X_t = \sum_{i\geqslant 1} \xi_i f_i(t), \tag{1}$$

où :

- la famille de fonctions $\{f_i\}_{i\geqslant 1}$ (facteurs principaux) forme un système orthonormal dans $L_2([0,T])$ et vérifie l'équation aux valeurs propres :

$$\int_0^T C(t,s) f_i(s) ds = \lambda_i f_i(t),\ \forall t \in [0,T], \tag{2}$$

  où $C(t,s) = E(X_t X_s)$, $\forall s,t \in [0,T]$.

- la famille des variables aléatoires $\{\xi_i\}_{i\geqslant 1}$ (composantes principales) définies par

$$\xi_i = \int_0^T X_t f_i(t) dt,$$

  forme un système orthogonal en $L_2(\Omega)$ tel que $Var(\xi_i) = \lambda_i$, $\forall i \geqslant 1$.

L'utilité des ces éléments (dérivés de la structure de $(X_t)_{t\in[0,T]}$) dans l'étude de la régression linéaire sur un processus est présentée de manière synthétique dans le paragraphe suivant.

2.1. Régression linéaire sur un processus

La meilleure prévision linéaire de Y à partir de $(X_t)_{t\in[0,T]}$ est la projection de Y sur $L_2(X)$. Soit $\hat{Y}$ cette projection. D'après le théorème de projection (équivalent ici au critère des moindres carrés) on a :

$$E(YX_t) = E(\hat{Y}X_t),\ \ \forall t \in [0,T]. \tag{3}$$

Si l'approximation $\hat{Y}$ se met sous la forme

$$\hat{Y} = \int_0^T \beta(t) X_t dt, \ \ \beta \in L_2([0,T]), \tag{4}$$

la fonction β vérifie alors l'équation de Wiener-Hopf

$$E(X_t Y) = \int_0^T C(t,s)\beta(s)ds, \ \ \forall t \in [0,T], \tag{5}$$

Le théorème de Picard ([12]) montre que l'équation (5) a une solution unique dans $L_2([0,T])$ si et seulement si :

$$\sum_{k\geqslant 1} \frac{c_k^2}{\lambda_k^2} < \infty, \ \ c_k = \int_0^T E(X_t Y) f_k(t) dt, \tag{6}$$

où $\{\lambda_k\}_{k\geqslant 1}$ et $\{f_k\}_{k\geqslant 1}$, sont respectivement la suite décroissante de valeurs propres, respectivement la suite de fonctions propres, de l'opérateur de covariance du processus $(X_t)_{t\in[0,T]}$. Si la condition (6) n'est pas satisfaite, en général β est une distribution ([12]). Cette difficulté est également mise en évidence dans le cas où l'approximation $\hat{Y}$ est cherchée à l'aide d'un échantillon aléatoire simple de taille N. En effet, soit $\{(Y_1, X_1), (Y_2, X_2), \ldots (Y_N, X_N)\}$ un échantillon aléatoire de taille N obtenu par tirages indépendants du couple (Y, X), où $X = (X_t)_{t\in[0,T]}$. Le système

$$Y_i = \int_0^T X_i(t)\beta(t)dt, \ \ \forall i = 1, ..., p, \tag{7}$$

admet une infinité de solutions $\hat{\beta}$ et donc l'erreur quadratique moyenne associée à la régression linéaire de Y sur $(X_t)_{t\in[0,T]}$ est nulle. Mais dans ce cas la fonction $\hat{\beta}$ n'apporte aucune information pertinente et le critère des moindres carrés ne fournit pas un estimateur consistant de β ([11]). Une solution possible est présentée dans la section suivante.

2.2. Régression sur les composantes principales

La famille de composantes principales $\{\xi_k\}_{k\geqslant 1}$ est une base pour $L_2(X)$ et, par le théorème de projection, la régression de Y sur $(X_t)_{t\in[0,T]}$ est équivalente à celle de Y sur l'ensemble de variables $\{\xi_k\}_{k\geqslant 1}$. On a donc :

$$\hat{Y} = \sum_{k\geqslant 1} \frac{E(Y\xi_k)}{\lambda_k} \xi_k. \tag{8}$$

La qualité de la régression est souvent faite en mesurant le carré du coefficient de corrélation entre Y et $\hat{Y}$:

$$R^2(Y,\hat{Y}) = \frac{1}{E(Y^2)} \sum_{k\geqslant 1} \frac{E^2(Y\xi_k)}{\lambda_k},$$

de sorte que

$$E(Y-\hat{Y})^2 = \Big(1 - R^2(Y,\hat{Y})\Big)\, E(Y^2).$$

Dans la pratique on choisit généralement une approximation d'ordre p, $p \geqslant 1$:

$$\hat{Y}^p = \sum_{k=1}^{p} \frac{E(Y\xi_k)}{\lambda_k}\xi_k. \tag{9}$$

Cependant, le choix des p composantes principales de la régression de Y sur $(X_t)_{t\in[0,T]}$ reste difficile à faire car, si le but est de maximiser le coefficient de corrélation entre Y et $\hat{Y}^p$ (cette situation convient lorsqu'on cherche à décrire l'ensemble des données), il est tout a fait possible que les p composantes principales choisies ne soient pas les p premières dans l'ordre décroissant des variances expliquées. On risque alors d'avoir des composantes de régression de variance faible (peu explicatives de $(X_t)_{t\in[0,T]}$) qui entraînent une forte instabilité des coefficients de régression ([7]). Le choix des p composantes est donc un compromis entre la stabilité du modèle linéaire et la grandeur du coefficient de corrélation entre Y et $\hat{Y}^p$. La régression PLS ([13]) donne une solution à ce compromis.

3. Rappels sur la régression PLS

Pour fixer notre propos, on va considérer le cas de la régression PLS d'un variable aléatoire réelle Y sur un vecteur aléatoire réel $\mathbf{X} = (X_1, X_2, \ldots, X_p)$ (régression PLS1, *cf.* [14]). Le principe de la régression PLS de Y sur $\mathbf{X}$ est de chercher de manière itérative un ensemble de variables aléatoires non corrélées $\{t_i\}_{1\leqslant i\leqslant p}$ (composantes PLS), combinaisons linéaires des variables $\{X_i\}_{1\leqslant i\leqslant p}$, vérifiant le critère d'optimalité de Tucker :

(T) Chercher $t_1 = \sum_{i=1}^{p} a_{1,i}X_i$, $a_1 \in \mathbf{R}^p$, $\|a_1\| = 1$, tel que $\mathrm{Cov}^2(Y,t_1)$ soit maximal.

Il est alors facile de voir ([13]) que $a_{1,i} = \dfrac{E(X_iY)}{\sqrt{\sum_{j=1}^{p} E^2(X_jY)}}$, $\forall i = 1,\ldots,p$.

De manière générale, si on pose $Y_0 = Y$et $\mathbf{X}_0 = \mathbf{X}$, au pas h de la régression PLS1 on construit la variable t_h qui maximise le critère de Tucker pour les groupes

de variables $\mathbf{X_{h-1}} = \{X_{h-1,i}\}_{i=1,\ldots,p}$ et $\{Y_{h-1}\}$:

$$t_h = \frac{\sum_{i=1}^{p} E(X_{h-1,i}Y_{h-1})X_{h-1,i}}{\sqrt{\sum_{i=1}^{p} E^2(X_{h-1,i}Y_{h-1})}}. \tag{10}$$

La variable Y_{h-1} est ensuite régressée sur t_h :

$$Y_{h-1} = c_h t_h + Y_h,$$

où $c_h = \dfrac{E(Y_{h-1}t_h)}{E(t_h^2)}$ et Y_h le résidu. $\mathbf{X}_h$ est alors le vecteur aléatoire des résidus de la régression des variables $\{X_{h-1,i}\}_{i=1,\ldots,p}$ sur la variable t_h :

$$\mathbf{X}_h = \mathbf{X}_{h-1} - \mathbf{p}_h t_h,$$

où $\mathbf{p}_h = (p_{h,i})_{i=1,\ldots,p}$ est le vecteur des coefficients de la variable t_h dans la régression de $X_{h-1,i}$ sur t_h, $i = 1, ..., p$.

L'ensemble de variables $\{t_i\}_{i=1,\ldots,h}$ forme un système orthogonal dans $L_2(\Omega)$ et pour tout h tel que $t_h \neq 0$, on a les propriétés de décomposition suivantes :

a) $Y = c_1 t_1 + c_2 t_2 + \ldots c_h t_h + Y_h$,

b) $X = p_1 t_1 + p_2 t_2 + \ldots p_h t_h + X_h$.

Remarque 1. – Si on note par $\mathcal{T}_h$ l'espace engendré linéairement par les variables $\{t_i\}_{i=1,\ldots,h}$, $h \geqslant 1$, et par $\hat{Y}$ la projection orthogonale de Y sur l'espace engendré linéairement par les variables $\{X_i\}_{i=1,\ldots,p}$, (soit $L_2(X)$), alors le nombre des pas de la régression PLS complète de Y sur $\mathbf{X}$ est le plus petit indice n tel que $\hat{Y}$ appartient à $\mathcal{T}_n$. Notons au passage que $\mathcal{T}_n$ ne s'identifie pas toujours avec $L_2(X)$ comme c'est le cas de l'espace engendré par les composantes principales du vecteur $\mathbf{X}$. □

On note par $\hat{Y}_h$ l'approximation de Y donnée par la régression PLS sur le vecteur $\mathbf{X}$ au pas h :

$$\hat{Y}_h = c_1 t_1 + c_2 t_2 + \ldots c_h t_h. \tag{11}$$

Notons enfin que la relation d'inclusion entre les espaces $\{\mathcal{T}_h\}_{h\geqslant 1}$ définis dans la Remarque 1, entraîne également :

$$R^2(Y, \hat{Y}_h) \leqslant R^2(Y, \hat{Y}_{h+1}) \leqslant R^2(Y, \hat{Y}), \ \ \forall h \geqslant 1. \qquad \square$$

Outre sa simplicité, la régression PLS s'avère efficace surtout dans le cas où le nombre d'observations de l'échantillon est inférieur au nombre de variables

explicatives. C'était ici la principale source des difficultés rencontrées dans la régression sur un processus ([11]). En plus, comme on l'a déjà évoqué, les coefficients de la régression PLS sont, dans la plupart des cas, plus stables que dans la régression classique (variances plus faibles), point essentiel lorsque le but de la régression est la prédiction.

4. Régression PLS univariée sur un processus

On considère ici que l'ensemble de variables explicatives est un processus $(X_t)_{t\in[0,T]}$ du second ordre, L_2-continu et la variable à expliquer est une variable aléatoire réelle unidimensionnelle Y de $L^2(\Omega)$. Supposons que $E(Y) = 0$ et $E(X_t) = 0, \forall t \in [0,T]$.

PROPOSITION 2.

$$\max_{\substack{w\in L_2([0,T])\\ \|w\|=1}} Cov^2(\int_0^T w(t)X_t dt, Y)$$

est réalisé pour w telle que :

$$w(t) = \frac{E(X_t Y)}{\sqrt{\int_0^T E^2(X_t Y)dt}}, \quad \forall t \in [0,T]. \tag{12}$$

Preuve.

$$Cov^2(\int_0^T w(t)X_t dt, Y) = \left[\int_0^T E(X_t Y)w(t)dt\right]^2 \leqslant \int_0^T E^2(X_t Y)dt.$$

L'inégalité de Schwarz devient égalité pour

$$w(t) = \frac{E(X_t Y)}{\sqrt{\int_0^T E^2(X_t Y)dt}}, \quad \forall t \in [0,T]. \qquad \square$$

On définit la première composante PLS de la régression de Y sur le processus $(X_t)_{t\in[0,T]}$ par la variable :

$$t_1 = \frac{\int_0^T E(X_t Y)X_t dt}{\sqrt{\int_0^T E^2(X_t Y)dt}} = \frac{\mathbf{W}Y}{\sqrt{E(Y\cdot \mathbf{W}Y)}}, \tag{13}$$

où $\mathbf{W}$ est l'opérateur d'Escoufier [2] ([5]), $\mathbf{W} : L_2(\Omega) \to L_2(\Omega)$, défini par :

$$\mathbf{W}Z = \int_0^T E(ZX_t)X_t dt, \ \ \forall Z \in L_2(\Omega).$$

La régression de Y sur t_1 s'écrit :

$$Y = c_1 t_1 + Y_1,$$

où $c_1 = \dfrac{E(Yt_1)}{E(t_1^2)}$ et Y_1 est le résidu. Notons par p_1 la fonction de $L_2([0,T])$,

$$p_1(t) = \frac{E(X_t t_1)}{E(t_1^2)}, \ \ \forall t \in [0,T].$$

$p_1(t)$ représente le coefficient de t_1 de la régression de X_t sur t_1. On obtient ainsi le processus résiduel $(X_{1,t})_{t\in[0,T]}$ de la régression de $(X_t)_{t\in[0,T]}$ sur t_1 :

$$X_{1,t} = X_t - p_1(t)t_1, \ \ \forall t \in [0,T].$$

La deuxième composante PLS1 est définie de la même manière que t_1 en considérant le processus $(X_{1,t})$ et la variable Y_1. De façon générale, si on pose $X_{0,t} = X_t$, $\forall t \in [0,T]$, et $Y_0 = Y$, on définit le pas h, $h \geqslant 1$, de la régression PLS1 de Y sur $(X_t)_{t\in[0,T]}$ par les étapes suivantes :

1) on construit la composante PLS1 t_h,

$$t_h = \frac{\mathbf{W}_{h-1}Y_{h-1}}{\sqrt{E(Y_{h-1} \cdot \mathbf{W}_{h-1}Y_{h-1})}}, \tag{14}$$

où $\mathbf{W}_{h-1}$ est l'opérateur d'Escoufier associé au processus $(X_{h-1,t})_{t\in[0,T]}$,

2) on pose le modèle linéaire :

$$Y_{h-1} = c_h t_h + Y_h, \quad \text{où } \ c_h = \frac{E(Y_{h-1}t_h)}{E(t_h^2)},$$

3) on construit le processus résiduel $(X_{h,t})_{t\in[0,T]}$ de la régression de $(X_{h-1,t})_{t\in[0,T]}$ sur t_h :

$$X_{h,t} = X_{h-1,t} - p_h(t)t_h,$$
$$p_h(t) = \frac{E(X_{h-1,t}t_h)}{E(t_h^2)}, \ \ \forall t \in [0,T].$$

[2] L'analyse spectrale de $\mathbf{W}$ donne les composantes principales de l'ACP du processus $(X_t)_{t\in[0,T]}$ ([12])

Comme dans le cas fini, on a le résultat suivant :

PROPOSITION 3. – $\forall h \geqslant 1$:

a) $\{t_h\}_{h\geqslant 1}$ *forment un système orthogonal dans* $L_2(X)$,

b) $Y = c_1 t_1 + c_2 t_2 + \ldots c_h t_h + Y_h$,

c) $X_t = p_1(t)t_1 + p_2(t)t_2 + \ldots p_h(t)t_h + X_{h,t}$,

d) $E(Y_h t_i) = 0, \forall i = 1, ..., h$,

e) $E(X_{h,t} t_i) = 0, \forall t \in [0,T], \forall i = 1, ..., h$,

f)
$$t_h = \frac{\mathbf{W}Y_{h-1} - \dfrac{t_1}{E(t_1^2)}E(t_1 \mathbf{W}Y_{h-1}) - \ldots - \dfrac{t_{h-1}}{E(t_{h-1}^2)}E(t_{h-1}\mathbf{W}Y_{h-1})}{E(Y_{h-1}\mathbf{W}Y_{h-1})},$$

où $\mathbf{W}$ *est l'opérateur d'Escoufier associé au processus* (X_t).

Preuve. Les démonstrations des points a) - e) sont identiques à celles de la régression sur un vecteur de dimension finie (voir [13], par exemple). Le point f) : $\forall h \geqslant 1$,

$$t_h = \frac{\mathbf{W}_{h-1}Y_{h-1}}{\sqrt{Y_{h-1}\mathbf{W}_{h-1}Y_{h-1}}}.$$

$$\mathbf{W}_{h-1}Y_{h-1} = \int_0^T E(X_{h-1,t}Y_{h-1})X_{h-1,t}dt \tag{15}$$

Or, en utilisant la formule de décomposition de X_t (point c)) et l'orthogonalité des variables Y_{h-1} et t_i, $i = 1, ..., h-1$, on a :

$$E(X_{h-1,t}Y_{h-1}) = E(X_t Y_{h-1}), \ \ \forall t \in [0,T].$$

On obtient,

$$X_{h-1,t}E(X_t Y_{h-1}) = X_t E(X_t Y_{h-1}) - \sum_{i=1}^{h-1} t_i p_i(t) E(X_t Y_{h-1}),$$

soit, compte tenu de l'expression de $p_i(t)$ et de l'égalité $E(X_{i-1,t}t_i) = E(X_t t_i)$, $\forall i = 1, ..., h$ (ceci découlant de la définition de $X_{i-1,t}$ et de l'orthogonalité des $\{t_i\}_{i=1,...,h}$),

$$X_{h-1,t}E(X_t Y_{h-1}) = X_t E(X_t Y_{h-1}) - \sum_{i=1}^{h-1} t_i \frac{E(X_t t_i)E(X_t Y_{h-1})}{E(t_i^2)}.$$

L'utilisation de cette dernière égalité en (15) termine la démonstration du point f). □

Remarque 4. – Le point f) de la Proposition 3 nous permet d'affirmer que la régression PLS de la variable aléatoire Y sur le processus $(X_t)_{t\in[0,T]}$ est équivalente

à la régression PLS de Y sur l'ensemble de composantes principales de $(X_t)_{t\in[0,T]}$ dans le sens suivant : à chaque pas h, $h \geqslant 1$, les deux régressions nous donnent la même composante PLS, t_h, et donc les mêmes formules de décomposition de Y. En effet, car les composantes principales du processus $(X_t)_{t\in[0,T]}$ sont vecteurs propres de l'opérateur $\mathbf{W}$ ([12]).

Notons encore par $\hat{Y}_h$, respectivement $\hat{Y}$, les approximations données par la régression PLS de Y sur $(X_t)_{t\in[0,T]}$ au pas h, respectivement par la régression de Y sur $(X_t)_{t\in[0,T]}$,

$$\hat{Y}_h = c_1 t_1 + c_2 t_2 + \dots c_h t_h. \tag{16}$$

PROPOSITION 5

$$\lim_{h\to\infty} E(\hat{Y}_h - \hat{Y})^2 = 0 \tag{17}$$

Preuve. – Si le nombre de composantes PLS est fini, la preuve est analogue à celle où X est un vecteur fini ([14]).

Considérons que la suite $\{t_h\}_{h\geqslant 1}$ est infinie. La suite $\{E(Y_h^2\}_{h\geqslant 1}$ est convergente. En effet on a :

$$E(Y^2) \geqslant E(Y_h^2) \geqslant E(Y_{h+1}^2) \geqslant 0.$$

En utilisant $E(Y_h^2) - E(Y_{h+1}^2) = c_{h+1}^2 E(t_{h+1}^2)$ on déduit que

$$\lim_{h\to\infty} c_h^2 E(t_h^2) = 0.$$

$\forall \varepsilon > 0$, $\exists h_0 \geqslant 1$ tel que

$$c_h^2 E(t_h^2) < \frac{\varepsilon}{V}, \ \ \forall h \geqslant h_0,$$

où V est la variance totale du processus $(X_t)_{t\in[0,T]}$, $V = \displaystyle\int_0^T E(X_t^2)dt$. Compte tenu de l'inégalité $E(t_h^2) \leqslant V, \forall h \geqslant 1$, (la première composante principale de $(X_t)_{t\in[0,T]}$ expliquant le plus de la variance totale V) et de l'expression de c_h, on obtient :

$$E^2(Y_{h-1}t_h) = c_h^2 E^2(t_h^2) < \frac{\varepsilon}{V} E(t_h^2) \leqslant \varepsilon, \ \ \forall h \geqslant h_0, \Longrightarrow \lim_{h\to\infty} E^2(Y_{h-1}t_h) = 0.$$

Or, $E^2(Y_{h-1}t_h) = \displaystyle\int_0^T E^2(X_{h-1,t}Y_{h-1})dt = \int_0^T E^2(X_t Y_{h-1})dt$. Donc,

$$\lim_{h\to\infty} \int_0^T E^2(X_t Y_{h-1})dt = 0.$$

La suite $\{Y_h\}_{h\geqslant 1}$ est convergente en $L_2(\Omega)$ car

$$E(Y_{h+p} - Y_h)^2 = c_{h+1}^2 E(t_{h+1}^2) + \ldots + c_{h+p}^2 E(t_{h+p}^2), \ \forall p \geqslant 1.$$

Soit

$$\lim_{h\to\infty} Y_{h-1} = Z.$$

Le lemme de Fatou appliquée à $\lim_{h\to\infty} \int_0^T E^2(X_t Y_{h-1})dt = 0$, entraîne :

$$E(X_t Z) = 0, \ \forall t \in [0,T].$$

$Z \perp L_2(X)$ et d'après la formule de décomposition (Proposition 3) :

$$\lim_{h\to\infty} \hat{Y}_h = Y - Z.$$

Le théorème de la projection rend la conclusion immédiate. □

5. Régression PLS multidimensionnelle sur un processus

La régression linéaire multidimensionnelle classique d'un vecteur aléatoire réel $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_p)$, $p > 1$, sur un processus $\mathbf{X} = (X_t)_{t\in[0,T]}$ se réduit, grâce au théorème de la projection, à la réalisation de p régressions linéaires simples des variables Y_i, $i = 1, ..., p$, sur $\mathbf{X}$. On n'utilise donc pas les corrélations entre les variables $\{Y_i\}_{i=1,\ldots,p}$.

Dans cette section nous allons développer la régression PLS multidimensionnelle du vecteur aléatoire $\mathbf{Y}$ sur le processus $(X_t)_{t\in[0,T]}$ vérifiant les mêmes hypothèses que dans la Section 4. Les cas selon lesquels la dimension du vecteur $\mathbf{Y}$ est finie ou infinie seront traités séparément.

Le cas fini. – Soient $(X_t)_{t\in[0,T]}$, $X_t : \mathbf{\Omega} \to \mathbf{R}, \forall t \in [0,T]$, un processus stochastique du second ordre, L_2-continu et $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_p)$, $p > 1$, un vecteur aléatoire réel défini sur le même espace de probabilité. Supposons que le processus est centré, $E(X_t) = 0, \forall t \in [0,T]$ et $E(Y_i) = 0, \forall i = 1, ..., p$.

On définit les opérateurs suivants :

$\mathbf{C}_{YX} : L_2([0,T]) \to \mathbf{R}^p$,

$$f \xrightarrow{\mathbf{C}_{YX}} x, \ \ x_i = \int_0^T E(X_t Y_i) f(t)dt, \ \ \forall i = 1, ..., p,$$

$\mathbf{C}_{XY} : \mathbf{R}^p \to L_2([0,T])$,

$$x \xrightarrow{\mathbf{C}_{XY}} f, \ \ f(t) = \sum_{i=1}^{p} E(X_t Y_i) x_i, \ \ \forall t \in [0,T]$$

Notons par $\mathbf{U}_X = \mathbf{C}_{XY} \circ \mathbf{C}_{YX}$ et $\mathbf{U}_Y = \mathbf{C}_{YX} \circ \mathbf{C}_{XY}$. Les propriétés suivantes sont immédiates :

a) $\mathbf{C}^*_{YX} = \mathbf{C}_{XY}$, où $\mathbf{C}^*_{YX}$ est l'adjoint de $\mathbf{C}_{YX}$,

b) $\mathbf{U}_X$ et $\mathbf{U}_Y$ sont opérateurs autoadjoints positifs, compacts (de rang $\leqslant p$) et ont le même spectre.

Soient également $\mathbf{W}^X$, respectivement $\mathbf{W}^Y$, les opérateurs d'Escoufier de $L_2(\Omega)$ associés aux vecteurs $\mathbf{X} = (X_t)_{t\in[0,T]}$, respectivement $\mathbf{Y} = (Y_i)_{i=1,\ldots,p}$ et définis par :

$$\mathbf{W}^X Z = \int_0^T E(X_t Z) X_t dt, \qquad \forall Z \in L_2(\Omega),$$

$$\mathbf{W}^Y Z = \sum_{i=1}^p E(Y_i Z) Y_i, \qquad \forall Z \in L_2(\Omega).$$

La composante PLS recherchée par le critère de Tucker est donnée par la proposition suivante :

PROPOSITION 6. – *Soit $w \in L_2([0,T])$ et $c \in \mathbf{R}^p$. Alors,*

$$\max_{\substack{w,c \\ \|w\|=1 \\ \|c\|=1}} Cov^2 \left(\int_0^T X_t w(t) dt, \sum_{i=1}^p c_i Y_i \right)$$

est réalisé pour w, respectivement c, les vecteurs propres correspondants aux plus grandes valeurs propres des opérateurs $\mathbf{U}_X$, respectivement $\mathbf{U}_Y$.

Preuve.

$$Cov^2 \left(\int_0^T X_t w(t) dt, \sum_{i=1}^p c_i Y_i \right) = \left[\int_0^T (\mathbf{C}_{XY} c)(t) w(t) dt \right]^2 = (\mathbf{C}_{XY} c | w)^2_{L_2([0,T])}$$

Avec les multiplicateurs de Lagrange μ_1 et μ_2, le Lagrangien s'écrit :

$$s = (\mathbf{C}_{XY} c | w)^2_{L_2([0,T])} - \mu_1 \left((w|w)_{L_2([0,T])} - 1 \right) - \mu_2 \left((c|c)_{R^p} - 1 \right).$$

Les conditions d'extremum s'expriment par :

$$\frac{\partial s}{\partial \mu_1} = - (w|w)_{L_2([0,T])} + 1 = 0,$$

$$\frac{\partial s}{\partial \mu_2} = - (c|c)_{R^p} + 1 = 0,$$

$$\frac{\partial s}{\partial w} = 2 (\mathbf{C}_{XY} c | w)_{L_2([0,T])} \mathbf{C}_{XY} c - 2\mu_1 w = 0,$$

$$\frac{\partial s}{\partial c} = 2\left(\mathbf{C}_{XY}c|w\right)_{L_2([0,T])}\mathbf{C}_{YX}w - 2\mu_2 c = 0.$$

En utilisant

$$\left(\mathbf{C}_{XY}c|w\right)_{L_2([0,T])} = \left(\mathbf{C}_{YX}w|c\right)_{R^p}$$

on obtient :

$$\mu_1 = \mu_2 = \lambda = \left(\mathbf{C}_{XY}c|w\right)^2_{L_2([0,T])}.$$

D'où les relations :

$$\mathbf{U}_Y c = \lambda c, \qquad \mathbf{U}_X w = \lambda w.$$

Ainsi, le maximum est réalisé pour w, respectivement c, vecteur propre associé à la plus grande valeur propre de $\mathbf{U}_X$, respectivement de $\mathbf{U}_Y$. □

Soit λ_1 la plus grande valeur propre de l'opérateur $\mathbf{U}_X$ et $w_1 \in L_2([0,T])$ une fonction propre de $\mathbf{U}_X$ associée à λ_1. On définit la première composante PLS de la régression du vecteur $\mathbf{Y}$ sur le processus $(X_t)_{t\in[0,T]}$ par la variable aléatoire :

$$t_1 = \int_0^T X_t w_1(t)dt \tag{18}$$

THÉORÈME 7. – *Soient $\mathbf{W}^X$, respectivement $\mathbf{W}^Y$, les opérateurs d'Escoufier associés aux vecteurs $\mathbf{X} = (X_t)_{t\in[0,T]}$, respectivement $\mathbf{Y}$. Alors t_1 est vecteur propre de l'opérateur $\mathbf{W}^X \circ \mathbf{W}^Y$ correspondant à la plus grande valeur propre.*

Preuve.

$$\begin{aligned}\mathbf{W}^Y t_1 &= \sum_{i=1}^{p} E(Y_i t_1)Y_i = \sum_{i=1}^{p} Y_i \int_0^T E(Y_i X_t)w_1(t)dt \\ &= \sum_{i=1}^{p} Y_i(\mathbf{C}_{YX}w_1)_i\end{aligned}$$

$$\begin{aligned}\mathbf{W}^X\mathbf{W}^Y t_1 &= \int_0^T X_t E(X_t \sum_{i=1}^{p} Y_i(\mathbf{C}_{YX}w_1)_i)dt = \int_0^T X_t \sum_{i=1}^{p} E(X_t Y_i)(\mathbf{C}_{YX}w_1)_i dt \\ &= \int_0^T X_t(\mathbf{U}_X w_1)(t)dt = \int_0^T \lambda X_t w_1(t)dt = \lambda_1 t_1.\end{aligned}$$

Les opérateurs $\mathbf{W}^X\mathbf{W}^Y$ et $\mathbf{U}_X$ ont le même spectre. En effet, si t est vecteur propre de l'opérateur $\mathbf{W}^X\mathbf{W}^Y$ correspondant à la valeur propre λ, alors la fonction définie par :

$$w(t) = E(X_t \cdot \mathbf{W}^Y t), \quad \forall t \in [0,T],$$

est fonction propre de l'opérateur $\mathbf{U}_X$ correspondant à la même valeur propre, λ.

On en déduit que λ_1 est la plus grande valeur propre de $\mathbf{W}^X\mathbf{W}^Y$. □

Soit $X_{0,t} = X_t, \forall t \in [0,T]$ et $Y_{0,i} = Y_i, \forall i = 1, ..., p$. Au pas h de la régression PLS de $\mathbf{Y}$ sur $(X_t)_{t\in[0,T]}$, $h \geqslant 1$, on calcule la composante t_h comme étant le vecteur propre associé à la plus grande valeur propre de l'opérateur $\mathbf{W}^X_{h-1}\mathbf{W}^Y_{h-1}$,

$$\mathbf{W}^X_{h-1}\mathbf{W}^Y_{h-1}t_h = \lambda_h t_h, \tag{19}$$

où $\mathbf{W}^X_{h-1}$, respectivement $\mathbf{W}^Y_{h-1}$ sont les opérateurs d'Escoufier associés aux vecteurs $\mathbf{X} = (X_{h-1,t})_{t\in[0,T]}$, respectivement $\mathbf{Y}_{h-1} = (Y_{h-1,i})_{i=1,...,p}$. On calcule ensuite les résidus :

$$X_{h,t} = X_{h-1,t} - p_h(t)t_h, \quad t \in [0,T],$$

$$Y_{h,i} = Y_{h-1,i} - c_{h,i}t_h, \quad i = 1, ..., p,$$

où $p_h(t) = \dfrac{E(X_{h-1}t_h)}{E(t_h^2)}, \forall t \in [0,T]$ et $c_{h,i} = \dfrac{E(Y_{h-1,i}, t_h)}{E(t_h^2)}, \forall i = 1, ..., p.$

On a les formules de décomposition suivantes :

$$X_t = p_1(t)t_1 + \ldots + p_h(t)t_h + X_{h,t}, \quad t \in [0,T],$$

$$Y_i = c_{1,i}t_1 + \ldots + c_{h,i}t_h + Y_{h,i}, \qquad i = 1, ..., p.$$

L'approximation de $\mathbf{Y}$ donnée par la régression PLS sur $(X_t)_{t\in[0,T]}$ au pas h, $h \geqslant 1$, est donnée par :

$$\hat{\mathbf{Y}}_h = c_1t_1 + \ldots + c_ht_h, \ \ c_i \in \mathbf{R}^p, i = 1, ..., p. \tag{20}$$

Si $\hat{\mathbf{Y}}$ est l'approximation de $\mathbf{Y}$ donnée par la régression linéaire classique sur $(X_t)_{t\in[0,T]}$, on montre (la démonstration est identique à celle du cas univarié) la convergence en moyenne quadratique de $\hat{\mathbf{Y}}_h$ vers $\hat{\mathbf{Y}}$.

Le cas continu. – Les résultats précédents restent valables dans le cas où $\mathbf{Y} = (X_t)_{t\in[T,T+a]}$, $a > 0$. En effet, grâce à la L_2 continuité du processus $(X_t)_{t\in[0,T+a]}$, $\mathbf{C}_{X,Y}$ et $\mathbf{C}_{Y,X}$ sont compacts ce qui entraîne la compacité des opérateurs $\mathbf{U}_X$, respectivement $\mathbf{U}_Y$. Le Théorème 7 est donc valable et dans ce cas.

On obtient alors les formules de décomposition suivantes :

$$X_t = \begin{cases} t_1p_1(t) + \ldots t_hp_h(t) + X_{h,t}, & \forall t \in [0,T], \\ t_1c_1(t) + \ldots t_hc_h(t) + X_{h,t}, & \forall t \in [T, T+a], \end{cases} \tag{21}$$

où $p_h(t) = \dfrac{E(X_{h-1,t}t_h)}{E(t_h^2)}, \forall t \in [0,T]$ et $c_h(t) = \dfrac{E(X_{h-1,t}t_h)}{E(t_h^2)}, \forall t \in [T, T+a].$

Pour tout $s \in [0, a]$, la prévision de X_{T+s} à l'aide du passé, $(X_t)_{t\in[0,T]}$, est donnée par

$$\hat{X}_{T+s} = t_1 c_1(T+s) + \ldots + t_h c_h(T+s). \tag{22}$$

Les propriétés relatives à la convergence de la régression PLS vers la régression linéaire restent valables et dans ce cas.

Remarque 8 (Méthode pratique)

a) En pratique, les coefficients $c_i(T+s)$ donnés en (22) sont estimés à partir d'un échantillon aléatoire fini; les opérateurs d'Escoufier se réduisent dans ce cas à des matrices ([12]) dont l'analyse spectrale nous fournit les composantes PLS. La prévision ponctuelle ainsi obtenue peut être accompagnée des intervalles de confiance comme en ([14]).

b) S'il existent $\{t_i\}_{i=0,\ldots,p}$, $0 = t_0 < t_1 < t_2 < \ldots t_{p-1} < t_p = T$ tels que le processus $(X_t)_{t\in[0,T]}$ a presque toutes les trajectoires constantes sur chaque intervalle $[t_i, t_{i+1})$, $\forall i = 0, ..., p-1$, (c'est-à-dire, $X_t = c_i$, $\forall t \in [t_i, t_{i+1})$, $c_i \in L_2(\Omega)$, $\forall i = 0, ..., p-1$) alors la régression PLS de $\mathbf{Y}$ sur $(X_t)_{t\in[0,T]}$ est équivalente à la régression PLS de $\mathbf{Y}$ sur l'ensemble $\{c_i\sqrt{(t_{i+1}-t_i)}\}_{i=0,\ldots,p-1}$ (*cf.* ([2]), cela revient à la régression PLS sur $\{c_i\}_{i=0,\ldots,p-1}$ utilisant comme métrique dans $\mathbf{R}^p$ $M = diag(t_1 - t_0, \ldots, t_p - t_{p-1})$.

c) Si les conditions du point b) ne sont pas satisfaites, alors une façon simple d'approcher la régression PLS sur le processus $(X_t)_{t\in[0,T]}$ est de considérer un découpage $\Delta = (0 = t_0 < t_1 < t_2 < \ldots t_{p-1} < t_p = T)$ de l'intervalle $[0, T]$ en p sous-intervalles et de réaliser la régression PLS sur le processus [3] (X_t^Δ) défini par :

$$X_t^\Delta = \frac{1}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} X_t dt, \quad \forall t \in [t_i, t_{i+1}[, \quad \forall i = 0, ..., p-1. \tag{23}$$

On est alors dans les conditions du point b) de cette remarque. □

6. Application sur des données boursières

Dans cette section, la régression PLS sur un processus présentée dans les sections précédentes sera utilisée pour prédire le comportement des actions boursières sur une certaine période de temps. De telles données constituent un bon exemple de réalisation d'un processus stochastique à temps continu pour lequel les hypothèses d'existence des moments de second ordre et de continuité en moyenne quadratique sont tout à fait raisonnables.

[3] Cette approximation (approximation par moyennage) a été traitée en détail dans le cas de l'analyse en composantes principales d'un processus dans ([9]).

Nous disposons d'un ensemble de 84 actions cotées à la Bourse de Paris pour lesquelles on connaît complètement le comportement de l'indice de croissance [4] sur un intervalle d'une heure (entre 10^{00}h – l'heure d'ouverture – et 11^{00}h). On connaît également l'évolution de l'indice de croissance d'une nouvelle action (notée 85) sur l'intervalle 10^{00}h - 10^{55}h. Le but est de prédire le comportement de cette action sur l'intervalle de 5 minutes entre 10^{55}h - 11^{00}h utilisant un modèle PLS construit à l'aide des 84 actions dont l'évolution est entièrement connue sur l'intervalle 10^{00}h - 11^{00}h.

Une action est susceptible de changer toutes les secondes : nous allons donc considérer les actions comme étant des réalisations indépendantes d'un processus stochastique $\{X_t : t \in [0, 3600]\}$ (l'intervalle de temps est exprimé ici en secondes). Avec les notations introduites en 5.2, il s'agit de la régression PLS de $\{X_t : t \in [T, T + a]\}$ sur $\{X_t : t \in [0, T]\}$ avec $T = 3300$ et $a = 300$.

Afin de donner au problème une dimension raisonnable – d'après la Remarque 8-b), le nombre total de variables serait de 1366 (il y a 1366 intervalles dans [0,3600] où les 85 trajectoires de $(X_t)_{t\in[0,3600]}$ sont constantes) – nous allons utiliser l'approximation donnée en (23) et la Remarque 8-b) en prenant un découpage équidistant de l'intervalle $[0, 3600]$ en 60 sous-intervalles. Soit $\{m_i\}_{i=1,...,60}$ l'ensemble des variables définies par :

$$m_i = \frac{1}{60} \int_{60\cdot(i-1)}^{60\cdot i} X_t dt, \quad i = 1, ..., 60. \tag{24}$$

La Figure 2 présente l'évolution de l'action 85 dans l'intervalle $[0, 3300]$ avant et après l'approximation par moyennage donnée en (23). Les prévisions obtenues vont donc correspondre au niveau moyen de l'indice de croissance de l'action 85 considéré sur chaque intervalle de 60 secondes, $[60 \cdot (i - 1), 60 \cdot i)$, $i = 56, ..., 60$.

Utilisant le logiciel SIMCA-P ([15]) nous allons construire plusieurs modèles PLS correspondant au nombre de composantes choisies pour la régression. Ainsi, nous allons noter par PLS(k), $k \geqslant 1$, le modèle avec k composantes PLS. Les prévisions fournies par ces modèles seront comparées avec celles données par la régression sur les composantes principales (modèles notés avec CP(k)) et l'algorithme NIPALS (voir [13] pour plus de détails). Pour juger la qualité de ces modèles nous allons comparer les prévisions obtenues par chaque modèle, notées avec $\{\hat{m}_i(85)\}$, aux vrais valeurs $\{m_i(85)\}$, $i = 56, ..., 60$, observées ultérieurement.

Le Tableau 1 présente les taux de variance expliquée par les trois premières composantes PLS $\{t_1, t_2, t_3\}$, respectivement les trois premières composantes principales $\{\xi_1, \xi_2, \xi_3\}$, dans l'espace de variables $\{m_i\}_{i=1,...,55}$.

Les prévisions obtenues par ces modèles sont présentées dans le Tableau 2 et la Figure 3.

Considérant l'indice SSE comme mesure globale de la qualité d'un modèle, la régression PLS utilisant les 2 premières composantes $\{t_1, t_2\}$ donne les meilleures prévisions dans ce cas. Les modèles PLS(3) et CP(3) s'avèrent être particulièrement

---

[4] Au moment t, l'indice de croissance d'une action ω est défini par $X_t(\omega) = \dfrac{v(t) - v(0)}{v(0)}$, où $v(t)$ est la valeur de la cotation de l'action ω à l'instant t et $v(0)$ sa valeur de l'ouverture.

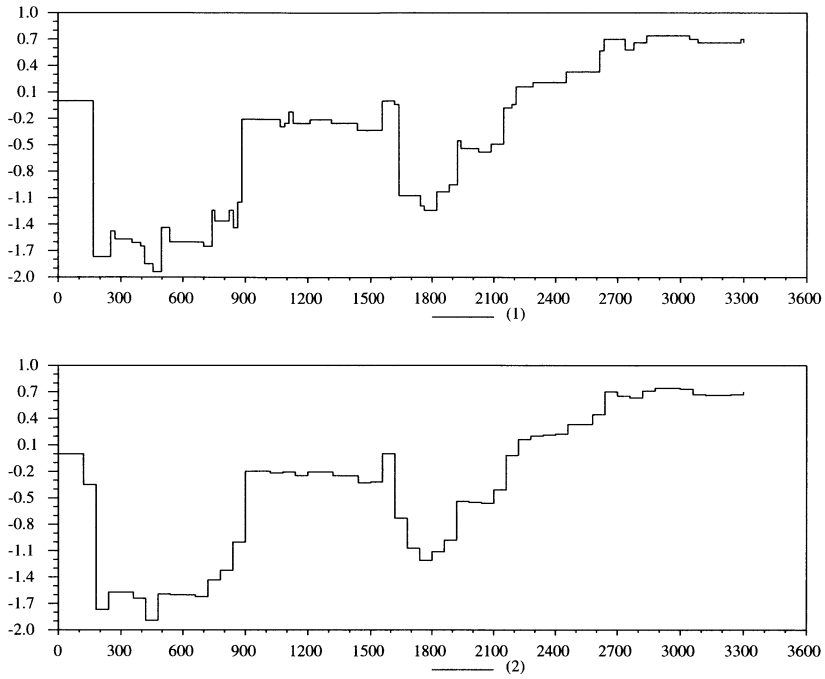

FIGURE 2

Évolution de l'action 85 : (1) – exacte, (2) – approximation

TABLEAU 1

Taux de variance expliquée

PLS	%	% cum.	ACP	%	% cum.
t_1	91.7	91.7	ξ_1	91.7	91.7
t_2	2.9	94.6	ξ_2	4.0	95.7
t_3	2.7	97.3	ξ_3	2.5	98.2

efficaces à un petit horizon (3300-3480) mais beaucoup moins performantes sur l'intervalle (3480-3600). L'algorithme NIPALS fournit également, en moyenne, des bonnes prévisions.

7. Conclusions

Nous avons développé dans cette article la régression PLS sur un processus L_2-continu. Le point clé de cette étude est l'exploitation des propriétés de l'opérateur

TABLEAU 2
Prévisions des différents modèles

	$\hat{m}_{56}(85)$	$\hat{m}_{57}(85)$	$\hat{m}_{58}(85)$	$\hat{m}_{59}(85)$	$\hat{m}_{60}(85)$	$SSE = \sum_{i=56}^{60} (\hat{m}_i - m_i)^2$
Vrai	**0.700**	**0.678**	**0.659**	**0.516**	**-0.233**	-
PLS(1)	-0.327	-0.335	-0.338	-0.325	-0.302	3.789
PLS(2)	0.312	0.355	0.377	0.456	0.534	0.928
PLS(3)	0.620	0.637	0.677	0.781	0.880	1.318
CP(1)	-0.356	-0.365	-0.368	-0.355	-0.331	4.026
CP(2)	-0.332	-0.333	-0.335	-0.332	-0.298	3.786
CP(3)	0.613	0.638	0.669	0.825	0.963	1.538
NIPALS	0.222	0.209	0.240	0.293	0.338	1.000

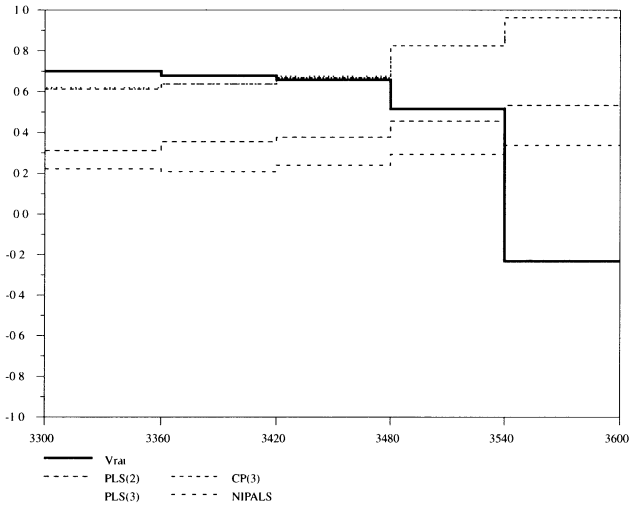

FIGURE 3
Prévisions pour l'action 85

d'Escoufier associé au processus, l'analogie avec l'analyse en composantes principales du processus étant évidente.

La régression PLS sur un processus offre une alternative à la régression sur les composantes principales. Elle donne une solution aux problèmes liés à la corrélation des prédicteurs et au cas où le nombre d'observations est inférieur au nombre de variables explicatives, comme il arrive souvent dans ce contexte.

Remerciements

Nous remercions le Groupe SBF de la Bourse de Paris qui nous a fourni gracieusement les données boursières que nous avons traitées dans cet article.

Bibliographie

[1] AGUILERA A.M., OCAÑA F., VALDERRAMA M.J. (1998), *An approximated principal component prediction model for continous-time stochastic process*, Applied Stochastic Models and Data Analysis, Vol. 13, p. 61-72.

[2] CAZES P. (1997), *Adaptation de la régression PLS au cas de la régression après Analyse des Correspondances Multiples*, Revue de Statistique Appliquée, XLIV (4), p. 35-60.

[3] DEVILLE J.C. (1974), *Méthodes statistiques et numériques de l'analyse harmonique*, Annales de l'INSEE, No. 15, p 3-101.

[4] DEVILLE J. C. (1978), *Analyse et prévision des séries chronologiques multiples non stationnaires*, Statistique et Analyse des Données, No. 3, p. 19-29.

[5] ESCOUFIER Y. (1970) *Echantillonnage dans une population de variables aléatoires réelles*, Publications de l'Institut de Statistique de l'Université de Paris, 19, Fasc. 4, p. 1-47.

[6] GREEN P.J., SILVERMAN B. W. (1994), *Nonparametric Regression and generalized linear models. A roughness penalty approach*, Monographs on statistic and applied probability, No. 58, Chapman & Hall.

[7] LEBART L., MORINEAU A., PIRON M. (1995), *Statistique exploratoire multidimensionnelle*, Dunod, Paris.

[8] PALM R., IEMMA A.F. (1995), *Quelques alternatives à la régression classique dans le cas de colinéarité*, Rev. Statistique Appliquée XLIII (2), p. 5-33.

[9] PREDA C. (1999), *Analyse factorielle d'un processus : problèmes d'approximation et de régression*, Thèse de doctorat de l'Université de Lille 1.

[10] RAMSAY J.O., DALZELL C.J. (1991), *Some tools for functional data analysis*, Journal of Royal Statistical Society (B), 53, No. 3, p. 539-572.

[11] RAMSAY J.O., SILVERMAN B.W. (1997), *Functional Data Analysis*, Springer Series in Statistics, Springer-Verlag, New York.

[12] SAPORTA G. (1981), *Méthodes exploratoires d'analyse de données temporelles*, Cahiers du B.U.R.O., No. 37-38, Université Pierre et Marie Curie, Paris.

[13] TENENHAUS M., GAUCHI J.P., MÉNARDO C. (1995), *Régression PLS et applications*, Revue de Statistique Appliquée, XLIII (1), p. 7-63.

[14] TENENHAUS M. (1998), *La régression PLS. Théorie et pratique*, Editions Technip, Paris.

[15] UMETRI A.B. (1996), *SIMCA-P for Windows, Graphical Software for Multivariate Process Modeling*, Umetri AB, Box 7960, S-90719 Umea, Sweden.