LES METHODES DE PREVISION PAR ANALYSE DE SERIES CHRONOLOGIQUES.

Anne BORSU

Introduction

Les méthodes de prévision sont élaborées sur l'idée fondamentale que les valeurs historiques des grandeurs à prévoir peuvent aider à dégager un certain comportement du phénomène étudié.

La prévision par analyse des séries chronologiques fait l'hypothèse que les données passées de la seule variable étudiée suffisent à découvrir la loi sous-jacente aux données.

Les autres méthodes quantitatives de prévision (telles que les modèles économétriques, d'input-output, la régression multiple) ainsi que les méthodes qualitatives ne font pas l'objet de cet article.

1. *Les techniques de lissage*

Ces techniques ont pour but de dégager la loi sous-jacente aux valeurs de la variable étudiée.

En lissant les valeurs historiques, on établit une distinction entre le contenu aléatoire de la variable et la loi de base supposée existante qui régit cette variable.

1.1. *Les moyennes mobiles*

Le succès de la méthode des moyennes mobiles est dû à l'extrême facilité de leur application.

Département de Méthodes Quantitatives, Faculté des Sciences économiques et sociales, Faculté Notre-Dame de la Paix, Namur

1.1.1. Les moyennes mobiles simples

Dans le cas d'une chronique constante, le modèle du processus est de la forme

$$Y_t = a + u_t \qquad (1.1)$$

où u_t est le terme aléatoire de moyenne nulle et de variance σ_u^2

a = constante.

Si on désigne par $\hat{Y}_t(\ell)$ la prévision faite à la fin de la période t pour la période $t + \ell$, on a :

$$\hat{Y}_t(\ell) = E(Y_{t+\ell}) = \hat{a} \quad \text{et en particulier} \qquad (1.2)$$

$$\hat{Y}_t(1) = \hat{a} = \hat{Y}_t(\ell)$$

La somme des carrés des écarts entre prévision et réalisation relatifs aux n données les plus récentes est

$$SS = \sum_{i=o}^{n-1} (Y_{t-i} - \hat{Y}_{t-i-1}(1))^2 = \sum_{i=o}^{n-1} (Y_{t-i} - \hat{a})^2 \qquad (1.3)$$

En minimisant cette somme, on obtient

$$\hat{Y}_t(1) = \hat{a} = \sum_{i=o}^{n-1} Y_{t-i} = \frac{1}{n} (Y_t + Y_{t-1} + \dots + Y_{t-n+1}) \qquad (1.4)$$

L'expression (1.4) est celle d'une moyenne mobile sur les n observations les plus récentes et représente la meilleure prévision pour les périodes futures.

On voit que le même poids $\frac{1}{n}$ est accordé aux n données antérieures et qu'on ne tient pas compte des observations antérieures à la période t-n+1. Il n'existe pas de méthode qui permette le choix optimal de n.

La méthode des moyennes mobiles simples ne peut être utile qu'à des situations à très court terme.

1.1.2. Les moyennes mobiles doubles

Cette méthode consiste à calculer les moyennes mobiles simples sur les données brutes et à recommencer le même processus sur la série des moyennes mobiles simples.

Cette procédure est adaptée aux cas où il existe une tendance linéaire dans la série chronologique.

Le modèle sous-jacent aux données est

$$Y_t = a + bt + u_t \tag{1.5}$$

a et b sont des constants.

On peut montrer que la prévision faite à la fin de la période t pour la période $t + \ell$ est

$$\hat{Y}_t(\ell) = a + b\ell$$

où $a = 2S_t - S_t^{(2)}$

$$b = \frac{2}{n-1}\,(S_t - S_t^{(2)})$$

$$S_t = \frac{1}{n}\sum_{i=o}^{n-1} Y_{t-i}$$

et $S_t^{(2)} = \sum_{i=o}^{n} S_{t-i}$

Les remarques faites à propos des moyennes mobiles simples et relatives au choix de n, à la pondération uniforme et à l'horizon de prévision s'appliquent aussi aux moyennes mobiles doubles.

1.2. Le lissage exponentiel {2} {7}

Les techniques de lissage n'accordent pas le même poids aux observations précédentes; celui-ci décroit exponentiellement dans le passé.

Partant de l'idée que les écarts prévisions-réalisations les plus récents ont une importance plus grande pour la prévision que les écarts plus anciens, Brown {2} a proposé de pondérer le carré des écarts avec un poids décroissant exponentiellement dans le temps.

Il s'agit donc de minimiser l'expression

$$WSS = \sum_{i=o}^{n} (1-\alpha)^{i} \left[Y_{t-i} - \hat{Y}_{t-i-1}(1)\right]^{2} \qquad o \leqslant \alpha \leqslant 1$$

$$= \sum_{i=o}^{n} \beta^{i} \left[Y_{t-i} - \hat{Y}_{t-i-1}(1)\right]^{2} \qquad \beta = 1 - \alpha \tag{1.2.1}$$

1.2.1. Le lissage exponentiel simple

Soit un processus constant de la forme $Y_t = a + u_t$

on a

$$\hat{Y}_t(\ell) = \hat{a} = \hat{Y}_t(1) \tag{1.2.2}$$

En annulant $\frac{\partial WSS}{\partial \hat{a}}$, il vient

$$\sum_i (1-\alpha)^{i} Y_{t-i} - \hat{a} \sum_i (1-\alpha)^{i} = o \tag{1.2.3}$$

Pour n suffisament grand

$\sum_i (1-\alpha)^{i} \approx \frac{1}{\alpha}$ et (1.2.3) devient

$\hat{a} = \alpha \sum_i (1-\alpha)^{i} Y_{t-i}$ et

$$\hat{Y}_t(1) = \alpha \sum_{i=o}^{n} \beta^{i} Y_{t-i} \quad \text{où} \quad \beta^{i} = (1-\alpha)^{i} \quad \text{ou} \tag{1.2.4}$$

$$\hat{Y}_t(1) = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + \alpha(1+\alpha)^{2} Y_{t-2} + \ldots\ldots$$

La forme générale utilisée pour le calcul d'une prévision et dérivée de (1.2.4) est

$$\hat{Y}_t(1) = \alpha\, y_t + (1-\alpha)\, \hat{Y}_{t-1}(1) \tag{1.2.5}$$

ce qui exprime que la nouvelle prévision est égale à une fraction α de la dernière observation plus une fraction $(1-\alpha)$ de l'ancienne prévision.

(1.2.5) peut encore s'écrire

$$\hat{Y}_t(1) = \hat{Y}_{t-1}(1) + \alpha(Y_t - \hat{Y}_{t-1}(1)) \tag{1.2.6}$$

La prévision pour la période t+1 est égale à la prévision faite à la période t-1 pour la période t plus α fois l'erreur de prévision (processus de correction par les erreurs). La faiblesse des techniques de lissage réside dans le fait qu'il n'existe pas de méthodologie générale pour le choix de la constante de lissage α. En général on conseille de prendre des valeurs de $\alpha < 0.3$ et on choisit après essais successifs le coefficient qui minimise les erreurs. Comme les moyennes mobiles simples et pour la même raison le lissage exponentiel simple ne permet de faire des prévisions qu'à un horizon très court.

1.2.2. Le lissage exponentiel double

Cette technique de lissage est plus adaptée aux processus présentant une tendance.

Soit un processus défini par un modèle linéaire

$$Y_t = a + bt + u_t \tag{1.2.6}$$

$$\hat{Y}_t(\ell) = \hat{a} + \hat{b}\ell \quad \text{où } \ell \text{ est l'horizon de la prévision} \tag{1.2.7}$$

En minimisant la somme pondérée des carrés des erreurs, on obtient

$$\hat{a} = 2\, S_t - S_t^{(2)} \tag{1.2.8}$$

$$\hat{b} = \frac{\alpha}{1-\alpha}\,(S_t - S_t^{(2)}) \tag{1.2.9}$$

où S_t est la valeur lissée de Y_t au temps t

$$S_t = \alpha Y_t + (1-\alpha) S_{t-1}$$

$S_t^{(2)}$ est la valeur lissée de la variable S_t

$$S_t^{(2)} = \alpha S_t + (1-\alpha)\ S_{t-1}^{(2)}$$

(1.2.7) devient donc

$$\hat{Y}_t(\ell) = 2S_t - S_t^{(2)} + \frac{\alpha}{1-\alpha} (S_t - S_t^{(2)})\ \ell \qquad (1.2.10)$$

Comme dans le cas du lissage exponentiel simple, le choix de la constante α nécessite plusieurs essais.

1.2.3 Le lissage exponentiel d'ordre supérieur

Les lissages d'ordre supérieur sont adaptés aux séries dont le processus sous-jacent est supposé être un polynôme d'ordre supérieur à 1.

1.3 Lissage direct ou adaptatif

Brown {2} a développé une méthode qui permet de mettre à jour les estimations des paramètres du modèle en tenant compte de l'erreur de prévision de la période courante.

On peut montrer que le lissage direct d'un modèle polynomial d'ordre k est équivalent au lissage multiple d'ordre k + 1. Un aperçu de cette méthode est donné au paragraphe 5.

1.4 Méthodes de lissage avec contrôle adaptatif

Ces méthodes ont pour but de modifier automatiquement la valeur de la constante de lissage α utilisée dans un lissage exponentiel. α est donc considéré comme un paramètre. L'adaptation du coefficient α permet au système de répondre à des changements brusques dans la série étudiée.

1.4.1 Cas d'une seule constante de lissage

Soit $Q(T)$ l'erreur lissée au temps T

$$Q(T) = \gamma\, e_1(T) + (1-\gamma)\,Q(T-1) \qquad o < \gamma < 1$$

Soit $\hat{\Delta}(T)$ l'erreur absolue lissée au temps T

$$\hat{\Delta}(T) = \gamma\,|e_1(T)| + (1-\gamma)\,|\hat{\Delta}(T-1)|$$

On définit le "tracking signal" par le rapport $\frac{Q(T)}{\hat{\Delta}(T)}$, valeur qui est comprise entre -1 et +1.

Trigg et Leach {11} proposent d'ajuster la valeur de α en posant

$$\alpha(T) = \left| \frac{Q(T)}{\hat{\Delta}(T)} \right|$$

La valeur de α augmente donc lorsque le processus est hors de contrôle et un plus grand poids est donné aux observations les plus récentes.

Chow {4} propose une autre méthode qui peut être étendue au cas de plusieurs constantes de lissage. Il calcule la valeur de la déviation moyenne absolue des erreurs de prévision pour trois valeurs de α

la valeur en cours α_o

une limite supérieure $\alpha_u = \alpha_o + \delta$

une limite inférieure $\alpha_\ell = \alpha_o - \delta$

δ = constante (=0.05)

A chaque période, la nouvelle valeur de α correspondra à la valeur α_i qui donne la moindre déviation moyenne absolue:

Cette procédure nécessite le calcul de trois prévisions à chaque période.

1.4.2 Cas de plusieurs constantes de lissage

Roberts et Reed {10} dans leur procédure appelée SAFT (Self Adaptive Forecasting Technique) appliquent la méthode d'analyse factorielle à deux niveaux, les paramètres de lissage étant les facteurs du plan d'expérimentation. Ce plan comprend $2^k + 1$ expériences. (k facteurs à 2 niveaux plus le point central). Les effets des facteurs sont mesurés par le carré moyen des erreurs de prévision, à chaque niveau. ($2^k + 1$ previsions).

La prévision retenue est celle évaluée au niveau central des paramètres. Ce niveau est ajusté si l'effet d'un des paramètres est significatif.

1.5 Le modèle de Winters {15}

Winters a appliqué la technique de lissage exponentiel simple au cas d'un modèle saisonnier. Les effets saisonnier et trend-cyclique (voir paragraphe 2 ci-après) sont supposés varier dans le temps.

Dans les formules ci-après, on a supposé un effet saisonnier multiplicatif et un effet trend-cyclique additif.

L'estimation de la valeur désaisonnalisée T_t de la variable Y_t au temps t est

$$T_t = \alpha \frac{Y'_t}{S_{t-L}} + (1-\alpha)\ (T_{t-1} + e_{t-1}) \qquad (1.5.1)$$

où α est une constante de lissage : $o < \alpha < 1$

Y_t est la valeur observée de la variable

S_{t-L} est le facteur saisonnier de la période t-L

L est la périodicité du facteur saisonnier

e_{t-1} est la variation de la tendance par unité de temps

On voit que cette formule pondère deux estimations de T_t. Cette valeur de T_t est alors utilisée pour estimer une nouvelle valeur de l'écart e.

$$e_t = \beta(T_t - T_{t-1}) + (1-\beta)e_{t-1} \quad \text{où} \quad o < \beta < 1 \qquad (1.5.2)$$

(moyenne pondérée de deux estimations)

De même, on estime la nouvelle valeur de S_t par

$$S_t = \gamma\,(\frac{Y_t}{T_t}) + (1-\gamma)S_{t-L} \qquad (1.5.3)$$

La prévision pour l'horizon t+h est donnée par

$$\hat{Y}_t(h) = (T_t + he_c) \cdot S_{t+h-L} \qquad \text{pour} \quad o < h \leqslant L \qquad (1.5.4)$$

Si $L < h \leqslant 2L$, on remplace S_{t+h-L} par S_{t+h-2L} ......

Les coefficients de pondération α, β et γ sont estimés empiriquement en retenant les valeurs qui minimisent l'erreur quadratique moyenne de prévision. Ce modèle est heuristique et est surtout adapté à des prévisions à court terme -Sa faiblesse réside dans la spécification à priori de la relation entre les composantes présumées de la variable.

2. *Décomposition des séries chronologiques*

Alors que les méthodes précédentes visaient à dégager la loi sous-jacente aux données, les méthodes de prévision par décomposition d'une série chronologique tentent d'identifier les composantes de cette loi qui sont

- le trend ou tendance généralement assimilé à une projection linéaire de la série (T_t)
- le facteur saisonnier qui se répète à intervalles tels que le mois, la semaine (S_t)
- le facteur cyclique supposé suivre une loi en forme d'ondulation qui se répète sur une période de temps plus longue (C_t)
- une composante aléatoire (I_t)

La forme mathématique qui représente une série est donc

$$Y_t = T_t \times C_t \times S_t \times I_t \qquad (2.1)$$

si les effets des facteurs sont supposés multiplicatits et

$$Y_t = T_t + C_t + S_t + I_t \qquad (2.2)$$

s'ils sont supposés additifs.

2.1 *Méthode classique*

La procédure suivie par les méthodes de décomposition comporte généralement les étapes suivantes (N = nombre d'années; les données sont supposées mensuelles).

a) Estimation du facteur saisonnier

On calcule les indices saisonniers en prenant le rapport moyen des données brutes au moyennes mobiles centrées sur 12 mois pour chaque mois.

On élimine ensuite la composante saisonnière des données brutes par le rapport (ou la différence)

$$\frac{TCSi}{S} = TCi$$

b) Estimation de la tendance

Le trend s'obtient en ajustant graphiquement une ligne droite à la série des valeurs des moyennes mobiles ou par régression simple sur les données brutes.

On élimine le trend par le rapport (ou la différence)

$$\frac{TCi}{T} = Ci$$

c) Détermination du cycle

Si la série des données est suffisamment étendue, il sera possible de dégager la composante cyclique des variations résiduelles par calcul d'une moyenne mobile pondérée sur la série Ci.

La prévision pour la période $t + \ell$ effectuée à la fin de la période t est donc

$$\hat{Y}_t(\ell) = \text{facteur saisonnier x tendance x (facteur cyclique)}.$$

Le facteur cyclique est l'effet le plus délicat à évaluer dans la pratique, surtout si on ne dispose pas d'un nombre suffisant de données.

2.2 *Méthode Census II* {5} {12}

Basée sur le même type de décomposition que les méthodes classiques, la méthode Census II offre plus de raffinements dans l'estimation des divers éléments de la série, tels que:

- les données mensuelles peuvent être corrigées en tenant compte du nombre de jours ouvrables

- les valeurs extrêmes des rapports saisonniers calculés une première fois sont modifiées ou éliminées

- la composante trend-cyclique est calculée par moyenne mobile de Spencer.

Le programme ordinateur fournit des éléments utiles à la prévision tels que:

- les pourcentages moyens de variation pour les lags 1 à 12 pour les séries originale, désaisonnalisée, irrégulière, la composante trend-cyclique ....

- la contribution relative de chaque composante à la variance de la série originale

- la prévision de l'indice saisonnier ...

Le système $F_o$ran {6} est basé sur la méthode Census II.

2.3 *La décomposition par régression* {9}

2.3.1 La méthode de régression multiple peut être appliquée au calcul de la tendance et des coefficients saisonniers en introduisant des variables fantômes pour ces derniers. Le modèle est alors

$$Y_t = a + \beta_t + \sum_i \gamma_i d_{it} + \varepsilon_t \quad \text{où } d_{it} \text{ sont des variables fantômes} \tag{2.3}$$

$$\gamma_{it} = 1 \quad \text{si l'observation t a été effectuée au mois i}$$

$$= o \quad \text{dans le cas contraire}$$

Afin d'éviter le problème de la multicollinéarité qui pourrait surgir dans le modèle (2.3) on pose

$$\sum_{h=1}^{i} \gamma_h = o \quad \text{où } i = 12 \text{ si les données sont mensuelles}$$

Il suffit donc d'introduire une variable auxiliaire en trop peu et d'estimer le coefficient de celle-ci par la valeur des autres coefficients.

2.3.2 Toute tendance linéaire ou non peut être recherchée par les méthodes de régression linéaire, non linéaire ou multiple.

3. *Le filtrage adaptatif* {14}

Comme les méthodes de moyennes mobiles et de lissage exponentiel, la méthode du filtrage adaptatif fonde ses prévisions sur une moyenne pondérée des observations précédentes. L'originalité de la méthode du filtrage adaptatif est de chercher à déterminer la meilleure pondération des données.

La prévision s'exprime par

$$\hat{Y}_t(1) = \sum_{i=t-N+1}^{t} w_i Y_i \qquad \text{avec } \sum_i w_i = 1 \qquad (3.1)$$

où $Y_t(1)$ est la prévision pour la période t+1

w_i est le poids accordé à la valeur observée Y_i

Y_i est la valeur observée à la période i

N est le nombre d'observations utilisées dans la pondération (= nombre de poids)

Le recherche de la pondération la meilleure se fait par itérations successives. Une itération d'exercise comporte un parcours de la série des données.

La formule de révision des poids au cours de ce parcours est

$$W' = W + 2k\, e\, Y \qquad (3.2)$$

où W' est le nouveau vecteur des poids

W est l'ancien vecteur des poids

k est une constante appelée constante d'apprentissage

Y est le vecteur des N observations

Au départ, il faut donc choisir la valeur initiale des poids, le nombre de poids N (ou nombre de termes intervenant dans la pondération) ainsi que la valeur de la constante d'apprentissage k.

Supposons qu'on applique cette méthode à une série de données mensuelles présentant une saisonnalité et qu'on prenne N = 12 poids.

On aura : $\hat{Y}_{12}(1) = \sum_{i=1}^{12} w_i Y_i$

L'erreur de cette prévision sera

$$e_{13} = Y_{13} - \hat{Y}_{12}(1)$$

on réajustera donc les poids par la formule (3.2)

$$w'_1 = w_1 + 2k\, e_{13} Y_1$$

$$w'_{12} = w_{12} + 2k\, e_{13} Y_{12}$$

Ces poids révisés (et réajustés pour que leur somme soit égale à 1) serviront à la prévision de Y_{14}.

$$\hat{Y}_{13}(1) = \sum_{i=2}^{13} w'_i Y_i$$

On calcule ensuite l'erreur e_{14} qui servira à la révision suivante des poids selon (3.2). Ce processus mené jusqu'à la fin de la série des données constitue la première itération. Cette itération est répétée jusqu'à ce que les adaptations successives ne donnent que des variations de poids faibles.

Il existe quelques règles de choix de N (qui correspond à l'ordre du polynôme qu'on doit ajuster) et de la constante k. (k décroit lorsque le contenu aléatoire de la série croît et lorsque N croît). Il est conseillé d'essayer plusieurs valeurs de k en se guidant par l'erreur quadratique moyenne et par le taux de changement de cette erreur après chaque itération. Quant aux poids initiaux, on peut utiliser comme valeur de départ les coefficients d'autocorrelation de la série chronologique. Cette technique a été peu utilisée jusqu'à présent.

4. *La méthode de Box-Jenkins* {8}

Cette technique n'implique pas le choix *a* priori d'une loi sous-jacente aux données.

La séquence des observations d'une variable est considérée comme une réalisation de variables aléatoires distribuées selon une fonction de probabilité jointe, supposée invariante dans le temps.

Box et Jenkins proposent trois types de modèles pouvant rendre compte de toutes les lois de comportement : auto-régressif, moyenne mobile et modèle mixte.

(i) Le modèle en moyenne mobile d'ordre q , désigné par MA (q) est de la forme

$$z_t = \mu + u_t - \Theta_1 u_{t-1} - \ldots\ldots - \Theta_q u_{t-q} \qquad (4.5)$$

où u_t est un terme d'erreur de moyenne nulle

$$(u_t = z_t - \hat{z}_{t-1}(1))$$

Θ_i sont les coefficients du processus. Le signe moins est introduit par convention

μ est une constante.

L'observation z_t est donc fonction des termes d'erreur.

La formule (4.5) s'écrit plus facilement si on introduit l'opérateur de retard B.

$$z = \mu + (1-\Theta_1 B - \Theta_2 B^2 - \ldots\ldots - \Theta_q B^q) u_t$$

où $B^i u_t = u_{t-i}$

(ii) Le modèle autorégressif d'ordre p désigné par AR(p) est de la forme

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots\ldots \phi_p z_{t-p} + \delta + u_t$$

$$(1-\phi_1 B - \ldots - \phi_p B^p)\, z_t = \delta + u_t$$

δ = constante

La valeur observée de z_t est fonction des valeurs antérieures prises par la variable.

(iii) Le modèle mixte, désigné par ARMA (p,q) où la variable z_t est fonction des observations antérieures et des termes d'erreur.

$$(1-\phi_1 B - \dots - \phi_p B^p) z_t = (1-\theta_1 B - \dots - \theta_q B^q) u_t + \delta \qquad (4.7)$$

Le modèle mixte adapté au cas d'une variable saisonnière s'écrit

$$(1 - \Gamma_1 B^s - \dots - \Gamma_P B^{sP})(1-\phi_1 B - \dots - \phi_p B^p)\ y_t =$$

$$(1 - \Delta_1 B^s - \dots - \Delta_Q B^{sQ})(1-\phi_1 B - \dots - \phi_q B^q)\ u_t + \delta \qquad (4.8)$$

où P est l'ordre du processus auto-régressif saisonnier
p est l'ordre du processus auto-régressif
Q est l'ordre du processus en moyenne mobile saisonnier
q est l'ordre du processus en moyenne mobile
s est le nombre de période du cycle (s = 12 pour des données mensuelles)

y_t est la variable rendue stationnaire par différentiations successives de degré d et différentiations saisonnières de degré D

$$y_t = (1-B^s)^D (1-B)^d z_t$$

L'application de la méthode de Box-Jenkins comprend essentiellement trois étapes : l'identification du modèle, l'estimation de ses paramètres et la prévision.

4.1 Identification du modèle

Les modèles présentés ci-dessus s'appliquant à des séries stationnaires, il s'agit tout d'abord de débarasser la série des données de toute tendance.

Cette stationnarité peut être atteinte par différenciation des données (ou de leur logarithme) : différentiation consécutive ou saisonnière. La série étant rendue stationnaire, il s'agit de reconnaître le type de modèle le plus adapté aux données. Ce choix est basé sur l'allure des corrélogrammes empiriques et graphiques d'autocorrélation partielle que l'on compare avec les corrélogrammes théoriques des différents types de processus.

Notons qu'une spécification incorrecte d'un modèle apparaîtra par l'examen des coefficients d'autocorrélation des résidus. (voir ci-après)

4.2 Estimation des paramètres

A ce stade, on "essaie" le modèle qui a semblé le plus adéquat au cours de la phase précédente. L'estimation des paramètres du modèle est faite par la méthode de Gauss Newton. Plusieurs tests permettent de juger de la bonne sélection du modèle, tels que:

- l'autocorrélation des résidus qui doit être nulle si le modèle est adéquat. Les coefficients d'autocorrélation significatifs servent de guide dans le choix d'un autre modèle

- tests en t de significativité des paramètres du modèle.

4.3 Prévision

La prévision pour ℓ périodes futures (en supposant que y_t représente la variable rendue stationnaire) est pour $\ell \leqslant q$.

$$\hat{y}_t(1) = \phi_1 y_t + \ldots + \phi_p\, y_{t-p+1} + \delta - \Theta_1 u_t - \ldots - \Theta_q u_{t-q+1}$$

$$y_t(2) \equiv \phi_1 \hat{y}_t(1) + \phi_2\, y_t + \ldots + \phi_p\, y_{t-p+2} + \delta - \Theta_2 u_t - \ldots - \Theta_q u_{t-q+2}$$
$$\vdots \qquad (4.9)$$

pour $\ell > q$

$$\hat{y}_t(\ell) = \phi_1\, \hat{y}_t(\ell-1) + \ldots + \phi_p\, \hat{y}_t\,(\ell-p) + \delta \qquad (4.10)$$

Les prévisions sont adaptées selon l'erreur de prévision précédente.

En effet, sachant que tout processus ARMA peut s'écrire sous la forme

$$z_t = y_t + \Psi_1\, u_{t-1} + \Psi_2\, u_{t-2} + \ldots \qquad (4.11)$$

où les Ψ_i sont fonction des coefficients ϕ_i et Θ_i

on peut écrire

$$\hat{z}_t(\ell) = \Psi_\ell u_t + \Psi_{\ell+1}\, u_{t-1} + \ldots$$

$$\hat{z}_{t-1}\,(\ell+1) = \Psi_{\ell+1}\, u_{t-1} + \Psi_{\ell+2}\, u_{t-2} + \ldots$$

d'où il vient que

$$\hat{z}_t(\ell) - \hat{z}_{t-1}(\ell+1) = \Psi_\ell u_t \quad \text{où} \quad u_t = \hat{z}_t - \hat{z}_{t-1}(1) \tag{4.12}$$

5. *Prévision par les séries de Fourier* {2} {7}

Ces méthodes tentent d'assimiler les observations aux valeurs prises par une fonction périodique. Cette fonction peut être décomposée en un certain nombre d'oscillations sinusoïdales de la forme

$$Y = a \sin \omega(t+\Theta) \quad \text{où} \tag{5.1}$$

a est l'amplitude

ω est la fréquence angulaire : $\omega = \frac{2\pi}{T}$

T est la période du cycle ($T = 12$ pour un cycle saisonnier et des données mensuelles)

Θ est la phase

(5.1) peut s'exprimer sous la forme

$$Y_t = a \cos \omega\Theta \,.\, \sin \omega t + a \;\; \sin \omega\Theta \,.\, \cos \omega t$$

$$= a_1 \;\; \sin \omega t + a_2 \;\; \cos \omega t \tag{5.2}$$

où $a_1 = a \cos \omega\Theta$

$a_2 = a \sin \omega\Theta$

Un modèle composé réunissant une tendance et une variation sinusoïdale peut s'écrire

$$Y_t = a_1 + a_2 t + a_3 \sin \omega t + a_4 \cos \omega t + a_5 \sin 2\omega t + a_6 \cos 2\omega t \;\ldots \tag{5.3}$$

Ce modèle comprend une harmonique du mouvement de base.

Les coefficients a_i sont estimés par la méthode des moindres carrés (régression linéaire multiple) ou, initialement, de manière empirique.

La théorie du lissage exponentiel direct permet une révision des estimations des coefficients a_i en tenant compte de l'erreur de prévision.

On peut montrer que, si on applique la méthode de Brown qui consiste à estimer les paramètres a_i en minimisant la somme pondérée des carrés des erreurs (voir formule 1.2.1)) on obtient les résultats suivants:

$$\underline{\hat{a}}(t) = \underline{L}' \ \underline{\hat{a}}(t-1) + \underline{h} \ e_1(t) \tag{5.4}$$

$\underline{\hat{a}}(t)$ est le vecteur k x 1 des coefficients au temps t

$\underline{L}$ est la matrice de transition k x k qui dépend du modèle choisi

$\underline{h}$ est un vecteur de lissage k x 1, fonction du coefficient β de (1.2.1)

$e_1(t)$ est l'erreur de prévision $e_1(t) = Y_t - \hat{Y}_{t-1}(1)$

k est le nombre de paramètres du modèle.

Les matrices de transition $\underline{L}$ et le vecteur $\underline{h}$ sont calculés pour plusieurs modèles du type (5.3) et plusieurs valeurs du coefficient de lissage β (voir (2)).

Les prévisions revisées en fonction de l'erreur s'expriment sous la forme

$$\hat{Y}_t(\ell) = \sum_{i=1}^{k} \hat{a}_i(t).\, f_i(\ell) \tag{5.5}$$

où $f_i(\ell)$ représente les différents fonctions du temps présentes dans le modèle.

Si on applique la formule (5.5) au modèle défini en (5.3), la prévision pour la période t + ℓ s'exprime par

$$\hat{Y}_t(\ell) = a_1(t) + a_2(t).\ell + a_3(t) \sin \frac{2\pi\ell}{T} + a_4(t) \cos \frac{2\pi\ell}{T}$$
$$+ a_5(t) \sin \frac{4\pi\ell}{T} + a_6(t) \cos \frac{4\pi\ell}{T} \tag{5.6}$$

Notons que cette méthode suppose qu'on connaisse a priori les périodes des cycles de base présents dans la série en se guidant par la nature du phénomène étudiée, par l'analyse spectrale, par l'examen du corrélogramme de la série.

6. *Conclusion*

Les techniques de lissage restent très largement utilisées dans la pratique, surtout pour les prévisions faites au niveau d'une entreprise. Leur application est facile; le lissage direct permet la révision des estimations des paramètres du modèle tandis que les techniques de contrôle de la prévision assurent l'adaptation des coefficients de lissage aux changements brusques du phénomène étudié.

La décomposition classique des séries chronologiques en composantes saisonnière, de tendance et cyclique est plus largement employée dans les séries de données économiques et notamment lorsqu'il s'agit de comparer des phénomènes économiques. L'application de ces méthodes démande une longue série de données historiques, ce qui n'est pas le cas pour le lissage exponentiel.

La méthode de Box-Jenkins a le grand mérite de ne pas imposer le choix a priori d'un modèle mais son application nécessite l'appel à un ordinateur. Son domaine potentiel d'application est cependant très large.

Bibliographie

(1) BROWN, Robert G.
"Decision Rules for Inventory Management"
Holt, Rinehart and Winston, Inc., 1967

(2) ———
"Smoothing, Forecasting and Prediction of Discrete Time series"
Prentice-Hall International, 1963

(3) CHAMBERS, John C., MULLICH, Satinder K., SMITH, Donald D.
"How to choose the right forecasting technique" in
Harvard Business Review, Vol. 65, N° 4, pp 45-74, July-August 1971

(4) CHOW, W.M., "Adaptive Control of the Exponential Smoothing Constant",
Journal of Industrial Engineering, vol. 16, N° 5, pp 314-317, 1965

(5) McLAUGHLIN, Robert L. "Time Series Forecasting", American Marketing Association, 1962

(6) McLAUGHLIN, Robert L., BOYLE, James J. "Short Term Forecasting", American Marketing Association, 1968

(7) MONTGOMERY, Douglas C., JOHNSON, L.A. "Forecasting and Time series Analysis", Mc Graw Hill, 1976

(8) NELSON, Charles R. "Applied Time Series Analysis for Managerial Forecasting", 1972

(9) PHLIPS, L., BLOMME, R. "Analyse chronologique" ed. Vander, Louvain, 1973

(10) ROBERTS, S.D. and REED, R. "The development of a Self Adoptive Forecasting Technique", AIIE Transactions, vol. 1, N° 4, pp 314-322, December 1969

(11) TRIGG, D.W. and LEACH, A.G. "Exponential Smoothing with an Adaptive Response Rate", in Operational Research Quarterly, vol. 18, N° 1, pp 53-59, March 1967

(12) "The X-11 Variant of the Census Method II. Seasonal Adjustment Program". US Department of Commerce. Technical Paper N° 15

(13) WHEELWRIGHT, Steven, C., MAKRIDAKIS, Spyros "Choix et valeur des méthodes de prévision", les éditions d'organisation, Paris, 1974

(14) ——— "An Examination of the Use of Adaptive Filtering in Forecasting" in Operational Research Quarterly, vol. 24, N° 1, 1973

(15) WINTERS, Peter, R. "Forecasting Sales by Exponentially Weighted Moving Averages", Management Science, vol. 6, N° 3, pp 324-342, April 1960.