

# TP 5 : Prédiction des processus ARMA

DESS d'Ingénierie Mathématique, Année 2002-2003

Module de Séries chronologiques

laurent.albera@fr.thalesgroup.com

## I. INTRODUCTION

Ce TP s'effectuera sous MATLAB. La première étape dans l'étude d'une série chronologique consiste à détecter la présence de facteurs *tendance* et *saisonnier*, et le cas échéant à les estimer pour pouvoir les retrancher aux données. On se trouve alors en possession de ce que l'on nomme *fluctuations résiduelles*, en d'autres mots, la partie aléatoire de la série chronologique. Or, bien que non déterministe, cette composante peut avoir une structure particulière qu'il est alors important d'identifier. Une solution consiste à modéliser cette partie aléatoire par un processus ARMA dont il faudra estimer les différents paramètres. Une fois cette seconde étape d'identification effectuée nous pouvons estimer avoir expliqué au mieux la série chronologique considérée, car alors seul le bruit blanc en entrée de l'ARMA reste inconnu. La phase explicative achevée, les données de la série chronologique peuvent alors être utilisées pour par exemple prédire soit des données futures, soit des données passées. Ce TP consistera donc à mettre en oeuvre la technique de prédiction d'un processus ARMA décrite en cours. L'estimation des paramètres d'un ARMA sera abordée dans un autre chapitre du cours.

## II. HYPOTHESES ET NOTATIONS

### A. Modèle

Rappelons l'expression d'un processus ARMA( $p,q$ ) noté  $\{X_i\}$  supposé centré (i.e. de moyenne nulle):

$$X_i - \phi_1 X_{i-1} - \dots - \phi_p X_{i-p} = Z_i + \theta_1 Z_{i-1} + \dots + \theta_q Z_{i-q} \quad (1)$$

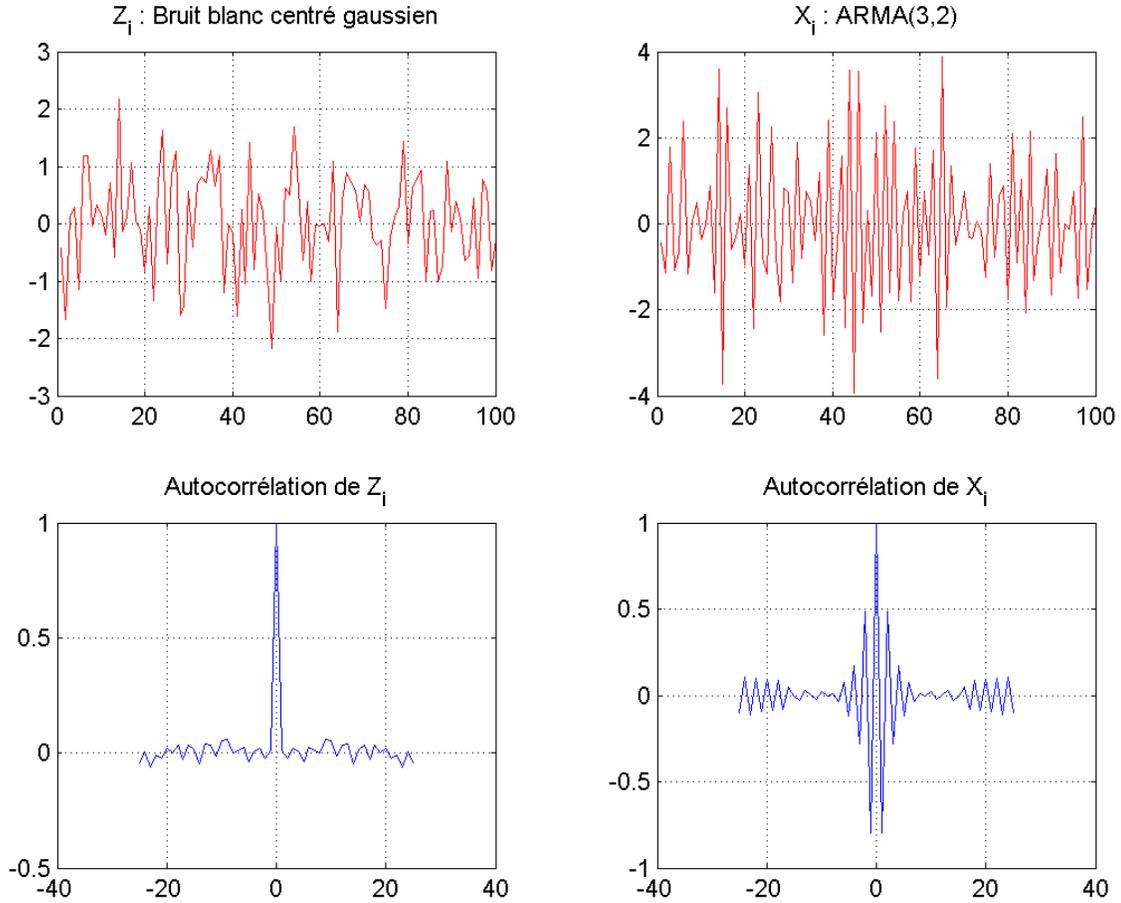
où  $\{Z_i\}$  est un processus aléatoire blanc centré de variance  $\sigma^2$  qui sera supposé gaussien dans la suite du TP, où  $\{\phi_k\}$ ,  $\{\theta_k\}$ ,  $p$  et  $q$  sont les paramètres du processus ARMA supposés connus (ou du moins estimés). Par ailleurs, un processus ARMA est dit *causal* s'il peut être écrit sous la forme d'un MA( $\infty$ ): les coefficients de ce processus MA seront notés  $\psi_k$  par la suite. Une autre expression équivalente à (1) est donnée à l'aide de l'opérateur retard  $B$ :

$$\phi(B)X_i = \theta(B)Z_i \quad (2)$$

où les polynômes  $\phi$  et  $\theta$ , de degrés respectifs  $p$  et  $q$ , ont pour coefficients les paramètres  $\{\phi_i\}$  et  $\{\theta_i\}$  cités dans (1) avec  $\phi_0 = \theta_0 = 1$ .

### B. Application

A titre d'exemple et d'application pour le TP, nous considérerons le processus ARMA(3,2)  $\{X_i\}$  défini par  $\phi(B) = -(B+3)(B+2)(B-7)/42$ ,  $\theta(B) = (B-5)(B-3)/15$  et  $\sigma^2 = 1$ . Tout d'abord, vérifier que le processus  $\{X_i\}$  est bien causal. Puis, sous MATLAB, en utilisant la fonction *filter*, générer le processus  $\{X_i\}$  à partir de 1000 échantillons ( $\{Z_i\}$  sera supposé gaussien). Représenter dans une seule fenêtre graphique, la courbe de  $\{Z_i\}$ , de  $\{X_i\}$  (en 100 points) et les courbes des autocorrélations respectives (en 50 points). Les fonctions d'autocorrélation seront calculées à partir des 1000 échantillons. Commenter alors les courbes (une tendance est-elle présente, un facteur saisonnier, peut-on différencier le processus ARMA du bruit blanc uniquement en observant les autocorrélations correspondantes?).



(a) Comparaison entre un bruit blanc et un ARMA

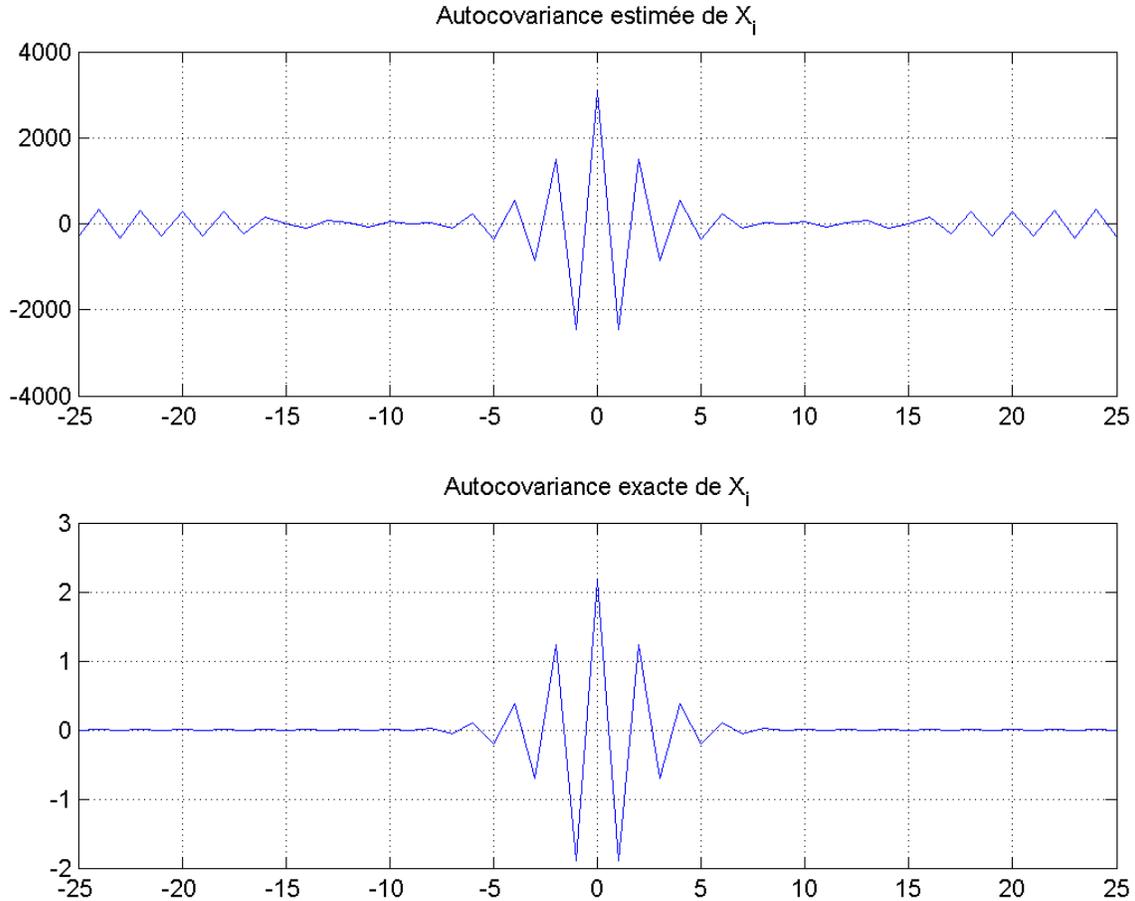
### III. CALCUL DE L'AUTOCOVARIANCE D'UN ARMA CAUSAL

La méthode utilisée dans ce TP est appelée méthode par *équation aux différences*.

#### A. Principe

La fonction d'autocovariance  $\gamma$  d'un processus ARMA( $p,q$ ) se calcule de manière récursive en utilisant l'équation suivante:

$$\sum_{i=0}^p \phi_i \gamma(k-i) = \begin{cases} \sigma^2 \sum_{j=k}^q \theta_j \psi_{j-k} & \text{si } 0 \leq k \leq q \\ 0 & \text{si } k > q \end{cases} \quad (3)$$



(b) Autocovariance d'un ARMA

Ce qui impose d'avoir préalablement calculé et ce de manière récursive les coefficients  $\psi_j$  du polynôme  $\psi = \theta/\phi$  pour  $0 \leq j \leq q$ . Pour cela, on peut utiliser les équations suivantes:

$$(\psi_j)_{0 \dots q} \leftarrow \begin{cases} \psi_0 = 1 \\ \psi_j - \sum_{k=1}^{\min\{p,j\}} \phi_k \psi_{j-k} = \theta_j & \text{si } 0 \leq j \leq q \\ \psi_j - \sum_{k=1}^{\min\{p,j\}} \phi_k \psi_{j-k} = 0 & \text{si } j > q \end{cases} \quad (4)$$

### B. Application

Calculer et représenter graphiquement la fonction d'autocovariance du processus  $\{X_i\}$  décrit ci-dessus, puis comparer-la avec celle estimée à l'aide de la fonction `xcov`.

#### IV. PRÉDICTION À UN PAS D'UN ARMA CAUSAL

##### A. Principe de l'algorithme des Innovations

La prédiction de la composante  $\widehat{X}_{n+1}$  à partir des  $n$  premières composantes  $X_i$  s'effectue de manière récursive en utilisant l'équation suivante:

$$\widehat{X}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{nj}(X_{n+1-j} - \widehat{X}_{n+1-j}) & \text{si } 1 \leq n < \max\{p, q\} \\ \sum_{j=1}^p \phi_j X_{n+1-j} + \sum_{j=1}^q \theta_{nj}(X_{n+1-j} - \widehat{X}_{n+1-j}) & \text{si } n \geq \max\{p, q\} \end{cases} \quad (5)$$

Les composantes  $\theta_{nj}$  se déduisent du système d'équations suivantes:

$$\begin{cases} v_0 = \kappa(1, 1) \\ \theta_{n, n-k} = v_k^{-1} \left( \kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k, k-j} \theta_{n, n-j} v_j \right) & \text{pour } 0 \leq k \leq n-1 \\ v_n = \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n, n-j}^2 v_j \end{cases} \quad (6)$$

où:

$$\kappa(i, j) = \begin{cases} \sigma^{-2} \gamma(i-j) & \text{si } 1 \leq i, j \leq \max\{p, q\} \\ \sigma^{-2} (\gamma(i-j) - \sum_{r=1}^p \phi_r \gamma(r - |i-j|)) & \text{si } \min\{i, j\} \leq \max\{p, q\} < \max\{i, j\} \leq 2\max\{p, q\} \\ \sum_{r=0}^q \theta_r \theta_{r+|i-j|} & \text{si } \min\{i, j\} > \max\{p, q\} \\ 0 & \text{sinon} \end{cases} \quad (7)$$

La fonction d'autocovariance  $\gamma$  du processus ARMA  $\{X_i\}$  doit donc être calculée avant d'engager la phase de prédiction.

##### B. Application

Séparer le processus ARMA généré dans la section (II) en deux processus. A partir du premier, nous allons prédire successivement les échantillons suivants. Le second processus va de ce fait nous permettre de mesurer la justesse de la prédiction. Nous réaliserons plusieurs fois cette expérience afin d'obtenir une quantité plus connue sous le nom d'*erreur quadratique moyenne*, celle-ci sera fonction de l'écart temporel entre l'instant de prédiction et l'instant de la dernière observation du processus.

#### V. CONCLUSION

Il est possible de prédire l'échantillon du processus à l'instant  $n+h$ , avec  $h > 1$  et où  $n$  est l'instant de la dernière observation du processus, sans avoir besoin de prédire les échantillons intermédiaires. Cette méthode présentée en cours repose également sur l'algorithme des Innovations.