

TP 9 : Représentation et estimation spectrale

DESS d'Ingénierie Mathématique, Année 2002-2003

Module de Séries chronologiques

laurent.albera@fr.thalesgroup.com

I. INTRODUCTION

Ce TP s'effectuera sous MATLAB. Il aura pour but de présenter le domaine *spectral* (ou fréquentiel) au même titre que son homologue, le domaine temporel. En effet, l'analyse de données requiert parfois le passage au domaine fréquentiel : c'est entre autres le cas lorsque l'on cherche à identifier un signal sinusoïdal noyé dans du bruit. Si la variance du signal d'intérêt est inférieure à celle du bruit, une visualisation des données dans le domaine temporel ne permet pas d'observer la sinusoïde. Par contre, dans le domaine fréquentiel, la conclusion est toute autre.

II. ESTIMATION NON PARAMÉTRIQUE DE LA DENSITÉ SPECTRALE

Un des outils majeurs d'analyse des signaux stationnaires dans le domaine spectral est la *densité spectrale*. Elle se définit comme étant la transformée de Fourier de la fonction d'autocovariance.

A. Périodogramme lissé

La densité spectrale d'un processus aléatoire stationnaire ergodique peut être estimée à l'aide du périodogramme $\hat{P}(e^{j\omega_k})$ défini par

$$\forall \omega_k = 2\pi k/N, \quad \hat{P}(e^{j\omega_k}) = \frac{1}{2\pi} \sum_{|h| < N} \hat{\gamma}(h) e^{-jh\omega_k} = \frac{1}{N} \left| \sum_{i=1}^N (x_i - \bar{x}) e^{-ji\omega_k} \right|^2 \quad (1)$$

où

$$\forall 0 \leq h \leq N-1, \quad \hat{\gamma}(h) = \frac{1}{N} \sum_{i=1}^{N-h} (x_i - \bar{x})(x_{i+h} - \bar{x}) \quad (2)$$

avec $\bar{x} = (\sum_{i=1}^N x_i)/N$. Bien que non biaisé, le périodogramme est un estimateur non consistant de la densité spectrale. Toutefois il peut le devenir en passant à travers un filtre $\{W_N(\cdot)\}$ dûment choisi

$$\forall \omega_k = 2\pi k/N, \quad \hat{f}(e^{j\omega_k}) = \sum_{|h| < m_N} W_N(h) \hat{P}(e^{j\omega_k+h}) \quad (3)$$

où m_N est un entier. L'estimateur $\hat{f}(e^{j\omega_k})$ de la densité spectrale peut également être directement défini à partir de l'estimée $\hat{\gamma}(h)$ de la fonction d'autocovariance

$$\forall \omega_k = 2\pi k/N, \quad \hat{f}(e^{j\omega_k}) = \frac{1}{2\pi} \sum_{|h| < \ell_N} w(h/\ell_N) \hat{\gamma}(h) e^{-jh\omega_k} \quad (4)$$

où les quantités $w(\cdot)$ et ℓ_N doivent vérifier

$$\begin{cases} w(0) = 1 \\ |w(x)| \leq 1 & \text{pour tout } x \\ w(x) = 0 & \text{pour tout } |x| > 1 \\ \ell_N/N \rightarrow 0 & \text{quand } N \rightarrow +\infty \end{cases} \quad (5)$$

B. Application

A titre d'application pour le TP, nous considérons le processus AR(2) $\{X_i\}$ défini par $\phi(B)X_i = \theta(B)Z_i$ où $\phi(B) = -(B+2)(B-7)/14$, $\theta(B) = 1$ et $\sigma^2 = 1$. Sous MATLAB, en utilisant la fonction *filter*, générer le processus $\{X_i\}$ à partir de $N = 500$ échantillons ($\{Z_i\}$ sera supposé gaussien). D'autre part, construire le processus $\{Y_i = X_i + \sin(0.3 * i * \pi)\}$. Représenter alors dans la même fenêtre MATLAB la densité spectrale estimée des deux processus $\{X_i\}$ et $\{Y_i\}$ (on utilisera alors le périodogramme lissé ainsi que la fonction *bartlett(.)* de MATLAB) et commenter les résultats.

III. ESTIMATION PARAMÉTRIQUE DE LA DENSITÉ SPECTRALE

A. Principe

On suppose que les données peuvent être modélisées par un ARMA de paramètres $(p, q, \sigma^2, \phi(B), \theta(B))$. Pour la suite, on se placera même dans le cas d'un AR de paramètres $(p, \sigma^2, \phi(B))$. La première étape consiste à estimer les paramètres de l'AR à partir des données et des méthodes temporelles usuelles d'estimation. L'ordre p du modèle étant inconnu, il rentre dans la liste des paramètres à estimer, il va de soi d'appliquer l'algorithme de Durbin-Levinson décrit en court. Rappelons les équations récursives employées ainsi que la procédure à suivre :

$$\begin{aligned}\hat{\phi}_{m,m} &= \left[\hat{\gamma}(m) - \sum_{i=1}^{m-1} \hat{\phi}_{m-1,i} \hat{\gamma}(m-i) \right] / \hat{v}_{m-1} \\ \begin{bmatrix} \hat{\phi}_{m,1} \\ \vdots \\ \hat{\phi}_{m,m-1} \end{bmatrix} &= \hat{\phi}_{m-1} - \hat{\phi}_{m,m} \begin{bmatrix} \hat{\phi}_{m-1,m-1} \\ \vdots \\ \hat{\phi}_{m-1,1} \end{bmatrix} \\ \hat{v}_m &= \hat{v}_{m-1} (1 - \hat{\phi}_{m,m}^2)\end{aligned}\quad (6)$$

où $\hat{\phi}_m = [\hat{\phi}_{m,1}, \dots, \hat{\phi}_{m,m}]^T$. Une estimation $\hat{p}$ de l'ordre p du modèle consiste à prendre la plus petite valeur m' telle que pour tout $m > m'$, $|\hat{\phi}_{m,m}| < 1.96N^{-1/2}$. Le vecteur $\hat{\phi}_{\hat{p}} = [\hat{\phi}_{\hat{p},1}, \dots, \hat{\phi}_{\hat{p},\hat{p}}]^T$ constitue donc un estimateur des coefficients du polynôme $\phi(B)$. Reste à estimer la variance σ^2 , il suffit alors de se référer à l'équation de Yule-Walker suivante :

$$\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}_{\hat{p}}^H \hat{\gamma}_{\hat{p}} \quad (7)$$

où $\hat{\gamma}_{\hat{p}} = [\hat{\gamma}(1), \dots, \hat{\gamma}(\hat{p})]^T$ et $\hat{\gamma}(0)$ la variance estimée des données.

Une fois les paramètres de l'AR estimés, on utilise le théorème de calcul de la densité spectral d'un processus ARMA $\{X_i\}$:

$$\forall -\pi \leq \lambda \leq \pi, \quad f_X(\lambda) = \frac{\sigma^2 |\theta(e^{-j\lambda})|^2}{2\pi |\phi(e^{-j\lambda})|^2} \quad (8)$$

où $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ et $\theta(z) = 1 - \theta_1 z - \dots - \theta_q z^q$. Néanmoins l'égalité (8) est soumise au respect de certaines conditions impliquant les deux polynômes ϕ et θ (rappeler ces conditions).

B. Application

Représenter à nouveau dans la même fenêtre MATLAB la densité spectrale estimée des deux processus $\{X_i\}$ et $\{Y_i\}$ (on utilisera cette fois l'approche paramétrique) et commenter les résultats. Comparer également avec les résultats obtenus dans la section (II-B).

IV. COMPLÉMENTARITÉ DE LA REPRÉSENTATION SPECTRALE

Reprendre l'exemple du processus $\{Y_i\}$, somme d'un AR(2) et d'une sinusoïde. Construire à présent le processus $\{U_i\}$ défini par $U_i - \alpha e^{j\xi} U_{i-1} = \alpha(Y_i - e^{j\xi} Y_{i-1})$ où $\alpha = 0.99$ et $\xi = 0.3$. Utiliser la fonction *filter* de MATLAB pour générer le processus $\{U_i\}$ à partir de $\{Y_i\}$. Représenter alors sous MATLAB les deux processus, tout d'abord dans le domaine temporel, puis dans le domaine fréquentiel. Commenter, que nous apporte la représentation spectrale? Quel est l'effet du filtrage de $\{Y_i\}$ pour obtenir $\{U_i\}$.

V. CONCLUSION

La représentation du signal dans le domaine spectral n'a d'intérêt que si le signal est stationnaire (au moins à l'ordre 2). Dans le cas contraire, la densité spectrale au sens où nous l'avons défini dépend du temps, comme c'est le cas pour les signaux de parole. Il devient alors nécessaire de rechercher une autre représentation du signal offrant les intérêts du domaine fréquentiel et tenant également compte de la non stationnarité du signal : une littérature conséquente répond à ce problème, entre autres celle traitant de la représentation *temps-fréquence*.