

Le choix optimal de la largeur de fenêtre spectrale pour un champ aléatoire

The optimal choice of the spectral bandwidth for a random field

par Mustapha RACHDI*, Rachid SABRE**

* Laboratoire AMS Université de Rouen, 76821 Mont Saint Aignan France,

E-mail : rachdi@univ-rouen.fr

** Laboratoire MAIS, ENESAD, 26 Docteur Petitjean 21000 Dijon France,

E-mail : r.sabre@enesad.fr

résumé et mots clés

Ce travail présente une procédure pour choisir la largeur de la fenêtre spectrale utilisée dans le lissage d'un périodogramme lors de l'estimation de la densité spectrale d'un champ aléatoire stationnaire. La procédure de validation croisée que nous proposons est basée, comme pour l'estimation de la densité de probabilité, sur l'estimation de l'erreur quadratique intégrée (ISE) en utilisant le principe du «Leave-out-I».

Densité spectrale, fenêtre spectrale, périodogramme, et validation croisée.

abstract and key words

In this paper, we give a criterion to choose the spectral bandwidth in the smoothing periodogram via spectral windows to estimate the spectral density for a stationary random field. As for the estimation of the probability density, the cross validation method proposed here is based on the integrated square error estimation using the principle of "leave-out-I".

Spectral density, spectral window, periodogram, and cross validation.

1. introduction

L'analyse de Fourier et l'estimation de la densité spectrale pour les champs aléatoires stationnaires ont une très longue histoire. Elles ont occupé une place importante dans le domaine de traitement du signal et traitement d'image. Elles ont prouvé leur utilité dans le filtrage linéaire et la théorie de prédiction en communication, en séismologie (où l'on cherche à déterminer la nature d'un événement sismique par l'estimation de la densité spectrale), en océanographie concernant la nature des ondes océaniques, en physique concernant les périodes des vibrations fondamentales de la terre, et dans divers domaines des sciences médicales.

Le but de ce travail est de présenter un critère de choix du paramètre de lissage utilisé lors de l'estimation de la densité spectrale d'un processus bidimensionnel (champ aléatoire). Nous considérons un champ aléatoire à temps discret, $X = \{X(n_1, n_2) :$

$n_1, n_2 \in Z\}$, centré et stationnaire c'est-à-dire : $E[X(n_1, n_2)] = 0$ $\forall n_1, n_2 \in Z$ et la covariance $\text{Cov}(X(n_1, n_2), X(n'_1, n'_2))$ est une fonction de $(n'_1 - n_1)$ et de $(n'_2 - n_2)$. La fonction autocorrélation, $R(s_1, s_2) = E[X^*(n_1, n_2)X(n_1 + s_1, n_2 + s_2)]$, admet, de la même manière que dans le cas unidimensionnel, une représentation spectrale : $R(s_1, s_2) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{is_1 u + is_2 v} dF(u, v)$.

La densité spectrale du champs aléatoire X , si elle existe, est définie par $f(u_1, u_2) = \frac{\partial^2 F(u_1, u_2)}{\partial u_1 \partial u_2}$. Priestley (1981) [7] a étudié l'estimation de la densité spectrale f en construisant le périodogramme suivant :

$$I_{N_1, N_2}(u_1, u_2) = \frac{1}{(2\pi)^2 N_1 N_2} \left| \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} X(n_1, n_2) e^{-iu_1 n_1 - iu_2 n_2} \right|^2. \quad (1)$$

Ce périodogramme est un estimateur asymptotiquement sans biais de $f(u_1, u_2)$, mais n'est pas consistant. Pour obtenir un estimateur consistant, nous lissons le périodogramme par deux fenêtres spectrales :

$$f_{N_1, N_2}(u_1, u_2) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} W_{N_1}^{(1)}(u_1 - x_1) W_{N_2}^{(2)}(u_2 - x_2) I_{N_1, N_2}(x_1, x_2) dx_1 dx_2$$

Les fenêtres spectrales $W_{N_1}^{(1)}$ et $W_{N_2}^{(2)}$ sont définies comme suit :

$W_{N_1}^{(1)}(x) = M_{N_1}^{(1)} W(x M_{N_1}^{(1)})$ et $W_{N_2}^{(2)}(x) = M_{N_2}^{(2)} W(x M_{N_2}^{(2)})$ où W est une fonction continue positive paire nulle en dehors de l'intervalle $[-1, 1]$ et telle que $\int_{-1}^1 W(x) dx = 1$, $M_{N_i}^{(i)} \rightarrow +\infty$ et $\frac{M_{N_i}^{(i)}}{N_i} \rightarrow 0$, $i = 1, 2$. Nous montrons que $f_{N_1, N_2}(u_1, u_2)$ est un estimateur asymptotiquement sans biais de $f(u_1, u_2)$ et consistant :

$$\text{Var} [f_{N_1, N_2}(u_1, u_2)] = O\left(\frac{M_{N_1}^{(1)} M_{N_2}^{(2)}}{N_1 N_2}\right). \quad (2)$$

Il est clair que le choix de $M_{N_1}^{(1)}$ et $M_{N_2}^{(2)}$ joue un rôle important puisque les vitesses de convergence dépendent de ces deux paramètres de lissage. Beltrao et Bloomfield [1] puis Hurvich et Beltrao [5] ont donné un critère de choix de h dans le cas unidimensionnel, ils se sont restreints au cas paramétrique. L'objectif de ce travail est de donner un critère de sélection de ces paramètres par des méthodes non paramétriques. Notons par $h_1 = \frac{1}{M_{N_1}^{(1)}}$ et

$h_2 = \frac{1}{M_{N_2}^{(2)}}$ les largeurs des deux fenêtres spectrales. Nous cherchons donc un critère $CV(h_1, h_2)$ nous permettant de sélectionner h_1 et h_2 minimisant l'erreur moyenne quadratique intégrée « mean integrated square error » ($MISE$), où

$$MISE(h_1, h_2) = \int \int E [f_{N_1, N_2}(x_1, x_2) - f(x_1, x_2)]^2 \rho(x_1, x_2) dx_1 dx_2, \quad (3)$$

ρ étant une fonction de poids que l'on suppose connue et nulle en dehors de $[0, 2\pi] \times [0, 2\pi]$.

Bien que $MISE(h_1, h_2)$ soit une bonne mesure de la qualité de f_{N_1, N_2} , elle ne peut pas nous aider pour choisir h_1 et h_2 , puisqu'elle dépend de la fonction inconnue f . Nous allons donc chercher à l'estimer. Pour cela, nous adoptons la méthode de la validation croisée qui a été proposée par Rudemo (1982) [10], et Bowman (1984) [2]. En effet, considérons l'erreur quadratique intégrée

« integrated square error » (ISE) définie par :

$$ISE(h_1, h_2) = \int \int [f_{N_1, N_2}(x_1, x_2) - f(x_1, x_2)]^2 \rho(x_1, x_2) dx_1 dx_2 = A - 2C + B$$

où $A = \int_0^{2\pi} \int_0^{2\pi} f_{N_1, N_2}^2(x_1, x_2) \rho(x_1, x_2) dx_1 dx_2$
 $C = \int_0^{2\pi} \int_0^{2\pi} f_{N_1, N_2}(x_1, x_2) f(x_1, x_2) \rho(x_1, x_2) dx_1 dx_2$
 $B = \int_0^{2\pi} \int_0^{2\pi} f^2(x_1, x_2) \rho(x_1, x_2) dx_1 dx_2.$

Puisque B est indépendant de h_1 et de h_2 , choisir h_1, h_2 minimisant $ISE(h_1, h_2)$ revient à choisir h_1, h_2 minimisant $A - 2C$. Le terme A est calculable puisque nous connaissons entièrement f_{N_1, N_2} , alors que, dans le terme C , il apparaît f qui est inconnue. Pour cela, nous procédons par le principe de « leave-out- I ». Auparavant, nous formulons deux hypothèses H_1 et H_2 que nous utiliserons par la suite :

$$X \text{ est strictement stationnaire et } \sum_{s_1} \sum_{s_2} s_1^2 s_2^2 |R(s_1, s_2)| < \infty \quad (H_1)$$

$$X \text{ admet des cumulants, qui sont bornés, jusqu'à l'ordre 4.} \quad (H_2)$$

2. construction de l'estimateur validé croisé

Dans cette section nous allons définir l'estimateur et donner quelques résultats sous forme de proposition ou théorème dont la démonstration se trouve en annexe dans la dernière section.

Soient $j, j' \in \{0, 1, \dots, n-1\}$. La construction de « leave-out- I » consiste à trouver un estimateur $f_{N_1, N_2}^{j, j'}(\omega_j, \omega_{j'})$ qui remplacera $f_{N_1, N_2}(\omega_j, \omega_{j'})$ dans l'expression de C et tel que $I_{N_1, N_2}(\omega_j, \omega_{j'})$ et $f_{N_1, N_2}^{j, j'}(\omega_j, \omega_{j'})$ soient asymptotiquement indépendants. Ainsi, on peut estimer C par :

$$\frac{1}{\bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} f_{N_1, N_2}^{j, j'}(\omega_j, \omega_{j'}) I_{N_1, N_2}(\omega_j, \omega_{j'}) \rho(\omega_j, \omega_{j'})$$

où

$$\omega_j = \frac{2\pi j}{N_1}, \omega_{j'} = \frac{2\pi j'}{N_2}, \bar{N}_1 = \left\lfloor \frac{N_1 - 1}{2} \right\rfloor, \bar{N}_2 = \left\lfloor \frac{N_2 - 1}{2} \right\rfloor,$$

$$f_{N_1, N_2}^{j, j'}(x_1, x_2) = \int_0^{2\pi} \int_0^{2\pi} I_{N_1, N_2}^{j, j'}(u_1, u_2) W_{N_1}^{(1)}(x_1 - u_1) W_{N_2}^{(2)}(x_2 - u_2) du_1 du_2$$

avec

$$I_{N_1, N_2}^{j, j'}(u_1, u_2) = I_{N_1, N_2}(u_1, u_2) \quad \text{si } (u_1, u_2) \notin A_{j, j'}$$

$$I_{N_1, N_2}^{j, j'}(u_1, u_2) = \theta_1(u_1, u_2) I_{N_1, N_2}(w_{j-1}, w_{j'-1}) + \theta_2(u_1, u_2) I_{N_1, N_2}(w_{j+1}, w_{j'-1}) + \theta_3(u_1, u_2) I_{N_1, N_2}(w_{j-1}, w_{j'+1}) + \theta_4(u_1, u_2) I_{N_1, N_2}(w_{j+1}, w_{j'+1}) \quad \text{sinon}$$

$A_{j, j'} =]w_{j-1}, w_{j+1}[\times]w_{j'-1}, w_{j'+1}[$. La construction de $I_{N_1, N_2}^{j, j'}(u_1, u_2)$ lorsque $(u_1, u_2) \in A_{j, j'}$ est faite comme si I_{N_1, N_2} était bilinéaire. Dans ce cas $\theta_1(u_1, u_2) = \alpha\beta$; $\theta_2(u_1, u_2) = (1-\alpha)\beta$; $\theta_3(u_1, u_2) = \alpha(1-\beta)$ et $\theta_4(u_1, u_2) = (1-\alpha)(1-\beta)$ où $\alpha = \frac{u_1 - w_{j+1}}{w_{j-1} - w_{j+1}}$ et $\beta = \frac{u_2 - w_{j'+1}}{w_{j'-1} - w_{j'+1}}$.

La proposition suivante, nous permet de montrer que $f_{N_1, N_2}^{j, j'}$ est un estimateur asymptotiquement sans biais de f .

Proposition 2.1

Sous les hypothèses H_1 et H_2 nous obtenons

$$E \left[f_{N_1, N_2}^{j, j'}(x_1, x_2) - f_{N_1, N_2}(x_1, x_2) \right] = O \left(\frac{1}{N_1 N_2} \right).$$

Suite à ce résultat, nous établissons notre critère, noté CV « cross validation », défini par :

$$CV(h_1, h_2) = CV_1(h_1, h_2) + \int_0^{2\pi} \int_0^{2\pi} f^2(u_1, u_2) \rho(u_1, u_2) du_1 du_2$$

$$\text{où } CV_1(h_1, h_2) = \int_0^{2\pi} \int_0^{2\pi} f_{N_1, N_2}^2(u_1, u_2) \rho(u_1, u_2) du_1 du_2 - \frac{2}{N_1 N_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} f_{N_1, N_2}^{j, j'}(w_j, w_{j'}) I_{N_1, N_2}(w_j, w_{j'}) \rho(w_j, w_{j'})$$

Les largeurs de fenêtres spectrales seront choisies aux points (\hat{h}_1, \hat{h}_2) minimisant le critère $CV(h_1, h_2)$:

$$(\hat{h}_1, \hat{h}_2) = \arg \min_{(h_1, h_2)} CV(h_1, h_2) = \arg \min_{(h_1, h_2)} CV_1(h_1, h_2) \quad (4)$$

Par la suite, pour faciliter l'écriture et sans perdre de généralité, nous prendrons $\rho(u_1, u_2) = \frac{1}{2\pi}$ sur $[0, 2\pi]^2$ et nulle ailleurs.

2.1. résultats de consistance et d'optimalité du critère

Dans cette section, nous établissons des résultats semblables à ceux donnés par G. Grégoire (1993) [3], concernant l'estimation des intensités d'un processus ponctuel. Il s'agit de montrer qu'en moyenne, quand N_1 et N_2 sont assez grands, le critère $CV(h_1, h_2)$ est approximativement égal à l'erreur quadratique intégrée $ISE(h_1, h_2)$ et que la variance de $CV(h_1, h_2)$ est asymptotiquement nulle. Cela nous permet de confirmer que les paramètres (\hat{h}_1, \hat{h}_2) minimisant le critère $CV(h_1, h_2)$ sont proches de ceux qui minimisent l'erreur quadratique intégrée (ISE) lorsque N_1 et N_2 sont assez grands. Ces résultats sont énoncés dans le théorème suivant :

Théorème 2.1

Sous les hypothèses H_1 et H_2 , nous avons

$$|E\{CV(h_1, h_2) - ISE(h_1, h_2)\}| = O \left(\frac{1}{N_1 N_2} \right).$$

$$\text{var}\{CV(h_1, h_2)\} = O \left(\frac{1}{N_1 N_2 h_1 h_2} \right)$$

Il s'ensuit aussitôt que :

$$E \left\{ [CV(h_1, h_2) - MISE(h_1, h_2)]^2 \right\} = \text{var}\{CV(h_1, h_2)\} + [E\{CV(h_1, h_2) - MISE(h_1, h_2)\}]^2 = O \left(\frac{1}{N_1 N_2 h_1 h_2} \right).$$

Les largeurs des fenêtres spectrales \hat{h}_1 et \hat{h}_2 obtenues par validation croisée, définies par (4), sont asymptotiquement optimales, c'est-à-dire que l'erreur quadratique intégrée aux points (\hat{h}_1, \hat{h}_2) converge en probabilité vers la plus petite erreur quadratique intégrée qui puisse exister :

Théorème 2.2

Sous les hypothèses du théorème 2.1, les largeurs des fenêtres spectrales \hat{h}_1, \hat{h}_2 obtenues par validation croisée sont asymptotiquement optimales :

$$\frac{ISE(\hat{h}_1, \hat{h}_2)}{ISE(\hat{h}_1, \hat{h}_2)} \rightarrow 1 \text{ en probabilité, où}$$

$$(\hat{h}_1, \hat{h}_2) = \arg \min_{h_1, h_2} CV(h_1, h_2) \text{ et}$$

$$(\hat{h}_1, \hat{h}_2) = \arg \min_{h_1, h_2} ISE(h_1, h_2).$$

Pour montrer ce résultat, nous adoptons les mêmes techniques de calcul utilisées dans le cas unidimensionnel par Sarda, Vieu (1989) [13] et Rachdi (1998) [8].

3. illustration de CV sur des données simulées

Nous étudions, sur des données simulées, la performance de la méthode de validation croisée appliquée à l'estimateur à noyau de la densité spectrale pour un processus bidimensionnel (champ aléatoire) stationnaire au sens strict. Nous reprenons la méthode utilisée dans [6] pour simuler un processus bidimensionnel de Gauss Markov centré dont la fonction de covariance $R(n_1, n_2)$ et la densité spectrale $f(x_1, x_2)$ sont définies par $R(n_1, n_2) = e^{-\sqrt{n_1^2 + n_2^2}}$ et $f(x_1, x_2) = \frac{1}{\pi(1 + x_1^2 + x_2^2)}$. Nous commençons par tracer la courbe de la fonction f (voir figure 1). L'estimation de la densité spectrale, f , est faite dans un premier temps par un lissage où les paramètres, (h_1, h_2) , sont pris au hasard (voir figure 2). En comparant les figures 1 et 2, nous constatons que le lissage de la courbe de l'estimateur n'est pas satisfaisant. Pour utiliser la méthode étudiée ci-dessus permettant de

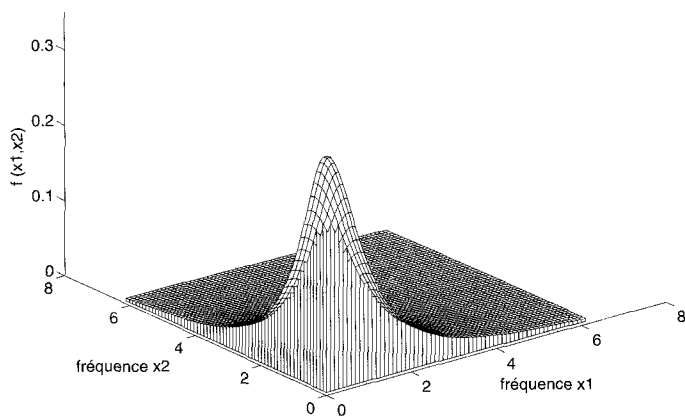


Figure 1. – La densité spectrale f .

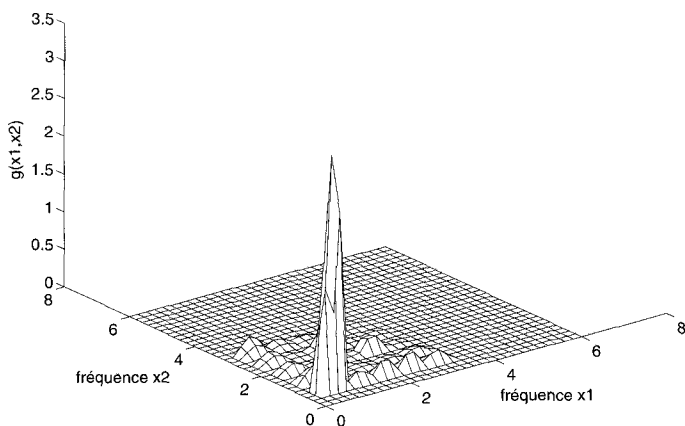


Figure 2. – L'estimateur lissé, $g = f_{N_1, N_2}$, de la densité f .

choisir les paramètres de lissage, (\hat{h}_1, \hat{h}_2) , minimisant le critère $CV_1(h_1, h_2)$, nous traçons d'abord ce critère (voir figure 3) puis nous déduisons (\hat{h}_1, \hat{h}_2) comme étant les arguments du minimum de $CV_1(h_1, h_2)$. Ensuite, nous construisons notre estimateur en utilisant les paramètres (\hat{h}_1, \hat{h}_2) , nous traçons sa courbe sur la figure 4. La similitude entre les figures 1 et 4 nous confirme la qualité des résultats obtenus grâce au choix optimal des paramètres de lissage par la méthode de validation croisée.

3.1. algorithme de calcul de CV_1 et résultats des simulations

La méthode de programmation de CV_1 suit les étapes suivantes :

Etape 1. – Se donner $(N_1 N_2)$ observations du processus : $\{X(n_1, n_2), n_1 = 1, \dots, N_1, n_2 = 1, \dots, N_2\}$.

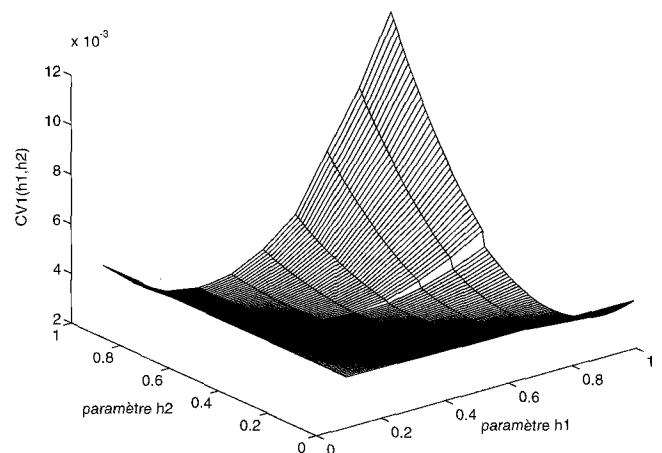


Figure 3. – Le critère de Validation-Croisée CV_1 .

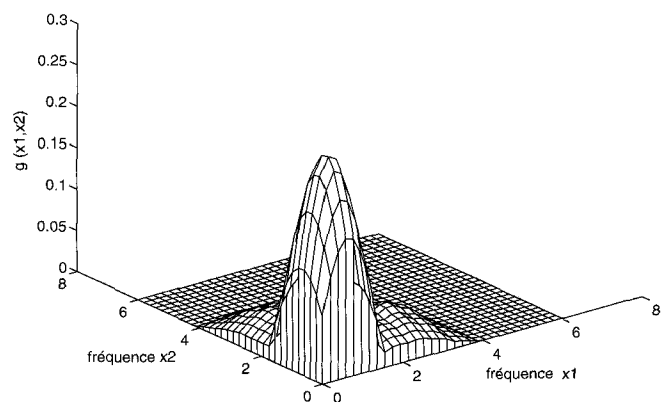


Figure 4. – L'estimateur lissé, $g = f_{N_1, N_2}$, de la densité f avec les paramètres sélectionnés.

Etape 2. – Calculer le périodogramme $\{I_N(w_j, w_{j'}), j = 1, \dots, \bar{N}_1, j' = 1, \dots, \bar{N}_2\}$, en utilisant le FFT2 (Fast Fourier Transform 2D).

Etape 3. – Calculer $f_{N_1, N_2}^{j, j'}(w_j, w_{j'})$, pour (h_1, h_2) dans un intervalle donné.

Etape 4. – Calculer la somme :

$$CV_1(h_1, h_2) = \frac{1}{2\pi \bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} f_{N_1, N_2}^2(w_j, w_{j'}) - \frac{2}{2\pi \bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} f_{N_1, N_2}^{j, j'}(w_j, w_{j'}) I_{N_1, N_2}(w_j, w_{j'}).$$

Etape 5. – Calculer $(\hat{h}_1, \hat{h}_2) = \arg \min_{h_1, h_2} CV_1(h_1, h_2)$

Etape 6. – Calculer l'estimateur $f_{N_1, N_2}(x_1, x_2)$ avec les paramètres (\hat{h}_1, \hat{h}_2) . Cet estimateur est noté $g(x_1, x_2)$ sur les figures 2 et 4.

4. annexe

Démonstration de la proposition 2.1 :

En partant de la définition de $f_{N_1, N_2}^{j, j'}$, et en introduisant la définition de $I_{N_1, N_2}^{j, j'}(u_1, u_2)$, nous obtenons

$$f_{N_1, N_2}^{j, j'}(\lambda_1, \lambda_2) - f_{N_1, N_2}(\lambda_1, \lambda_2) = \int_0^{2\pi} \int_0^{2\pi} W_{N_1}^{(1)}(\lambda_1 - u_1) W_{N_2}^{(2)}(\lambda_2 - u_2) \mathbf{1}_{A_{j, j'}}(u_1, u_2) \gamma_{N_1, N_2}(u_1, u_2) du_1 du_2,$$

où

$$\gamma_{N_1, N_2}(u_1, u_2) = \theta_1(u_1, u_2) I_{N_1, N_2}(w_{j-1}, w_{j'-1}) + \theta_2(u_1, u_2) I_{N_1, N_2}(w_{j+1}, w_{j'-1}) + \theta_3(u_1, u_2) I_{N_1, N_2}(w_{j-1}, w_{j'+1}) + \theta_4(u_1, u_2) I_{N_1, N_2}(w_{j+1}, w_{j'+1}) - I_{N_1, N_2}(u_1, u_2)$$

I_{N_1, N_2} étant un estimateur asymptotiquement sans biais de f et puisque $\omega_{j+1} - \omega_{j-1} = 4\pi/\bar{N}_1$ et $\omega_{j'+1} - \omega_{j'-1} = 4\pi/\bar{N}_2$ nous déduisons facilement que :

$$\mathbb{E}\{f_{N_1, N_2}^{j, j'}(\lambda_1, \lambda_2) - f_{N_1, N_2}(\lambda_1, \lambda_2)\} = O\left(\frac{1}{\bar{N}_1 \bar{N}_2}\right),$$

ce qui achève la démonstration de la proposition.

Démonstration du théorème 2.1 :

L'égalité suivante se démontre de la même manière que dans le cas unidimensionnel :

$$\int_0^{2\pi} \int_0^{2\pi} f_{N_1, N_2}(\lambda_1, \lambda_2) f(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 = \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} f_{N_1, N_2}(w_j, w_{j'}) f(w_j, w_{j'}) + O\left(\frac{1}{\bar{N}_1 \bar{N}_2}\right).$$

Ce résultat nous permet de déduire que :

$$\mathbb{E}\{ISE(h_1, h_2) - CV(h_1, h_2)\} = R_1 + R_2 + O\left(\frac{1}{\bar{N}_1 \bar{N}_2}\right),$$

où

$$R_1 = \frac{1}{\pi \bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} \mathbb{E}\left\{\left[f_{N_1, N_2}^{j, j'}(w_j, w_{j'}) - f_{N_1, N_2}(w_j, w_{j'})\right]\right\} f(w_j, w_{j'}),$$

$$R_2 = \frac{1}{\pi \bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} \mathbb{E}\left\{f_{N_1, N_2}^{j, j'}(w_j, w_{j'}) [I_{N_1, N_2}(w_j, w_{j'}) - f(w_j, w_{j'})]\right\}.$$

D'après la proposition 2.1 et le fait que f soit continue, nous avons

$$R_1 = O\left(\frac{1}{\bar{N}_1 \bar{N}_2}\right). \text{ D'autre part, puisque } f_{N_1, N_2}^{j, j'}(w_j, w_{j'}) \text{ et } I_{N_1, N_2}(w_j, w_{j'}) \text{ sont asymptotiquement indépendants, nous avons } R_2 = O\left(\frac{1}{\bar{N}_1 \bar{N}_2}\right). \text{ Ainsi } \mathbb{E}\{ISE(h_1, h_2) - CV(h_1, h_2)\} = O\left(\frac{1}{\bar{N}_1 \bar{N}_2}\right).$$

D'autre part, nous exprimons la variance de $CV(h_1, h_2)$ en fonction des covariances de la manière suivante : $\text{var}\{CV(h_1, h_2)\} = \frac{1}{(2\pi)^2} (\psi_1 + \psi_2 + \psi_3)$, où

$$\psi_1 = \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \text{cov}\{f_{N_1, N_2}^2(\lambda_1, \lambda_2), f_{N_1, N_2}^2(\mu_1, \mu_2)\} d\lambda_1 d\lambda_2 d\mu_1 d\mu_2$$

$$\psi_2 = \frac{-4}{\bar{N}_1 \bar{N}_2} \sum_{j=1}^{\bar{N}_1} \sum_{j'=1}^{\bar{N}_2} \int_0^{2\pi} \int_0^{2\pi} \text{cov}\left\{f_{N_1, N_2}^2(\lambda_1, \lambda_2), f_{N_1, N_2}^{j, j'}(w_j, w_{j'}) I_{N_1, N_2}(w_j, w_{j'})\right\} d\lambda_1 d\lambda_2$$

$$\psi_3 = \frac{4}{\bar{N}_1 \bar{N}_2} \sum_{j, j'=1}^{\bar{N}_1, \bar{N}_2} \sum_{k, k'=1}^{\bar{N}_1, \bar{N}_2} \text{cov}\left\{f_{N_1, N_2}^{j, j'}(w_j, w_{j'}), f_{N_1, N_2}^{k, k'}(w_k, w_{k'}) I_{N_1, N_2}(w_k, w_{k'})\right\}.$$

Etudions le comportement asymptotique de chaque terme ψ_s , pour $s = 1, 2, 3$.

– **Traitement du terme ψ_1**

En utilisant les relations existantes entre la covariance et les cumulants, nous décomposons ψ_1 de la façon suivante :

$$\begin{aligned} \psi_1 &= \sum_{t=1}^5 \psi_{1,t}, \quad \text{où} \\ \psi_{1,1} &= 2 \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \text{cov}\{f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\mu_1, \mu_2)\}^2 d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 \\ \psi_{1,2} &= 2 \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \mathbb{E}\{f_{N_1, N_2}(\lambda_1, \lambda_2)\} \times \\ &\quad \text{cum}\{f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\mu_1, \mu_2), f_{N_1, N_2}(\mu_1, \mu_2)\} \\ &\quad d\lambda_1 d\lambda_2 d\mu_1 d\mu_2, \\ \psi_{1,3} &= 2 \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \mathbb{E}\{f_{N_1, N_2}(\mu_1, \mu_2)\} \times \\ &\quad \text{cum}\{f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\mu_1, \mu_2)\} \\ &\quad d\lambda_1 d\lambda_2 d\mu_1 d\mu_2, \\ \psi_{1,4} &= 4 \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \mathbb{E}\{f_{N_1, N_2}(\lambda_1, \lambda_2)\} \\ &\quad \mathbb{E}\{f_{N_1, N_2}(\mu_1, \mu_2)\} \times \text{cov}\{f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\mu_1, \mu_2)\} \\ &\quad d\lambda_1 d\lambda_2 d\mu_1 d\mu_2, \\ \psi_{1,5} &= \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \text{cum}\{f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\lambda_1, \lambda_2), \\ &\quad f_{N_1, N_2}(\mu_1, \mu_2), f_{N_1, N_2}(\mu_1, \mu_2)\} d\lambda_1 d\lambda_2 d\mu_1 d\mu_2. \end{aligned}$$

D'après le résultat de Priestley (1981) [7] sur le comportement asymptotique de la covariance de l'estimateur lissé, nous déduisons que

$$\psi_{1,1} = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right)^2 \quad \text{et} \quad \psi_{1,4} = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right).$$

D'autre part, $\psi_{1,2}$ et $\psi_{1,3}$ se comportent de la même manière. Nous traitons donc un seul de ces deux termes. En effet, en utilisant le fait que

$$\text{cum}\{f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\lambda_1, \lambda_2), f_{N_1, N_2}(\mu_1, \mu_2)\} = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right),$$

et le fait que f soit continue sur $[0, 2\pi] \times [0, 2\pi]$, nous obtenons $\psi_{1,2} = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right)$ et $\psi_{1,5} = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right)$. Ainsi nous déduisons que $\psi_1 = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right)$.

En utilisant les propriétés des cumulants et en exploitant les mêmes techniques et mêmes arguments utilisés pour montrer la convergence de ψ_1 vers zéro, nous montrons que $\psi_2 = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right)$ et $\psi_3 = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right)$. Ainsi nous obtenons

$$\text{var}\{CV(h_1, h_2)\} = O\left(\frac{1}{N_1 N_2 h_1 h_2}\right).$$

BIBLIOGRAPHIE

- [1] Beltrao K.I, Bloomfield P, (1987), Determining the bandwidth of a kernel spectrum estimate, *Journal of time series analysis*, 8(1), 21–38.
- [2] Bowman A.W. (1984), an alternative method of cross-validation for smoothing of density estimates, *Biometrika*, 71(2), 253–360.
- [3] Gregoire G, (1993), Least squares cross-validation for counting process intensities, *Scand. J. Statist.* 20, 343–360.
- [4] Hart J, Vieu P, (1990), Data-driven bandwidth choice for density estimation based on dependent data, *Ann. Stat.* 18, 873–890.
- [5] Hurvich C.M, Beltrao K.L, (1990), Cross-validators choice of a spectrum estimate and its connection with aic, *J. Time Series. Analy.*, 11(2), 121–137.
- [6] Messaci F, (1986), Estimation de la densité spectrale d'un processus en temps continu par échantillonnage poissonnien, *Thèse de 3^{ème} cycle à l'univ de Rouen*.
- [7] Priestley MP, (1981), Spectral analysis and time series, *Probability and Mathematical Statistics*, Academic Press.
- [8] Rachdi M., (1998), Choix de la largeur de fenêtre spectrale par validation croisée. Analyse spectrale p -adique, *Thèse de doctorat à l'université de Rouen*.
- [9] Rachdi M., (1998), Choix de la largeur de fenêtre spectral par validation croisée pour un processus stationnaire à temps continue. *C.R. Acad. Sci, Paris*, t. 327, série I, 777–780.
- [10] Rudemo M., (1982), Emperical choice of histograms and kernel density estimators, *Scand. j. Stat.*, 9, 65–78.
- [11] Sabre R., (1994), Estimation de la densité de la mesure spectrale mixte pour un processus symétrique stable strictement stationnaire, *C. R. Acad. Sci.*, Paris, t. 319, Série I, p. 1307–1310.
- [12] Sabre R., (1995), Spectral density estimation for stationary stable random fields, *Journal Applicationes Mathematicae*, 23, 2, p. 107–133.
- [13] Sarda P, Vieu P., (1989), Estimation non paramétrique de la fonction de hasard, *Neuvième rencontre franco-belge de Statisticiens*, cahier du CERO, 31, (3-4), 241–265.

Manuscrit reçu le 18 mars 1999.

LES AUTEURS

Mustapha RACHDI



Docteur de l'université de Rouen en statistiques depuis 1998. Professeur contractuel à la même université. Ses travaux de recherche concernent la statistique des processus. Il travaille en collaboration avec l'équipe de statistique de l'ENESAD à Dijon.

Rachid SABRE



Docteur de l'université de Rouen en statistiques depuis 1993. Maître de conférences à l'ENESAD. Il est responsable du laboratoire MAIS (Mathématiques Appliquées à l'Informatique et aux Statistiques). Il mène une recherche orientée vers la statistique des processus et l'analyse spectrale ainsi que leurs applications en Agronomie et Agroalimentaire.