

A

Simulation de variables aléatoires

Nous présentons ici quelques méthodes classiques de simulation d'une v.a.. Pour une approche plus approfondie, le lecteur pourra consulter le livre de Devroye [59].

Supposons que l'on dispose d'un générateur U de loi uniforme $\mathcal{U}([0, 1])$ sur $[0, 1]$ et que, placé dans une boucle, ce générateur retourne des réalisations (U_n) i.i.d. $\mathcal{U}([0, 1])$.

A.1 La méthode d'inversion

Soit X une v.a.r. de fonction de répartition $F : \mathbb{R} \rightarrow [0, 1]$ définie, si $x \in \mathbb{R}$, par $F(x) = P(X \leq x)$: F est croissante, cadlag (continue à droite, limitée à gauche), de limite 0 en $-\infty$, 1 en $+\infty$. De plus, F est continue partout dès que X admet une densité g ; dans ce cas, $F(x) = \int_{-\infty}^x g(u)du$.

Notons $F^{-1} : [0, 1] \rightarrow \mathbb{R}$ la *pseudo-inverse* de F définie par :

$$F^{-1}(u) = \inf\{x : F(x) \geq u\} \text{ si } u \in [0, 1].$$

La propriété suivante est à la base de la méthode de simulation par inversion : si U est uniforme sur $[0, 1]$, alors $X = F^{-1}(U)$ a pour fonction de répartition F . En effet :

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x). \quad (\text{A.1})$$

Chaque fois que F est explicite, $(F^{-1}(U_n), n \geq 1)$ est une suite i.i.d. de v.a. de loi X . Donnons quelques exemples d'utilisation.

Simulation d'une variable de Bernoulli de paramètre p

$X \in \{0, 1\}$ est de loi $p = P(X = 1) = 1 - P(X = 0)$. On tire alors une variable $U \sim \mathcal{U}([0, 1])$: si $U < 1 - p$ on retient $X = 0$; sinon on prend $X = 1$. La procédure reste valable pour simuler toute loi à valeurs binaires $X \in \{a, b\}$, $a \neq b$.

Simulation d'une variable prenant un nombre fini de valeurs

Supposons que X prenne K valeurs $\{a_1, a_2, \dots, a_K\}$ avec les probabilités $p_k = P(X = a_k)$, $k = 1, \dots, K$, $\sum_1^K p_k = 1$. On construit une partition de $[0, 1[$ en K intervalles adjacents $I_k = [c_{k-1}, c_k[$, $k = 1, \dots, K$ où $c_0 = 0$, $c_k = \sum_1^k p_i$ pour $k = 1, \dots, K$ et on tire $U \sim \mathcal{U}([0, 1])$: si $U \in I_k$, on retient $X = a_k$. Cette méthode réalise bien la simulation de X puisque

$$P(c_{k-1} \leq U < c_k) = p_k = P(X = a_k) \quad \text{pour } k = 1, \dots, K. \quad (\text{A.2})$$

Variable prenant un nombre infini dénombrable de valeurs

Si $X \in E$ prend un nombre infini dénombrable de valeurs, on commencera par déterminer, pour un seuil $\alpha > 0$ petit (par exemple $\alpha = 10^{-3}$) un sous-ensemble fini $E_\alpha \subset E$ concentrant $\geq 1 - \alpha$ de la "masse" totale de X , c-à-d tel que $P(X \in E_\alpha) \geq 1 - \alpha$. Puis on simulera X sur l'ensemble fini des modalités $E_\alpha \cup \{E \setminus E_\alpha\}$, $\{E \setminus E_\alpha\}$ regroupant toutes les modalités en dehors de E_α en un unique état.

Par exemple, pour simuler une loi de Poisson de paramètre λ , on commence par déterminer la valeur n_0 telle que $P(X > n_0) < 10^{-3}$, puis on simule X sur $\{0, 1, 2, \dots, n_0\} \cup \{\text{plus}\}$ de la façon décrite ci-dessus, pour les probabilités :

$$P(X = n) = p_n = e^{-\lambda} \frac{\lambda^n}{n!} \quad \text{si } 0 \leq n \leq n_0 \quad \text{et} \quad P(X \in \{\text{plus}\}) = 1 - \sum_0^{n_0} p_n.$$

Remarquons que là aussi, les modalités de X peuvent être qualitatives.

Simulation d'une loi exponentielle

L'identité (A.1) permet de simuler une loi dont l'espace d'état $E \subseteq T$ est continu chaque fois que F^{-1} est explicite. Si $X \sim \text{Exp}(\lambda)$ suit une loi exponentielle de paramètre $\lambda > 0$, alors $F(x) = 0$ si $x < 0$ et $F(x) = 1 - e^{-\lambda x}$ si $x \geq 0$ et on en déduit que, pour tout $u \in [0, 1]$,

$$F^{-1}(u) = -\frac{\log(1-u)}{\lambda}.$$

U comme $1 - U$ étant uniforme sur $[0, 1]$ si $U \sim \mathcal{U}([0, 1])$, $X = -\log(U)/\lambda$ simule une loi exponentielle de paramètre λ . De même,

$$X = -\frac{\sum_{n=1}^N \log U_n}{\lambda}$$

réalise la simulation d'une loi $\Gamma(\lambda, N)$ pour tout paramètre entier $N \geq 1$. On va voir plus loin comment simuler une telle loi si ce paramètre n'est pas entier.

A.2 Simulation d'une chaîne de Markov à nombre fini d'état

Soit $X = (X_0, X_1, X_2, \dots)$ une chaîne de Markov homogène à valeur dans l'espace d'état fini $E = \{a_1, a_2, \dots, a_K\}$ et de transition $P = (p_{ij})$ [120; 103; 229] :

$$p_{ij} = P(X_{n+1} = j \mid X_n = i).$$

Avec la convention $p_{i,0} = 0$, on définit pour chaque $i = 1, \dots, K$ la partition de $[0, 1[$ en K intervalles adjacents $I_i(j)$:

$$I_i(j) = [c_{j-1}(i), c_j(i)[, \quad j = 1, \dots, K$$

où $c_j(i) = \sum_{l=0}^j p_{i,l}$.

Considérons alors l'application $\Phi : E \times [0, 1] \rightarrow E$ définie, pour tout $i = 1, \dots, K$ et $u \in [0, 1]$ par :

$$\Phi(a_i, u) = a_j, \quad \text{si } u \in I_i(j).$$

Si $U \sim \mathcal{U}([0, 1])$, on a alors, pour tout $i, j = 1, \dots, K$:

$$P(\Phi(a_i, U) = a_j) = P(U \in I_i(j)) = p_{ij}.$$

La suite $\{x_0, X_{n+1} = \Phi(X_n, U_n), n \geq 0\}$ réalise la simulation d'une chaîne de Markov de transition P de condition initiale $X_0 = x_0$. Si $X_0 \sim \nu_0$ est simulée à partir de (A.2), la suite $(U_n, n \geq 0)$ permet de simuler la chaîne de Markov de loi initiale ν_0 et de transition P .

A.3 La méthode de rejet

Cette méthode réalise la simulation d'une v.a. X de densité f sur \mathbb{R}^p à condition de connaître une loi de densité g sur \mathbb{R}^p , facile à simuler et telle que, pour tout x , $f(x) \leq cg(x)$ pour un $c < \infty$. Considérons en effet Y une v.a. de densité g et soit $U \sim \mathcal{U}([0, 1])$ indépendante de Y . Alors, la variable conditionnelle suivante est de loi X :

$$(Y \mid \text{si } cUg(Y) < f(Y)) \sim X.$$

En effet, puisque $c = \int cg(y)dy \geq \int f(y)dy > 0$ et que $f(y) = 0$ si $g(y) = 0$, on a :

$$\begin{aligned} P(Y \in [x, x + dx] \mid cUg(Y) < f(Y)) &= \frac{g(x)dxP(U < \frac{f(x)}{cg(x)})}{P(U < \frac{f(Y)}{cg(Y)})} \\ &= \frac{g(x)\frac{f(x)}{cg(x)}dx}{\int \frac{f(y)}{cg(y)}g(y)dy} \\ &= f(x)dx = P(X \in [x, x + dx]). \end{aligned}$$

Pour que la méthode soit utilisable, Y doit être facile à simuler et c ne doit pas être trop grand (rejet pas trop fréquent); quand il est possible, le choix optimal de c est $c = \sup_x f(x)/g(x)$.

Exemple : simulation de lois Gamma et Beta

Une loi Gamma $\Gamma(\lambda, a)$, $\lambda > 0$, $a > 0$, a pour densité

$$f(x) = \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbf{1}(x > 0),$$

où $\Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$.

Si $Y \sim \Gamma(1, a)$, alors $X = Y/\lambda \sim \Gamma(\lambda, a)$; d'autre part, la somme $X + Z$ de deux v.a. indépendantes de loi $\Gamma(\lambda, a)$ et $\Gamma(\lambda, b)$ suit une loi $\Gamma(\lambda, a + b)$. Il suffit donc de savoir simuler une loi $\Gamma(1, a)$ pour $a \in]0, 1]$ pour simuler toutes les lois $\Gamma(\lambda, a^*)$ pour $\lambda > 0$, $a^* > 0$.

La simulation de Y , une v.a. de loi $\Gamma(1, a)$, $a \in]0, 1]$, s'obtient par méthode de rejet : on remarque que pour la densité $f(x) = \Gamma(a)^{-1} e^{-x} x^{a-1} \mathbf{1}(x > 0)$ de Y ,

$$f \leq \frac{a+e}{ae\Gamma(a)} g,$$

où $g(x) = (a+e)^{-1} [e g_1(x) + a g_2(x)]$ avec $g_1(x) = a x^{a-1} \mathbf{1}(0 < x < 1)$ et $g_2(x) = e^{-x+1} \mathbf{1}(1 < x < \infty)$. g_1 est une densité sur $]0, 1[$, g_2 sur $]1, +\infty[$, l'une et l'autre simulables par la méthode d'inversion, et g est le mélange de ces deux lois pour les poids $(e/(a+e), a/(a+e))$. Une première loi uniforme permet de choisir si la réalisation de g est dans $]0, 1[$ ou dans $]1, +\infty[$, après quoi on simule la variable retenue de densité g_i par méthode d'inversion, la simulation de Y étant alors obtenue par la méthode de rejet : trois lois $\mathcal{U}([0, 1])$ sont donc utilisées dans cette simulation. D'autres méthodes peuvent faire l'économie d'une variable uniforme, en particulier en simulant directement la loi de densité g par la méthode d'inversion.

Une loi Beta de paramètres a et $b > 0$, notée $\beta(a, b)$, a pour densité

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbf{1}(0 < x < 1).$$

La simulation d'une telle loi découlera de la simulation de lois Γ au travers de la propriété suivante : si $X \sim \Gamma(\lambda, a)$ et $Y \sim \Gamma(\lambda, b)$ sont indépendantes, alors $X/(X+Y) \sim \beta(a, b)$.

A.4 Simulation d'une loi gaussienne

Simulation d'une gaussienne réduite $\mathcal{N}(0, 1)$

Les logiciels fournissent un générateur de la loi normale réduite. Celui-ci peut s'obtenir à partir de deux lois U_1 et $U_2 \sim \mathcal{U}([0, 1])$ indépendantes de la façon

suivante (méthode de Box-Muller) : les deux variables X_1 et X_2 suivantes sont normales réduites

$$X_1 = \sqrt{-2 \log U_1} \cos(2\pi U_2), \quad X_2 = \sqrt{-2 \log U_2} \cos(2\pi U_1).$$

La simulation de $Y \sim \mathcal{N}(m, \sigma^2)$ s'obtiendra de façon affine, $Y = m + \sigma X$, à partir de celle d'une normale réduite X . Les lois log-normale, du χ^2 , de Student et de Fisher se simulent facilement à partir du générateur gaussien réduit.

Simulation d'une loi gaussienne multivariée : $Z \sim \mathcal{N}_p(\mu, \Sigma)$

Si Σ est de rang plein, et si p n'est pas trop grand ($p < 5,000$), on recherchera une décomposition (par exemple de Choleski) $\Sigma = A {}^t A$ de Σ : alors, si $X = {}^t(X_1, X_2, \dots, X_p)$ est un p -échantillon gaussien réduit, $\mu + AX$ réalise la simulation de Z . Si $r = \text{rang}(\Sigma) < p$, on commence par rechercher le sous-espace S , de dimension r , support de Z , puis, en se plaçant dans une base orthonormée de S , on simule le vecteur gaussien par la méthode précédente.

B

Théorèmes limites pour un champ aléatoire

B.1 Ergodicité et lois des grands nombres

Nous rappelons ici quelques résultats d'ergodicité et de lois des grands nombres (LGN) pour les processus spatiaux, une loi forte (LFGN) pour un résultat en convergence presque sûr (p.s.) et faible pour une convergence dans L^2 .

B.1.1 Ergodicité et théorème ergodique

Soit $X = \{X_s, s \in S\}$ un champ aléatoire réel sur $S = \mathbb{R}^d$ ou $S = \mathbb{Z}^d$. L'ergodicité est une propriété qui renforce la notion de stationnarité et qui permet d'obtenir la convergence *p.s.* d'une moyenne empirique spatiale lorsque le domaine d'observation "tend vers l'infini". Si on se limite à une convergence dans L^2 , une ergodicité au second ordre suffira. L'ergodicité est importante en statistique car elle permet d'établir la consistance *p.s.* d'estimateur explicite comme une moyenne spatiale. Cependant il n'est pas indispensable de disposer de cette propriété, très forte, d'ergodicité pour établir une telle consistance ; des conditions de sous-ergodicité, voire des conditions L^2 , peuvent suffire (consistance de l'estimation par PVC ou par codage, cf. Th. 5.4 et Th. 5.6 ; consistance de l'estimation par Minimum de Contraste, cf. Appendice C, Th. C.1).

Notant $\Omega = \mathbb{R}^S$ l'espace des états de X et \mathcal{E} sa tribu borélienne, on dira qu'un événement $A \in \mathcal{E}$ est *invariant* par translation si pour tout $i \in S$, $\tau_i(A) = A$ où τ_i est la i -translation sur Ω ainsi définie : pour tout $i \in S$ et $\omega \in \Omega$, $(\tau_i(\omega))_j = \omega_{j-i}$. Les événements invariants forment une sous-tribu $\mathcal{I} \subseteq \mathcal{E}$.

Un processus stationnaire X est caractérisé par le fait que sa loi P est invariante par translation : pour tout événement $A = \{\omega : X_{s_1}(\omega) \in B_1, X_{s_2}(\omega) \in B_2, \dots, X_{s_n}(\omega) \in B_n\}$ générateur de la tribu borélienne, on a :

$$\forall i \in S : P(A) = P(\tau_i(A)).$$

L'ergodicité renforce cette propriété.

Définition B.1. *Le processus stationnaire X est ergodique si pour tout événement invariant $A \in \mathcal{I}$, $P(A) = 0$ ou $P(A) = 1$.*

Si X est ergodique, la tribu \mathcal{I} des invariants est triviale. L'interprétation heuristique est que si un événement invariant A est de probabilité > 0 , alors il est de probabilité 1, l'ensemble des translatés de A engendrant l'espace de toutes les trajectoires.

L'exemple de base d'un processus ergodique est celui d'une suite de variables *i.i.d.*. L'ergodicité peut être vue comme une propriété jointe de stationnarité et d'indépendance asymptotique.

Enonçons alors le théorème ergodique. Si $B(x, r)$ est la boule centrée en x de rayon r , on définit le diamètre intérieur $d(D)$ de $D \subseteq \mathbb{R}^d$ par :

$$d(D) = \sup\{r : B(x, r) \subseteq D\}.$$

Théorème B.1. *(Birkhoff [32] si $d = 1$; Tempelman [211] et Nguyen et Zessin [164] si $d \geq 2$)*

Supposons que X soit un processus réel sur \mathbb{R}^d , stationnaire, de L^p pour un $p \geq 1$ et que (D_n) soit une suite croissante de convexes bornés telle que $d(D_n) \rightarrow \infty$. Alors :

1. $\bar{X}_n = |D_n|^{-1} \int_{D_n} X_u du \rightarrow E(X_0 | \mathcal{I})$ dans L^p et *p.s.* si $S = \mathbb{R}^d$.
2. Si de plus X est ergodique, $\bar{X}_n \rightarrow E(X_0)$ *p.s.*

Sur $S = \mathbb{Z}^d$, on a : $\bar{X}_n = \#D_n^{-1} \sum_{i \in D_n} X_i \rightarrow E(X_0 | \mathcal{I})$ dans L^p et *p.s.* si $S = \mathbb{Z}^d$, la limite *p.s.* étant $E(X_0)$ si X est ergodique.

B.1.2 Exemples de processus ergodiques

Donnons quelques exemples de processus ergodiques.

$X = \{X_i, i \in S\}$ est une suite de v.a. réelles *i.i.d.* de L^1 sur un ensemble dénombrable S .

$Y = \{Y_i = g(X \circ \tau_i), i \in S\}$ où X est ergodique sur S et $g : E^V \rightarrow \mathbb{R}$, pour $V \subset S$ fini, est une application mesurable.

X est un champ stationnaire fortement mélangeant [116; 67] :

$$\lim_{\|h\| \rightarrow \infty} P(\tau_h(A) \cap B) = P(A)P(B), \quad \forall A, B \in \mathcal{E}.$$

X est un champ stationnaire m -dépendant : $\forall U \subset S$ et $V \subset S$ distants d'au moins m , alors les variables $\{X_u, u \in U\}$ et $\{X_v, v \in V\}$ sont indépendantes.

X est un champ gaussien stationnaire sur \mathbb{R}^d (\mathbb{Z}^d) dont la covariance C tend vers 0 à l'infini [3] :

$$\lim_{\|h\| \rightarrow \infty} C(h) = 0.$$

X est un champ de Gibbs sur \mathbb{Z}^d à potentiel invariant par translation et satisfaisant la condition d'unicité de Dobrushin [85].

$N = \{N([0, 1^{d+i}), i \in \mathbb{Z}^d\}$ où $N(A)$ est le nombre de points dans A d'un PP ergodique sur \mathbb{R}^d (par exemple un PPP homogène, un PP de Neyman-Scott homogène ou un PP de Cox d'intensité un champ Λ ergodique [204].

Y est un sous-ensemble aléatoire de $\mathbb{R}^d : Y = \cup_{x_i \in X} B(x_i, r)$ où X est un PPP homogène (un PP ergodique) sur \mathbb{R}^d ; Y est un champ booléen sur \mathbb{R}^d , $Y_s = 1$ si $s \in Y$ et 0 sinon [43].

L'ergodicité des PP (et des ensembles aléatoires fermés de \mathbb{R}^d qui incluent les PP) est étudiée dans [125; 204; 56] et [135] (cas des PPP). Heinrich [111] et Guan et Sherman [95] utilisent ces propriétés pour prouver la consistance d'estimateurs paramétriques d'un PP. Une façon d'approcher l'ergodicité d'un PP X est de lui associer le processus latticiel \tilde{X} de ses réalisations sur la partition $\mathbb{R}^d = \cup_{i \in \mathbb{Z}^d} A_i$, où $A_i = [i, i + \mathbf{1}[$, $\mathbf{1}$ étant le vecteur de composantes 1 de \mathbb{R}^d , et de vérifier l'ergodicité de \tilde{X} .

B.1.3 Ergodicité et LGN faible dans L^2

Supposons que X soit un champ stationnaire au second ordre sur \mathbb{R}^d de covariance C . On dira que X est ergodique dans L^2 si pour toute suite (D_n) de convexes bornés telle que $d(D_n) \rightarrow \infty$, alors

$$\overline{X}_n = \frac{1}{|D_n|} \int_{D_n} X_u du \rightarrow E(X_0) \text{ dans } L^2.$$

Soit F la mesure spectrale de X . On a la loi faible des grans nombres suivante [227, Ch. 3; 96, Ch. 3; 43]

Théorème B.2. *On a les équivalences :*

- (i) X est ergodique dans L^2 .
- (ii) F n'a pas d'atome en 0 : $F(\{0\}) = 0$.
- (iii) $|D_n|^{-1} \int_{D_n} C(u) du \rightarrow 0$.

En particulier, ces conditions sont satisfaites si $\lim_{\|h\| \rightarrow \infty} C(h) = 0$ ou si F est absolument continue.

Ces résultats se transposent directement sur \mathbb{Z}^d : par exemple, si F est absolument continue sur T^d ,

$$\lim_n \overline{X}_n = E(X_0) \text{ dans } L^2.$$

Si de plus la densité spectrale f est bornée et continue en 0, alors

$$\lim_{n \rightarrow \infty} \#D_n \text{Var}(\overline{X}_n) = (2\pi)^d f(0),$$

la convergence de $\sqrt{\#D_n} \overline{X}_n$ étant gaussienne si X est gaussien.

B.1.4 LFGN sous conditions L^2

On dispose aussi de LFGN sous des conditions L^2 . Un premier résultat (cf. [32], §3.4) concerne les suites de variables indépendantes mais non (nécessairement) équidistribuées (cf. lemme 5.1) : si $X = \{X_i, i \in \mathbb{N}\}$ sont des v.a.r. indépendantes, centrées et dans L^2 , alors $\sum_{i=1}^n X_i \rightarrow 0$ p.s. dès que $\sum_{i=1}^\infty \text{Var}(X_i) < \infty$.

Un autre résultat concerne les estimateurs empiriques de la moyenne μ et de la covariance $C_X(k)$ d'un processus X stationnaire au second ordre sur \mathbb{Z}^d ([96], §3.2) : si $\sum_{h \in \mathbb{Z}^d} |C_X(h)| < \infty$, alors $\bar{X}_n \rightarrow \mu$ p.s.; si de plus X est stationnaire au 4^{ème} ordre, $Y = \{Y_i = (X_i - \mu)(X_{i+k} - \mu), i \in \mathbb{Z}^d\}$ vérifiant, $\sum_{h \in \mathbb{Z}^d} |C_Y(h)| < \infty$, alors $\bar{Y}_n \rightarrow C_X(k)$ p.s.; si X est gaussien, cette dernière condition est satisfaite dès que X est à densité spectrale de carré intégrable.

B.2 Coefficient de mélange fort

Soit $Z = \{Z_i, i \in S\}$ un champ sur un réseau S muni d'une métrique d , A et B deux parties de S , $\mathcal{F}(Z, H)$ la tribu engendrée par Z sur H , une partie de S . Le coefficient de *mélange fort* $\alpha^Z(E, F)$ de Z sur E, F , défini par Ibragimov et Rozanov ([117]; cf. aussi Doukhan [67]) est :

$$\alpha^Z(E, F) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}(Z, E), B \in \mathcal{F}(Z, F)\}.$$

Le champ est α -mélangeant si $\alpha^Z(E, F) \rightarrow 0$ dès que $\text{dist}(E, F) \rightarrow \infty$.

On utilise également les coefficients α paramétrés par les tailles k et $l \in \mathbb{N} \cup \{\infty\}$ de E et de F ; pour $n \in \mathbb{R}$:

$$\alpha_{k,l}(n) = \alpha_{k,l}^Z(n) = \sup\{\alpha^Z(E, F) : \#E \leq k, \#F \leq l \text{ et } \text{dist}(E, F) \geq n\}.$$

Exemples de champ α -mélangeant

1. Si Z est α -mélangeant, toute fonctionnelle mesurable $Z_W = (f_i(Z_{W_i}))$ qui dépend localement de Z est encore un champ α -mélangeant : plus précisément, si pour tout $i \in S$, (W_i) est une famille de parties t.q. $i \in W_i$ et le diamètre $\delta(W_i)$ de W_i est uniformément borné par Δ , alors $Z_W = (f_i(Z_{W_i}))$ est α -mélangeant où $\alpha_{k,l}^{Z_W}(n) = \alpha_{k,l}^Z(n - 2\Delta)$ dès que $n > 2\Delta$.
2. *Champ R -dépendant* : Z est un champ R -dépendant si, pour tout couple (i, j) à distance $> R$, Z_i et Z_j sont indépendants. Pour un champ R -dépendant, $\alpha^Z(E, F) = 0$ si $d(E, F) \geq R$ et, pour tout k, l , $\alpha_{k,l}^Z$ est nulle si $n \geq 2R$.
3. *Champ gaussien* : soit Z un champ gaussien stationnaire sur \mathbb{Z}^d de densité spectrale f , $f(\lambda) \geq a > 0$ sur le tore \mathbb{T}^d . Notons $D_K(f)$ la distance entre f et sa meilleure approximation par un polynôme trigonométrique de degré $(K - 1)$.

$$D_K(f) = \inf\{|f - P| : P(\lambda) = \sum_{|t| < K} c_t \exp\langle \lambda, t \rangle\} \text{ où } |t| = \sum_{i=1}^d |t_i|.$$

On a alors ([117] pour $d = 1$; [96, §1.7], pour les champs) :

$$\alpha_{\infty, \infty}(k) \leq \frac{1}{a} D_K(f).$$

Si f est continue, ce coefficient tend vers 0 si $k \rightarrow \infty$, exponentiellement vite si f est analytique (par exemple, si Z est un *ARMA*). Si Z admet la représentation linéaire $Z_t = \sum_{\mathbb{Z}^d} b_{t-s} \varepsilon_t$ pour un BB gaussien ε , on a, plus précisément :

$$\alpha_{\infty, \infty}(k) \leq \frac{2}{a} \|b\|_{\infty} \left\{ \sum_{|s| \geq k/2} |sb_s| \right\}.$$

On dispose aussi d'évaluations du coefficient de mélange pour des champs linéaires non gaussiens [67; 96].

4. *Champ de Gibbs sous la condition d'unicité de Dobrushin* (Dobrushin [66]; Georgii [85] Guyon [96]).

La mesure d'influence de Dobrushin $\gamma_{a,b}(\pi)$ d'un site a sur le site b , $a \neq b$ d'une spécification de Gibbs π ($\gamma_{a,b}(\phi)$ si la spécification dérive d'un potentiel ϕ) est définie par :

$$\gamma_{a,b}(\pi) = \sup \frac{1}{2} \|\pi_b(\cdot | \omega) - \pi_b(\cdot | \omega')\|_{VT},$$

où $\|\cdot\|_{VT}$ est la norme en variation totale, le sup étant pris sur les configurations ω et ω' identiques partout sauf au site a . Si $a = b$, on pose $\gamma_{a,b}(\pi) = 0$. On dit que le potentiel de Gibbs satisfait la condition de Dobrushin si

$$(D) : \alpha(\phi) = \sup_{a \in S} \sum_{b \in S} \gamma_{a,b}(\phi) < 1 \tag{B.1}$$

(D) est une condition suffisante (mais non-nécessaire) assurant qu'il y a au plus une mesure de Gibbs dans $\mathcal{G}(\phi)$. Par exemple, pour le modèle d'Ising sur \mathbb{Z}^2 , isotropique et aux 4-ppv, de spécification en i ,

$$\pi_i(z_i | z^i) = \frac{\exp \beta z_i v_i}{\exp -\beta v_i + \exp \beta v_i}, \quad v_i = \sum_{j: \|i-j\|_1=1} z_j$$

la condition exacte d'unicité (Onsager [166]; Georgii [85]) est $\beta < \beta_c = \frac{1}{2} \log(1 + \sqrt{2}) \simeq 0.441$ alors que la condition de Dobrushin, qui s'explique facilement dans ce cas, est $\beta < \frac{1}{4} \log 2 \simeq 0.275$.

Si de plus l'espace d'état du champ de Gibbs est polonais (par exemple fini ou compact), muni d'une mesure de référence λ positive et finie et

si le potentiel ϕ est sommable, il y aura *existence* et unicité de la mesure de Gibbs μ associée à $\pi : \mathcal{G}(\phi) = \{\mu\}$; dans ce cas, si ϕ est borné et de portée bornée, l'unique mesure de Gibbs μ vérifie une condition de mélange uniforme exponentiel,

$$\varphi(A, B) \leq C (\#A) \alpha^{d(A,B)},$$

où $\varphi(A, B) = \sup\{|\mu(E|F) - \mu(E)|, E \in \mathcal{F}(A), F \in \mathcal{F}(B), \mu(F) > 0\}$. μ est également fortement mélangeant puisqu'on a toujours $2\alpha(\cdot) \leq \varphi(\cdot)$ [67].

Une difficulté de la statistique asymptotique pour un champ de Gibbs μ connu par sa spécification $\pi(\phi)$ est qu'on ne sait pas en général s'il y a unicité ou non de μ dans $\mathcal{G}(\phi)$. On ne dispose donc pas en général de propriété de faible dépendance. De même, si le potentiel ϕ est invariant par translation ($S = \mathbb{Z}^d$), on ne sait pas si μ est ergodique, voir si μ est stationnaire : ainsi, les outils classiques de la statistique asymptotique (ergodicité, faible dépendance, TCL) ne sont pas toujours utilisables.

5. *Mélange pour un processus ponctuel spatial.*

Un PP de Poisson X , homogène ou non, est α -mélangeant puisque X est indépendant. Cette propriété de mélange s'étend aux PP de Neyman-Scott si les lois de descendance sont spatialement bornées par R : dans ce cas, le PP est $2R$ -dépendant.

D'autres exemples de champs α -mélangeant sont présentés dans Doukhan [67].

B.3 TCL pour un champ mélangeant

Soit (D_n) une suite strictement croissante de sous-ensembles finis de \mathbb{Z}^d , Z un champ centré réel et de variance finie, $S_n = \sum_{D_n} Z_i$ et $\sigma_n^2 = \text{Var}(S_n)$. Notons $(\alpha_{k,l}(\cdot))$ les coefficients de mélange de Z . On a le résultat suivant (Bolthausen [30]; Gyon [96], sans stationnarité) :

Proposition B.1. *TCL pour un champ réel sur \mathbb{Z}^d*

Supposons vérifiées les conditions suivantes :

(i) $\sum_{m \geq 1} m^{d-1} \alpha_{k,l}(m) < \infty$ si $k + l \leq 4$ et $\alpha_{1,\infty}(m) = o(m^{-d})$.

(ii) Il existe $\delta > 0$ t.q. : $\sup_{i \in S} \mathbb{E} |Z_i|^{2+\delta} < \infty$ et

$$\sum_{m \geq 1} m^{d-1} \alpha_{1,1}(m)^{\delta/2+\delta} < \infty.$$

(iii) $\liminf_n (\#D_n)^{-1} \sigma_n^2 > 0$.

Alors, $\sigma_n^{-1} S_n \xrightarrow{\text{loi}} \mathcal{N}(0, 1)$.

Commentaires

1. Les conditions sur le mélange sont satisfaites pour un champ gaussien stationnaire de densité suffisamment régulière ou pour un champ de Gibbs sous la condition d'unicité de Dobrushin (B.1).
2. Si on veut utiliser un seul coefficient de mélange, on peut se limiter à $\alpha_{2,\infty}$ et à la condition :

$$\sum_{m \geq 1} m^{d-1} \alpha_{2,\infty}(m) < \infty.$$

3. Le résultat du théorème reste inchangé si $S \subset \mathbb{R}^d$ est un réseau infini dénombrable localement fini : $\exists \delta_0 > 0$ t.q. pour tout $i \neq j$, deux sites de S , $\|i - j\| \geq \delta_0$. En effet, dans ce cas et comme pour le réseau régulier \mathbb{Z}^d , la propriété clé est que pour tout $i \in S$, la boule centrée en i de rayon m vérifie, uniformément en i et m : $\#\{B(i, m) \cap S\} = O(m^d)$.
4. Comme pour tout TCL, la condition de positivité (iii) peut être délicate à obtenir.
5. Si $Z \in \mathbb{R}^k$ est un champ multidimensionnel, si $\Sigma_n = Var(S_n)$, et si (iii) est remplacée par

$$(iii)' : \liminf_n (\#D_n)^{-1} \Sigma_n \geq I_0 > 0,$$

pour une matrice d.p. I_0 , alors : $\Sigma^{-1/2} S_n \xrightarrow{loi} \mathcal{N}_k(0, I_k)$.

6. Les conditions données par Bolthausen sur le mélange semblent minimales. Signalons cependant le travail de Lahiri [137] qui établit un TCL à partir des seuls coefficients $\alpha_k^*(n) \equiv \alpha_{k,k}(n)$, $k < \infty$. Les conditions requises sont du type de celles de Bolthausen mais sans l'évaluation des coefficients $\alpha_{k,\infty}(\cdot)$. Ce travail propose également des TCL pour des choix de sites d'observations aléatoires avec la possibilité de densification des sites d'observation (*mixed increasing-domain infill asymptotics*).

B.4 TCL pour une fonctionnelle d'un champ de Markov

Supposons que Z soit un champ de Markov sur $S = \mathbb{Z}^d$, à valeur dans E et de spécification π invariante par translation. Notons $V_i = \{Z_j, j \in \partial i\}$ les m valeurs dont dépend la spécification locale, $\pi_i(\cdot | Z^{\{i\}}) \equiv \pi_i(\cdot | V_i)$ et considérons une fonctionnelle mesurable $A : E^{m+1} \rightarrow \mathbb{R}$ telle que $Y_i = A(Z_i, V_i)$ vérifie, pour tout i , la condition de *centrage conditionnel* :

$$E(Y_i | Z_j, j \neq i) = 0. \tag{B.2}$$

L'exemple type où cette condition est vérifiée est le cas où Y est le gradient (en θ) de la pseudo-vraisemblance conditionnelle d'un champ de Markov.

Pour simplifier, supposons que $D_n = [-n, n]^d$ soit la suite des domaines d'observation de Z et notons $S_n = \sum_{D_n} Y_i$.

Nous présentons deux TCL pour la fonctionnelle Y : le premier [99; 123] utilise la propriété d'ergodicité du champ Z , propriété qui remplace celle de mélange. Le deuxième (Comets et Janzura [47]), plus général, suppose uniquement l'invariance par translation de la spécification conditionnelle π de Z et donne une version studentisée du TCL sous une condition d'uniforme positivité de la variance de S_n par observation. Le résultat de Comets et Janzura s'applique donc sans hypothèse d'ergodicité ni de stationnarité du champ, qu'il y ait ou non transition de phase pour la spécification π .

Proposition B.2. [99; 123; 96]

Supposons que Z soit un champ de Markov sur \mathbb{Z}^d de spécification invariante par translation, ergodique, la fonctionnelle a étant bornée et vérifiant la condition de centrage (B.2). Alors, si $\sigma^2 = \sum_{j \in \partial 0} E(Y_0 Y_j) > 0$, on a :

$$(\#D_n)^{-1/2} S_n \xrightarrow{loi} \mathcal{N}(0, 1).$$

Comme pour chaque TCL, la condition $\sigma^2 > 0$ peut être délicate à établir. Cette condition est établie dans [99] pour le modèle d'Ising isotropique et aux 4-ppv sur \mathbb{Z}^d .

Venons-en au résultat de [47]. Définissons $A_n = \sum_{i \in D_n} \sum_{j \in \partial i} Y_i Y_j$ et pour $\delta > 0$, $A_n^\delta = \max\{A_n, \delta \times \#D_n\}$. Remarquons que A_n estime sans biais la variance de S_n .

Proposition B.3. (Comets-Janzura [47])

Supposons que Z est un champ de Markov sur \mathbb{Z}^d de spécification invariante par translation, que la fonctionnelle a vérifie : $\sup_i \|Y_i^4\| < \infty$ ainsi que la condition de centrage (B.2). Définissons la version studentisée de S_n :

$$\zeta_n = A_n^{-1/2} S_n \text{ si } A_n > 0 \text{ et } \zeta_n = 0 \text{ sinon.}$$

Alors, sous la condition :

$$\exists \delta > 0 \text{ tel que } (\#D_n)^{-1} E |A_n - A_n^\delta| \xrightarrow{n \rightarrow \infty} 0 \tag{B.3}$$

$$\zeta_n \xrightarrow{loi} \mathcal{N}(0, 1).$$

Estimation par minimum de contraste

Nous présentons ici la méthode d'estimation d'un modèle paramétrique (ou semi-paramétrique) par *minimum de contraste* (Dacunha-Castelle et Duflo [54]). Certains auteurs parlent encore d'estimation par *pseudo-vraisemblance* (Whittle [222]; Besag [25]), ou par *quasi-vraisemblance* (McCullagh et Nelder [155]), ou encore d'*estimateurs d'extremum* (Amemiya [6]; Gourieroux et Monfort [92]). Ces différentes terminologies recouvrent un même principe : la valeur estimée du paramètre maximise une fonctionnelle de "pseudo-vraisemblance" (PV). Cette fonctionnelle se substitue à la vraisemblance quand celle-ci est indisponible, soit parce que le modèle est semi-paramétrique et incomplètement spécifié, soit parce que cette vraisemblance est numériquement incalculable.

Sous de "bonnes conditions" sur la fonctionnelle de pseudo-vraisemblance et sur le schéma d'observations, l'estimation par maximum de pseudo-vraisemblance (MPV) a de bonnes propriétés statistiques : convergence, normalité et test asymptotique pour le paramètre d'intérêt. Ces fonctionnelles, une fois pénalisée par la dimension du modèle, permettent également d'identifier un modèle.

Deux fonctionnelles de pseudo-vraisemblance jouent un rôle central en statistique des processus spatiaux :

1. La *pseudo-vraisemblance gaussienne* pour un modèle du second-ordre. Elle s'obtient en calculant la vraisemblance (ou une approximation de celle-ci) en faisant l'hypothèse que le modèle est gaussien. Ce contraste a été introduit par Whittle [222] pour les séries temporelles et pour les champs sur \mathbb{Z}^2 (cf. § 5.3.1).
2. La *pseudo-vraisemblance conditionnelle* (PVC) d'un champ markovien sur un réseau (cf. § 5.4.2), produit des densités conditionnelles en chaque site (Besag [25]; Guyon [96]). Si ce produit est limité à un sous-ensemble de codage, on obtient la *pseudo-vraisemblance par codage*. La notion de PVC existe aussi pour les processus ponctuels de Markov (cf. 5.5.5).

L'exemple standard de contraste est la fonctionnelle de moindres carrés (estimation d'une régression spatiale (cf. 5.3.4), d'un modèle de variogramme (cf. §5.1.3), d'un processus ponctuel (cf. 5.5.4)). Des contrastes pour modèles spatio-temporels sont proposés dans [41; 98; 100].

D'une façon générale, une fonctionnelle de pseudo-vraisemblance devra répondre aux objectifs suivants :

1. Capter simplement les informations que l'on privilégie sur le modèle.
2. Être numériquement facile à calculer.
3. Rendre identifiable les paramètres du modèle.
4. Permettre (si possible) un contrôle statistique de l'estimateur.

C.1 Définitions et exemples

Considérons $X = \{X_i, i \in S\}$ un processus défini sur un ensemble fini ou dénombrable de sites S . Que notre connaissance sur X soit partielle (modèle semi-paramétrique) ou totale (modèle paramétrique), on notera $\theta \in \Theta \subseteq \mathbb{R}^p$ le paramètre spécifiant l'information qui nous intéresse.

L'objectif est d'estimer θ à partir d'observations $X(n) = \{X_i, i \in D_n\}$, où D_n est un sous-ensemble fini de sites. L'étude asymptotique est associée à une suite (D_n) strictement croissante de domaines d'observation. Pour simplifier, on supposera que θ , la vraie valeur inconnue du paramètre, est un point intérieur à Θ , un compact de \mathbb{R}^p ; $\alpha \in \Theta$ dénote un point courant de Θ .

Une *fonction de contraste* pour θ est une fonction déterministe

$$K(\cdot, \theta) : \Theta \rightarrow \mathbb{R}, \quad \alpha \mapsto K(\alpha, \theta) \geq 0,$$

admettant un unique minimum en $\alpha = \theta$. La valeur $K(\alpha, \theta)$ s'interprète comme une pseudo-distance entre le modèle sous θ et celui sous α .

Un *processus de contraste* associé à la fonction de contraste $K(\cdot, \theta)$ et aux observations $X(n)$ est une suite $(U_n(\alpha), n \geq 1)$ de variables aléatoires adaptée à $X(n)$, $U_n(\alpha) = U_n(\alpha, X(n))$, définie pour tout $\alpha \in \Theta$, et telle que :

$$\forall \alpha \in \Theta : \liminf_n [U_n(\alpha) - U_n(\theta)] \geq K(\alpha, \theta), \text{ en } P_\theta\text{-probabilité.} \quad (\text{C.1})$$

Cette condition *sous-ergodique* (C.1) traduit que la différence $U_n(\alpha) - U_n(\theta)$ qui estime le contraste de α sur θ sur la base de $X(n)$ sépare asymptotiquement les paramètres. La condition (C.1) peut être renforcée par la condition "ergodique",

$$\lim_n [U_n(\alpha) - U_n(\theta)] = K(\alpha, \theta), \text{ en } P_\theta\text{-probabilité.} \quad (\text{C.2})$$

Définition C.1. *L'estimateur du minimum de contraste U_n est une valeur $\hat{\theta}_n$ de Θ minimisant U_n :*

$$\hat{\theta}_n = \underset{\alpha \in \Theta}{\operatorname{argmin}} U_n(\alpha).$$

Donnons quelques exemples.

Exemple C.1. Vraisemblance d'un modèle de Bernoulli

Supposons que les X_i soient des v.a. de Bernoulli indépendantes de paramètres $p_i = p(\alpha, Z_i) = (1 + \exp \alpha Z_i) / (\exp \alpha Z_i)$, où Z_i est une variable exogène réelle, $\alpha \in \mathbb{R}$. Le contraste de la vraisemblance est l'opposé de la vraisemblance, $U_n(\alpha) = -\sum_1^n \log f_i(X_i, \alpha)$. Si les exogènes (Z_i) et α restent bornés, la fonction de contraste

$$K(\alpha, \theta) = \liminf_n \frac{1}{n} \sum_1^n \log \frac{f_i(X_i, \theta)}{f_i(X_i, \alpha)}$$

satisfait (C.1) dès que $\liminf_n n^{-1} \sum_{i=1}^n Z_i^2 > 0$.

Exemple C.2. Contraste des moindres carrés d'une régression

Considérons un modèle de régression (linéaire ou non),

$$X_i = m(Z_i, \theta) + \varepsilon_i, \quad i = 1, \dots, n$$

expliquant $X_i \in \mathbb{R}$ à partir de l'exogène Z_i , les résidus (ε_i) formant un BBF de variance $\sigma^2 < \infty$. Ce modèle est *semi-paramétrique* car aucune hypothèse n'est faite sur les lois des résidus en dehors du fait que ε soit un BBF. Le contraste des moindres carrés ordinaires (MCO) est défini par :

$$U_n(\alpha) = SCR_n(\alpha) = \sum_{i=1}^n (X_i - m(Z_i, \alpha))^2.$$

Définissons $K(\alpha, \theta) = \liminf_n \sum_{i=1}^n \{m(x_i, \alpha) - m(x_i, \theta)\}^2 / n$. Si le dispositif expérimental $\mathcal{Z} = \{Z_i, i = 1, 2, \dots\}$ est tel que $K(\alpha, \theta) > 0$ pour $\alpha \neq \theta$, alors (U_n) est un processus de contraste associé à la fonction de contraste $K(\cdot, \theta)$. Cette condition est par exemple vérifiée si :

1. Les Z_i sont i.i.d. de loi $Z : \mathcal{Z}$ est ergodique.
2. \mathcal{Z} rend le modèle $\theta \mapsto m(\cdot, \theta)$ identifiable, c-à-d :

$$\text{Si } \theta \neq \theta', \text{ alors } P_Z \{z : m(z, \theta) \neq m(z, \theta')\} > 0.$$

Si les résidus sont gaussiens, $U_n(\alpha)$ est, à une constante multiplicative près, l'opposé de la log-vraisemblance.

Si les résidus sont corrélés de matrice de variance R_n connue et inversible, le contraste des *moindres carrés généralisés* (MCG) est,

$$U_n^{MCG}(\alpha) = \|X(n) - m_n(\alpha)\|_{R_n^{-1}}^2,$$

où $m_n(\alpha) = \{m(Z_1, \alpha), \dots, m(Z_n, \alpha)\}$ et $\|u\|_G^2 = {}^t u \Gamma u$ est la norme associée à la matrice d.p. Γ . Le contraste des *moindres carrés pondérés* correspond à la norme associée à la matrice de variance diagonale,

$$U_n^P(\alpha) = \sum_{i=1}^n \frac{(X_i - m(Z_i, \alpha))^2}{\text{Var}(\varepsilon_i)}.$$

Exemple C.3. Méthode de moments, PV marginale

Supposons que X_1, X_2, \dots, X_n soient des observations réelles d'une même loi μ_θ dépendant d'un paramètre $\theta \in \mathbb{R}^p$. Notons $(\mu_k(\theta), k = 1, \dots, r)$ les r premiers moments de la loi commune, $(\hat{\mu}_{n,k}(\theta), k = 1, \dots, r)$ les estimations empiriques de ces moments. Si D est une distance sur \mathbb{R}^p , un contraste pour l'estimation de θ est

$$U_n(\theta) = D((\mu_k), (\hat{\mu}_{n,k}(\theta))).$$

Une condition nécessaire pour que ce contraste rende θ identifiable est que $r \geq p$. Pour construire un contraste rendant le paramètre identifiable, on peut être amené à considérer, au-delà de la loi marginale μ_θ de X_1 , des lois de couples, voir de triplets La méthode s'étend au cas où X est à valeur dans un espace d'état E général.

De même, des densités marginales peuvent être utilisées pour construire des estimateurs de θ . Si, par exemple, les couples (X_i, X_{i+1}) sont de même loi, et si ce couple permet d'identifier θ , on pourra utiliser la PV marginale des couples :

$$l_n(\theta) = \sum_{i=1}^{n-1} \log f(x_i, x_{i+1}; \theta).$$

Exemple C.4. Contraste gaussien d'un processus au second ordre

Supposons que $X = (X_t, t \in \mathbb{Z})$ soit une série temporelle réelle, centrée, stationnaire à l'ordre 2, de densité spectrale f_θ . Le spectrogramme associé aux observations $X(n) = (X_1, X_2, \dots, X_n)$ est l'estimation de la densité spectrale associée aux covariances empiriques $\hat{r}_n(k)$:

$$I_n(\lambda) = \frac{1}{2\pi} \sum_{k=-n+1}^{n-1} \hat{r}_n(k) e^{i\lambda k}$$

où $\hat{r}_n(k) = \hat{r}_n(-k) = n^{-1} \sum_{i=1}^{n-|k|} X_i X_{i+k}$.

Le spectrogramme $I_n(\lambda)$ est un mauvais estimateur de $f(\lambda)$. Par contre, le contraste de Whittle défini par la régularisation suivante de I_n :

$$U_n(\alpha) = \frac{1}{2\pi} \int_0^{2\pi} \left\{ \log f_\alpha(\lambda) + \frac{I_n(\lambda)}{f_\alpha(\lambda)} \right\} d\lambda$$

est une bonne fonctionnelle pour estimer θ . Sous hypothèse gaussienne, $-2U_n(\alpha)$ approxime la log-vraisemblance. Sans hypothèse gaussienne, U_n conduit encore à une bonne estimation sous de bonnes conditions générales (Dalhaus et Künsch [57]; Guyon [96]; cf. 5.3.1). La fonction de contraste associée à U_n est :

$$K(\alpha, \theta) = \frac{1}{2\pi} \int_0^{2\pi} \left\{ \log \frac{f_\alpha(\lambda)}{f_\theta(\lambda)} - 1 + \frac{f_\alpha(\lambda)}{f_\theta(\lambda)} \right\} d\lambda, \quad K(\alpha, \theta) > 0 \text{ si } \alpha \neq \theta.$$

La condition $K(\alpha, \theta) \neq 0$ si $\alpha \neq \theta$ est assurée si la paramétrisation de f_θ en θ est identifiable.

Exemple C.5. Pseudo-vraisemblance conditionnelle (PVC) d'un champ de Markov

Si la vraisemblance d'une chaîne de Markov se calcule bien par récursivité, ce n'est plus le cas pour un champ de Markov spatial qui est fondamentalement un modèle non-causal. La raison tient en partie à la complexité du calcul de la constante de normalisation. La même difficulté se présente pour un modèle de Gibbs général.

C'est la raison pour laquelle Besag a proposé, dans le contexte d'un champ de Markov sur un réseau, d'utiliser la *pseudo-vraisemblance conditionnelle* (PVC), produit des densités conditionnelles en chaque i de D_n ,

$$l_n^{PV}(\theta) = \prod_{D_n} \pi_i(x_i | x_{\partial i}, \theta). \quad (\text{C.3})$$

Si on limite le produit à un ensemble C de codage, on parlera de *pseudo-vraisemblance sur le codage C* . Pour l'une ou l'autre de ces fonctionnelles, on devra s'assurer de la condition de sous-ergodicité (C.1). Si l'ergodicité du champ est suffisante, elle n'est pas nécessaire. Cette remarque est importante car en général on ne sait pas si un champ de Gibbs est ergodique.

Exemple C.6. Estimation par moindres carrés d'un modèle de variogramme

Si des données géostatistiques sont modélisées par un processus intrinsèque de variogramme $\gamma(\cdot, \theta)$, une façon classique d'estimer θ est de minimiser le contraste des moindres carrés,

$$U_n^{MC}(\theta) = \sum_{i=1}^k (\hat{\gamma}_n(h_i) - \gamma(h_i, \theta))^2.$$

Dans cette expression, $\hat{\gamma}_n(h_i)$ sont des estimations empiriques du variogramme en h_i , ceci pour k vecteurs (h_i) préalablement choisis (cf. §5.1.3). Cette méthode de moindres carrés est également utilisée pour l'estimation de modèles paramétriques de processus ponctuels spatiaux (cf. §5.5.4).

Exemple C.7. Pseudo-transition marginale d'une dynamique de système de particules

Considérons, à titre d'exemple, une dynamique de présence/absence $X_i(t) \in \{0, 1\}$, $i \in \mathbb{Z}^2$, $t = 0, 1, \dots$ d'une espèce végétale sur le réseau spatial \mathbb{Z}^2 . Durrett et Levin [74] proposent d'étudier un processus de contact à temps discret à deux paramètres (γ, λ) ainsi caractérisé (cf. Fig. C.1) : soient x et y les configurations aux instants successifs t et $(t + 1)$,

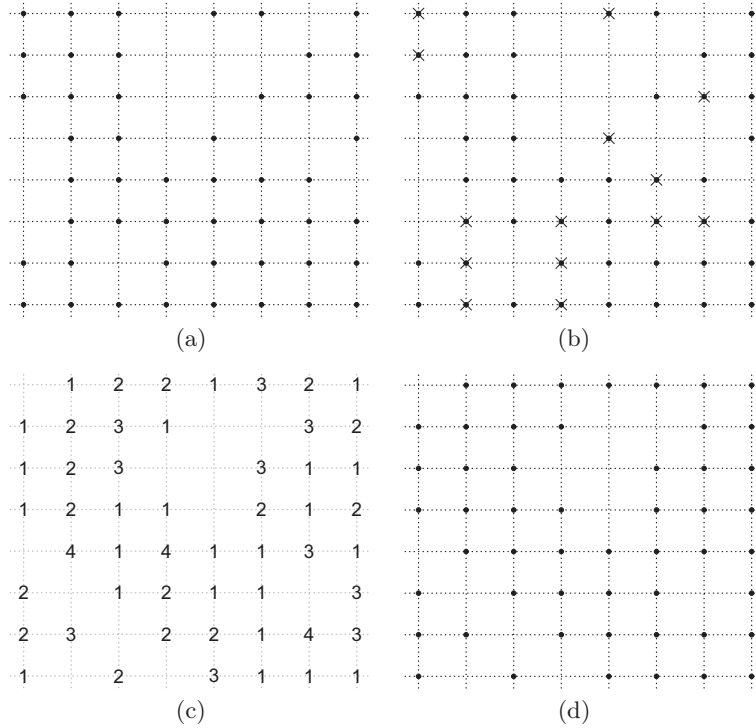


Fig. C.1. Un exemple d'évolution de la dynamique $t \rightarrow t+1$ d'un processus de contact sur un réseau régulier : (a) configuration $X(t)$ au temps t (les \bullet représentent les plantes vivantes); (b) plantes (\times) qui meurent avec probabilité $\gamma = 0.3$; (c) nombre de semis dans une parcelle après dissémination dans chacune des 4 parcelles voisines de façon i.i.d. avec probabilité $\lambda = 0.25$; (d) configuration $X(t+1)$ au temps $t+1$.

1. Une plante en un site i survit avec la probabilité $(1 - \gamma)$.
2. Si une plante survit, elle essaime dans chacune des 4 parcelles voisines de façon i.i.d. avec la probabilité λ .
3. $y_i = 1$ si au moins une plante est présente sur la parcelle i au temps $(t+1)$.

On suppose de plus que tous les tirages définissant cette dynamique sont indépendants dans l'espace et dans le temps.

Après les étapes 1. et 2., notons $z_i \in \{0, 1, 2, 3, 4, 5\}$ le nombre de plantes présentes en i au temps $(t+1)$; par exemple $z_i = 5$ si $x_i = 1$ et survit et si chaque plante en un site j voisin de i existe, survit et essaime en i . Le point 3. nous dit alors que $y_i = \mathbf{1}\{z_i \geq 1\} = 1$. Si la simulation d'une telle dynamique est simple, le calcul de la transition $P_S(x, y)$ est impossible dès que S est grand : en effet, d'une part, il faut examiner chaque configuration particulière x pour calculer la probabilité jointe de la réalisation $z = \{z_i, i \in S\}$,

problème qui est de complexité $2^{\#(S)}$; d'autre part, on est en situation de données incomplètes puisque seules les variables $\mathbf{1}\{z_i \geq 1\}$ sont observées et le calcul de la loi de l'observation au temps $(t + 1)$ s'en trouve compliquée.

Guyon et Pumo [101] proposent alors de remplacer cette transition par la *pseudo-transition marginale*

$$M_S(x, y) = \prod_{i \in I(x, S)} P_\theta(X_i(t + 1) = y_i | x), \quad (\text{C.4})$$

produit des transitions marginales sur $I(x, S) = \{s \in S : x_s + \sum_{j \in \partial_s} x_j \geq 1\}$. Les probabilités en jeu sont faciles à calculer,

$$P_\theta(X_i(t + 1) = y_i | x) = P_\theta(X_i(t + 1) = y_i | x_{\{i\} \cup \partial_i}).$$

$I(x, S)$ est l'ensemble des seuls sites où x apporte de l'information puisqu'en effet, si $x_s + \sum_{t \in \partial_s} x_t = 0$ au temps t , alors $X_{t+1}(s) = 0$.

La *pseudo-vraisemblance marginale* des observations temporelles $\{x_S(0), x_S(1), \dots, x_S(T)\}$ est alors définie par le produit,

$$M_{S,T}(\theta) = \prod_{t=0}^{T-1} M_S(x(t), x(t + 1)).$$

Conditionnellement à la survie du processus, Guyon et Pumo [101] établissent la consistance et la normalité asymptotique de l'estimateur associé à cette PV marginale.

Exemple C.8. Autres exemples

Dans la définition (C.3), $\log \pi_i(x_i | x^i, \alpha)$ peut être remplacée par d'autres fonctionnelles $h_i(x(V_i); \alpha)$ mieux adaptées et permettant d'identifier le modèle. Si la dépendance conditionnelle π_i est non-bornée et/ou non-explicite, on peut la remplacer par une fonctionnelle ad hoc : par exemple si $X_i \in \mathbb{R}$ et si on sait calculer $g_i(x(W_i), \alpha) = E_\alpha(X_i | x(W_i))$, alors

$$h_i(x(V_i), \alpha) = \{x_i - g_i(x(W_i), \alpha)\}^2 \quad \text{pour } V_i = \{i\} \cup W_i$$

conduit à une pseudo-vraisemblance de *moindres carrés conditionnels* (MCC). Une *pseudo-vraisemblance marginale* pour laquelle $h_i(x(V_i), \alpha)$ est la log-densité marginale de $X(V_i)$ sous α constitue un autre exemple de contraste.

C.2 Propriétés asymptotiques

Les résultats dans leur version "ergodique" sont dus à Dacunha-Castelle et Duflo [54]. D'autres preuves sont données par Amemiya [6] et Gourieroux-Montfort [92]. La version non-ergodique des propriétés asymptotiques de l'estimation par minimum de contraste est exposée dans ([96], Ch. 3).

C.2.1 Convergence de l'estimateur

On a le résultat de convergence suivant (Hardouin [108]; Guyon [96]) :

Théorème C.1. $\widehat{\theta}_n \xrightarrow{\text{Pr}} \theta$ sous les conditions suivantes :

(C1) $\alpha \mapsto K(\alpha, \theta)$ est continue ainsi que, P_α -p.s., les contrastes $\alpha \mapsto U_n(\alpha)$.

(C2) (U_n) vérifie la condition de sous-ergodicité (C.1).

(C3) Si $W_n(\eta)$ est le module de continuité de $U_n(\cdot)$, $\exists \varepsilon_k \downarrow 0$ t.q. pour chaque k :

$$\lim_n P_\theta(W_n(1/k) \geq \varepsilon_k) = 0. \quad (\text{C.5})$$

Corollaire C.1. Soit $U_n = \sum_{i=1}^p a_{n,i} U_{n,i}$, $a_{n,i} \geq 0$ un processus de contraste tel que :

1. $U_{n,1}$ satisfait (C1-C2-C3).
2. Chaque $U_{n,i}$ satisfait (C3).
3. $a = \liminf_n a_{n,1} > 0$.

Alors l'estimateur du minimum de contraste U_n est convergent.

Ce corollaire explique pourquoi les conditions assurant la convergence de l'estimation d'un champ de Markov pour un codage C impliquent celle de l'estimateur par pseudo-vraisemblance conditionnelle U_n (cf. § 5.4.2). En effet, dans ce cas, $U_n = U_{n,C} + \sum_l U_{n,C_l}$ pour une partition $\{C, (C_l)\}$ de S en sous-ensembles de codage et les conditions du corollaire sont bien satisfaites.

Pour les processus de contrastes convexes, on a le résultat de convergence p.s. suivant (Senoussi [196]; Guyon [96]) :

Proposition C.1. Si Θ est un ouvert convexe de \mathbb{R}^p , si les contrastes $\theta \mapsto U_n(\theta)$ sont convexes et si (C.2) est vérifiée, alors $\widehat{\theta}_n \xrightarrow{p.s.} \theta$.

Obtention d'un estimateur efficace par algorithme Newton-Raphson

Considérons un estimateur $\widehat{\theta}_n$ de $\theta \in \mathbb{R}^p$ solution d'un système de p équations :

$$F(x(n); \theta) = 0, \theta \in \mathbb{R}^p \quad (\text{C.6})$$

où F à valeur dans \mathbb{R}^p . Les estimateurs du minimum de contraste sont de ce type pour $F_n(\theta) = U_n^{(1)}(\theta)$, le gradient de U_n .

La recherche de la solution de (C.6) n'étant pas toujours facile, il est utile d'utiliser l'algorithme de Newton-Raphson initialisé avec un "bon" estimateur $\widetilde{\theta}_n$ facile à obtenir, estimateur conduisant en un pas à θ_n^* donné par :

$$\theta_n^* = \widetilde{\theta}_n - \mathcal{F}^{-1}(\widetilde{\theta}_n) F(\widetilde{\theta}_n), \quad (\text{C.7})$$

où $\mathcal{F}(\alpha)$ est la matrice $p \times p$ d'éléments $F_{i,\alpha_j}^{(1)}(\alpha)$, dérivée en α_j de la composante i de F , $i, j = 1, \dots, p$.

Si F est assez régulière et si $\tilde{\theta}_n$ est consistant à une vitesse suffisante, Dzharidze [75] montre que θ_n^* est asymptotiquement équivalent à $\hat{\theta}_n$. Précisons ce résultat. Soit $(v(n))$ une suite réelle tendant vers l'infini : on dira qu'un estimateur $\bar{\theta}_n$ de θ est $v(n)$ -consistant si,

$$v(n)(\bar{\theta}_n - \theta) = O_P(1).$$

Deux estimateurs $\bar{\theta}_n$ et θ_n^* sont asymptotiquement $v(n)$ -équivalents si,

$$\lim_n v(n)(\bar{\theta}_n - \theta_n^*) = o_P(1).$$

Posons alors les conditions :

(DZ-1) L'équation (C.6) admet une solution $\hat{\theta}_n$ qui est $\tau(n)$ -consistante.

(DZ-2) F est de classe $\mathcal{C}^2(\mathcal{V}(\theta))$, pour $\mathcal{V}(\theta)$ un voisinage de θ et $\exists W(\theta)$ régulière non-stochastique t.q.,

$$\lim_n (\mathcal{F}(\theta) - W(\theta)) \stackrel{P_\theta}{=} 0.$$

(DZ-3) Les dérivées secondes vérifient :

$$\forall \delta > 0, \exists M < \infty \text{ t.q. } \lim_n P_\theta \{ \sup \{ \|F^{(2)}(\alpha)\|, \alpha \in \mathcal{V}(\theta) \} < M \} \geq 1 - \delta.$$

Proposition C.2. (Dzharidze [75]) *Supposons satisfaites les conditions (DZ). Alors, si $\tilde{\theta}_n$ est un estimateur initial $\tilde{\tau}(n)$ -consistant de θ à la vitesse $\tilde{\tau}(n) = o(\sqrt{\tau(n)})$, l'estimateur θ_n^* donné par (C.7) est asymptotiquement $\tau(n)$ -équivalent à $\hat{\theta}_n$:*

$$\lim_n \tau(n)(\hat{\theta}_n - \theta_n^*) = o_P(1).$$

C.2.2 Normalité asymptotique

Quelques notations

Si h est une fonction réelle de classe \mathcal{C}^2 au voisinage de θ , $h^{(1)}(\theta)$ dénote le gradient de h (vecteur des dérivées premières en θ), et $h^{(2)}(\theta)$ la matrice hessienne des dérivées secondes en θ . Si A et B sont deux matrices $p \times p$ symétriques, on notera :

$$\|A - B\| = \sum_{i,j} |A_{ij} - B_{ij}|.$$

$A \geq B$ (resp. $A > B$) si $A - B$ est s.d.p. (resp. d.p.).

Si $A > 0$ admet la décomposition spectrale $A = PD^tP$, où P est orthogonale et D diagonale, on choisira $R = PD^{\frac{1}{2}}$ comme matrice racine carrée de A : $R^tR = {}^tRR = A$.

Hypothèses (N) assurant la normalité asymptotique

(N1) Il existe un voisinage V de θ sur lequel U_n est de classe \mathcal{C}^2 et une variable aléatoire réelle h , P_θ -intégrable, tels que :

$$\forall \alpha \in V, \|U_{n,\alpha^2}(\alpha, x)\| \leq h(x).$$

(N2) Les matrices $J_n = \text{Var}(\sqrt{a_n}U_n^{(1)}(\theta))$ existent et il existe une suite $(a_n) \rightarrow \infty$ telle que :

(N2-1) Il existe J d.p. telle que à partir d'un certain rang, $J_n \geq J$.

(N2-2) $\sqrt{a_n}J_n^{-1/2}U_n^{(1)}(\theta) \xrightarrow{\text{loi}} \mathcal{N}_p(0, I_p)$.

(N3) Il existe une suite de matrices non-stochastiques (I_n) telle que :

(N3-1) Il existe I d.p. et non-stochastique telle que, à partir d'un certain rang, $I_n \geq I$.

(N3-2) $(U_n^{(2)}(\theta) - I_n) \xrightarrow{\text{Pr}} 0$.

Théorème C.2. *Normalité asymptotique de $\hat{\theta}_n$ [108; 96]*

Si $\hat{\theta}_n$ est convergent et si les conditions (N) sont vérifiées, alors :

$$\sqrt{a_n}J_n^{-1/2}I_n(\hat{\theta}_n - \theta) \xrightarrow{\text{loi}} \mathcal{N}_p(0, I_p).$$

Corollaire C.2. *Si U_n est le contraste de la vraisemblance, alors :*

1. *Sous les hypothèses du théorème précédent, on peut choisir pour $I_n = J_n$, la matrice d'information de Fisher :*

$$\sqrt{a_n}I_n^{-1/2}(\hat{\theta}_n - \theta) \xrightarrow{\text{loi}} \mathcal{N}_p(0, I_p).$$

2. *Sous la condition d'ergodicité $I_n \xrightarrow{\text{Pr}} I(\theta) > 0$,*

$$\sqrt{a_n}(\hat{\theta}_n - \theta) \xrightarrow{\text{loi}} \mathcal{N}_p(0, I(\theta)^{-1}).$$

Commentaires

1. J_n et I_n sont des matrices de pseudo-information qu'il faudra "minorer" positivement.
2. La convergence gaussienne (N2-2) résultera d'un TCL pour des variables faiblement dépendantes ou pour un champ conditionnellement centré (cf. Th. B.2).

Exemple C.9. Contraste additif pour un champ latticiel mélangeant [108; 96]

Les conditions (A1–A3) ci-dessous impliquent (N). Elles sont relatives à un champ faiblement dépendant défini sur un réseau S discret de \mathbb{R}^d , non-nécessairement régulier, et au contraste additif :

$$U_n(\alpha) = \frac{1}{d_n} \sum_{s \in D_n} g_s(X(V(s)), \alpha).$$

U_n est la somme de fonctionnelles locales $\{g_s, s \in S\}$, où $\{V(s), s \in S\}$ est une famille de voisinages bornés de $s \in S$ et d_n est le cardinal de D_n . Par exemple, g_s est (l'opposé) du log d'une (pseudo) densité, marginale ou conditionnelle.

(A1) *Sur le lattice* $S \subset \mathbb{R}^d$:

S est infini, localement fini : $\forall s \in S$ et $\forall r > 0$, $\#\{B(s, r) \cap S\} = O(r^d)$.

(A2) *Sur le processus* X : X est α -mélangeant, de coefficient de mélange $\alpha(\cdot) = \alpha_{\infty, \infty}(\cdot)$ (Doukhan [67] ; Guyon [96], § B.2) satisfaisant :

$$(A2-1) \exists \delta > 0 \text{ t.q. } \sum_{i, j \in D_n} \alpha(d(i, j))^{\frac{\delta}{2+\delta}} = O(d_n).$$

$$(A2-2) \sum_{l \geq 0} l^{d-1} \alpha(l) < \infty.$$

(A3) *Sur les fonctionnelles* $(g_s, s \in S)$.

(A3-1) (N1) est vérifiée uniformément sur V par les $g_s, s \in S$.

$$(A3-2) \forall s \in S, E_{\theta}(g_s^{(1)}(\theta)) = 0 \text{ et } \sup_{s \in S, \alpha \in \Theta, k=1,2} \|g_s^{(k)}(\alpha)\|_{2+\delta} < \infty.$$

(A3-3) il existe deux matrices symétriques I et J d.p. telles que pour n grand :

$$J_n = \text{Var}(\sqrt{d_n} U_n^{(1)}(\theta)) \geq J > 0 \text{ et } I_n = E_{\theta}(U_n^{(2)}(\theta)) \geq I > 0.$$

Test asymptotique du rapport de pseudo-vraisemblance

Soit (H_p) l'hypothèse de base $\theta \in \Theta \subset \mathbb{R}^p$, de dimension p , et (H_q) , $q < p$, une sous-hypothèse définie par la spécification fonctionnelle :

$$(H_q) : \alpha = r(\varphi), \varphi \in \Lambda \text{ un ouvert de } \mathbb{R}^q, \theta = r(\phi) \tag{C.8}$$

où $r : \Lambda \rightarrow \Theta$ est de classe $\mathcal{C}^2(W)$ sur W , un voisinage de ϕ , la vraie valeur du paramètre sous (H_q) . On supposera que (H_q) est de dimension q et que $R = \frac{\partial r}{\partial \alpha}(\phi)$ est de rang q .

Il y a deux façons d'aborder le problème du test de la sous-hypothèse (H_q) : l'une construit un test de différence de contraste ; l'autre est le test de Wald testant que $\alpha = r(\varphi)$ s'exprime sous la forme contrainte $C(\theta) = 0$.

Statistique de différence de contraste

Notons $\bar{U}_n(\varphi) = U_n(r(\varphi))$ le contraste sous (H_q) , $\hat{\varphi}_n$ l'estimateur du minimum de contraste associé, $\hat{\theta}_n = r(\hat{\varphi}_n)$. Le test de différence de contraste utilise la statistique :

$$\Delta_n = 2a_n \left[U_n(\bar{\theta}_n) - U_n(\hat{\theta}_n) \right].$$

Soient $\bar{I}_n, \bar{J}_n, \bar{I}$ et \bar{J} les matrices définies de façon analogue à I_n, J_n, I et J mais pour \bar{U}_n , et :

$$A_n = J_n^{1/2} (I_n^{-1} - R \bar{I}_n^{-1} {}^t R) J_n^{1/2}.$$

A_n est une matrice s.d.p. de rang $(p - q)$ dont on note $\{\lambda_{i,n}, i = 1, \dots, p - q\}$ les valeurs propres > 0 . Notons $(F_n \stackrel{\mathcal{L}}{\sim} G_n)$ pour deux suites de fonctions de répartition telles que, pour tout $x \in \mathbb{R}$, $\lim_n (F_n(x) - G_n(x)) = 0$.

Théorème C.3. *Test asymptotique de différence de contraste ([22; 96])*

Supposons que $\hat{\theta}_n$ soit convergent, que (U_n) (resp. (\bar{U}_n)) vérifie les hypothèses (N) sous (H_p) (resp. sous (H_q)). Alors, pour n grand et pour des χ_1^2 indépendants :

$$\text{sous } (H_q) : \Delta_n = 2a_n \left[U_n(\bar{\theta}_n) - U_n(\hat{\theta}_n) \right] \stackrel{\mathcal{L}}{\sim} \sum_{i=1}^{p-q} \lambda_{i,n} \chi_{i,1}^2.$$

Commentaires

1. Si on peut choisir $I_n = J_n$, A_n est idempotente de rang $(p - q)$: on retrouve alors, dans une situation non-nécessairement ergodique, le test du χ_{p-q}^2 du rapport de vraisemblance. On a $I_n = J_n$ pour les contrastes suivants :
 - (a) La vraisemblance d'observations indépendantes.
 - (b) La vraisemblance d'une chaîne de Markov non-nécessairement homogène.
 - (c) Le contraste de codage d'un champ de Markov.
 - (d) Plus généralement, un modèle de variables conditionnellement indépendantes, avec une condition de sous-ergodicité sur les variables conditionnantes.
2. Si le modèle est ergodique, notant I (resp. J, \bar{I}) la limite de (I_n) (resp. $(J_n), (\bar{I}_n)$), et $\{\lambda_i, i = 1, \dots, p - q\}$ les valeurs propres > 0 de $A = J^{1/2} (I^{-1} - R \bar{I}^{-1} {}^t R) J^{1/2}$, alors $\Delta_n \stackrel{\mathcal{L}}{\sim} \sum_{i=1}^{p-q} \lambda_i \chi_{i,1}^2$ sous (H_q) .

Test de spécification sous contrainte

Supposons que (H_q) s'écrive sous une forme contrainte :

$$(H_q) : \psi = C(\theta) = 0,$$

où $C : \mathbb{R}^p \rightarrow \mathbb{R}^{p-q}$ est une contrainte de classe \mathcal{C}^2 au voisinage de θ de rang $(p - q)$ en θ , $\text{rang}(C_\alpha^{(1)}(\theta)) = p - q$. Une façon directe de tester (H_q) est d'utiliser la statistique de Wald associée à la contrainte : ψ est estimée par $\hat{\psi}_n = C(\hat{\theta}_n)$ et le test de (H_q) repose sur la statistique

$$\Xi_n = {}^t \hat{\psi}_n \Sigma_n^{-1} \hat{\psi}_n \sim \chi_{p-q}^2 \text{ sous } (H_q)$$

où $\Sigma_n = C^{(1)}(\hat{\theta}_n) \widehat{\text{Var}}(\hat{\theta}_n) {}^t C^{(1)}(\hat{\theta}_n)$ est la variance estimée de $\hat{\psi}_n$ sous (H_p) .

C.3 Identification d'un modèle par contraste pénalisé

Supposons que l'espace des paramètres vérifie $\Theta \subseteq \mathbb{R}^M$ où \mathbb{R}^M correspond à un modèle majorant, $M < \infty$. Un choix habituel pour une famille \mathcal{E} de modèles envisageables est la famille de parties non-vides de $M = \{1, 2, \dots, M\}$,

$$\delta = \{\theta = (\theta_i)_{i \in M} \text{ t.q. } \theta_i = 0 \text{ si } i \notin \delta\},$$

ou encore une suite croissante d'espaces δ . D'autres choix peuvent se révéler utiles pour intégrer des hypothèses d'isotropie d'un champ.

Identifier le modèle, c'est, au vu de la réalisation $X(n)$, estimer par $\hat{\delta}_n$ le support $\delta \in \mathcal{E}$ du modèle. Si (a_n) est la vitesse associée au contraste (U_n) (cf. (N2)), on utilise comme fonction de décision le *contraste* (pseudo-vraisemblance) *pénalisé* à la vitesse $c(n)$ par la dimension $|\delta|$ du modèle :

$$W_n(\alpha) = U_n(\alpha) + \frac{c(n)}{a_n} |\delta(\alpha)|.$$

Notons :

$$\begin{aligned} \overline{W}_n(\delta) &= \overline{U}_n(\delta) + \frac{c(n)}{a_n} |\delta| \\ \text{avec } \overline{U}_n(\delta) &= \underset{\alpha \in \Theta_\delta}{\text{argmin}} U_n(\alpha). \end{aligned}$$

Le choix répondant au principe de parcimonie de Akaike [4] est :

$$\hat{\delta}_n = \underset{\delta \in \mathcal{E}}{\text{argmin}} \overline{W}_n(\delta).$$

Ce choix réalise un compromis entre un bon ajustement (qui exige un modèle assez "grand", $\overline{U}_n(\delta)$ étant décroissante de δ) et un modèle simple et interprétable (modèle explicatif de "petite taille").

On dira qu'un tel critère identifie le vrai modèle si $\hat{\delta}_n \rightarrow \delta_0$: par exemple, une vraisemblance pénalisée permet d'identifier un modèle convexe d'observations i.i.d. [196]. De même, sous de bonnes conditions, le contraste de Whittle d'une série temporelle stationnaire identifie le modèle [199; 107; 34]. L'identification est aussi possible sans hypothèse d'ergodicité.

La preuve de ce type de résultat d'identifiabilité repose d'une part sur le contrôle des probabilités de mauvaise identification de modèle, d'autre part sur une condition du type Loi du Logarithme Itéré pour le gradient $U_n^{(1)}$ du processus de contraste. Ces conditions et les résultats sont présentés dans un cadre général dans [96, §3.4]. De façon plus précise, Guyon et Yao [102] donnent pour une large classe de modèles et de contrastes associés (régression et moindres carrés, AR et contraste de Whittle, champ de Markov et pseudo-vraisemblance conditionnelle, modèle à variance infinie), une description des ensembles de sur- et de sous-paramétrisation de modèle, l'évaluation de leurs probabilités, évaluation dont découle les conditions assurant la consistance du critère pour chaque type de modèle. Pour l'identification d'un modèle de champ de Markov par pseudo-vraisemblance conditionnelle pénalisée, on pourra aussi consulter [124] et [52].

C.4 Preuve de deux résultats du chapitre 5

C.4.1 Variance de l'estimateur du MV d'une régression gaussienne

Soit l la log-vraisemblance d'une régression gaussienne (§ 5.3.4), $l^{(1)} = (l_\delta^{(1)}, l_\theta^{(1)})$; on vérifie directement que $l_\delta^{(1)} = {}^t Z \Sigma^{-1} (Z\delta - X)$. Utilisant l'identité $\partial/\partial\theta_i \log(|\Sigma|) = \text{tr}(\Sigma^{-1} \Sigma_i)$, on obtient :

$$(l_\theta^{(1)})_i = 2^{-1} \{ \text{tr}(\Sigma^{-1} \Sigma_i) + {}^t (X - Z\delta) \Sigma^i (X - Z\delta) \}.$$

Identifions alors les 4 blocs de la matrice des dérivées secondes ainsi que leurs espérances.

- (i) $l_{\delta^2}^{(2)} = {}^t Z \Sigma^{-1} Z$ est constante, égale à J_δ .
- (ii) La i -ième colonne de $l_{(\delta, \theta)}^{(2)}$ valant ${}^t Z \Sigma^i (Z\delta - X)$, $E(l_{(\delta, \theta)}^{(2)}) = 0$.
- (iii) $(l_{\theta^2}^{(2)})_{ij} = 2^{-1} \{ \text{tr}(\Sigma^{-1} \Sigma_{ij} + \Sigma^i \Sigma_j) + {}^t (X - Z\delta) \Sigma^{ij} (X - Z\delta) \}$; mais

$$\begin{aligned} E\{ {}^t (X - Z\delta) \Sigma^{ij} (X - Z\delta) \} &= \sum_{kl} \Sigma^{ij}(k, l) \text{cov}(X_k, X_l) \\ &= \sum_{kl} \Sigma^{ij}(k, l) \Sigma(k, l) = \text{tr}(\Sigma^{ij} \Sigma). \end{aligned}$$

On en déduit que $E(l_{\theta^2}^{(2)})_{ij} = 2^{-1} \text{tr}(\Sigma^{-1} \Sigma_{ij} + \Sigma^i \Sigma_j + \Sigma^{ij} \Sigma)$. Dérivant en $\theta_i \times \theta_j$ le produit $\Sigma^{-1} \Sigma \equiv I$, on obtient

$$\Sigma^{-1} \Sigma_{ij} + \Sigma^i \Sigma_j + \Sigma^j \Sigma_i + \Sigma^{ij} \Sigma = 0$$

et donc : $E(l_{\theta^2}^{(2)})_{ij} = -2^{-1} \text{tr}(\Sigma^j \Sigma_i)$.

□

C.4.2 Consistance du MV pour un champ de Markov stationnaire

Nous donnons ici la preuve de la consistance du MV pour un champ de Markov stationnaire sur \mathbb{Z}^d (cf. 5.4.1). La démonstration consiste à vérifier les conditions générales assurant la convergence d'un estimateur du minimum de contraste (cf. C.2). Pour cela, on utilise plusieurs propriétés d'un champ de Gibbs. La première dit que si $\mu = P_\theta \in \mathcal{G}_s(\pi_\theta)$ est stationnaire, alors μ s'exprime comme une combinaison linéaire convexe des éléments extrémaux μ^* de $\mathcal{G}_s(\pi_\theta)$, lois qui sont ergodiques [85, § 7.3 et § 14.2]. Si on montre la consistance pour chaque composante μ^* , on en déduira la μ -consistance du MV car les éléments extrémaux sont étrangers entre eux. Il suffira donc de démontrer la consistance pour une telle loi μ^* (qui est ergodique), c-à-d démontrer la consistance du MV si μ est stationnaire et ergodique. A cette fin, notant U_n le contraste égal à l'opposé de la log-vraisemblance de X sur D_n ,

$$U_n(x; \alpha) = \frac{1}{\#D_n} \{ \log Z_n(x_{\partial D_n}; \alpha) - H_n(x; \alpha) \},$$

on prouvera la consistance du MV en :

1. Identifiant la fonction de contraste : $K(\mu, \alpha) = \lim_n U_n(\alpha)$.
2. Vérifiant sa continuité en α et que $\alpha = \theta$ est son unique minimum.
3. Constatant que la condition sur le module de continuité de U_n est satisfaite.

Afin d'étudier le comportement limite de U_n , faisons la remarque préliminaire suivante : pour un potentiel général $\phi = (\phi_A, A \in S)$, définissons l'énergie moyenne par site par,

$$\bar{\phi}_i = \sum_{A:i \in A} \frac{\phi_A}{\#A}, \quad i \in S.$$

On vérifie que l'énergie $H_\Lambda(x) = \sum_{A:A \cap \Lambda \neq \emptyset} \phi_A(x)$ de x sur Λ vérifie :

$$H_\Lambda(x) = \sum_{i \in \Lambda} \bar{\phi}_i(x) + \varepsilon_\Lambda(x)$$

où

$$\varepsilon_\Lambda(x) = - \sum_{A:A \cap \Lambda \neq \emptyset \text{ et } A \not\subseteq \Lambda} \phi_A(x) \left\{ 1 - \frac{\#(A \cap \Lambda)}{\#A} \right\}.$$

Pour la spécification (5.23) engendrée par les $\{\Phi_{A_k}, k = 1, \dots, p\}$, on a :

$$|\varepsilon_\Lambda(x)| \leq \sum_{A:A \cap \Lambda \neq \emptyset \text{ et } A \not\subseteq \Lambda} |\phi_A(x)| \leq p \times (\#\partial\Lambda) \times \sup_k \|\Phi_{A_k}\|_\infty. \quad (\text{C.9})$$

En effet, les potentiels sont bornés et $\#\{A : \phi_A \neq 0, A \cap \Lambda \neq \emptyset \text{ et } A \not\subseteq \Lambda\} \leq p \times \#\partial\Lambda$. D'autre part, si ϕ est invariant par translation, $\bar{\phi}_i(x) \equiv \bar{\phi}_0(\tau_i(x))$. On en déduit que :

$$U_n(x; \alpha) = p_n(x; \alpha) - \frac{1}{\#D_n} \sum_{i \in D_n} \bar{\Phi}_i(x; \alpha) + \frac{1}{\#D_n} \varepsilon_{D_n}(x; \alpha), \quad (\text{C.10})$$

où $p_n(x; \alpha) = \frac{1}{\#D_n} \log Z_n(x; \alpha)$. Le premier terme de (C.10) tend vers la pression $p(\alpha)$ du potentiel ϕ_α , indépendante de x (cf. [85], §15.3). D'autre part, d'après la majoration (C.9), le troisième terme tend vers 0 uniformément en x puisque $\#\partial D_n / \#D_n \rightarrow 0$. Quant au deuxième terme, il tend vers $-E_\mu(\bar{\Phi}_{\{0\}}(\alpha))$ puisque μ est ergodique et que le potentiel est borné. On obtient donc :

$$U_n(x; \alpha) - U_n(x; \theta) \longrightarrow K(\mu, \alpha) = p(\alpha) - E_\mu(\bar{\Phi}_{\{0\}}(\alpha)) + h(\mu) \geq 0$$

où $h(\mu)$ est l'entropie spécifique de μ ([85], 15.4). La représentation $\alpha \mapsto \pi_{\{0\}, \alpha}$ étant propre, $\mathcal{G}(\pi_\theta) \cap \mathcal{G}(\pi_\alpha) = \emptyset$ si $\alpha \neq \theta$: le principe variationnel ([85], §15.4) dit alors que $K(\mu, \alpha) > 0$ si $\alpha \neq \theta$.

Reste à vérifier la continuité de $\alpha \mapsto K(\mu; \alpha)$ et la condition sur le module de continuité de U_n . Contrôlons d'abord le terme p_n . On a :

$$\begin{aligned} p_n(x; \alpha) - p_n(x; \beta) &= \frac{1}{\#D_n} \log \frac{\int_{E^{D_n}} \exp H_n(x; \alpha) \lambda_n(dx_{D_n})}{\int_{E^{D_n}} \exp H_n(x; \beta) \lambda_n(dx_{D_n})} \\ &= \frac{1}{\#D_n} \log \int_{E^{D_n}} \exp\{H_n(x; \alpha) - H_n(x; \beta)\} \pi_{D_n}(dx_{D_n}/x; \beta) \\ &\geq \frac{1}{\#D_n} E_{\pi_{D_n}(\cdot/x; \beta)}\{H_n(x; \alpha) - H_n(x; \beta)\} \quad (\text{Jensen}) \\ &= {}^t(\alpha - \beta) \bar{h}_n(x; \beta), \end{aligned}$$

où $\bar{h}_n(x; \beta) = {}^t(\bar{h}_{k,n}(x; \beta), k = 1, \dots, p)$, $\bar{h}_{k,n}(x; \beta) = \frac{1}{\#D_n} E_{\pi_{D_n}(\cdot/x; \beta)}\{h_{k,n}(x)\}$. On en déduit :

$${}^t(\alpha - \beta) \bar{h}_n(x; \beta) \leq p_n(x; \alpha) - p_n(x; \beta) \leq {}^t(\alpha - \beta) \bar{h}_n(x; \alpha).$$

Notant que si $a \leq u \leq b$, $|u| \leq \max\{|a|, |b|\}$ et que, pour tout $(x; \beta)$, $\bar{h}_{k,n}(x; \beta) \leq \|\Phi_k\|_\infty$, on a :

$$|p_n(x; \alpha) - p_n(x; \beta)| \leq \sum_{k=1}^p |\alpha_k - \beta_k| \|\Phi_k\|_\infty. \quad (\text{C.11})$$

La même majoration valant pour la limite $p(\alpha) - p(\beta)$, $K(\mu, \alpha)$ est continue en α . Quant à la condition sur le module de continuité de U_n , elle résulte de la continuité uniforme en (α, x) de U_n , qui est elle-même une conséquence de (C.11) et de l'écriture de l'énergie comme produit scalaire,

$$H_n(x; \alpha) - H_n(x; \beta) = {}^t(\alpha - \beta) h_n(x).$$

□

D

Logiciels

Nous avons utilisé trois logiciels pour effectuer les calculs illustrant les exemples d'application : *R*, *OpenBUGS* et *AntsInFields*.

R est un système d'analyse statistique et graphique créé par Ihaka et Gentleman [118]. C'est à la fois un logiciel et un langage issu du logiciel *S* créé par AT&T Bell Laboratories. Plus précisément, *R* est un langage *orienté-objet* interprété.

R est distribué librement sous les termes de la GNU General Public Licence (GPL) (voir www.r-project.org). Son développement et sa distribution sont assurés par le *R Development Core Team*. *R* est disponible sous plusieurs formes : le code est écrit principalement en C (et certaines programmes en Fortran) surtout pour les machines Unix et Linux ou des exécutables précompilés pour Windows, Macintosh et Alpha Unix. Le code et les exécutables sont distribués à partir du site internet cran.r-project.org du Comprehensive *R Archive Network* (*CRAN*).

Plusieurs manuels sont distribués avec *R* dans les répertoires du *CRAN*. Pour se faire une idée rapide de *R*, nous conseillons de lire “*R* pour les débutants” d’Emmanuel Paradis. Certaines fonctions pour l’analyse des données sont contenues dans le package de base mais la majorité des méthodes statistiques pour l’analyse de données spatiales par *R* sont distribuées sous forme de *packages* complémentaires.

Les *packages* constituent l’un des points forts de *R* : ce sont des recueils de fonctions développés le plus souvent par des statisticiens qui ont proposé la méthodologie correspondante. Pour nos exemples de données spatiales, nous avons utilisé les *packages* suivants : *geoR*, *RandomFields*, *spatstat*, *spdep*. Il existe d’autres *packages* : pour en savoir plus, on pourra consulter la page : <http://cran.r-project.org/web/views/Spatial.html>.

Donnons un court exemple de commandes afin de se faire une idée de la syntaxe de *R*. Considérons l'exemple (5.16) étudiant la répartition spatiale de la tulipe aquatique. On suppose que les données sont enregistrées dans un fichier *nyssa.txt*, les deux premières colonnes donnant les coordonnées spatiales et la troisième le genre (mâle ou femelle) de la tulipe. La lecture des fichiers de données structurées en colonnes, en format ASCII standard ou

CSV (texte délimité, largeur fixe, etc.) est réalisée par la fonction `read.table`. Ce qui se crée après l'application de la fonction est un "objet" `nyssa` de type tableau de données (*data.frame*). Cet objet est constitué d'une liste :

```
nyssa <- read.table("nyssa.txt", header = TRUE)
```

Chaque composant de la liste est l'équivalent d'un vecteur et les différents éléments de ces composants correspondent aux individus. Pour connaître les noms de ces composantes, il suffit d'invoquer la fonction `names` :

```
names(nyssa)
```

```
[1] "x"      "y"      "genre"
```

et pour une composante (par exemple la coordonnée x), on écrit `nyssa$x`.

L'utilisateur va agir sur un objet avec des opérateurs (arithmétiques, logiques et de comparaison) et des fonctions (qui sont elles-mêmes des objets) :

```
library(spatstat)
X <- ppp(nyssa$x, nyssa$y, marks = nyssa$genre,
        window = owin(xrange = c(-1, 53),
                      yrange = c(0, 53)))
plot(X)
```

La première ligne rend disponible le *package* `spatstat`. Le package `spatstat`, réalisé par Baddeley et Turner [15], contient les fonctions de manipulation, représentation et certaines fonctions statistiques pour l'analyse de données ponctuelles. La deuxième ligne crée un objet `X` de la classe `ppp` représentant une répartition spatiale bidimensionnelle dans une fenêtre $[-1, 53] \times [0, 53]$. Enfin, on applique à cet objet la fonction `plot` qui donne une représentation de la configuration (cf. Fig. D.1).

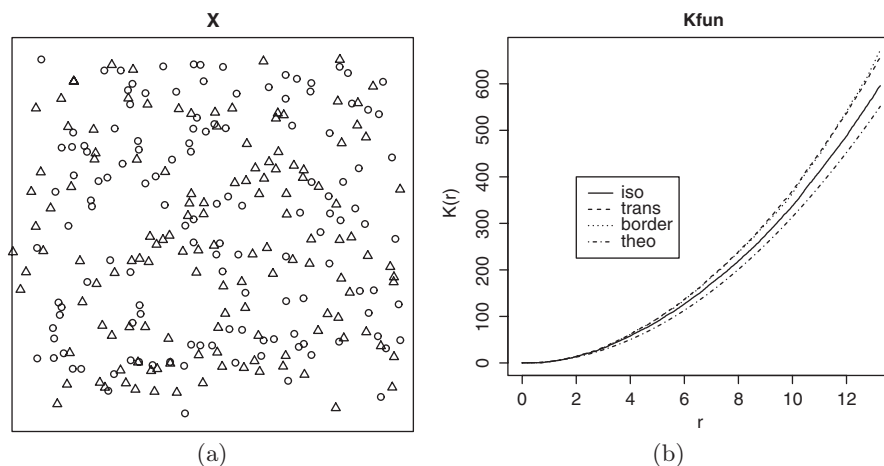


Fig. D.1. (a) Répartition spatiale de l'espèce *Nyssa aquatica* selon les caractères mâles (Δ) ou femelles (\circ); (b) fonction théorique K pour un PPP homogène et estimations pour différentes corrections de l'effet de bord.

Pour calculer la fonction K de Ripley, on utilise la fonction `Kest`. La souplesse de R est qu'une nouvelle application de la fonction `plot` à un objet différent donne une représentation de la fonction K . La dernière ligne précise la légende.

```
Kfun <- Kest(X)
plot(Kfun)
legend(2, 400, legend = c("iso", "trans", "border",
"theo"))
```

BUGS (*Bayesian inference Using Gibbs Sampling*) et sa version pour Windows *WinBUGS* [146] est un logiciel développé pour l'analyse statistique bayésienne en utilisant les méthodes de simulation *MCMC*, notamment l'échantillonneur de Gibbs. Nous avons utilisé la version à code ouvert *OpenBUGS* [212] (disponible sur mathstat.helsinki.fi/openbugs). Il est commode de spécifier et d'estimer un modèle hiérarchique spatial avec *BUGS* en utilisant le module *GeoBUGS*. A titre d'exemple, étudions les spécifications de modèle (5.52), (5.53) et (5.54) de l'exemple 5.23 sur le cancer du poumon en Toscane.

```
model {
  gamma[1:N] ~ car.normal(adj[], weights[], num[],
                          kappa)
  for (i in 1 : N) {
    alpha[i] ~ dnorm(0.0, tau)
    0[i] ~ dpois(mu[i])
    log(mu[i]) <- log(E[i]) + beta1 + gamma[i] +
                alpha[i]
    SMRhat[i] <- mu[i]/E[i]
  }
  beta1 ~ dnorm(0.0, 1.0E-5)
  tau ~ dgamma(0.5, 0.0005)
  kappa ~ dgamma(0.5, 0.0005)
}
```

Il existe aussi des *packages* de R , `BRugs` et `R2WinBUGS` qui fournissent des interfaces entre R et *BUGS*.

AntsInFields est un logiciel développé par Felix Friedrich pour la simulation et l'estimation des champs de Gibbs sur réseau ainsi que pour le traitement d'image. C'est à la fois un bon support didactique pour la simulation et l'estimation des champs de Gibbs et une aide à la recherche. Orienté-objet et modulaire, il est disponible au site www.antsinfields.de sous les termes de la GNU Less General Public Licence (LGPL). *AntsInFields* permet la simulation par échantillonneur de Gibbs, par dynamique de Metropolis-Hastings, la simulation exacte et l'optimisation par recuit simulé. Il réalise aussi l'estimation par PVC pour les modèles d'Ising, de Potts et les auto-modèles de Besag et est ouvert aux méthodes bayésiennes en reconstruction d'image.

Littérature

- [1] Aarts, E., Korst, J. : Simulated Annealing and Boltzman Machines : Stochastic Approach to Combinatorial and Neural Computing. Wiley, New York (1989)
- [2] Abramowitz, M., Stegun, I.A. (eds.) : Handbook of Mathematical Functions. Dover, New York (1970)
- [3] Adler, R.J. : The Geometry of Random Fields. Wiley, New York (1981)
- [4] Akaike, H. : Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics* **21**, 243–247 (1969)
- [5] Alfö, M., Postiglione, P. : Semiparametric modelling of spatial binary observations. *Statistical Modelling* **2**, 123–137 (2002)
- [6] Amemiya, T. : *Advanced Econometric*. Basil Blackwell, Oxford (1985)
- [7] Anselin, L. : *Spatial Econometrics : Methods and Models*. Kluwer, Dordrecht (1988)
- [8] Arnold, B.C., Castillo, E., Sarabia, J.M. : *Conditional Specification of Statistical Models*. Springer, New York (1999)
- [9] Augustin, N.H., McNicol, J.W., Marriott, C.A. : Using the truncated auto-poisson model for spatially correlated counts of vegetation. *Journal of Agricultural, Biological & Environmental Statistics* **11**, 1–23 (2006)
- [10] Azencott, R. (ed.) : *Simulated Annealing : Parallelization Techniques*. Wiley, New York (1992)
- [11] Baddeley, A.J., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (eds.) : *Case Studies in Spatial Point Processes Modeling*. Lecture Notes in Statistics 185. Springer, New York (2006)
- [12] Baddeley, A.J., Møller, J. : Nearest-neighbour Markov point processes and random sets. *International Statistical Review* **57**, 90–121 (1989)
- [13] Baddeley, A.J., Møller, J., Waagepetersen, R.P. : Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica* **54**(329–350) (2000)
- [14] Baddeley, A.J., Turner, R. : Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statistics* **42**, 283–322 (2000)

- [15] Baddeley, A.J., Turner, R. : Spatstat : an R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**, 1–42 (2005)
- [16] Baddeley, A.J., Turner, R., Møller, J., Hazelton, M. : Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society, Series B* **67**, 617–666 (2005)
- [17] Baddeley, A.J., van Lieshout, M.N.M. : Area-interaction point processes. *Annals of the Institute of Statistical Mathematics* **46**, 601–619 (1995)
- [18] Banerjee, S., Carlin, B.P., Gelfand, A.E. : Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall/CRC, Boca Raton : FL (2004)
- [19] Barker, A.A. : Monte Carlo calculations of the radial distribution functions for a proton-electron plasma. *Australian Journal of Physics* **18**, 119–133 (1965)
- [20] Bartlett, M.S. : Physical nearest-neighbour models and non-linear time series (I). *Journal of Applied Probability* **8**, 222–232 (1971)
- [21] Bartlett, M.S. : Physical nearest-neighbour models and non-linear time series (II). *Journal of Applied Probability* **9**, 76–86 (1972)
- [22] Bayomog, S., Guyon, X., Hardouin, C., Yao, J. : Test de différence de contraste et somme pondérée de Chi 2. *Canadian Journal of Statistics* **24**, 115–130 (1996)
- [23] Benveniste, A., Métivier, M., Priouret, P. : Adaptive Algorithms and Stochastic Approximations. Springer, New York (1990)
- [24] Besag, J. : On the correlation structure of some two dimensional stationary processes. *Biometrika* **59**, 43–48 (1972)
- [25] Besag, J. : Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B* **36**, 192–236 (1974)
- [26] Besag, J. : Efficiency of pseudo likelihood estimation for simple Gaussian fields. *Biometrika* **64**, 616–618 (1977)
- [27] Besag, J., Moran, P.A.P. : On the estimation and testing of spatial interaction for Gaussian lattice processes. *Biometrika* **62**, 555–562 (1975)
- [28] Besag, J., York, J., Mollié, A. : Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* **43**, 1–59 (1991)
- [29] Bochner, N. : Lectures on Fourier Integrals. Princeton University Press, Princeton : NJ (1959)
- [30] Bolthausen, E. : On the central limit theorem for stationary mixing random fields. *Annals of Probability* **10**, 1047–1050 (1982)
- [31] Bouthemy, P., Hardouin, C., Piriou, G., Yao, J. : Mixed-state auto-models and motion texture modeling. *Journal of Mathematical Imaging and Vision* **25**, 387–402 (2006)
- [32] Breiman, L. : Probability. SIAM Classics in Applied Mathematics 7, Philadelphia : PA (1992)

- [33] Brillinger, D.R. : Estimation of the second-order intensities of a bivariate stationary point process. *Journal of the Royal Statistical Society, Series B* **38**, 60–66 (1976)
- [34] Brockwell, P.J., Davis, R.A. : *Time Series Analysis : Theory and Methods*. Springer, New York (1992)
- [35] Brook, D. : On the distinction between the conditional probability and joint probability approaches in the specification of nearest neighbour systems. *Biometrika* **51**, 481–483 (1964)
- [36] Brown, P.E., Kåresen, K.F., Roberts, G.O., Tonellato, S. : Blur-generated non-separable space-time models. *Journal of the Royal Statistical Society, Series B* **62**, 847–860 (2000)
- [37] Cairoli, R., Walsh, J.B. : Stochastic integral in the plane. *Acta Mathematica* **134**, 111–183 (1975)
- [38] Casson, E., Coles, S.G. : Spatial regression models for extremes. *Extremes* **1**, 449–468 (1999)
- [39] Catelan, D., Biggeri, A., Dreassi, E., Lagazio, C. : Space-cohort Bayesian models in ecological studies. *Statistical Modelling* **6**, 159–173 (2006)
- [40] Catoni, O. : Rough large deviation estimates for simulated annealing : application to exponential schedules. *Annals of Probability* **20**, 1109–1146 (1992)
- [41] Chadoeuf, J., Nandris, D., Geiger, J., Nicole, M., Pierrat, J.C. : Modélisation spatio-temporelle d’une épidémie par un processus de Gibbs : estimation et tests. *Biometrics* **48**, 1165–1175 (1992)
- [42] Chalmond, B. : *Éléments de modélisation pour l’analyse d’image*. Springer, Paris (2000)
- [43] Chilès, J.P., Delfiner, P. : *Geostatistics*. Wiley, New York (1999)
- [44] Christensen, O.F., Roberts, G.O., Sköld, M. : Robust Markov chain Monte Carlo methods for spatial generalised linear mixed models. *Journal of Computational and Graphical Statistics* **15**, 1–17 (2006)
- [45] Cliff, A.D., Ord, J.K. : *Spatial Processes : Models and Applications*. Pion, London (1981)
- [46] Comets, F. : On consistency of a class of estimators for exponential families of Markov random fields on a lattice. *Annals of Statistics* **20**, 455–468 (1992)
- [47] Comets, F., Janzura, M. : A central limit theorem for conditionally centered random fields with an application to Markov fields. *Journal of Applied Probability* **35**, 608–621 (1998)
- [48] Cressie, N.A.C. : *Statistics for Spatial Data*, 2nd edn. Wiley, New York (1993)
- [49] Cressie, N.A.C., Hawkins, D.M. : Robust estimation of the variogram, I. *Journal of the International Association of Mathematical Geology* **12**, 115–125 (1980)
- [50] Cressie, N.A.C., Huang, H.C. : Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association* **94**, 1330–1340 (1999)

- [51] Cross, G.R., Jain, A.K. : Markov field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **5**, 155–169 (1983)
- [52] Csiszar, I., Talata, Z. : Consistent estimation of the basic neighborhood of Markov random fields. *Annals of Statistics* **34**, 123–145 (2006)
- [53] Cuzick, J., Edwards, R. : Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society, Series B* **52**, 73–104 (1990)
- [54] Dacunha-Castelle, D., Duflo, M. : *Probabilités et Statistiques, Tome 2 : Problèmes à temps mobile*. Masson, Paris (1993)
- [55] Dacunha-Castelle, D., Duflo, M. : *Probabilités et Statistiques. Tome 1 : Problèmes à temps fixe*. Masson, Paris (1994)
- [56] Daley, D., Vere-Jones, D. : *An Introduction to the Theory of Point Processes, Vol. I, Elementary Theory and Methods*, 2nd edn. Springer, New York (2003)
- [57] Dalhaus, R., Künsch, H.R. : Edge effect and efficient parameter estimation for stationary random fields. *Biometrika* **74**, 877–882 (1987)
- [58] De Iaco, S., Myers, D.E., Posa, T. : Nonseparable space-time covariance models : some parametric families. *Mathematical Geology* **34**, 23–42 (2002)
- [59] Devroye, L. : *Non Uniform Random Variable Generation*. Springer, New York (1986)
- [60] Diaconis, P., Freedman, D. : Iterated random functions. *SIAM Review* **41**, 45–76 (1999)
- [61] Diaconis, P., Graham, R., Morrison, J. : Asymptotic analysis of a random walk on an hypercube with many dimensions. *Random Structure Algorithms* **1**, 51–72 (1990)
- [62] Diggle, P.J. : *Statistical Analysis of Spatial Point Patterns*. Oxford University Press, Oxford (2003)
- [63] Diggle, P.J., Ribeiro, P.J. : *Model-based Geostatistics*. Springer, New York (2007)
- [64] Diggle, P.J., Tawn, J.A., Moyeed, R.A. : Model-based geostatistics (with discussion). *Applied Statistics* **47**, 299–350 (1998)
- [65] Dobrushin, R.L. : Central limit theorems for non stationary Markov chains I, II. *Theory of Probability and its Applications* **1**, 65–80, 329–383 (1956)
- [66] Dobrushin, R.L. : The description of a random field by means of conditional probabilities and condition of its regularity. *Theory of Probability and its Applications* **13**, 197–224 (1968)
- [67] Doukhan, P. : *Mixing : Properties and Examples*. Lecture Notes in Statistics 85. Springer, Berlin (1994)
- [68] Dreesbeke, J.J., Fine, J., Saporta, G. (eds.) : *Méthodes bayésiennes en statistique*. Technip, Paris (2002)
- [69] Dreesbeke, J.J., Lejeune, M., Saporta, G. (eds.) : *Analyse statistique des données spatiales*. Technip, Paris (2006)

- [70] Dubois, G., Malczewski, J., De Cort, M. : Spatial interpolation comparison 1997. *Journal of Geographic Information and Decision Analysis* **2** (1998)
- [71] Duflo, M. : Algorithmes stochastiques. Mathématiques et applications. Springer, Paris (1996)
- [72] Duflo, M. : Random Iterative Models. Springer, New York (1997)
- [73] Durrett, R. : Ten lectures on particle systems. In : P. Bernard (ed.) *École d'Été de St. Flour XXIII, Lecture Notes in Mathematics* 1608, pp. 97–201. Springer, New York (1995)
- [74] Durrett, R., Levin, S.A. : Stochastic spatial models : a user's guide to ecological applications. *Philosophical Transactions of the Royal Society of London, series B* **343**, 329–350 (1994)
- [75] Dzhaparidze, K.O. : On simplified estimators of unknown parameters with good asymptotic properties. *Theory of Probability and its Applications* **19**, 347–358 (1974)
- [76] Pfeifer, P.E., Deutsch, S.J. : Identification and interpretation of first order space-time arma models. *Technometrics* **22**, 397–408 (1980)
- [77] Eriksson, M., Siska, P.P. : Understanding anisotropy computations. *Mathematical Geology* **32**, 683–700 (2000)
- [78] Ferrandiz, J., Lopez, A., Llopis, A., Morales, M., Tejerizo, M.L. : Spatial interaction between neighbouring counties : cancer mortality data in Valencia (Spain). *Biometrics* **51**, 665–678 (1995)
- [79] Fuentes, M. : Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association* **102**, 321–331 (2007)
- [80] Gelman, A., Rubin, D.B. : Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–511 (1992)
- [81] Geman, D. : Random fields and inverse problem in imaging. In : P.L. Hennequin (ed.) *École d' Eté de Probabilités de Saint-Flour XVIII, Lecture Notes in Mathematics* 1427, pp. 113–193. Springer, New York (1990)
- [82] Geman, D., Geman, S. : Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741 (1984)
- [83] Geman, S., Graffigne, C. : Markov random fields models and their applications to computer vision. In : A.M. Gleason (ed.) *Proceedings of the International Congress of Mathematicians 1986*, pp. 1496–1517. American Mathematical Society, Providence : RI (1987)
- [84] Georgii, H.O. : Canonical and grand canonical Gibbs states for continuum systems. *Communications of Mathematical Physics* **48**, 31–51 (1976)
- [85] Georgii, H.O. : Gibbs measure and phase transitions. De Gruyter, Berlin (1988)

- [86] Geyer, C.J. : On the convergence of Monte Carlo maximum likelihood calculations. *Journal of the Royal Statistical Society, Series B* **56**, 261–274 (1994)
- [87] Geyer, C.J. (1999). Likelihood inference for spatial point processes. In *Stochastic geometry : likelihood and computation* (eds O.E. Barndorff-Nielsen, W.S. Kendall & M.N.M. van Lieshout), 79–140, Chapman & Hall/CRC, Florida
- [88] Geyer, C.J., Møller, J. : Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics* **21**, 359–373 (1994)
- [89] Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (eds.) : *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London (1996)
- [90] Gneiting, T. : Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association* **97**, 590–600 (2002)
- [91] Gneiting, T., Genton, M.G., Guttorp, P. : Geostatistical space-time models, stationarity, separability and full symmetry. In : B. Finkenstadt, L. Held, V. Isham (eds.) *Statistical Methods for Spatio-Temporal Systems*, pp. 151–175. Chapman & Hall/CRC, Boca Raton : FL (2007)
- [92] Gourieroux, C., Monfort, A. : *Statistiques et modèles économétriques*, Tomes 1 et 2. Economica, Paris (1992)
- [93] Green, P.J., Richardson, S. : Hidden Markov models and disease mapping. *Journal of the American Statistical Association* **97**, 1055–1070 (2002)
- [94] Greig, D.M., Porteous, B.T., Seheult, A.H. : Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B* **51**, 271–279 (1989)
- [95] Guan, Y., Sherman, M. : On least squares fitting for stationary spatial point processes. *Journal of the Royal Statistical Society, Series B* **69**, 31–49 (2007)
- [96] Guyon, X. : *Random Fields on a Network : Modeling, Statistics and Applications*. Springer, New York (1995)
- [97] Guyon, X., Hardouin, C. : The Chi-2 difference of coding test for testing Markov random field hypothesis. In : P. Barone, A. Frigessi, M. Piccioni (eds.) *Stochastic Models, Statistical Methods and Algorithms in Image Analysis, Lecture Notes in Statistics* 74, pp. 165–176. Springer, Berlin (1992)
- [98] Guyon, X., Hardouin, C. : Markov chain Markov field dynamics : models and statistics. *Statistics* **36**, 339–363 (2002)
- [99] Guyon, X., Künsch, H.R. : Asymptotic comparison of estimator of the Ising model. In : P. Barone, A. Frigessi, M. Piccioni (eds.) *Stochastic Models, Statistical Methods and Algorithms in Image Analysis, Lecture Notes in Statistics* 74, pp. 177–198. Springer, Berlin (1992)

- [100] Guyon, X., Pumo, B. : Estimation spatio-temporelle d'un modèle de système de particule. *Comptes rendus de l'Académie des sciences Paris* **I-340**, 619–622 (2005)
- [101] Guyon, X., Pumo, B. : Space-time estimation of a particle system model. *Statistics* **41**, 395–407 (2007)
- [102] Guyon, X., Yao, J.F. : On the underfitting and overfitting sets of models chosen by order selection criteria. *Journal of Multivariate Analysis* **70**, 221–249 (1999)
- [103] Häggström, O. : *Finite Markov Chains and Algorithmic Applications*. Cambridge University Press, Cambridge (2002)
- [104] Häggström, O., van Lieshout, M.N.M., Møller, J. : Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes. *Bernoulli* **5**, 641–658 (1999)
- [105] Haining, R. : *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge (1990)
- [106] Hajek, B. : Cooling schedules for optimal annealing. *Mathematics of Operations Research* **13**, 311–329 (1999)
- [107] Hannan, E.J. : The estimation of the order of an ARMA process. *Annals of Statistics* **8**, 1071–1081 (1980)
- [108] Hardouin, C. : Quelques résultats nouveaux en statistique des processus : contraste fort, régressions à rélog-périodogramme. Ph.D. thesis, Université Paris VII, Paris (1992)
- [109] Hardouin, C., Yao, J. : Multi-parameter auto-models and their application. *Biometrika* (2008). Forthcoming
- [110] Hastings, W. : Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109 (1970)
- [111] Heinrich, L. : Minimum contrast estimates for parameters of spatial ergodic point processes. In : *Transactions of the 11th Prague Conference on Random Processes, Information Theory and Statistical Decision Functions*, pp. 479–492. Academic Publishing House, Prague (1992)
- [112] Higdon, D. : Space and space-time modeling using process convolutions. In : C. Anderson, V. Barnett, P.C. Chatwin, A. El-Shaarawi (eds.) *Quantitative Methods for Current Environmental Issues*, pp. 37–56. Springer, London (2002)
- [113] Higdon, D.M., Swall, J., Kern, J. : Non-stationary spatial modeling. In : J.M. Bernardo, J.O. Berger, A.P. Dawid, A.F.M. Smith (eds.) *Bayesian Statistics 6*, pp. 761–768. Oxford University Press, Oxford (1999)
- [114] Hoeting, A., Davis, A., Merton, A., Thompson, S. : Model selection for geostatistical models. *Ecological Applications* **16**, 87–98 (2006)
- [115] Huang, F., Ogata, Y. : Improvements of the maximum pseudo-likelihood estimators in various spatial statistical models. *Journal of Computational and Graphical Statistics* **8**, 510–530 (1999)
- [116] Ibragimov, I.A., Linnik, Y.V. : *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen (1971)

- [117] Ibragimov, I.A., Rozanov, Y.A. : Processus aléatoires gaussiens. MIR, Moscou (1974)
- [118] Ihaka, R., Gentleman, R. : R : a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**, 299–314 (1996)
- [119] Illig, A. : Une modélisation de données spatio-temporelles par AR spatiaux. *Journal de la société française de statistique* **147**, 47–64 (2006)
- [120] Isaacson, D.L., Madsen, R.Q. : *Markov Chains : Theory and Application*. Wiley, New York (1976)
- [121] Jensen, J., Møller, J. : Pseudolikelihood for exponential family of spatial point processes. *Annals of Applied Probability* **3**, 445–461 (1991)
- [122] Jensen, J.L. : Asymptotic normality of estimates in spatial point processes. *Scandinavian Journal of Statistics* **20**, 97–109 (1993)
- [123] Jensen, J.L., Künsch, H.R. : On asymptotic normality of pseudolikelihood estimates for pairwise interaction processes. *Annals of the Institute of Statistical Mathematics* **46**, 475–486 (1994)
- [124] Ji, C., Seymour, L. : A consistent model selection procedure for Markov random fields based on penalized pseudo-likelihood. *Annals of Applied Probability* **6**, 423–443 (1996)
- [125] Jolivet, E. : Central limit theorem and convergence of empirical processes for stationary point processes. In : P. Bastfai, J. Tomko (eds.) *Point Processes and Queuing Problems*, pp. 117–161. North-Holland, Amsterdam (1978)
- [126] Jones, R., Zhang, Y. : Models for continuous stationary space-time processes. In : T.G. Gregoire, D.R. Brillinger, P.J. Diggle, E. Russek-Cohen, W.G. Warren, R.D. Wolfinger (eds.) *Modelling Longitudinal and Spatially Correlated Data*, *Lecture Notes in Statistics* 122, pp. 289–298. Springer, New York (1997)
- [127] Kaiser, M.S., Cressie, N.A.C. : Modeling Poisson variables with positive spatial dependence. *Statistics and Probability Letters* **35**, 423–432 (1997)
- [128] Keilson, J. : *Markov Chain Models : Rarity and Exponentiality*. Springer, New York (1979)
- [129] Kemeny, G., Snell, J.L. : *Finite Markov Chains*. Van Nostrand, Princeton : NJ (1960)
- [130] Kendall, W.S., Møller, J. : Perfect simulation using dominating processes on ordered state spaces, with application to locally stable point processes. *Advances in Applied Probability* **32**, 844–865 (2000)
- [131] Klein, D. : Dobrushin uniqueness techniques and the decay of correlation in continuum statistical mechanics. *Communications in Mathematical Physics* **86**, 227–246 (1982)
- [132] Koehler, J.B., Owen, A.B. : Computer experiments. In : S. Ghosh, C.R. Rao (eds.) *Handbook of Statistics*, Vol 13, pp. 261–308. North-Holland, New York (1996)
- [133] Kolovos, A., Christakos, G., Hristopoulos, D.T., Serre, M.L. : Methods for generating non-separable spatiotemporal covariance models with po-

- tential environmental applications. *Advances in Water Resources* **27**, 815–830 (2004)
- [134] Krige, D. : A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa* **52**, 119–139 (1951)
- [135] Kutoyants, Y.A. : *Statistical Inference for Spatial Poisson Processes*. Springer, New York (1998)
- [136] Kyriakidis, P.C., Journel, A.G. : Geostatistical space-time models : a review. *Mathematical Geology* **31**, 651–684 (1999)
- [137] Lahiri, S.N. : CLT for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhya A* **65**, 356–388 (2003)
- [138] Lahiri, S.N., Lee, Y., Cressie, C.N.A. : On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters. *Journal Statistical Planning and Inference* **103**, 65–85 (2002)
- [139] Lantuèjoul, C. : *Geostatistical Simulation*. Springer, Berlin (2002)
- [140] Laslett, M. : Kriging and splines : and empirical comparison of their predictive performance in some applications. *Journal of the American Statistical Association* **89**, 391–409 (1994)
- [141] Lawson, A.B. : *Statistical Methods in Spatial Epidemiology*. Wiley, New York (2001)
- [142] Le, N.D., Zidek, J.V. : *Statistical Analysis of Environmental Space-Time Processes*. Springer, New York (2006)
- [143] Lee, H.K., Higdon, D.M., Calder, C.A., Holloman, C.H. : Efficient models for correlated data via convolutions of intrinsic processes. *Statistical Modelling* **5**, 53–74 (2005)
- [144] Lee, Y.D., Lahiri, S.N. : Least square variogram fitting by spatial subsampling. *Journal of the Royal Statistical Society, Series B* **64**, 837–854 (2002)
- [145] Loève, M. : *Probability Theory II*. Springer, New York (1978)
- [146] Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D.J. : WinBUGS - a Bayesian modelling framework : concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337 (2000)
- [147] Ma, C. : Families of spatio-temporal stationary covariance models. *Journal of Statistical Planning and Inference* **116**, 489–501 (2003)
- [148] Mardia, K.V., Goodall, C., Redfern, E.J., Alonso, F.J. : The Kriged Kalman filter (with discussion). *Test* **7**, 217–252 (1998)
- [149] Mardia, K.V., Marshall, J. : Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71**, 289–295 (1984)
- [150] Marroquin, J., Mitter, S., Poggio, T. : Probabilistic solution of ill posed problem in computational vision. *Journal of the American Statistical Association* **82**, 76–89 (1987)
- [151] Mase, S. : Marked Gibbs processes and asymptotic normality of maximum pseudo-likelihood estimators. *Mathematische Nachrichten* **209**, 151–169 (1999)

- [152] Matheron, G. : *Traité de géostatistique appliquée*, Tome 1. Mémoires du BRGM, n. 14. Technip, Paris (1962)
- [153] Matheron, G. : The intrinsic random function and their applications. *Advances in Applied Probability* **5**, 439–468 (1973)
- [154] Matérn, B. : *Spatial Variation : Stochastic Models and their Applications to Some Problems in Forest Surveys and Other Sampling Investigations*, 2nd edn. Springer, Heidelberg (1986)
- [155] McCullagh, P., Nelder, J.A. : *Generalized Linear Models*. Chapman & Hall, London (1989)
- [156] Mercer, W.B., Hall, A.D. : The experimental error of field trials. *The experimental error of field trials* **4**, 107–132 (1973)
- [157] Meyn, S.P., Tweedie, R.L. : *Markov Chains and Stochastic Stability*. Springer, New York (1993)
- [158] Mitchell, T., Morris, M., Ylvisaker, D. : Existence of smoothed process on an interval. *Stochastic Processes and their Applications* **35**, 109–119 (1990)
- [159] Møller, J., Syversveen, A.R., Waagepetersen, R.P. : Log-gaussian Cox processes. *Scandinavian Journal of Statistics* **25**, 451–82 (1998)
- [160] Møller, J., Waagepetersen, R.P. : *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton : FL (2004)
- [161] Møller, J., Waagepetersen, R.P. : Modern statistics for spatial point processes. *Scandinavian Journal of Statistics* **34**, 643–684 (2007)
- [162] Moyeed, R.A., Baddeley, A.J. : Stochastic approximation of the MLE for a spatial point pattern. *Scandinavian Journal of Statistics* **18**, 39–50 (1991)
- [163] Neyman, J., Scott, E.L. : Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society, Series B* **20**, 1–43 (1958)
- [164] Nguyen, X.X., Zessin, H. : Ergodic theorems for spatial processes. *Probability Theory and Related Fields* **48**, 133–158 (1979)
- [165] Nguyen, X.X., Zessin, H. : Integral and differential characterization of the Gibbs process. *Mathematische Nachrichten* **88**, 105–115 (1979)
- [166] Onsager, L. : Crystal statistics I : A two dimensional model with order-disorder transition. *Physical Review* **65**, 117–149 (1944)
- [167] Ord, J.K. : Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* **70**, 120–126 (1975)
- [168] Papangelou, F. : The conditional intensity of general point processes and application to line processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **28**, 207–227 (1974)
- [169] Penttinen, A. : Modelling interaction in spatial point patterns : parameter estimation by the maximum likelihood method. *Jyväskylän Studies in Computer Science, Economics and Statistics* **7** (1984)
- [170] Perrin, O., Meiring, W. : Identifiability for non-stationary spatial structure. *Journal of Applied Probability* **36**, 1244–1250 (1999)

- [171] Perrin, O., Senoussi, R. : Reducing non-stationary random fields to stationary and isotropy using space deformation. *Statistics and Probability Letters* **48**, 23–32 (2000)
- [172] Peskun, P. : Optimum Monte Carlo sampling using Markov chains. *Biometrika* **60**, 607–612 (1973)
- [173] Peyrard, N., Calonnec, A., Bonnot, F., Chadoeuf, J. : Explorer un jeu de données sur grille par test de permutation. *Revue de Statistique Appliquée* **LIII**, 59–78 (2005)
- [174] Pfeifer, P.E., Deutsch, S.J. : A three-stage iterative procedure for space-time modeling. *Technometrics* **22**, 93–117 (1980)
- [175] Ploner, A. : The use of the variogram cloud in geostatistical modelling. *Environmentrics* **10**, 413–437 (1999)
- [176] Preston, C. : *Random Fields. Lecture Notes in Mathematics* 534. Springer, Berlin (1976)
- [177] Propp, J.G., Wilson, D.B. : Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms* **9**, 223–252 (1996)
- [178] R Development Core Team : *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2007). URL <http://www.R-project.org>
- [179] Rathbun, S.L., Cressie, N.A.C. : Asymptotic properties of estimators for the parameters of spatial inhomogeneous Poisson point processes. *Advances in Applied Probability* **26**, 122–154 (1994)
- [180] Revuz, D. : *Probabilités*. Herman, Paris (1997)
- [181] Ribeiro, P., Diggle, P.J. : geoR : a package for geostatistical analysis. *R-NEWS* **1**, 14–18 (2001)
- [182] Richardson, S., Guihenenc, C., Lasserre, V. : Spatial linear models with autocorrelated error structure. *The Statistician* **41**, 539–557 (1992)
- [183] Ripley, B.D. : The second-order analysis of stationary point processes. *Journal of Applied Probability* **13**, 255–266 (1976)
- [184] Ripley, B.D. : *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge (1988)
- [185] Ripley, B.D. : *Spatial Statistics*. Wiley, New York (1991)
- [186] Ripley, B.D., Kelly, F.P. : Markov point processes. *Journal of the London Mathematical Society* **15**, 188–192 (1977)
- [187] Robert, C.P. : *L'analyse statistique bayésienne*. Economica, Paris (1992)
- [188] Robert, C.P., Casella, G. : *Monte-Carlo Statistical Methods*. Springer, New York (1999)
- [189] Rue, H., Held, L. : *Gaussian Markov Random Fields : Theory and Applications*. Chapman & Hall/CRC, Boca Raton : FL (2005)
- [190] Ruelle, D. : *Statistical Mechanics*. Benjamin, New York (1969)
- [191] Saloff-Coste, L. : Lectures on finite Markov chains. In : P. Bernard (ed.) *Lectures on Probability Theory and Statistics. Ecole d'été de Probabilité de St. Flour XXVI, Lecture Notes in Mathematics* 1665, pp. 301–408. Springer (1997)

- [192] Sampson, P., Guttorp, P. : Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association* **87**, 108–119 (1992)
- [193] Santer, T.J., Williams, B.J., Notz, W.I. : *The Design and Analysis of Computers Experiments*. Springer, New York (2003)
- [194] Schabenberger, O., Gotway, C.A. : *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC, Boca Raton : FL (2004)
- [195] Schlather, M. : Introduction to positive definite functions and to unconditional simulation of random fields. Tech. Rep. ST 99-10, Lancaster University, Lancaster (1999)
- [196] Senoussi, R. : Statistique asymptotique presque sûre des modèles statistiques convexes. *Annales de l'Institut Henri Poincaré* **26**, 19–44 (1990)
- [197] Serra, J. : *Image Analysis and Mathematical Morphology*. Academic Press, New York (1982)
- [198] Shea, M.M., Dixon, P.M., Sharitz, R.R. : Size differences, sex ratio, and spatial distribution of male and female water tupelo, *nyssa aquatica* (nyssaceae). *American Journal of Botany* **80**, 26–30 (1993)
- [199] Shibata, R. : Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika* **63**, 117–126 (1976)
- [200] Stein, M.L. : *Interpolation of Spatial Data : Some Theory for Kriging*. Springer, New York (1999)
- [201] Stein, M.L. : Statistical methods for regular monitoring data. *Journal of the Royal Statistical Society, Series B* **67**, 667–687 (2005)
- [202] Storvik, G., Frigessi, A., Hirst, D. : Stationary space-time gaussian fields and their time autoregressive representation. *Stochastic Modelling* **2**, 139–161 (2002)
- [203] Stoyan, D., Grabarnik, P. : Second-order characteristics for stochastic structures connected with Gibbs point processes. *Mathematische Nachrichten* **151**, 95–100 (1991)
- [204] Stoyan, D., Kendall, W.S., Mecke, J. (eds.) : *Stochastic Geometry and its Applications*, 2nd edn. Wiley, New York (1995)
- [205] Strathford, J.A., Robinson, W.D. : Distribution of neotropical migratory bird species across an urbanizing landscape. *Urban Ecosystems* **8**, 59–77 (2005)
- [206] Strauss, D.J. : A model for clustering. *Biometrika* **62**, 467–475 (1975)
- [207] Strauss, D.J. : Clustering on colored lattice. *Journal of Applied Probability* **14**, 135–143 (1977)
- [208] Stroud, J.R., Müller, P., Sansó, B. : Dynamic models for spatio-temporal data. *Journal of the Royal Statistical Society, Series B* **63**, 673–689 (2001)
- [209] Sturtz, S., Ligges, U., Gelman, A. : R2WinBUGS : a package for running WinBUGS from R. *Journal of Statistical Software* **2**, 1–16 (2005)
- [210] Sweeting, T.J. : Uniform asymptotic normality of the maximum likelihood estimator. *Annals of Statistics* **8**, 1375–1381 (1980)

- [211] Tempelman, A.A. : Ergodic theorems for general dynamical systems. *Transactions of the Moscow Mathematical Society* **26**, 94–132 (1972)
- [212] Thomas, A., O’ Hara, B., Ligges, U., Sturtz, S. : Making BUGS open. *R News* **6**, 12–17 (2006)
- [213] Thomas, M. : A generalisation of Poisson’s binomial limit for use in ecology. *Biometrika* **36**, 18–25 (1949)
- [214] Tierney, L. : Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics* **22**, 1701–1762 (1994)
- [215] Tierney, L. : A note on Metropolis-Hastings kernels for general state space. *Annals of Applied Probability* **3**, 1–9 (1998)
- [216] Tukey, J.W. : *Spectral Analysis Time Series*. Wiley, New York (1967)
- [217] Van Lieshout, M.N.M. : *Markov Point Processes and their Applications*. Imperial College Press, London (2000)
- [218] Van Lieshout, M.N.M., Baddeley, A.J. : Indices of dependence between types in multivariate point patterns. *Scandinavian Journal of Statistics* **26**, 511–532 (1999)
- [219] Ver Hoef, J., Barry, R.P. : Constructing and fitting models for cokriging and multivariable spatial prediction. *Journal of Statistical Planning and Inference* **69**, 275–294 (1998)
- [220] Waagepetersen, R.P. : An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics* **63**, 252–258 (2007)
- [221] Wackernagel, H. : *Multivariate Geostatistics : An Introduction with Applications*, 3rd edn. Springer, New York (2003)
- [222] Whittle, P. : On stationary processes in the plane. *Biometrika* **41**, 434–449 (1954)
- [223] Wikle, C.K., Cressie, N.A.C. : A dimension-reduced approach to space-time Kalman filtering. *Biometrika* **86**, 815–829 (1999)
- [224] Winkler, G. : *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*, 2nd edn. Springer, Berlin (2003)
- [225] Wolpert, R.L., Ickstadt, K. : Poisson/Gamma random fields models for spatial statistics. *Biometrika* **85**, 251–267 (1998)
- [226] Wu, H., Huffer, F.W. : Modelling the distribution of plant species using the autologistic regression model. *Environmental and Ecological Statistics* **4**, 49–64 (1997)
- [227] Yaglom, A.M. : *Correlation Theory of Stationary and Related Random Functions. Volume I : Basic Results*. Springer, New York (1987)
- [228] Yao, J.F. : On constrained simulation and optimisation by Metropolis chains. *Statistics and Probability Letters* **46**, 187–193 (2000)
- [229] Ycart, B. : *Modèles et algorithmes markoviens. Mathématiques et Applications*. Springer, Paris (2002)
- [230] Younes, L. : Estimation and annealing for Gibbsian fields. *Annales de l’Institut Henri Poincaré (B). Probabilités et Statistiques* **2**, 269–294 (1988)

- [231] Zhang, H., Zimmerman, D.L. : Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika* **92**, 921–936 (2005)
- [232] Zimmerman, D., Zimmerman, M. : A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. *Technometric* **33**, 77–91 (1991)

Index

- abréviations, liste de, XV
- Admissible
 - densité, 94, 100
 - potentiel, énergie, 58, 76, 96
- Algorithme
 - de Metropolis, 123, 239
 - de Metropolis inhomogène, 130
 - de Monte Carlo monotone, 141
 - de Propp et Wilson, 138
 - de recuit simulé, 125, 130
 - markovien, 120
- AMM, 69, 70, 128
- AMMX : AMM avec covariables, 73
- Analyse de la variance, 39, 59, 184
- Analyse variographique, 154
- Anisotropie
 - d'un variogramme, 14, 156
 - de texture, 61
 - géométrique, stratifiée, 14
- Apériodique (chaîne), 117, 120, 123
- Approximation Monte Carlo
 - d'un quantile, 113, 220
 - d'une vraisemblance, 228
- AR, 28
 - conditionnelle (CAR), 31, 182
 - estimation d'une, 181
 - factorisante, 30, 48
 - simultanée (SAR), 28, 181
- ARMA, 25
 - non-stationnaire, 34
 - stationnaire, 26
- Asymptotique, 224
 - du minimum de contraste, 275
 - du MV d'un champ de Markov, 195
 - du MV d'un PP, 230
 - du MV pour un PPP, 214
 - en géostatistique, 161
 - extensive, 153
 - intensive, 153, 162
 - mixte, 162
 - pour l'indice de Moran, 172
 - pour la PVC d'un champ de Markov, 198
 - pour la PVC d'un PP, 227
 - pour un CAR gaussien, 183
 - pour un champ stationnaire, 180
 - pour un régression spatiale, 187
 - pour une spécification invariante, 201
- Attraction entre points d'un PP, 89
- Auto-modèle, 69
 - à état mixte, 73
 - avec covariables, 73
 - binomial, 71
 - exponentiel, 72
 - gaussien, 73
 - logistique, 70, 247
 - poissonnien, 71
- Autosimilaire (variogramme), 13
- Balayage
 - séquentiel, aléatoire, 121
- bandes tournantes (Méthode des), 143
- Bayésienne
 - imagerie, 63
 - reconstruction, 63
 - restauration, 80
 - statistique, 235

- Bootstrap paramétrique, 163, 203, 224
- Brownien (mouvement, drap), 3, 8
- Bruit
 - blanc, 5, 186
 - coloré, 31
- CAR, 31, 32
 - estimation, 182, 201, 246
 - gaussien markovien, 36
 - identification d'un, 181
 - intrinsèque, 242
 - modèle de Gibbs, 73
 - non-stationnaire, 36
- Carte de prédiction, 43, 165
- Centrage conditionnel d'un champ (condition), 267
- CFTP : coupling from the past, 138
- Chaîne de Markov, 81, 115, 146
 - apériodique, 117
 - de Champ de Markov, 76
 - inhomogène, 130, 252
 - irréductible, 116
 - réversible, 118
 - transition d'une, 115
- Champ de Gibbs, 57, 79, 96
- Champ de Markov, 66, 79, 81
 - dynamique d'un, 76
- Champ gaussien markovien, 31, 36
- Choleski (décomposition de), 129, 142
- Clique d'un graphe, 67, 79, 105, 127 d'un PP, 105
- Codage, 203
 - ensemble de, 198
 - test du chi-2 de, 204, 248
- Coefficient de mélange, 264
- Cohérente (famille de lois), 58
- Compétition spatiale, 171
- Contiguité spatiale (matrice de), 170
 - normalisée, 170
- Continuité (en m.q.), 16
- Contraste
 - additif, 279
 - Estimateur du minimum, 251, 269
 - gaussien, 272
 - pénalisé, 163, 181, 207, 281
 - Processus de, 270
- Convolution (Modèle de), 19
- Coopération spatiale, 171
- Corrélation repondérée d'un PP, 88, 219
- Couplage
 - de chaînes de Markov, 138
 - depuis le passé (CFTP), 140
 - temps de, 140
- Covariance, 2
 - de Matérn, 12
 - empirique, 178
 - exponentielle, 7
 - extension, restriction, 49
 - isotropique, 5, 6
 - représentation spectrale, 6
 - séparable, 19, 23
 - spatio-temporelle, 22
 - sphérique, 12
 - stationnaire, 3, 4
- Cox (PP de), 93
- Croissance de tache (modèle), 131, 249
- CSR (Complete Spatial Randomness), 88, 90, 100, 212
- Décalage spatial (modèle à), 38, 42
- Densité d'un PP, 132, 225
 - conditionnelle, 89
 - inconditionnelle, 94
- Densité spectrale, 6, 7, 24, 31, 179, 263, 272
 - d'un ARMA, 27
 - des cumulants d'ordre 4, 181
- Différentiabilité (en m.q.), 16
- Dispositif expérimental, 47, 185, 271
- Distance aux p.p.v., 101, 110, 215
- Durbin spatial (modèle de), 38, 53
- Dynamique
 - d'un champ de Markov, 76, 252
 - d'un système de particules, 273
 - d'une chaîne de Markov, 114
 - de Baker, 124
 - de Metropolis, 124
 - de relaxation site par site, 148
 - estimation d'une, 248
 - spatio-temporelle, 148
- Echange de spins, 127
- Echantillonneur de Gibbs, 121, 127
 - aléatoire, séquentiel, 121
- Effacement de points d'un PPP, 92
- Effets de bord, 154, 178, 217
- Efficacités comparées, 207

- Energie, 58
 - admissible, 58, 71
 - moyenne, 195, 283
- EQNM, 163
- Equations simultanées, 29, 30
- Ergodicité, 118, 261
- Ergodicité géométrique, 120
- Espace exponentiel d'un PP, 85
- Estimation
 - d'un champ de Markov, 193
 - d'un champ gaussien, 183
 - d'un champ stationnaire, 178
 - d'un processus ponctuel, 212
 - d'une AR, 181
 - d'une auto-corrélation, 170
 - d'une régression spatiale, 184
 - d'une texture, 196
- Estimation d'un variogramme
 - empirique, 156
 - paramétrique, 158
- Expérience numérique, 47

- Fonction K de Ripley, 102
 - estimation, 217
- Frontière libre (conditions à), 194

- Gaussien
 - continuité d'un processus, 16
 - processus, 3
- Geary (Indice d'auto-corrélation de), 174
- Géostatistique, 1
 - estimation, 154
 - modélisation, 11
 - simulation, 142
- Gibbs (modèle de), 55
 - bivarié, 147
 - latticiel, 57, 79
 - ponctuel, 96
- Graphe d'influence, 35, 170
- Graphe de Markov, 37, 67

- Hammersley-Clifford, 67, 105, 107
- héréditaire (densité d'un PP), 96, 132
- Hiérarchique (modèle), 66
 - linéaire généralisé, 238
- Homogène (PP), 92

- Identifiabilité, 30, 35, 58, 157, 160, 198, 226
- Identification d'un modèle, 281
 - CAR, 181
 - de Markov, 207
- Imagerie bayésienne, 63
- Indice d'auto-corrélation
 - de Moran, 245
- Information de Fisher, 278
 - conditionnelle, 204
 - d'un champ de Markov, 195
 - pour une régression spatiale, 187
- Intensité d'un PP, 87
 - conditionnelle, 96, 104, 225
 - d'ordre 2, 88
- Intrinsèque (processus), 8, 151, 154
 - CAR, 242
- Ising (modèle d'), 59
 - attractif, 141
 - Estimation, 205
 - Simulation, 127, 148
 - sur graphe irrégulier, 200
- Isotropique, 5, 60, 85
- Krigeage, 43, 51, 163
 - bayésien, 236
 - carte de, 46, 165
 - ordinaire, 45
 - simple, 43, 146
 - universel, 44, 237
- Kronecker (produit de), 23
- Lissage par noyau, 155, 215
 - localement finie (partie), 84
- Log-linéaire (modèle), 74, 92, 93, 202
- Logit (Modèle), 70
- Loi des grands nombres, 113, 120, 228, 261
 - au second ordre, 199
- Mélange
 - d'un champ, 161, 180, 264
 - d'un PP, 224
- Méthode de rejet, 257
 - pour un PPP, 92
- MA, 26, 33
- MAP (Maximum a posteriori), 63
- Marche aléatoire, 116
- Matérn (variogramme de), 11
- Matrice de transition, 115
 - primitive, 117

- Maximum de vraisemblance
 - d'un champ gaussien, 183
 - d'une régression gaussienne, 187
 - pour un champ de Markov, 193
 - pour un PPP, 213
- MCMC, 117, 236
- Mesure spectrale, 6, 143, 263
- Mesures de Palm d'un PP, 101
- Metropolis, 123, 124, 239
 - par échange de spins, 148
- Moëbius (formule de), 59
- Modèle hiérarchique, 150, 235
- Modèle linéaire généralisé (MLG), 202
 - hiérarchique, 238
- Moindres carrés, 271
 - quasi généralisés (MCQG), 186
 - conditionnels (MCC), 275
 - généralisés (MCG), 45, 159, 186, 271
 - ordinaires (MCO), 30, 158
 - pondérés (MCP), 159, 271
 - pour un PP, 223
 - pour un variogramme, 158
 - pour une régression spatiale, 184
- Moment réduit d'ordre 2, 102
 - estimation, 217
- Moments d'un PP, 87
- Monte Carlo, 113, 163, 174, 220, 228
- Moran (Indice d'auto-corrélation de), 171
- Mosaïque de Voronoi, 108
- MPM : Maximum posterior mode, 82

- Neyman-Scott (PP de), 90
- notations, liste de, XV
- Noyau de convolution, 19, 155, 215
- Noyau dur, 147
 - PP à, 90, 97, 250
 - sur réseau discret, 119
- Nuée variographique, 154

- Optimalité de Metropolis, 126
- Ordre lexicographique, 29

- Papangelou (intensité conditionnelle de), 96, 104
- Pépitique (variogramme), 11
- Plan d'échantillonnage, 47, 240
- Plan d'expérience, 250
- Potentiel, 57, 68, 249
 - admissible, 57
 - d'interaction, 57, 79
 - de portée bornée, 58
 - identifiabilité, 58
- Potts (Modèle de), 62, 80, 82, 242
- PP stationnaire, 85
 - pour la corrélation repondérée, 89
- Processus
 - au second ordre, 2, 178
 - continuité, 16
 - de contact, 273
 - de naissance et de mort, 116
 - différentiabilité, 16
 - gaussien, 3
 - intrinsèque, 8
 - linéaire, 26
 - stationnaire, 4
- Processus booléen, 86
- Processus ponctuel, 83, 92, 212, 263
 - à interaction d'aire, 100
 - à interaction de connexité, 100
 - à noyau dur, 90, 97
 - binomial, 90
 - caractéristique d'ordre 2, 87
 - de Cox, 93, 134, 224
 - de fibres, 86
 - de Gibbs, 96
 - de Neyman-Scott, 90, 224
 - de saturation de Geyer, 99
 - de Strauss, 97, 226
 - de Thomas, 94
 - densité inconditionnelle, 94
 - doublement poissonien, 94, 134
 - log-gaussien, 93, 134
 - multivarié, 86
- Processus ponctuel de Markov, 104
 - de Baddeley-Møller, 106
 - de Ripley-Kelly, 104
 - intensité conditionnelle, 96, 104
- Processus ponctuel de Poisson
 - estimation, 213, 253
 - homogène, 92
 - inhomogène, 92
 - intensité, 92
- Processus ponctuel marqué, 86
- pseudo-information (Matrice de), 202, 228, 278
- Pseudo-transition marginale, 275
- Pseudo-vraisemblance, 269

- conditionnelle, 196, 251
 - d'un CAR gaussien, 197
 - d'un processus ponctuel, 225
 - gaussienne, 179
 - marginale, 252, 272
 - pénalisée, 208, 281
 - poissonnienne, 215
- Quadrats (méthode des), 212
- Régression spatiale, 39, 205
 - estimation d'une, 183
- Régularisation (paramètre de), 63
- Répulsion entre points d'un PP, 89
- Réseau discret, 1, 25, 55, 119, 170, 267
- Régression spatiale
 - estimation bayésienne, 236
- Résidus d'un PP, 231
- Rabotage des données, 178
 - profil du rabot, 178
- Recollement de lois, 56, 78
- Recuit simulé, 65
- Représentation spectrale, 6
- Robuste (variogramme), 158
- Séparable (covariance), 23
- SAR, 28, 32
 - aux plus proches voisins, 37
 - estimation, 181
 - factorisant, 51
- SARX (SAR avec exogène), 38
- Segmentation d'image, 63, 79, 82
- Semi définie positive, 2, 49
- Semi-causal (modèle), 28
- Shot noise process, 94
- Simulation, 113, 255
 - conditionnelle d'un PP, 132
 - d'un auto-modèle, 128
 - d'un champ de Markov, 126
 - d'un champ gaussien, 142
 - d'un PP, 132
 - d'une chaîne de Markov, 257
 - d'une dynamique de champ, 131
 - exacte, 138
 - gaussienne conditionnelle, 146
 - par bandes tournantes, 143
 - par la méthode d'inversion, 255
 - par la méthode de rejet, 257
 - sous contrainte, 130, 149
- Sous ergodicité, 199, 204, 261, 276
- Spécification de Gibbs, 58, 59, 96
 - invariante par translation, 66, 193, 201, 208, 267
- Spatio-temporel, 21, 22, 28, 76
- Spectrogramme, 179, 272
- STARMA, 28
- Stationnaire (champ), 3
 - au second ordre, 4
 - estimation, 178
 - strict, 5
- Strauss (PP de), 97
 - à noyau dur, 99
- Télédétection, 66, 130
- TCL, 113
 - pour champ mélangeant, 266
 - pour fonctionnelle de champ, 267
 - pour une chaîne de Markov, 120
- Test
 - d'homogénéité spatiale, 212
 - d'indépendance spatiale, 174
 - d'isotropie, 205, 248
 - de minimum de contraste, 280
 - de Monte Carlo, 176, 220
 - de permutation, 174
 - de rapport de PV, 279
 - de sous-hypothèse, 160, 180, 188, 206
 - du Chi-2 de codage, 204
- Texture, 62, 66, 79, 130, 148, 149
 - estimation d'une, 196
- Transformée
 - de Fourier, 6
 - de Hankel, 6
- Transition, 115, 121
 - de proposition de changement, 123, 126
 - invariante, 116
 - irréductible, 116
 - matrice de, 115
 - primitive, 117
 - réversible, 118
- Transition de phase, 58, 60
- Triangulation de Delaunay, 108
- Validation croisée, 163
- Validation d'un modèle
 - de PP, 224, 230
 - de variogramme, 163

Variation totale (distance en), 117, 134, 265

Variogramme, 8

anisotropique, 14

autosimilaire, 13

de Matérn, 11, 164

estimation, 156

exponentiel, 11

gigogne, 14

pépitique, 11

palier d'un, 10

portée d'un, 10

puissance, 11

robustifié, 158

sphérique, 11

Whittle

pseudo-vraisemblance, 179, 208

Yule-Walker (équations de), 28, 32, 52

Déjà parus dans la même collection

1. T. CAZENAVE, A. HARAUX : Introduction aux problèmes d'évolution semi-linéaires. 1990
2. P. JOLY : Mise en œuvre de la méthode des éléments finis. 1990
- 3/4. E. GODLEWSKI, P.-A. RAVIART : Hyperbolic systems of conservation laws. 1991
- 5/6. PH. DESTUYNDER : Modélisation mécanique des milieux continus. 1991
7. J. C. NEDELEC : Notions sur les techniques d'éléments finis. 1992
8. G. ROBIN : Algorithmique et cryptographie. 1992
9. D. LAMBERTON, B. LAPEYRE : Introduction au calcul stochastique appliqué. 1992
10. C. BERNARDI, Y. MADAY : Approximations spectrales de problèmes aux limites elliptiques. 1992
11. V. GENON-CATALOT, D. PICARD : Eléments de statistique asymptotique. 1993
12. P. DEHORNOY : Complexité et décidabilité. 1993
13. O. KAVIAN : Introduction à la théorie des points critiques. 1994
14. A. BOSSAVIT : Électromagnétisme, en vue de la modélisation. 1994
15. R. KH. ZEYTOUNIAN : Modélisation asymptotique en mécanique des fluides Newtoniens. 1994
16. D. BOUCHE, F. MOLINET : Méthodes asymptotiques en électromagnétisme. 1994
17. G. BARLES : Solutions de viscosité des équations de Hamilton-Jacobi. 1994
18. Q. S. NGUYEN : Stabilité des structures élastiques. 1995
19. F. ROBERT : Les systèmes dynamiques discrets. 1995
20. O. PAPINI, J. WOLFMANN : Algèbre discrète et codes correcteurs. 1995
21. D. COLLOMBIER : Plans d'expérience factoriels. 1996
22. G. GAGNEUX, M. MADAUNE-TORT : Analyse mathématique de modèles non linéaires de l'ingénierie pétrolière. 1996
23. M. DUFLO : Algorithmes stochastiques. 1996
24. P. DESTUYNDER, M. SALAUN : Mathematical Analysis of Thin Plate Models. 1996
25. P. ROUGEE : Mécanique des grandes transformations. 1997
26. L. HÖRMANDER : Lectures on Nonlinear Hyperbolic Differential Equations. 1997
27. J. F. BONNANS, J. C. GILBERT, C. LEMARÉCHAL, C. SAGASTIZÁBAL : Optimisation numérique. 1997
28. C. COCOZZA-THIVENT : Processus stochastiques et fiabilité des systèmes. 1997
29. B. LAPEYRE, É. PARDOUX, R. SENTIS : Méthodes de Monte-Carlo pour les équations de transport et de diffusion. 1998
30. P. SAGAUT : Introduction à la simulation des grandes échelles pour les écoulements de fluide incompressible. 1998
31. E. RIO : Théorie asymptotique des processus aléatoires faiblement dépendants. 1999
32. J. MOREAU, P.-A. DOUDIN, P. CAZES (EDS.) : L'analyse des correspondances et les techniques connexes. 1999
33. B. CHALMOND : Eléments de modélisation pour l'analyse d'images. 1999
34. J. ISTAS : Introduction aux modélisations mathématiques pour les sciences du vivant. 2000
35. P. ROBERT : Réseaux et files d'attente : méthodes probabilistes. 2000
36. A. ERN, J.-L. GUERMOND : Eléments finis : théorie, applications, mise en œuvre. 2001
37. S. SORIN : A First Course on Zero-Sum Repeated Games. 2002
38. J. F. MAURRAS : Programmation linéaire, complexité. 2002

39. B. YCART : Modèles et algorithmes Markoviens. 2002
40. B. BONNARD, M. CHYBA : Singular Trajectories and their Role in Control Theory. 2003
41. A. TSYBAKOV : Introduction à l'estimation non-paramétrique. 2003
42. J. ABDELJAOUED, H. LOMBARDI : Méthodes matricielles – Introduction à la complexité algébrique. 2004
43. U. BOSCAIN, B. PICCOLI : Optimal Syntheses for Control Systems on 2-D Manifolds. 2004
44. L. YOUNES : Invariance, déformations et reconnaissance de formes. 2004
45. C. BERNARDI, Y. MADAY, F. RAPETTI : Discrétisations variationnelles de problèmes aux limites elliptiques. 2004
46. J.-P. FRANÇOISE : Oscillations en biologie : Analyse qualitative et modèles. 2005
47. C. LE BRIS : Systèmes multi-échelles : Modélisation et simulation. 2005
48. A. HENROT, M. PIERRE : Variation et optimisation de formes : Une analyse géométrique. 2005
49. B. BIDÉGARAY-FESQUET : Hiérarchie de modèles en optique quantique : De Maxwell-Bloch à Schrödinger non-linéaire. 2005
50. R. DÁGER, E. ZUAZUA : Wave Propagation, Observation and Control in 1 – d Flexible Multi-Structures. 2005
51. B. BONNARD, L. FAUBOURG, E. TRÉLAT : Mécanique céleste et contrôle des véhicules spatiaux. 2005
52. F. BOYER, P. FABRIE : Éléments d'analyse pour l'étude de quelques modèles d'écoulements de fluides visqueux incompressibles. 2005
53. E. CANCÈS, C. L. BRIS, Y. MADAY : Méthodes mathématiques en chimie quantique. Une introduction. 2006
54. J-P. DEDIEU : Points fixes, zeros et la methode de Newton. 2006
55. P. LOPEZ, A. S. NOURI : Théorie élémentaire et pratique de la commande par les régimes glissants. 2006
56. J. COUSTEIX, J. MAUSS : Analyse asymptotique et couche limite. 2006
57. J.-F. DELMAS, B. JOURDAIN : Modèles aléatoires. 2006
58. G. ALLAIRE : Conception optimale de structures. 2007
59. M. ELKADI, B. MOURRAIN : Introduction à la résolution des systèmes polynomiaux. 2007
60. N. CASPARD, B. LECLERC, B. MONJARDET : Ensembles ordonnés finis : concepts, résultats et usages. 2007
61. H. PHAM : Optimisation et contrôle stochastique appliqués à la finance. 2007
62. H. AMMARI : An Introduction to Mathematics of Emerging Biomedical Imaging. 2008
63. C. GAETAN, X. GUYON : Modélisation et statistique spatiales. 2008