

INVITED PAPER/TRABAJO INVITADO

MODELACIÓN PARA LA ESTADÍSTICA ESPACIAL

Xavier Guyón* *

*SAMOS, Université Paris 1, Panthéon-Sorbonne
France

ABSTRACT

We present the main spacial models used in spacial statistics: (i) second order model; intrinsic model and geo-statistic; (ii) spacial model of Gibbs-Markov and AR-models on a discrete network; (iii) model of point processes

KEY WORDS: geo-statistics, spacial auto-regression, Gibbs-Markov Fields, estimation process, spacial simulation.

MSC: 90C59

RESUMEN

Presentamos los principales modelos espaciales usados en estadística: (i) modelo de segundo orden; modelo intrínscico y geo-estadístico; (ii) modelo espacial de Gibbs-Markov y AR-modelos en una red discreta ; (iii) modelo de procesos puntuales.

1 Introducción

La estadística espacial estudia fenómenos aleatorios $X = \{X_s, s \in S\}$ indizados por un conjunto espacial S , X_s con estado en E . La localización del sitio $s \in S$ puede ser determinista (geo-estadística, campo de Markov, AR espacial) o aleatorio (proceso puntual). Clásicamente, $S \subseteq \mathbb{R}^2$ pero se puede tener $S \subseteq \mathbb{R}$ (cromatografía, experimentos agronómicos) o también $S \subseteq \mathbb{R}^3$ (prospección mineral, ciencias del suelo, imágenes 3-D). En cuanto a una dinámica espacial, el índice espacio-temporal es $(s, t) \in \mathbb{R}^2 \times \mathbb{R}^+$. La estadística espacial conoce un desarrollo importante en numerosas ramas tales que ingeniería minera, ciencias del medio ambiente y de la tierra, ecología, biología, geografía y economía, epidemiología, agronomía, forestales, procesamiento de imagen, etc. Hay tres tipos de datos espaciales :

- los *datos geo-estadísticos* : $S \subset \mathbb{R}^2$ es un *sub-conjunto continuo*, X es con *valor real*, observado sobre un sub-conjunto finito $\mathcal{O} = \{s_1, s_2, \dots, s_n\}$ regular o no. La geo-estadística se preocupa de la identificación y estimación del modelo, luego de la predicción (kriging) en los sitios no observados.
- los *datos sobre una red (un lattice)* : S es discreto estructurado por un grafo de vecindad \mathcal{G} , X es con *valores reales o no*, observado sobre $\mathcal{O} = \{s_1, s_2, \dots, s_n\}$. Los sitios representan unidades geográficas (número de enfermos en un

*guyon@univ-paris1.fr

departamento, datos integrados sobre una superficie) o un sitio espacial, pixeles de una imagen o una red de estaciones de mediciones de lluvia (pluviometría). Dominan dos familias de modelos : las Auto-Regresiones (AR) espacial y los campos de Markov. modelación, estimación, control de la correlación espacial, restauración de imagen o problemas inversos son algunos de los objetivos de estudio para esos datos.

- los *datos puntuales* (Proceso Puntual, PP) : aquí se observan las posiciones aleatorias $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^2$ de los casos del fenómeno : \mathbf{x} es la realización de un proceso puntual espacial, PP marcado (PPM) si además se observa una marca en cada x_i . Una pregunta central es de saber si la repartición espacial es “al azar” (PP de Poisson) o no, y si no es de Poisson, cual es el modelo de dependencia espacial, si presenta clusters, etc.

Como las series de tiempo, la estadística espacial se diferencia de la estadística clásica por la dependencia de las observaciones : X es un proceso espacial o un *campo aleatorio* sobre S . Pero la estadística espacial es radicalmente distinta de la de las series de tiempo porque se basa en *modelaciones no causales*. Por ejemplo, una modelación espacial de Markov se define en términos de vecindad espacial en “todas las direcciones”.

La tipología en 3 tipos de datos espaciales estructura este artículo. Para cada categoría, presentaremos los principales modelos utilizados y algunas de sus propiedades. Estudios más exhaustivos y ejemplos de aplicaciones se pueden encontrar en libros como los de Chiles y Delfiner (Geo-estadística, [17]), Chalmond (análisis de imagen, [16]), Cressie [19], Diggle (PP, [22]), Diggle y Rivero (Geo-estadística)[23], Gaetan et Guyon ([27]; [28]), Geman (procesamiento de imágenes, [29], [30]), Guyon [32], Lantuéjoul (simulación geo-estadística, [35]), Möller et Waagepetersen (PP, [42]), Ripley ([49],[50]), Rue et Held (campo gaussiano, [52]), Schabenberger y Gotway [53], Van Lieshout (PP, [61]), Wakernagel (Geo-estadística, [64]), Winkler (campo de Markov y análisis de imágenes, [66]). También se encontrara información y referencias en R [46] en particular con sus paquetes espaciales, `geoR` para la geo-estadística [47], `spdep` para la dependencia espacial y la modelación AR, `spatstat` para los P.P. (cf. [8]) y otros tales `RandomFields` (simulación y análisis de RF) o `mcmc` por la simulación espacial por dinámica de cadena de Markov. Damos al final una bibliografía de base limitada pero que permite ir más adelante en todos los tópicos que presentamos aquí.

2 Modelo de segundo orden y geo-estadística

Aquí, $X = \{X_s, s \in S\}$ es un campo sobre $S \subseteq \mathbb{R}^d$ con valores en $E \subseteq \mathbb{R}$ (o en $E \subseteq \mathbb{R}^d$). Este capítulo es dedicado a los *campos de segundo orden*, es decir campos X t.q. $Var(X_s) < \infty$ para todo s , y su modelación es realizada via la *covarianza*. Luego presentamos la clase mayor de los *campos intrínsecos*, los cuales son de crecimiento de varianza finita, y su modelación por la vía del *variograma*. Otra forma de definir un campo de orden dos es por la vía de las Auto-Regresiones espaciales simultáneas (SAR), S siendo discreto y equipado de una relación (orientada) de vecindad. Los SAR son adaptados a los datos integrando una medida sobre una “unidad espacial” s . Las AR Condicionales (CAR) extienden esos modelos. Primero, recordamos algunas definiciones de base.

Definition 2.1 X es un proceso de orden dos (se denota $X \in L^2$) si para todo $s \in S$, $E(X_s^2) < \infty$. La media de X es la función $m : s \mapsto m(s) = E(X_s)$ y la covarianza la función $c : (s, t) \mapsto c(s, t) = Cov(X_s, X_t)$. X es centrado si $m(s) = 0$ en todo s .

La propiedad característica de una covarianza es la de ser *semi-definida positiva* (s.d.p.), es decir que para todo $m \geq 1$, $a \in \mathbb{R}^m$ y $(s_1, s_2, \dots, s_m) \in S^m$, se tiene : $\sum_{i,j=1,m} a_i a_j c(s_i, s_j) \geq 0$. La covarianza es *definida positiva* (d.p.) si además $\sum_{i=1}^m \sum_{j=1}^m a_i a_j c(s_i, s_j) > 0$ para $a \neq 0$. Los *procesos gaussianos* forman una sub-clase importante de procesos de L^2 .

Definition 2.2 X es un proceso gaussiano si para toda parte finita $\Lambda \subset S$ y toda sucesión real $a = (a_s, s \in \Lambda)$, $\sum_{s \in \Lambda} a_s X_s$ es gaussiana.

Si $m_\Lambda = E(X_\Lambda)$ y $\Sigma_\Lambda = Cov(X_\Lambda)$ es inversible, X_Λ tiene como densidad :

$$f_\Lambda(x_\Lambda) = (2\pi)^{-\#\Lambda/2} (\det \Sigma_\Lambda)^{-1/2} \exp \left\{ -1/2 {}^t(x_\Lambda - m_\Lambda) \Sigma_\Lambda^{-1} (x_\Lambda - m_\Lambda) \right\},$$

donde $\#\Lambda$ es el cardinal de Λ y x_Λ la realización de X_Λ . El teorema de Kolmogorov asegura que para toda función m y toda covarianza c d.p., existe un campo (gaussiano) de media m y covarianza c .

2.1 Proceso estacionario

Supongamos que S es el grupo aditivo \mathbb{R}^d o \mathbb{Z}^d .

Definition 2.3 X es un campo estacionario de L^2 si la media es constante y la covarianza c invariante por traslación :

$$\forall s, t \in S : E(X_s) = m \text{ y } c(s, t) = Cov(X_s, X_t) = C(t - s).$$

$C : S \rightarrow \mathbb{R}$ es la covarianza estacionaria de X . La función de correlación es $h \mapsto \rho(h) = C(h)/C(0)$. Se tiene :

Proposition 1 Sea X un proceso estacionario de L^2 de covarianza C . Entonces :

1. $\forall h \in S, |C(h)| \leq C(0) = Var(X_s)$.
2. $\forall m \geq 1, a \in \mathbb{R}^m$ y $\{t_1, t_2, \dots, t_m\} \subseteq S : \sum_{i=1}^m \sum_{j=1}^m a_i a_j C(t_i - t_j) \geq 0$.
3. Si $A : \mathbb{R}^d \rightarrow \mathbb{R}^d$ es lineal, $X^A = \{X_{As}, s \in S\}$ es estacionario de covarianza $C^A(s) = C(As)$.
4. Si C es continua en 0, entonces C es uniformemente continua dondequiera.
5. Si C_1, C_2 son covarianzas estacionarias, también lo son $C(h) = a_1 C_1(h) + a_2 C_2(h)$ si $a_1, a_2 \geq 0$ y $C(h) = C_1(h)C_2(h)$.

X es un *Ruido Blanco Fuerte* (RBF) si los $\{X_s, s \in S\}$ son centrados, independientes y idénticamente distribuidos (i.i.d.). X es un *Ruido Blanco Debil* (RBD) si los $\{X_s, s \in S\}$ son centrados, incorrelacionados y de varianzas finitas constantes : $\forall s \neq t, Cov(X_s, X_t) = 0$ y $Var(X_s) = \sigma^2 < \infty$.

Sea $\|\cdot\|$ la norma euclidiana sobre \mathbb{R}^d : $\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^d x_i^2}$ si $x = (x_1, x_2, \dots, x_d)$.

Definition 2.4 La covarianza es *isotrópica* si $Cov(X_s, X_t)$ depende únicamente de $\|s - t\|$:

$$\exists C_0 : \mathbb{R}^+ \rightarrow \mathbb{R} \text{ t.q. } : \forall t, s \in S, c(s, t) = C_0(\|s - t\|) = C(s - t).$$

Una covarianza isotrópica es estacionaria pero, como vamos a verlo, la isotropía impone restricciones.

2.1.1 Representación espectral de una covarianza

Para esta noción y resultados asociados, no hay diferencias fundamentales entre proceso sobre \mathbb{R} y campo sobre \mathbb{R}^d . La teoría de Fourier y el teorema de Bochner ponen en biyección una covarianza estacionaria C sobre S y su medida espectral F .

El caso $S = \mathbb{R}^d$: se asocia a C una medida $F \geq 0$ simétrica y acotada sobre $\mathcal{B}(\mathbb{R}^d)$, $C(h) = \int_{\mathbb{R}^d} e^{i^t hu} F(du)$ donde ${}^t hu = \sum_{i=1}^d h_i u_i$. Si C es integrable, F admite una densidad f , la *densidad espectral* (d.es.) de X , que verifica :

$$f(u) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i^t hu} C(h) dh. \quad (2.1)$$

Si X es isotrópico, f también es isotrópica y recíproco mente. Sea $h = (r, \theta)$ et $u = (\rho, \alpha)$ las representaciones polares de h y u , $c_d(r) = C(h)$ y $f_d(\rho) = f(u)$ las covarianza y d.es. isotrópicas asociadas. Se tiene [67], $C(h) = c_d(r) = \int_{]0, \infty[} \Lambda_d(r\rho) \rho^{d-1} f_d(\rho) d\rho$ donde $\Lambda_d(v) = \Gamma(d/2)(v/2)^{-(d-2)/2} \mathcal{J}_{(d-2)/2}(v)$, \mathcal{J}_κ la función de Bessel de primera especie y de orden κ . Para $n = 2$ y 3 :

$$c_2(r) = 2\pi \int_{]0, \infty[} \rho J_0(\rho r) f_2(\rho) d\rho \quad \text{y} \quad c_3(r) = \frac{2}{r} \int_{]0, \infty[} \rho \sin(\rho r) f_3(\rho) d\rho.$$

En particular, tenemos las cotas inferiores para una correlación isotrópica [67, 50] : $\rho_0(\|h\|) \geq -0.403$ sobre \mathbb{R}^2 , $\rho_0(\|h\|) \geq -0.218$ sobre \mathbb{R}^3 , $\rho_0(\|h\|) \geq -0.113$ sobre \mathbb{R}^4 y $\rho_0(\|h\|) \geq 0$ sobre \mathbb{R}^N .

Ejemplo : covarianza exponencial sobre \mathbb{R}^d .

Para $t \in \mathbb{R}$, $C_0(t) = b \exp(-\alpha |t|)$, $\alpha, b > 0$, tiene como transformada de Fourier $f(u) = \frac{1}{2\pi} \int_{-\infty, \infty[} b e^{-\alpha|t| - iut} dt = \frac{\alpha b}{\pi(\alpha^2 + u^2)}$. Como $f \geq 0$ es integrable, es un d.es. y C_0 es una covarianza sobre \mathbb{R} . Por otra parte, la identidad,

$$\int_{]0, \infty[} e^{-\alpha x} \mathcal{J}_\kappa(ux) x^{\kappa+1} dx = \frac{2\alpha(2u)^\kappa \Gamma(\kappa + 3/2)}{\pi^{1/2}(\alpha^2 + u^2)^{\kappa+3/2}},$$

dice que $\phi(u) = \frac{\alpha b \Gamma[(d+1)/2]}{[\pi(\alpha^2 + u^2)]^{(d+1)/2}}$ es una d.es. isotrópica sobre \mathbb{R}^d asociada a la covarianza :

$$C(h) = C_0(\|h\|) = b \exp(-\alpha \|h\|).$$

Para toda dimensión d , C es una covarianza, denominada covarianza exponencial.

El caso $S = \mathbb{Z}^d$: sea $\mathbb{T}^d = [0, 2\pi[^d$ el toro de dimensión d . A C estacionaria sobre \mathbb{Z}^d se asocia una medida $F \geq 0$ acotada sobre $\mathcal{B}(\mathbb{T}^d)$ tal que :

$$C(h) = \int_{\mathbb{T}^d} e^{i^t uh} F(du). \quad (2.2)$$

Si $\sum_{h \in \mathbb{Z}^d} C(h)^2 < \infty$, F admite una densidad $f(u) = (2\pi)^{-d} \sum_{h \in \mathbb{Z}^d} C(h) e^{-i^t uh}$ en $L^2(\mathbb{T}^d)$. Si además $\sum_{h \in \mathbb{Z}^d} |C(h)| < \infty$, f es continua.

2.2 Proceso intrínseco y variograma

2.2.1 Definición, ejemplos y propiedades

Una manera de debilitar la hipótesis de estacionalidad L^2 es de considerar, para todo h , el proceso de los incrementos $\{\Delta X_s^{(h)} = X_{s+h} - X_s, s \in S\}$ de X . Esos incrementos pueden ser estacionarios en L^2 sin estacionalidad de X o sin que X sea en L^2 .

Definition 2.5 X es un proceso intrínseco si, $\forall h \in S$, $\Delta X_s^{(h)} = \{ \Delta X_s^{(h)} = X_{s+h} - X_s : s \in S \}$ es estacionario en L^2 . El semi-variograma de X es la función $\gamma : S \rightarrow \mathbb{R}$ definida por :

$$2\gamma(h) = \text{Var}(X_{s+h} - X_s).$$

Todo X estacionario de L^2 y de covarianza C es intrínseco de variograma $2\gamma(h) = 2(C(0) - C(h))$. Pero el recíproco es falsa : el movimiento browniano B es de variograma $\gamma(h) = \frac{1}{2}|h| = \text{Var}(B_{t+h} - B_t)$ pero no es estacionario. También, un proceso con media afín y residuo estacionario es intrínseco. En lo que sigue, nos limitaremos a estudiar procesos intrínsecos con crecimientos centrados : $\forall s, h$, $m(h) = E(X_{s+h} - X_s) = 0$.

Proposition 2 *Propiedades de un variograma γ*

1. $\gamma(h) = \gamma(-h)$, $\gamma(h) \geq 0$ y $\gamma(0) = 0$.
2. γ es condicionalmente definido negativo (c.d.n.) : $\forall a \in \mathbb{R}^n$ t.q. $\sum_{i=1}^n a_i = 0$, $\forall \{s_1, \dots, s_n\} \subseteq S$, $\sum_{i,j=1}^n a_i a_j \gamma(s_i - s_j) \leq 0$.
3. Si A es una transformación lineal sobre \mathbb{R}^d , $h \mapsto \gamma(Ah)$ es un variograma.
4. La propiedad (5) de la covarianza C de la i (1) se mantiene para γ .
5. Si γ es continua en 0, γ es continua en todo s donde γ es localmente acotada.
6. Si γ es acotada en una vecindad de 0, $\exists a$ et $b \geq 0$ t.q. $\forall x : \gamma(x) \leq a\|x\|^2 + b$.

Al contrario de una covarianza, un variograma no es necesariamente acotado (i.e. $\gamma(h) = |h|$ para el movimiento browniano). Pero la proposición indica que γ crece al infinito a lo más como $\|h\|^2$. Un ejemplo de tal crecimiento $\gamma(t) = \sigma^2 t^2$ es el variograma de $X_t = Z_0 + tZ_1$, $t \in \mathbb{R}$, donde Z_0 y Z_1 son centradas independientes y $\text{Var}(Z_1) = \sigma^2 > 0$.

Existen caracterizaciones asegurando que γ es un variograma. Una es la siguiente ([17]) : si γ es continua y si $\gamma(0) = 0$, entonces γ es un variograma si y solamente si (ssi) para todo $u > 0$, $t \mapsto \exp\{-u\gamma(t)\}$ es una covarianza. Por ejemplo, como $t \mapsto \exp\{-u\|t\|^2\}$, $u > 0$, es una covarianza sobre \mathbb{R}^d para todo d , $\gamma(t) = \|t\|^2$ es un variograma sobre \mathbb{R}^d .

Si X es estacionario de covarianza C , X es intrínseco de variograma $2\gamma(h) = 2(C(0) - C(h))$. En particular, el variograma de un proceso estacionario es acotado. Matheron ([39] y [40]) probó el recíproco parcial siguiente : si el variograma de un proceso intrínseco X es acotado, entonces $X_t = Z_t + Y$ donde Z es estacionario en L^2 y Y es una v.a.real.

Si $C(h) \rightarrow 0$ cuando $\|h\| \rightarrow \infty$, $\gamma(h) \rightarrow C(0)$ si $\|h\| \rightarrow \infty$: el variograma γ tiene un nivel de alcance $C(0) = \text{Var}(X)$ si $\|h\| \rightarrow \infty$.

Ejemplos de variogramas isotrópicos Los ejemplos siguientes son variogramas isotrópicos sobre \mathbb{R}^d muy clásicos (para otros ejemplos, cf. [67], [17], [64] y el artículo de revista [54]). Los 5 primeros ejemplos de variograma son asociados a una covarianza estacionaria $C(h) = C(0) - \gamma(h)$, con alcance $a > 0$ y nivel de alcance σ^2 .

Pepitico : $\gamma(h; \sigma^2) = \sigma^2$ si $h > 0$, $\gamma(0) = 0$, asociado a un RB.

Exponencial : $\gamma(h; a, \sigma^2) = \sigma^2 \{1 - \exp(-\|h\|/a)\}$.

Esférico ($d \leq 3$) : $\gamma(h; a, \sigma^2) = \sigma^2 \{1.5\|h\|/a - 0.5(\|h\|/a)^3\}$ si $\|h\| \leq a$, $= 0$ sino.

Exponencial generalizado : $\gamma(h; a, \sigma^2, \alpha) = \sigma^2(1 - \exp(-(\|h\|/a)^\alpha))$ si $0 < \alpha \leq 2$; $\alpha = 2$ es el modelo gaussiano.

Matérn [38]:

$$\gamma(h; a, \sigma^2, \nu) = \sigma^2 \left\{ 1 - \frac{2^{1-\nu}}{\Gamma(\nu)} (\|h\|/a)^\nu \mathcal{K}_\nu(\|h\|/a) \right\},$$

donde $\mathcal{K}_\nu(\cdot)$ es la función de Bessel modificada de segunda especie y de parámetro $\nu > -1$ [1, 56, 67].

Potencia : $\gamma(h; b, c) = b\|h\|^c$, $0 < c \leq 2$.

Comentarios

1. Una interpretación de la covarianza esférica es la siguiente : el volumen $V(a, r)$ de la intersección de dos esferas de \mathbb{R}^3 de mismo diámetro a y cuya distancia al centro r es, $V(a, r) = \nu(S_a) \{1 - 1.5(r/a) + 0.5(r/a)^3\}$ si $r \leq a$ y $V(a, r) = 0$ sino, donde $\nu(S_a)$ es el volumen de la esfera de radio a . Un proceso que realiza esta covarianza es $X_s = N(S_a(s))$ que cuenta el número de puntos de un PP de Poisson homogéneo (cf. §4.4) de intensidad $\sigma^2/\nu(S_a)$ en la esfera $S_a(s)$ de diámetro a centrada en $s \in \mathbb{R}^3$.
2. La restricción de una covarianza a todo sub-espacio sigue siendo una covarianza. Al contrario, la extensión de una covarianza isotrópica de \mathbb{R}^d a \mathbb{R}^p , $p > d$, no es siempre una covarianza.
3. El interés en la covarianza de Matérn está en su parámetro ν : ν controla la regularidad de γ en 0; esta regularidad luego controla la regularidad en media cuadrática del campo X y de su predicción \hat{X} por kriging : si ν es mayor, más regular es γ en 0 y más regular es el campo X y la superficie de kriging \hat{X} [56].
4. El modelo potencia es auto-similar, es decir : $\forall s > 0$, $\gamma(sh) = s^\alpha \gamma(h)$; no es acotado (i.e. el movimiento browniano).
5. Esos modelos se extienden por combinación lineal positiva : si $X = \sum_j X_j$ es suma de procesos intrínsecos incorrelacionados, $\gamma_X(h) = \sum_j \gamma_{X_j}(h)$.

Anisotropías Si \vec{e} es una dirección de \mathbb{R}^d , $\|\vec{e}\| = 1$, el variograma \vec{e} -direccional es $2\gamma(h) = \text{Var}(X_{s+h\vec{e}} - X_s)$. Hay *anisotropía* si dos variogramas direccionales son distintos. Se distinguen dos tipos de anisotropía : la *anisotropía geométrica* está asociada a una deformación lineal del espacio; la segunda está asociada a una *estratificación del variograma* sobre varios sub-espacios de \mathbb{R}^d [17, 53].

2.2.2 Continuidad y diferenciabilidad de X

Equipando L^2 con la norma en media cuadrática (m.c.), la proposición siguiente caracteriza la continuidad en m.c. de X .

Proposition 3 *Sea X un proceso de L^2 centrado y de covarianza $C(s, t) = \text{Cov}(X_s, X_t)$. Entonces X es continuo en m.q. ssi su covarianza es continua sobre la diagonal de $S \times S$.*

Un proceso X sobre \mathbb{R}^1 es diferenciable en m.c. en s si existe una v.a.r. \dot{X}_s tal que $\lim_{h \rightarrow 0} \frac{X_{s+h} - X_s}{h} = \dot{X}_s$ en L^2 .

Proposition 4 *Sea X un proceso de L^2 centrado de covarianza $C(s, t) = \text{Cov}(X_s, X_t)$. Si $\frac{\partial^2}{\partial s \partial t} C(s, t)$ existe y es finita en la diagonal de $S \times S$, entonces X es diferenciable en m.c. dondequiera, $\frac{\partial^2}{\partial s \partial t} C(s, t)$ existe dondequiera y la covarianza de \dot{X} es $\text{Cov}(\dot{X}_s, \dot{X}_t) = \frac{\partial^2}{\partial s \partial t} C(s, t)$.*

El caso de un campo estacionario. Se deduce de esas proposiciones que :

- un proceso intrínseco (resp. estacionario de L^2) es continuo en m.c. si γ (resp. C) es continuo en $h = 0$; en este caso, γ (resp. C) es continuo sobre todo conjunto donde γ es acotado (resp. es continuo donde quiera).

- $t \mapsto X_t$ es diferenciable en m.c. sobre \mathbb{R} si $\gamma''(0)$ existe. En este caso, γ'' existe donde quiera y \dot{X} es estacionario de covarianza γ'' , el proceso bivariate $(X, \dot{X}) \in L^2$ verificando [67] : $E(\dot{X}_{s+\tau} X_s) = \gamma'(\tau)$ et $E(X_s \dot{X}_{s+\tau}) = -\gamma'(\tau)$. En particular, X_s et \dot{X}_s son incorrelacionados para todo s . Además, si X es estacionario con $C(s, t) = c(s - t)$, entonces $C''_{s,t}(s, t) = -c''(s - t)$; si $c''(0)$ existe, \dot{X} es estacionario de covarianza $-c''$.

[2] da condiciones de continuidad y diferenciabilidad casi segura para un campo (campo gaussiano) X .

2.3 Las Auto-Regresiones espaciales

2.3.1 Modelo MA

Sea $(c_s, s \in \mathbb{Z}^d)$ un asucesión de $l^2(\mathbb{Z}^d)$ y η un RB sobre \mathbb{Z}^d de varianza σ_η^2 . Un $MA(\infty)$ sobre \mathbb{Z}^d es un proceso lineal definido sobre L^2 por :

$$X_t = \sum_{s \in \mathbb{Z}^d} c_s \eta_{t-s}. \quad (2.3)$$

X es la *media mobil infinita* de η por los pesos c . La covarianza de X es $C(h) = \sigma_\eta^2 \sum_{t \in \mathbb{Z}^d} c_t c_{t+h}$ y su d.es. $f(u) = \frac{\sigma_\eta^2}{(2\pi)^d} \left| \sum_{t \in \mathbb{Z}^d} c_t e^{i \cdot t u} \right|^2$.

2.3.2 Modelo ARMA

Esos modelos extienden los *ARMA* temporales. Sea P y Q dos polinomios de la variable compleja $z \in \mathbb{C}^d$,

$$P(z) = 1 - \sum_{s \in R} a_s z^s \text{ et } Q(z) = 1 + \sum_{s \in M} c_s z^s,$$

donde R y M son dos subconjuntos finitos de $\mathbb{Z}^d \setminus \{0\}$, $z^s = z_1^{s_1} \cdots z_d^{s_d}$ si $s = (s_1, s_2, \dots, s_d)$. Si $B^s X_t = X_{t-s}$, un *ARMA* verifica :

$$\forall t \in \mathbb{Z}^d : P(B)X_t = Q(B)\eta_t. \quad (2.4)$$

Sea $\mathbb{T} = \{\xi \in \mathbb{C}, |\xi| = 1\}$ en toro unidimensional.

Proposition 5 Si $P(z) \neq 0$ sobre \mathbb{T}^d , (2.4) admite una solución estacionaria en L^2 de d.es., si $e^{iu} = (e^{iu_1}, \dots, e^{iu_d})$:

$$f(u) = \frac{\sigma^2}{(2\pi)^d} \left| \frac{Q}{P}(e^{iu}) \right|^2.$$

Como las series de tiempo, los *ARMA* espaciales acercan todo campo que admite una d.es. continua. La d.es. de un *ARMA* siendo racional, su covarianza decrece exponencialmente a 0 en el infinito. Pero, al contrario de las series de tiempo, un *ARMA* espacial ($d \geq 2$) no admite en general una representación unilateral (o representación causal) finita para el orden lexicográfico.

Dos tipos de AR, los *SAR* y los *CAR*, son muy populares y útiles en modelación espacial. Miramos primero el caso de modelos estacionarios.

2.3.3 Auto-Regresión Simultánea (SAR) estacionaria

Supongamos que X es centrado y $R \subset \mathbb{Z}^d \setminus \{0\}$ es finito t.q. $P(z) = 1 - \sum_{s \in R} a_s z^s \neq 0$ sobre \mathbb{T}^d . Un modelo estacionario SAR relativo al RB η y los $a = \{a_s, s \in R\}$ es el modelo :

$$X_t = \sum_{s \in R} a_s X_{t-s} + \eta_t. \quad (2.5)$$

(2.5) son ecuaciones AR Simultáneas en el sentido de la econometría : los $\{X_{t-s}, s \in R\}$ son variables endógenas “atrasadas espacialmente”, un sitio u influyendo t si $t - u \in R$. Esta relación define el grafo orientado \mathcal{R} del SAR. Ilustramos con algunos ejemplos.

Modelo semi-causal Espacio×Tiempo $s \in \mathbb{Z}$ es un sitio espacial, $t \in \mathbb{N}$ el tiempo; una dinámico markoviano en t y local en s es :

$$\forall t \in \mathbb{N} \text{ y } s \in \mathbb{Z} : X_{s,t} = \alpha X_{s,t-1} + \beta(X_{s-1,t} + X_{s+1,t}) + \varepsilon_{s,t}.$$

La vinculación $(s, t-1) \rightarrow (s, t)$ está orientada, $(s, t) \longleftrightarrow (s \mp 1, t)$ no lo es. La representación causal lexicográfica de este SAR es infinita [13].

Modelo SAR isotrópico a los 4-vecinos más cercanos (vmc) sobre \mathbb{Z}^2

$$X_{s,t} = \alpha(X_{s-1,t} + X_{s+1,t} + X_{s,t-1} + X_{s,t+1}) + \varepsilon_{s,t}.$$

Aquí, \mathcal{R} es simétrico y X existe ssi, $\forall \lambda, \mu \in [0, 2\pi[$, $P(\lambda, \mu) = 1 - 2\alpha(\cos \lambda + \cos \mu) \neq 0$ (una condición suficiente es $|\alpha| < 1/4$). En este caso, la d.es. de X es $f(\lambda, \mu) = \sigma_\varepsilon^2 P(\lambda, \mu)^{-2} \in L^2(\mathbb{T}^2)$.

SAR(1) factorizante

$$X_{s,t} = \alpha X_{s-1,t} + \beta X_{s,t-1} - \alpha\beta X_{s-1,t-1} + \varepsilon_{s,t}, \quad |\alpha| \text{ et } |\beta| < 1 \quad (2.6)$$

es el AR asociado a $P(z_1, z_2) = (1 - \alpha z_1)(1 - \beta z_2)$. Su covarianza es separable, $c(s-s', t-t') = \sigma^2 \alpha^{|s-s'|} \beta^{|t-t'|}$, con $\sigma^2 = \sigma_\varepsilon^2 (1 - \alpha^2)^{-1} (1 - \beta^2)^{-1}$. Es fácil generalisar esos modelos AR al orden $p = (p_1, p_2)$ y a una dimensión $d \geq 2$. La factorización del polinomio AR y de la covarianza hacen más fáciles las manipulaciones numéricas de esos modelos. En particular, si se observa X sobre un rectángulo, la matriz de covarianza de X es un producto de Kronecker.

Los modelos SAR son de fácil uso y parsimoniosos en el número de parámetros. Pero hay que cuidar dos cosas :

1. Sin restricciones, un SAR no es en general identificable (i.e. $SAR(\theta) \equiv SAR(\theta')$ para un $\theta \neq \theta'$ es posible).
2. La estimación de un SAR por mínimos cuadrados ordinarios (MCO) sobre los residuales no es convergente.

2.3.4 Auto-Regresión condicional (CAR) estacionaria

Sea X un proceso sobre \mathbb{Z}^d , centrado, estacionario en L^2 con d.es. f . Si $f^{-1} \in L^1(\mathbb{T}^d)$, entonces X admite una representación lineal no causal ([32], Th. 1.2.2), $X_t = \sum_{s \in \mathbb{Z}^d \setminus \{0\}} c_s X_{t-s} + e_t$, donde $c_s = c_{-s}$ para todo s y, además, para todo $s \neq t$, $Cov(e_t, X_s) = 0$. Eso conduce a la definición siguiente de los campos L -markovianos o CAR(L). Sea L una parte finita y simétrica de $\mathbb{Z}^d \setminus \{0\}$, L^+ la parte positiva de L para el orden lexicográfico sobre \mathbb{Z}^d .

Definition 2.6 Un $CAR(L)$ estacionario de L^2 se escribe, para todo $t \in \mathbb{Z}^d$,

$$X_t = \sum_{s \in L} c_s X_{t-s} + e_t \text{ con } c_s = c_{-s} \text{ y tal que : } \forall s \neq t, Cov(e_t, X_s) = 0, E(e_t) = 0. \quad (2.7)$$

Esa definición traduce que $\sum_{s \in L} c_s X_{t-s}$ es la *mejor predicción lineal* en L^2 de X_t sobre todas las demás variables $\{X_s, s \neq t\}$: X es L -markoviano para la predicción lineal. Si X es gaussiano, es la mejor predicción: X es un campo gaussiano L -markoviano. Varias cosas diferencian los CAR de los SAR :

1. Un CAR exige restricciones paramétricas: L es simétrico y $\forall s: c_s = c_{-s}$.
2. Los residuales condicionales $\{e_t\}$ no forman un ruido blanco: $\{e_t\}$ es un *ruido colorado*.
3. Los residuales e_t no están correlacionados a los X_s si $s \neq t$.

Proposition 6 X definido por (2.7) existe en L^2 si $P^*(\lambda) = (1 - 2 \sum_{s \in L^+} c_s \cos(\lambda s))$ no se anula sobre \mathbb{T}^d . En este caso, la d.es. vale $f_X(u) = \sigma_e^2 (2\pi)^{-d} P^*(\lambda)^{-1}$ y los e forman un ruido colorado de covarianza:

$$Cov(e_t, e_{t+s}) = \begin{cases} \sigma_e^2 & \text{si } s = 0, \\ -\sigma_e^2 c_s & \text{si } s \in L \end{cases} \text{ y } Cov(e_t, e_{t+s}) = 0 \text{ sino.}$$

Donde la d.es. de un CAR es proporcional a $P^*(u)^{-1}$, la d.es. de un SAR es proporcional a $|P(u)|^{-2}$. Como para un SAR , las ecuaciones de Yule-Walker sobre la covarianza de un CAR se obtienen multiplicando por X_s la ecuación (2.7) que define X_t , luego tomando la esperanza: por ejemplo, para el CAR isotrópico a los 4-vmc sobre \mathbb{Z}^2 , las ecuaciones de YW son:

$$\forall s: c(s) = \sigma_e^2 \delta_0(s) + a \sum_{t: \|t-s\|_1=1} c(t).$$

Tres razones pueden justificar una modelación CAR :

1. una representación CAR es intrínseca: es la “mejor” predicción lineal de X_t sobre las otras variables $\{X_s, s \neq t\}$;
2. la estimación de un CAR por MCO es convergente;
3. la familia de los CAR estacionarios contiene estrictamente la de los SAR si $d \geq 2$.

Proposition 7 Modelos SAR y modelos CAR estacionarios sobre \mathbb{Z}^d .

1. Todo SAR es un CAR . Sobre \mathbb{Z}^1 , las dos clases coinciden.
2. Si $d \geq 2$, la familia de los CAR es mayor que la de los SAR .

La representación CAR del SAR , $P(B)X_t = \varepsilon_t$ se obtiene por identificación a partir de la forma analítica de la d.es. f de X ,

$$f(u) = \frac{\sigma_e^2}{(2\pi)^d |P(e^{iu})|^2} = \frac{\sigma_e^2}{(2\pi)^d C(e^{iu})}, \text{ con } c_0 = 1.$$

Si $A - B = \{i - j : i \in A \text{ et } j \in B\}$, se obtiene $L = \{R^* - R^*\} \setminus \{0\}$, donde $R^* = R \cup \{0\}$ y :

$$\text{si } s \in L : c_s = (\sigma_e^2 / \sigma_\varepsilon^2) \sum_{v, v+s \in R} a_v a_{v+s} \text{ si } s \neq 0 \text{ et } 0 \text{ si no.}$$

Para $d = 1$, la identidad de las clases *SAR* y *CAR* resulta del teorema de Fejer que dice que un polinomio trigonométrico $P^*(e^{i\lambda}) \geq 0$ es el módulo al cuadrado de un polinomio trigonométrico : si $P^*(e^{i\lambda}) \geq 0$, existe P t.q. $P^*(e^{i\lambda}) = |P(e^{i\lambda})|^2$.

Por ejemplo, el *SAR* factorizante, $X_{s,t} = \alpha X_{s-1,t} + \beta X_{s,t-1} - \alpha\beta X_{s-1,t-1} + \varepsilon_{s,t}$, $|\alpha|, |\beta| < 1$, es un *CAR* a los 8-vmc,

$$c_{1,0} = \alpha(1 + \alpha^2)^{-1}, c_{0,1} = \beta(1 + \beta^2)^{-1}, c_{1,1} = c_{-1,1} = -c_{1,0} \times c_{0,1}, \\ \sigma_e^2 = \sigma_\varepsilon^2 \kappa^2 \text{ où } \kappa^2 = (1 + \alpha^2)^{-1} (1 + \beta^2)^{-1}.$$

$\kappa^2 < 1$ es el factor de reducción en varianza de la predicción *CAR* de X comparada a la predicción *SAR*.

Un proceso *MA* de soporte finito es de covarianza con alcance acotado. Si $d = 1$, el teorema de Fejer asegura que el recíproco es verdad : un proceso sobre \mathbb{Z} de covarianza de alcance finito es un *MA*. Eso falla si $d \geq 2$: por ejemplo, para $d = 2$, el campo de correlación ρ tiene distancia 1, $|\rho| < \frac{1}{4}$, y 0 de distancia > 1 no admite una representación *MA*; eso se basa sobre argumentos similares a los utilizados en la prueba de (2) de la proposición anterior.

2.3.5 AR non-estacionario sobre un red finita S

Un campo real sobre $S = \{1, 2, \dots, n\}$ es una v.a. $X^* \in \mathbb{R}^n$. La no estacionalidad de X^* puede ser sobre la red S , sobre la media $\mu = E(X^*)$ o sobre la covarianza $\Sigma = Cov(X^*)$. Consideramos aquí la no estacionalidad en covarianza, trabajando sobre el campo centrado $X = X^* - \mu$ y su covarianza $\Sigma = Cov(X^*) = Cov(X)$. Sea $\varepsilon = (\varepsilon_t, t \in S)$ un RB de L^2 : representaciones, sitio por sitio o globales, *MA*, *AR* y *ARMA* de X en base de ε son definidas por las ecuaciones :

$$MA : X_t = \sum_{s \in S} b_{t,s} \varepsilon_s, \text{ o } X = B\varepsilon,$$

$$AR : X_t = \sum_{s \in S: s \neq t} a_{t,s} X_s + \varepsilon_t, \text{ o } AX = \varepsilon,$$

$$ARMA : X_t = \sum_{s \in S: s \neq t} a_{t,s} X_s + \sum_{s \in S} b_{t,s} \varepsilon_s, \text{ o } AX = B\varepsilon,$$

donde, para $s, t \in S$, $B_{t,s} = b_{t,s}$, $A_{s,s} = 1$, $A_{t,s} = -a_{t,s}$ si $t \neq s$. La representación *MA* está siempre definida; las representaciones *AR* y *ARMA* existen si A es inversible. Si $\Gamma = Cov(\varepsilon)$, esos modelos son caracterizados al orden dos por sus covarianzas, $\Sigma = B\Gamma^t B$ para un *MA*, $\Sigma = A^{-1}\Gamma^t(A^{-1})$ para un *AR* y $\Sigma = (A^{-1}B)\Gamma^t(A^{-1}B)$ para un *ARMA*.

Consideramos ε un RB de varianza 1 ($\Gamma = I_n$) y sea $<$ un orden total (arbitrario) de enumeración de S . Si X es centrado de covarianza Σ inversible, X admite una única representación *AR* causal relativa a ε y a $<$. Esta representación está asociada a la matriz triangular inferior A^* de la factorización de Cholesky, $\Sigma = {}^t A^* A^*$. A^* , como Σ , depende de $n(n+1)/2$ parámetros, lo que confirma la identificabilidad del modelo *AR* causal. Las representaciones *AR* equivalentes, pero en general no identificables, son $\tilde{A}X = \eta$, donde, para P orthogonal, $\eta = P\varepsilon$ y $\tilde{A} = PA^*$ (η es todavía un RB de varianza 1).

En la práctica, un modelo AR es asociado a un grafo de influencia \mathcal{R} no necesariamente simétrico : $s \rightarrow t$ es una arista orientada de \mathcal{R} si X_s influya sobre X_t , la vecindad de t siendo $\mathcal{N}_t = \{s \in S : s \rightarrow t\}$.

Ejemplo : una modelación SAR a los vmc (econometría, epidemiología).

Sea $W = (w_{t,s})_{t,s \in S}$ una matriz de influencia de s sobre t , t.q., $\forall t, w_{t,t} = 0$. Por ejemplo, W es la matriz de contiguidad espacial de un grafo; otras elecciones de W son presentadas en Cliff y Ord [18]). Por ejemplo, un SAR a los vmc con un parámetro de dependencia espacial ρ es :

$$X_t = \rho \sum_{s:s \neq t} w_{t,s} X_s + \varepsilon_t, \text{ ou } X = \rho W X + \varepsilon.$$

Representación markoviana CAR .

Sea X el CAR : $\forall t \in S : X_t = \sum_{s \in S: s \neq t} c_{t,s} X_s + e_t$ con $E(e_t) = 0$, $Var(e_t) = \sigma_t^2 > 0$ y $Cov(X_t, e_s) = 0$ si $t \neq s$. Esta representación está asociada al grafo de vecindad \mathcal{G} de arista $s \rightarrow t$ si $c_{t,s} \neq 0$. Como vamos a ver, \mathcal{G} es simétrico. Sea C la matriz de coeficientes $C_{s,s} = 0$ y $C_{t,s} = c_{t,s}$ si $s \neq t$, D la matriz diagonal con $D_{t,t} = \sigma_t^2$. Las ecuaciones de Yule-Walker se escriben $\Sigma = C\Sigma + D$ y Σ verifica : $(I - C)\Sigma = D$. Si Σ es inversible, $\Sigma^{-1} = D^{-1}(I - C)$ es simétrica y d.p.. En particular, C y \mathcal{G} son simétricos y la representación CAR verifica :

$$\forall t \neq s \in S : c_{t,s} \sigma_s^2 = c_{s,t} \sigma_t^2. \quad (2.8)$$

Campo gaussiano markoviano

Supongamos que X es un campo gaussiano sobre $S = \{1, 2, \dots, n\}$, $X \sim \mathcal{N}_n(\mu, \Sigma)$ donde Σ es inversible y que S está equipado de un grafo \mathcal{G} simétrico, denotando $\langle s, t \rangle$ si s y t son vecinos. Se dice que X es un campo \mathcal{G} -markoviano si, denotando $Q = (q_{s,t}) = \Sigma^{-1}$ la matriz de precisión de X , entonces $q_{st} = 0$ salvo $\langle s, t \rangle$. Luego, se tiene para todo $t \in S$,

$$\mathcal{L}(X_t | X_s, s \neq t) \sim \mathcal{N}(\mu_t - q_{t,t}^{-1} \sum_{s: \langle t, s \rangle} q_{t,s} (X_s - \mu_s), q_{t,t}^{-1}) \quad (2.9)$$

y X es el CAR : $X_t - \mu_t = -q_{t,t}^{-1} \sum_{s: \langle t, s \rangle} q_{t,s} (X_s - \mu_s) + e_t$, $Var(e_t) = q_{t,t}^{-1}$. Sea $[Q]$ la matriz $n \times n$ de diagonal 0, $[Q]_{t,s} = q_{t,s}$ si $t \neq s$ y $Diag(Q)$ la matriz diagonal de diagonal la de Q : X se escribe,

$$X - \mu = -(Diag(Q))^{-1} [Q] (X - \mu) + e.$$

Como lo veremos en el §3, los CAR gaussianos son modelos de Gibbs con potenciales cuadráticos [52].

Grafo de Markov \mathcal{G} de un SAR

El SAR gaussiano $AX = \varepsilon$ existe si A^{-1} existe. Su covarianza inversa es $\Sigma^{-1} = Q = \sigma^{-2}({}^t A A)$ y el grafo \mathcal{R} del SAR es : $\langle t, s \rangle_{\mathcal{R}} \iff a_{t,s} \neq 0$. La representación CAR equivalente tiene como coeficientes $c_{t,s} = -q_{t,s}/q_{t,t}$ donde $q_{t,s} = \sum_{l \in S} a_{l,t} a_{l,s}$ y un grafo de Markov \mathcal{G} definido por :

$$\langle t, s \rangle_{\mathcal{G}} \iff \begin{cases} \text{sea } \langle t, s \rangle_{\mathcal{R}}, \text{ sea } \langle s, t \rangle_{\mathcal{R}}, \\ \text{sea } \exists l \in S \text{ t.q. } \langle l, t \rangle_{\mathcal{R}} \text{ y } \langle l, s \rangle_{\mathcal{R}} \end{cases} .$$

\mathcal{G} es no es orientado de alcance “doble” del de \mathcal{R} .

2.3.6 Auto-Regresión con covariables : SARX

Esos modelos se han desarrollado en particular en econometría espacial. Sea Z una matriz exógena real $n \times p$. Un SARX (X para exógena) propone además de la SAR una Regresión de X sobre Z :

$$X = \rho W X + Z\beta + W Z\gamma + \varepsilon, \rho \in \mathbb{R}, \beta \text{ et } \gamma \in \mathbb{R}^p. \quad (2.10)$$

Tres factores explican X : las exógenas $Z\beta$, las endógenas $\rho W X$ y las exógenas atrasadas $W Z\gamma$. El modelo de Durbin espacial es $(I - \rho W)X = (I - \rho W)Z\beta + \varepsilon$, el con desfase espacial $X = \rho W X + Z\beta + \varepsilon$. Las medias para esos tres SARX son respectivamente :

$$E(X) = (I - \rho W)^{-1}[Z\beta + W Z\gamma], E(X) = Z\beta \text{ et } E(X) = (I - \rho W)^{-1}Z\beta,$$

para una misma covarianza $\Sigma^{-1} \times \sigma^2 = (I - \rho^t W)(I - \rho W)$ si $\varepsilon \sim RB(\sigma^2)$. Se deduce sencillamente la estimación de los SARX por MV gaussiano.

2.4 El modelo de Regresión espacial

Es el modelo con residual un campo aleatorio ε :

$$X_s = m(s) + \varepsilon_s, E(\varepsilon_s) = 0.$$

Existen numerosas maneras de modelar la media $s \mapsto m(s)$: regresión (lineal o no, log-lineal), análisis de varianza (exógena cualitativa), análisis de la covarianza. Para el residual ε , Cressie [19] propone la descomposición $\varepsilon_s = W_s + \eta_s + e_s$ donde W_s es una componente de ruido a larga escala y regular, η_s una componente a pequeña escala y e_s un ruido blanco, error de medida o ruido de pepita.

Si X se observa en n sitios $s_i \in S$, el modelo se escribe $X_{s_i} = {}^t z_{s_i} \delta + \varepsilon_{s_i}, i = 1, n$ con z_{s_i} y $\delta \in \mathbb{R}^p, z_{s_i}$ una covariable observada y $\varepsilon = (\varepsilon_{s_i}, i = 1, n)$ un residual centrado y correlacionado espacialmente. Si $X = {}^t(X_{s_1}, \dots, X_{s_n}), \varepsilon = {}^t(\varepsilon_{s_1}, \dots, \varepsilon_{s_n}), Z = {}^t(z_{s_1}, \dots, z_{s_n})$:

$$X = Z\delta + \varepsilon \text{ con } E(\varepsilon) = 0 \text{ y } Cov(\varepsilon) = \Sigma. \quad (2.11)$$

El segundo paso es la modelación de Σ .

2.5 Predicción con covarianza conocida o krigage

Se desea construir un mapa de predicción de X sobre todo S cuando X es observado en n puntos de S . La terminología krigage es de Matheron [39] en referencia a los trabajos de Krige [34], un ingeniero de minas de Sur Africa. Se observa $\{X_{s_1}, \dots, X_{s_n}\}$ y se desea hacer una predicción lineal de X_{s_0} cuando la covarianza (el variograma) de X es conocido. Sea $X_0 = X_{s_0}, X = {}^t(X_{s_1}, \dots, X_{s_n}), \Sigma = Cov(X), \sigma_0^2 = Var(X_0)$ y $c = Cov(X, X_0) \in \mathbb{R}^n$ las características (conocidas o estimadas) de orden dos de X . Buscamos un predictor sin sesgo de $X_0, \hat{X}_0 = \sum_{i=1}^n \lambda_i X_{s_i} = {}^t \lambda X$ que minimiza $EQM(s_0) = E\{(X_0 - \hat{X}_0)^2\}$.

El krigage sencillo examina la situación con $m(\cdot)$ conocida. Sin restricción, supongamos X centrado.

Proposition 8 Krigage sencillo : la predicción lineal y sin sesgo de X_0 optimal para el EQM y la varianza del error de predicción son,

$$\hat{X}_0 = {}^t c \Sigma^{-1} X \text{ y } \tau^2(s_0) = \sigma_0^2 - {}^t c \Sigma^{-1} c. \quad (2.12)$$

Si X es gaussiano, $\hat{X}_0 = E(X_0 | X_{s_1}, \dots, X_{s_n})$ y $X_0 - \hat{X}_0 \sim \mathcal{N}(0, \tau^2(s_0))$.

Si $X = Z\delta + \varepsilon$, si z_0 y Σ son conocidos pero no δ , el resultado del *krigeage universal* es el siguiente.

Proposition 9 *Krigeage universal: la predicción lineal optimal de X_0 es*

$$\hat{X}_0 = \{ {}^t c \Sigma^{-1} + {}^t (z_0 - {}^t Z \Sigma^{-1} c) ({}^t Z \Sigma^{-1} Z)^{-1} {}^t Z \Sigma^{-1} \} X. \quad (2.13)$$

con varianza del error de predicción $\tau^2(s_0) = \sigma_0^2 - {}^t c \Sigma^{-1} c + {}^t (z_0 - {}^t Z \Sigma^{-1} c) ({}^t Z \Sigma^{-1} Z)^{-1} (z_0 - {}^t Z \Sigma^{-1} c)$.

La interpretación del *krigeage universal* es la siguiente : $\hat{X}_0 = {}^t z_0 \hat{\delta} + c \Sigma^{-1} (X - Z \hat{\delta})$ es la suma de la estimación optimal ${}^t z_0 \hat{\delta}$ (con $\hat{\delta} = ({}^t Z \Sigma^{-1} Z)^{-1} {}^t Z \Sigma^{-1} X$) de $E(X_0) = {}^t z_0 \delta$ y del *krigeage sencillo* $c \Sigma^{-1} (X - Z \hat{\delta})$ de los residuales $\hat{\varepsilon} = (X - Z \hat{\delta})$ estimados por MCG.

Si $X_s = m + \varepsilon_s$, se habla de *krigeage ordinario*. Las fórmulas de *krigeage* se escriben de manera similar en términos del variograma si ε es intrínseco porque la estacionalidad no es necesaria en la obtención de esos resultados. El *krigeage* es un *interpolador exacto*, es decir que $\hat{X}_{s_0} = X_{s_0}$ si s_0 es un sitio de observación.

Regularidad de la superficie de *krigeage*

La regularidad en 0 de la covarianza C (del variograma γ) determina la regularidad de la superficie de *krigeage* $s \mapsto \hat{X}_s$. Por el modelo pepitico, la predicción es en todo s la media aritmética de los (X_{s_i}) si $s_0 \neq s_i$, con discontinuidades $\hat{X}_{s_0} = X_{s_i}$ en $s_0 = s_i$. Más generalmente, si $C(h)$ no es continua en 0, $s \mapsto \hat{X}_s$ no es continua en los puntos de observación s_i . Si $C(h)$ es continua en 0, $s \mapsto \hat{X}_s$ es continua donde quiera pero no es derivable en los puntos de observación. Si $C(h)$ es parabólica en 0, $s \mapsto \hat{X}_s$ es continua y derivable donde quiera.

Si en la dimensión $d = 1$ el *krigeage* utiliza una covarianza cúbica, el interpolador es una función spline cúbica. En dimensión superior y para una covarianza separable cúbica, las predicciones son cúbicas por parte en cada variable ([17], p. 272). Una comparación empírica entre las previsiones mediante funciones splines o por *krigeage* es dada en [36].

3 Campo de Gibbs-Markov sobre una red S

Sea $X = (X_i, i \in S)$ un campo aleatorio sobre S , un *conjunto discreto* de sitios, $X \in \Omega = E^S$ con E un *espacio de estado general*. S puede ser regular ($S \subseteq \mathbb{Z}^2$, análisis de imagen, radiografía) o no (econometría, epidemiología). Notamos π la ley de X o su densidad. Nos interesamos aquí en la descripción de π a partir de leyes condicionales, en particular contestaremos a la pregunta : si $\{\nu_i(\cdot | x^i), i \in S, x^i \in E^{S \setminus \{i\}}\}$ es una familia de leyes sobre E indexadas por x^i , la configuración de x fuera de i , cuando esas leyes son las condicionales de una ley conjunta π ? Si es el caso, se dice que la leyes condicionales se encolan en π . Si la dependencia de $\nu_i(\cdot | x^i)$ en x^i es "local", se dice que X es un *campo de Markov (Markov Random Field (MRF))* en inglés).

Sin condiciones, las $\{\nu_i\}$ no se "encolan". La familia de las *especificaciones de Gibbs* es una familia para la cual las condicionales se pegan sin condición. La importancia de los modelos de Gibbs es reforzada por el teorema de Hammersley-Clifford que dice que un campo de Markov es un campo de Gibbs con potenciales locales. Los *auto-modelos markovianos (AMM)* de Besag forman una sub-clase de campos de Gibbs sobre \mathbb{R} muy útil en estadística espacial.

En este capítulo, el espacio de estados E es general (recordamos que $E \subseteq \mathbb{R}$ en geo-estadística y para los *SAR*). Eso abre muchas posibilidades de aplicaciones : $E = \mathbb{R}^+$ para observaciones positivas (campo exponencial o campo Gamma, meteorología), $E = \mathbb{N}$ para una variable contadora (campo de Poisson, epidemiología), $E = \{a_0, a_1, \dots, a_{m-1}\}$ para un campo categórico (repartición de m especies vegetales, ecología), $E = \{0, 1\}$ para un campo binario (presencia/ausencia de una enfermedad), $E = \Lambda \times \mathbb{R}^p$ para un campo cruzando una etiqueta ν

cualitativo con una modalidad cuantitativa x (en teledetección, ν es una textura de paisaje y x una firma multiespectral), $E = \{0\} \cup]0, +\infty[$ para un estado mixto (en pluviometría, $X = 0$ si no llueve, $X > 0$ sino), sin olvidar $E = \mathbb{R}$ o $E = \mathbb{R}^p$ para un campo gaussiano.

En lo que sigue, (E, \mathcal{E}) está equipado de una medida de referencia $\lambda > 0$; \mathcal{E} es la σ -álgebra boreliana y λ la medida de Lebesgue si $E \subset \mathbb{R}^p$; \mathcal{E} es $\mathcal{P}(E)$ y λ la medida contadora si E es discreto.

Encolamiento de leyes condicionales

Sin condiciones, leyes condicionales $\{\nu_i\}$ no se encolan. Arnold, Castillo y Sarabia [3] dan una condición para que dos familias $(X|y)$ et $(Y|x)$, $x \in S(X)$ y $y \in S(Y)$ se encolen : sean μ y ν dos medidas sobre $S(X)$ y $S(Y)$, $a(x, y)$ la densidad de $(X = x|y)$ con respecto a μ , $b(x, y)$ la de $(Y = y|x)$ con respecto a ν , $N_a = \{(x, y) : a(x, y) > 0\}$ y $N_b = \{(x, y) : b(x, y) > 0\}$. Entonces las familias se encolan ssi $N_a = N_b (\doteq N)$ y si :

$$\forall (x, y) \in N : a(x, y)/b(x, y) = u(x)v(y), \quad (3.1)$$

donde $\int_{S(X)} u(x)\mu(dx) < \infty$. Así, las leyes gaussianas $(X|y) \sim \mathcal{N}(a + by, \sigma^2 + \tau^2 y^2)$ y $(Y|x) \sim \mathcal{N}(c + dx, \sigma'^2 + \tau'^2 x^2)$, σ^2 et $\sigma'^2 > 0$ no se encolan si $\tau\tau' \neq 0$; si $\tau = \tau' = 0$, se encolan si $d\sigma^2 = b\sigma'^2$, la ley conjunta siendo gaussiana (cf. (2.8); para ver cuales son las condiciones asegurando que leyes condicionales se encolan, cf. [3] y el lemma de Brook [15]).

3.1 Campo de Gibbs sobre S

Sea S denumerable (i.e. $S = \mathbb{Z}^d$ si S es regular), $\mathcal{S} = \mathcal{P}_F(S)$ el conjunto de las partes finitas de S , $x_A = (x_i, i \in A)$ la configuración de x sobre A , $x^A = (x_i, i \notin A)$ la configuración sobre $S \setminus A$, $x^i = x^{\{i\}}$, $\Omega_A = \{x_A, y \in \Omega\}$, $\Omega^A = \{x^A, x \in \Omega\}$, $\mathcal{F} = \mathcal{E}^{\otimes S}$ la σ -álgebra sobre Ω y dx_Λ la medida $\lambda^{\otimes \Lambda}(dx_\Lambda)$ sobre $(E^\Lambda, \mathcal{E}^{\otimes \Lambda})$.

3.1.1 Potencial de interacción y especificación de Gibbs

Un campo de Gibbs es asociado a una familia π^Φ de leyes condicionales definidas por potenciales de interacción Φ .

Definition 3.1 *Potencial de interacción, energía, admisibilidad, especificación de Gibbs*

1. *Un potencial de interacción es un familia $\Phi = \{\Phi_A, A \in \mathcal{S}\}$ de aplicaciones medibles $\Phi_A : \Omega_A \mapsto \mathbb{R} \text{ t.q. } \forall \Lambda \in \mathcal{S}$, la suma siguiente existe,*

$$U_\Lambda^\Phi(x) = \sum_{A \in \mathcal{S} : A \cap \Lambda \neq \emptyset} \Phi_A(x). \quad (3.2)$$

U_Λ^Φ es la energía de Φ sobre Λ , Φ_A el potencial sobre A .

2. Φ es admisible si $\forall \Lambda \in \mathcal{S}$ y $y^\Lambda \in \Omega^\Lambda$, $Z_\Lambda^\Phi(x^\Lambda) = \int_{\Omega_\Lambda} \exp U_\Lambda^\Phi(x_\Lambda, x^\Lambda) dx_\Lambda < +\infty$.

3. Si Φ es admisible, la especificación de Gibbs π^Φ es la familia de leyes condicionales :

$$\pi_\Lambda^\Phi(x_\Lambda | x^\Lambda) = \{Z_\Lambda^\Phi(x^\Lambda)\}^{-1} \exp U_\Lambda^\Phi(x), \Lambda \in \mathcal{S}. \quad (3.3)$$

La familia π^Φ es *coherente*, es decir : si $\Lambda \subset \Lambda^*$, π_Λ es la restricción a Λ de π_{Λ^*} . La sumabilidad (3.2) se verifica si S es un espacio métrico y si Φ es de *alcance acotado* : $\exists R < \infty$ t.q. $\Phi_A \equiv 0$ si $\delta(A) = \sup_{i,j \in A} d(i, j) > R$.

Especificación π^Φ y medidas de Gibbs $\mathcal{G}(\Phi)$ Una *medida de Gibbs* $\mu \in \mathcal{G}(\Phi)$ es un ley μ sobre (Ω, \mathcal{F}) t.q. sus leyes condicionales coinciden con π^Φ . Si S es finito, $\pi(x) = \pi_S(x)$ y $\mathcal{G}(\Phi) = \{\pi\}$. Si S es infinito, una pregunta es si existe tal medida y, si existe, si hay unicidad. Eso no es asegurado sin condiciones particulares. Contestar a eso es uno de los objetivos de la *mecánica estadística* (cf. Georgii, [31]). Si E es un espacio polaco (por ejemplo un cerrado de \mathbb{R}^d , o un conjunto finito) y si λ es finita, Dobrushin ([25]; [31]) demostró que $\mathcal{G}(\Phi) \neq \emptyset$. Pero esta ley no es necesariamente única : si $\sharp\mathcal{G}(\Phi) > 1$, se dice que hay *transición de fase*, dos leyes distintas generando las mismas leyes condicionales. Eso quiere decir que si S es infinito, una especificación de Gibbs no especifica siempre un modelo conjunto sobre S . Esta es una de las dificultades de la estadística asintótica para los campos de Gibbs. Dobrushin ha dado una condición suficiente asegurando que no hay transición de fase.

Identificabilidad de un potencial Φ . Sin restricciones, $\Phi \mapsto \pi^\Phi$ no es identificable : por ejemplo, cualquiera sea un real c , si cambiamos Φ_A en $\tilde{\Phi}_A = \Phi_A + c$, entonces $\pi^\Phi \equiv \pi^{\tilde{\Phi}}$. Una manera de hacer Φ identificable es la siguiente : sea τ un estado de referencia de E ; entonces las condiciones siguientes hacen Φ identificable :

$$\forall A \neq \emptyset, \Phi_A(x) = 0 \text{ si } \exists i \in A \text{ t.q. } x_i = \tau. \quad (3.4)$$

Es una consecuencia de la fórmula de inversión de Moëbius : $\forall A, \Phi_A(x_A) = \sum_{V \subseteq A} (-1)^{\sharp(A \setminus V)} U(x_A, \tau^A)$, que asocia unívocamente a U los potenciales que verifican (3.4).

3.1.2 Ejemplos de especificación de Gibbs

Un modelo de Gibbs se caracteriza por los potenciales $\Phi = \{\Phi_A, A \in \mathcal{C}\}$ indizados por la familia $\mathcal{C} \subseteq \mathcal{P}(S)$. Si los Φ_A se escriben $\Phi_A(x) = \theta_A \phi_A(x)$, con ϕ_A conocidas, π pertenece a la *familia exponencial* $\pi(x) = Z^{-1}(\theta) \exp {}^t \theta T(x)$ de parámetro $\theta = (\theta_A, A \in \mathcal{C})$ y estadística exhaustiva $T(x) = (\phi_A(x), A \in \mathcal{C})$. Conocer la forma explícita de las leyes condicionales permite :

- (i) utilizar algoritmos MCMC de simulación (muestreador de Gibbs, algoritmo de Metrópolis);
- (ii) usar estimación mediante pseudo-verosimilitud condicional (*SV C*) fáciles de poner en marcha numericamente y con buenas propiedades asintóticas cuando el *MV* es de difícil obtención (Gaetan et Guyon [27],[28]).

Modelos de Ising. Introducidos por los físicos para modelar configuraciones espaciales de spins $x_i \in \{-1, +1\}$ sobre la red \mathbb{Z}^3 , esos modelos binarios son populares en el análisis de imagen o en la estadística espacial. Si $S \subseteq \mathbb{Z}^2$, el modelo de *Ising isotrópico a los 4-vmc* tiene como potenciales $\Phi_{\{i\}}(x) = \alpha x_i$, $i \in S$ y $\Phi_{\{i,j\}}(x) = \beta x_i x_j$ si $\langle i, j \rangle$, es decir $\|i - j\|_1 = |i_1 - j_1| + |i_2 - j_2| = 1$. Si $\Lambda \subseteq \mathbb{Z}^2$ es finita, la energía sobre Λ es $U_\Lambda(x_\Lambda | x^\Lambda) = \alpha \sum_{i \in \Lambda} x_i + \beta \sum_{i \in \Lambda, j \in S : \langle i, j \rangle} x_i x_j$ y π_Λ^Φ se escribe :

$$\pi_\Lambda(x_\Lambda | x^\Lambda) = Z_\Lambda^{-1}(\alpha, \beta; x^\Lambda) \exp U_\Lambda(x_\Lambda; x^\Lambda),$$

donde el *constante de normalización* vale $Z(\alpha, \beta; x^\Lambda) = \sum_{x_\Lambda \in E^\Lambda} \exp U_\Lambda(x_\Lambda; x^\Lambda)$. El papel de los parámetros α y β es el siguiente : α modula la ley marginal, β la correlación espacial. $Z_\Lambda(\alpha, \beta; x^\Lambda)$, suma de $2^{\sharp\Lambda}$ términos, es rápidamente incalculable : por ejemplo, si Λ es un (pequeño) cuadrado 10×10 , hay $2^{100} \simeq 1.27 \times 10^{30}$ términos en Z_Λ ! Al contrario, la ley condicional en i es explícita,

$$\pi_i(x_i | x^i) = \frac{\exp x_i(\alpha + \beta v_i(x))}{2 \text{ch}(\alpha + \beta v_i(x))} \text{ con } v_i(x) = \sum_{j: \langle i, j \rangle} x_j,$$

función de los 4-vmc. De aquí sale la simulación de π_Λ por el muestreador de Gibbs o la estimación de los parámetros (α, β) por *SVC*.

El modelo de Ising se generaliza en varias direcciones, cubriendo así un ancho espectro de situaciones : se puede introducir una anisotropía, aumentar la relación de vecindad $\langle i, j \rangle$, introducir potenciales más allá que de pares, aumentar el número de estados de E (modelo de Potts, Strauss [59]) con estados cualitativos (variedades en ecología) o cuantitativos (nivel de gris en análisis de imagen). También, esos modelos se pueden definir sobre una red S sin regularidad, sin estacionalidad para el grafo de vecindad \mathcal{G} y sin estacionalidad para los potenciales.

Especificación gaussiana sobre $S = \{1, 2, \dots, n\}$. Si $\Sigma^{-1} = Q$ existe, una ley gaussiana $X = (X_i, i \in S) \sim \mathcal{N}_n(\mu, \Sigma)$ es un campo de Gibbs de energía $U(x) = \frac{1}{2} {}^t(x - \mu)Q(x - \mu)$ y potenciales :

$$\Phi_{\{i\}}(x) = x_i \left\{ \sum_{j:j \neq i} q_{ij} \mu_j \right\} - \frac{1}{2} q_{ii} x_i^2 \text{ y } \Phi_{\{i,j\}}(x) = -q_{ij} x_i x_j \text{ si } i \neq j.$$

Las leyes condicionales $\mathcal{L}_A(X_A | x^A)$ se deducen de la energía condicional $U_A(\cdot | x^A)$. X es un campo \mathcal{G} -markoviano si, para todo $i \neq j : q_{ij} \neq 0 \iff \langle i, j \rangle$ es una arista de \mathcal{G} .

3.2 Campo de Markov y campo de Gibbs

El interés de los campos de Gibbs tiene entre sus capacidades el definir sencillamente y de manera flexible especificaciones condicionales coherentes. Pero también los campos de Gibbs son *campos de Markov*. Definamos esta noción.

Sea $S = \{1, 2, \dots, n\}$ un conjunto de sitios equipado de un grafo de vecindad \mathcal{G} simétrico y sin lazo, $\langle i, j \rangle$ la relación de vecindad asociada, $\partial A = \{i \in S, i \notin A : \exists j \in A \text{ t.q. } \langle i, j \rangle\}$ la frontera de vecindad de A y $\partial i = \partial \{i\}$ si $i \in S$.

Definition 3.2 *Campo de Markov, cliques de un grafo*

1. X es un campo de Markov (*Markov Random Field (MRF) en inglés*) sobre S para \mathcal{G} si, $\forall A \subset S$ y $x^A \in \Omega^A$, la ley de X sobre A condicional a x^A depende unicamente de $x_{\partial A} : \pi_A(x_A | x^A) = \pi_A(x_A | x_{\partial A})$.
2. Una parte C no vacía de S es una *clique* del grafo \mathcal{G} si o C es un *singleton* o si los elementos de C son dos a dos vecinos para \mathcal{G} . El conjunto de las cliques de \mathcal{G} es denotado $\mathcal{C}(\mathcal{G})$.

Por ejemplo, para el grafo a los 4-vmc sobre \mathbb{Z}^2 , las cliques son los singletons $\{i\}$ y los pares $\{i, j\}$ t.q. $\|i - j\|_1 = 1$. Para la relación a los 8-vmc, hay además las partes $A = \{i, j, k\}$ y $A = \{i, j, k, l\}$ t.q. $\forall u, v \in A, \|u - v\|_\infty \leq 1$.

A una familia \mathcal{C} de partes de S que incluya todos los singletons de S se associa el grafo $\mathcal{G}(\mathcal{C})$ así definido : $i \neq j$ son vecinos de $\mathcal{G}(\mathcal{C})$ si $\exists C \in \mathcal{C}$ t.q. $\{i, j\} \subset C$. La propiedad siguiente identifica un campo de Markov a un campo de Gibbs.

Theorem 1 *Teorema de Hammersley-Clifford (Besag, [14]; [27], [28]).*

1. Sea π un campo \mathcal{G} markoviano sobre E t.q. $\pi_A(x_A | x^A) > 0, \forall A \subset S$ y $x \in E^S$. Entonces existe un potencial $\Phi = \{\Phi_A, A \in \mathcal{C}\}$ definido sobre \mathcal{C} , las cliques de \mathcal{G} , t.q. π es un campo de Gibbs de potencial Φ .
2. Recíprocamente, sea \mathcal{C} una familia de partes de S incluyendo los singletons. Entonces, un campo de Gibbs de potenciales $\Phi = \{\Phi_A, A \in \mathcal{C}\}$ es un campo de Markov para el grafo de vecindades $\mathcal{G}(\mathcal{C})$.

Una sub-clase importante de campos de Markov es la de los auto-modelos de Markov de Besag.

3.3 Auto-modelos markovianos de Besag (AMM)

Estos modelos de Markov son caracterizados por densidades condicionales en una familia exponencial. Empezamos con el resultado básico de su definición.

Sea π un campo de Markov sobre $S = \{1, 2, \dots, n\}$ con *potenciales de singletones y de pares*, $\pi(x) = C \exp\{\sum_{i \in S} \Phi_i(x_i) + \sum_{\{i,j\}} \Phi_{ij}(x_i, x_j)\}$. Para todo $x \in E^n$, $\pi(x) > 0$. Además se supone que se verifica la condición de identificabilidad (3.4) con $0 \in E$ el estado de referencia. La propiedad siguiente permite identificar los potenciales de X a partir de las leyes condicionales de X .

Theorem 2 (Besag [14];[27],[28]) *Supongamos que cada $\pi_i(\cdot|x^i)$ pertenece a la familia exponencial :*

$$\log \pi_i(x_i|x^i) = A_i(x^i)B_i(x_i) + C_i(x_i) + D_i(x^i), \quad (3.5)$$

donde $B_i(0) = C_i(0) = 0$. Entonces :

1. Para todo $i, j \in S$, $i \neq j$, existe α_i y $\beta_{ij} = \beta_{ji}$ tales que :

$$A_i(x^i) = \alpha_i + \sum_{j \neq i} \beta_{ij} B_j(x_j), \quad \Phi_i(x_i) = \alpha_i B_i(x_i) + C_i(x_i), \quad \Phi_{ij}(x_i, x_j) = \beta_{ij} B_i(x_i) B_j(x_j). \quad (3.6)$$

2. A la inversa, leyes condicionales que verifican (3.5) y (3.6) se encolan en una ley conjunta que es un campo de Markov de potenciales dados por (3.6).

Una sub-clase importante de campo de Markov es la de los campo de Gibbs con valores en $E \subseteq \mathbb{R}$ y potenciales no mayores que los potenciales de pares, siendo los potenciales de pares cuadráticos.

Definition 3.3 *Auto-Modelo de Markov (AMM) : X es un AMM si X es con valores en \mathbb{R} y si su ley π se escribe :*

$$\pi(x) = Z^{-1} \exp\left\{\sum_{i \in S} \Phi_i(x_i) + \sum_{\langle i,j \rangle} \beta_{ij} x_i x_j\right\} \text{ con, } \forall i, j, \beta_{ij} = \beta_{ji} \quad (3.7)$$

Auto-modelo logístico : $E = \{0, 1\}$ y $\forall i$, $\pi_i(\cdot|x^i) \sim \text{Logit}(\theta_i(x^i))$, $\theta_i(x_i) = \{\alpha_i + \sum_{j:\langle i,j \rangle} \beta_{ij} x_j\}$, t.q. $i \neq j$, $\beta_{ij} = \beta_{ji}$:

$$\pi_i(x_i|x^i) = \frac{\exp x_i \{\alpha_i + \sum_{j:\langle i,j \rangle} \beta_{ij} x_j\}}{1 + \exp\{\alpha_i + \sum_{j:\langle i,j \rangle} \beta_{ij} x_j\}}.$$

Esas leyes se encolan en una ley conjunta de energía $U(x) = \sum_i \alpha_i x_i + \sum_{\langle i,j \rangle} \beta_{ij} x_i x_j$.

Auto-modelo binomial : $E = \{0, 1, 2, \dots, N\}$. Consideramos leyes $\pi_i(\cdot|x_i) \sim \text{Bin}(N, \theta_i(x_i))$ t.q. $\theta_i(x_i)$ verifica :

$$A_i(x^i) = \log\{\theta_i(x_i)/(1 - \theta_i(x_i))\} = \alpha_i + \sum_{j:\langle i,j \rangle} \beta_{ij} x_j.$$

Si $\forall i \neq j$, $\beta_{ij} = \beta_{ji}$, esas leyes se encolan en una conjunta de energía $U(x) = \sum_i (\alpha_i x_i + \log \binom{N}{x_i}) + \sum_{\langle i,j \rangle} \beta_{ij} x_i x_j$. El parámetro binomial condicional es $\theta_i(x_i) = [1 + \exp\{-\{\alpha_i + \sum_{j \in S:\langle i,j \rangle} \beta_{ij} x_j\}]\}^{-1}$.

Para esos dos ejemplos, E es finito y luego U es siempre admisible. No es el caso en los ejemplos siguientes.

Auto-modelo de Poisson : $E = \mathbb{N}$. Supongamos que $\pi_i(\cdot|x^i) \sim \mathcal{P}(\lambda_i(x^i))$, $i \in S$ con parámetro log-lineales:

$$\log \lambda_i(x^i) = A_i(x^i) = \alpha_i + \sum_{j:(i,j)} \beta_{ij} x_j.$$

Si $\forall i \neq j$, $\beta_{ij} = \beta_{ji} \leq 0$, la energía vale $U(x) = \sum_i (\alpha_i x_i + \log(x_i!)) + \sum_{j:(i,j)} \beta_{ij} x_i x_j$. La energía U es admisible ssi $\beta_{ij} \leq 0$ para todo $i \neq j$. Esto traduce una *competencia entre sitios vecinos*. Se puede permitir *cooperación* de dos maneras : (i) acotando E por un $K < \infty$ (censura derecha); (ii) trabajando con las leyes de Poisson condicionales a $\{X \leq K\}$ [33, 4].

Auto-modelo exponencial : $E =]0, +\infty[$. Son modelos con condicionales exponencial :

$$\pi_i(x_i|x^i) \sim \mathcal{Exp}(\mu_i(x^i)), \quad \mu_i(x^i) = \{\alpha_i + \sum_{j:(i,j)} \beta_{ij} x_j\}.$$

Si, $\forall i \neq j$, $\alpha_i > 0$ et $\beta_{ij} = \beta_{ji} \geq 0$, esa leyes se encolan en una ley conjunta de energía $U(x) = -\sum_i \alpha_i x_i - \sum_{(i,j)} \beta_{ij} x_i x_j$. La positividad de $\beta_{ij} \geq 0$ traduce una *competencia entre los sitios*. Se puede autorizar *cooperación* de forma análoga a los *AMM* de Poisson.

Auto-modelo gaussiano : $E = \mathbb{R}$. Sea $X \sim \mathcal{N}_n(\mu, \Sigma)$ sobre $S = \{1, 2, \dots, n\}$ tal que $Cov(X) = \Sigma$ es inversible, $Q = \Sigma^{-1}$. X es un modelo de Gibbs de energía $U(x) = -(1/2)^t (x - \mu)Q(x - \mu)$, con potenciales $\Phi_i(x_i) = -\frac{1}{2} q_{ii} x_i^2 + \alpha_i x_i$, $\alpha_i = \sum_j q_{ij} \mu_j$, y $\Phi_{ij}(x_i, x_j) = -q_{ij} x_i x_j$. La energía U es admisible ssi Q es d.p.. Como $U_i(x_i|x^i) = \Phi_i(x_i) + \sum_{j \neq i} \Phi_{ij}(x_i, x_j)$, la ley condicional $\mathcal{L}_i(X_i|x^i)$ es gaussiana de varianza q_{ii}^{-1} y media $\mu_i(x^i) = -q_{ii}^{-1} \sum_{j \neq i} q_{ij}(x_j - \mu_j)$. Este modelo es un *CAR* gaussiano (cf. (2.9)).

Auto-modelo markoviano con covariables (AMMX) : sin restricciones, un modelo no estacionario es de dimensión demasiado grande para ser útil en modelación estadística. Un camino para reducir la dimensión es modelar los parámetros $\theta = \{(\alpha_i, \beta_{ij}), i \neq j, i, j \in S\}$ a partir de covariables exógenas $z = (z_i, i \in S)$ y de pesos dados $\{(a_i), (w_{ij})\}$, w simétrica. Por ejemplo, el modelo $\beta_{ij} = \delta w_{ij}$ y $\alpha_i = \sum_{j=1}^p \gamma_j a_i z_{ij}$ donde $z_i = {}^T(z_{i1}, \dots, z_{ip}) \in \mathbb{R}^p$ es observable, es de dimensión $p + 1$. Un ejemplo de *AMMX* de Poisson es estudiado en Ferrándiz et altri [26] para la modelación epidemiológica espacial de varios tipos de cánceres en la región de Valencia, España.

4 Proceso puntual espacial

Los procesos puntuales (PP) se utilizan en varias ramas (Diggle, [22]) tales que ecología y forestales (Matern, [38]), epidemiología espacial ([37]), ciencias de los materiales ([55]), geofísica, astrofísica ([43]), etc. La teoría probabilística de los PP necesita definiciones (i.e. medidas de Palm) y justificaciones técnicas que no tratamos aquí y que se pueden encontrar en los libros de Daley y Veres-Jones [21], Stoyan, Kendall y Mecke [57], van Lieshout [61], Møller y Waagepetersen [42].

4.1 Definiciones y notaciones

Sea S un conjunto de \mathbb{R}^d , \mathcal{B} (resp. $\mathcal{B}(S)$, resp. $\mathcal{B}_b(S)$) los borelianos de \mathbb{R}^d (resp. de S , resp. acotados de S), ν la medida de Lebesgue sobre \mathcal{B}_d . La realización x de un PP X sobre S es dado por una colección *localmente finita* de puntos de S , $x = \{x_1, x_2, \dots\}$, $x_i \in S$, es decir una parte $x \subset S$ t.q. $x \cap B$ es finita para toda parte $B \in \mathcal{B}_b(S)$. Denotamos por E el espacio de las configuraciones localmente finitas $,x, y, \dots$ tales configuraciones sobre S , x_i, y_i, ξ, η puntos de esas configuraciones.

4.1.1 Espacio exponencial de las configuraciones

Si $S \in \mathcal{B}_b(S)$, el espacio E de las configuraciones del PP X sobre S es la unión de los espacios E_n de las configuraciones con n puntos, $n \geq 0$: $E = \bigcup_{n \geq 0} E_n$ es el *espacio exponencial* de las configuraciones, equipado de la σ -álgebra \mathcal{E} que hace medibles las variables contadoras $N(A) : E \rightarrow \mathbb{N}$ para todos los $A \in \mathcal{B}_b(S)$. La ley de X es la probabilidad P imagen sobre (E, \mathcal{E}) de \mathbb{P} . Esta es caracterizada por las leyes finitas de esas variables.

Definition 4.1 *Proceso puntual.*

1. Un proceso puntual sobre S es una aplicación X de un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$ en E t.q., $\forall A \in \mathcal{B}_b(S)$, el número de puntos $N(A) = N_X(A)$ de X en A es medible.
2. La ley de un PP X es definida para todo $m \geq 1$ y (A_1, A_2, \dots, A_m) de $\mathcal{B}_b(S)$ por las leyes finitas sobre \mathbb{N}^m de $(N(A_1), N(A_2), \dots, N(A_m))$.

Si $S \in \mathcal{B}_b(S)$, $N_X(S)$ es casi-seguramente finito. Aquí, se estudiará únicamente PP *sencillos*, es decir PP sin repetición de puntos x_i : en tal caso, una realización x coincide con un sub-conjunto de S .

PP estacionario, PP isotrópico. Un PP sobre \mathbb{R}^d es *estacionario* si, $\forall \xi \in \mathbb{R}^d$, la ley del PP trasladado $X_\xi = \{X_i + \xi\}$ es la misma que la de X . El PP es *isotrópico* si la ley de ρX , la ρ -rotación de X , es la misma que la ley de X , eso para toda rotación ρ .

PP marcado (PPM). Sea K un espacio métrico, por ejemplo $K \subseteq \mathbb{R}^m$. Un *proceso puntual de marcado* (X, M) sobre $S \times K$ es un PP sobre $S \times K$ t.q. X sea un PP sobre S : $(x, m) = \{(x_1, m_1), (x_2, m_2), \dots\}$ donde $m_i \in K$ es la marca en x_i . Ejemplos de marcas son : $K = \{m_1, m_2, \dots, m_K\}$ (K tipos de puntos), $K = \mathbb{R}^+$ (la marca es una medida $r \geq 0$), $K = [0, 2\pi[\times \mathbb{R}^+$ (la marca es un segmento centrado en x , de orientación $\theta \in [0, 2\pi[$ y longitud $l \geq 0$). Examinemos algunos ejemplos de PPM.

- Sea $B(x, r) \subset S$ la bola de \mathbb{R}^2 de centro x y de radio $r > 0$. Un ejemplo de PP marcado es el dato de los centros x_i de un PP X y de marcas bolas $B(x_i, r_i)$ centradas en x_i y de radios i.i.d. independientes de X . En términos de morfología matemática [55, 57], $\mathcal{X} = \bigcup_{x_i \in X} B(x_i, r_i)$ es un *proceso booleano*.
- Un PP *marcado de fibras* [57] es asociado a marcas curvilíneas m_i , por ejemplo un segmento centrado en x_i , de longitud $l_i \sim \text{Exp}(l^{-1})$ y orientación θ_i , (θ_i, l_i) i.i.d. y independientes de X ; tal PPM se utilizan en ciencias de los suelos para modelar la repartición espacial de una red de raíces en $S \subset \mathbb{R}^3$.
- Un PP *multivariado* $X = (X(1), X(2), \dots, X(M))$ se puede ver como un PPM con M marcas : $X(m)$ es el sub-conjunto de S de las posiciones de la variedad m . X se identifica con $\tilde{X} = \bigcup_{m=1}^M X(m)$, el PPM superposición de los $X(m)$.

4.2 Momentos de un proceso puntual

Si los momentos de orden 1 (esperanza) y 2 (covarianza) son la base del estudio de los procesos con valores reales, las nociones adoptadas para los PP son las *medidas de momento de orden $p \geq 1$* (en particular $p = 1$ y $p = 2$), *medidas* sobre $(S, \mathcal{B}(S))^p$ definida por :

$$\mu_p(B_1 \times \dots \times B_p) = E(N(B_1) \dots N(B_p)).$$

Momento de orden 1, intensidad ρ de un PP. La medida de intensidad λ de X es la medida de momento de orden 1 :

$$\lambda(B) = \mu_1(B) = E(N(B)) = E\left\{\sum_{\xi \in X} \mathbf{1}(\xi \in B)\right\}.$$

En general, $\lambda(d\xi)$, la probabilidad de que haya un punto de X en el volumen $d\xi$ alrededor de ξ , se expresa a partir de una *densidad de intensidad* $\rho(\xi)$, $\lambda(d\xi) = \rho(\xi)d\xi$. Si X es estacionario, λ es invariante por traslación, es decir $\lambda(B) = \tau\nu(B)$ donde τ , la intensidad constante de X , es el número medio de puntos de X en un volumen unidad.

Momento factorial y intensidad ρ_2 de orden 2. La covarianza de variables contadoras se obtiene a partir de la medida del momento de orden 2, $\mu_2(B_1 \times B_2) = E(N(B_1)N(B_2))$. Sin embargo, anotando $\sum_{\xi, \eta \in X}^{\neq}$ la suma extendida a los sitios $\xi \neq \eta$ distintos de X , la descomposición,

$$\mu_2(B_1 \times B_2) = E\left\{\sum_{\xi \in X} \mathbf{1}(\xi \in B_1 \cap B_2)\right\} + \mathbb{E}\left\{\sum_{\xi, \eta \in X}^{\neq} \mathbf{1}((\xi, \eta) \in (B_1, B_2))\right\},$$

enseña que μ_2 tiene una componente medida sobre $\mathcal{B}(S)$ y otra medida sobre $\mathcal{B}(S) \times \mathcal{B}(S)$. Esta singularidad desaparece considerando la medida de *momento factorial* α_2 de orden 2 definida por :

$$\alpha_2(B_1 \times B_2) = E\left\{\sum_{\xi, \eta \in X}^{\neq} \mathbf{1}((\xi, \eta) \in (B_1, B_2))\right\} = \mu_2(B_1 \times B_2) - \lambda(B_1 \cap B_2).$$

α_2 coincide con μ_2 sobre el producto de eventos disjuntos.

Si $h_1 : \mathbb{R}^d \rightarrow [0, \infty)$ y $h_2 : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ son medibles, se tiene ([42]) :

$$E\left\{\sum_{\xi \in X} h(\xi)\right\} = \int_{\mathbb{R}^d} h_1(\xi)\lambda(d\xi) \text{ y } E\left\{\sum_{\xi, \eta \in X}^{\neq} h_2(\xi, \eta)\right\} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(\xi, \eta)\alpha_2(d\xi, d\eta). \quad (4.1)$$

Si B_1 y B_2 son acotados, $Cov(N(B_1), N(B_2)) = \alpha_2(B_1 \times B_2) + \mu_1(B_1 \cap B_2) - \mu_1(B_1)\mu_1(B_2)$. Si $\xi \neq \eta$, $\alpha_2(d\xi \times d\eta)$ es la probabilidad que X presenta un punto en $d\xi$ y un punto en $d\eta$. Si α_2 es absolutamente continua con respecto de la medida de Lebesgue sobre $(S, \mathcal{B}(S))^2$, su densidad $\rho_2(\xi, \eta)$ es la *densidad de intensidad de orden 2* de X . Si X es estacionario (resp. isotrópico), $\rho_2(\xi, \eta) = \rho_2(\xi - \eta)$ (resp. $\rho_2(\xi, \eta) = \rho_2(\|\xi - \eta\|)$).

Correlación de pares reponderada. Una pregunta central en el estudio de los PP es de saber si los puntos de la realización \mathbf{x} de un PP tienen una tendencia a atraerse (conglomerados o clusters), o al contrario a repelerse (regularidad), o a no influirse (independencia). Esta última hipótesis, denominada CSR para *Complete Spatial Randomness*, traduce que los puntos de \mathbf{x} se reparten en S independientemente unos de los otros, pero no necesariamente de manera uniforme sobre S . Esto define la clase de los PP de Poisson (PPP) con densidad $\rho(\cdot)$; los PPP estacionarios son PPP de densidad ρ constante.

Sin suponer la densidad estacionaria, Baddeley, Møller y Waagepetersen [7] definen la *correlación de pares reponderada* $g(\xi, \eta)$ que es un indicador al orden 2 de la dependencia espacial :

$$\text{si } \rho(\xi) > 0 \text{ y } \rho(\eta) > 0 : g(\xi, \eta) = \frac{\rho_2(\xi, \eta)}{\rho(\xi)\rho(\eta)}. \quad (4.2)$$

Como $g(\xi, \eta) = 1 + \frac{Cov(N(d\xi), N(d\eta))}{\rho(\xi)\rho(\eta)d\xi d\eta}$, la interpretación de g es la siguiente : (i) $g(\xi, \eta) = 1$: independencia entre los puntos (hipótesis CSR); (ii) $g(\xi, \eta) > 1$: atracción entre los puntos (la covarianza es > 0); (iii) $g(\xi, \eta) < 1$: repulsión entre los puntos (la covarianza es < 0).

Se dice que X es *estacionario de orden dos para la correlación reponderada* si $g(\xi, \eta) = g(\xi - \eta)$. Un PP puede ser estacionario al orden 2 para esta noción sin estacionalidad al orden 1 (cf. §4.5.1). Esta noción permite construir una prueba de independencia espacial sin suponer la estacionalidad de la densidad espacial.

4.3 Ejemplos de proceso puntual

Modelo condicional de $n(\mathbf{x}) = n \geq 0$. Si S es acotada, la ley de X puede definirse a partir de,

1. las probabilidades p_n que una configuración \mathbf{x} tiene n puntos, $n \geq 0$;
2. la *densidad condicional* g_n sobre E_n , las configuraciones con n puntos, $n \geq 1$.

La densidad g_n sobre E_n es en biyección con una densidad f_n sobre S^n invariante para las permutaciones de las coordenadas : $f_n(x_1, x_2, \dots, x_n) = \frac{1}{n!} g_n(\{x_1, x_2, \dots, x_n\})$. Si $B_n^* = \{(x_1, x_2, \dots, x_n) \in S^n \text{ t.q. } \{x_1, x_2, \dots, x_n\} \in B_n\}$ es asociado a $B_n \in \mathcal{E}_n$, la probabilidad de $B = \bigcup_{n \geq 0} B_n$ es :

$$P(X \in B) = \sum_{n \geq 0} p_n \int_{B_n^*} f_n(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

Especificar los p_n es poco realista en la práctica. Esos p_n serán implícitos si X está definido por una *densidad incondicional* (cf. §4.6)

Puntos independientes : proceso Binomial de n puntos. Sea S acotada de volumen $0 < \nu(S) < \infty$. Un *PP binomial* sobre S con n puntos coincide con la realización de n puntos i.i.d. uniformes sobre S : si $\{A_1, A_2, \dots, A_k\}$ es una partición boreliana de S , $(N(A_1), N(A_2), \dots, N(A_k))$ es la ley multinomial $\mathcal{M}(n; q_1, q_2, \dots, q_k)$ de parámetros $q_i = \nu(A_i)/\nu(S)$, $i = 1, k$. Eso se extiende a reparticiones i.i.d. no necesariamente con densidad uniforme $\rho(\cdot)$: eso define la repartición de un PP de Poisson de intensidad ρ con n puntos en S (cf. §4.4).

Más regularidad espacial : un modelo con núcleo duro. Una manera de “regularizar” x es prohibir puntos demasiado cercanos. Eso es adoptado cada vez que un individuo i en x_i necesita un espacio propio alrededor de él : esferas no penetrables en física, árboles en una selva, animales sobre un territorio de caza, centros de células de un tejido celular, panaderías en una ciudad. Esos modelos son casos particulares del modelo de Strauss que es un modelo de Gibbs definido por su densidad incondicional (cf. §4.6.2).

Repartición con clusters : el PP de Neyman-Scott. Examinamos la modelación siguiente de una dinámica de población :

1. las posiciones X de los padres forman un PP de Poisson;
2. a la generación siguiente, cada padre i en x_i genera hijos Y_{x_i} en número K_{x_i} y posiciones Y_{x_i} alrededor de x_i , con (K_{x_i}, Y_{x_i}) i.i.d. independientes de X .

El PP de Neyman-Scott [43] es la superposición $D = \cup_{x_i \in x} Y_{x_i}$ de los hijos de la primera generación. D presenta clusters cerca de los padres. Las generalizaciones son múltiples: otra repartición espacial de los padres, dependencias entre los hijos (competencia), leyes de descendencia no idénticas (variabilidad de fertilidad), etc. Esos modelos son de la clase de los PP de Cox (cf. §4.5).

Presentamos ahora las tres grandes clases de PP :

1. Los PP de Poisson (PPP) que modelan reparticiones espaciales “al azar” y son caracterizados por su intensidad $\rho(\cdot)$;
2. Los PP de Cox que son PPP condicionales a un contexto aleatorio. modelan reparticiones espaciales *menos regulares* y pueden presentar clusters como el PP de Neyman-Scott.
3. Los PP de Gibbs son definidos a partir de la densidad de una especificación condicional. Sirven para modelar reparticiones espaciales *más regulares* que las de un PP de Poisson, como la de un modelo con núcleo duro.

4.4 Proceso puntual de Poisson

Sea λ una medida positiva sobre $(S, \mathcal{B}(S))$ de densidad ρ , λ finita sobre $\mathcal{B}_b(S)$. Un *PP de Poisson (PPP) de medida de intensidad* $\lambda(\cdot) > 0$ y densidad $\rho(\cdot)$ se caracteriza por :

1. $\forall A \in \mathcal{B}_b(S)$ t.q. $\lambda(A) < \infty$, $N(A)$ sigue una ley de Poisson de parámetro $\lambda(A)$;
2. Condicionalmente a $N(S) = n$, los puntos de $x \cap A$ son i.i.d. con densidad proporcional a $\rho(\xi)$, $\xi \in A$:

$$P(N(A) = n) = e^{-\lambda(A)} \frac{(\lambda(A))^n}{n!} \text{ y } g_n(\{x_1, x_2, \dots, x_n\}) \propto \rho(x_1)\rho(x_2) \cdots \rho(x_n).$$

Si $A_i, i = 1, p$ son p borelianos disjuntos de $\mathcal{B}_b(S)$, las v.a. $N(A_i)$ son independientes. El PPP es *homogéneo* de intensidad ρ si $\lambda(\cdot) = \rho\nu(\cdot)$: a la independencia espacial se agrega la uniformidad de la repartición espacial.

$\rho(\cdot)$ se puede modelar de manera log-lineal $\log \rho(\xi) = {}^t z(\xi)\beta$, $\beta \in \mathbb{R}^p$ con covariables conocidas $z(\xi)$, $\xi \in S$.

Simulación de un PPP : Se utiliza el algoritmo de rechazo, o borrado de puntos : si $\rho(x) < c < \infty$ sobre S ,

1. simular $X_h = \{x_i\}$, un *PPP* homogéneo sobre S de intensidad c ;
2. borrar independientemente los x_i con la probabilidad $(1 - \frac{\rho(x_i)}{c})$.

4.5 Proceso puntual de Cox

Sea $\Lambda = (\Lambda(\xi))_{\xi \in S}$ un proceso ≥ 0 sobre S t.q., casi-seguramente y $\forall B \in \mathcal{B}_b(S)$, $\int_B \Lambda(\xi) d\xi < \infty$. Un PP de Cox X dirigido por $\Lambda = (\Lambda(\xi))_{\xi \in S}$ es un PPP de densidad aleatoria Λ ; Λ modela un contexto aleatorio. Si la densidad de Λ es estacionaria, X es estacionario.

4.5.1 Proceso de Cox log-gaussiano

Introducido por Møller, Syversveen et Waagepetersen [41], esos modelos admiten una intensidad Λ log-lineal con efecto aleatorio :

$$\log \Lambda(\xi) = {}^t z(\xi)\beta + \Psi(\xi). \quad (4.3)$$

Aquí, $\Psi = (\Psi(\xi))_{\xi \in S}$ es un proceso gaussiano centrado de covarianza $c(\xi, \eta) = Cov(\Psi(\xi), \Psi(\eta))$ t.q. c asegura la integrabilidad local de Λ . Las medidas de momentos de esos procesos se manipulan bien : en particular, $\log \rho(\xi) = {}^t z(\xi)\beta + c(\xi, \xi)/2$ y la correlación de pares reponderada $g(\xi, \eta) = \exp(c(\xi, \eta))$. La correspondencia $c \longleftrightarrow g$ es biyectiva. Si Ψ es estacionario (resp. isotrópico), X es estacionario (resp. isotrópico) al orden 2 para la correlación reponderada g .

4.5.2 PP doblemente de Poisson

También llamado *shot noise process*, son PP de Cox X de intensidad,

$$\Lambda(\xi) = \sum_{(c, \gamma) \in \varphi} \gamma k(c, \xi). \quad (4.4)$$

En esta escritura, φ es la realización de un PPP sobre $S \times \mathbb{R}^+$, $k(c, \cdot)$ una densidad sobre S centrada en c . Así, X es la superposición de PPP independientes de intensidades $\gamma k(c, \cdot)$ para una repartición de Poisson de los (c, γ) . Esos PP se pueden utilizar para modelar una dinámica de población.

Para el PP de Neyman-Scott ([43] y §4.3), los centros c forman un PPP estacionario de intensidad τ y γ es el número medio de hijos nacidos de cada padre; este PP es estacionario de intensidad $\tau\gamma$. El PP de Thomás [60] es asociado a una ley de dispersión gaussiana $k(c, \cdot) \sim \mathcal{N}_d(c, \sigma^2 I)$; el PP de Thomás es isotrópico de correlación reponderada [42],

$$g(\xi, \eta) = g(\|\xi - \eta\|) \text{ con } g(r) = 1 + \frac{1}{4\pi\kappa\sigma^2} \exp\left\{-\frac{r^2}{4\sigma^2}\right\}.$$

Se puede también considerar un modelo inhomogéneo donde Λ depende de covariables $z(\xi)$ [63],

$$\Lambda(\xi) = \exp({}^t z(\xi)\beta) \sum_{(c, \gamma) \in \varphi} \gamma k(c, \xi).$$

De misma correlación reponderada g que (4.4), este PP permite estudiar propiedades de dependencia al orden 2 en presencia de una inhomogeneidad al orden 1.

4.6 Densidad de un proceso puntual

Para $S \in \mathcal{B}_b(\mathbb{R}^d)$, vamos a definir la densidad de probabilidad f de un PP X sobre S con respecto a una medida de referencia sobre (E, \mathcal{E}) , la de un $PPP(1)$. Como es el caso muchas veces, la densidad de un PP será conocida salvo un factor multiplicativo cercano, $f(x) = cg(x)$, g conocida. Eso no tiene consecuencia para la simulación MCMC de un PP vía el algoritmo de Metrópolis para el cual basta conocer g (cf. Møller et Waagepetersen [42]; [27], Ch. 4; [28]). Pero para la estimación paramétrica de θ del modelo $f_\theta(x) = c(\theta)g_\theta(x)$ por MV necesitará conocer $c(\theta)$, constante analíticamente incalculable que tendrá que ser numéricamente aproximada mediante un método MCMC ([42]; [27], Ch. 5; [28]). Otro método de estimación que evita esta dificultad es el método de Máxima Seudo-Verosimilitud Condicional (MSVC) para el cual basta conocer $g_\theta(x)$.

4.6.1 Definición

Sea Y_ρ un $PPP(\rho)$. El desarrollo de Poisson siguiente, para un evento $F \in \mathcal{E}$:

$$P(Y_\rho \in F) = \sum_{n=0}^{\infty} \frac{e^{-\lambda(S)}}{n!} \int_{S^n} \mathbf{1}\{\mathbf{x} \in F\} \rho(x_1) \rho(x_2) \cdots \rho(x_n) dx_1 dx_2 \cdots dx_n,$$

y $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ permite definir la densidad de un PP con respecto a la medida sobre E de $Y_1 \sim PPP(1)$.

Definition 4.2 X es de densidad f con respecto a Y_1 si para cada $F \in \mathcal{E}$:

$$P(X \in F) = E[\mathbf{1}\{Y_1 \in F\} f(Y_1)] = \sum_{n=0}^{\infty} \frac{e^{-\nu(S)}}{n!} \int_{S^n} \mathbf{1}\{\mathbf{x} \in F\} f(x) dx_1 dx_2 \cdots dx_n.$$

La probabilidad de E_n , las configuraciones con n puntos, es $p_n = \frac{e^{-\nu(S)}}{n!} \int_{S^n} f(x) dx_1 dx_2 \cdots dx_n$. Además, condicionalmente a $n(\mathbf{x}) = n$, los puntos de \mathbf{x} son de densidad $f_n(x_1, x_2, \dots, x_n) \propto f(\{x_1, x_2, \dots, x_n\})$: f_n es conocida salvo un factor multiplicativo cercano, los p_n son dados implícitamente pero numéricamente incalculables. Una de las pocas situaciones donde esas cantidades se explicitan es la de un $PPP(\rho)$ para el cual $f(\mathbf{x}) = e^{\nu(S) - \lambda(S)} \prod_{i=1}^n \rho(x_i)$.

Intensidad condicional de Papangelou. Una densidad f es *hereditaria* si se tiene:

$$\forall \mathbf{x} \in E \text{ t.q. } f(\mathbf{x}) > 0 \text{ y si } \mathbf{y} \subset \mathbf{x}, \text{ entonces } f(\mathbf{y}) > 0. \quad (4.5)$$

Las densidades usuales son hereditarias. Si f es hereditaria, la *intensidad de Papangelou* de $\xi \notin \mathbf{x}$ condiciona a \mathbf{x} es [44]:

$$\lambda(\xi, \mathbf{x}) = \frac{f(\mathbf{x} \cup \{\xi\})}{f(\mathbf{x})} \text{ si } f(\mathbf{x}) > 0, \quad \lambda(\xi, \mathbf{x}) = 0 \text{ sinon.} \quad (4.6)$$

Si $f(x) = cg(x)$, $\lambda(\xi, x)$ no depende de c . Papangelou (4.6) establece que la correspondencia entre f y λ es biyectiva: así, un PP puede modelarse a partir de su intensidad condicional. Esta intensidad está en la base de las simulaciones MCMC de un PP definido por su densidad o de la estimación paramétrica por MSVC.

La interpretación de $\lambda(\xi, x)$ es la probabilidad de presencia de un punto de X en ξ si la realización fuera de ξ es x . Se tiene: $E(\lambda(\xi, X)) = \rho(\xi)$ si la densidad de X es ρ . Para un $PPP(\rho)$, $\lambda(\xi, x) = \rho$ no depende de x . Para los PP de Markov (cf. 4.9.1), $\lambda(\xi, x) = \lambda(\xi, x \cap \partial\xi)$ depende únicamente de la configuración de x al vecindad $\partial\xi$ de ξ .

4.6.2 Proceso puntual de Gibbs

Supongamos que X es de densidad $f(x) = \exp\{-U(x)\}/Z$ con una energía $U(x)$ *admisibile*, es decir:

$$\forall n \geq 0 : q_n = \int_{S^n} \exp\{-U(x)\} dx_1 dx_2 \cdots dx_n < \infty \text{ y } \sum_{n=0}^{\infty} \frac{e^{-\nu(S)}}{n!} q_n < \infty.$$

Una condición suficiente de admisibilidad es que $\exists n_0 < \infty$ t.q. $\forall x, n(x) \leq n_0$ y que U sea admisibile condicionalmente para todo $n \leq n_0$ (i.e. X es un modelo de núcleo duro). Otra condición suficiente de admisibilidad es que U esté acotada. Si $U(x) = \sum_{y \subset x} \phi(y)$, se dice que X es un *PP de Gibbs de potencial* ϕ .

PP de Strauss, PP con núcleo duro. El ejemplo de base es la energía con *potenciales de singletons y de pares* :

$$U(x) = \sum_{i=1}^n \varphi(x_i) + \sum_{i=1}^n \sum_{j>i}^n \psi(\|x_i - x_j\|).$$

Si $r > 0$ es un radio dado, el PP de *Strauss* [58] tiene como potenciales :

$$\varphi(x_i) = a \text{ y } \psi_{\{x_i, x_j\}}(x) = b \mathbf{1}(\|x_i - x_j\| \leq r).$$

Su densidad es $f_\theta(x) = c(\theta) \exp(t\theta T(x))$ con $\theta = (a, b) \in \mathbb{R}^2$, $T_1(x) = n(x)$ y $T_2(x) = s(x) = \sum_{i<j} \mathbf{1}(\|x_i - x_j\| \leq r)$: T_2 cuenta el número de pares “ r -vecinos” de x . Si $\beta = e^a$ y $\gamma = e^b$,

$$f_\theta(x) = c(\beta, \gamma) \beta^{n(x)} \gamma^{s(x)}.$$

El PPP homogéneo corresponde a $\gamma = 1$; el PP con *núcleo duro* a $\gamma = 0$, con densidad $f_{\beta,r}(x) = c\beta^{n(x)} \mathbf{1}\{\forall i \neq j, \|x_i - x_j\| > r\}$ y como S está acotado, $n(x)$ es acotado y la energía de PP con núcleo duro es admisible.

Algunas distribuciones asociadas al PP de Strauss.

- *condicionalmente* a $n(x) = n$, $f_{\theta,n}(x) \propto \gamma^{s(x)}$;
- si $\gamma < 1$, X es más regular que el PP binomial, todavía más si γ está cerca de 0;
- $\gamma = 1$ da el PP Binomial a n puntos;
- si $\gamma > 1$, X es menos regular que un PP binomial : clusters aparecen.
- *Incondicionalmente*, f_θ es admisible ssi $\gamma \leq 1$ [51].

Algunas generalizaciones del PP de Strauss : se puede,

1. prohibir pares con distancia menor que r_0 , $0 < r_0 < r$: $s(x) = \sum_{i<j} \mathbf{1}(r_0 \leq \|x_i - x_j\| \leq r)$ (PP de Strauss con núcleo duro);
2. “saturar” los $n_{x_i}(x) = \sum_{\xi \in x: \xi \neq x_i} \mathbf{1}_{\|x_i - \xi\| \leq r}$ a un $\delta < \infty$ dada (PP de saturación de Geyer);
3. modelar el potencial de pares con una función con k escalones :

$$\phi(x_i, x_j) = \gamma_k \text{ si } d(x_i, x_j) \in (r_{k-1}, r_k], \phi(x_i, x_j) = 0 \text{ sino, } k = 1, p,$$

para $r_0 = 0 < r_1 < r_2 < \dots < r_p = r$ dados. Si $s_k(x)$ es el número de pares con distancia en $(r_{k-1}, r_k]$, $k = 1, \dots, p$:

$$f_\theta(x; \beta, \gamma_1, \dots, \gamma_p) = c(\theta) \beta^{n(x)} \prod_{k=1}^p \gamma_k^{s_k(x)};$$

f_θ est admisible ssi $\gamma_1 \leq 1$.

4. con potenciales de *tripletas* (o más) de puntos dos a dos con distancia $\leq r$: si $t(x)$ cuenta esas tripletas, $f_\theta(x) = c(\theta) \beta^{n(x)} \gamma^{s(x)} \delta^{t(x)}$ es admisible ssi γ y $\delta \leq 1$.

PP con interacción de área o con interacción de conectividad. Son otras alternativas que permiten modelaciones, sin restricciones paramétricas, de reparticiones más o menos regulares que la de un PPP ([61]; [?]). Tienen como densidad,

$$f_{\theta}(x) = c\beta^{n(x)}\gamma^{h(x)}$$

donde h es para definir. Para $r > 0$, sea $B(x) = \bigcup_{x_i \in x} B(x_i, r/2) \cap S$ la unión (limitada a S) de las bolas de centros $x_i \in x$ y radios r , $a(x)$ la superficie de $B(x)$ y $c(x)$ el número de sus componentes conexas. El PP con *interacción de área* corresponde a $h(x) = a(x)$, el PP a *interacción de conectividad* a $h(x) = c(x)$. S siendo acotado y $r > 0$, las funciones a y c son acotadas y las densidades son admisibles sin restricciones sobre los parámetros (β, γ) . Para cada modelo, la repartición espacial va a ser más (resp. menos) regular si $\gamma < 1$ (resp. $\gamma > 1$) y $\gamma = 1$ da el PPP(β). Una dificultad en la utilización de esos modelos es el cálculo numérico de $a(x)$ y de $c(x)$.

4.7 Distancias a los vecinos más cercanos (vmc)

Esas estadísticas son útiles para construir pruebas de hipótesis “CSR” de la independencia espacial de un PP. Los resultados sobre sus leyes utilizan la noción de medidas de Palm del PP, P_{ξ} para $\xi \in S$. Heurísticamente, P_{ξ} est la ley del PP condicional a la presencia de un punto $\xi \in x$ (cf. [21]; [61]; [42]). Se definen dos distancias a los vmc de X .

Distancia de $\xi \in X$ a su vmc en X : es la distancia $d(\xi, X \setminus \{\xi\})$ de un punto $\xi \in X$ a su vecino más cercano en X . Esta distancia tiene como función de repartición $G_{\xi}(r) = P_{\xi}(d(\xi, X \setminus \{\xi\}) \leq r)$, $r \geq 0$. Si X es estacionario, $G_{\xi} = G$ es independiente de ξ . Si X es un PPP homogéneo sobre \mathbb{R}^2 , $G(r) = P(N(B(\xi, r)) = 0) = 1 - \exp\{-\lambda\pi r^2\}$, con esperanza $(2\sqrt{\lambda})^{-1}$ y varianza $\lambda^{-1}(\pi^{-1} - 0.25)$.

Distancia de un punto corriente $u \in S$ a su vmc en X : es la distancia $d(u, X)$ a un punto corriente $u \in S$ al vecino más cercano de X . Tiene como función de repartición $F_u(r) = P(d(u, X) \leq r)$, $r \geq 0$. Si X es estacionario, $F_u = F$. Si X es un PPP homogéneo, $G(r) = F(r)$. Por lo tanto, un *indicador del carácter PPP homogéneo* de X es la estadística :

$$J(r) = \frac{1 - G(r)}{1 - F(r)}.$$

Si X es un PP estacionario, $J > 1$, $J = 1$ y $J < 1$ indican respectivamente que X presenta más, tanto y menos regularidad que un PPP homogéneo.

4.8 Momento reducido de orden 2

4.8.1 El momento K de Ripley [48]

Sea X un PP *isotrópico* sobre \mathbb{R}^d con intensidad ρ . Un otro indicador de orden 2 de la repartición espacial de X es la función K de Ripley o *momento reducido de orden 2* :

$$K(h) = \frac{1}{\rho} E_{\xi}[N(B(\xi, h) \setminus \{\xi\})], \quad h \geq 0, \quad (4.7)$$

donde E_{ξ} es la esperanza para la ley de Palm P_{ξ} . Hay dos interpretaciones de K :

1. $\rho K(h)$ es proporcional al número medio de puntos de X en la bola $B(\xi, h)$ si $\xi \in X$ pero sin considerar a ξ ;
2. $\rho^2 K(h)/2$ es el número medio de pares de puntos distintos con distancia $\leq h$, uno de esos puntos perteneciendo a un sub-conjunto fijado A de superficie unidad.

Si $\rho_2(\xi, \eta) = \rho_2(\|\xi - \eta\|)$ es la densidad de orden dos de X y b_d el volumen de la esfera unidad de \mathbb{R}^d , $K(h) = \frac{d \times b_d}{\rho} \int_0^h \xi^{d-1} \rho_2(\xi) d\xi$. En particular, para $d = 2$,

$$\rho^2 K(h) = 2\pi \int_0^h u \rho_2(u) du \text{ y } \rho_2(h) = \frac{\rho^2}{2\pi h} \frac{dK}{dh}(h). \quad (4.8)$$

Si X es un PPP homogéneo sobre \mathbb{R}^d , $K(h) = b_d \times h^d$. Si X es un PP de Neyman-Scott con padres siguiendo un $PPP(\rho)$ homogéneo, cada padre x_i generando N hijos de leyes de dispersión centrado en x_i e isotrópica alrededor de x_i y de la densidad $k(\xi) = g(\|\xi\|)$, entonces, con $G(h) = \int_0^h g(u) du$, se tiene [19],

$$K(h) = b_d h^d + \frac{E(N(N-1))G(h)}{\rho[E(N)]^2}.$$

Otro indicador del carácter de Poisson de un PP es dado por la función $L(h) = \left\{ \frac{K(h)}{b_d} \right\}^{1/d}$: si X es un PPP homogéneo, $h \mapsto L(h) = h$ es una recta; L cóncava indica una repartición con clusters; L convexa indica una repartición más regular que la de un PPP.

4.8.2 Momento K_{BMW}

Si X es un PP estacionario de orden dos para la correlación reponderada g (4.2), Baddeley-Møller-Waagepetersen [7] han extendido la función K de Ripley al momento:

$$K_{BMW}(h) = \int_{\mathbb{R}^d} \mathbf{1}\{\|\xi\| \leq h\} g(\xi) d\xi = \frac{1}{\nu(B)} E \left[\sum_{\xi, \eta \in X \cap B}^{\neq} \frac{\mathbf{1}\{\|\xi - \eta\| \leq h\}}{\rho(\xi)\rho(\eta)} \right]. \quad (4.9)$$

Una vez estimada, K_{BMW} permite construir estadísticas para la prueba de hipótesis del carácter ‘‘CSR’’ de Poisson de un PP *sin suponer* la estacionalidad al orden 1 de X .

4.9 Proceso puntual de Markov

La noción de *PP de Markov* fue introducida por Ripley et Kelly [51]. Baddeley y Møller [6] la generalizaron a la propiedad de *Markov a los vecinos más cercanos* (vmc).

4.9.1 Propiedad de Markov de Ripley-Kelly

Sea X un PP sobre S de densidad f con respecto a un $PPP(\lambda)$, λ una medida con densidad positiva y finita sobre $\mathcal{B}_b(S)$. Sea $\xi \sim \eta$ una relación de vecindad simétrica sobre S (i.e. $\xi \sim_r \eta$ si $\|\xi - \eta\| \leq r$). Sea $\partial A = \{\eta \in S \text{ et } \eta \notin A : \exists \xi \in A \text{ t.q. } \xi \sim \eta\}$ la vecindad de $A \subset S$ y $\partial\{\xi\} = \partial\xi$ si $\xi \in S$.

Definition 4.3 *Un proceso X de densidad f hereditaria (cf. (4.5)) es de Markov para la relación \sim si, $\forall x$ con $f(x) > 0$, la intensidad condicional de Papangéou $\lambda(\xi, x) = f(x \cup \{\xi\})/f(x)$ depende unicamente de ξ y de $\partial\xi \cap x$:*

$$\lambda(\xi, x) = \frac{f(x \cup \{\xi\})}{f(x)} = \lambda(\xi; x \cap \partial\xi).$$

La propiedad de Markov, local en ξ , se extiende a todo $A \subseteq S$: si X es de Markov, la ley de $X \cap A$ condicional a $X \cap A^c$ depende unicamente de $X \cap \partial A \cap A^c$. Algunos ejemplos son :

1. Si $X \sim PPP(\rho)$, $\lambda(u, x) \equiv \rho(\xi)$: un PPP es de Markov para cualquier relación de vecindad sobre S .
2. Para el PP de Strauss (4.6.2), $\lambda(\xi, x) = \beta \exp\{\log \gamma \times \sum_i \mathbf{1}\{\|x_i - \xi\| \leq r\}\}$: el PP de Strauss es de Markov para la relación \sim_r ; sus generalizaciones (cf. §4.6.2) son también \sim_r markovianas.
3. El PP con núcleo duro tiene por intensidad condicional $\lambda(\xi, x) = \beta \mathbf{1}\{\partial \xi \cap x = \emptyset\}$: es \sim_r -markoviano.

Al contrario, el PP con interacción de conexidad no es \sim_r markoviano : la razón intuitiva de eso es la siguiente : dos puntos de S pueden ser conexos aunque estén arbitrariamente lejos para la distancia euclidiana. Eso es una de la justificación de la extensión a la noción de PP de Markov a los vmc (cf. §4.9.2).

Como para los campos de Markov sobre una red discreta, un teorema de Hammersley-Clifford caracteriza la densidad de un PP de Markov a partir de potenciales definidos sobre las cliques del grafo de vecindad : una clique de la configuración $x = \{x_1, x_2, \dots, x_n\}$ para \sim es una sub-configuración $y \subseteq x$ t.q. para toda $\{x_i, x_j\} \subseteq y$, $x_i \sim x_j$ (más la convención que los singletons son cliques). Sea \mathcal{C} la familia de las cliques de (S, \sim) .

Proposition 10 [51, 61] *Un PP de densidad f es de Markov para la relación \sim ssi existe una función $\Phi : E \rightarrow (\mathbb{R}^+)^*$ medible t.q. :*

$$f(x) = \prod_{y \subset x, y \in \mathcal{C}} \Phi(y) = \exp \sum_{y \subset x, y \in \mathcal{C}} \phi(y).$$

$\phi = \log \Phi$ es el potencial de interacción de Gibbs del PP; su intensidad condicional de Papangelou es $\lambda(u, x) = \prod_{y \subset x, y \in \mathcal{C}} \Phi(y \cup \{u\})$. Un ejemplo de densidad de PP de Markov con interacciones de pares es :

$$f(x) = \alpha \prod_{x_i \in x} \beta(x_i) \prod_{x_i \sim x_j, i < j} \gamma(x_i, x_j).$$

El PP de Strauss corresponde a los potenciales $\beta(x_i) = \beta$ y $\gamma(x_i, x_j) = \gamma \mathbf{1}\{\|x_i - x_j\| \leq r\}$.

Propiedad de Markov para un PP marcado. La definición de la propiedad de Markov y el resultado de Hammersley-Clifford no cambian si $Y = (X, M)$ es un PP marcado sobre $S \times K$ para \sim , una relación de vecindad sobre $S \times K$. Si (X, M) es con marcas independientes y si X es de Markov para la relación \sim sobre S , (X, M) es un PPM de Markov para la relación $(x, m) \sim (y, o) \iff x \sim y$. Un ejemplo de modelo isotrópico con interacciones de pares y número finito de marcas $M = \{1, 2, \dots, K\}$ es dado por la densidad en $y = \{(x_i, m_i)\}$, $f(y) = \alpha \prod_i \beta_{m_i} \prod_{i < j} \gamma_{m_i, m_j}(\|x_i - x_j\|)$: si, para $r_{kl} > 0$, $k, l \in K$, $k \neq l$, K reales dados, $\gamma_{kl}(d) \equiv 1$ para $d > r_{kl}$, Y es un PPM de Markov para la relación de vecindad $(\xi, m) \sim (\xi', m') \iff \|\xi - \xi'\| \leq r_{m, m'}$.

Ejemplo : interacción de recubrimiento en forestales. Supongamos que la zona de influencia de un árbol centrado en x_i sea el disco $B(x_i; m_i)$, con los $m_i \leq m < \infty$. Una interacción de competencia entre dos árboles i y j puede modelarse por $\Phi_2((x_i; m_i), (x_j; m_j)) = b \times \nu(B(x_i; m_i) \cap B(x_j; m_j))$. En cuanto a los potenciales de singletons, podemos tomar, para K umbrales dados $0 = r_0 < r_1 < r_2 < \dots < r_{K-1} < r_K = m < \infty$:

$$\Phi_i(x_i; m_i) = \alpha(m_i) = a_k \text{ si } r_{k-1} < m_i \leq r_k, k = 1, K.$$

La energía asociada es admisible si $b < 0$ (competencia entre los árboles) y el PP marcado asociado es de Markov sobre $\mathbb{R}^2 \times \mathbb{R}^+$ para la relación $(x, m) \sim (x', m') \iff \|x - x'\| \leq 2m$. Su intensidad condicional es,

$$\lambda((u, h); (x, m)) = \exp\{\alpha(h) + b \sum_{j: \|x_j - u\| \leq 2m} \nu(B(u; h) \cap B(x_j; m_j))\}$$

[42, 5] presentan otros ejemplos.

4.9.2 Propiedad de Markov a los vmc

Una propiedad de Markov más general que la de Ripley-Kelly es la propiedad de *Markov a los vecinos más cercanos* (vmc), noción desarrollada por Baddeley y Møller [6]. La presentamos brevemente con un ejemplo, el del PP con interacción de conexidad. Este PP no es de Markov al sentido de Ripley-Kelly pero lo es para una otra relación de vecindad, la relación \sim_x a los x -vmc, ella es la misma función de la configuración x .

Sea x una configuración sobre S , $r > 0$ dado y $B(x) = \bigcup_{x_i \in x} B(x_i, r/2) \cap S$: dos puntos ξ y η de x son conectados por x si ξ y η son en la misma componente conexa de $B(x)$; se denota $\xi \sim_x \eta$: \sim_x es una relación sobre los puntos de x dependiente de x .

Sea $c(x)$ el número de componentes conexas de $B(x)$. Si S es acotado, la densidad del PP con interacción de conexidad vale $f(x) = c\alpha^{n(x)}\beta^{c(x)}$, α y $\beta > 0$ y la intensidad condicional de Papangelou es $\lambda(\xi, x) = \alpha\beta^{c(x \cup \{\xi\}) - c(x)}$. Si $\eta \in x$ no está conectado a ξ en $B(x \cup \{\xi\})$, η no contribuye a la diferencia $c(x \cup \{\xi\}) - c(x)$ y $\lambda(\xi, x)$ no depende de η : X es de Markov para la relación \sim_x a los vmc para las componentes conexas.

Ilustramos esta propiedad de Markov a los vmc con otro ejemplo, la de la relación con los vmc para la *triangulación de Delaunay*. Sea x una configuración localmente finita de puntos de $S \subset \mathbb{R}^2$. A cada $x_i \in x$ se asocia su *zona de influencia* $\mathcal{P}_i(x)$ definida por: $\xi \in \mathcal{P}_i(x) \iff \forall j \neq i, \|\xi - x_i\| \leq \|\xi - x_j\|$. Si $x = \{x_1, x_2, \dots, x_n\}$, la descomposición $S = \bigcup_{i=1}^n \mathcal{P}_i(x)$ se llama la *mosaica de Voronoi* de x . Salvo eventualmente zonas $\mathcal{P}_i(x)$ en la frontera de S , $\mathcal{P}_i(x)$ es un polígono convexo. A esta mosaica se asocia la *triangulación de Delaunay* $t(x)$ de x , dos puntos x_i y x_j distintos de x siendo vecinos si $\mathcal{P}_i(x)$ y $\mathcal{P}_j(x)$ tienen un lado común. Eso define una relación $\sim_{t(x)}$, la relación a los vmc para la triangulación de Delaunay $t(x)$ de x . Si ϕ y ψ son potenciales acotados de singletones y de pares, la densidad de Gibbs siguiente es de Markov a los vmc para la relación $\sim_{t(x)}$,

$$f(x) = c \exp\left\{ \sum_{i=1, n} \phi(x_i) + \sum_{x_i \sim_{t(x)} x_j} \psi(x_i, x_j) \right\}.$$

RECEIVED DECEMBER, 2008.

REVISED APRIL 2009.

References

- [1] ABRAMOWITZ, M. & STEGUN, I. A., (EDS.) [1970], *Handbook of Mathematical Functions*, Dover, New York,
- [2] ADLER, R.J. [1981], *The Geometry of Random Fields*, Wiley, New York,
- [3] ARNOLD, B.C., CASTILLO, E. & SARABIA, J.M. [1999], *Conditional Specification of Statistical Models*, Springer, New York,
- [4] AUGUSTIN, N.H., MCNICOL, J.W. & MARRIOTT, C.A. [2006], Using the truncated auto-Poisson model for spatially correlated counts of vegetation, *Journal of Agricultural, Biological & Environmental Statistics*, **11**, 1-23.

- [5] BADDELEY, A.J., GREGORI, P., MATEU, J., STOICA, R. & STOYAN, D., EDS. [2006], *Case studies in Spatial Point Processes Modeling*, Lecture Notes in Statistics 185, New York, Springer, New York,
- [6] BADDELEY, A.J. & MØLLER, J. [1989], Nearest-neighbour Markov point processes and random sets, *International Statistical Review*, **57**, 90-121.
- [7] BADDELEY, A.J., MØLLER, J. & WAAGEPETERSEN, R. [2000], Non- and semi-parametric estimation of interaction in inhomogeneous point patterns, *Statistica Neerlandica*, **54**, 329-350.
- [8] BADDELEY, A.J. & TURNER, R. [2005], Spatstat: an R package for analyzing spatial point patterns, *Journal of Statistical Software*, **12**, pp. 1-42.
- [9] BADDELEY, A.J. & VAN LIESHOUT, M.N.M. [1995], Area-interaction point processes, *Annals of the Institute of Statistical Mathematics*, **46**, 601-619.
- [10] BANERJEE, S., CARLIN, B.P. & GELFAND, A.E. [2004], *Hierarchical Modeling and Analysis for Spatial Data*, Chapman and Hall/CRC Press, Boca Raton: FL.
- [11] BARTLETT, M.S. [1971], Physical nearest-neighbour models and non-linear time series (I), *Journal of Applied Probability*, **8**, 222-232.
- [12] BARTLETT, M.S. [1972], Physical nearest-neighbour models and non-linear time series (II), *Journal of Applied Probability*, **9**, 76-86.
- [13] BESAG, J. [1972], On the correlation structure of some two dimensional stationary processes, *Biometrika*, **59**, 43-48.
- [14] BESAG, J. [1974], Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society, Series B*, **36**, pp. 192-236.
- [15] BROOK, D. [1964], On the distinction between the conditional probability and joint probability approaches in the specification of nearest neighbour systems, *Biometrika*, **51**, 481-483.
- [16] CHALMOND, B. [2000]. *Eléments de modélisation pour l'analyse d'image*. Springer, Paris.
- [17] CHILÈS, J.-P. & DELFINER, P. [1999], *Geostatistics*, Wiley, New York.
- [18] CLIFF, A.D. & ORD, J.K. [1981], *Spatial Processes: Models and Applications*, Pion, London.
- [19] CRESSIE, N. [1993], *Statistics for Spatial Data*, Wiley, New York.
- [20] CROSS, G.R. & JAIN, A.K. [1983], Markov field texture models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **5**, 155-169.
- [21] DALEY, D.J. & VERE-JONES, D. [2003], *An Introduction to The Theory of Point Processes, Vol. I, Elementary Theory and Methods*, 2nd edn., Springer, New York.
- [22] DIGGLE, P.J. [2003], *Statistical Analysis of Spatial Point Patterns*, Oxford University Press, Oxford.
- [23] DIGGLE, P.J. & RIBEIRO P. [2007], *Model-based Geostatistics*, Springer, New York.
- [24] DOBRUSHIN, R.L. [1956], Central limit theorems for non stationary Markov chains I, II, *Theory of Probability and its Applications*, **1**, 65-80, 329-383.

- [25] DOBRUSHIN, R.L. [1968], The description of a random field by mean of conditional probabilities and condition of its regularity, *Theory of Probability and its Applications*, **13**, 197-224.
- [26] FERRANDIZ, J., LOPEZ, A., LLOPIS, A., MORALES, M. & TEJERIZO, M. L., [1995], Spatial interaction between neighbouring counties : cancer mortality data in Valencia (Spain), *Biometrics*, **51**, 665-678.
- [27] GAETAN, C. & GUYON X. [2008], *Modélisation et Statistique Spatiales*, Springer, Paris.
- [28] GAETAN, C. & GUYON X. [2009], *Modelation and Spatial Statistiques*, Springer, Paris.
- [29] GEMAN, D. & GEMAN, S. [1984], Stochastic relaxation, Gibbs distributions and the bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721-741.
- [30] GEMAN, D. [1990], *Random fields and inverse problem in Imaging*, in Hennequin, P.L. (ed.), *Lecture Notes in Mathematics*, Springer, New York 113-193.
- [31] GEORGII, H.O. [1988], *Gibbs measure and phase transitions*, De Gruyter, Berlin.
- [32] GUYON, X. [1995], *Random Fields on a Network: Modeling, Statistics and Applications*, Springer, New York.
- [33] KAISER, M.S. & CRESSIE, N. [1997], Modeling Poisson variables with positive spatial dependence, *Statistics & Probability Letters*, **35**, 423-432.
- [34] KRIGE, D.G. [1951], A statistical approach to some basic mine valuation problems on the Witwatersrand, *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 119-139.
- [35] LANTUËJOU, C. [2002], *Geostatistical Simulation*, Springer, Berlin.
- [36] LASLETT, M. [1994] Kriging and splines: and empirical comparison of their predictive performance in some applications, *Journal of the Americal Statistical Association*, **89**, 391-409.
- [37] LAWSON, A.B. [2001], *Statistical Methods in Spatial Epidemiology*, Wiley, New York.
- [38] MATÉRN B. [1960], *Spatial Variation: Stochastic Models and their Applications to Some Problems in Forest Surveys and Other Sampling Investigations*, 2nd Edition (1986), Springer, Heidelberg.
- [39] MATHERON, G. [1973], The intrinsic random function and their applications, *Advances in Applied Probability*, **5**, 439-468.
- [40] MATHERON, G. [1975], *Random Sets and Integral Geometry*, Wiley, New-York.
- [41] MØLLER, J., SYVERSVEEN, A.R. & WAAGEPETERSEN, R.P. [1998], Log-gaussian Cox processes, *Scandinavian Journal of Statistics*, **25**, 451-82.
- [42] MØLLER, J. & WAAGEPETERSEN, R.P. [2004] *Statistical Inference and Simulation for Spatial Point Processes*, Chapman & Hall/CRC, Boca Raton: FL,.
- [43] NEYMAN, J. & SCOTT, E.L. [1958], Statistical approach to problems of cosmology, *Journal of the Royal Statistical Society, Series B*, **20**, 1-43.
- [44] PAPANGELOU, F. [1974], The conditional intensity of general point processes and application to line processes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **28**, 207-227.

- [45] PFEIFER P.E. & DEUTSCH, S.J. [1980], Identificación and interpretation of first order space-time ARMA models, *Technometrics*, **22**, 3,97-408.
- [46] R DEVELOPMENT CORE TEAM [2005], *R: A language and environment for statistical computing*, Vienna, Austria, R Foundation for Statistical Computing, <http://www.R-project.org>.
- [47] RIBEIRO JR., P.J. & P.J. DIGGLE [2001] geoR: a package for geostatistical analysis, *R-NEWS*, **1**, 15–18.
- [48] RIPLEY, B.D. [1976], The second-order analysis of stationary point processes, *Journal of Applied Probability*, **13**, 255-266.
- [49] RIPLEY, B.D. [1981], *Spatial Statistics*, Wiley, New York.
- [50] RIPLEY, B.D. [1988], *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge.
- [51] RIPLEY, B.D. & KELLY, F.P. [1977], Markov point processes, *Journal of the London Mathematical Society*, **15**, 188-192.
- [52] RUE, H. & HELD, L. [2005], *Gaussian Markov Random Fields, Theory and Applications*, Chapman & Hall, London.
- [53] SCHABENBERGER, O. & GOTWAY, C.A. [2004], *Statistical Methods for Spatial Data Analysis*, Chapman and Hall, Boca Raton: FL.
- [54] SCHLATHER, M. [1999], Introduction to positive definite functions and to unconditional simulation of random fields, *Technical Report ST 99-10*, Lancaster University.
- [55] SERRA, J. [1982], *Image Analysis and Mathematical Morphology*, New York, Academic Press.
- [56] STEIN, M.L. [1999], *Interpolation of Spatial Data: Some Theory for Kriging*, Springer, New York.
- [57] STOYAN, D., KENDALL, W.S. & MECKE, J. [1995], *Stochastic Geometry and its Applications*, 2nd ed., Wiley, New York.
- [58] STRAUSS, D.J. [1975], A model for clustering, *Biometrika*, **62**, 467-475.
- [59] STRAUSS, D.J. [1977], Clustering on colored lattice, *Journal of Applied Probability*, **14**, 135-143.
- [60] THOMAS, M. [1949], A generalisation of Poisson's binomial limit for use in ecology. *Biometrika*, **36**, 18-25.
- [61] VAN LIESHOUT, M.N.M. [2000], *Markov Point Processes and their Applications*, Imperial College Press, London.
- [62] VAN LIESHOUT, M.N.M. & BADDELEY, A. J. [1999], Indices of dependence between types in multivariate point patterns, *Scandinavian Journal of Statistics*, **26**, 511–532.
- [63] WAAGEPETERSEN, R. [2007] An estimating function approach to Inference for inhomogeneous Neyman-Scott processes, *Biometrics*, **63**, 252-258.
- [64] WACKERNAGEL, H. [2003], *Multivariate Geostatistics: An Introduction with Applications*, Springer, New York.
- [65] WHITTLE, P. [1954], On stationary processes in the plane, *Biometrika*, **41**, 434-449.

- [66] WINKLER, G. [2003], *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*, Springer, New York.
- [67] YAGLOM, A.M. [1987], *Correlation Theory of Stationary and Related Random Functions. Volume I: Basic Results*, Springer, New York.