Multi-parameter auto-models with applications to cooperative systems

C. Hardouin [a,1], J. Yao [b]

[a] *CES/SAMOS-MATISSE/Université de Paris 1*
[b] *IRMAR/Université de Rennes 1*

Abstract

We propose in this paper an extension of Besag's auto-models to exponential families with multi-dimensional parameters. This extension is necessary for the treatment of spatial models like the ones with Beta conditional distributions. A family of cooperative auto-models is proposed.

Résumé

Auto-modèles à paramètres multiples et applications aux systèmes coopératifs. Nous proposons dans ce travail une extension des auto-modèles de Besag aux familles exponentielles de paramètres multiples. Cette extension est nécessaire dans plusieurs applications comme la construction des modèles coopératifs dont les lois conditionnelles sont des lois Beta.

1. Introduction

Let us consider a random field $X = \{X_i, i \in S\}$ on a finite set of sites $S = \{1, \ldots, n\}$. For a *site i*, let us denote $p_i(x_i|\cdot) = p_i(x_i|x_j, j \neq i)$, the conditional density function of X_i given the event $\{X_j = x_j, j \neq i\}$. An important approach in stochastic modelling consists in specifying the family of all these conditional distributions $\{p_i(x_i|\cdot)\}$, and then to determine a joint distribution P of the system, which is compatible with this family, i.e. the p_i's are exactly the conditional distributions associated to P.

In this paper, we focus on auto-models introduced by J. Besag ([2]). These auto-models are constructed under two assumptions: first, the dependence between sites is pairwise and secondly, the collection of conditional distributions from the sites belongs to a one-parameter exponential family. We propose in this work an extension to exponential families involving a multi-dimensional parameter. As an application

of this general approach, we address the particular problem of building *cooperative* spatial models. We consider for instance a class of Beta conditionals, which has the advantage to be able to exhibit spatial cooperation as well as spatial competition according to a suitable choice of its parameter values.

Moreover, to assess the quality of the pseudo likelihood estimator, we also present an exploratory simulation study in case of auto-models with Beta conditional distributions. Although theoretical results are still missing, the simulation results indicate that the pseudo-likelihood estimates are consistent.

2. Multi-parameter auto-models

Let us consider a set of sites $S = \{1, \ldots, n\}$, a measurable state space $(E, \mathcal{E}, m)$ (usually a subset of $\mathbb{R}^d$). We let the configuration space $\Omega = E^S$ be equipped with the σ-algebra and the product measure $(\mathcal{E}^{\otimes S}, \nu := m^{\otimes S})$. For simplicity, we shall consider $\Omega = E^S$, but all the following results hold equally with a more general configuration space $\Omega = \prod_{i \in S} E_i$, where each individual space $(E_i, \mathcal{E}_i)$ is equipped with some measure m_i.

A random field is specified by a probability distribution μ on Ω, and we will assume throughout the paper the positivity condition: namely, μ has an everywhere positive density P with respect to ν i.e. $\mu(dx) = P(x)\nu(dx)$, $P(x) = Z^{-1} \exp Q(x)$, where Z is a normalization constant. The Hammersley-Clifford's Theorem gives a characterization of $Q(x)$ as a sum of potentials G deduced from a set of cliques ([2]). The basic assumptions of the present setting are the following.

[**B1**] The dependence between the sites is pairwise-only,

$$Q(x) = \sum_{i \in S} G_i(x_i) + \sum_{\{i,j\}} G_{ij}(x_i, x_j) .$$

[**B2**] For all $i \in S$, $\log p_i(x_i|\cdot) = \langle A_i(\cdot), B_i(x_i) \rangle + C_i(x_i) + D_i(\cdot)$, $\quad A_i(\cdot) \in \mathbb{R}^d$, $\quad B_i(x_i) \in \mathbb{R}^d$.

We fix a *reference configuration* $\tau = (\tau_i) \in \Omega$. The potential functions are fully identified if we assume that, for all i, j and x it happens that $G_{ij}(\tau_i, x_j) = G_{ij}(x_i, \tau_j) = G_i(\tau_i) = 0$. The main result of the paper is the following theorem.

Theorem 2.1 *Let us assume that the two conditions [**B1**]-[**B2**] are satisfied with the normalization $B_i(\tau_i) = C_i(\tau_i) = 0$ in [**B2**], as well as the following condition*

[**C**]: *for all $i \in S$, $\mathrm{Span}\{B_i(x_i), x_i \in E\} = \mathbb{R}^d$.*

Then there exists for $i, j \in S$, $i \neq j$, a family of d-dimensional vectors $\{\alpha_i\}$ and a family of $d \times d$ matrices $\{\beta_{ij}\}$ satisfying $\beta_{ij}^T = \beta_{ji}$, such that

$$A_i(\cdot) = \alpha_i + \sum_{j \neq i} \beta_{ij} B_j(x_j) . \tag{1}$$

And the potentials are given by

$$G_i(x_i) = \langle \alpha_i, B_i(x_i) \rangle + C_i(x_i) , \tag{2}$$
$$G_{ij}(x_i, x_j) = B_i^T(x_i) \beta_{ij} B_j(x_j) . \tag{3}$$

A model satisfying the assumptions of the theorem is called a multi-parameter auto-model. Therefore, Theorem 2.1 determines the necessary form of local canonical parameters $\{A_i(.)\}$ which allow the conditional specifications to "reconstruct" together a joint distribution.

2

The following proposition is useful, giving a converse to the previous theorem. It also provides a practical way to choose the parameters for a well-defined multi-parameter auto-model. Indeed, the only additional condition one must check in practice is that the energy function Q is *admissible* in the sense that $\int_\Omega e^{Q(x)} \nu(dx) < \infty$.

Proposition 2.2 *Assume that the energy function Q is defined by [B1] with potentials G_i, G_{ij} given in (2)-(3), and that it is moreover admissible. Then the family of conditional distributions $p_i(x_i|\cdot)$ belong to an exponential family of type [B2] whose sufficient statistics $A_i(\cdot)$ satisfy (1).*

3. A special class of auto-models with Beta conditionals

Several common one-parameter auto-models necessarily imply *spatial competition* between neighbouring sites. For instance, this is the case for the auto-exponential and auto-Poisson schemes. This competition behaviour is clearly inadequate for many spatial systems ([2]). By using Beta conditional distributions, we get interesting solution to this problem. Note that the advantages of such auto-models have been already certified in a previous work [4].

Let us write the density of a Beta distribution on [0,1] with parameters $p, q > 0$ as

$$f_\theta(x) = \kappa(p,q)x^{p-1}(1-x)^{q-1} = \exp\left\{\langle\theta, B(x)\rangle - \psi(\theta)\right\} \ , 0 < x < 1$$

with $\theta = (p-1, q-1)^T$, $B(x) = [\log(2x), \log(2(1-x))]^T$ and $\psi(\theta) = (p+q-2)\log 2 + \log\kappa(p,q)$. We recall that $\kappa(p,q) = \Gamma(p+q)/[\Gamma(p)\Gamma(q)]$. Here the reference state is $\tau = \frac{1}{2}$ ensuring $B(\tau) = 0$.

We now consider a random field X with such Beta conditional distributions. Clearly, Condition [C] is satisfied. From Theorem 2.1, there exists for $i, j \in S$ and $i \neq j$ some vectors $\alpha_i = (a_i, b_i)^T \in \mathbb{R}^2$ and

(2×2)-matrices $\beta_{ij} = \begin{pmatrix} c_{ij} & d_{ij} \\ d_{ij}^* & e_{ij} \end{pmatrix}$ verifying $\beta_{ij} = \beta_{ji}^T$, such that

$$A_i(\cdot) = \alpha_i + \sum_{j\neq i} \beta_{ij}B(x_j) = \alpha_i + \sum_{j\neq i} \beta_{ij}\begin{pmatrix} \log(2x_j) \\ \log(2(1-x_j)) \end{pmatrix} \ .$$

The energy function Q can be written as

$$Q(x_1,\ldots,x_n) = \sum_{i\in S}\langle\alpha_i, B(x_i)\rangle + \sum_{\{i,j\}} B(x_i)^T\beta_{ij}B(x_j) \ .$$

Finally the reference configuration is $\tau = (\frac{1}{2},\ldots,\frac{1}{2})$ satisfying $Q(\tau) = 0$. Let the conditions
[T1]   (i)   for all $\{i,j\}$, c_{ij}, d_{ij}, d_{ij}^* and e_{ij} are all nonpositive ;
           (ii)   for all i, $1 + a_i + (\log 2)\sum_{j\neq i}\{c_{ij}+d_{ij}^*\} > 0$ and $1 + b_i + (\log 2)\sum_{j\neq i}\{d_{ij}^*+e_{ij}\} > 0$.

Proposition 3.1 *Under Conditions [T1], the family of conditional distributions $\{p_i(x_i|\cdot) \ , i \in S\}$ is everywhere well-defined, and the energy function Q is admissible.*

3.1. *Spatial cooperation versus spatial competition*

We now examine the spatial competition or cooperation behaviour of this model. At each site i, the mean of the conditional distribution $p_i(x_i|\cdot)$ is $E(X_i|\cdot) = \frac{1+A_{i,1}(\cdot)}{2+A_{i,1}(\cdot)+A_{i,2}(\cdot)}$. This conditional mean increases with $A_{i,1}(\cdot)$ and decreases with $A_{i,2}(\cdot)$. Besides the model is spatially cooperative if at each i the above

3

conditional mean increases with each neighbouring value x_j, $j \neq i$. This is possible by requiring for all i, j, $c_{ij} = e_{ij} = 0$.

Alternatively, if we adopt the constraints $d_{ij} = d_{ij}^* = 0$ for all pairs $i \neq j$, the above conditional mean becomes a decreasing function on any of its neighbouring value x_j. There is then a spatial competition between neighbouring sites.

3.2. *Estimation for a Beta cooperative model*

Let us consider the four-nearest-neighbours system on a two-dimensional lattice $S = [1, M] \times [1, N]$: each site $i \in S$ has the four neighbours denoted as $\{i_e = i + (1, 0), i_w = i - (1, 0), i_n = i + (0, 1), i_s = i - (0, 1)\}$ (with obvious correction on the boundary). We assume spatial symmetry which implies $d_{ij} = d_{ij}^*$ but allow possible anisotropy between the horizontal and vertical directions. The system is also required to be spatially cooperative and stationary. Then the model involves 4 parameters (a, b, d_1, d_2). The conditions [**T1**] become

$$d_1 \leq 0, \quad d_2 \leq 0 \; ; \quad 1 + a + 2(d_1 + d_2)\log 2 > 0 \; ; \quad 1 + b + 2(d_1 + d_2)\log 2 > 0 \; . \tag{4}$$

The associated local conditional distributions are Beta-distributed with canonical parameters

$$A_i(\cdot) = \begin{pmatrix} a + d_1[\log(2(1 - x_{i_e})) + \log(2(1 - x_{i_w}))] + d_2[\log(2(1 - x_{i_n})) + \log(2(1 - x_{i_s}))] \\ b + d_1[\log(2x_{i_e}) + \log(2x_{i_w})] + d_2[\log(2x_{i_n}) + \log(2x_{i_s})] \end{pmatrix} \; . \tag{5}$$

If we denote by ϕ the vector of all model parameters, the pseudo-likelihood is defined as $L(x; \phi) = \prod_{i \in S} p_i(x_i | x_j, \; j \neq i)$.

We refer to e.g. [3] for theoretical results on the pseudo-likelihood estimator in the general framework of a Markov random field. The study of the theoretical properties of this estimator for the multi-parameter case is still an open problem, but we think it should have a good behaviour, and we propose here to assess its performance on the basis of simulation experiments.

We consider the auto-model (5) and run 600 scans of the Gibbs sampler for each simulation on a square lattice of size 64×64. The mean and the standard deviation of the pseudo-likelihood estimates are computed from 100 independent simulations. The following table presents these results where the standard deviations are given in brackets. These results support favorably the consistency of the pseudo-likelihood estimation in the present situation.

Parameter	a	b	d_1	d_2
True values	16.6	18.9	-4.5	-4.5
Mean	16.6004	19.0062	-4.4725	-4.5093
(st. deviation)	(0.5847)	(0.5872)	(0.2742)	(0.3153)

References

[1] B. C. Arnold, E. Castillo et J. M. Sarabia, 1999. *Conditional Specification of Statistical Models*. Springer-Verlag, New York

[2] J. Besag, 1974. Spatial interactions and the statistical analysis of lattice systems. *J. Roy. Statist. Soc.* **B** 148, 1-36

[3] X. Guyon, 1995. *Random Fields on a Network: Modeling, Statistics, and Applications*. Springer-Verlag, New York

[4] M. S. Kaiser, N. Cressie and J. Lee, 2002. Spatial mixture models based on exponential family conditional distributions. *Statistica Sinica* **12**, 449-474