Représentation bayésienne avec M-splines de la mesure spectrale d'une distribution bivariée aux valeurs extrêmes

Khader Khadraoui [1] & Pierre Ribereau [2]

[1] *Laval University, Department of Mathematics and statistics, Quebec city G1V 0A6, Canada, khader.khadraoui@mat.ulaval.ca*

[2] *Université Lyon 1, Institut Camille Jordan ICJ UMR 5208 CNRS, 69622 Lyon, France, pierre.ribereau@univ-lyon1.fr*

Résumé. Nous considérons une estimation bayésienne de la mesure spectrale d'une distribution bivariée aux valeurs extrêmes. La queue d'une distribution bivariée F dans le domaine d'attraction du maximum d'une distribution bivariée aux valeurs extrêmes G est caractérisée par une mesure de probabilité d'espérance égale à 0.5 appelée mesure spectrale et par deux indices des valeurs extrêmes. Cette mesure spectrale détermine la structure de dépendance de la queue de F. L'estimation de la mesure spectrale est proposée grâce à un estimateur nonparamétrique bayésien garantissant la contrainte de moyenne. Le problème du calcul de la loi a posteriori adaptative pour tout les paramètres de la mesure spectrale est adressé par une technique MCMC à sauts réversibles. Nous présentons quelques résultats théoriques pour la consistance de la loi a posteriori et la convergence en variation totale du schéma des simulations. Une étude numérique est présentée pour valider le bon comportement de l'estimateur par rapport à deux autres procédures proposées dans la littérature.

Mots-clés. Estimation bayésienne, M-splines, Mesure spectrale, MCMC.

Abstract. We consider a Bayesian estimation with M-splines of the spectral measure of an extreme-value distribution. The tail of a bivariate distribution function F in the max-domain of attraction of an extreme-value distribution function G may be approximated by that of its extreme value attractor. The function G is characterized by a probability measure with expectation equal to $1/2$, called the spectral measure, and two extreme-value indices. This spectral measure determining the tail dependence structure of F. The estimation of the spectral measure is proposed thanks to a Bayesian nonparametric estimator that guaranteed to satisfy the moment constraint. The problem of routine calculation of posterior distributions for both coefficients and knots of M-spline is adressed using the Markov chain Monte Carlo (MCMC) simulation technique of reversible jumps. We give some theoretical results for the consistency of the posterior and for the convergence of the MCMC scheme. A simulation study shows that the Bayesian estimator resulted from the M-spline prior settin provides significant improvement over other two procedures proposed in the literature.

Keywords. Bayesian estimation, M-splines, Spectral measure, MCMC.

1 Introduction

Supposons qu'on observe un échantillon aléatoire (X_{i1}, X_{i2}), $i = 1, \ldots, n$, à partir d'une distribution bivariée inconnue F dans le domaine d'attraction du maximum d'une distribution bivariée aux valeurs extrêmes G. Il est bien connu que la queue de F est bien approximée par la queue de G sauf, bien sûr, dans le cas d'une indépendance asymptotique. Chaque distribution marginale de G est caractérisée par trois paramètres. La structure de dépendance de G est caractérisée par une mesure spectrale qui est une mesure σ-finie sur un compact. Clairement, l'inférence pour approximer la queue de F peut être réalisée à partir de l'inférence sur les six paramètres et la mesure spectrale. L'estimation des six paramètres est bien compris alors que l'estimation de la mesure spectrale reste un problème qui mérite une étude plus approfondie, bien que certains travaux apparaissent dans la littérature pour répondre à cette question.

La littérature qui porte sur la mesure spectrale se concentre habituellement sur l'inférence fréquentiste avec des approches paramétriques; voir par exemples Coles and Tawn (1991, 1994); Joe et al. (1992); Smith (1994); Ledford and Tawn (1996); Einmahl et al. (2008); Boldi and Davison (2007) et des procédures nonparamétrique; voir par exemples de Haan and de Rond (1998); de Haan and Sinha (1999); Einmahl et al. (2006, 2001); Schmidt and Stadtmüller (2006); Einmahl and Segers (2009). Récemment, en utilisant la méthode des multiplicateurs de Lagrange, des efforts importants ont été réalisés dans Einmahl and Segers (2009) afin d'intégrer la contrainte de moment caractérisant la classe des mesures spectrales. Une synthèse sur la fonction de la dépendance et les résultats pour les estimateurs de mesure spectrale peut être trouvée dans les monographies Coles (2001); de Haan and Ferreira (2006). Les approches bayésienne sur la mesure spectrale sous contrainte de moment sont plutôt rares, bien que certains travaux apparaissent dans la littérature tel que Guillotte et al. (2011).

Le but de ce travail est d'obtenir une estimation nonparamétrique de la mesure spectrale en utilisant une base M-splines. Cette estimation est proposée via une approche bayésienne qui garantie la satisfaction d'une contrainte de moment grâce à la distribution a priori. La cohérence du paradigme bayésien avec l'inférence univariée et multivariée des extrêmes a été justifiée dans la littérature (Guillotte et al., 2011; Aitchidson and Dunsmore, 1975; Coles and Tawn, 1996, 2005; Guillotte and Perron, 2008). Les contributions de ce travail sont quatre: d'abord, de proposer un estimateur nonparamétrique bayésien pour la mesure spectrale qui remplit à la fois la contrainte de moment et la forme monotone; deuxièmement, d'utiliser une base M-splines dans la construction d'un estimateur lisse et monotone qui, à notre connaissance, c'est le seul exemple d'un estimateur contraint avec M-splines dans le cadre de la mesure spectrale; troisièmement, d'étudier la consistance de la loi a posteriori en présence de contraintes de valeur et de forme; quatrièmement, de vérifier la convergence du schéma MCMC des simulations proposé. Le problème des simulations de routine suivant la distribution a posteriori pour à la fois les cœfficients et les nœuds de la spline est adressé en utilisant une méthode de Monte Carlo

par chaîne de Markov à sauts réversibles (Green, 1995). Nous étudions les propriétés asymptotiques de notre approche, y compris la convergence du schéma numérique. Notre étude est basée sur des outils standards pour une analyse asymptotique des approches bayésiennes, donnés dans Ghosal et al. (2000), à savoir la quantité d'intérêt est la probabilité a priori autour (au sens d'une métrique précise; la distance en variation totale ici) de la vraie mesure spectrale, et une sorte de mesure de l'entropie de la distribution a priori. Les détails techniques diffèrent cependant, comme nous utilisons une base M-splines et des contraintes de forme et de valeur.

L'article est organisé de la manière suivante. Dans la section 2, nous présentons la construction du sous-espace des mesures spectrales. Section 3 est consacrée à l'inférence bayésienne et la sélection de distribution a priori.

2 Construction de la mesure spectrale

Dans cette section, nous construisons la structure de dépendance d'une queue à deux variables à partir de la mesure spectrale H introduite dans De Haan and Resnick (1977). Nous considérons un vecteur (X_1, X_2) observé d'une réalisation de la distribution continue F ayant deux distributions marginales F_1 et F_2. Nous allons faire aucune hypothèse sur les marginales F_1 et F_2 sauf pour la continuité. Cependant, nous considérons F sur un quadrant de la forme $[u_1, \infty) \times [u_2, \infty)$ (où u_1 et u_2 sont deux seuils élevés). Alors, le domaine d'attraction d'une distribution bivariée G donne une bonne approximation de F. Nous notons que F est seulement bien approximée sur une partie de son support. Précisément, s'il existe deux séquences réelles $a_{n\ell} > 0$ et $b_{n\ell}$, pour $\ell \in \{1, 2\}$, et une fonction de répartition bivariée G avec des marges non dégénérés, alors, pour tout $x, y \in \mathbb{R}$

$$F^n(a_{n1}x + b_{n1}, a_{n2}y + b_{n2}) \overset{n\to\infty}{\longrightarrow} G(x, y) = \exp\Big[- l\{-\log G_1(x), -\log G_2(y)\}\Big], \quad (1)$$

où l est *la fonction stable de dépendance de la queue* (appelée aussi *la copule de la queue*) et peut être exprimée en terme de la mesure spectrale H par

$$l(x_1, x_2) = 2 \int_{[0,1]} \max(wx_1, (1 - w)x_2) dH(w)$$

pour $(x_1, x_2) \in [0, \infty)^2$. La mesure spectrale H étant une mesure de probabilité sur $[0, 1]$ avec une moyenne $1/2$.

On se donnant deux seuils suffisamment élevés u_1 et u_2, les fonctions de répartitions marginales peuvent être données par

$$-\log\{G_\ell(x_\ell|\delta_\ell)\} = \zeta_\ell \Big(1 + \eta_\ell \frac{x_\ell - u_\ell}{\sigma_\ell}\Big)^{-1/\eta_\ell}, \quad (2)$$

pour x_ℓ tel que $\eta_\ell(x_\ell - u_\ell) + \sigma_\ell > 0$, où η_ℓ est un paramètre de forme (l'indice des valeurs extrêmes), σ_ℓ est un paramètre d'échelle et $\delta_\ell = (\zeta_\ell, \eta_\ell, \sigma_\ell)$. Nous notons que $0 < \zeta_\ell =$

$-\log\{G_\ell(u_\ell|\delta_\ell)\}$ et comme u_ℓ est large, nous approchons ζ_ℓ par la probabilité marginale de dépassement du seuil ($\zeta_\ell \approx 1 - G_\ell(u_\ell|\delta_\ell)$). Ainsi, la fonction G est caractérisée par ses paramètres marginaux les vecteurs δ_1 et δ_2 et sa mesure spectrale H d'espérance égale à $1/2$.

Maintenant, nous allons construire une classe $\mathcal{H}$ de mesures spectrales lisses dont les seuls atomes, le cas échéant, sont à $\{0, 1\}$. Notre classe $\mathcal{H}$ sera pris en charge sur l'ensemble des fonctions lisses construites à partir d'une base M-spline. Fixons un certain ordre q, un nombre naturel, un certain K, un autre nombre naturel qui croit avec n, et la partition de l'intervalle $[0, 1)$ en sous-intervalles $[(k-1)/K, k/K)$ pour $k = 1, \ldots, K$. Considérons l'espace linéaire des splines d'ordre q par rapport à cette partition, qui est, toutes les fonctions $s : [0, 1) \mapsto \mathbb{R}$ qui sont polynômiales par morceaux de degré $< q$ et qui sont, dans le cas où $q \geq 2$, $q-2$ fois différentiables. Une présentation complète des splines est donnée dans de Boor (2001). Donc, on a construit un espace vectoriel de dimension $J = (q + K - 1)$. Une base pratique dans notre étude est l'ensemble des M-splines. Soit $\mathbf{t} = (t_1, \ldots, t_{K-1+2q})^T$ une séquence croissante de nœuds, de telle sorte que

$$\mathbf{t} := \Big(\frac{-(q-1)}{K}, \ldots, \frac{-1}{K}, 0, \frac{1}{K}, \frac{2}{K}, \ldots, \frac{K-1}{K}, 1, \frac{(K+1)}{K}, \ldots, \frac{K+(q-1)}{K}\Big),$$

et soit $M_{1,q}, \ldots, M_{J,q}$ les fonctions M-splines d'ordre q et de vecteur nodal $\mathbf{t}$ (nœuds intérieurs et extérieurs). Pour $\beta \in \mathbb{R}^J$, nous définissons la mesure spectrale $H_J^\beta : [0, 1] \mapsto \mathbb{R}^1$ par une combinaison linéaire des M-splines, à savoir une fonction de la forme:

$$H_J^\beta(w) = \sum_{j=1}^{J} \beta_j M_{j,q}(w). \tag{3}$$

Comme la contrainte de moment est $\int_0^1 w dH_J^\beta(w) = 1/2$, puis, en utilisant (3) et la formule des dérivées des fonctions M-spline comme indiqué dans (de Boor, 2001, Ch. X), nous pouvons écrire

$$\int_0^1 w dH_J^\beta(w) = \int_0^1 w d\Big(\sum_{j=1}^{J} \beta_j M_{j,q}(w)\Big) = \sum_{j=1}^{J} \beta_j M_{j,q}(1) - \sum_{j=1}^{J} \beta_j = 1/2, \tag{4}$$

où la fonction $M_{j,q}(1) = \boldsymbol{w}_{j,q}(1) M_{j,q-1}(1) + \{1 - \boldsymbol{w}_{j+1,q}(1)\} M_{j+1,q-1}(1)$ et où $\boldsymbol{w}_{j,q}(w) = \frac{w - t_j}{t_{j+q-1} - t_j}$ si $t_j < t_{j+q-1}$ et 0 sinon. Nous rappelons le lecteur que la fonction M-spline d'ordre 1 est $M_{j,1}(w) = K \mathbf{1}_{[t_j, t_{j+1})}(w)$ où $\mathbf{1}_A$ note la fonction indicatrice d'un ensemble A. Afin de remplir la contrainte de moment et de forme monotone, il sera commode de travailler avec le paramétrage suivant: Soit

$$\Theta = \bigcup_{J \geq 2} \Big(\{J\} \times \Theta_J\Big), \tag{5}$$

où Θ_J est défini comme suit: il existe une constante $\mathcal{C}_{\beta,K}$ qui dépend de β et K telle que

$$\Theta_J = \Big\{(\beta_1, \ldots, \beta_{J-2}, \beta_J) : \beta_{J-1} = \mathcal{C}_{\beta,K} \text{ et } \beta_1 \leq \cdots \leq \beta_J\Big\}, \tag{6}$$

avec le cœfficient β_{J-1} étant fonction de $\beta_1, \ldots, \beta_{J-2}, \beta_J$ et le nombre de sous-intervalles K à travers la restriction de moyenne (4):

$$\beta_{J-1} = \frac{1/2 + \sum_{j\neq J-1}^{J} \beta_j\{1 - M_{j,q}(1)\}}{M_{J-1,q}(1) - 1} = \mathcal{C}_{\beta,K}. \tag{7}$$

3 Inférence Bayésienne

Le modèle de l'approximation de la queue de la distribution F est spécifiée grâce à une mesure spectrale $H \in \mathcal{H}$ et les paramètres marginaux $(\delta_1, \delta_2) \in \mathcal{T}^2$ où $\delta_\ell = (\zeta_\ell, \eta_\ell, \sigma_\ell)$, $\ell \in \{1, 2\}$ et $\mathcal{T} = (-\infty, \infty) \times (0, \infty) \times (0, \infty)$. Comme l'espace des paramètres de $\mathcal{H}$ est donné par $\Theta = \cup_{J\geq 2}(\{J\} \times \Theta_J)$, alors l'espace complet des paramètres est $\Omega = \Theta \times \mathcal{T}^2$. Dans la suite, on pose $\theta = (J, \beta)$ et le modèle est paramétrisé par $(\theta, \delta_1, \delta_2)$ ce qui définit F à travers la factorisation

$$\begin{aligned} F : \Big(\Big\{ \cup_{J=2}^{\infty} \{J\} \times \Theta_J \Big\} \times (-\infty, \infty)^2 \times (0, \infty)^4\Big) &\to \mathcal{F} \\ (\theta, \delta_1, \delta_2) &\to F(x, y) \end{aligned}$$

où $\mathcal{F}$ désigne l'ensemble des distributions extrêmes bivariées et on pose

$$\left\{ \begin{array}{rcl} F(x,y) & = & \exp\Big[- l\{-\log F_1(x), -\log F_2(y)\}\Big], \text{ pour } (x,y) \in \mathbb{R}^2, \\ F_\ell(x) & = & \exp\Big\{ - \zeta_\ell\Big(1 + \eta_\ell \frac{x-u_\ell}{\sigma_\ell}\Big)^{-1/\eta_\ell}\Big\}, \text{ pour } \ell \in \{1,2\}, \\ l(x_1,x_2) & = & 2\int_{[0,1]} \max(wx_1, (1-w)x_2) dH(w), \text{ pour } (x_1,x_2) \in [0,\infty)^2. \end{array} \right.$$

On spécifie maintenant la loi a priori nonparamétrique π pour $(\theta, \delta_1, \delta_2) \in (\{\cup_{J=2}^{\infty}\{J\} \times \Theta_J\} \times (-\infty, \infty)^2 \times (0, \infty)^4)$. Sous la paramétrisation décrite précédemment, la distribution a priori π est exprimée comme un prior trans-dimensionnel sur le vecteur $(\theta, \delta_1, \delta_2)$ où par convenance elle se factorise $\pi_J(J)\pi_\beta(\beta|J)\pi_{\delta_1}(\delta_1)\pi_{\delta_2}(\delta_2)$. La distribution a priori admet une densité par rapport à la mesure de Lebesgue ou la mesure de comptage qui est spécifiée comme suit:

$$\left\{ \begin{array}{l} (\beta|J, \tau^2) \sim \mathcal{N}_J^{\Theta_J}(\mathbf{m}, \tau^2\mathbf{V}) \text{ avec densité proportionnelle à } \frac{(2\pi\tau^2)^{-J/2}}{|\mathbf{V}|^{1/2}} \exp\Big\{ - \frac{\beta'\mathbf{V}^{-1}\beta}{2\tau^2}\Big\}\mathbf{1}_{\{\beta\in\Theta_J\}}; \\ \tau^2 \sim \mathcal{IG}(\tau_1, \tau_2) \text{ avec densité égale à } \frac{\tau_2^{\tau_1}}{\Gamma(\tau_1)}(\tau^2)^{\tau_1+1} \exp\{-\frac{\tau_2}{\tau^2}\}; \\ J \sim \mathcal{P}(\lambda_1) \text{ avec densité égale à } \exp(\lambda_1)\frac{(\lambda_1)^J}{J!}, \end{array} \right.$$

où $\mathbf{m} = (0, \ldots, 0)$ est l'espérance a priori et $\mathbf{V}$ est la matrice de covariance a priori de dimension $J \times J$. Concernant les paramètres marginaux, on considère des priors indépendants pour les deux marges:

$$\pi_{\delta_\ell}(\delta_\ell) \propto \exp\{-\frac{\zeta_\ell}{2}\} \exp\{-\lambda_2 \eta_\ell\} \sigma_\ell \exp\{-\frac{\sigma_\ell}{\lambda_3}\}, \text{ pour } \ell \in \{1, 2\},$$

où ζ_ℓ, η_ℓ et σ_ℓ suivent respectivement une distribution normale, exponentielle et gamma. On peut introduire à ce niveau la vraisemblance pour terminer l'exposition de l'inférence bayésienne. Dans ce contexte, on adopte une approche par censure. On pose $(X_1^*, X_2^*) = (X_1 \vee u_1, X_2 \vee u_2)$ et $\mathbb{I} = (\mathbf{1}_{[u_1,\infty)}(X_1), \mathbf{1}_{[u_2,\infty)}(X_2))$ ce qui permet de définir:

$$f^*(X_1^*, X_2^*|\theta, \tau^2, \delta_1, \delta_2) := \begin{cases} F(u_1, u_2|\theta, \tau^2, \delta_1, \delta_2), \text{ si } \mathbb{I} = (0,0), \\ \frac{\partial}{\partial X_1} F(X_1, u_2|\theta, \tau^2, \delta_1, \delta_2), \text{ si } \mathbb{I} = (1,0), \\ \frac{\partial}{\partial X_2} F(u_1, X_2|\theta, \tau^2, \delta_1, \delta_2), \text{ si } \mathbb{I} = (0,1), \\ \frac{\partial^2}{\partial X_1 \partial X_2} F(X_1, X_2|\theta, \tau^2, \delta_1, \delta_2), \text{ si } \mathbb{I} = (1,1). \end{cases}$$

Soit $X = \{(X_{i1}, X_{i2}) : i = 1, \ldots, n\}$ pour désigner un échantillon provenant de F et soit $X^* = \{(X_{i1}^*, X_{i2}^*) : i = 1, \ldots, n\}$ pour désigner un échantillon censuré. Alors, la vraisemblance censurée est donnée par

$$L(X^*|\boldsymbol{\theta}) = \prod_{i=1}^{n} f^*(X_{i1}^*, X_{i2}^*|\boldsymbol{\theta}), \tag{8}$$

où $\boldsymbol{\theta} = (\theta, \tau^2, \delta_1, \delta_2)$. Il est naturel que la vraisemblance censurée (8) dépend des seuils u_1 et u_2. La loi a posteriori de $\boldsymbol{\theta}$ est finalement déduite par la règle de Bayes comme suit:

$$\pi(\beta, \tau^2, J, \delta_1, \delta_2|X^*) \propto L(X^*|\boldsymbol{\theta}) \pi_\beta(\beta|\tau, J) \pi_\tau(\tau^2) \pi_J(J) \pi_{\delta_1}(\delta_1) \pi_{\delta_2}(\delta_2), \tag{9}$$

où des simulations suivant cette loi a posteriori (9) peuvent être obtenus par un algorithme de type Metropolis-Hastings à sauts réversibles.

References

J. Aitchidson and I. R. Dunsmore. Non-parametric bayesian inference on bivariate extremes. *J. R. Statist. Soc., B*, 73:377–406, 1975.

M. O. J. Boldi and A. C. Davison. A mixture model for multivariate extremes. *J. R. Statist. Soc, B*, 69:217–229, 2007.

S. Coles. *An Introduction to Statistical Modelling of Extreme Values*. Springer, New York, 2001.

S. G. Coles and J. A. Tawn. Modelling extreme multivariate events. *J. R. Statist. Soc., B*, 53:377–392, 1991.

S. G. Coles and J. A. Tawn. Statistical methods for multivariate extremes: an application to structural design (with discussion). *Appl. Statist.*, 43:1–48, 1994.

S. G. Coles and J. A. Tawn. A bayesian analysis of extreme rainfall data. *Appl. Statist.*, 45:463–478, 1996.

S. G. Coles and J. A. Tawn. Bayesian modelling of extreme surges on the uk east coast. *Phil. Trans. R. Soc. Lond., A*, 363:1387–1406, 2005.

C. de Boor. *A practical guide to splines*. Springer, New York, 2001.

L. de Haan and J. de Rond. Sea and wind: multivariate extremes at work. *Extremes*, 1: 7–45, 1998.

L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer, New York, 2006.

L. De Haan and S. Resnick. Limit theory for multidimensional sample extremes. *Z. Wahrsch. Verw. Gebiete*, 40:317–337, 1977.

L. de Haan and A. K. Sinha. Estimating the probability of a rare event. *Ann. Stat.*, 27: 732–759, 1999.

J. H. J. Einmahl and J Segers. Maximum empirical likelihood estimation of the spectral measure of an extreme-value distribution. *Ann. Stat.*, 37:2953–2989, 2009.

J. H. J. Einmahl, L. de Haan, and V. Piterbarg. Nonparametric estimation of the spectral measure of an extreme value distribution. *Ann. Stat.*, 29:1401–1423, 2001.

J. H. J. Einmahl, L. de Haan, and D. Li. Weighted approximations to tail copula processes with application to testing the bivariate extreme value condition. *Ann. Stat.*, 34:1987–2014, 2006.

J. H. J. Einmahl, A. Krajina, and J. Segers. A method of moments estimator of tail dependence. *Bernoulli*, 14:1003–1026, 2008.

S. Ghosal, J. Ghosh, and A. van der Vaart. Convergence rates of posterior distributions. *Ann. Stat.*, 28:500–531, 2000.

P. J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82:711–732, 1995.

S. Guillotte and F. Perron. A bayesian estimator for the dependence function of a bivariate extreme-value distribution. *Can. J. Statist.*, 36:383–396, 2008.

S. Guillotte, F. Perron, and J. Segers. Non-parametric bayesian inference on bivariate extremes. *J. R. Statist. Soc., B*, 73:377–406, 2011.

H. Joe, R. L. Smith, and I. Weissman. Bivariate threshold methods for extremes. *J. R. Statist. Soc., B*, 54:171–183, 1992.

A. W. Ledford and J. A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187, 1996.

R. Schmidt and U. Stadtmüller. Nonparametric estimation of tail dependence. *Scan. J. Statist.*, 33:307–335, 2006.

R. L. Smith. *Multivariate threshold methods. In Extreme Value Theory and Applications.* Dordrecht, Kluwer, 1994.