CaMUS **7**, 33 – 44

Estimation de la copule gaussienne dans le cadre Bayésien

Marwa Hamdi

RÉSUMÉ Parmi les différentes familles de copules, on étudie en particulier dans cet article la copule gaussienne. Après une introduction générale, on montrera comment estimer efficacement les paramètres de la copule gaussienne à l'aide de l'approche bayésienne, en se basant sur l'algorithme de Metropolis Hastings.

1 Introduction

La dépendance est un concept fondamental en statistique. Souvent, des mesures comme le coefficient de corrélation, le tau de Kendall et le rho de Spearman sont utilisées pour évaluer la force de la dépendance. Aussi, les modèles de régression, comme la régression linéaire, permettent de relier une variable réponse et plusieurs variables explicatives. Cependant, les mesures de dépendance et les modèles de régression sont restreints.

Les copules sont des fonctions qui modélisent la structure de dépendance entre deux ou plusieurs variables aléatoires. Elles ont été développées à partir d'un problème de probabilité, énoncé par Maurice Fréchet [Fré51] dans le cadre des espaces métriques aléatoires et comme solution introduite par Abde Sklar en 1959 [Skl59] dans la théorie des lois multidimensionnelles. L'idée principale des copules a été introduite par ce dernier. Elle consiste à séparer la modélisation des lois marginales et de la structure de dépendance d'un vecteur aléatoire. En raison de sa flexibilité à modéliser des relations complexes entre les variables, les copules ont plusieurs applications dans multiples domaines comme : l'analyse de la survie, l'actuariat, la finance, le marketing, etc.

Comme l'idée des copules est récente, la plupart des recherches se sont concentrées sur le développement, les propriétés de ses fonctions et leur performance pour résoudre des problèmes. Cependant, moins d'attention a été portée sur comment estimer efficacement les paramètres des copules.

L'approche bayésienne à l'avantage de permettre l'estimation simultanée des lois marginales et de la copule des modèles multidimensionnelles. Le but de ce

En tout premier lieu, je tiens à remercier le Professeur Bernard Colin pour ses précieux conseils dans la structuration et la supervision de ce travail. Je tiens également à remercier mon directeur de maîtrise, le Professeur Taoufik Bouezmarni, pour ses différents commentaires et suggestions visant à améliorer la qualité de ce travail.

projet est de présenter brièvement les copules, en particulier la copule gaussienne, et d'appliquer l'approche bayésienne (méthode MCMC) pour estimer ses paramètres de la copule gaussienne.

2 Définition et propriétés de la Copule

2.1 Copules : définitions et propriétés

Définition 2.1. Une *copule bidimensionnelle*, notée C, est une fonction de répartition définie sur $I = [0,1]^2$ dont les lois marginales sont égales à la loi uniforme sur $[0,1]$, c'est-à-dire que pour tout $u, v \in [0,1]^2$, on a :

$$C(u,v) = P(U \leq u, V \leq v),$$

où U et V sont deux variables aléatoires uniformes sur $[0,1]$.

Autrement dit, une copule est une fonction C définie sur $[0,1]^2$ vérifiant les caractéristiques suivantes :

1. $\forall u, v \in [0,1]$, $C(u,0) = 0 = C(0,v)$;
2. $\forall u, v \in [0,1]$, $C(u,1) = u$ et $C(1,v) = v$;
3. $\forall (u_1,u_2), (v_1,v_2) \in [0,1]^2$ tels que $u_1 \leq v_1$ et $u_2 \leq v_2$ on a :
$$C(v_1,v_2) - C(v_1,u_2) - C(u_1,v_2) + C(u_1,u_2) \geq 0.$$

Dans la littérature, plusieurs familles de copules ont été introduites : les copules elliptiques (par exemple, la copule gaussienne et de Student) et les copules archimédiennes (par exemple, la copule de Clayton, Gumbel et Frank). L'exemple suivant présente la copule comonotone, notée M, qui est la borne supérieure de Fréchet-Hoeffding, c'est-à-dire que pour toute copule C on a

$$C(u,v) \leq M(u,v).$$

Exemple 2.2.
Soit la fonction $M(u,v) = \min(u,v)$, pour $u,v \in [0,1]$. On vérifie que M définit une copule. On a,

1. $\forall u, v \in [0,1]$, $\min(u,0) = 0 = \min(0,v)$;
2. $\forall u, v \in [0,1]$, $C(u,1) = u$ et $C(1,v) = v$;
3. $\forall (u_1,u_2), (v_1,v_2) \in [0,1]^2$ tels que $u_1 \leq v_1$ et $u_2 \leq v_2$ on a :
$$\min(v_1,v_2) - \min(v_1,u_2) - \min(u_1,v_2) + \min(u_1,u_2) \geq 0.$$

En effet, $\min(v_1,v_2) - \min(v_1,u_2) - \min(u_1,v_2) + \min(u_1,u_2)$ est égal à

- $v_1 - v_1 - u_1 + u_1\,(\geq 0)$ si $u_1 \leq v_1 \leq u_2 \leq v_2$;
- $v_2 - u_2 - v_2 + u_2\,(\geq 0)$ si $u_2 \leq v_2 \leq u_1 \leq v_1$;
- $v_1 - u_2 - v_1 + v_1\,(\geq 0)$ si $u_1 \leq u_2 \leq v_1 \leq v_2$;
- $v_2 - u_2 - u_1 + v_2\,(\geq 0)$ si $u_2 \leq u_1 \leq v_2 \leq v_1$.

2.2 Théorème de Sklar (1959)

L'idée du théorème est de décomposer la fonction de répartition conjointe en deux composantes : la copule et ses lois marginales. Ceci permet de construire les fonctions de répartitions multidimensionnelles en choisissant séparément les lois marginales et la structure de dépendance.

Théorème 2.3 (Sklar(1958) [Skl59])**.** *Soit $\boldsymbol{X} = (X_1, \ldots ,X_d)$ un vecteur aléatoire et F une fonction de répartition conjointe sur celui-ci, dont les lois marginales sont F_1, $\ldots$, F_d. Alors, il existe une fonction de copule C, telle que :*

$$\forall (x_1, \ldots ,x_d) \in \mathbb{R}^d,\ F(x_1, \ldots ,x_d) = C(F_1(x_1), \ldots ,F_d(x_d)).$$

Remarque 2.4. Si les lois marginales, F_1, F_2, $\ldots$, F_d sont continues, alors C est unique. Sinon elle est unique seulement sur $\prod_{i=1}^{d} Im(F_i)$, où $Im(F_i)$ est l'image de la fonction F_i.

Réciproquement, si C est une fonction de copule et $F_1, \ldots ,F_d$ sont des fonctions de répartition, alors :

$$\forall u_i \in [0,1],\ C(u_i) = F(F_1^{-1}(x_1), \ldots ,F_d^{-1}(x_d)),$$

où F_i^{-1} est la fonction réciproque de F_i, appelée aussi la fonction quantile.

L'exemple suivant montre comment on peut construire une copule à partir d'une fonction de répartition bidimensionnelle. Ici, la copule construite est la copule de Clayton de paramètre $\phi = 1$.

Exemple 2.5. On considère la distribution bidimensionnelle

$$F(x,y) = (1 + e^{-x} + e^{-y})^{-1},$$

où x, $y \in [0,\infty)$. Les fonctions marginales sont égales à

$$F_1(x) = (1 + e^{-x})^{-1} \quad \text{et} \quad F_2(y) = (1 + e^{-y})^{-1}.$$

Alors, d'après le théorème de Sklar, on a que

$$\begin{aligned} C(u,v) &= F(F_1^{-1}(u), F_2^{-1}(v)) \\ &= \left(1 + e^{\ln\left(\frac{1}{u}-1\right)} + e^{\ln\left(\frac{1}{v}-1\right)}\right) \\ &= \frac{uv}{u + v - uv}, \end{aligned}$$

qui définit bien une copule.

2.3 Densité d'une copule

On peut modéliser la structure de dépendance d'un vecteur aléatoire en utilisant la densité de la copule C qui est définie par :

$$c(u_1,\ldots,u_d) \quad = \quad \frac{\partial^d}{\partial u_1 \ldots \partial u_d} C(u_1,\ldots,u_d).$$

On peut exprimer la densité conjointe f du vecteur aléatoire $\boldsymbol{X}$ en fonction de sa densité de copule, de ses fonctions de répartitions et de ses densités marginales. En effet, d'après le théorème de Sklar et la définition de la densité de copule, f peut s'écrire de la façon suivante :

$$\begin{aligned} f(x_1,...,x_d) &= \frac{\partial^d}{\partial x_1 \ldots \partial x_d} F(x_1,\ldots,x_d) \\ &= c(F_1(x_1),\ldots,F_d(x_d)) \prod_{i=1}^{d} f_i(x_i). \end{aligned}$$

En utilisant la densité conjointe f, on peut réexprimer la densité de copule :

$$c(u_1,\ldots,u_d) = \frac{f(F_1^{-1}(u_1),\ldots,F_d^{-1}(u_d))}{\prod_{i=1}^{d} f_i(F_i^{-1}(u_i))}.$$

2.4 Mesure de dépendance

Plusieurs coefficients ont été développés pour mesurer la dépendance entre les variables aléatoires. Parmi ces mesures, on trouve le coefficient de corrélation de Pearson entre deux variables aléatoires. La proposition suivante établit que le coefficient de corrélation de Pearson peut être déduit à partir de la fonction de copule et de ses distributions marginales. Notons que les deux autres coefficients cités dans l'introduction, le tau de Kendall et le rho de Spearman, peuvent être calculés en utilisant seulement la fonction de copule.

Proposition 2.6.
Soit (X_1,X_2) un vecteur aléatoire de copule C et de fonctions de répartition marginales F_1 et F_2. Le coefficient de corrélation entre deux variables aléatoires X_1 et X_2 est donné par :

$$\rho(X_1,X_2) \quad = \quad \frac{1}{\sqrt{var(X_1)var(X_2)}} \int_0^1 \int_0^1 (C(u_1,u_2) - u_1u_2) dF_1^{-1}(u_1) dF_2^{-1}(u_2).$$

En effet,

$$\begin{aligned} \rho(X_1,X_2) &= \frac{cov(X_1,X_2)}{\sqrt{var(X_1)var(X_2)}} \\ &= \frac{1}{\sqrt{var(X_1)var(X_2)}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_1,x_2) - F_1(x_1)F_2(x_2)) \mathrm{d}x_1 \mathrm{d}x_2 \\ &= \frac{1}{\sqrt{var(X_1)var(X_2)}} \int_0^1 \int_0^1 (C(u_1,u_2) - u_1u_2) \mathrm{d}F_1^{-1}(u_1) \mathrm{d}F_2^{-1}(u_2). \end{aligned}$$

On trouve la dernière équation en appliquant le théorème de Sklar et le changement de variables $u_1 = F_1(x_1)$ et $u_2 = F_2(x_2)$.

2.5 Modèles de copules

Parmi les modèles de copules les plus importants, on trouve les copules archimédiennes et les copules elliptiques. Dans cette section, nous allons présenter brièvement les copules les plus utilisées.

2.5.1 Copules archimédiennes

La classe des copules archimédiennes permet de construire de nombreuses familles de copules qui sont capables de modéliser plusieurs structures de dépendance. Les intéressantes propriétés et la facilité de construire les copules archimédiennes les rendent très attractives en pratique. Pour plus de détails, voir [Nel06].

Définition 2.7.
Soit la fonction $\varphi : [0,1] \to \mathbb{R}^+$, dite la fonction génératrice, telle que :

$$\varphi(1) = 0,\ \varphi'(u) < 0 \text{ et } \varphi''(u) > 0.$$

La *copule archimédienne avec le générateur* φ est alors définie par :

$$C(u_1, \ldots, u_d) = \begin{cases} \varphi^{-1}\left(\sum\limits_{i=1}^{d} \varphi(u_i)\right) & \text{si } \sum\limits_{i=1}^{d} \varphi^{-1}(u_i) \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

Exemple 2.8. Dans la littérature, on trouve plusieurs copules archimédiennes. Dans cet exemple, on en cite quelques-unes.

Copule d'indépendance

La copule d'indépendance, notée $\Pi(u,v)$, fait partie des copules archimédiennes. Soit la fonction génératrice $\varphi(t) = -\log(t)$. Son inverse est donnée par la fonction $\varphi^{-1}(t) = \exp(-t)$. En se basant sur la définition, on trouve :

$$\begin{aligned} C(u,v) &= \exp\left(-\left[-\log(t) + (-\log(v))\right]\right) \\ &= uv \\ &:= \Pi(u,v). \end{aligned}$$

Copule de Clayton

Pour cette famille, le générateur est défini pour un paramètre réel $\phi > -1, \phi \neq 0$ et $u \in]0,1]$ par :

$$\varphi^{Cl}(u) = \phi^{-1}(u^{-\phi} - 1).$$

Ceci permet de générer la copule de Clayton dans le cas bidimensionnel de la façon suivante :

$$C_{\phi}^{Cl}(u_1,u_2) = (u_1^{-\phi} + u_2^{-\phi} - 1)^{-1/\phi}.$$

Dans l'Exemple 2.5, on a construit cette copule, avec $\phi = 1$, à partir d'une fonction de répartition bidimensionnelle. Comme cette fonction de copule est différentiable, elle possède une fonction de densité donnée par la forme suivante :

$$c_{\phi}^{Cl}(u_1,u_2) = (\phi+1)(u_1u_2)^{-\phi-1}(u_1^{-\phi} + u_2^{-\phi} - 1)^{\frac{-1}{\phi}-2}.$$

Copule de Frank

Pour $\phi \neq 0$ et pour $u \in]0,1]$, le générateur de la copule de Frank est donné par :

$$\varphi^F(u) = -\ln\left(\tfrac{e^{-\phi u}-1}{e^{-\phi}-1}\right).$$

La copule de Frank est donnée par :

$$C_{\phi}^{F}(u_1,u_2) = -\tfrac{1}{\phi}\ln\left(1 + \tfrac{(e^{-\phi}u_1-1)(e^{-\phi}u_2-1)}{e^{-\phi}-1}\right),$$

et sa densité de copule est égale à :

$$c_{\phi}^{F}(u_1,u_2) = \frac{\phi(1-e^{-\phi})e^{-\phi(u_1+u_2)}}{[(1-e^{-\phi}) - (e^{-\phi u_1}-1)(e^{-\phi u_2}-1)]^2}.$$

Copule de Gumbel

La dernière copule archimédienne que nous présentons est la copule de Gumbel. Sa fonction génératrice est donnée, pour $u \in]0,1]$, par :

$$\varphi^G(u) = (-\ln(u))^{\phi} \quad \text{avec} \quad \phi > 0.$$

Donc, la copule de Gumbel s'écrit sous la forme :

$$C_{\phi}^{G}(u_1,u_2) = e^{-[(-\ln(u_1))^{\phi}+(-\ln(u_2))^{\phi}]^{1/\phi}}.$$

Remarque 2.9. Notons que, C_{ϕ}^{Cl}, C_{ϕ}^{F} et C_{ϕ}^{G} convergent vers la copule comonotone introduite dans l'Exemple 2.2, si ϕ tendent vers l'infini. Aussi, C_{ϕ}^{Cl} et C_{ϕ}^{F} $\left(\text{resp. } C_{\phi}^{G}\right)$ tendent vers la copule d'indépendance si ϕ converge vers 0 (resp. 1.)

2.6 Copules elliptiques

Dans cette classe de copules se trouve la copule gaussienne et la copule de Student. Nous allons nous limiter à la présentation de la copule gaussienne qui nous intéresse dans cet article. Cependant, la méthodologie que nous présenterons dans la Section 3 peut être adaptée pour estimer les paramètres des autres copules.

Copule gaussienne

La copule gaussienne est la plus connue des familles des copules elliptiques et surtout la plus utilisée en pratique (voir [Smi11]). La copule gaussienne d'un vecteur $\boldsymbol{X}$ de dimension d s'écrit sous la forme suivante :

$$C(u_1,\ldots,u_d;\Gamma) \quad = \quad \boldsymbol{\Phi}_d\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_d);\Gamma\right), \tag{1}$$

où $\boldsymbol{\Phi}_d$ (respectivement Φ) est la distribution normale multidimensionnelle de moyenne $0_d = (0,\ldots,0)^\top$ et de matrice de corrélation Γ (respectivement la distribution normale unidimensionnelle de moyenne 0 et de variance 1). Sa densité de copule est donnée par :

$$\begin{aligned} c(\boldsymbol{u};\boldsymbol{\Phi}) &= \frac{\partial}{\partial u_1 \ldots \partial u_d} C(\boldsymbol{u};\boldsymbol{\Phi}) \\ &= |\Gamma|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{x}^\top(\Gamma^{-1} - I)\boldsymbol{x}\right\}, \end{aligned}$$

où $\boldsymbol{u} = (u_1,\ldots,u_d)$ et $\boldsymbol{x} = (\Phi^{-1}(u_1,1),\ldots,\Phi^{-1}(u_d,1))^\top$.

3 Estimation des copules

3.1 Inférence bayésienne

L'approche bayésienne est cohérente pour la résolution des problèmes d'inférence statistique. Elle permet la modélisation et l'analyse complète des incertitudes. [PCK06] et [Smi11] ont suggérés l'approche bayésienne pour estimer les paramètres de la copule gaussienne en se basant sur l'algorithme de simulation des méthodes de Monte-Carlo par chaine de Markov (MCMC). Un des objectifs de l'estimation des paramètres des copules est de construire une inférence sur les mesures de dépendances. Dans le cas de la copule gaussienne, l'inférence est construite sur le coefficient de corrélation de Pearson.

Estimation

Il existe plusieurs méthodes pour estimer les paramètres d'une copule. Une de ces méthodes, l'approche bayésienne, est basée sur la méthode MCMC.

Fonction de vraisemblance

Soit $\{\boldsymbol{y}_i = (y_{i1},\ldots,y_{id})\}_{i=1}^n$, n réalisations indépendantes et identiquement distribuées d'un vecteur aléatoire continue $\boldsymbol{Y} = (Y_1,\ldots,Y_d)$. On note par $f_j(.;\boldsymbol{\theta}_j)$, $j = 1, \ldots, d$, la densité marginale de Y_j et par $f(.;\boldsymbol{\theta},\boldsymbol{\phi})$ la fonction de densité conjointe de $\boldsymbol{Y}$, où $\boldsymbol{\phi}$ désigne le vecteur des paramètres de la copule de $\boldsymbol{Y}$ et $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_d)$. On suppose que la copule de $\boldsymbol{Y}$ est gaussienne définie par (1).

Dans ce cas, la fonction de vraisemblance, avec $\boldsymbol{\phi} = \Gamma$, est de la forme suivante (voir [Smi11]) :

$$\begin{aligned} L &= \prod_{i=1}^{n} f(\boldsymbol{y}_i;\boldsymbol{\theta},\boldsymbol{\phi}) \\ &= \prod_{i=1}^{n} \left(c(\boldsymbol{u}_i;\Gamma) \prod_{j=1}^{d} f_j(y_{ij};\boldsymbol{\theta}_j) \right) \\ &= |\Gamma|^{-n/2} \prod_{i=1}^{n} \left[\exp\left\{ -\frac{1}{2} \boldsymbol{x}_i^\top (\Gamma^{-1} - I) \boldsymbol{x}_i \right\} \prod_{j=1}^{d} f_j(y_{ij};\boldsymbol{\theta}_j) \right], \end{aligned}$$

pour $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{id})$ et $\boldsymbol{x}_i = (\Phi^{-1}(u_{i1}), \ldots, \Phi^{-1}(u_{id}))^\top$ avec $u_{ij} = F_j(y_{ij};\boldsymbol{\theta}_j)$, où F_j est la fonction de distribution de Y_j.

3.2 Méthode de simulation de Monte-Carlo

L'estimation bayésienne peut être rendue difficile si ce n'est pas impossible, pour deux raisons :

1. Le calcul explicite de $\pi(\theta|x)$ peut être impossible.
2. Lorsque l'espace des paramètres et l'espace des décisions sont de grandes dimensions. Même si $\pi(\theta|x)$ est connue, les calculs nécessitent un temps considérable.

Pour résoudre ce problème, une solution est donnée par les méthodes de simulation.

Méthode de MCMC

Le principe de la méthode MCMC est de générer une chaîne de Markov qui suit une loi asymptotiquement distribuée selon la loi a posteriori de θ, puis approximer $\mathbb{E}(g(\theta|x))$, où g est une fonction réelle. Pour construire une telle chaîne, il existe deux algorithmes : "Algorithme de Gibbs" et "Algorithme de Metropolis-Hastings (MH)". Dans notre étude, on s'intéresse à l'algorithme de MH comme suggéré par [Smi11].

Algorithme de Metropolis-Hastings (MH)

Une chaîne de Markov peut être construite par l'algorithme (MH). Cet algorithme nécessite une connaissance partielle de la fonction de densité. Étant donné la densité $\pi(\theta|x)$, on choisit une densité conditionnelle instrumentale $q(.|x)$. Pour plus de détails sur le choix de la loi de proposition symétrique q, voir [Ra11]. L'algorithme est donné comme suit :

1. Étant donné θ^p, simuler $\tilde{\theta}$ à partir de $q(\tilde{\theta}|\theta^p)$.

2. Prendre :

$$\theta^{p+1} = \begin{cases} \tilde{\theta} & \text{avec probabilité} \quad \alpha \\ \boldsymbol{\theta}^p & \text{avec probabilité} \quad 1-\alpha \end{cases}$$

où α est la probabilité d'acceptation donnée par :

$$\alpha = \min\left\{1; \frac{\pi(\tilde{\theta}|x)q(\theta^p|\tilde{\theta})}{\pi(\theta^p|x)q(\tilde{\theta}|\theta^p)}\right\}.$$

Metropolis-Hastings pour la copule gaussienne

On utilise la statistique bayésienne pour estimer les paramètres de la copule gaussienne avec la méthode de MCMC. La méthode MCMC consiste en deux étapes :

1. D'abord, on génère à partir de la densité $f(\theta_j|\{\boldsymbol{\theta}/\theta_j\},\Gamma,\boldsymbol{y})$, où $\{\boldsymbol{\theta}/\theta_j\}$ désigne tous les éléments de $\boldsymbol{\theta}$ sauf θ_j;

2. ensuite on génère à partir de $f(\Gamma|\boldsymbol{\theta},\boldsymbol{y})$.

Pour la première étape, la loi a posteriori $f(\theta_j|\{\boldsymbol{\theta}/\theta_j\},\Gamma,y)$ est donnée par :

$$\begin{aligned} f(\theta_j|\{\boldsymbol{\theta}/\theta_j\},\Gamma,\boldsymbol{y}) &= \frac{f(\theta_j,\{\boldsymbol{\theta}/\theta_j\},\Gamma,\boldsymbol{y})}{f(\{\boldsymbol{\theta}/\theta_j\},\Gamma,\boldsymbol{y})} \\ &\propto f(\boldsymbol{y}|\boldsymbol{\theta},\Gamma)f(\boldsymbol{\theta},\Gamma) \\ &\propto f(\boldsymbol{y}|\boldsymbol{\theta},\Gamma)f(\Gamma|\boldsymbol{\theta}/\theta_j)\pi(\theta_j) \\ &\propto f(\boldsymbol{y}|\boldsymbol{\theta},\Gamma)\pi(\theta_j) \\ &\propto |\Gamma|^{-n/2}\prod_{i=1}^{n}\left[\exp\left\{-\frac{1}{2}\boldsymbol{x}_i^\top(\Gamma^{-1}-I)\boldsymbol{x}_i\right\}\prod_{j=1}^{d}f_j(y_{ij};\boldsymbol{\theta}_j)\right]\pi(\theta_j), \end{aligned} \tag{2}$$

où $\pi(\theta_j)$ est la loi marginale a priori. Notons que $\boldsymbol{x}_i = (\Phi^{-1}(u_{i1}),\ldots,\Phi^{-1}(u_{id}))^\top$, où $\boldsymbol{u}_i = (u_{i1},\ldots,u_{id})$ avec $u_{ij} = F_j(y_{ij};\boldsymbol{\theta}_j)$, donc $\boldsymbol{x}_i$ dépend de θ_j. La vraisemblance dans (2) n'est pas une forme standard et donc est difficile à utiliser en pratique.
Pour construire l'algorithme de Metropolis-Hastings, [PCK06] ont suggéré la loi Student multivariée $T_\nu(\hat{\theta}_j,V)$, où $V = -H^{-1}$,

$$H = \frac{\partial^2\log(f(\theta_j|\{\boldsymbol{\theta}/\theta_j\},\Gamma,\boldsymbol{y}))}{\partial\theta_j(\partial\theta_j)^\top}\Big|_{\theta_j=\hat{\theta}_j}.$$

La probabilité d'acceptation est donnée par :

$$\alpha = \min\left\{1, \frac{f(\hat{\theta}_j|\boldsymbol{\theta}/\theta_j\Gamma,\boldsymbol{y})T_\nu(\theta_j^p|\hat{\theta}_j,V)}{f(\theta_j^p|\boldsymbol{\theta}/\theta_j,\Gamma,\boldsymbol{y})T_\nu(\hat{\theta}_j|\theta_j^p,V)}\right\}.$$

Dans la deuxième étape où la matrice de corrélation Γ est générée conditionnellement par rapport à θ et $\boldsymbol{y}$, plusieurs suggestions ont été proposées ([PCK06], [DP09], [Smi11]). Chaque méthode dépend de la paramétrisation de la matrice de corrélation et la fonction a priori. Une de ces suggestions sur la paramétrisation de la matrice de corrélation est celle de [DP09] donnée par :

$$\lambda_{t,s} = Corr(Y_t,Y_s|Y_{t-1},...,Y_{s+1}), \quad s < t.$$

Pour assurer une paramétrisation unique de la matrice de corrélation, on choisit

$$\Lambda = \{\lambda_{t,s} : t = 2,...m; s < t\}.$$

On utilise la méthode du MCMC comme suggérée par [Smi11], pour générer Γ à l'aide de la paramétrisation de [DP09].

- **Étape** 1 : Générer à partir de $f(\theta_j|\{\boldsymbol{\theta}/\theta_j\}, \Gamma, y)$, tel que $j = 1, \ldots, m$.
- **Étape** 2 : Générer à partir de $f(\tilde{\lambda}_{t,s},\gamma_{t,s}|\Theta, \{\tilde{\Lambda}/\tilde{\lambda}_{t,s}\}, \{\gamma/\gamma_{t,s}\}, y)$, tel que $t = 2, \ldots, m, s < t$.
- **Étape** 3 : Calculer Λ à partir de $(\tilde{\Lambda},\gamma)$, où $\gamma = \{\gamma_{t,s} : t = 2, \ldots, m;\ s < t\}$, $\tilde{\lambda}_{t,s} = 0 \Leftrightarrow \gamma_{t,s} = 0$, telle que $\tilde{\lambda}_{t,s}$ et $\tilde{\Lambda}$ sont des variables latentes de $\lambda_{t,s}$ et Λ respectivement.

Pour la deuxième étape, la loi de $(\tilde{\lambda}_{t,s},\gamma_{t,s})$ est $q(\tilde{\lambda}_{t,s},\gamma_{t,s}) = q_1(\gamma_{t,s})q_2(\tilde{\lambda}_{t,s})$ qui a comme probabilité d'acceptation p

$$p = \min\{1,\alpha\tfrac{\Pi(\lambda^{\ell+1})}{\Pi(\lambda^{\ell})}\kappa\}.$$

Si la densité q_2 est symétrique sur $(-1,1)$ avec une distribution Q_2, alors :

$$\kappa = \tfrac{Q_2(1-\tilde{\lambda}^{\ell})-Q_2(-1-\tilde{\lambda}^{\ell})}{Q_2(1-\tilde{\lambda}^{\ell+1})-Q_2(-1-\tilde{\lambda}^{\ell+1})}$$

$$\alpha = \begin{cases} 1 & \text{si } (\tilde{\lambda}^{\ell},\gamma^{\ell} = 0) \to (\tilde{\lambda}^{\ell+1},\gamma^{\ell+1} = 0) \\ \tfrac{L(\tilde{\lambda}^{\ell+1},\gamma^{\ell+1}=1)\delta_1}{L(0,\gamma^{\ell}=0)\delta_0} & \text{si } (\tilde{\lambda}^{\ell},\gamma^{\ell} = 0) \to (\tilde{\lambda}^{\ell+1},\gamma^{\ell+1} = 1) \\ \tfrac{L(0,\gamma^{\ell+1}=0)\delta_0}{L(\tilde{\lambda}^{\ell},\gamma^{\ell}=1)\delta_1} & \text{si } (\tilde{\lambda}^{\ell},\gamma^{\ell} = 1) \to (\tilde{\lambda}^{\ell+1},\gamma^{\ell+1} = 0) \\ \tfrac{L(\tilde{\lambda}^{\ell+1},\gamma^{\ell+1}=1)}{L(\lambda^{\ell},\gamma^{\ell}=0)} & \text{si } (\tilde{\lambda}^{\ell},\gamma^{\ell} = 1) \to (\tilde{\lambda}^{\ell+1},\gamma^{\ell+1} = 1) \end{cases},$$

où $\delta_0 = \Pi(\gamma_{t,s} = 0|\{\gamma/\gamma_{t,s}\})$ et $\delta_1 = \Pi(\gamma_{t,s} = 1|\{\gamma/\gamma_{t,s}\})$ sont les lois a priori de $\gamma_{t,s}$.

Une fois que Λ est généré, on peut calculer Γ (voir l'équation (2) dans l'article de [DP09] et [Joe06]).

Approcher la moyenne a posteriori est le résultat attendu. En utilisant la méthode du MCMC dans le cas de la copule gaussienne, l'itération de la méthode est :

$$\{(\Gamma^{[1]},\Theta^{[1]}),\ldots,(\Gamma^{[L]},\Theta^{[L]})\}$$

avec $(\Gamma^{[L]},\Theta^{[L]}) \sim f(\Gamma,\Theta|y)$

$$\tfrac{1}{L}\sum_{\ell=1}^{L}\theta_k^{[\ell]} \overset{L\to\infty}{\longrightarrow} \mathbb{E}(\theta_k|y)$$

et

$$\tfrac{1}{L}\sum_{\ell=1}^{L}\Gamma^{[\ell]} \overset{L\to\infty}{\longrightarrow} \mathbb{E}(\Gamma|y)$$

4 Conclusion

Dans cet article nous avons travaillé sur l'estimation bayésienne de la copule gaussienne. Le domaine des copules est récent, donc l'estimateur bayésien n'est pas assez développé dû à la difficulté causé par le grand nombre de paramètres à estimer. Cela n'empêche pas d'appliquer ces approches sur d'autres copules.

Références

[DP09] Michael J. Daniels et Mohsen Pourahmadi : Modeling covariance matrices via partial autocorrelations. *Journal of Multivariate Analysis*, 100:2352–2363, 2009.

[Fré51] Maurice Fréchet : Sur les tableaux dont les marges et des bornes sont données. *Ann.Univ. Lyon, Science*, 1951.

[Joe06] Harry Joe : Generating random correlation matrices based on partial correlations. *Journal of Multivariate Analysis*, 97:2177–2189, 2006.

[Nel06] Roger B. Nelsen : *An Introduction to Copulas.* Springer, 2006.

[PCK06] Michael Pitt, David Chan et Robert Kohn : Efficient bayesian inference for gaussian copula regression models. *Biometrika*, 93:537–554, 2006.

[Ra11] Christian P. Robert et George Casella (auth.) : *Méthodes de Monte-Carlo avec R.* Springer Paris, 2011.

[Skl59] Abe Sklar : Fonctions de répartition à n dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 1959.

[Smi11] Michael Stanley Smith : Bayesian approaches to copula modelling. *Methodology (stat.ME)*, 2011.

Marwa Hamdi
Département de mathématiques, Université de Sherbrooke
Courriel: Marwa.Hamdi@USherbrooke.ca