Détection de ruptures multiples dans un modèle de copules

Olivier Lopez [1,2]

[1] *ENSAE Paris-Tech & CREST, Laboratoire de Finance et d'Assurance, 3 avenue Pierre Larousse, 92245 Malakoff Cedex, France, olivier.lopez@ensae.fr*
[2] *Sorbonne Universités, UPMC Université Paris VI, EA 3124, LSTA, 4 place Jussieu 75005 Paris, France.*

Résumé. Dans ce travail, on s'intéresse à l'évolution temporelle d'une structure de dépendance entre plusieurs variables aléatoires. Cette structure de dépendance est modélisée par une copule appartenant à une famille paramétrique. Le paramètre de la copule subit des évolutions au cours du temps. Cette évolution est envisagée via un modèle de ruptures successives dans la valeur du paramètre. Le nombre de ruptures est inconnu, et une méthode de sélection du nombre de ruptures adapté à la description des données est proposée. A partir d'inégalités de concentration, des résultats sur la qualité d'estimation de la dynamique suivie par le paramètre de copule, ainsi que sur la procédure de sélection, sont démontrés. Ces résultats sont obtenus de manière non asymptotique, et sont valides pour une large gamme de familles de copules.

Mots-clés. Copules, détection de rupture, inégalité de concentration, minimax

Abstract. In this work, we are interested in the evolution through time of a dependence structure between several random variables. This dependence structure is modeled by a copula function, belonging to a parametric family. The copula parameter evolves through time. This evolution is considered through a multiple change-point model in the values of the copula parameter. The number of changes is assumed to be unknown, and a selection method for the adequate number of changes to describe data is proposed. Using concentration inequalities, results are shown on the quality of estimation of the procedure, and on the model selection strategy. These results are not asymptotic, and are valid for a large class of copula families.

Keywords. Copulas, multiple change-point, concentration inequalities, minimax

1 Description du modèle

Pour simplifier la description, nous nous bornons au cas de la dimension 2 (2 variables aléatoires), même si nos résultats se généralisent en plus grande dimension. Considérons deux variables aléatoires continues $X_{t,1}$ et $X_{t,2}$ de fonctions de répartitions notées $F_{t,j}(x) =$

$P(X_{t,j} \leq x)$, pour $j = 1, 2$. Le Théorème de Sklar stipule que, si $F_t(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$, on a

$$F_t(x_1, x_2) = \mathbf{C}_t(F_{t,1}(x_1), F_{t,2}(x_2)),$$

où C_t est l'unique fonction copule (i.e. une fonction de répartition sur $[0,1]^2$ de marginales uniformes) décrivant la structure de dépendance entre $X_{t,1}$ et $X_{t,2}$. Notre but, à partir de l'observation de réalisations indépendantes de $(X_{t,1}, X_{t,2})$ pour $t = 1, ..., n$, est d'identifier $\mathbf{C}_t$. Cette fonction copule est supposée appartenir à une famille de copules $\mathcal{C} = \{C_\theta : \theta \in \Theta\}$, où $\Theta \subset R^k$, et on note $\theta(t)$ le paramètre correspondant à la copule $\mathbf{C}_t$. On suppose de plus que

$$\theta(t) = \sum_{i=1}^{k+1} \theta_i^* \mathbf{1}_{\tau_{i-1}^* \leq t < \tau_i^*},$$

où $\theta_i \in \Theta$ pour tout i, et $\tau_0^* = 1 \leq \tau_1^* \leq ... \leq \tau_{k+1}^* = n$. On ajoute également des conditions d'identifiabilité usuelles sur les paramètres. L'étude de ruptures dans des modèles de copules porte essentiellement sur la détection d'une seule rupture, voir Dias et Embrechts (2009), notre approche étant quant à elle plus rétrospective.

Par souci de simplicité, nous décrivons la procédure dans le cas où $F_{t,j} = F_j$ pour tout t et pour tout j, même si nos résultats s'étendent au cadre où une dynamique est estimée sur les distributions marginales. Dans le cas le plus simple, F_j est estimée par la fonction de répartition empirique de $X_{t,j}$ notée $\hat{F}_j$. L'estimation des paramètres $(\tau_1^*, ..., \tau_k^*, \theta_1^*, ..., \theta_k^*, \theta_{k+1}^*)$ est faite par maximum de vraisemblance, c'est à dire en maximisant

$$\sum_{t=1}^{n} \sum_{i=1}^{k+1} \log\left(c_{\theta_i}(\hat{U}_{t,1}, \hat{U}_{t,2})\right) \mathbf{1}_{\tau_{i-1} \leq t < \tau_i},$$

avec $\hat{U}_{t,j} = \hat{F}_j(X_{t,j})$, et c_θ désignant la densité de copule de la copule C_θ. Pour des valeurs fixées de $(\tau_1, ..., \tau_k)$, cette procédure revient à effectuer sur chacun des plages ainsi découpés une estimation via la procédure semi-paramétrique définie par Genest, Ghoudi et Rivest (1995).

2 Résumé des résultats obtenus

Sous des conditions standards, on étudie l'écart entre l'estimateur du maximum de vraisemblance et les paramètres du modèle, en comparant la qualité d'estimation au risque minimax. Contrairement aux résultats du type de ceux obtenus par Csörgö et Horváth (1997) pour les problèmes de détection de rupture, les résultats présentés sont valides de manière non asymptotique, ce qui est particulièrement appréciable dans des problématiques de détection de rupture, où l'enjeu est de pouvoir détecter des ruptures les plus petites possibles peu de temps après leur occurrence. On utilise une technique de pénalisation pour sélectionner le nombre de ruptures, et des résultats sont obtenus sur la consistance de cette méthode de sélection.

Bibliographie

[1] Csörgö, M. et Horváth, L. (1997), *Limit Theorems in Change-Point Analysis*, Wiley, Chichester.
[2] Dias, A. et Embrechts, P. (2009) , Testing for structural changes in exchange rates dependence beyond linear correlation, *European Journal of Finance*, 15, 619–637.
[3] Genest, C., Ghoudi, K. et Rivest L.-P. (1995), A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, 82, 543–552.