

Utilisation de la notion de copule en tomographie

Doriano-Boris POUGAZA, Ali MOHAMMAD-DJAFARI, Jean-François BERCHER

Laboratoire des Signaux et Systèmes
UMR 8506 (CNRS-SUPELEC-Univ Paris Sud 11)
SUPELEC, Plateau de Moulon

3 rue Joliot Curie, 91192 Gif-sur-Yvette Cedex, France

Doriano.Pougaza@lss.supelec.fr, Ali.Djafari@lss.supelec.fr
Jean-Francois.Bercher@lss.supelec.fr

Résumé – Un problème important en statistique est la détermination d’une loi de probabilité jointe à partir de ses lois marginales. Dans le cas bidimensionnel, les lois de probabilité marginales $f_1(x)$ et $f_2(y)$ sont reliées à la loi jointe $f(x, y)$ par les intégrales suivant les lignes horizontale et verticale (les deux axes x et y). Ainsi, le problème de la détermination de $f(x, y)$ connaissant $f_1(x)$ et $f_2(y)$ est un problème inverse mal posé. En statistique la notion de *copule* est introduite pour obtenir une solution à ce problème. Un problème similaire en tomographie à rayon X est la reconstruction d’une image $f(x, y)$ représentant la répartition de la densité d’une quantité à l’intérieur de l’objet à partir de ses deux projections horizontale et verticale, $f_1(x)$ et $f_2(y)$. Il existe aussi un grand nombre de méthodes pour de tels problèmes fondées sur la transformée de Radon. Dans cet article, nous montrons les liens entre la notion de copule et celle de la tomographie à rayon X et voyons si on peut utiliser les méthodes d’un domaine à l’autre.

Abstract – An important problem in statistics is to determine a joint probability distribution from its marginals. In 2D case, the marginal probability density functions $f_1(x)$ and $f_2(y)$ are related to their joint distribution $f(x, y)$ via the horizontal and vertical line integrals. So, the problem of determining $f(x, y)$ from $f_1(x)$ and $f_2(y)$ is an ill-posed inverse problem. In statistics the notion of *copula* is exactly introduced to obtain a solution to this problem. Interestingly, this is also a problem encountered in X ray tomography image reconstruction where $f(x, y)$ is an image representing the distribution of the material density and $f_1(x)$ and $f_2(y)$ are the horizontal and vertical line integrals. In this paper, we try to link the notion of copula to X ray Computed Tomography (CT) and to see if we can use the methods used in each domain to the other one.

1 Introduction

Le mot *copule* vient du latin signifiant *lien, union*. En statistique, une copule est une fonction qui relie une loi multivariée et ses lois marginales [9]. Les copules sont utilisées dans plusieurs domaines, en mathématiques financières, sciences environnementales [6], en bioinformatique [5]. Il s’avère que les copules sont de puissants outils pour modéliser la structure de dépendance entre deux ou plusieurs quantités et offrent plusieurs possibilités dans la construction de la loi jointe, particulièrement avantageuse dans le cas des modèles non gaussiens [3, 10, 11].

Dans le cas 2D, en considérant la densité de probabilité jointe $f(x, y)$ comme une image, les densités de probabilité marginales $f_1(x)$ et $f_2(y)$ sont les intégrales suivant les lignes horizontale et verticale :

$$f_1(x) = \int f(x, y) dy \text{ et } f_2(y) = \int f(x, y) dx \quad (1)$$

On observe bien que le problème de la détermination de $f(x, y)$ à partir de $f_1(x)$ et $f_2(y)$ est un problème inverse mal posé [4]. Nous verrons plus tard que toutes les fonctions de la forme

$$f(x, y) = f_1(x) f_2(y) \Omega(x, y) \quad (2)$$

où $\Omega(x, y) = c(F_1(x), F_2(y))$, avec $F_1(x)$ et $F_2(y)$ les fonctions de répartition de $f_1(x)$ et de $f_2(y)$ et où $c(u, v)$ étant une densité de *copule*, sont des solutions à ce problème.

En 1917, Johann Radon introduit la transformée de Radon [8] qui sera utilisée plus tard en tomographie [1] à rayons X . Si

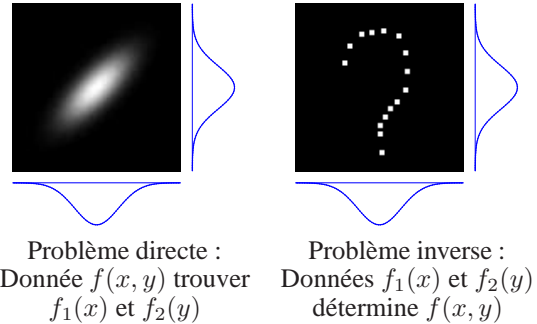


FIG. 1 – Problèmes directe et inverse

nous représentons par $f(x, y)$ la répartition spatiale de la densité d’une section de corps, un modèle simple, liant une projection $p_\theta(r)$ dans une direction θ à $f(x, y)$ est donnée par la transformée de Radon :

$$p_\theta(r) = \int_{L_{r, \theta}} f(x, y) dl \quad (3)$$
$$= \iint_{\mathcal{R}^2} f(x, y) \delta(r - x \cos \theta - y \sin \theta) dx dy.$$

Si maintenant, nous considérons seulement deux projections : horizontale $\theta = 0$ et verticale $\theta = \pi/2$, nous aboutissons à des relations entre les deux problèmes. L’objectif principal de cet article est de montrer en détails ces relations. Notre travail

est alors organisé comme suit : Dans la deuxième partie, nous présentons brièvement une définition, des propriétés et aussi une méthode de génération d'une copule. Dans la troisième partie, nous présentons les méthodes classiques de reconstruction d'images en tomographie, fondées sur l'inversion de la transformée de Radon. La quatrième partie se focalise sur la présentation de la nouvelle approche que nous proposons, c'est-à-dire le lien entre les deux précédentes parties.

Enfin, on présentera quelques résultats préliminaires comparant les différentes méthodes obtenues à l'aide de notre logiciel *Copula-Tomography* en développement.

2 Copule

Nous donnons quelques définitions et propriétés importantes des copules pour la suite. Premièrement, on notera par $F(x, y)$ une fonction de répartition bivariable, par $f(x, y)$ sa densité de probabilité, par $F_1(x)$, $F_2(y)$ ses fonctions de répartition marginales et $f_1(x)$, $f_2(y)$ leurs densités de probabilités respectives et les relations suivantes :

$$\begin{aligned} F(x, y) &= \int_{-\infty}^x \int_{-\infty}^y f(s, t) \, ds \, dt; \\ f(x, y) &= \frac{\partial^2 F(x, y)}{\partial x \partial y}; \\ F_1(x) &= \int_{-\infty}^x f_1(s) \, ds = F(x, \infty); \\ F_2(y) &= \int_{-\infty}^y f_2(t) \, dt = F(\infty, y); \\ f_1(x) &= \frac{\partial F_1(x)}{\partial x} = \int f(x, y) \, dy; \\ f_2(y) &= \frac{\partial F_2(y)}{\partial y} = \int f(x, y) \, dx. \end{aligned}$$

qui lient ces différentes fonctions.

Copule bivariable : Une copule bivariable C est une fonction de $[0, 1]^2$ dans $[0, 1]$ telle que :

- $\forall u, v \in [0, 1], C(u, 0) = 0 = C(0, v)$;
- $\forall u, v \in [0, 1], C(u, 1) = u$ et $C(1, v) = v$;
- $\forall u_1, u_2, v_1, v_2$ avec $0 \leq u_1 \leq u_2 \leq 1$ et $0 \leq v_1 \leq v_2 \leq 1$, on a :

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

Théorème de Sklar : Soit F une fonction de répartition de deux variables aléatoires X et Y ayant comme fonctions marginales univariées F_1 et F_2 . Alors il **existe** une copule C telle que : $F(x, y) = C(F_1(x), F_2(y))$. **Réciproquement**, pour des

fonctions univariées F_1 , F_2 et une copule C , la fonction F est une fonction de répartition dont les marginales sont F_1 et F_2 . De plus, si les fonctions marginales sont continues, alors la copule C est **unique**, et donnée par

$$C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v)). \quad (4)$$

Définition 2.0.1. Densité de copule :

A partir du théorème précédent et, en dérivant (4), la densité d'une copule bivariable est

$$c(u, v) = \frac{\partial^2 C}{\partial u \partial v} = \frac{f(F_1^{-1}(u), F_2^{-1}(v))}{f_1(F_1^{-1}(u)) f_2(F_2^{-1}(v))}, \quad (5)$$

et donc

$$f(x, y) = f_1(x) f_2(y) c(F_1(x), F_2(y)) \quad (6)$$

où $x = F_1^{-1}(u)$ et $y = F_2^{-1}(v)$.

Un exemple de copules usuelles est la copule d'**indépendance** : $C(u, v) = uv$ avec sa densité $c(u, v) = 1$, $(u, v) \in [0, 1]^2$.

Proposition 1. Une copule quelconque $C(u, v)$, satisfait l'inégalité

$$W(u, v) \leq C(u, v) \leq M(u, v), \quad (7)$$

où $M(u, v)$ est la **copule de Fréchet-Hoeffding minimale** (ou copule comonotone) :

$$M(u, v) = \min(u, v), \quad (u, v) \in [0, 1]^2, \quad (8)$$

et $W(u, v)$ est la **copule de Fréchet-Hoeffding maximale** (ou copule contra-monotone) :

$$W(u, v) = \max\{u + v - 1, 0\}, \quad (u, v) \in [0, 1]^2. \quad (9)$$

Il y a plusieurs familles de copules parmi lesquelles, les **copules**

archimédiennes qui forment une des classes les plus importantes (voir [7] page 109) généralisant les copules usuelles et simple d'utilisation car s'ajustant à plusieurs comportements de dépendance. Elles peuvent s'écrire sous la forme

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad (10)$$

où $\varphi(\cdot)$ est une fonction strictement décroissante et convexe telle que $\varphi(1) = 0$ appelé la génératrice. La fonction de répartition peut aussi s'écrire dans ce cas sous la forme

$$F(x, y) = \varphi^{-1}(\varphi(F_1(x)) + \varphi(F_2(y))). \quad (11)$$

Construction des copules par la méthode d'inversion : C'est une méthode directement basée sur le théorème de Sklar. Etant donnée $F(x, y)$ la fonction de répartition jointe de deux variables aléatoires X et Y avec $F_1(x)$ et $F_2(y)$ leurs fonctions marginales, toutes sont supposées continues, la copule correspondante peut être construite en utilisant l'unique transformation inverse (Transformation quantile) $x = F_1^{-1}(u)$, $y = F_2^{-1}(v)$, par

$$C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v)), \quad (12)$$

avec u, v uniformes sur $[0, 1]$.

3 Tomographie

En tomographie à rayons X pour une géométrie parallèle, on a respectivement la transformée de Radon (TR) et son inverse :

$$\begin{aligned} p(r, \theta) &= \iint_{\mathcal{R}^2} f(x, y) \delta(r - x \cos \theta - y \sin \theta) \, dx \, dy; \\ f(x, y) &= \frac{1}{2\pi} \int_0^\pi \int_0^\infty \frac{\partial p(r, \theta)}{\partial r} \, dr \, d\theta \quad (13) \end{aligned}$$

qui sont les principaux outils utilisés, pour la reconstruction d'image. Nous présentons ici brièvement les principales méthodes classiques en tomographie à rayons X.

En commençant par décomposer la transformée inverse de Radon à l'aide des opérateurs suivants :

$$\text{Dérivation } \mathcal{D} : \tilde{p}_\theta(r) = \frac{\partial p(r, \theta)}{\partial r},$$

$$\text{Hilbert Transform } \mathcal{H} : \tilde{\tilde{p}}(r', \theta) = \frac{1}{\pi} \int_0^\infty \frac{\tilde{p}(r, \theta)}{(r - r')} dr,$$

$$\text{Rétroprojection } \mathcal{B} : f(x, y) = \frac{1}{2\pi} \int_0^\pi \tilde{\tilde{p}}(x \cos \theta + y \sin \theta, \theta) d\theta.$$

Ensuite, en définissant

$$P(\Omega, \theta) = \int p(r, \theta) \exp[-j\Omega r] dr$$

et en utilisant les propriétés de la transformée de Fourier \mathcal{F} et de la dérivation \mathcal{D} :

$$\tilde{P}(\Omega, \theta) = \Omega P(\Omega, \theta)$$

et les relations entre la transformée de Hilbert \mathcal{H} et \mathcal{F} :

$$\tilde{\tilde{P}}(\Omega, \theta) = \text{sign}(\Omega) \Omega \tilde{P}(\Omega, \theta) = |\Omega| P(\Omega, \theta),$$

nous obtenons facilement les relations suivantes :

$$f(x, y) = \mathcal{B} \mathcal{H} \mathcal{D} p(r, \theta) = \mathcal{B} \mathcal{F}_1^{-1} |\Omega| \mathcal{F}_1 p(r, \theta) \quad (14)$$

et la méthode classique de *Rétroprojection Filtrée* (RPF) :



De même, si nous définissons

$$b(x, y) = \frac{1}{2\pi} \int_0^\pi p(x \cos \theta + y \sin \theta, \theta) d\theta \quad (15)$$

alors, il est montré que

$$b(x, y) = f(x, y) * h(x, y) \quad (16)$$

où $*$ est l'opération de convolution en 2D et

$$h(x, y) = 1/\sqrt{x^2 + y^2}.$$

Notant que $h(x, y)$ décroît très rapidement avec $r = \sqrt{x^2 + y^2}$, on peut utiliser $b(x, y)$ comme une bonne estimation de f . C'est la méthode de rétroprojection (simple RP). En générale, dans les problèmes de tomographie, si nous avons un grand nombre de projections uniformément distribuées sur un intervalle d'angle $[0, \pi]$, la méthode de la rétroprojection filtrée (RPF) ou même de la rétroprojection (RP) simple, fournissent de bonnes solutions. Mais, quand il y a seulement deux projections ces méthodes ne fournissent pas de solutions adéquates.

4 Lien entre la notion de copule et la tomographie

Maintenant, considérons le cas particulier où il y a seulement deux projections $\theta = 0$ et $\theta = \pi/2$. Alors

$$p_0(r) = \iint f(x, y) \delta(r - x) dx dy = \int f(r, y) dy$$

$$p_{\pi/2}(r) = \iint f(x, y) \delta(r - y) dx dy = \int f(x, r) dx$$

et si on note $f_1 = p_0$ et $f_2 = p_{\pi/2}$ nous déduisons des relations suivantes pour le problème inverse concernant la détermination de $f(x, y)$ à partir de $f_1(x)$ et $f_2(y)$:

Rétroprojection :

$$f(x, y) = \frac{1}{2}(f_1(x) + f_2(y)). \quad (17)$$

Rétroprojection filtrée :

$$f(x, y) = \frac{1}{2} \left(\int \frac{\partial f_1(x')}{\partial x'} dx' + \int \frac{\partial f_2(y')}{\partial y'} dy' \right) \quad (18)$$

qui peut être aussi implémentée dans le domaine de Fourier :

$$\begin{aligned} f(x, y) &= \int e^{+jux} \text{sign}(u) \left(\int e^{-jux} f_1(x) dx \right) du + \\ &+ \int e^{+jvy} \text{sign}(v) \left(\int e^{-jvy} f_2(y) dy \right) dv \\ &= \frac{1}{2} \int e^{+jux} |u| \left(\int e^{-jux} f_1(x) dx \right) du + \\ &+ \frac{1}{2} \int e^{+jvy} |v| \left(\int e^{-jvy} f_2(y) dy \right) dv. \end{aligned}$$

5 Utilisation des copules en tomographie

La définition et la notion de copule nous donne une possibilité pour proposer une nouvelle méthode pour la tomographie à rayon X. Considérons le cas de deux projections. Dans ce cas, on peut proposer une première utilisation de la notion de copule correspondant à l'une des copules usuelles, la copule indépendante. La méthode qui en découle est celle de la *Rétroprojection Multiplicative* (RPM) :

$$f(x, y) = f_1(x) f_2(y). \quad (19)$$

Ce nom est choisi naturellement en comparant les deux équations (17) et (19). Les figures (voir Fig.2) comparant ces méthodes montrent qu'au moins une image obtenue par RPM est mieux que la RP classique et s'ajuste très bien avec les marginales. On peut faire encore mieux si nous utilisons une autre copule que la copule indépendante par la méthode de la *Rétroprojection Copule* (RPCO) :

$$f(x, y) = f_1(x) f_2(y) c(F_1(x), F_2(y)). \quad (20)$$

Ici, le choix de copule est équivalent à l'introduction d'une information *a priori* sur la fonction à reconstruire. Nous sommes entrain de chercher le lien entre les différentes copules et la nature de l'information *a priori* correspondante.

À titre d'exemple, nous fournissons l'expression de la copule gaussienne :

$$C_\rho(u, v) = \frac{A}{2\pi} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp \left\{ \frac{-(s^2 - 2\rho st + t^2)}{2(1 - \rho^2)} \right\} ds dt$$

avec les valeurs du coefficient de corrélation $\rho = -1, 0, 1$ correspondant respectivement aux copules $W(u, v)$, $\Pi(u, v)$ et $M(u, v)$, et aussi $A = \frac{1}{\sqrt{1 - \rho^2}}$.

La densité de la copule gaussienne est :

$$c_\rho(u, v) = A \exp \left\{ \frac{-A^2}{2} ((\rho x)^2 - 2\rho xy + (\rho y)^2) \right\}.$$

Et donc la fonction $f(x, y)$ recherchée dans ce cas est :

$$f(x, y) = A f_1(x) f_2(y) \exp \left\{ -\frac{(\rho^2 x^2 - 2\rho xy + \rho^2 y^2)}{2(1 - \rho^2)} \right\}$$

avec $\Phi^{-1}(u) = x$ et $\Phi^{-1}(v) = y$. Dans ce cas, l'information a priori apportée dans la procédure de l'inversion est que la fonction recherchée est une gaussienne avec un coefficient de corrélation égale à ρ . C'est une information *a priori* très forte et pas réaliste en pratique.

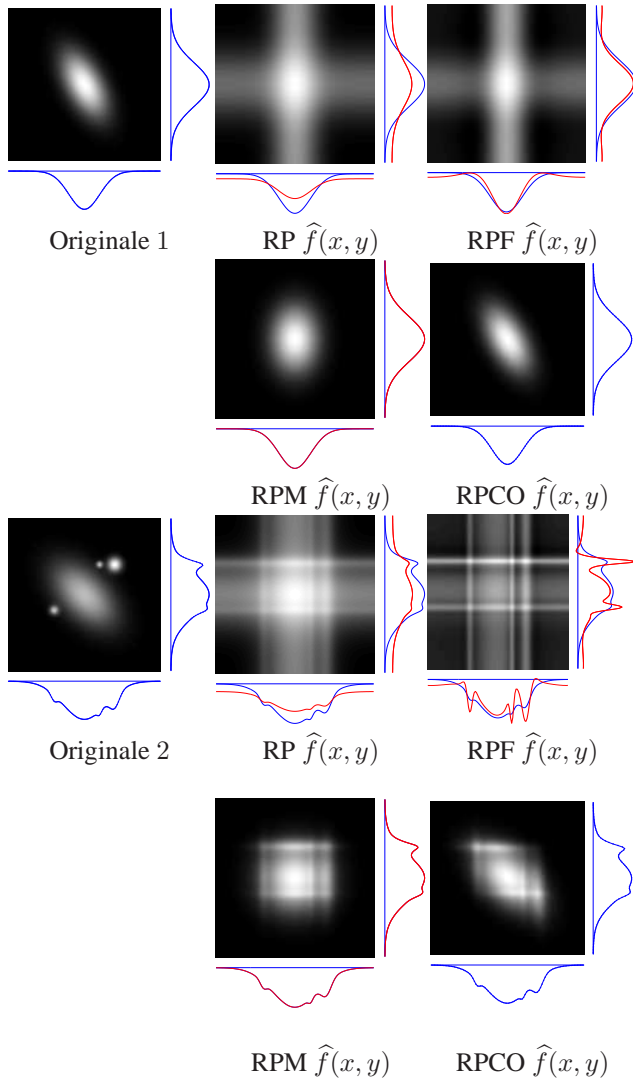


FIG. 2 – Comparaison respective entre RP, RPF, RPM et RPCO sur deux exemples.

Nous constatons que les images reconstruites avec la RPM et la RPCO sont mieux que celles obtenues avec la RP et la RPF. De plus, les deux marginales sont aussi modélisées correctement. L'image originale est bien reconstruite dans le premier

cas avec la RPCO où on utilise l'information a priori portant sur la connaissance du coefficient de corrélation et le choix de la copule gaussienne. Mais comme on peut l'observer, dans le cas du deuxième exemple avec seulement deux projections on ne peut pas reconstruire un objet de forme complexe. Nous avons besoin de plus de projections. Nous sommes en train d'étendre l'utilisation de copule pour le cas où il y a plus de deux projections. Une autre voie à explorer est l'utilisation de la notion d'entropie comme information *a priori* qui a été déjà utilisé en [12].

6 Conclusion

Notre contribution principale dans cet article est de trouver un lien entre la notion de *copule* en statistique et la tomographie par rayon X. Pour cela, nous avons présenté brièvement les copules bivariées et le problème de la reconstruction d'image en tomographie. Nous avons alors établi un lien entre les deux problèmes de i) détermination d'une densité de probabilité jointe bivariées à partir de ses deux marginales et ii) la reconstruction d'image à partir de seulement deux projections, horizontale et verticale. Nous avons montré que dans les deux cas, nous avons le même problème inverse de détermination d'une fonction de densité bivariée (une image) à partir des intégrales linéiques. Nous travaillons toujours sur cette idée en essayant de la généraliser, au cas de plusieurs projections.

Références

- [1] Cormack, A. M. (1963). Representation of a function by its line integrals with some radiological application. *J. Appl. Physics*, 34 :2722–2727.
- [3] Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12 :347–368.
- [4] Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13 :49–52.
- [5] Kim, J.M. and Jung, Y.S. and Sungur, E.A. and Han, K.H. and Park, C. and Sohn, I. (2008). A copula method for modeling directional dependence of genes. *BMC Bioinformatics*, 9 :225.
- [6] Joe, H. (1994). Multivariate extreme-value distributions with applications to environmental data. *The Canadian Journal of Statistics*, 22 :47–64.
- [7] Nelsen, R. B. (2006). *An Introduction to copulas*. Springer Science+ Business Media, Inc.
- [8] Radon, J. (1917). Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Ber. Verh. Säch. Akad. Wiss. Leipzig, Math. Nat. Kl*, 69 :262–277.
- [9] Sklar, A. (1959). Fonctions de repartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris* 8, pages 229–231.
- [10] Yan, J. (2007). Enjoy the joy of copulas : With a package copula. *Journal of Statistical Software*, 21, Issue 4.
- [11] Kallenberg, Wilber C.M. (2008). Modelling dependence. *Insurance : Mathematics and Economics*, 42 : 127–146.
- [12] Mohammad-Djafari, Ali (1994). Maximum d'entropie et problèmes inverses en imagerie. *Traitement du signal*, pages 87–116.