Department of Applied Mathematics, University of Venice

WORKING PAPER SERIES

Marta Cardin , Maddalena Manzi

Aggregation functions: an approach using copulae

Working Paper n. 159/2007
November 2007

ISSN: 1828-6887

Aggregation functions:
an approach using copulae

Marta Cardin
<mcardin@unive.it>
Dept. of Applied Mathematics
University of Venice

Maddalena Manzi
<mmanzi@math.unipd.it>
Dept. of Mathematics
University of Padua

(november 2007)

Abstract. In this paper we present the extension of the copula approach to aggregation functions. In fact we want to focus on a class of aggregation functions and present them in the multilinear form with marginal copulae. Moreover we will define also the joint aggregation density function.

Keywords: Aggregation function, n-increasing function, copula, joint aggregation density function, absolutely continuity.

JEL Classification Numbers: C02.

MathSci Classification Numbers: 90B50, 91B82, 60A10, 60E15.

Correspondence to:

Marta Cardin	Dept. of Applied Mathematics, University of Venice
	Dorsoduro 3825/e
	30123 Venezia, Italy
Phone:	[++39] (041)-234-6929
Fax:	[++39] (041)-522-1756
E-mail:	mcardin@unive.it

Introduction

In many decision situations, we are faced with multiple and conflicting attributes, that is with the problem of aggregating a collection of numerical readings to obtain a typical value. So, aggregation functions are used in multicriteria decision making problems to obtain a global score, but also in many different domains. We give some formal definitions related to the problem of aggregation and in particular we present a new unified approach to copula-based modeling and characterizations of aggregation functions. The concept of copula introduced by Sklar in 1959 is now common in the statistical literature, but only recently its potential for applications has become clear. Copulae permit to represent joint distribution functions by splitting the marginal behavior, embedded in the marginal distributions, from the dependence captured by the copula itself. So, the natural application of this function in the problem of modeling interaction between attributes is really an interesting question in the theory of aggregation functions.
The paper is structured as follows. In the next section we introduce briefly the general background of aggregation functions. After a brief overview about some properties of copulae, we present the copula approach for studying aggregation problems, considering the case $n = 3$. In fact we have already studied the bivariate case, which can be extended to n dimensions, but for the sake of simplicity in this paper we are considering the trivariate one. The last section concludes with a discussion of perspectives for future developments.

1 Definitions and properties

The aggregation operators are mathematical objects that have the function of reducing a set of numbers into a unique representative number.
For example, the arithmetic mean as an aggregation function is defined by

$$AM(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{1}$$

We introduce some properties which could be desirable for the aggregation of criteria.
If we consider the behavior of the aggregation in the best and in the worst case we expect that an aggregation satisfies the following boundary conditions :

$$A(0, \dots, 0) = 0 \qquad \text{and} \qquad A(1, \dots, 1) = 1$$

These conditions mean that if we observe only completely bad (or satisfactory) criteria the total aggregation has to bee completely bad (or satisfactory). We consider aggregation functions that satisfy the boundary conditions.
Increasingness is another property, which is often required for aggregation and commonly accepted for functions used to aggregate preferences.

So, we can define an aggregation operator as a function

$$A : \bigcup_{n \in \mathbb{N}} [0,1]^n \to [0,1]$$

that satisfies:

- $A(0, \ldots, 0) = 0$ and $A(1, \ldots, 1) = 1$
- $A(x_1, \ldots, x_n) \leq A(y_1, \ldots, y_n)$ if $(x_1, \ldots, x_n) \leq (y_1, \ldots, y_n)$

Associativity is also an interesting property for aggregation functions. The associativity property concerns the " clustering " character of an aggregation function. The properties that could be required for an aggregation function are generally based on natural considerations corresponding to the idea of an aggregated value. In deciding on the form of the aggregation operator, another elementary mathematical property can be *continuity* in the usual sense. We extend our analysis to the continuous case and in particular we suppose that A has nth-order derivatives to define $\frac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} A$ as a joint aggregation density function.
When modeling multivariate distributions, one has to take into account the effects of the marginal distributions as well as the dependence between them. This can be achieved by using the copula approach, which allows to deal with the margins and the dependence structure separately. Although almost 50 years old copulae have only recently been applied in a variety of areas. Now we briefly introduce some definitions and properties that lie within the scope of this article.

Definition 1.1 *A function f: $\mathbb{R}^n \to \mathbb{R}$ is n-increasing if $V_f(B) \geq 0$ for all n-box $B = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subseteq [0,1]^n$ with $a_i \leq b_i$, $i = 1, 2, \ldots, n$*

If f has nth-order derivatives, n-increasing is equivalent to $\frac{\partial^n}{\partial X_1 \partial X_2 \ldots \partial X_n} f \geq 0$.
This definition is the multivariate extension of the concept of "increasing" for a univariate function when we interpret "increasing" as "increasing as a distribution function". A function f: $\mathbb{R}^n \to \mathbb{R}$ is called *absolutely continuous* if it has a joint density given by $\partial^n f^n(x_1, \ldots, x_n)/\partial x_1, \partial x_2, \ldots, \partial x_n$.

An *n-copula* is the restriction to the unit cube $[0,1]^n$ of a multivariate cumulative distribution function, whose marginals are uniform on $[0,1]$.
More precisely, an n-copula is a function $C : [0,1]^n \to [0,1]$ that satisfies:

(a) $C(\mathbf{x})$ is increasing in each component x_i;

(b) $C(\mathbf{x}) = 0$ if $x_i = 0$ for any $i = 1, \ldots, n$, that is C is *grounded*;

(c) $C(\mathbf{x}) = x_i$ if all coordinates of $\mathbf{x}$ are 1 except x_i, that is C has *uniform one-dimensional marginals*;

(d) C is *n-increasing.*

Conditions (b) and (c) are known as *boundary conditions* whereas condition (d) is known as *monotonicity.*
Various properties of copulae have been studied in literature, but most part of the research concentrates on the bivariate case, since multivariate extensions are generally not easily to be done. Moreover, for any n-copula:

$$W(x_1, \dots, x_n) \le C(x_1, \dots, x_n) \le M(x_1, \dots, x_n).$$

The upper function M is an n-copula for any $n \in \mathbb{N}$, the lower function W is not an n-copula for any $n > 2$.
Now we recall also Sklar's theorem:

Theorem 1.1 (Sklar 72) *If $X_1, \dots, X_n$ are random variables defined on a common probability space, with the one-dimensional cdf's $F_{X_k}(x_k) = P(X_k \le x_k)$ and the joint cdf $F_{X_1,\dots,X_n}(x_1, \dots, x_n) = P(X_1 \le x_1, \dots, X_n \le x_n)$, then there exists an n-dimensional copula $C_{X_1,\dots,X_n}(u_1, \dots, u_n)$ such that*
$F_{X_1,\dots,X_n}(x_1, \dots, x_n) = C_{X_1,\dots,X_n}(F_{X_1}(x_1), \dots, F_{X_n}(x_n))$ for all $x_k \in \mathbf{R}$,
$k = 1, \dots, n$.

A rigorous mathematical description of copulae and their features is available in [10].

2 Dependence modeling with copulae

The subject of assessing probabilistic dependence between one-dimensional distribution functions to construct a joint distribution function is an important task in probability theory and statistics. The copula function captures the dependence relationships among the individual random variables as each multivariate distribution can be represented in terms of its marginals through a given copula structure.
The aim of this section is to present the copula approach for studying aggregation problems. It can be extended to n dimensions, but for the sake of simplicity in this paper we are considering the case $n = 3$. We are studying a class of aggregation functions that can be expressed in terms of marginal functions using the method of copulae. Elimination of marginals through copulae helps to model and understand dependence structure between variables more effectively, as the dependence has nothing to do with the marginal behavior.
The next proposition introduces an analogy between probability distributions and this class of aggregation functions, but first of all we define the following one-dimensional marginal functions:
$F_1(x_1) = A(x_1, 1, 1), \quad F_2(x_2) = A(1, x_2, 1), \quad F_3(x_3) = A(1, 1, x_3)$

Proposition 2.1 (Particular case) *If A is a 3-increasing continuous trivariate aggregation function such that*

$$A(x_1, 0, 0) = A(0, x_2, 0) = A(0, 0, x_3) = 0 \quad (2)$$

$$A(x_1, x_2, 0) = A(x_1, 0, x_3) = A(0, x_2, x_3) = 0 \quad (3)$$

there exists a copula C, such that

$$A(x_1, x_2, x_3) = C(u_1, u_2, u_3)$$

Proof.

Thanks to the previous observations, we can define the following element for three dimensions:

$$a(x_1, x_2, x_3) \triangleq \frac{\partial^3 A(x_1, x_2, x_3)}{\partial x_1, \partial x_2, \partial x_3}$$

So, from the properties of derivatives, the integral of a joint aggregation density function is the following:

$$H(x_1, x_2, x_3) \triangleq \int_0^{x_3} \int_0^{x_2} \int_0^{x_1} a(u_1, u_2, u_3) du_1 du_2 du_3 =$$

$$\int_0^{x_3} \int_0^{x_2} [\frac{\partial^2 A(x_1, u_2, u_3)}{\partial u_2, \partial u_3} - \frac{\partial^2 A(0, u_2, u_3)}{\partial u_2, \partial u_3}] du_2 du_3 =$$

$$\int_0^{x_3} [\frac{\partial A(x_1, x_2, u_3)}{\partial u_3} - \frac{\partial A(x_1, 0, u_3)}{\partial u_3}] du_3 - \int_0^{x_3} [\frac{\partial A(0, x_2, u_3)}{\partial u_3} - \frac{\partial A(0, 0, u_3)}{\partial u_3}] du_3 =$$

$$A(x_1, x_2, x_3) - A(x_1, x_2, 0) - A(x_1, 0, x_3) + A(x_1, 0, 0) +$$

$$-A(0, x_2, x_3) + A(0, x_2, 0) + A(0, 0, x_3) - A(0, 0, 0) =$$

$$A(x_1, x_2, x_3) - A(x_1, x_2, 0) - A(x_1, 0, x_3) - A(0, x_2, x_3) + A(x_1, 0, 0) + A(0, x_2, 0) + A(0, 0, x_3)$$

because $A(0, 0, 0) = 0$ for the property of aggregation function.
So, for our hypotesis $A(x_1, x_2, x_3) = H(x_1, x_2, x_3)$ and it remains to prove that $H(x_1, x_2, x_3)$ is a copula. Surely H is 3-increasing, because this is our hypotesis for A. So, we must prove that H is grounded and it has uniform one-dimensional marginals. We can observe thanks to the lemma 2.10.4 in [10] that

$$|A(x_1, x_2, x_3) - A(y_1, y_2, y_3)| \leq$$

$$\leq |F_1(x_1) - F_1(y_1)| + |F_2(x_2) - F_2(y_2)| + |F_3(x_3) - F_3(y_3)|$$

for all $\mathbf{x}, \mathbf{y} \in [0, 1]$. Then, if $F_1(x_1) = F_1(y_1)$, $F_2(x_2) = F_2(y_2)$ and $F_3(x_3) = F_3(y_3)$ it follows that $A(\mathbf{x}) = A(\mathbf{y})$. So, it is well-defined a function C whose domain is $[0, 1]^3$ with range $[0, 1]$, such that $C(u_1, u_2, u_3) = A(x_1, x_2, x_3)$, where $u_1 = F_1(x_1), u_2 = F_2(x_2)$ and $u_3 = F_3(x_3)$. Therefore we can prove that $A(\mathbf{x}) = C(F_1(x_1), F_2(x_2), F_3(x_3))$ has uniform one-dimensional marginals. In fact we have

$$C(1, 1, u_3) = C(1, 1, F_3(x_3)) = A(1, 1, x_3) = u_3.$$

Verifications of the other conditions are similar.
Now we give the more general proposition:

Proposition 2.2 (General case) *If A is a 3-increasing continuous trivariate aggregation function, there exist three increasing and continuous functions H_1, H_2, H_3, three 2-increasing and continuous functions $H_{1,2}$, $H_{2,3}$, $H_{1,3}$ and one copula C, such that*

$$A(x_1, x_2, x_3) = H_{1,2}(x_1, x_2) + H_{1,3}(x_1, x_3) + H_{2,3}(x_2, x_3) +$$

$$-H_1(x_1) - H_2(x_2) - H_3(x_3) + C(u_1, u_2, u_3)$$

Proof.

Thanks to the previous observations, it is enough to define:

$$H_1(x_1) := A(x_1, 0, 0), \quad H_2(x_2) := A(0, x_2, 0), \quad H_3(x_3) := A(0, 0, x_3) \quad (4)$$
$$H_{1,2}(x_1, x_2) := A(x_1, x_2, 0), \quad H_{1,3}(x_1, x_3) := A(x_1, 0, x_3), \quad H_{2,3}(x_2, x_3) := A(0, x_2, x_3) (5)$$

and to prove that (4) and (5) are increasing and 2-increasing, respectively.
We have already said that increasingness is a property required to aggregation preferences. It remains to prove 2-increasingness. By using lemma 2.1 in [9] $H_{1,2}(x_1, x_2) - H_1(x_1)$ is 2-increasing, that is $\frac{\partial^2 A(x_1,x_2)}{\partial x_1, \partial x_2} - \frac{\partial A(x_1)}{\partial x_1} \geq 0$, but $\frac{\partial A(x_1)}{\partial x_1} \geq 0$ and so $\frac{\partial^2 A(x_1,x_2)}{\partial x_1, \partial x_2}$ is 2-increasing. At last, we must prove that $C(u_1, u_2, u_3) = H(x_1, x_2, x_3)$ is a copula. In fact H satisfies the hypotesis of proposition (2.1) for its construction. So, there exists a copula C and three one-dimensional marginals $G_1 = u_1$, $G_2 = u_2$ and $G_3 = u_3$, such that $H(x_1, x_2, x_3) = C(G_1(x_1), G_2(x_2), G_3(x_3))$.
As a consequence, we have the following result.

Corollary 2.1 *If A is a 3-increasing continuous trivariate aggregation function, there exist three functions A_1, A_2, A_3, three constants α, β, γ and four copulae $C_{1,2}$, $C_{2,3}$, $C_{1,3}$ and C, such that*

$$A(x_1, x_2, x_3) = A_1(x_1) + A_2(x_2) + A_3(x_3) +$$

$$+\alpha C_{1,2}(x_1, x_2) + \beta C_{1,3}(x_1, x_3) + \gamma C_{2,3}(x_2, x_3) + C(x_1, x_2, x_3)$$

Proof.

We use the bivariate case, that is we have studied that

$$A(x_1, x_2, 0) = H_{1,2}(x_1, x_2) = F_\alpha(x_1) + G_\alpha(x_2) + k_\alpha C_\alpha(F(x_1), G(x_2))$$

where C_α is a copula, k_α is a constant, while F and G are two increasing and continuous functions. Similarly,

$$H_{1,3}(x_1, x_3) = F_\beta(x_1) + G_\beta(x_3) + k_\beta C_\beta(F(x_1), G(x_3))$$

and

$$H_{2,3}(x_2, x_3) = F_\gamma(x_2) + G_\gamma(x_3) + k_\gamma C_\gamma(F(x_2), G(x_3))$$

So,

$$A(x_1, x_2, x_3) = H_{1,2}(x_1, x_2) + H_{1,3}(x_1, x_3) + H_{2,3}(x_2, x_3) +$$

$$-H_1(x_1) - H_2(x_2) - H_3(x_3) + C(u_1, u_2, u_3) =$$

$$F_\alpha(x_1) + G_\alpha(x_2) + k_\alpha C_\alpha(F(x_1), G(x_2)) - H_1(x_1) +$$

$$F_\beta(x_1) + G_\beta(x_3) + k_\beta C_\beta(F(x_1), G(x_3)) - H_2(x_2) +$$

$$F_\gamma(x_2) + G_\gamma(x_3) + k_\gamma C_\gamma(F(x_2), G(x_3)) - H_3(x_3) + C(u_1, u_2, u_3)$$

By assuming $A_1(x_1) = F_\alpha(x_1) + F_\beta(x_1) - H_1(x_1)$, $A_2(x_2) = G_\alpha(x_2) + F_\gamma(x_2) - H_2(x_2)$ and $A_3(x_3) = G_\beta(x_3) + G_\gamma(x_3) - H_3(x_3)$, we have our thesis.

3 Concluding remarks

In this work we have analized the copula approach to aggregation functions for $n = 3$, continuing the study of our previous work [2]. The decomposition of a multivariate distribution function between its marginal distribution functions and its dependence structure facilitates its analysis and understanding. In this paper, a new method to construct multivariate aggregation functions is introduced. In fact our method can be used in the other sense and it is trivial to prove that the sum of several bivariate copulae and a trivariate one in a linear combination is an aggregation function. This approach can be generalized to n dimensions, studying n-increasing functions and in particular those ones which are n-copulae. However, this generalization remains an open problem, as well as the application to discrete copulae and quasi-copulae.

References

[1] T. Calvo, G. Mayor, R. Mesiar (Eds.), (2002). *Aggregation operators. New Trends and Applications*, Physica-Verlag, Heidelberg.

[2] M. Cardin, M. Manzi (2007). "Supermodular binary aggregation operators and copulae", *Social Science Research Network (SSRN Archive)*, ID1024444

[3] F. Durante R. Mesiar, P.L. Papini, C. Sempi (2007). "2-increasing binary aggregation operators", *Infor.Sciences.* **177**, 111-129.

[4] F. Durante, R. Mesiar, C. Sempi (2006). "On a family of copulas constructed from the diagonal section", *Soft Computing* **10**, 490-494.

[5] H.Joe (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.

[6] G. Mayor, J. Suñer, J. Torrens. (2005). "Copula-Like Operations on Finite Settings", *IEEE Transactions on fuzzy systems,* **13**, *n.4*, 468-476.

[7] E. P. Klement, R. Mesiar, E. Pap (2002). "Invariant copulas", *Kybernetika (Prague),* **38**, 275-285.

[8] Mesiar R., Saminger S. (2004). "Domination of ordered weighted averaging operators over t-norms", *Soft Computing* **8**, 562-570.

[9] Morillas M. P. (2005) "A characterization of absolutely monotonic (Δ) functions of a fixed order", *Publications de l'institut mathématique*, Nouvelle série, tome **78(92)**, 93-105.

[10] Roger B. Nelsen (1999). *An Introduction to Copulas*, in: Lecture Notes in Statistics, Vol. **139**, Springer, New York.

[11] J. Von Neumann (1950). *Functional operators, Volume I: Measures and integrals*, Princeton University Press, Princeton, NJ.