Modeling dependencies between decision variables and objectives with copula models

Abdelhakim Cheriet
Kasdi Merbah University
Ouargla, Algeria
abdelhakim.cheriet@univ-ouargla.dz

Roberto Santana
University of the Basque Country (UPV/EHU)
San Sebastian, Spain
roberto.santana@ehu.es

ABSTRACT

Probabilistic modeling in multi-objective optimization problems (MOPs) has mainly focused on capturing and representing the dependencies between decision variables in a set of selected solutions. Recently, some works have proposed to model also the dependencies between the objective variables, which are represented as random variables, and the decision variables. In this paper, we investigate the suitability of copula models to capture and exploit these dependencies in MOPs with a continuous representation. Copulas are very flexible probabilistic models able to represent a large variety of probability distributions.

CCS CONCEPTS

• **Mathematics of computing** → *Probabilistic algorithms*; • **Theory of computation** → *Mathematical optimization*;

KEYWORDS

multi-objective optimization, copulas, probabilistic modeling

ACM Reference Format:
Abdelhakim Cheriet and Roberto Santana. 2018. Modeling dependencies between decision variables and objectives with copula models. In *GECCO '18 Companion: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan.* ACM, New York, NY, USA, Article 4, 3 pages. https://doi.org/10.1145/3205651.3205694

1 INTRODUCTION

Usually, optimization problems exhibit strong interactions between the variables. These interactions should be taken into account at the time of searching for the optimal solutions. In multi-objective optimization problems (MOPs), interactions are determined by the different objectives involved in the problem. Two variables can have a strong interaction with respect to one objective, and be independent with respective to a second objective. This complex fabric of interactions can be exploited by optimization methods that problem structure.

Estimation of distribution algorithms (EDAs) [2] are probabilistic models that capture the interactions between the variables of the problems by identifying and representing the probabilistic dependencies in the problem. EDAs have been applied in the multi-objective domain with good results. In MOPs, the probabilistic model is generally learned from the set of selected solutions considering exclusively the decision variables of the problem. A number of recent approaches have proposed to explicitly model the dependencies between the objectives and between decision variables and objectives [1]. The rationale behind this approach is that we can effectively exploit the knowledge of how variables are affected by the objective values, as captured by the model. For example, it is possible, during the sampling step, bias the sampling of solutions by introducing as evidence ideal values of the objectives [1].

The question remains of which are the most appropriate models to represent these dependencies. In principle, we would like to count with models that are sufficiently flexible to represent the large variety of relationships that arise in MOPs. One of the best candidates, are copula-based models. Copulas [3] are joint probability distributions that satisfy the constraint that the marginal distribution of each of the variables are uniform. Copula-based models such as vine copulas are formed by coupling a graphical representation with a family of copulas, one copula for each edge in the graph. Copulas have been applied in EDAs [4], but not to represent dependencies between objectives and variables.

In this paper, we present preliminary results on using vine-copulas to represent the dependencies between objectives and decision variables in solutions that belong to the PS of bi-objective functions. Our goal is to evaluate how accurate are vine-models that use information about the relationship between decision-variables and objectives with respect to those that only represent the dependencies between the decision variables.

2 COPULAS AND VINES

Definition 2.1. A function $C(u, v) : [0, 1]^2 \rightarrow [0, 1]$ is a *copula* if and only if;

(1) For every $0 \leq u \leq 1$ $C(u, 0) = C(0, u) = 0$
(2) For every $0 \leq u \leq 1$ $C(u, 1) = u$ and $C(1, u) = u$
(3) For every $0 \leq u_1 \leq u_2 \leq 1$ and every $0 \leq v_1 \leq v_2 \leq 1$ $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$

Copulas therefore satisfy the conditions of zero-grounded bivariate distribution functions of U and V with uniform margins. Hence a probabilistic interpretation may be given in the same way as any other joint cumulative distribution function (JCF) $C(u, v) = \Pr(U \leq u, V \leq v)$. Then the unique joint probability density function (JDF) $c(u, v)$ assocciated to C is such that $C(u, v) = \int_{-\infty}^{u} \int_{-\infty}^{v} c(v, v) dv dv$.

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

ACM ISBN 978-1-4503-5764-7/18/07.
https://doi.org/10.1145/3205651.3205694

The relevance and utility of copulas are due to Sklar's theorem [3]. Thus, it is possible to separate the marginal behaviour due to the individual contributions of the random variables X, Y, described by its margins F_X and F_Y respectively, and the dependence structure, which is given by the copula (C *couples* X and Y). Moreover, a key feature of copulas is that they are invariant under strictly monotone transformations of their random variables (U and V).

3 A C-VINE MODEL FOR COPULAS AND VINES

Copulas can represent a wide variety of distributions between pairs of variables. They have been also extended to represent multivariate distributions. However, multivariate copulas are not flexible to represent distributions in which not all pairs of variables share the same type of dependence. Pair copula constructions are an effective way to build multivariate dependence models using bivariate copulas and they can be represented using a particular type of graphical model called vine.

A vine on n variables is a nested set of trees $T_1, \ldots, T_{n-1}$, where the edges of tree j are the nodes of tree $j+1$ with $j = 1, \ldots, n-2$. One of the special cases of vines is canonical vines (C-vines). In C-vines, in each tree T_j there is a unique node connected to $n-j$ edges. The C-vine density is given by

$$\prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|i,\ldots,j-1}, \tag{1}$$

We will evaluate different vines that correspond to different models of the dependencies between objectives and variables. To learn the vines we will use solutions from a true Pareto set of a bi-objective function. We compare three vine models:

(1) A C-vine that comprises only the decision variables. Variables are sampled using the C-vine structure without considering objective information.
(2) A C-vine that comprises decision variables and objectives. We learn one C-vine for each objective. The root of each C-Vine is an objective. Sampling is implemented in two ways:
   (a) All variables in a given solution are sampled from the same C-vine, alternating the two C-vines at the time of sampling the solutions.
   (b) Variables in solutions can be generated from the two C-vines, alternating between the two C-vines.

For learning the models from the (PS,PF) we use the decision variables and the objectives but in order to generate new solutions, we start by sampling from the first node of the vine (the objective). Sampling is done randomly also but we fix an objective value in the root and then sample the other variables according to the vine.

For the experiments, we use the unconstrained functions UC1, UC2, UC3, UC4, UC5, and UC6 from the CEC2009 bi-objective benchmark. We calculate an approximation of the Pareto optimal using MOEA/D after a number of generation equal to 50, then we use the solutions as an input for the three variants of the algorithm (Vine, Obj, and Mixed_Obj). For each one, we computed the IGD of the vines approximations obtained in 30 runs. We conduct this experiment for the 6 functions decision space dimensions $n \in \{5, 10\}$. Finally, we make a statistical analysis of the differences and create

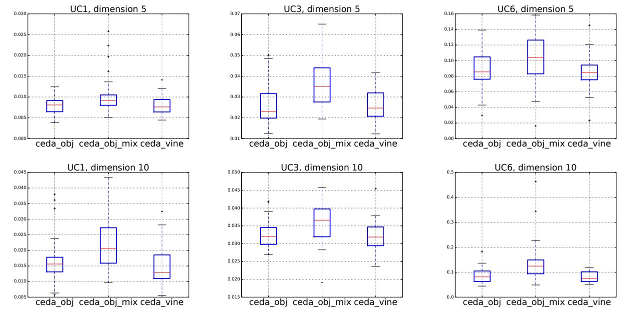

Figure 1: Three vine approximations comparison.

box-plots of the obtained results. Figure 1 shows the results for functions UC1, UC3, and UC6 with $n = 5$ and $n = 10$, respectively.

The analysis of the experiments did not show that using information about the objectives could improve the quality of the generated solutions. In general, there were no statistical differences between Vine and Obj. However, our results show that mixing in a single solution samples from the two vines was not a good idea. The quality of the solutions decreased when this strategy was used, as illustrated in Figure 1.

4 CONCLUSIONS

In this work, we have addressed for the first time the use of C-vine copulas to model the dependencies between decision variables for bi-objective problems. We also investigated how can the objective values help to generate new solutions that are close to the Pareto front. The preliminary results presented in this paper show that there are no statistical differences between the approach that considers the objective information and the regular C-vine learned from the decision variables. We could hypothesize that large datasets might be required to learn the vine models. Also, if the decision space dimension is large then objective modelling might not be useful.

ACKNOWLEDGMENTS

A. Cheriet acknowledges the computing resources provided on the high-performance computing cluster operated by the technical intensive computing platforms of Kasdi Merbah University Ouargla. R. Santana acknowledges support from the IT-609-13 (Basque Government) and TIN2016-78365-R (Spanish Ministry of Economy, Industry and Competitiveness) programs.

REFERENCES

[1] Marcella SR Martins, Myriam RBS Delgado, Ricardo Lüders, Roberto Santana, Richard Aderbal Gonçalves, and Carolina Paula de Almeida. 2017. Hybrid multi-objective Bayesian estimation of distribution algorithm: a comparative analysis for the multi-objective knapsack problem. *Journal of Heuristics* (2017). In press.
[2] H. Mühlenbein and G. Paaß. 1996. From recombination of genes to the estimation of distributions I. Binary parameters. In *Parallel Problem Solving from Nature - PPSN IV (Lectures Notes in Computer Science)*, Vol. 1141. Springer, Berlin, 178–187.
[3] Abe Sklar. 1973. Random variables, distribution functions, and copulas. *Kybernetica* (1973), 449–460.
[4] Marta Soto, Yasser Gonzalez-Fernandez, and Alberto Ochoa. 2012. Modeling with copulas and vines in estimation of distribution algorithms. *CoRR* abs/1210.5500 (2012). http://arxiv.org/abs/1210.5500