Derivative-free optimization with copula-based models for probability maximization problems

PPGM UFPR Programa de Pós-Graduação em Matemática

Emerson Butyn[a], **Elizabeth Karas**[a] **and Welington de Oliveira**[b]

a. Federal University of Paraná b. Mines ParisTech - CMA

emerson.butyn@ufpr.br

Abstract

In this work, we assess the numerical performance of several optimization solvers applied to the Cash Matching problem via a probability maximization formulation. Due to the evidenced difficulty in evaluating numerically the probability function and its gradient, we propose a derivative-free trust-region method with copula-based models for dealing with general probability maximization problems.

The probability maximization problem (PMP)

For a fixed target $\tau > 0$, consider the problem

$$\begin{array}{ll} \text{maximize} & \mathbb{P}[Ax \geq \xi] \\ \text{subject to} & c^T x \geq \tau \end{array} \iff \begin{array}{ll} \text{minimize} & \phi(x) \\ \text{subject to} & c^T x \geq \tau, \end{array} \tag{1}$$

where $x \in \mathbb{R}^n$ is the decision variable, $A \in \mathbb{R}^{m \times n}$, $\xi \in \mathbb{R}^m$ is a random vector, $\mathbb{P}$ is its associated probability measure, $c \in \mathbb{R}^n$ and

$$\phi(x) := -log(\mathbb{P}[Ax \geq \xi]). \tag{2}$$

Under the assumption that ξ follows a Gaussian probability distribution with a nonsingular covariance matrix, function ϕ is smooth and convex. In general, ϕ is difficult to be assessed numerically.

Example: the Cash Matching problem

The pension fund of a company has to make certain payments for the next m time periods, which shall be financed by buying n types of bonds. The goal is to maximize the probability of positive cash in all time periods while satisfying that the sum of the bonds yields, at the end of the period, reach a minimal target τ.

Numerical experiments

Considering the data from [3], the Cash Matching problem was solved in Matlab by the `fmincon` routine and other ones from OPTI toolbox. The Cutting Plane algorithm was used as benchmark. The gradient of the probability function was either supplied to the solvers, or approximated by finite-differences.

To compare the solvers, we considered six values for the target τ and set the CPU time limit to 2 hours.

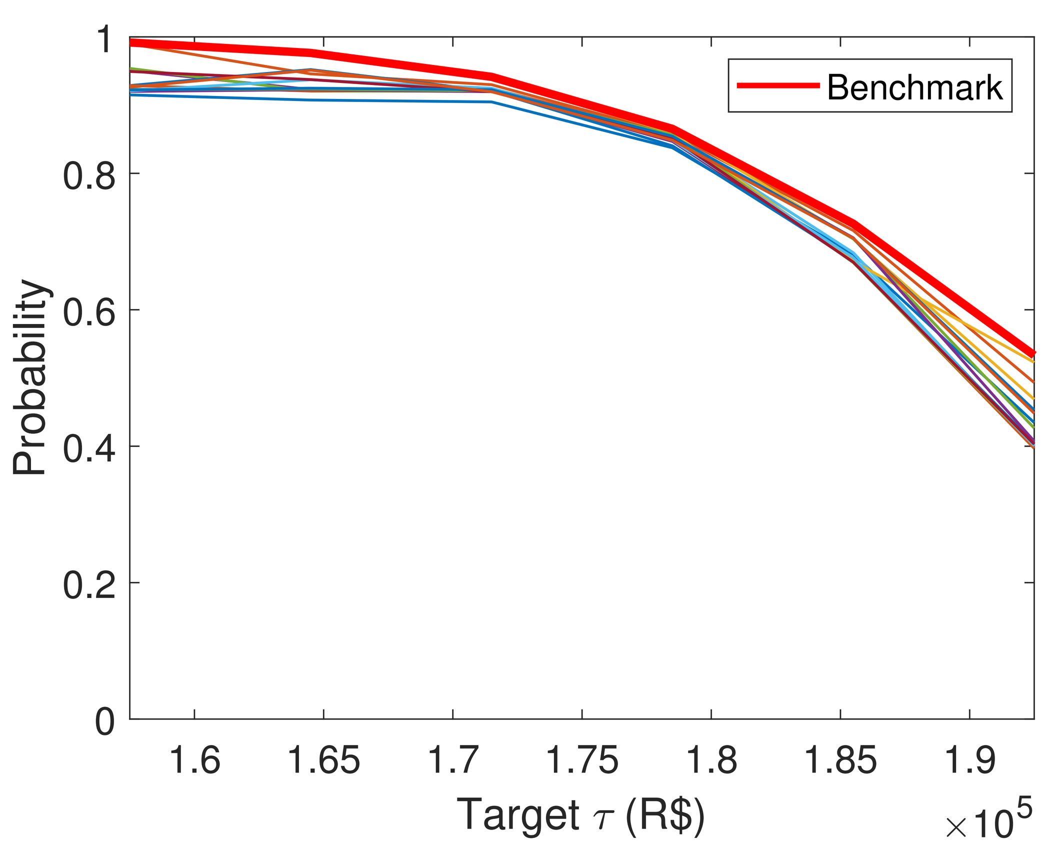

Figure: Optimal probability values from numerical experiments.

Difficulties

In our numerical experiments, almost all the CPU time was spent to compute the oracle information: function values and gradients. As the probability function is assessed via Monte-Carlo simulation and numerical integration techniques, its computed values and gradients are noisy.

Proposal

The idea is to propose a derivative-free trust-region method with copula-based models to solve it. To guarantee its global convergence we must prove that such models satisfy the hypotheses used in [2].

Copulas

Definition: For every $d \geq 2$, a *d-dimensional copula* (a d-copula) is a d-dimensional distribution function concentrated on $\mathbb{I}^d$ whose univariate marginals are uniformly distributed on $\mathbb{I}$.

Sklar's theorem: Given a random vector $\xi = (\xi_1, \ldots, \xi_d)$ in a probability space $(\Xi, \mathcal{F}, \mathbb{P})$, let

$$H(x) := \mathbb{P}(\xi_1 \leq x_1, \ldots, \xi_d \leq x_d)$$

be the joint distribution function of ξ, and for $j = 1, \ldots, d$, let $F_j(x_j) = \mathbb{P}(\xi_j \leq x_j)$ be its marginals. Then there exists a d-copula $C = C_\xi$ such that, for every point $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) =: C^F(x).$$

If the marginals are continuous, then the copula C is uniquely defined.

The model

By *Sklar's theorem*, there exists a copula C^* which coincides to the objective function of problem (1). This means that $-\log(C^*(x)) = \phi(x)$. However, estimating the correct copula for a general probability distribution is a difficulty task. We propose to work of a class C_i^F, $i = 1, \ldots, r$, of copulas and estimate, at every iteration k of our algorithm, the probability distribution via

$$H^k(x) := \sum_{i=1}^{r} \alpha_i^k C_i^F(x),$$

where the coefficients α_i^k are obtained via least-squares:

$$\alpha^k \in \arg\min \sum_{j=0}^{k} \left[\sum_{i=1}^{r} \alpha_i C_i^F(x^j) - H(x^j) \right]^2.$$

General trust-region algorithm

Given $x^0 \in \mathbb{R}^n$, $\Delta_0 > 0$ and set $k = 0$

1. Construct a model H^k
2. Find an approximate solution of the trust-region subproblem

$$\begin{array}{ll} \underset{x}{\text{minimize}} & -\log(H^k(x)) \\ \text{subject to} & c^T x \geq \tau \\ & \|x - x^k\| \leq \Delta_k. \end{array}$$

3. Update Δ_{k+1} and x^{k+1} based on the ratio between actual reduction and predicted reduction.
4. Set $k = k + 1$ and go back to step 1.

References

[1] F. Durante and C. Sempi. *Principles of copula theory*. CRC Press, 2016.

[2] P. D. Conejo et. al. Global convergence of trust-region algorithms for convex constrained minimization without derivatives. *Appl Math Comput*, 220:324–330, 2013.

[3] R. Henrion. Introduction to chance-constrained programming. *Tutorial paper for the Stochastic Programming Community*, 2004.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 1689507.