A Bayesian Approach to Constrained Multi-objective Optimization

Paul Feliot, Julien Bect, Emmanuel Vazquez

A Bayesian approach to constrained multi-objective optimization

Paul FELIOT, Julien BECT, Emmanuel VAZQUEZ

IRT SystemX, Palaiseau, France & SUPELEC, Gif-sur-Yvette, France

firstname.lastname@irt-systemx.fr or firstname.lastname@supelec.fr

Abstract. This paper addresses the problem of derivative-free multi-objective optimization of real-valued functions under multiple inequality constraints. Both the objective and constraint functions are assumed to be smooth, nonlinear, expensive-to-evaluate functions. As a consequence, the number of evaluations that can be used to carry out the optimization is very limited. The method we propose to overcome this difficulty has its roots in the Bayesian and multi-objective optimization literatures. More specifically, we make use of an extended domination rule taking both constraints and objectives into account under a unified multi-objective framework and propose a generalization of the expected improvement sampling criterion adapted to the problem. A proof of concept on a constrained multi-objective optimization test problem is given as an illustration of the effectiveness of the method.

1 Introduction

This paper addresses the problem of derivative-free multi-objective optimization of real-valued functions under multiple inequality constraints:

$$\begin{cases} \text{Minimize} \quad f(x) \\ \text{Subject to } x \in \mathbb{X} \text{ and } c(x) \leq 0 \end{cases}$$

where $f = (f_j)_{1 \leq j \leq p}$ is a vector of objective functions to be minimized, $\mathbb{X} \subset \mathbb{R}^d$ is the search domain and $c = (c_i)_{1 \leq i \leq q}$ is a vector of constraint functions. Both the objective functions f_j and the constraint functions c_i are assumed to be smooth, nonlinear functions that are expensive to evaluate. As a consequence, the number of evaluations that can be used to carry out the optimization is very limited. This setup typically arises when the values $f(x)$ and $c(x)$ for a given $x \in \mathbb{X}$ correspond to the outputs of a computationally expensive computer program.

In this work, we consider a Bayesian approach to this optimization problem. The objective and constraint functions are modelled using a vector-valued Gaussian process and $\mathbb{X}$ is explored using a sequential Bayesian design of experiments approach. More specifically, we focus on the Expected Improvement (EI) sampling criterion. This criterion was originally introduced in the context of single-objective, unconstrained optimization [10,13]. It was later extended to handle constraints [7,16,18,20,21] and to address unconstrained multi-objective problems [4,17,23,9]. However, to the best of our knowledge, the general case of a constrained multi-objective problem has only been addressed very recently by [22]. In their paper, Shimoyama et al. consider three different Bayesian criteria for unconstrained multi-objective optimization and study the effect of multiplying the criteria by a probability of feasibility in order to handle the constraints.

The approach we propose to handle the constraints is based on an extended domination rule, in the spirit of [6,15,19], which takes both objectives and constraints into account under a unified framework. The extended domination rule makes it possible to derive a new expected improvement criterion to deal with constrained multi-objective optimization problems. Section 2 introduces the proposed method, while Section 3 presents a proof of concept on a classical test case from the literature. Results and future works are briefly discussed at the end of Section 3.

2 An expected improvement criterion for constrained multi-objective optimization

In this section, we present our extended domination rule and introduce a new expected improvement criterion suitable for constrained and unconstrained multi-objective problems. The new criterion is equivalent to the original EI on unconstrained single-objective problems and to Schonlau's extension to the constrained case [21] once a feasible point has been found. It is also similar to the formulation of [23] for unconstrained multi-objective problems and to that of [22] in the constrained case once a feasible point has been found. As such, it can be seen as a generalization of the above-mentioned criteria.

Denote by $\mathbb{F} \subset \mathbb{R}^p$ and $\mathbb{C} \subset \mathbb{R}^q$ the objective and constraint spaces respectively, and let $\mathbb{Y} = \mathbb{F} \times \mathbb{C}$. We shall say that $y_1 \in \mathbb{Y}$ dominates $y_2 \in \mathbb{Y}$, which will be denoted by $y_1 \lhd y_2$, if $\psi(y_1)$ dominates $\psi(y_2)$ in the usual Pareto sense, where

$$\begin{aligned} \psi : \quad \mathbb{F} \times \mathbb{C} \;&\to\; \overline{\mathbb{R}}^p \times \mathbb{R}^q \\ (y_f, y_c) \;&\mapsto\; \begin{cases} (y_f, 0) & \text{if } y_c \leq 0, \\ (+\infty, \max(y_c, 0)) & \text{otherwise,} \end{cases} \end{aligned}$$

In the above system of equations, $\overline{\mathbb{R}}$ denotes the extended real line. For unconstrained problems, we simply take the usual domination rule on $\mathbb{F}$. Figure 1 illustrates this extended domination rule in different cases.

Assume now that $\mathbb{Y}$ is bounded. Much like [4,23,17], we define the improvement yielded by a new observation as the increase of the dominated hypervolume:

$$I_N\left(x_{N+1}\right) = \left|H_{N+1}\right| - \left|H_N\right|,$$

where H_N is the subset of $\mathbb{Y}$ dominated by the solutions observed so far $\left(f(x_1), c(x_1)\right), \ldots, \left(f(x_N), c(x_N)\right)$ and $|\cdot|$ denotes the usual (Lebesgue) volume measure in $\mathbb{R}^{p+q}$. The corresponding expected improvement criterion can be written as

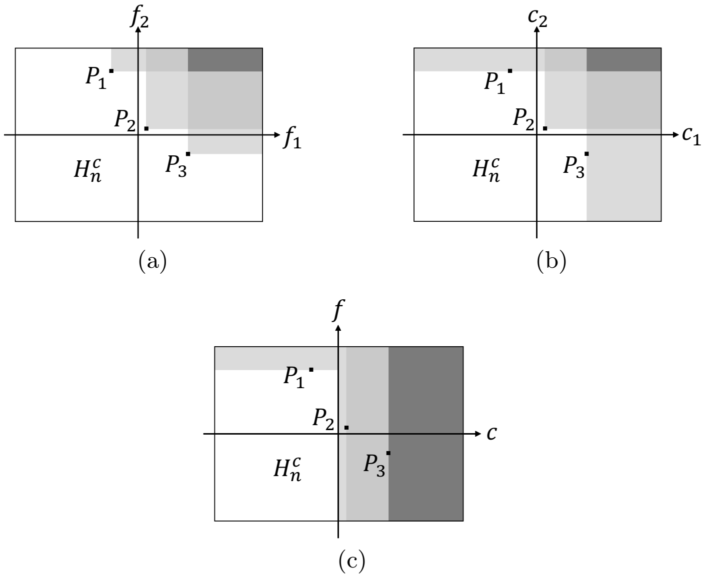

Fig. 1. Illustration of the extended domination rule in different situations. The region dominated by each point is represented by a shaded area. Darker shades of gray indicate overlapping regions. (a) Feasible solutions are compared with respect to their objective values using the usual domination rule in the objective space. (b) Non-feasible solutions are compared component-wise with respect to their constraint violations using the usual domination rule applied in the constraint space. (c) Feasible solutions always dominate non-feasible solutions; other cases are handled as in the first two figures.

$$
\begin{aligned}
\mathrm{EI}_N(x_{N+1}) &= \mathbb{E}_N\left((I_N(x_{N+1})\right) \\
&= \mathbb{E}_N\left(\int_{\mathbb{Y}\setminus H_N} \mathbb{1}_{\xi(x_{N+1})\lhd y}\,\mathrm{d}y\right) \\
&= \int_{\mathbb{Y}\setminus H_N} \mathbb{P}_N(\xi(x_{N+1})\lhd y)\,\mathrm{d}y
\end{aligned}
$$

where $\mathbb{P}_N$ denotes the probability conditional to the observations and ξ is a vector-valued Gaussian model for (f, c).

Even though the integrand of the EI formula can be readily computed analytically, its integration is not trivial due to the combinatorial nature of the problem [8,2,5]. To overcome this difficulty, we propose to use a Sequential Monte Carlo (SMC) approximation [3,11,12,1]:

$$
\mathrm{EI}_N(x_{N+1}) \approx \sum_{i=1}^{n} w_i\,\mathbb{P}_N(\xi(x_{N+1})\lhd y_i),
$$

where $\mathcal{Y}_N = (w_i, y_i)_{1 \le i \le n}$ is a weighted sample that targets the uniform density on $\mathbb{Y} \setminus H_N$.

3 Proof of concept

In this paper, we illustrate the behavior of our new optimization strategy using the Osyczka and Kundu test problem [14] for constrained multi-objective optimization ($d = 6$, $p = 2$, $q = 6$). The algorithm is initialized using a Latin Hypercube sample of 18 samples and proceeds using the above mentionned criterion. Figure 2 shows the convergence of the algorithm at different steps of the optimization.

We are also able to report good results on other challenging test cases from the literature and future communications will include a comparison of our method to reference optimization methods. More details about the SMC procedure will also be proposed.

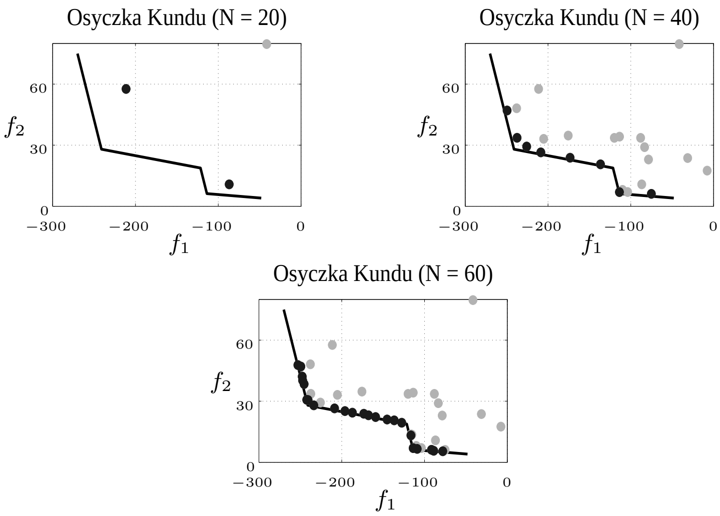

Fig. 2. Test results on Osyczka and Kundu test problem with, from left to right, $N = 20$, 40 and 60 evaluations. Only feasible points are shown on the figures. The dark dots represent non-dominated observations while the light gray dots represent dominated ones. The dark curve represents the target Pareto front.

Acknowledgements. This research work has been carried out in the frame of the Technological Research Institute SystemX, and therefore granted with public funds within the scope of the French Program *Investissements d'Avenir*.

References

1. Au, S.K., Beck, J.L.: Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics 16(4), 263–277 (2001)
2. Bader, J., Zitzler, E.: Hype: An algorithm for fast hypervolume-based many-objective optimization. Evolutionary Computation 19(1), 45–76 (2011)
3. Benassi, R., Bect, J., Vazquez, E.: Bayesian optimization using sequential Monte Carlo. In: Learning and Intelligent Optimization. 6th International Conference, LION 6, Paris, France, January 16-20, 2012, Revised Selected Papers, Lecture Notes in Computer Science, vol. 7219, pp. 339–342. Springer (2012)
4. Emmerich, M.T.M., Giannakoglou, K.C., Naujoks, B.: Single- and multi-objective evolutionary optimization assisted by Gaussian random field metamodels. IEEE Transactions on Evolutionary Computation 10(4), 421–439 (2006)
5. Emmerich, M., Klinkenberg, J.W.: The computation of the expected improvement in dominated hypervolume of Pareto front approximations. Rapport technique, Leiden University (2008)
6. Fonseca, C.M., Fleming, P.J.: Multiobjective optimization and multiple constraint handling with evolutionary algorithms. I. A unified formulation. IEEE Transactions on Systems, Man and Cybernetics. Part A: Systems and Humans 28(1), 26–37 (1998)
7. Gramacy, R.L., Lee, H.: Optimization under unknown constraints. In: Bayesian Statistics 9. Proceedings of the Ninth Valencia International Meeting. pp. 229–256. Oxford University Press (2011)
8. Hupkens, I., Emmerich, M., Deutz, A.: Faster computation of expected hypervolume improvement. arXiv preprint arXiv:1408.7114 (2014)
9. Jeong, S., Minemura, Y., Obayashi, S.: Optimization of combustion chamber for diesel engine using kriging model. Journal of Fluid Science and Technology 1(2), 138–146 (2006)
10. Jones, D.R., Schonlau, M., Welch, W.J.: Efficient global optimization of expensive black-box functions. Journal of Global Optimization 13(4), 455–492 (1998)
11. Li, L., Bect, J., Vazquez, E.: Bayesian Subset Simulation: a kriging-based subset simulation algorithm for the estimation of small probabilities of failure. In: Proceedings of PSAM 11 & ESREL 2012, 25-29 June 2012, Helsinki, Finland. IAPSAM (2012)
12. Liu, J.S.: Monte Carlo strategies in scientific computing. Springer (2008)
13. Mockus, J.: Application of bayesian approach to numerical methods of global and stochastic optimization. Journal of Global Optimization 4(4), 347–365 (1994)
14. Osyczka, A., Kundu, S.: A new method to solve generalized multicriteria optimization problems using the simple genetic algorithm. Structural Optimization 10(2), 94–99 (1995)
15. Oyama, A., Shimoyama, K., Fujii, K.: New constraint-handling method for multi-objective and multi-constraint evolutionary optimization. Transactions of the Japan Society for Aeronautical and Space Sciences 50(167), 56–62 (2007)
16. Parr, J.M., Keane, A.J., Forrester, A.I.J., Holden, C.M.E.: Infill sampling criteria for surrogate-based optimization with constraint handling. Engineering Optimization 44(10), 1147–1166 (2012)
17. Picheny, V.: Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction. Statistics and Computing pp. 1–16 (2014)

18. Picheny, V.: A stepwise uncertainty reduction approach to constrained global optimization. In: Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS), 2014, Reykjavik, Iceland. vol. 33, pp. 787–795. JMLR: W&CP (2014)
19. Ray, T., Tai, K., Seow, K.C.: Multiobjective design optimization by an evolutionary algorithm. Engineering Optimization 33(4), 399–424 (2001)
20. Sasena, M.J., Papalambros, P., Goovaerts, P.: Exploration of metamodeling sampling criteria for constrained global optimization. Engineering Optimization 34(3), 263–278 (2002)
21. Schonlau, M., Welch, W.J., Jones, D.R.: Global versus local search in constrained optimization of computer models. In: New Developments and Applications in Experimental Design: Selected Proceedings of a 1997 Joint AMS-IMS-SIAM Summer Conference. IMS Lecture Notes-Monographs Series, vol. 34, pp. 11–25. Institute of Mathematical Statistics (1998)
22. Shimoyama, K., Sato, K., Jeong, S., Obayashi, S.: Updating kriging surrogate models based on the hypervolume indicator in multi-objective optimization. Journal of Mechanical Design 135(9), 094503 (2013)
23. Wagner, T., Emmerich, M., Deutz, A., Ponweiser, W.: On expected-improvement criteria for model-based multi-objective optimization. In: Parallel Problem Solving from Nature, PPSN XI. 11th International Conference, Krakov, Poland, September 11-15, 2010, Proceedings, Part I. Lecture Notes in Computer Science, vol. 6238, pp. 718–727. Springer (2010)