# M23: Derivative Free Optimization (In English, Summer 2024)

Masters Course

Hall TBA, Rudower Chausee 25

Day/Time: TBA, Apr. 2024 - Sep. 2024

Aswin Kannan (aswin.kannan@hu-berlin.de)
Nachwuchsgruppenleiter, Inst. for Math., Humboldt Univ.

This course is aimed for the target audience of graduate students. Several engineering and scientific problems lead to optimization settings, where the objectives and constraints are governed by complex and time-consuming simulations. Additionally, there exist no-closed form expressions for these functions (objectives and/or constraints). Further, several of these simulators are proprietary, thus making applicability of automatic differentiation less conducive. The last two decades have made significant inroads by looking at either model based methods or direct search techniques to solve such problems. More notably, both these classes of algorithms have established convergence theory under some mild problem specific assumptions. We focus on the following questions as part of this course. We are interested in solving problems of the form:

$$\min_{\theta} \quad f(\theta), \qquad \text{where,} \qquad \theta - \text{inputs}, \qquad \text{and} \qquad f(\theta) - \text{outputs}.$$


Figure 1: Blackbox - Basics.

- Applications: Scientific settings from chemical engineering [ASZ10] and hydraulic modeling [KW12] to automotive design and imaging entail such formulations. Say, problems in automotive design entail maximization of fuel economy and minimization of acceleration time. Some underlying design variables are engine and gear specifications. The evaluation of the objectives require solving several systems of ordinary and partial differential equations for fixed design specifications. Similar examples can be easily noticed in hydraulics, where "*well*" installation locations are the design variables. The objectives refer to fluid flow levels, represented by flux. These similarly involve solving multiple differential systems. Some simulators deal with numerical integration, look-up tables, and solving large nonlinear systems of equations. We will focus on a plethora of such applications (both modeling and algorithmic aspects).

- Model based methods: We will discuss about model based algorithms belonging to the following two classes.

  - Bayesian Optimization [Kno06]: For problems of the form, $\min_{\theta \in \Theta} f(\theta)$, the theory rests on the assumption that $f(\theta)$ follows Gaussian processes: $f(\theta) \sim \mathcal{N}(\mu(\theta), \sigma^2(\theta, \theta))$. The heart of these schemes lie in estimating the parameters associated with $\mu$ and $\sigma$ by Gaussian regression. The related acquisition functions are optimized in place of the objective $f(\theta)$. One example of the acquisition function is $UCB(\mu, \sigma) = \mu(\theta) + \rho\sigma(\theta)$, where $\rho > 0$ is a user-defined scalar.
  - Trust region type methods [Pow06]: For the problem $\min_\theta f(\theta)$, quadratic surrogate models are formed in a sequential sense. Noting that finite differences can be very expensive, coarse approximations to gradients and hessians are obtained based on existing function evaluations. Model parameters $(g^f, H^f)$ are obtained by solving a multivariate interpolation problem

$$f(\theta^{\text{old}} + d) \approx m^f(\theta^{\text{old}} + d) = f(\theta^{\text{old}}) + d^T g^f + \frac{1}{2} d^T H^f d.$$

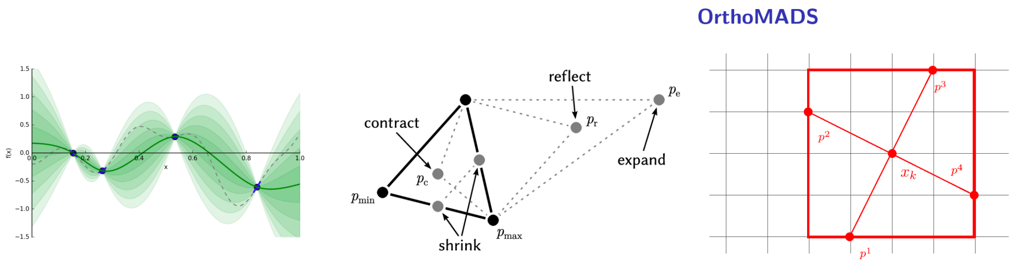


Figure 2:

Figure 3: Left - Bayesian Methods (Image Source - towardsdatascience.com). Middle - Nelder-Mead 2D example (Wikipedia - maximization problem shown). Right - Orthogonal MADS, source Polytechnique Montreal (Flip x with $\theta$ for notation).

Trust region type methods are deployed sequentially to aid in objective descent.

$$\min_{d} \left\{ m(d) = d^T g + \frac{1}{2} d^T H d : \ \theta + d \in \Theta; \ \|d\|_\infty \leq \Delta \right\}.$$

In addition to the above basics, we will study about trust region type algorithms for constrained and noisy versions of the original simulation based problem of interest.

- Search based methods [AD06]: Both the above model based methods work well when the simulation based function $f$ is inherently smooth. In cases of nonsmooth (but continuous) settings, search-type methods have been very beneficial. They are also very good performance-wise when the variables are discrete.

By the end of the course, students would be able to spot real-world situations where DFO arises. They would also be equipped with the basics on the algorithmic and implementation side to help in solving such problems. Several research works (publications) will be distributed during the course of the study. The following texts may be adhered [CSV09] to as key references to get started with the concepts. The tentative outline of the course is as follows.

- Class 1 – Basics of DFO.
- Classes 2-5 – Algorithms.
- Classes 6-8 – Applications.
- Classes 9-12 – Student presentations.

**Grading:** Grades will be a weighted combination of the following three components.

- DFO Summary (10 percent): A one page summary of your understanding of DFO and possible future research avenues. This will include class participation/questions.
- Project (50 percent): You will be split into teams of two and will be presented a real-instance DFO problem. You will work on modeling/solving this problem. This will be graded based on the final report.
- Presentation (40 percent): You will be given a research article based on your interests (algorithms/applications). You should present your understanding of the paper in one of the classes. This will be individual and you will not be split into teams.

# References

[AD06] Charles Audet and J. E. Dennis. Mesh adaptive direct search algorithms for constrained optimization. *SIAM Journal on Optimization*, 17(1):188–217, 2006.

[ASZ10] Charles Audet, Gilles Savard, and Walid Zghal. A mesh adaptive direct search algorithm for multiobjective optimization. *European Journal of Operational Research*, 204(3):545–556, 2010.

[CSV09] Andrew R. Conn, Katya Scheinberg, and Luís Nunes Vicente. Introduction to derivative-free optimization. In *MPS-SIAM series on optimization*, 2009.

[Kno06] Joshua Knowles. Parego: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1):50–66, 2006.

[KW12] Aswin Kannan and Stefan M. Wild. Benefits of deeper analysis in simulation-based groundwater optimization problems. In *Proceedings of the XIX International Conference on Computational Methods in Water Resources (CMWR 2012)*, June 2012. Available at `http://www.mcs.anl.gov/~wild/papers/2012/AKSW12.pdf`.

[Pow06] M. J. D. Powell. *The NEWUOA software for unconstrained optimization without derivatives*, pages 255–297. Springer US, Boston, MA, 2006.