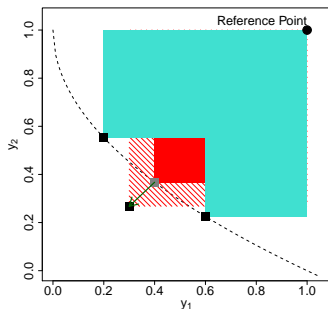
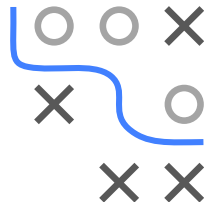


Optimization in Machine Learning

Bayesian Optimization

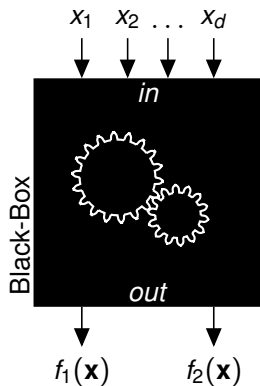
Multicriteria Bayesian Optimization



Learning goals

- Multicriteria Optimization
- Taxonomy
- ParEGO, SMS-EGO, EHI

MULTICRITERIA BAYESIAN OPTIMIZATION



$$f : \mathcal{S} \rightarrow \mathbb{R}^m$$

$$\min_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$$

- A configuration \mathbf{x} **dominates** (\prec) $\tilde{\mathbf{x}}$ if

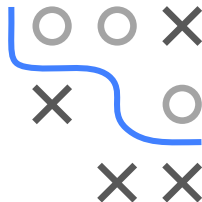
$$\forall i \in \{1, \dots, m\} : f_i(\mathbf{x}) \leq f_i(\tilde{\mathbf{x}})$$

$$\text{and } \exists j \in \{1, \dots, m\} : f_j(\mathbf{x}) < f_j(\tilde{\mathbf{x}})$$

- Set of non-dominated solutions:

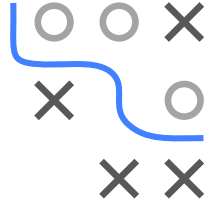
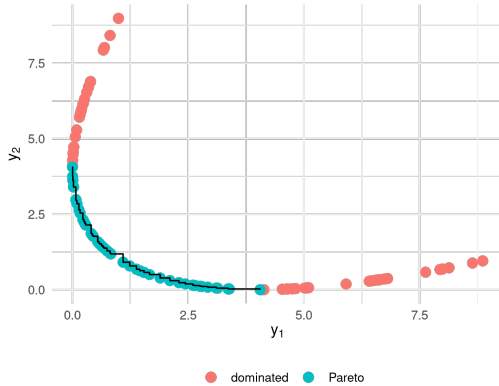
$$\mathcal{P} := \{\mathbf{x} \in \mathcal{S} \mid \nexists \tilde{\mathbf{x}} \in \mathcal{S} : \tilde{\mathbf{x}} \prec \mathbf{x}\}$$

- Pareto set \mathcal{P} , Pareto front $\mathcal{F} = f(\mathcal{P})$
- Goal: Find $\hat{\mathcal{P}}$ of non-dominated points that estimates the true Pareto set \mathcal{P}

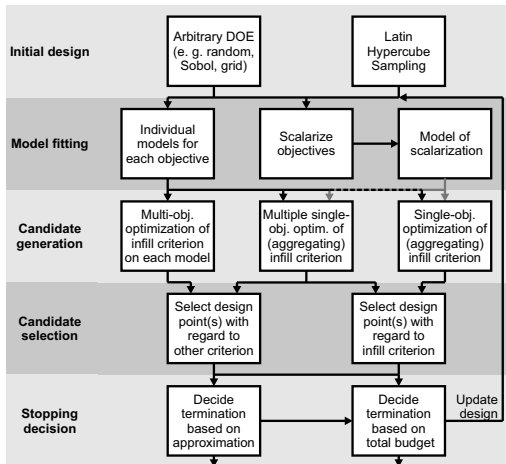


MULTICRITERIA BAYESIAN OPTIMIZATION / 2

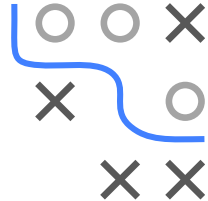
Example Pareto front:



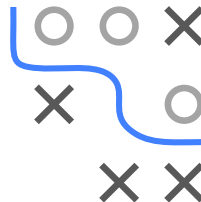
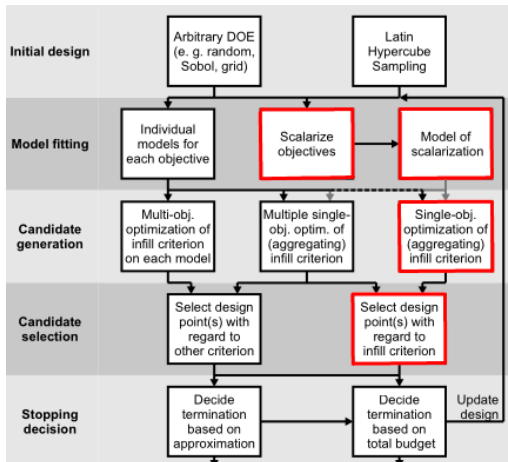
TAXONOMY



Horn, Wagner, Bischl et al. (2014).



PAREGO



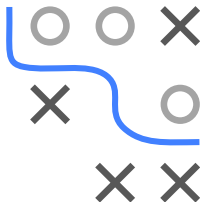
PAREGO / 2

- 1 Scalarize standardized objectives using the augmented Tchebycheff norm

$$\max_{i \in \{1, \dots, m\}} w_i f_i(\mathbf{x}) + \rho \sum_{i=1}^m w_i f_i(\mathbf{x})$$

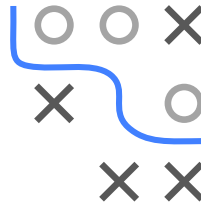
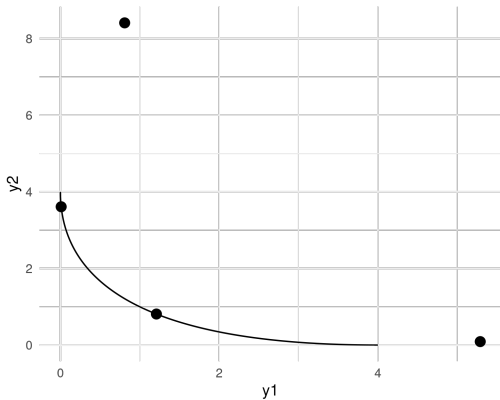
with weight vector \mathbf{w} drawn uniformly from the set of evenly distributed weight vectors \mathcal{W}

- 2 Fit SM on the scalarized objective function
- 3 Proceed to use any standard single-objective acquisition function (EI, PI, LCB, ...)



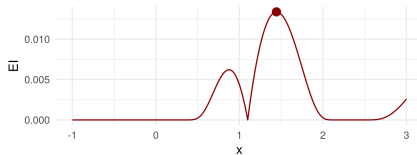
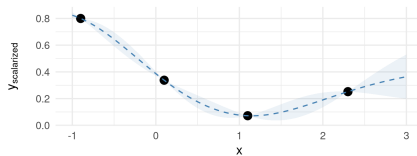
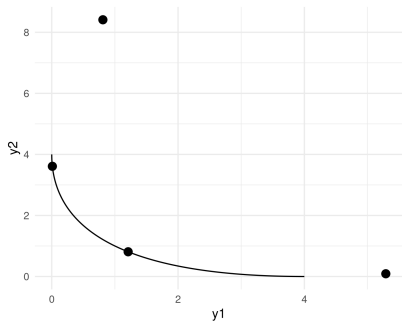
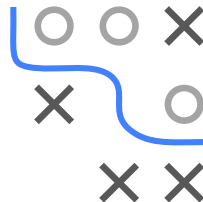
PAREGO / 3

ParEGO Example, initial design and true Pareto front in black ...

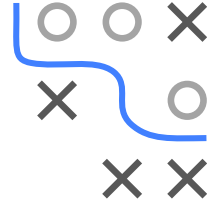
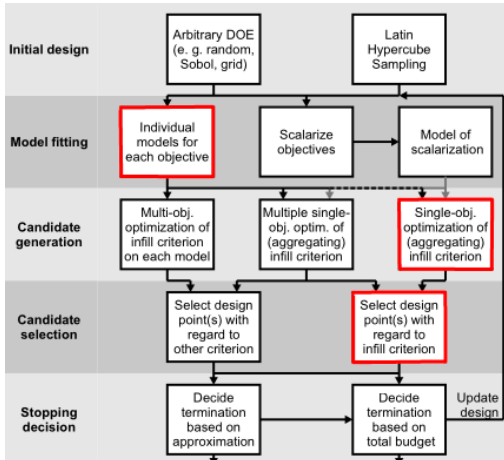


PAREGO / 4

... standardize objectives, obtain scalarized objective via augmented Tchebycheff norm, fit SM and optimize EI ...



SMS-EGO

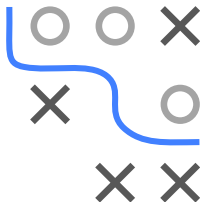
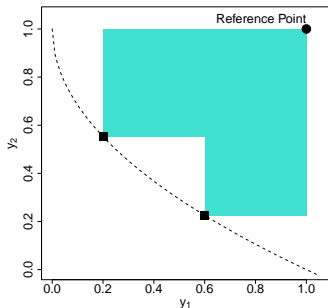


SMS-EGO

Individual models for each objective f_i

Single-objective optimization of aggregating acquisition function:
Calculate contribution of the confidence bound of candidate to the current front approximation

- Calculate LCB for each objective
- Measure contribution with regard to the hypervolume improvement
- For ϵ -dominated (\succ_{ϵ}) solutions, a penalty is added

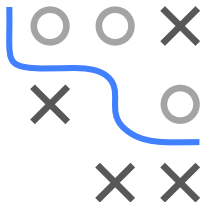
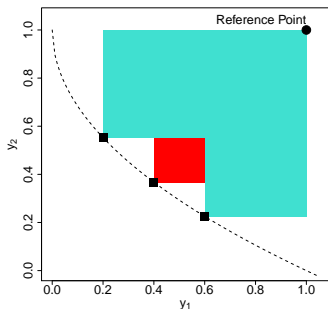


SMS-EGO

Individual models for each objective f_i

Single-objective optimization of aggregating acquisition function:
Calculate contribution of the confidence bound of candidate to the current front approximation

- Calculate LCB for each objective
- Measure contribution with regard to the hypervolume improvement
- For ϵ -dominated (\succ_{ϵ}) solutions, a penalty is added

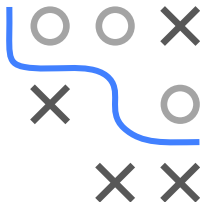
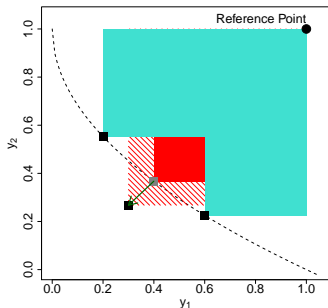


SMS-EGO

Individual models for each objective f_i

Single-objective optimization of aggregating acquisition function:
Calculate contribution of the confidence bound of candidate to the current front approximation

- Calculate LCB for each objective
- Measure contribution with regard to the hypervolume improvement
- For ϵ -dominated (\succ_{ϵ}) solutions, a penalty is added



OUTLOOK

Many more options exist:

- Expected Hypervolume improvement
- Multi-Ego
- Entropy based: PESMO, MESMO
- ...

