Constrained Bayesian Optimisation with Knowledge Gradient

Juan Ungredda

University of Warwick

Constrained KG

Non-constrained KG:

$$KG(x) = \mathbb{E}[\max_{x'}\{\mu^{n+1}(x')\}|x^{n+1} = x]$$

Constrained KG with Probability of Feasibility (pf):

$$c-KG(x) = \mathbb{E}[\max_{x'}\{pf^{n+1}(x')\mu^{n+1}(x')\}|x^{n+1} = x]$$

Constrained KG corrected*:

$$c-KG(x) = \mathbb{E}[\max_{x'}\{pf^{n+1}(x')\mu^{n+1}(x') + (1-pf^{n+1}(x'))M\}|x^{n+1} = x]$$

where $M \in \mathbb{R}$ is the penalisation for sampling points in an infeasible region. It's commonly assumed to be zero.

Constrained KG

Benefits

- Takes into account that constraints change for each possible x^{n+1} considered.
- Assuming $M = 0$ may give "benefit" to infeasible regions. A more general approach avoids that problem.

Limitations

- Computationally expensive compared to constrained Expected Improvement. However, it's possible to make an efficient implementation by using gradients.
- M may be need to chosen by the decision maker.

Results

benchmark Method:

- Constrained Expected Improvement. i.e. Expected Improvement times Probability of feasibility.

Test Functions:

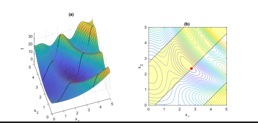
- New Branin, test function 2, and Mistery function. Non-noisy and constrained test functions.

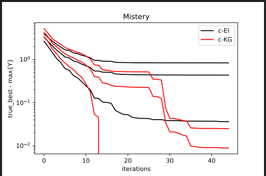
Important Remark:

- In the last evaluation of our approach, the function is sampled according to Expected Improvement times the probability of feasibility.

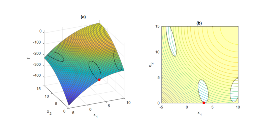
Plots

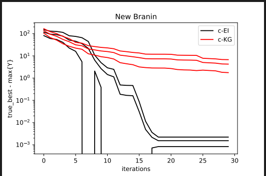
Mistery



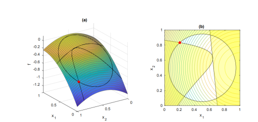


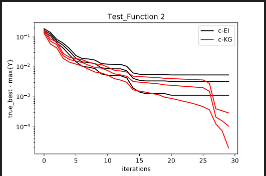
New Branin





Test function 2





Current Work

- Include Gradient Information.
- Include noisy test functions.
- Set dynamically the M penalization.