Sequential designs for computer experiments

Pierre BARBILLON

AgroParisTech / INRA MIA UMR 518

SPADRO

Outline

Context

- A computer experiment is an evaluation of a determinist expensive black-box function describing a physical system:

$$f : \mathbf{x} \in E \subset \mathbb{R}^d \mapsto \mathbb{R} .$$

Ref. : Fang et al. (2006), Koehler et Owen (1996), Santner et al. (2003) .

- Expensive $\Leftrightarrow$ N possible calls to f.
- Uncertainties on inputs $\Rightarrow$ modelled by a random variable $\mathbf{X}$ (known distribution).

Goal

- Vizualisation,
- Optimization,
- Output distribution: mean, median, quantile, probability of rare events...
- Inverse problem.

Metamodelling: prior distribution on f

Sacks et al. (1989).
f realization of a Gaussian process F:
$\forall \mathbf{x} \in E$,

$$F(\mathbf{x}) = \sum_{k=1}^{Q} \beta_k h_k(\mathbf{x}) + \zeta(\mathbf{x}) = H(\mathbf{x})^T \boldsymbol{\beta} + \zeta(\mathbf{x}),$$

where

- $h_1, \ldots, h_Q$ regression functions and $\boldsymbol{\beta}$ parameters vector,
- ζ centred Gaussians process with covariance function:

$$\mathrm{Cov}(\zeta(\mathbf{x}), \zeta(\mathbf{x}')) = \sigma^2 K(\mathbf{x}, \mathbf{x}'),$$

where K is a correlation kernel.

Hypotheses

- parameters $\boldsymbol{\beta}, \sigma^2$ and those of K assumed fixed;
- process F independent of $\mathbf{X}$.

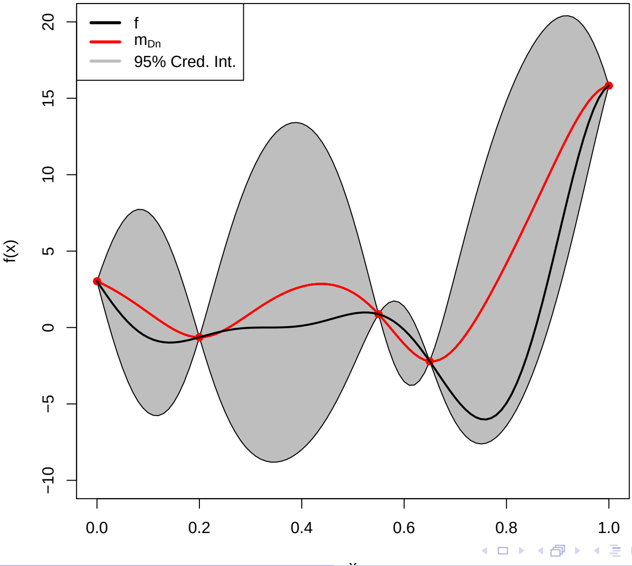
Metamodelling: posterior

- $y_1 = f(\mathbf{x}_1), \ldots, y_n = f(\mathbf{x}_n)$ evaluations of f on a design D_n ($n \leq N$).
- Conditioning F to $\{F(\mathbf{x}_1) = y_1, \ldots, F(\mathbf{x}_n) = y_n\} = F(D_n)$.
  Gaussian Process with mean $m_{D_n}(\mathbf{x})$ and covariance $C_{D_n}(\mathbf{x}, \mathbf{x}')$ $\forall \mathbf{x}, \mathbf{x}'$.

  For all $\mathbf{x} \in E$,
  - $m_{D_n}(\mathbf{x})$ approximates $f(\mathbf{x})$,
  - $C_{D_n}(\mathbf{x}, \mathbf{x})$ uncertainty on this approximation.

For all $\mathbf{x}_i \in D_n$,

- $m_{D_n}(\mathbf{x}_i) = f(\mathbf{x}_i)$,
- $C_{D_n}(\mathbf{x}_i, \mathbf{x}_i) = 0$.


f
m_{Dn}
95% Cred. Int.
f(x)
20
15
10
5
0
−5
−10
0.0
0.2
0.4
0.6
0.8
1.0

Outline

Principle

1. Construct a first exploratory design: D_n s. t. $n \leq N$,
2. For $i = n + 1 ... N$ do $D_i = D_{i-1} \cup \{\mathbf{x}_i\}$ where

$$\mathbf{x}_i \in \arg\max Crit(D_{i-1}, f) \,.$$

$Crit(D_{i-1}, f)$ can be adapted to the applied goal (optimization, estimation of probability of rare event).

Plan

Expected Improvement criterion

- Goal: Find the global extremum (here minimum e.g.) of f,
- Expected improvement criterion proposed by Jones et al. (1998):

$$EI_n(\mathbf{x}) = \mathbb{E}((min_n - F(\mathbf{x}))^+ | F(D_n)) \,,$$

where min_n is the current minimum value:

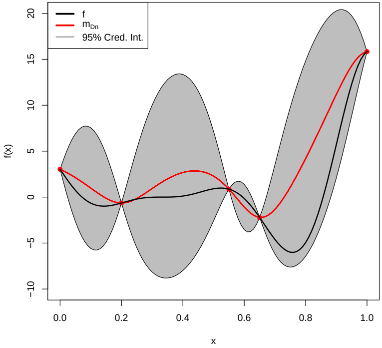
$$min_n = \min_{1,\dots,n} f(\mathbf{x}_i)$$

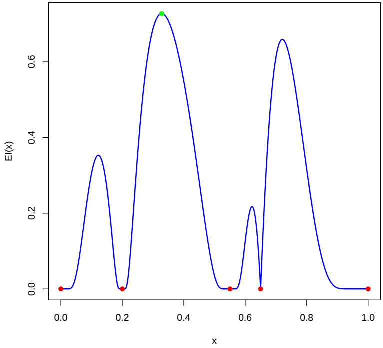
- Closed-form computation:

$$EI_n(\mathbf{x}) = (min_n - m_{D_n}(\mathbf{x}))\Phi\left(\frac{min_n - m_{D_n}(\mathbf{x})}{\sqrt{C_{D_n}(\mathbf{x},\mathbf{x})}}\right) + \sqrt{C_{D_n}(\mathbf{x},\mathbf{x})}\phi\left(\frac{min_n - m_{D_n}(\mathbf{x})}{\sqrt{C_{D_n}(\mathbf{x},\mathbf{x})}}\right)$$

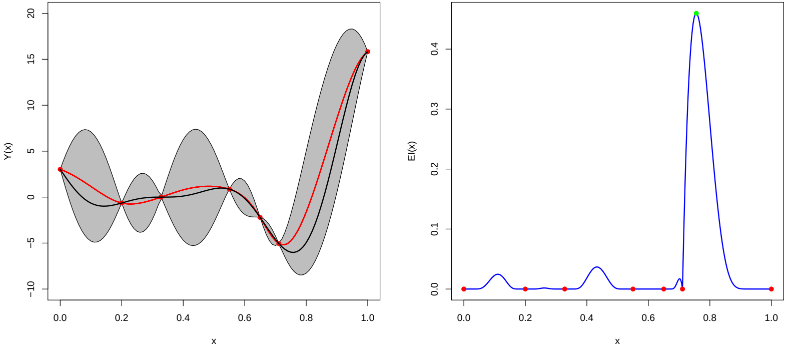
where Φ and ϕ are respectively the cdf and the pdf of $\mathcal{N}(0,1)$.

Example step 1

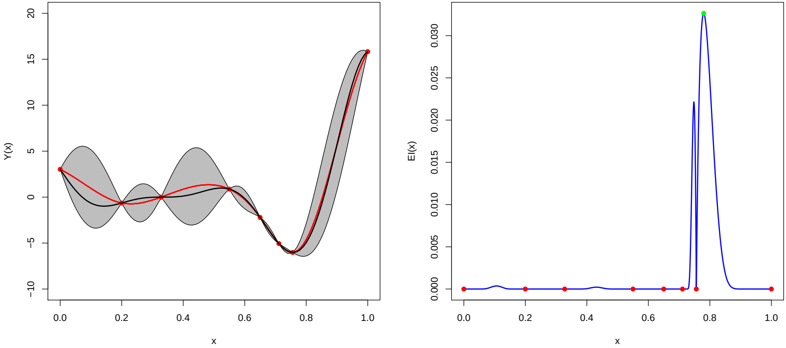




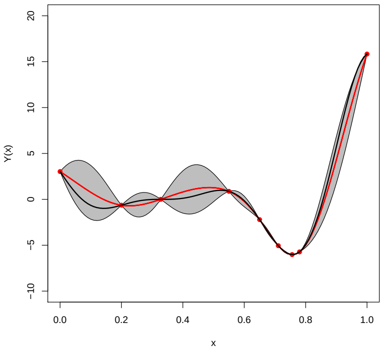
Example step 2

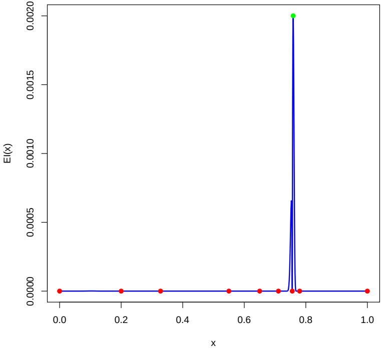


Example step 3

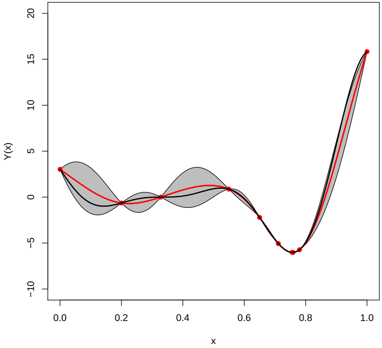


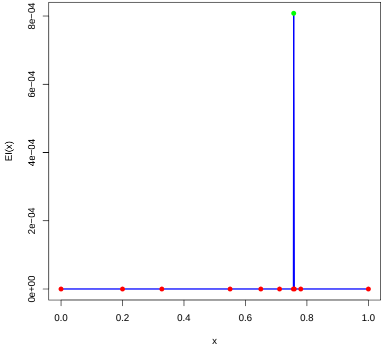
Example step 4





Example step 5





GP-UCB

Gaussian Process - Upper Confidence Bound (Srinivas et al., 2010):

- Add point $\mathbf{x}_i$ s. t.

$$\mathbf{x}_i \in \arg\max_{\mathbf{x}} m_{D_{i-1}}(\mathbf{x}) + \beta_i^{1/2} \cdot \sqrt{C_{D_{i-1}}(\mathbf{x}, \mathbf{x})}$$

- Bounds on cumulative regret

$$R_T = \sum_{i=1}^{T} f(\mathbf{x}^*) - f(\mathbf{x}_i)$$

  for a well chosen sequence $(\beta_i)_i$ and depending on the covariance kernel of the GP.

- in the pratical cases studied in the paper, similar performance of EI and $GP - UCB$.

Plan

Principle

- Uncertainties on $\mathbf{x} \Rightarrow$ random variable $\mathbf{X}$,
- Goal: estimate the probability $\alpha = \mathbb{P}(f(\mathbf{X}) > c)$ under the constraint of a limited number N of calls to f.
- For a design D_n, estimation of α:

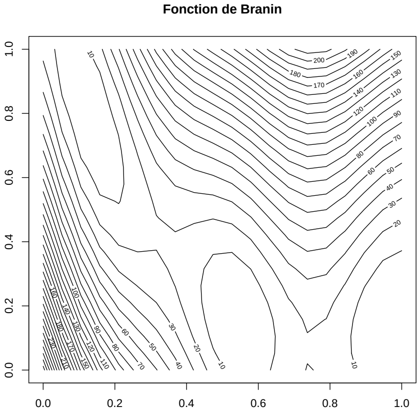
$$\hat{\alpha}_n = \mathbb{E}\left[\mathbb{P}(F(\mathbf{X}) > c) | F(D_n)\right] = \int \Phi\left(\frac{m_{D_n}(\mathbf{x}) - c}{\sqrt{C_{D_n}(\mathbf{x}, \mathbf{x})}}\right) d\mathbb{P}_{\mathbf{X}}(x)$$

- Sequential design criterion:

$$\mathbf{x}_i \in \arg\max_{\mathbf{x}} \mathbb{E}\left[\mathbb{E}((\hat{\alpha}_i - \alpha)^2 | F(D_{i-1}) \cap \mathbf{x}_i = \mathbf{x}) | F(D_{i-1})\right] ,$$

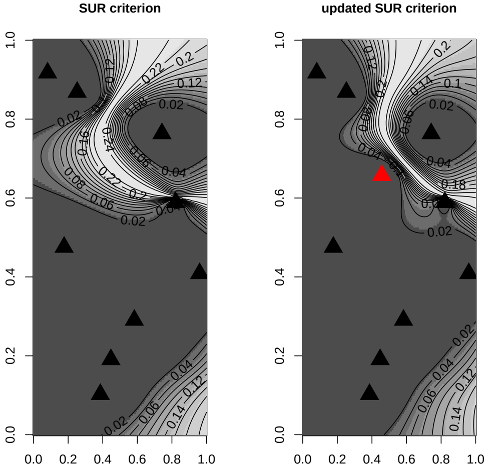
proposed in Chevalier et al. (2012).

Example



- $\mathbf{X} \sim \mathcal{U}([0,1]^2)$,
- estimate $\mathbb{P}(f(\mathbf{X}) > 80)$.

SUR criterion



point-wise probability of exceeding the threshold