

Lecture 23: Bayesian global optimization

Professor Ilias Bilonis

Expected improvement - No observation noise

Most commonly used IAF

Derivation of expected improvement

→ $x_{1:n} = (x_1, \dots, x_n); y_{1:n} = (y_1, \dots, y_n); \sigma = 0$

$f(\cdot) \sim \mathcal{GP}(m(\cdot), c(\cdot, \cdot)) \Rightarrow$ posterior point predictive distribution

→ $p(y|x, x_{1:n}, y_{1:n}) = \mathcal{N}(y | \mu_n^*(x), \sigma_n^{*2}(x))$

→ Current best observed output: $\bar{y}_n = \max_{1 \leq i \leq n} y_i$

Assume make hypothetical obs. (x, y) :

Improvement: $I(x, y) = \begin{cases} 0, & y \leq \bar{y}_n \\ y - \bar{y}_n, & y > \bar{y}_n \end{cases} = \max\{0, y - \bar{y}_n\}$

Exp. Improvement: $E I(x) \stackrel{\text{def.}}{=} E[I(x, y) | x, x_{1:n}, y_{1:n}]$

(since y has not actually been measured)

$= \int_{-\infty}^{\infty} I(x, y) p(y|x, x_{1:n}, y_{1:n}) dy$

expectation over y

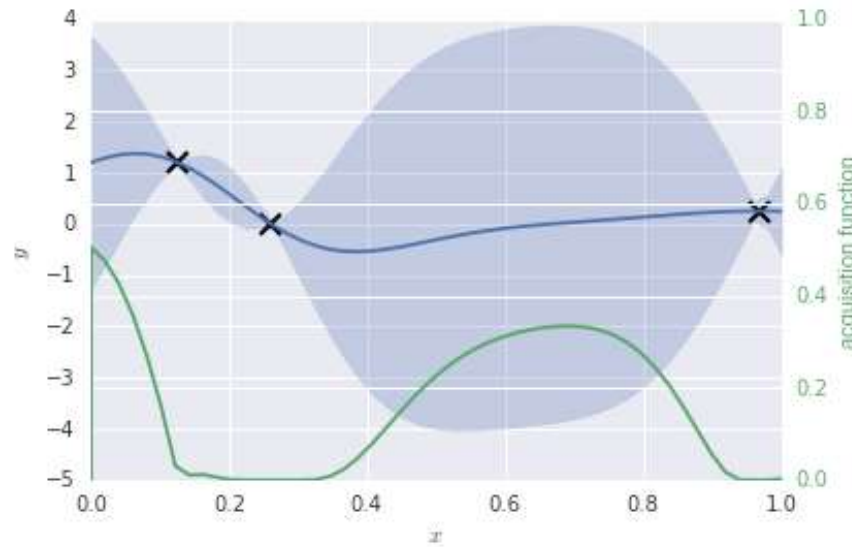
CDF $\mathcal{N}(0, 1)$

$= (\mu_n^*(x) - \bar{y}_n) \Phi\left(\frac{\mu_n^*(x) - \bar{y}_n}{\sigma_n^*(x)}\right) + \sigma_n^*(x) \phi\left(\frac{\mu_n^*(x) - \bar{y}_n}{\sigma_n^*(x)}\right)$

Exploitation

pdf $\mathcal{N}(0, 1)$

Expected Improvement



Automatic exploration vs exploitation...