Bayesian Optimization

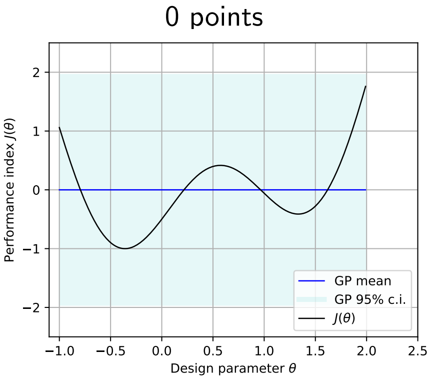
Overview

- BO is a tool to solve
$$\min_{\theta} J(\theta)$$
when θ is unknown, but a set of measurements $J(\theta_1), J(\theta_2), \ldots, J(\theta_N)$ can be collected;
- It iteratively updates a Bayesian surrogate model of $J(\theta)$;
- The measurements are collected to favor points with estimated good performance $\rightarrow$ exploitation and/or high variance $\rightarrow$ exploration

Bayesian Optimization

Gaussian Process

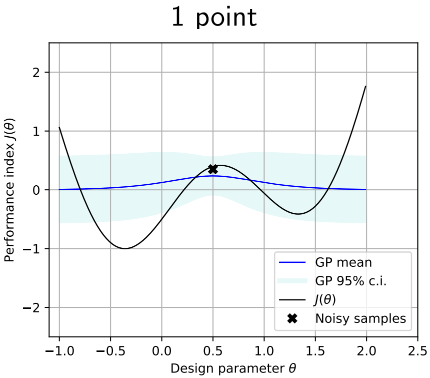
- The function $J(\theta)$ assumed Gaussian with prior mean $E[J(\theta)] = \mu(\theta)$ and covariance $\mathrm{cov}[J(\theta_1), J(\theta_2)] = \kappa(\theta_1, \theta_2)$.
- The posterior mean and covariance given a new observation (θ_i, J_i) is obtained in closed form



Bayesian Optimization

Gaussian Process

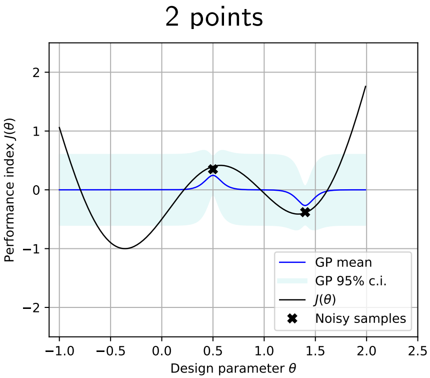
- The function $J(\theta)$ assumed Gaussian with prior mean $E[J(\theta)] = \mu(\theta)$ and covariance $\mathrm{cov}[J(\theta_1), J(\theta_2)] = \kappa(\theta_1, \theta_2)$.
- The posterior mean and covariance given a new observation (θ_i, J_i) is obtained in closed form



Bayesian Optimization

Gaussian Process

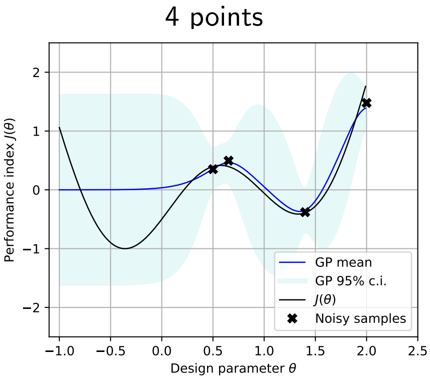
- The function $J(\theta)$ assumed Gaussian with prior mean $E[J(\theta)] = \mu(\theta)$ and covariance $\text{cov}[J(\theta_1), J(\theta_2)] = \kappa(\theta_1, \theta_2)$.
- The posterior mean and covariance given a new observation (θ_i, J_i) is obtained in closed form



Bayesian Optimization

Gaussian Process

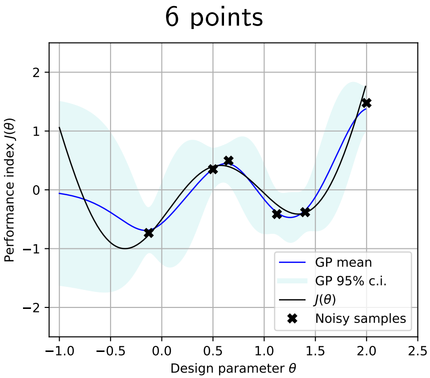
- The function $J(\theta)$ assumed Gaussian with prior mean $E[J(\theta)] = \mu(\theta)$ and covariance $\mathrm{cov}[J(\theta_1), J(\theta_2)] = \kappa(\theta_1, \theta_2)$.
- The posterior mean and covariance given a new observation (θ_i, J_i) is obtained in closed form



Bayesian Optimization

Gaussian Process

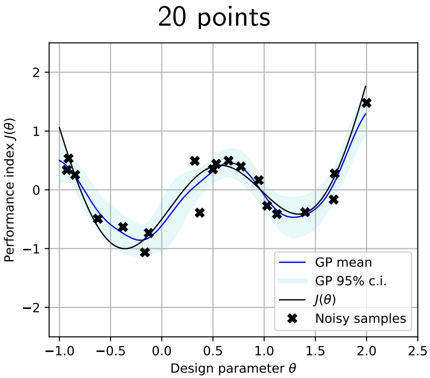
- The function $J(\theta)$ assumed Gaussian with prior mean $E[J(\theta)] = \mu(\theta)$ and covariance $\text{cov}[J(\theta_1), J(\theta_2)] = \kappa(\theta_1, \theta_2)$.
- The posterior mean and covariance given a new observation (θ_i, J_i) is obtained in closed form



Bayesian Optimization

Gaussian Process

- The function $J(\theta)$ assumed Gaussian with prior mean $E[J(\theta)] = \mu(\theta)$ and covariance $\mathrm{cov}[J(\theta_1), J(\theta_2)] = \kappa(\theta_1, \theta_2)$.
- The posterior mean and covariance given a new observation (θ_i, J_i) is obtained in closed form



Bayesian Optimization

Acquisition function

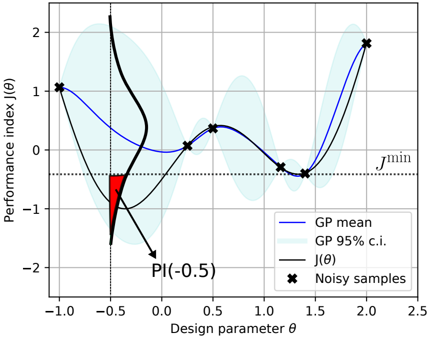
The GP provides the probability distribution of $J(\theta)$ for each parameter θ.
This probability is used to define an acquisition function, *e.g.*,

Probability of Improvement

$$A(\theta) = \mathrm{PI}(\theta) = p(J(\theta) \leq J^{\mathrm{min}})$$

Expected improvement

$$A(\theta) = \mathrm{EI}(\theta) = \mathbb{E}[\max\left(0, J^{\mathrm{min}} - J(\theta)\right)]$$



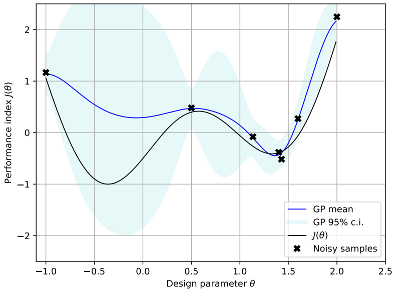
Bayesian Optimization

Overview

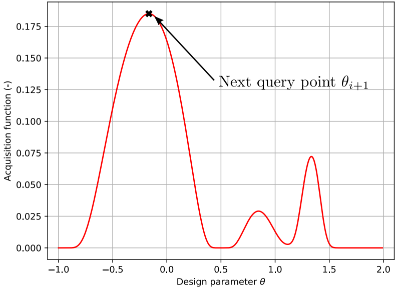
Steps of BO: for $i = 1, 2, \ldots i_{\max}$

1. **Execute** experiment with θ_i, measure $J_i = J(\theta_i) + e_i$
2. **Update** the GP model $\theta \to J(\theta)$ with (θ_i, J_i)
3. **Construct** acquisition function $A(\theta)$
4. **Maximize** $A(\theta)$ to obtain next query point θ_{i+1}

GP at iteration i



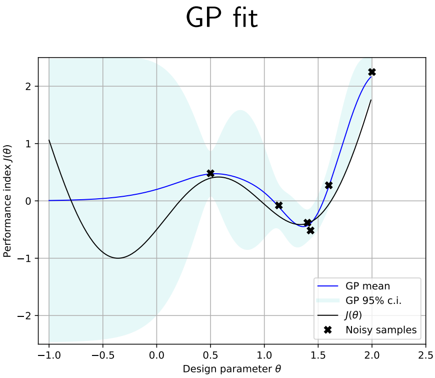
$A(\theta)$ at iteration i

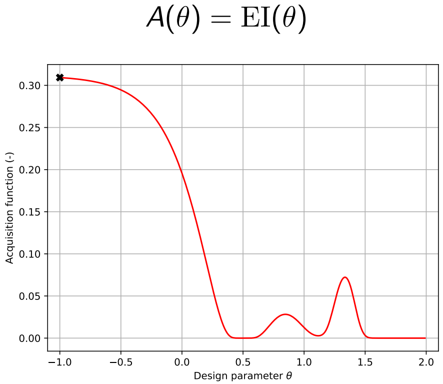


Bayesian Optimization

Example

iteration 6

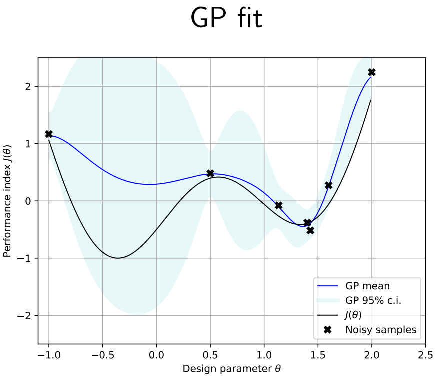


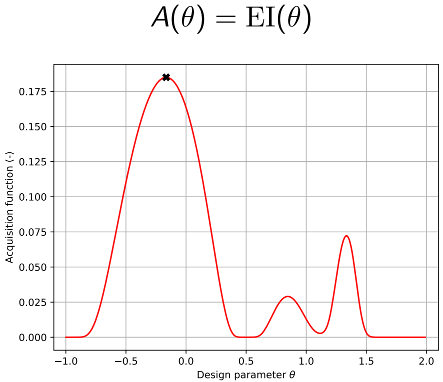


Bayesian Optimization

Example

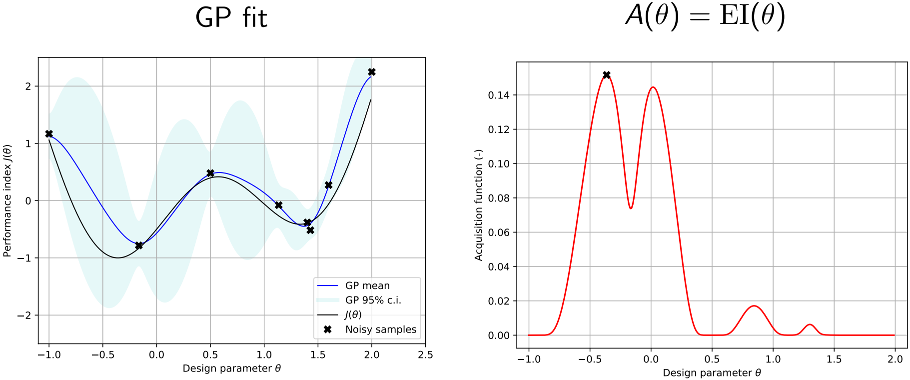
iteration 7





Bayesian Optimization
Example

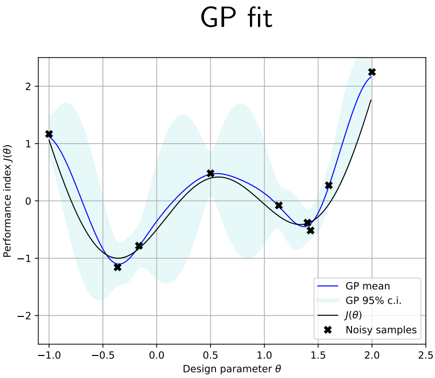
iteration 8

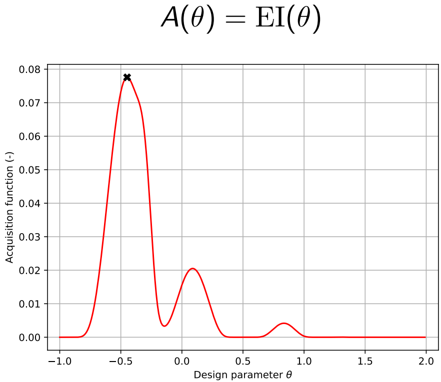


Bayesian Optimization

Example

iteration 9

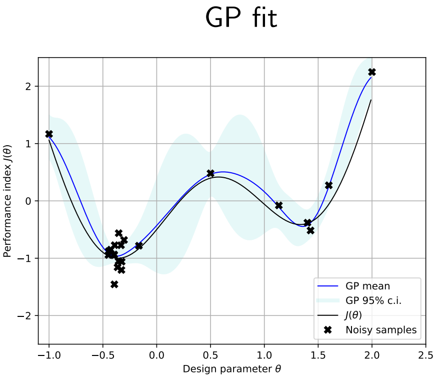


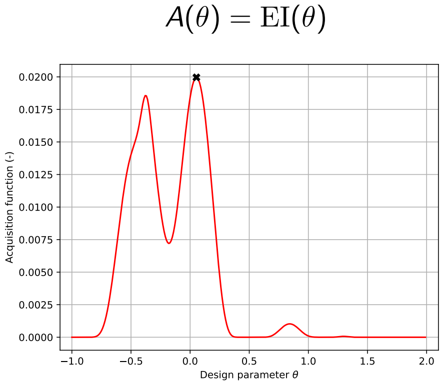


Bayesian Optimization

Example

iteration 20





Bayesian Optimization

Example

iteration 20