Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design

Niranjan Srinivas, Andreas Krause[†], Sham Kakade[††], Matthias Seeger[†††]

†: California Institute of Technology
††: University of Pennsylvania
†††: Saarland University

January 9, 2023

目次

Next Section

背景

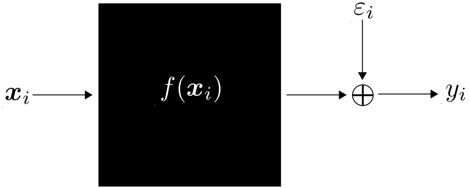
- 多くの実問題は，目的関数に対する最適変数探索問題として定式化できる
  - 耐久性の高いロボットの開発

ブラックボックス関数

- ブラックボックス関数 f

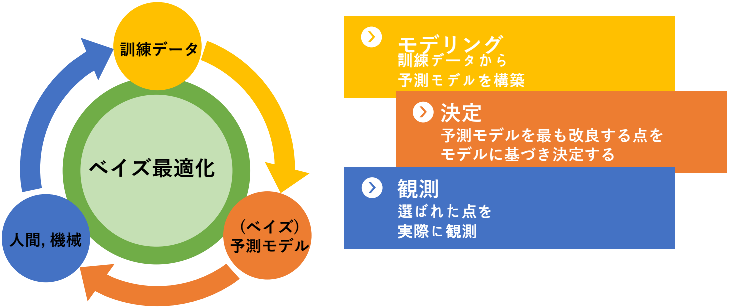
$$y_i = f(\boldsymbol{x}_i) + \varepsilon_i$$

入力：$\boldsymbol{x}_i$，　出力：y_i，　誤差：ε_i



- 実応用で扱う対象はブラックボックス関数であることが多い
- ブラックボックス関数は，主に以下の 2 つの性質を持つ
  - 関数の具体的な形状が不明
  - 各入力における関数値を得るには，大きなコストがかかる

ベイズ最適化



- 多くの実問題は，ブラックボックス関数最適化と等価
- できるだけ少ない関数評価回数で最適化を見つけたい
- 有効な手法としてベイズ最適化がある

本論文の貢献

- 本研究の目的
- 本研究の意義
    -
    -

Next Section

問題設定

- 候補入力 $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ が与えられている
- 関数 f を評価して出力 $y_i = f(\boldsymbol{x}_i)$ を得るにはコストがかかる
- できるだけ少ないコストで関数 f を最大化するパラメータ $\boldsymbol{x}$ を求めたい

$$\boldsymbol{x}^* = \operatorname*{arg\,max}_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$$

  - ガウス過程は，確率分布を用いた複雑で非線形な関数を扱うためのモデルである．
  - このような複雑で非線形な関数を扱う場合には，一般的な最適化手法では，解決が困難になることがある．

ガウス過程

任意の入力 $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ に対して $\{f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n)\}$ が n 次元正規分布に従うなら，f はガウス過程に従う．

- 関数 f がガウス過程に従うことを $f(\boldsymbol{x}) \sim \mathcal{GP}(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}'))$
- 平均関数: $\mu(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})]$
- 共分散関数 (カーネル関数):
  $k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - \mu(\boldsymbol{x}))(f(\boldsymbol{x}') - \mu(\boldsymbol{x}'))]$

ベイズ最適化 1

- f にガウス過程事前分布を仮定する
- 訓練データに基づき，f の事後分布を求める
- 事後分布に基づき最も最大値となりそうな点を次に観測する
- 観測した (x_{next}, y_{next}) を訓練データに追加し，再び事後分布を求める

獲得関数

-

従来の獲得関数

variance only

$$\boldsymbol{x} = \arg\max_{\boldsymbol{x}} \ \sigma_T^2(\boldsymbol{x})$$

mean only

$$\boldsymbol{x} = \arg\max_{\boldsymbol{x}} \ \mu_T(\boldsymbol{x})$$

-

Next Section

- 探索・活用に特化した

GP-UCB: Gaussian Process Upper Confidence Bound

$$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}} \ \mu_{t-1}(\boldsymbol{x}) + \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}).$$

- β_t は，探索の度合いを表す定数
  - β_t が小さい $\rightarrow$ 探索
  - β_t が大きい

GP-UCB のアルゴリズム

- 初期化: 最初に，f の予測分布を設定する．これは f の予測分布を表すカーネルを選択し，そのカーネルに対応するガウス過程を設定することで行われる．
- アクションの選択: 次に，f の予測分布から，次のような式を用いてアクションを選択する
  - $a_t = argmax_{a \in D}(\mu_t(a) + \beta_t \sigma_t(a))$

  ここで、$\mu_t(a)$ は、時刻 t において、a が選択されると予測される f の平均値を表します。$\sigma_t(a)$ は、時刻 t において、a が選択されると予測される f の不確実性を表します。β_t は、時刻 t においてのアクションの選択におけるリスク係数を表します
- 報酬の観測: 選択したアクション a_t に対して報酬 r_t を観測する．
- 予測分布の更新: 次に，最新のアクション a_t を使用して f の予測分布を更新する. これは，ガウス過程のアップデートによって行われる．

Next Section

実験手法

-

実験対象データ

- 人工的に作成した合成データ
  - 長さスケールパラメータ 0.2 の二乗指数カーネルからランダムな関数をサンプリング
  - サンプリングノイズ分散
- 46 個のセンサーを用いて 5 日間に渡って収集された温度データ (Intel Research Berkeley)
  - 具体的に説明
- カリフォルニア州 I-880 South 高速道路に設置された交通センサーのデータ
  - 具体的に説明

実験結果

-

Next Section

結論

-