

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 7, 2023

- * 本スライドでは，背景 (導入) として，実例について説明する
- 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
- ここに，新規タンパク質の開発や，ロボットの開発などに関する，具体的な例を図とともに示す

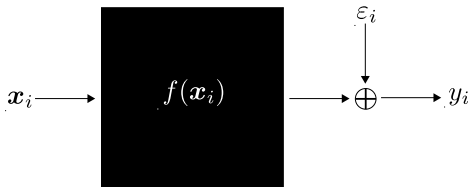
ブラックボックス関数

* 本スライドでは、ブラックボックス関数について説明する (前提知識の提供)

- ブラックボックス関数 f

$$y_i = f(x_i) + \varepsilon_i$$

入力: 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

- * 本スライドでは、ベイズ最適化について説明する (前提知識の提供)
 - 初めに前スライドで説明したブラックボックス関数の最適化を行うための手法として、ベイズ最適化があることを述べる
- 本スライドに、ベイズ最適化のフローに関する図を示す
 - 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

- * 本スライドでは、本論文の貢献として、ベイズ最適化の一つの手法である GP-UCB を提案したことを示す
- * また、その手法が、探索や活用のバランスが取れた手法であり、実験的に良い性能を示すことを述べる
- 本研究の目的
 -
 -
- 本研究の意義
 -
 -

- * 本スライドでは、問題設定を定式化して示す
- 候補入力 $\mathcal{X} = \{x_1, \dots, x_n\}$ が与えられている
- 関数 f を評価して出力 $y_i = f(x_i)$ 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 x を求めたい

$$x^* = \arg \max_{x \in \mathcal{X}} f(x)$$

- 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。

- * 本スライドでは，ガウス過程の定義を示す (前提知識の提供)

ガウス過程

- * ガウス過程の定義を記述する

- * 本スライドでは、(ガウス過程を用いた) ベイズ最適化のアルゴリズムを示す
- 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

- * 本スライドでは，獲得関数とは何かについて示す

- * 本スライドでは，従来の獲得関数について示す (探索重視のものと活用重視のもの)

- * 本スライドでは，GP-UCB の定義を示す

- * 本スライドでは、GP-UCB のアルゴリズムを示す
- 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
 - $a_t = \operatorname{argmax}_{a \in D} (\mu_t(a) + \beta_t \sigma_t(a))$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.

- * 本スライドでは，実験手法を示す (性能を比較する手法や regret の簡単な説明)

- * 本スライドでは、実験設定 (実験を行うデータセット) を示す
- 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

- * 本スライドでは，実験結果を示す
- 本スライドに，論文中の Figure4 の図を挿入する

- * 本スライドでは，論文のまとめを行う