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

December 28, 2022

目次

Next Section

ブラックボックス関数

- 序論として，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)
- 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
- 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
  - 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ために，大きなコストを要する

ガウス過程

- 序論として，ガウス過程について説明する (前提知識の提供)
  - 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．

ベイズ最適化

- 序論として，ベイズ最適化について説明する (前提知識の提供)
  - ベイズ最適化は，確率論的アプローチを用いた最適化手法である
  - そのため，不確実性を含む問題を解決するために用いらていることが多い
  - そのような問題を扱うために，事前分布を用いて事後分布を求めることができる
  - 論文中では，このようなベイズ最適化を用いて，ガウス過程を用いた最適化問題を解決するアルゴリズムを提案している．
  - また，そのアルゴリズムがバンディット設定でも有効であることを示している．

バンディット問題

- 序論として，バンディット問題について説明する (前提知識の提供)
  - バンディット設定とは，決定をする際に報酬を得ることができるが，そのためには，その決定をする前にその詳細を知ることができない設定を指す
  - このような設定では，報酬を最大化することを目的として，決定をする場所を最適に選択することが重要になる
  - バンディット問題の具体例: 投資をする場合や，商品を購入する場合など．このような場面では，最適な決定をすることで，最大の報酬を得ることができるため，重要である．

導入

- 本研究の目的
  -
  -
- 本研究の意義
  -
  -
- 計算機科学や統計学において，最適化問題は非常に重要であることを強調する
- ガウス過程を用いた最適化は，複雑で非線形な関数を扱う場合に有用であることを紹介する
- バンディット設定において，最適な決定をすることが重要であることを紹介する
- 今までにバンディット設定でのガウス過程最適化については，リグレット制限が証明されていなかった
- 本論文では，バンディット設定におけるガウス過程最適化において，リグレット制限を証明することで，最適な決定をするアルゴリズムを提案する

Next Section

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

ベイズ最適化 2

ブロック

a

- テスト

ベイズ最適化 3

- テスト

Next Section

GP-UCB1

- GP-UCB は，バンディット設定でのガウス過程最適化において用いられるアルゴリズムである．
- バンディット設定では，報酬を最大化することを目的として，決定をする場所を最適に選択することが重要である．
- GP-UCB はこのような場合に，ガウス過程を用いて最適な決定をするためのアルゴリズムである
-

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

実験設定

- 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

- テスト

Next Section

まとめ 1

- We analyze GP-UCB, an intuitive algorithm for GP optimization, when the function is either sampled from a known GP, or has low RKHS norm.
- We bound the cumulative regret for GP-UCB in terms of the information gain due to sampling, establishing a novel connection between experimental design and GP optimization.
- By bounding the information gain for popular classes of kernels, we establish sublinear regret bounds for GP optimization for the first time. Our bounds depend on kernel choice and parameters in a fine-grained fashion.
- We evaluate GP-UCB on sensor network data, demonstrating that it compares favorably to ex- isting algorithms for GP optimization.

まとめ 2

- テスト