6R-02

ベイズ最適化に基づく最適パラメータ探索法の研究動向と課題に関する一考察

良川　太河† 　小林　学† 　後藤　正幸†
早稲田大学†

1. まえがき

ベイズ最適化は，最適化を行いたい目的関数が陽に与えられず，実験的に結果が得られるような場合に，ガウス過程回帰によって実験データ点から関数形を学習しつつ，その関数の大域的最適解の探索を行う手法である．この手法は，実験結果を反映した効果的な実験候補の探索や，機械学習モデルのハイパーパラメータの決定などで効果を挙げている．一般にベイズ最適化は計算量が大きくなるという実用上の問題点があるとされているが，計算技術の発展や近似解法による計算量の削減に伴い，近年様々な問題設定への拡張が行われている．ベイズ最適化の研究領域は，今後も様々な方向へ発展を遂げると共に，ビジネス領域の施策最適化といった応用分野に対しても影響を与える可能性がある．本稿では，動的な実験計画法，並びにパラメータ最適化におけるベイズ最適化の問題設定と最近の研究動向をまとめ，今後の研究課題について考察を与え，展望を述べる．

2. 事前準備

2.1 ガウス過程回帰[1]

ガウス過程回帰は，パラメータなどの入力$\boldsymbol{x}$と，推定精度などの出力yの間にある関数$y = f(\boldsymbol{x})$を推定する手法の一つである．ガウス過程回帰の最も重要な特徴として，入力の集合$\boldsymbol{X}^n = \{\boldsymbol{x}_i\}_{i=1}^n$が与えられたとき，対応する出力$\boldsymbol{Y}^n = \{y_i\}_{i=1}^n$が多変量ガウス分布に従うという点が挙げられる．

ガウス過程回帰の例を以下の図 1 に示す．

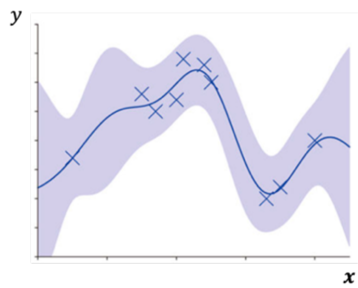

図 1：ガウス過程回帰の例

ガウス過程回帰による関数では，入力$\boldsymbol{X}^n$に対する出力$\boldsymbol{Y}^n$が多変量ガウス分布に従う形で出力されるため，上記の図 1 のように出力yの推定幅が算出される．ベイズ最適化ではこの推定幅を用いて次の入力の候補点を算出する．

A Study on Research Trends and Issues in Optimal Parameter Search Methods Based on Bayesian Optimization
† Waseda University

2.2 ベイズ最適化[2]

ベイズ最適化は，ガウス過程回帰により未知の関数をデータから学習しつつ，少ない試行回数でその関数の大域的最適解の推定を行う手法である．ベイズ最適化では，ガウス過程回帰によって得られた関数に対し，獲得関数と呼ばれる関数を用いて大域的最適解の推定において最適な入力を算出する．獲得関数としては様々な関数が提案されている[2]が，主に用いられている関数としては，新しい入力による改善可能性と改善量を考慮した Expected improvement などが挙げられる．

3. 最近の研究動向

ベイズ最適化は前述の通り，情報の少ない未知の関数の大域的最適解を探索したい場合に有効な手法である．この有効性から，近年様々な発展研究が行われている．ベイズ最適化の発展領域としては，計算量の削減と，問題設定の拡張の 2 つに分けられる．これらについて，本節で詳しく説明する．

3.1 計算量の削減

ベイズ最適化はガウス過程回帰を用いて出力yの事後分布を求めており，その計算過程ではカーネル行列やその逆行列を求める必要がある．しかし，この計算には$O(N^3)$の計算量がかかるため，データの次元数Nが大きい場合，計算量が膨大になってしまうという問題点が挙げられる．

そこで従来より，補助変数法[1]と呼ばれる手法を用いることで，ガウス過程回帰における計算量の削減を行う試みがなされている．補助変数法では，直接観測されない次元数$M(\ll N)$の変数を用いて元の行列をゼロ要素が多い疎行列で近似する．これにより，従来より大幅に計算量が少ない$O(M^3)$まで計算量を減らすことができる．他にも，フーリエ特徴を用いて高速な近似を行う手法も提案されている[3]が，予測分布の精度が低いことが欠点となっている．これらの手法は問題設定ごとに有効な手法が異なり，お互いに補完性のある手法となっているため，近年提案された Efficiently Sampling[4]では，事後分布を事前分布と更新分布の和に分解し，上記の 2 つの手法を上手く組み合わせることで，高速に推定精度の良い事後分布を求めている．

3.2 問題設定の拡張

ベイズ最適化は，未知の関数の大域的最適解を探索する手法であり，少ないデータからも不確実性を

考慮して探索を行うことができるため，大域的最適解の探索のみならず，様々な問題設定への拡張が近年研究されている．以下に，主な発展研究の例と，その概要を示す．

Conditioning[5]では，パラメータ最適化におけるハイパーパラメータとデータセットなどの，最適化される変数が複数存在する場合の最適化方法を定式化している．Knowing What[6]では，達成したい推定精度などの，目標とする出力値が予め決まっている場合に，この事前情報を活用して最適な入力を求める方法を定式化している．SafeOpt[7]では，一定の閾値を下回ることがない入力の領域を安全領域と定義し，安全領域内での探索を行うことで，一定の出力を保証した探索を提案している．一方，InfoBAX[8]では，入力に対して任意のアルゴリズムを実行した後の出力をガウス過程回帰でモデル化することにより，変曲点や Top-k 探索などの様々な問題設定への拡張を可能にしている．

4. 施策最適化問題におけるベイズ最適化の活用

本章では，施策最適化問題におけるベイズ最適化の可能性を考察する．

初めに，ベイズ最適化の定式化を行う．ベイズ最適化では，目的関数$f(\boldsymbol{x})$と最適化したい入力$\boldsymbol{x}(\in \mathbb{R}^D)$が定義された場合に，式(1)を満たす大域的最適解$\boldsymbol{x}^*$を探索することを目的としている．

$$\boldsymbol{x}^* = \arg\max_{x \in \mathbb{R}^D} f(\boldsymbol{x}) \tag{1}$$

ベイズ最適化では，この$\boldsymbol{x}^*$を推定するために，ガウス過程回帰により推定した関数$\hat{f}(\boldsymbol{x})$を用いて，$\boldsymbol{x}^*$に近い推定値である$\hat{\boldsymbol{x}}$を算出する．このとき，与えられた学習データで構築したモデルが，信頼できる性能を発揮できるデータ領域を Applicability Domain（以下，$\mathrm{AD}(\boldsymbol{X}^n, \boldsymbol{Y}^n)$）と表す[9]と，この条件は以下の式(2)で表すことができる．

$$\hat{\boldsymbol{x}} = \underset{x \in \mathrm{AD}(\boldsymbol{X}^n, \boldsymbol{Y}^n)}{\operatorname{argmax}} \hat{f}(\boldsymbol{x}) \tag{2}$$

ベイズ最適化は，2.2 節で挙げた獲得関数を適切に設定し，学習データの追加に用いることで$\boldsymbol{x}^*$が存在する可能性が高い領域のデータを探索し，$\mathrm{AD}(\boldsymbol{X}^n, \boldsymbol{Y}^n)$の範囲を広げていく手法であると考えられる．

一方，施策最適化にベイズ最適化を活用する場合には，目的関数$f(\boldsymbol{x})$と最適化したい入力$\boldsymbol{x}$を現実の施策の目的や内容に合わせ，適切に設定する必要がある．ただし，これらの問題設定で扱うデータは高次元であることが多く，そのまま適用してしまうと$\boldsymbol{x}$の空間$\mathbb{R}^D$が非常に高次元となる．この場合，$\hat{f}(\boldsymbol{x})$を求める際の計算が不可能なレベルになる上に，高次元空間上から$\mathrm{AD}(\boldsymbol{X}^n, \boldsymbol{Y}^n)$を効率良く広げるデータを探索することが難しくなってしまう．そこで，3 章で紹介した各手法では，このような問題に対して計算量の削減や，効率良く$\mathrm{AD}(\boldsymbol{X}^n, \boldsymbol{Y}^n)$を広げるデータの探索法を提案していると考えることができ，施策最適化における適用可能性を高めている．

ただし，未だに施策最適化における問題設定をベイズ最適化の枠組みに落とし込むためのフレームワークは存在していない．そのため，今後の検討が必要であると考えられる．

5. まとめ

本論文では，実験計画法・パラメータ最適化におけるベイズ最適化の問題設定と最近の研究動向をまとめ，今後の研究課題を示した．ベイズ最適化は実験計画法やパラメータ最適化のみならず，ビジネスの場における施策最適化問題といった発展事例に対しても有効な手法である可能性が考えられる．しかし，これらの問題設定をそのままベイズ最適化に適用することは難しい．そこで，近年これらの適用に関する様々な手法が提案されているが，未だに施策最適化における問題設定をベイズ最適化の枠組みに適用するためのフレームワークは存在していない．施策最適化は重要な課題であるため，今後の研究が期待される．

謝辞

本研究の一部は，日本学術振興会科学研究費 No.21H04600,No.19K04914 の助成を受けたものです．

参考文献

[1] 持橋大地，大羽成征．(2019)．ガウス過程と機械学習． 講談社.

[2] Frazier, P. I. (2018). A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811*.

[3] Rahimi, A., & Recht, B. (2007, December). Random Features for Large-Scale Kernel Machines. In *NIPS* (Vol. 3, No. 4, p. 5).

[4] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. Journal of Global optimization, Vol. 13, No. 4, pp. 455–492, 1998.

[5] Pearce, M., Klaise, J., & Groves, M. (2020). Practical Bayesian Optimization of Objectives with Conditioning Variables. *arXiv preprint arXiv:2002.09996*.

[6] Nguyen, V., & Osborne, M. A. (2020, November). Knowing the what but not the where in Bayesian optimization. In *International Conference on Machine Learning* (pp. 7317-7326). PMLR.

[7] Sui, Y., Gotovos, A., Burdick, J., & Krause, A. (2015, June). Safe exploration for optimization with Gaussian processes. In *International Conference on Machine Learning* (pp. 997-1005). PMLR.

[8] Neiswanger, W., Wang, K. A., & Ermon, S. (2021). Bayesian Algorithm Execution: Estimating Computable Properties of Black-box Functions Using Mutual Information. *arXiv preprint arXiv:2104.09460*.

[9] 金子弘昌．(2021)．Python で学ぶ実験計画法入門 ベイズ最適化によるデータ解析．講談社.