

ベイズ最適化と反応障壁計算を組み合わせた合成実験数削減法の開発: クライゼン転位を対象とした数値検証

¹北大院総化, ²北大院理, ³WPI-ICReDD, ⁴JST-ERATO

○岡田 拓明¹, 前田 理^{2,3,4}

【序】

反応開発において、反応条件の最適化は重要である。化学者は、自身の勘や経験を元にスクリーニングを行い、最適な反応条件を見出してきた。しかし、試行錯誤に基づく実験検証は多くのコストを必要とするため、最小の実験回数で反応条件を最適化できる手法が必要である。そこで本研究では、ベイズ最適化[1]と反応障壁計算を組み合わせた合成実験数削減法を考案した。本手法では、ベイズ最適化が提案した系における反応性を障壁計算で予測することで、必要な実験回数を削減する。しかし、障壁計算を行うことにより、具体的にどれだけ実験回数を減らせるかは明らかでない。そこで本研究では、10万分子のクライゼン転位における理論的反応障壁のデータセット[2]を用いた数値検証を行い、本手法の性能評価を行った。

【方法】

ベイズ最適化の実装では、ガウス過程をサロゲートモデルとし、カーネル関数にはRBFカーネルを、獲得関数にはEI関数を用いた。ガウス過程の実装にあたっては、GPyTorchライブラリ[3]を使用した。また、性能評価において実験で反応性を調べる操作は、対応する反応障壁をデータセットから取得することで実行した。一方、計算で反応性を調べる操作は、データセットの障壁にノイズを加えることで、計算誤差の影響を仮想的に考慮した。

【結果・考察】

ある実験回数において、反応障壁の小ささが上位10個の分子のいずれかを本手法が発見できた確率を調べた。その結果、いずれの検討条件化でも、本手法は通常のベイズ最適化よりも高い性能を示し、計算結果を利用して実験回数を削減できることがわかった。続いて、計算のバッチサイズ: m ごとに本手法の性能を比較した。ここでは、 m が5、10および15の場合の結果を示す。計算誤差として加えたノイズの大きさに対し、必要となった実験回数をFigure 1に示した。その結果、いずれの m においても、計算誤差が大きくなるにつれて必要な実験回数が増える傾向が確認された。詳細な結果や解析については当日報告する。

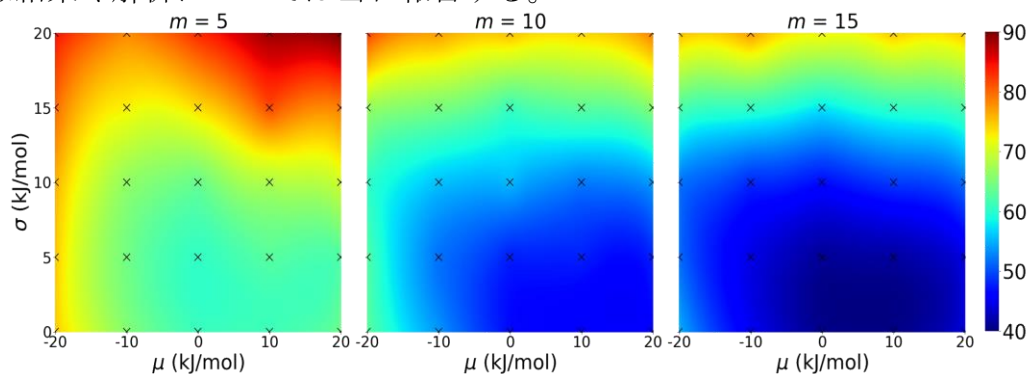


Figure 2. 計算誤差の大きさに対して必要となる実験回数

【参考文献】

- [1] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. D. Freitas, *Proc. IEEE* **2016**, 104, 148.
- [2] H. Okada, S. Maeda, *Mol. Inf.* **2021**, 41, e2100216.
- [3] J. R. Gardner, G. Pleiss, D. Bindel, K. Q. Weinberger, A. G. Wilson, *Adv. Neural Inf. Process. Syst.* **2018**, 31, 7587.