再生核ヒルベルト空間とベイズ最適化

Amuta (Twitter: @Amuta151)

Abstract

未知の関数の最小化は重要な問題である．たとえば機械学習におけるパラメータチューニングや事業におけるコストの最小化などがあげられる．ベイズ最適化は最小化問題の方法論の一つであり, 実用上有用なだけでなく関数解析に基く興味深い理論を持つことが知られている．

ベイズ最適化はガウス過程 (Gaussian Process : GP) を用いた最適化である．GP は無限次元のガウス分布, すなわちランダムに関数を返す確率変数と捉えることができる (図 1)．このイメージを多次元正規分布から類推してみよう．多次元正規分布の確率密度関数の概形は分散共分散行列 Σ によって決まる．特に Σ のスペクトルと固有ベクトルによって決定されることがわかる (図 2)．実は GP に関してこれと同様の性質が成り立つことが知られている．すなわち, 分散共分散行列に対応するカーネル関数 $k(s,t)$ ($s,t \in T, T \subset \mathbb{R}^d$: compact) をスペクトル分解することが可能 (Mercer の定理) である:

$$k(s,t) = \sum_{i=1}^{\infty} \lambda_i \phi_i(s)\phi_i(t).$$

そして固有関数 $\{\phi_i\}_{i\in\mathbb{N}}$ が生成する再生核ヒルベルト空間が GP の確率分布としての特徴を決定する．

これらの理論的背景について詳しく触れながらベイズ最適化について紹介する．

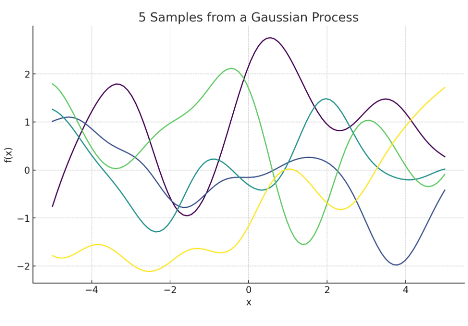

図 1: GP から発生させた関数 5 つ

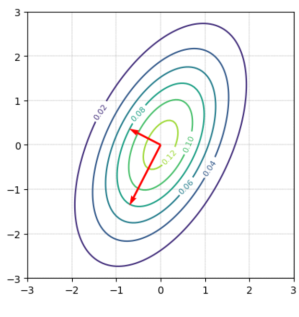

図 2: 二次元正規分布の等高線図

参考文献

[1] Ghosal, Subhashis, and Aad Van der Vaart. Fundamentals of nonparametric Bayesian inference. Vol. 44. Cambridge University Press, 2017.

[2] Rasmussen, Carl Edward, and Christopher KI Williams. Gaussian processes for machine learning. Vol. 1. Cambridge, MA: MIT press, 2006.

[3] 藤田宏, 黒田成俊, 伊藤清三. 「関数解析」. 岩波書店, 1991.