

確率場のハイパーパラメータ推定の高速化を目的とした次元削減モデリング Dimensionality reduction modeling for faster estimation of hyperparameters of random fields

○柴田達哉*, Koch Michel Conrad**, 藤澤和謙*

○Tatsuya Shibata, Michael Conrad Koch, Kazunori Fujisawa

1. はじめに

地盤・構造物の内部を推定する逆解析では、観測誤差や観測データ不足により、推定結果は不確実性を持つ。ベイズ推定に基づく逆解析では、推定結果は確率分布として得られるため、推定結果の不確かさを定量評価できる手法として注目されている。ベイズ推定では、逆問題の推定モデルを簡潔に記述するために、事前確率場をガウス確率場とし、Karhunen-Loève (KL) 展開を用いて低次元モデル化するアプローチが有効である。ガウス確率場を特徴づける2つのハイパーパラメータ（分散と相関長）は未知であるため、推定すべきだが、推定の過程でこれらのパラメータが変更される度に、KL 展開の再構築する必要があるため、計算コストが膨大になる。本研究では、この問題を打破できる手法として、理論解を活用した KL 展開を提案する。

2. 理論解を用いた Karhunen-Loève 展開

本研究では、ガウス確率場 $u(z): \mathcal{D} \rightarrow \mathbb{R}$ を扱う。共分散カーネルはガウシアンカーネル

$$C(z, z', \sigma, l) = \sigma^2 \exp\left(-\left(\frac{z - z'}{l}\right)^2\right) \quad (1)$$

とする。ここで、 σ^2 は分散、 l は相関長である。また、簡単のため1次元とする。

通常、KL 展開は、ガウス確率場 u の定義域 \mathcal{D} 上の積分固有値問題

$$\int_{\mathcal{D}} C(z, z', \sigma, l) \phi_i(z', l) dz' = \lambda_i(\sigma, l, s) \phi_i(z, l) \quad (2)$$

の固有値・固有関数 $\{\lambda_i, \phi_i\}_{i=1}^{\infty}$ を用いて定義される関数 $\psi_i(z, \sigma, l) = \sqrt{\lambda_i(\sigma, l)} \phi_i(z, l)$ によって

$$u(z) = \bar{u}(z) + \sum_{i=1}^{\infty} \psi_i(z, \sigma, l) \theta_i \quad (3)$$

と表される。ここで $\theta_i \sim \mathcal{N}(0, 1)$ である。式(3)の無限和の影響の大きい上位 M 項のみを採用し、

$$\hat{u}(z) = \bar{u}(z) + \sum_{i=1}^M \psi_i(z, \sigma, l) \theta_i \quad (4)$$

とすることで、ガウス確率場を M 次元ベクトル $\boldsymbol{\theta} = [\theta_1, \dots, \theta_M]^T$ で近似でき、次元が削減される (truncated KL 展開)。なお、式(4)は、 M 次元ベクトルを用いた近似の中で、平均二乗誤差を最小にする近似であることが知られている[1]。すなわち、以下を最小にする。

$$\varepsilon = \int_{\mathcal{D}} \text{Var}(u(z) - \hat{u}(z)) dz = 1 - \frac{1}{|\mathcal{D}|} \sum_{i=1}^M \int_{\mathcal{D}} \psi_i(z, \sigma, l)^2 dz \quad (5)$$

一般に、式(2)は行列固有値問題に離散化して数値的に解かれる。ハイパーパラメータ σ, l の推定では、 σ, l が変更される度に、KL 展開の再構築のために、この固有値問題を解き直す必要があるため、ハイパーパラメータ推定の計算負荷が大きくなるという問題がある。

一方、本研究では、領域 $[-\infty, \infty]$ 上で定義された以下の積分固有値問題の理論解を用いる。

$$\int_{-\infty}^{\infty} C(z, z', \sigma, l) \phi_i^*(z', l, s) d\mu(z', s) = \lambda_i^*(\sigma, l, s) \phi_i^*(z, l, s) \quad (6)$$

ここで、 $d\mu(z', s) = p(z', s) dz'$ は測度であり、 $p(z', s) = \mathcal{N}(z' | 0, s^2)$ である。 $p(z', s)$ は重みと見なせ、積分固有値問題における重みの有無 (式(2)と式(6)) は、行列固有値問題における固有値

*京都大学 (Kyoto University), **オーストラリア国立大学 (Australian National University) 数値解析, 逆解析

問題と一般化固有値問題の違いに対応する．固有値と固有関数の理論解は以下のである[2]．

$$\lambda_i^*(\sigma, l, s) = \lambda_i^*(\sigma, \gamma(l, s)) = \frac{2\sigma^2(\gamma - 1)^{l-1}}{(\gamma + 1)^i} \quad (7)$$

$$\phi_i^*(z, l, s) = \phi_i^*(z, \gamma(l, s)) = \frac{\sqrt[4]{\gamma}}{\sqrt{(i-1)!}} \exp\left(-\frac{\gamma-1}{4} \frac{z^2}{s^2}\right) \text{He}_{i-1}\left(\sqrt{\gamma} \frac{z}{s}\right) \quad (8)$$

ただし, $\gamma(l, s) = \sqrt{1 + 8s^2/l^2}$, $\text{He}_i(\cdot)$ は確率論者のエルミート多項式である．これらの固有値・固有関数の理論解を用いて, $\psi_i(z, \sigma, l) = \sqrt{\lambda_i^*(\sigma, l, s)} \phi_i^*(z, l, s)$ とした場合でも, \mathcal{D} 上のガウス確率場 u は, 式(3)の無限和の形に KL 展開できる(証明は紙面の都合上割愛)．これらの理論解を用いた KL 展開では, σ, l の更新に伴う固有値問題の解き直しを回避でき, 計算コストが大幅に削減される．ただし, 理論解を用いた truncated KL 展開による近似は, 平均二乗誤差を最小にする近似ではないため, 所望の精度を得るために必要な KL 展開の項数は, 通常の KL 展開, すなわち, 式(2)の固有値・固有関数を用いた KL 展開よりも大きくなることに注意する．

3. 数値計算

理論解を用いた KL 展開による近似の精度を確認するために, 通常の KL 展開との比較を行った．ガウス確率場の定義域

$$\mathcal{D} = \{z: -1 \leq z \leq 1\}$$

とした．(定義域 \mathcal{D} は, ψ_i の拡大・縮小および平行移動により変更できるため, \mathcal{D} の選び方は本質ではない)．また, ガウス確率場のハイパーパラメータは $(\sigma, l) = (1, 0.2), (1, 0.4), (1, 1.0)$ の3組とした．通常の KL 展開では, 固有値・固有関数は Nyström 法によって数値的に解いた．理論解を用いた KL 展開では, 代表値として $s = 0.3, 0.5$ を選んだ．図 1 にそれぞれの KL 展開の誤差 ε (式7)を示す．いずれの場合も, 同じ項数の truncated KL 展開による誤差 ε は通常の KL 展開(黒線)で最小となり, 平均二乗誤差を最小にする性質を反映している．また, 理論解を用いた KL 展開では, s の選び方により誤差 ε の挙動が変わる．これは, できるだけ少ない項数 M で所望の精度を満たすには, ハイパーパラメータの値および許容する誤差の大きさに対して, 適切な s を選ぶ必要があることを意味する．例えば, $(\sigma, l) = (1, 0.2)$ のとき, $\varepsilon \leq 10^{-3}$ を満足するには, $s = 0.3$ だと $M = 19$, $s = 0.5$ だと $M = 23$ となる(図 1 左)ため, $s = 0.3$ を選ぶべきである．なお, $(\sigma, l) = (1, 0.4), (1, 1.0)$ のグラフから, $l \geq 0.4$ では, 適切な s を選べば, 通常の KL 展開とほとんど同じ項数で所望の精度を満たせることが確認できる．

4. 結言

本研究では, ガウスカーネルを持つガウス確率場を, 理論解を用いた KL 展開でモデル化する手法を提案した．この KL 展開は, 通常の KL 展開とは異なり, 平均二乗誤差を最小にする性質を持たないが, 数値計算による検証により, 重み $p(z', s)$ に関するパラメータ s を適切に選択すれば, 通常の KL 展開と大差ない項数で所望の精度を満たせることが確認された．なお, 本研究では, 単純な 1 次元領域 $\mathcal{D} = \{z | -1 \leq z \leq 1\}$ を扱ったが, 高次元の任意領域への拡張は容易である[3,4]ことに言及しておく．今後は, ハイパーパラメータ σ, l の推定を行い, 計算時間の改善について検証する予定である．

参考文献 [1] Ghanem, Roger G., and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Courier Corporation, 2003. [2] Williams, Christopher KI, and Carl Edward Rasmussen. *Gaussian processes for machine learning*. Vol. 2. No. 3. Cambridge, MA: MIT press, 2006. [3] Pranesh, Srikara, and Debraj Ghosh. "Faster computation of the Karhunen–Loève expansion using its domain independence property." *Computer Methods in Applied Mechanics and Engineering* 285 (2015): 125-145. [4] Basmaji, A. A., et al. "Karhunen–Loève expansion based on an analytical solution over a bounding box domain." *Probabilistic Engineering Mechanics* 74 (2023): 103519.