

グリーン関数理論に基づく共分散行列のガウス過程回帰への応用

永井 朋子

工学院大学 学習支援センター

1. はじめに

機械学習におけるカーネル法[1]は、非線形データ解析の手法である。近年カーネル法と再生核ヒルベルト空間と呼ばれる関数の集合との関連が指摘され、線形微分方程式の初期値問題における基本解(再生核)との対応が指摘された[2]。一方微分方程式の境界値問題のグリーン関数は外力に対する応答関数として物理・工学において重要で、適切なヒルベルト空間に対して任意の関数を再生する再生核であることが指摘されている[3]。本研究では、機械学習⇔再生核⇔グリーン関数の繋がりに注目し、2階常微分方程式のグリーン関数をガウス過程回帰[4,5]の新しいカーネル関数として提案する。

2. グリーン関数

本研究では、最も基本的、かつ弦のたわみなど工学的に重要な2階常微分方程式の境界値問題を考える。

$$\begin{cases} -\frac{d^2u}{dx^2} + a^2u = w(x) & (0 < x < 1) \\ u(0) = u(1) = 0 \end{cases} \quad (1)$$

ここで a は0以上の定数、 $w(x)$ は外力項である。解の要となるグリーン関数 $G(x, y)$ は $w(x)$ に対する解 $u(x) = \int_0^1 G(x, y) w(y) dy$ を生成し、工学の分野においてはインパルス応答関数と呼ばれ、次のように表される。

$$G(x, y) = G(a; x, y) = \begin{cases} \frac{\sinh(a \min(x, y)) \sinh(a(1 - \max(x, y)))}{a \sinh a} & (a > 0) \\ \min(x, y)(1 - \max(x, y)) & (a = 0) \end{cases} \quad (2)$$

また、 $a > 0$ の L_1 ノルムは次のようになる。

$$\begin{aligned} L_1 &= L_1(y) = \int_0^1 |G(x, y)| dx = \int_0^1 G(x, y) dx \\ &= \frac{1}{a^2} \left(1 - \cosh(ay) + \frac{(\cosh a - 1) \sinh(ay)}{\sinh a} \right) \end{aligned} \quad (3)$$

本研究では L_1 ノルムで規格化されたグリーン関数

$$\tilde{G}(x, y) = \frac{G(x, y)}{L_1(y)} \quad (0 < y < 1) \quad (4)$$

をガウス過程回帰[4,5]の共分散行列の各成分を構成するカーネル関数として提案する。

3. グリーン関数を用いたガウス過程回帰

ガウス過程[4,5]では、入力変数 $X_N = \{x_1, x_2, \dots, x_N\} (x_i \in \mathbf{R})$ に対する関数値 $\mathbf{f}_N = (f(x_1), f(x_2), \dots, f(x_N))^T (f(x_i) \in \mathbf{R})$ の分布は、平均関数 $\mu(x) = 0$ 、 $N \times N$ の共分散行列 V_{NN} のガウス分布、 $\mathbf{f}_N \sim N(\mathbf{0}, V_{NN})$ でモデル化することができる。共分散行列の各成分はカーネル法のカーネル関数と同じもので、本研究では V_{NN} のカーネル関数として式(4)の規格化されたグリーン関数を提案する。従って V_{NN} の (i, j) 成分は $(V_{NN})_{ij} = \tilde{G}(x_i, x_j)$ となる。更に関数値 $f(x_i)$ に $\varepsilon \sim N(0, \rho)$ のガウスノイズが加わって、観測値 $y_i = f(x_i) + \varepsilon$ は、 $\mathbf{y} \sim N(\mathbf{0}, V_{NN} + \rho I_N)$ と表される。ここで、問題： N 個の観測データ D

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \quad (5)$$

が与えられたときの、新しい M 個の入力変数

$X_M^* = \{x_1^*, x_2^*, \dots, x_M^*\}$ に対する出力 $\mathbf{y}^* = (y_1^*, y_2^*, \dots, y_M^*)^T$ の予測分布を求める。

このために、 X_M^* に対する関数値 $\mathbf{f}_M = (f(x_1^*), f(x_2^*), \dots, f(x_M^*))^T$ の事後分布を求め、ガウスノイズを加えて $\mathbf{y}^*$ の分布を求める。まず $\mathbf{y}$ と $\mathbf{f}_M$ の同時分布は次のようになる。

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{f}_M \end{pmatrix} \sim N \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} H_{NN} + \rho I_N & H_{NM} \\ H_{MN} & H_{MM} \end{pmatrix} \right) \quad (6)$$

$$(H_{NN})_{ij} = \tilde{G}(x_i, x_j) \quad (i, j = 1, 2, \dots, N) \quad (7)$$

The covariance matrix of Green functions and its application to Gaussian process regression

Tomoko Nagai, Kogakuin University

$$(H_{MM})_{ij} = \tilde{G}(x_i^*, x_j^*) (i, j = 1, 2, \dots, M) \quad (8)$$

$$(H_{NM})_{ij} = \tilde{G}(x_i, x_j^*) (i = 1, 2, \dots, N, j = 1, 2, \dots, M) \quad (9)$$

$$(H_{MN})_{ij} = \tilde{G}(x_i^*, x_j) (i = 1, 2, \dots, M, j = 1, 2, \dots, N) \quad (10)$$

ここで共分散行列 $H_{NN}, H_{MM}, H_{NM}, H_{MN}$ の (i, j) 成分は式(4)の規格化されたグリーン関数で構成されている。 $\mathbf{y}$ が観測されたときの $\mathbf{f}_M$ の条件付分布 [4] に、ガウスノイズを加えた $\mathbf{y}^*$ の予測分布は次のようなガウス分布に従う。

$$p(\mathbf{y}^* | X_M^*, D) = N(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) \quad (11)$$

$$\hat{\boldsymbol{\mu}} = H_{MN}(H_{NN} + \rho I_N)^{-1} \mathbf{y} \quad (12)$$

$$\hat{\boldsymbol{\Sigma}} = H_{MM} + \rho I_M - H_{MN}(H_{NN} + \rho I_N)^{-1} H_{NM} \quad (13)$$

式(13)の共分散行列 $\hat{\boldsymbol{\Sigma}}$ の対角成分は分散ベクトル $\mathbf{V} = (V_1, V_2, \dots, V_M)^T$ で、その i 成分 ($i = 1, 2, \dots, M$) は次のようになる。

$$V_i = \tilde{G}(x_i^*, x_i^*) + \rho - \mathbf{h}_{MNI}^T (H_{NN} + \rho I_N)^{-1} \mathbf{h}_{NMi} \quad (14)$$

$$\mathbf{h}_{MNI}^T = (\tilde{G}(x_i^*, x_1), \tilde{G}(x_i^*, x_2), \dots, \tilde{G}(x_i^*, x_N)) \quad (15)$$

$$\mathbf{h}_{NMi} = (\tilde{G}(x_1, x_i^*), \tilde{G}(x_2, x_i^*), \dots, \tilde{G}(x_N, x_i^*))^T \quad (16)$$

$\mathbf{y}^*$ の予測分布は、定数 a とノイズ分散 ρ をハイパーパラメータとして対数尤度最大化 [4, 5] により得られる。本研究の対数尤度 l は次のようになる。

$$l = -\frac{1}{2} \mathbf{y}^T (H_{NN} + \rho I_N)^{-1} \mathbf{y} - \frac{1}{2} \log \det (H_{NN} + \rho I_N) - \frac{N}{2} \log 2\pi \quad (17)$$

4. 数値実験

数値実験は $N = 15$, $M = 99$ とし、 $x_i = 0.1 + \frac{0.8(i-1)}{14} (i = 1, 2, \dots, 15)$, $x_j^* = 0.01j (j = 1, 2, \dots, 99)$ 、次のカスプのある観測データ D_1, D_2 で行った。

$$D_1: y_i = 7(1 - \exp(-5|x_i - 0.5|)) \quad (18)$$

$$D_2: y_i = 7(1 - \exp(-5|x_i - 0.5|)) + \text{randomness} \quad (19)$$

グリーン関数カーネルの場合の、 D_2 に対する $\mathbf{y}^*$ の予測分布と対応する共分散行列を図 1 に示す。

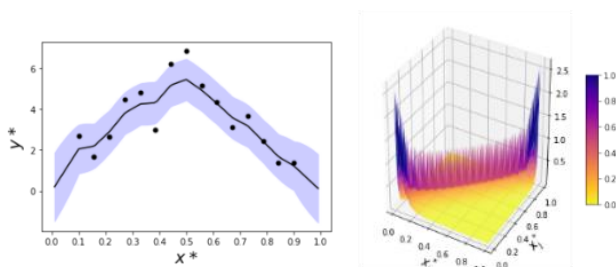


図 1. グリーン関数の場合の予測分布(左)と共分散行列(右)。

図 1(左)予測分布の実線が平均、影の部分は平均 ± 標準偏差を表す。(右)共分散行列の対角線上の値は分散を表す。

更にカーネル関数として規格化されたガウス関数を用い、 $(V_{NN})_{ij} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - x_j)^2}{2\sigma^2}\right)$ の共分散行列における、定数 σ とノイズ分散 ρ をハイパーパラメータとした計算と比較した。対数尤度 l の比較を表 1 に示す。表 1 より、グリーン関数の対数尤度 l は、広く利用されるガウス関数の場合 [4, 5] より大きい、同程度である。

data	グリーン関数	ガウス関数
D_1	$l = -24.79$ $a = 2.88 (\rho = 0 \text{ fixed})$	$l = -26.35$ $\sigma = 0.07 (\rho = 0 \text{ fixed})$
D_1	$l = -24.79$ $(a, \rho) = (2.88, 0.0)$	$l = -21.64$ $(\sigma, \rho) = (0.12, 0.1)$
D_2	$l = -30.99$ $a = 5.61 (\rho = 0 \text{ fixed})$	$l = -32.52$ $\sigma = 0.061 (\rho = 0 \text{ fixed})$
D_3	$l = -30.00$ $(a, \rho) = (0.01, 0.8)$	$l = -30.29$ $(\sigma, \rho) = (0.20, 1.0)$

表 1. 対数尤度 l の比較。

5. まとめ

2 階常微分方程式の境界値問題のグリーン関数を規格化し、ガウス過程回帰の新しいカーネル関数として提案した。ベイズ推定に基づき、観測データに対する条件付き分布として予測分布を求めた。カスプのある場合には、対数尤度は広く利用されるガウスカーネルの場合より、大きなあるいは同程度の値を示し、よい回帰結果が得られた。

参考文献

- [1] B. Schölkopf and A. J. Smola, *Learning with kernels* (The MIT Press, 2002).
- [2] 福水健次, カーネル法入門ー 正定値カーネルによるデータ解析 (朝倉書店, 2010).
- [3] Y. Kametaka, K. Watanabe and A. Nagai, Proc. Japan. Acad. **81** Ser. A, 57 (2005).
- [4] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
- [5] 持橋大地・大羽成征, ガウス過程と機械学習 (講談社, 2019).