

ガウス過程回帰を用いた生体時系列データのモデル化

Gaussian Process Regression Model for Physiological Time Series Data

幸島 明男 車谷 浩一
Akio Sashima Koichi Kurumatani

産業技術総合研究所 人間情報研究部門

Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology

Gaussian process regression is a flexible analysis framework for modeling nonlinear time-series data based on a stochastic process. In this paper, we apply the framework to model physiological signals which are obtained by a mobile physiological sensor. First, we describe an outline of Gaussian process regression. Then, we propose a framework based on the Gaussian process regression for modeling the physiological sensing data. We also show experimental results to apply the framework to modeling electrocardiogram.

1. はじめに

近年、スマートフォンと連携するウェアラブルデバイスを用いた健康管理・情報サービスが注目を集めている。腕時計型の端末や被服装着型の小型端末を使用して自らの活動状態を計測・記録し、その計測データに基づいて普段の活動量や活動パターンを継続的に把握することで、利用者の健康の維持に役立てるサービスが提供されるようになってきている。

我々は、これまで、高齢者の活動状況や生体信号を遠隔地から見守る遠隔生体見守りサービスとそのサービスのためのモバイル生体センサ(図1参照)の開発を行ってきた[幸島 11][幸島 13]。このような見守りサービスでは、活動や生体情報のセンシングを常時継続的にを行い、その結果をオンラインでモニタリングする機能が重視される。そのため、省電力化のための通信量の削減や、体動・通信状況等に起因するノイズの除去など、目的に応じて適切かつ効率的に生体時系列データを処理する手法の開発が必要とされている。

生体時系列データの処理に関しては、これまで、FFT や適応フィルタなど、一般的な信号処理のフレームワークにより処理する手法の研究が多かった。しかしながら、心電位などの生体時系列データには固有の時系列パターンが存在する。この時系列パターンに関する情報を用いて、データをモデル化できれば、より強力な時系列データ処理が可能になる。

本研究では、生体時系列データに固有の時系列パターンをモデル化するための一手法として、ガウス過程回帰 (Gaussian Process Regression)[Rasmussen 05][Bishop 06] によるモデル化を提案する。ガウス過程回帰は、時系列データの各値の同時分布がガウス分布にしたがうものと仮定することで、非線形時系列データの回帰モデルを構築する手法である。本論文では、ガウス過程回帰を用いて、生体センサから得られた生体時系列データについて回帰モデルを構築する手法を提案する。そして、提案手法を実際の心電位データに適用した結果について考察する。

2. ガウス過程回帰によるモデル化

生体センシングデータのモデル化に関して、近年、時系列データの生成モデルを仮定し、そのモデルに基づいてデー



図 1: モバイル生体センサ

タの処理を行うというアプローチの研究が提案されている。[McSharry 03][Clifford 05] は、心電位データを時間遅れのあるガウス分布型の波形の重ね合わせとしてモデル化することで、フィルタリングや圧縮などの処理を行う手法を提案している。この生成モデルを用いて、[Nabar 11] は、スマートフォン向けの心電位データ処理を提案している。また、[Sameni 07] は、このような生成モデルと非線形ベイズフィルターを組み合わせることで、心電位データのノイズ除去を可能にする手法を提案している。

生成モデルを事前知識として用いる時系列データの非線形解析フレームワークは、事前知識を利用しない手法に比べてより強力な解析が可能であるが、一方で、あらかじめ設定したモデルから外れた入力データの場合には、解析処理が困難になる恐れがある。

そこで、本研究では、入力データから適応的にモデルを学習する手法として、ガウス過程回帰による手法について検討する。以下では、ガウス過程回帰の概要について述べる[Rasmussen 05][Bishop 06]。

一般に時系列データの回帰処理は、入力 (時刻) $t_1, t_2, \dots, t_N$ における N 個の出力値: $y_1, y_2, \dots, y_N$ が与えられたとき、任意の入力 t_k に対する回帰モデル $y_i = f(t_k)$ を推定する処理である。ガウス過程回帰では、時系列データ $y = y_1, y_2, \dots, y_n$ からなる N 次元のベクトル y の同時分布がガウス分布にしたがうものと仮定する。そして、この同時分布を与えるようなガウス過程は、平均 $E[y]$ 、共分散 $Cov[y]$ によって記述できるので、この平均と共分散を推定することで、与えられた時系列

データを生成する確率モデルを構築する。

ガウス過程回帰において、平均 $E[y]$ と共分散 $Cov[y]$ を、それぞれ $E[y] = 0, Cov[y] = K$ とすると、共分散行列 K の各要素 $K_{i,j}$ はカーネル関数 $k(t_i, t_j)$ を要素とするグラム行列となる。 y に対して観測ノイズの存在を仮定すると、目的とする回帰モデルは、このグラム行列を構成するカーネル関数 $k(t_i, t_j)$ と y の観測ノイズ $\beta^{-1}\delta_{i,j}$ との和を要素とする共分散行列 C_N により計算できる。 $\delta_{i,j}$ は $i = j$ のとき 1 となるデルタ関数である。

新たな入力 t_{N+1} に対する出力値 $\hat{y}_{N+1}$ を推定する場合は、入力 t_{N+1} におけるガウス分布の平均を求めることになり、入力 t_{N+1} と各 $t_i (i = 1, \dots, N)$ によるカーネル関数 $k(t_i, t_{N+1})$ の値を要素とする長さ N のベクトル k および時系列データベクトル y とにもとづいて、

$$\hat{y}_{N+1} = k^T C_N^{-1} y$$

を計算することで求められる。同様に、 t_{N+1} における共分散は

$$\hat{\sigma}_{N+1}^2 = k(t_{N+1}, t_{N+1}) + \beta^{-1} - k^T C_N^{-1} k$$

となる。

ガウス過程回帰のカーネル関数 $k(t_i, t_j)$ を設定する際に、対象とする時系列データに適したハイパーパラメータ θ を設定することで、より良いモデル化が可能になる。与えられた入力データから、モデル化に適したハイパーパラメータを推定するには、対数尤度

$$\ln p(y|\theta) = -\frac{1}{2} \ln |C_N| - \frac{1}{2} y^T C_N^{-1} y - \frac{N}{2} \ln 2\pi$$

を最大化するようなハイパーパラメータを勾配降下法などにより推定することによって行う [Bishop 06]。

3. 生体時系列データのガウス過程回帰

3.1 生体時系列データ

実際の生体センシングデータに対してガウス過程回帰を適用し、得られた確率モデルを用いて、時系列データの予測を行う実験を行った。

対象とする生体センシングデータは、我々が開発したモバイル生体センサを用いて取得を行った。(図 1 参照)。本モバイル生体センサは、小型のマルチセンサユニットであり、MPU、フラッシュメモリ、5 種類のセンサ（心電位、3 軸加速度、気圧、温度、湿度）、リチウムイオンバッテリー、無線通信モジュール等を本ユニット内に内蔵している。心電位をセンシングする場合は、電極付きのベルトやシール型電極などで、本体の計測部を利用者の胸部に装着して計測を行う。一方で、心電位の計測の必要が無い場合は、胸ポケットなどに入れて携帯することで、動作の計測が可能である。センシングしたデータは無線通信モジュールを用いて、スマートフォンやパーソナルコンピュータ (PC) へとリアルタイムで送信することができる。本研究では、モバイル生体センサと PC とを無線で接続し、通信により取得した心電位データについて実験を行った。

3.2 生体時系列データの正規化

生体時系列データをガウス過程回帰により学習するために、時系列データの正規化処理を行った。まずモバイル生体センサを用いて心電位の時系列データを取得した。時系列データは 100Hz のセンシングレートで得られたスカラー値で表現されている。このデータに関して基線の位置がゼロになるような調



図 2: ガウス過程回帰によるセンシングデータの推定

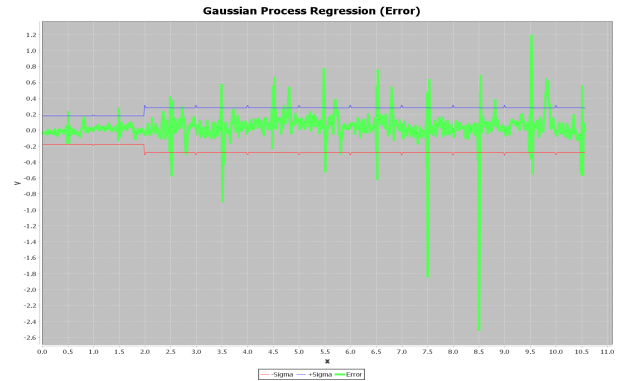


図 3: ガウス過程回帰による分散の推定と誤差

整と、心拍間隔のゆらぎの影響を排除するため、時系列データの R-R 間隔が 1.0 になるように時間軸のスケールについて正規化を行った。

この正規化した時系列データから、最初の 2 拍分の時系列データを学習データとして切り出し、その学習データについてガウス過程回帰による学習を行った。そして、この学習データから得られた回帰モデルを用いて、学習データ以降の時系列データについての予測を行い、実測値との誤差を計測した。

3.3 カーネル関数の設計とハイパーパラメータの推定

対象とする時系列データの周期性を考慮して、今回の実験では、以下のカーネル関数を用いて学習を行った。

$$k(t_1, t_2) = \theta_1^2 \exp\left(-\frac{1}{2} \frac{\sin\left(\frac{\pi|t_1 - t_2|}{\theta_2}\right)^2}{\theta_3^2}\right)$$

ハイパーパラメータ $\theta_i (i = 1, 2, 3)$ の設定は、勾配の方向に基づいてパラメータの最適化を行うアルゴリズムである RPROP アルゴリズム [Riedmiller 93][Blum 13] を用いて決定した。初期値をランダムに設定した後に、RPROP アルゴリズムによるパラメータ推定処理を、何度か繰り返すことにより、対数尤度を最大化するハイパーパラメータを決定した。

アルゴリズムの実装は、Java の行列演算ライブラリである JAMA^{*1} を用いて行った。擬逆行列の計算のためのコレスキー分解など、一部の処理については、例外処理を回避するために独自の実装を行っている。

*1 <http://math.nist.gov/javanumerics/jama/>

3.4 実験結果と考察

モバイル生体センサを用いて取得した心電位データについて、ガウス過程回帰により学習を行った。その実験の結果について述べる。

実験の対象とした時系列データの長さ $l = 1000$ である。そのうち、学習データの長さは2拍分で $n = 182$ であった。

ガウス過程回帰により学習したモデルを用いて生成した波形と実際の波形のグラフを図2として示す。また、ガウス過程回帰で推定された共分散および実測値と予測値の誤差をプロットしたグラフを図3として示す。プロットに用いたハイパーパラメータは $\theta = [0.75, 1.0, 0.05]$ 、観測ノイズは $\beta = 20$ とした。

図2から、学習の結果、実測値をほぼ近似する波形の生成が可能であることがわかる。周期性をもつカーネル関数を用いることで、生体時系列データを近似するパターンを生成できることがわかった。

図3の“+Sigma”と“-Sigma”はモデルから推定された共分散 σ^2 に基づく $\pm\sigma$ の大きさを表している。R波の部分で誤差が大きいのは、R波の頂点の位置の時間軸上のずれに起因するものと考えられる。今回の実験で用いた時系列データは、R波の時間軸上の位置が等間隔になるように正規化した時刻を入力値として学習を行った。この処理により、心電位データのパターンの周期性をガウス過程回帰によりモデル化できることが示されたが、一方で、R波の性質上、少しの時間軸上のずれであっても、誤差が大きくなることも示された。R波の高さ、R-R間隔、基線の位置など、実際の心電位データにはゆらぎがある。生体時系列データにおける規則的な周期性ではとらえられないゆらぎのモデル化と処理については、今後の課題である。

4. おわりに

本研究では、確率過程に基づく生体時系列データのモデル化の手法について提案した。そして、生体センサから得られた心電位データについての回帰モデルを構築した。周期性をもつカーネル関数を用いてガウス過程回帰を行うことで、生体時系列データのパターンを生成できることを示した。

今後は、本モデルを用いた、ノイズのフィルタリング処理や効率的な通信のための圧縮処理などについて検討を行う予定である。スマートフォンでの処理を念頭に、ガウス過程回帰の近似解法など、学習・推定処理の効率化についても検討をすすめたい。また、3軸加速度など心電位以外の生体時系列データの処理に対する、確率過程に基づくモデル化のフレームワークについて検討したいと考えている。

謝辞

本研究の一部は、JSPS 科研費 (26330125) の助成を受けた。

参考文献

- [Bishop 06] Bishop, C. M.: *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
- [Blum 13] Blum, M. and Riedmiller, M.: Optimization of Gaussian Process Hyperparameters using Rprop, in *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (2013)

- [Clifford 05] Clifford, G. D., Shoeb, A., McSharry, P. E., and Janz, B. A.: Model-based filtering, compression and classification of the ECG, in *Proceedings of Bioelectromagnetism and 5th International Symposium on Non-invasive Functional Source Imaging within the Human Brain and Heart (BEM & NFSI)*, Minnesota, USA (2005)
- [McSharry 03] McSharry, P., Clifford, G., Tarassenko, L., and Smith, L.: A dynamical model for generating synthetic electrocardiogram signals, *Biomedical Engineering, IEEE Transactions on*, Vol. 50, No. 3, pp. 289–294 (2003)
- [Nabar 11] Nabar, S., Banerjee, A., Gupta, S., and Pooven-dran, R.: GeM-REM: Generative Model-Driven Resource Efficient ECG Monitoring in Body Sensor Networks, in *Body Sensor Networks (BSN), 2011 International Conference on*, pp. 1–6 (2011)
- [Rasmussen 05] Rasmussen, C. E. and Williams, C. K. I.: *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press (2005)
- [Riedmiller 93] Riedmiller, M. and Braun, H.: A direct adaptive method for faster backpropagation learning: the RPROP algorithm, in *Neural Networks, 1993., IEEE International Conference on*, pp. 586–591 vol.1 (1993)
- [Sameni 07] Sameni, R., Shamsollahi, M., Jutten, C., and Clifford, G.: A Nonlinear Bayesian Filtering Framework for ECG Denoising, *Biomedical Engineering, IEEE Transactions on*, Vol. 54, No. 12, pp. 2172–2185 (2007)
- [幸島 11] 幸島 明男, 久我 幸史, 池田 剛, 栗原 正紀, 車谷 浩一: 携帯型生体センサを用いた動作・動線計測システムの構築, 情報処理学会研究報告. UBI, [ユビキタスコンピューティングシステム], Vol. 2011, No. 2, pp. 1–8 (2011)
- [幸島 13] 幸島 明男, 池田 剛, 山本 晃, 車谷 浩一: データサーバレス方式による遠隔生体情報モニタリングシステムの構築, IT ヘルスケア, Vol. 8, No. 1, pp. 54–57 (2013)