

多層パーセプトロンを用いた効率的な潜在空間推定への取り組み

理学専攻 情報科学コース

1640638

大山まりほ

1 はじめに

近年、ロボティクスの分野における人工知能分野のめまぐるしい発達が注目されており、人間と共存するロボットの需要が高まっている。人と共存するロボットは複数のモダリティの情報を使った直接的な教示から学び、人と同じような振る舞いを学習できる能力が必要とされる。このことを実現するために、本研究ではまず人の動作の認識に着目した。ロボットが人の動作を認識し、真似をするためには、ロボット自身の動作と人間の動作の対応関係を捉えねばならない。ロボットや人の動作は時系列データとして表すことができ、そのデータは非常に高次元且つ異なる次元となる。そのため本研究は、異なる高次元な複数の時系列データ同士の対応関係を効率よく取得する手法の開発、実世界において利用可能な技術を確認することを目的とする。複数の異なる次元の時系列データの対応を考える際、それらのデータを共有する空間上で同じ次元に揃えて比較可能にする必要がある。本研究では、複数の時系列データを対象に、効率的に共有の潜在空間に圧縮する手法の提案を行う。

2 提案手法概要

複数の時系列データを共有する潜在空間上に圧縮する。

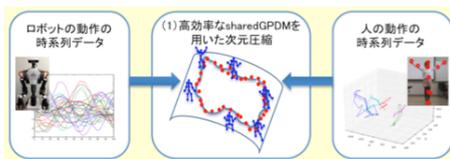


図 1: 提案手法の概要

図 1 に提案手法の概要を示す。圧縮には GPLVM を拡張した sharedGPLVM、GPDm を拡張した sharedG-PDM を用いる。

3 GPLVM

3.1 GPLVM の概要

本研究では、GPLVM[1] を用いてガウス過程で表現される高次元の時系列データを潜在空間に圧縮する。GPLVM とは、確率的成分分析にガウス過程を導入したものである。観測データ全体を Y 、圧縮先の潜在空間上のデータ全体を X とする。まず Principal component analysis(PCA) を用いて Y を圧縮し、 X の初期値を設定する。そして X, Y それぞれをガウス過程に従っていると仮定した正規化を行う。その後、以下の式 (1) で表される $p(Y|X, \alpha, \beta, \gamma)$ の対数尤度が最大となるよう式 (2) で表されるカーネル関数のパラメータ α, β, γ を更新する。更新したパラメータを当

てはめ、 X を再設定する。

$$p(Y|X, \alpha, \beta, \gamma) = \frac{1}{(2\pi)^{DN/2} |\mathbf{K}|^{D/2}} \times \exp\left(-\frac{1}{2} \text{tr}(\mathbf{K}^{-1} \mathbf{Y} \mathbf{Y}^T)\right) \quad (1)$$

$$k(x_n, x_m) = \alpha \exp\left(-\frac{\gamma}{2} (x_n - x_m)^T (x_n - x_m)\right) + \delta_{nm} \beta^{-1} \quad (2)$$

3.2 GPLVM with MLP

GPLVM では、カーネルのパラメータの更新を T 回繰り返すことで X が更新され、高次元の Y に対して低次元の潜在空間 X が求まる。しかし、このアルゴリズムでは EM アルゴリズムを採用しており、予め定めた回数 T 回 X の更新を繰り返しており、収束判定を行っていない。そのため最適な X が求められた保証がされていない。そのため収束判定を行うために MLP を組み込み、収束判定をするアルゴリズムを作成し、実験を行い、GPLVM との違いを検証した。

3.3 実験と考察

構築したアルゴリズムの正当性を検証するために、Ek[2] において sharedGPLVM の検証用に使用されていたのと同じサンプルデータを作成し、動作結果を比較した。3 式を基底としてガウス分布に従う重みをかけ足し合わせ、ノイズを付加し 20 次元の観測データ Y を作成する。20 次元の観測データ Y を 3 次元の潜在空間 X に圧縮する。GPLVM のみの場合と、GPLVM に MLP を組み込んだ場合とで実験を行ったところ、いずれも基底に近い曲線を得ることに成功した。それぞれの実行時間を比較したところ以下の表のようになった。

表 1: elapsed time

original GPLVM	GPLVM with MLP
5872.89[sec]	25.27[sec]

Lawrence[1] によって提案された元々のアルゴリズムでは、カーネル関数 K のパラメータを求める際の X の初期値および繰り返しの回数によって得られる X がガウス過程に従っているか定かではないケースが多く観測された。そこで本研究では、GitHub 上で公開されているソースコード 1 を参考に、対数尤度の微分値の収束を観測できる MLP を採用し、高次元の動作データを低次元に圧縮し良い結果を得ることができた。表 1 より、MLP を導入することで圧倒的に実行時間が短縮されたことがわかる。MLP の導入によって、精度の向上と計算量の削減を実現した。提案するアルゴリズムは精度の向上と実行時間削減を実現する。

4 sharedGPLVM

対応関係の取得を行うために2つの観測空間で1つの潜在空間を共有することで、2つの観測空間の関係を学習することの出来る sharedGPLVM[2] を用いる。sharedGPLVM の概要図を図3に示す。sharedGPLVM は大別して学習 (共有する潜在空間の構築) とマッピング (取得された潜在空間を通じた2つの時系列データの対応) の2つの段階に分けられる。本研究では学習フェーズに着目した。sharedGPLVM にも MLP を組み込み、精度の向上、効率化を図った。

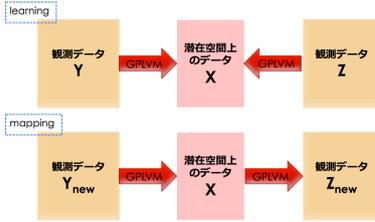


図 2: sharedGPLVM の概要図

4.1 実験と考察

3.3 節と同様のサンプルデータを使用し、多層パーセプトロンを組み込んだ sharedGPLVM を用いて2つの観測データから共有の3次元の潜在空間へ圧縮を行ったところ圧縮結果も基底と類似した形をしていることが分かった。ガウス過程に従う状態で圧縮が出来ており、理想的な学習結果が得られた。このことから、提案するアルゴリズムにより、実際に人の動作空間とロボットの動作空間に用いることで潜在空間を通して対応関係の学習が可能だと考えられる。

5 GPDM

GPLVM では潜在空間 X でのダイナミクスについては考慮されていないため、さらなる精度の向上を図り、潜在空間のダイナミクスを扱うことが出来る GPDM[3] を用いる。GPLVM や sharedGPLVM の際と同様に精度の向上のため MLP を組み込む。

5.1 実験と考察

3.3 節と同様のサンプルデータを使用し、オリジナルの GPDM と多層パーセプトロンを組み込んだ GPDM を用いて観測データから潜在空間への圧縮を行った。

表 2: elapsed time

original GPDM	GPDM with MLP
11262.35[sec]	17.42[sec]

GPDM を用いて高次元の時系列データを潜在空間に圧縮することが出来た。GPLVM との差異を視認することは難しいが、GPDM では潜在空間でのダイナミクスが考慮された形で圧縮がされている。以下の表2より、MLP を導入することで圧倒的に実行時間が短縮されたことがわかる。MLP の導入によって、精度の向上と計算量の削減を実現した。よって GPLVM の場合と同様に GPDM に MLP を導入することは有効であると考えられる。

6 sharedGPDM

識別を行うために、高次元の時系列データである2つの観測空間が1つの潜在空間を共有することで、2つの観測空間の関係を学習することができ、GPDM を基にした sharedGPDM[4] を用いる。sharedGPLVM の場合と同様に、sharedGPDM にも MLP を導入する。

6.1 実験と考察

2つの高次元な時系列データを sharedGPDM を用いて共有の低次元でありダイナミクスを考慮した潜在空間に圧縮することが出来た。

7 動作データへの適用

MLP を導入した sharedGPLVM と sharedGPDM の実データへの適用を試みた。Microsoft 社製 Kinect カメラを用いて人の関節の動きを取得し、15次元 (5関節 \times 3軸) のデータと、2関節分削除した9次元のデータとを共有する潜在空間への圧縮を行ったところ、共に2つの高次元な時系列データを用いて共有の潜在空間に圧縮することが出来た。実データへの適用をした場合でも観測データを2次元に圧縮した場合に滑らかな曲線が描かれていることを確認した。本手法の実データへの適用は有効だと考えられる。

8 おわりに

本研究は、ガウス過程に従う高次元の時系列データを対象にし、GPLVM を用いて低次元の潜在空間に圧縮する手法の開発を行った。また、従来の手法を効率と精度の2つの面で改良する手法として、EM アルゴリズムを使用する代わりに、非線形最適化手法である多層パーセプトロンを導入する手法を提案した。複数の時系列データの対応関係を取得することが出来る sharedGPLVM にも本手法を導入し、sharedGPLVM においても精度よく効率的な潜在空間の推定が出来ることを確認した。潜在空間上でのダイナミクスを考慮した圧縮が可能な GPDM や sharedGPDM にも本手法を導入し、実用性を調査した。本手法は、GPLVM や sharedGPLVM, GPDM, sharedGPDM のような代表的なガウス過程に基づく次元圧縮手法を精度、効率の面で向上させることに成功した。多層パーセプトロンを組み込んだ本手法を動作データに実際に適用し、潜在空間への圧縮を行い、実用性を確認した。今後の課題として、圧縮のさらなる精度の向上と、観測データの双方向への変換があげられる。

参考文献

- [1] Neil D. Lawrence, "Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data", NIPS, 2004.
- [2] Carl Henrik Ek., "Shared Gaussian Process Latent Variables Models", Ph.D thesis, Oxford Brookes University, 2009.
- [3] Wang, J. M., Fleet, D. J., Hertzmann, A. "Gaussian Process Dynamical Models for Human Motion", In IEEE Transactions on Pattern Recognition and Machine Intelligence, pp. 283-298., February, 2008.
- [4] Jixu Chen, Minyoung Kim, Yu Wang, Qiang Ji, "Switching Gaussian Process Dynamic Models for Simultaneous Composite Motion Tracking and Recognition", 2009 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2009.