ガウス過程による潜在変数モデル

Gaussian Process Latent Variable Model (GPLVM)

明治大学 理工学部 応用化学科
データ化学工学研究室　金子 弘昌

ガウス過程による潜在変数モデル [1,2] とは？

✓各特徴量 x が、潜在変数 z からのガウス過程回帰により生成されると仮定し、潜在変数を計算する

✓カーネル関数により非線形の潜在変数を計算可能

✓z のカーネル関数であり、x のカーネル関数を計算するカーネル主成分分析とは異なる

✓主成分分析 + カーネル関数 + ガウス分布

✓次元削減手法の一つ

✓事前に潜在変数の数を決める

✓潜在変数の事前分布により、様々な潜在変数を計算可能

- ガウス過程回帰 (Gaussian Process Regression, GPR) https://datachemeng.com/gaussianprocessregression/ を理解していることが前提の説明になります

[1] N. Lawrence, Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data, NIPS Proceedings, 2003
[2]持橋 大地, 大羽 成征,ガウス過程と機械学習, 講談社, 2019

GPR と GPLVM

✓GPR で念頭にあったこと：
説明変数 x の値が似ている (近い) サンプル同士は、
目的変数 y の値も似ている (近い)

→ サンプル間における y の値の関係は、x の値の関係から計算できる

✓GPLVM で念頭にあること：
潜在変数 z の値が似ている (近い) サンプル同士は、
特徴量 x の値も似ている (近い)

→ サンプル間における x の値の関係は、z の値の関係から計算できる

GPR で導いたこと

✓i 番目のサンプルの y の平均値は 0

✓i 番目のサンプルと j 番目のサンプルとの間の y の共分散 ($j = i$ のときは分散) は、i 番目のサンプル の x の値と j 番目のサンプルの x の値から計算される

- 具体的には、 $K\left(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}\right)$ (K はカーネル関数)

カーネル関数の例 (他には、GPR のスライドの p.28-30 参照):

$$K\left(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}\right) = \theta_0 \exp\left\{-\frac{\theta_1}{2}\left\|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}\right\|^2\right\} + \theta_2 + \theta_3 \sum_{k=1}^{m} x_k^{(i)} x_k^{(j)}$$

GPLVM でもそのまま使います！

✓i 番目のサンプルの x の平均値は 0

✓i 番目のサンプルと j 番目のサンプルとの間の x の共分散 ($j = i$ のときは分散) は、i 番目のサンプル の z の値と j 番目のサンプルの z の値から計算される

- 具体的には、 $K\left(\mathbf{z}^{(i)}, \mathbf{z}^{(j)}\right)$ (K はカーネル関数)

z でカーネル関数を計算 (カーネル主成分分析では x で計算)

全サンプルの確率

✓潜在変数が z のとき、サンプル 1, 2, …, i, …, n の x の値が生じる確率

- n : サンプル数

$$p(\mathrm{x} \mid \mathrm{z}) = \prod_{i=1}^{n} p\left(\mathbf{x}^{(i)} \mid z\right)$$

$$= \prod_{i=1}^{n} N\left(\mathbf{x}^{(i)} \mid 0, K\left(\mathbf{z}^{(i)}, \mathbf{z}^{(j)}\right)\right)$$

平均 0、分散共分散行列が $K\left(\mathbf{z}^{(i)}, \mathbf{z}^{(j)}\right)$ の

正規分布

この確率が大きくなる z の値を求めたいが、z の事前分布 $p(\mathrm{z})$ は？

z の事前分布 p(z)

✓GPLVM における基本の事前分布 (いわゆる GPLVM)
- 各潜在変数の平均値 0, 分散 1,
  潜在変数間の共分散 0 の多次元正規分布

✓infinite Warped Mixture Model (iWMM) [1]
- 混合ガウス分布 (Gaussian Mixture Models, GMM)
  https://datachemeng.com/gaussianmixturemodel/

✓Gaussian Process Dynamical Model (GPDM) [2]
- 潜在変数 z が時系列データであること (時間発展すること) を仮定
- 非線形なマルコフ過程 (ある時刻の z の値は 1 時刻前の z の値のみに依存)

$$p(\mathrm{z}) = \prod_{i=2}^{n} p\left(\mathbf{z}^{(i)} \mid \mathbf{z}^{(i-1)}\right)$$ ⇨ カーネル関数で表現

[1] T. Iwata, et al., https://arxiv.org/abs/1206.1846
[2] J. M. Wang, et al., IEEE Transactions on Pattern Recognition and Machine Intelligence, 30, 283-298, 2008

GPLVMではX,Zの同時確率を最大化

同時確率 $p(\mathrm{x}, \mathrm{z}) = p(\mathrm{x} \mid \mathrm{z}) \times p(\mathrm{z})$ を最大化する z を求める

意味合い (式変形後)

x の分散共分散行列と、z のグラム行列 (カーネル関数の値の行列) の逆行列が等しくなるように、z を求める

潜在変数の高次元空間において直交化 (主成分分析) を行っているイメージ

GPLVM などを実行するためのコード

✓GPLVM

- GPy (Python): https://sheffieldml.github.io/GPy/
  - 変分ベイズ法を用いてハイパーパラメータも学習する Bayesian GPLVM が実装されており便利
- vargplvm (MATLAB): https://github.com/SheffieldML/vargplvm

✓iWMM

- warped-mixtures(MATLAB): https://github.com/duvenaud/warped-mixtures

✓GPDM

- http://www.dgp.toronto.edu/~jmwang/gpdm/ (MATLAB)