

変分行列分解に基づくベイズ的動的モード分解

川島貴大 (総合研究大学院大学 統計科学専攻 博士課程3年)

背景: 動的モード分解

動的モード分解 (DMD) は行列分解に基づいて多次元 (時) 系列データからダイナミクスを抽出する手法である [1]。\$D\$ 次元のデータ系列 \$\{\mathbf{y}_t\}_{t=1}^T\$ の時間発展が未知のダイナミクス \$\mathbf{f}: \mathbb{C}^D \rightarrow \mathbb{C}^D\$ によって \$\mathbf{y}_t = \mathbf{f}(\mathbf{y}_{t-1})\$ と書けるとすると, DMD は \$\mathbf{f}\$ を近似する線形ダイナミクス \$\mathbf{A} \in \mathbb{C}^{D \times D}\$ の \$K\$ 組の固有対 \$\{\lambda_k, \mathbf{w}_k\}_{k=1}^K\$ を効率的に計算することができる。

DMDのアルゴリズム

- データ行列の作成
\$D \times (T-1)\$ 行列 \$\mathbf{Y}_0 = (\mathbf{y}_1, \dots, \mathbf{y}_{T-1})\$, \$\mathbf{Y}_1 = (\mathbf{y}_2, \dots, \mathbf{y}_T)\$ を定義。線形ダイナミクスは \$\mathbf{Y}_1 = \mathbf{A}\mathbf{Y}_0\$ と書ける。
- 主部分空間とその上でのダイナミクスの推定
\$K\$ ランクの SVD を実行し (\$\mathbf{Y}_0 \approx \mathbf{U}_K \mathbf{L}_K \mathbf{V}_K^*\$), 以下の \$K \times K\$ 行列 \$\tilde{\mathbf{A}}\$ を考える:

$$\tilde{\mathbf{A}} := \mathbf{U}_K^* \mathbf{Y}_1 \mathbf{V}_K \mathbf{L}_K^{-1} (\approx \mathbf{U}_K^* \mathbf{Y}_1 \mathbf{Y}_0^{-1} \mathbf{U}_K \approx \mathbf{U}_K^* \mathbf{A} \mathbf{U}_K).$$

\$\tilde{\mathbf{A}}\$ は主部分空間上の時間発展行列と解釈できることに注意。

- 固有値分解
主部分空間上の時間発展行列の固有値分解 \$\tilde{\mathbf{A}} \tilde{\mathbf{w}}_k = \lambda_k \tilde{\mathbf{w}}_k\$ を計算。
- もとの空間上の固有対に変換
\$\mathbf{A}\$ の \$K\$ 組の支配的な固有対を \$\{(\lambda_k, \mathbf{U}_K \tilde{\mathbf{w}}_k)\}_{k=1}^K\$ として得る。

事前準備: 変分行列分解

変分行列分解 (VMF) は SVD ライクな行列分解を変分ベイズ的に実現する方法である [3]。VMF の確率モデルは行列複素正規分布 \$MCN\$ によって

$$p(\mathbf{Y} | \mathbf{U}_K, \mathbf{V}_K) = MCN(\mathbf{Y} | \mathbf{U}_K \mathbf{V}_K^*, \mathbf{I}_D, s^2 \mathbf{I}_T),$$
$$p(\mathbf{U}_K) \sim MCN(\mathbf{U}_K | \mathbf{O}, \mathbf{I}, \mathbf{C}_U), \quad p(\mathbf{V}_K) \sim MCN(\mathbf{V}_K | \mathbf{O}, \mathbf{I}, \mathbf{C}_V),$$

と与えられ, ここで \$\mathbf{C}_U, \mathbf{C}_V\$ は対角行列, \$s^2\$ は等方的なノイズの強さを決める分散パラメータである。\$p(\mathbf{U}_K, \mathbf{V}_K | \mathbf{Y}) \approx r(\mathbf{U}_K) r(\mathbf{V}_K)\$ と事後分布の独立性を仮定することで変分事後分布 \$r(\mathbf{U}_K), r(\mathbf{V}_K)\$ を推定するための次の更新則を導出できる:

$$r(\mathbf{U}_K) = MCN(\mathbf{U}_K | \tilde{\mathbf{U}}_K, \mathbf{I}_D, \tilde{\Sigma}_U), \quad r(\mathbf{V}_K) = MCN(\mathbf{V}_K | \tilde{\mathbf{V}}_K, \mathbf{I}_T, \tilde{\Sigma}_V),$$
$$\tilde{\Sigma}_U^{-1} = \mathbf{C}_U^{-1} + s^{-2}(T \tilde{\Sigma}_V + \tilde{\mathbf{V}}_K^* \tilde{\mathbf{V}}_K), \quad \tilde{\Sigma}_V^{-1} = \mathbf{C}_V^{-1} + s^{-2}(D \tilde{\Sigma}_U + \tilde{\mathbf{U}}_K^* \tilde{\mathbf{U}}_K),$$
$$\tilde{\mathbf{U}}_K = s^{-2} \mathbf{Y} \tilde{\mathbf{V}}_K \tilde{\Sigma}_U, \quad \tilde{\mathbf{V}}_K = s^{-2} \mathbf{Y}^* \tilde{\mathbf{U}}_K \tilde{\Sigma}_V.$$

この更新則によって得られる \$r(\mathbf{U}_K)\$ は主部分空間への射影行列の確率表現と解釈できる。また更新則は少し複雑になるものの, VMF は欠測のあるデータにも適用可能である。

提案法: BDMD-VMF

本研究では VMF に基づく DMD のベイズ的拡張, BDMD-VMF を開発した。DMD は \$t\$ 時点目のデータ \$\mathbf{y}_t\$ に対して

$$\mathbf{y}_t \approx \sum_{k=1}^K \lambda_k^{t-1} \mathbf{w}_k b_k, \quad t = 1, \dots, T,$$

の近似が良くなるような \$\lambda_k \in \mathbb{C}, \mathbf{w}_k \in \mathbb{C}^D\$ を求めることになる。ここで \$b_k\$ は \$k\$ 番目のモードの強度を決める係数である。DMD のアルゴリズムでは, \$\tilde{\mathbf{A}} \in \mathbb{C}^{K \times K}\$ の固有対 \$\{(\lambda_k, \tilde{\mathbf{w}}_k)\}_{k=1}^K\$ によってもとの空間におけるダイナミクス \$\mathbf{A} \in \mathbb{C}^{D \times D}\$ の \$K\$ 組の固有対を \$\{(\lambda_k, \mathbf{U}_K \tilde{\mathbf{w}}_k)\}_{k=1}^K\$ と得た。そこで上の \$\mathbf{y}_t\$ の近似式を

$$\mathbf{y}_t \approx \mathbf{U}_K \sum_{k=1}^K \lambda_k^{t-1} \tilde{\mathbf{w}}_k, \quad t = 1, \dots, T, \quad (1)$$

と書き換える。ここで \$\tilde{\mathbf{w}}_k \in \mathbb{C}^K\$ は \$\mathbf{w}_k b_k = \mathbf{U}_K \tilde{\mathbf{w}}_k\$ を満たすベクトルである。まず (1) 式を平均とする複素正規分布 \$\mathcal{CN}\$ によって尤度を定義する:

$$p(\mathbf{y}_t | \mathbf{U}_K, \{\lambda_k\}, \{\tilde{\mathbf{w}}_k\}, \sigma^2) = \mathcal{CN}\left(\mathbf{y}_t \mid \mathbf{U}_K \sum_{k=1}^K \lambda_k^{t-1} \tilde{\mathbf{w}}_k, \sigma^2 \mathbf{I}\right). \quad (2)$$

\$\mathbf{U}_K\$ は主部分空間への射影行列であるため, 事前分布として VMF の変分事後分布 \$r(\mathbf{U}_K)\$ を採用する。ガウス型の前分布 \$r(\mathbf{U}_K)\$ によって尤度 (2) から \$\mathbf{U}_K\$ の解析的な周辺化が可能となり, 以下の BDMD-VMF の尤度を得る。

BDMD-VMFの尤度 (欠測のない場合)

データ行列 \$\mathbf{Y} \in \mathbb{C}^{D \times T}\$ に欠測がない場合, BDMD-VMF の尤度は以下のように書ける:

$$p(\mathbf{Y} | \{\lambda_k\}, \{\tilde{\mathbf{w}}_k\}, \sigma^2) = \prod_{t=1}^T \mathcal{CN}(\mathbf{y}_t | \tilde{\mathbf{y}}_t, \sigma_t^2 \mathbf{I}_D).$$

ここで各パラメータは

$$\mathbf{g}_t = \sum_{k=1}^K \lambda_k^{t-1} \tilde{\mathbf{w}}_k, \quad \sigma_t^{-2} = \sigma^{-2} (1 - \mathbf{g}_t^* \tilde{\Sigma}_U^{-1} \mathbf{g}_t),$$
$$\tilde{\mathbf{y}}_t = \sigma^{-2} \tilde{\mathbf{U}}_K \tilde{\Sigma}_U^{-1} \tilde{\Sigma}_U^{(t)} \mathbf{g}_t, \quad \tilde{\Sigma}_U^{(t-1)} = \sigma^{-2} \mathbf{g}_t \mathbf{g}_t^* + \tilde{\Sigma}_U.$$

BDMD-VMF にはいくつかの特筆すべき性質がある。

- パラメータ数は \$2K(K+1)+1\$
 - 入力次元 \$D\$ にもサンプル数 \$T\$ にも依存しない
- 欠測のあるデータにも適用可能
 - VMF が欠測をうまく扱えるため
 - ただし尤度の式は少し複雑になる
- 時刻 \$t\$ に依存しないパラメータによる条件付き独立性 \$\mathbf{y}_t \perp \mathbf{y}_{t'} | \{\lambda_k\}, \{\tilde{\mathbf{w}}_k\}, \sigma^2\$ が \$t \neq t'\$ において成り立つ
 - タイトな予測分布をもたらす

実験

非線形偏微分方程式である非線形シュレーディンガー方程式およびバーガース方程式からデータを生成し, DMD, Bayesian DMD [2], BDMD-VMF によるダイナミクスの推定と再構成を行った (Figure 1, 2)。

Bayesian DMD がバーガース方程式において再構成に失敗している一方, BDMD-VMF はうまくダイナミクスを捉えられていることがわかる。BDMD-VMF のモデルは最低限のパラメータのみによって成り立つため, 安定した推定に多くの観測値を必要としないことがこの違いの要因である。

次に HuGaDB データセット [4] から, 脚に着けられたジャイロセンサによって得た自転車運転のデータを用いて欠測を含むデータを作成した。BDMD-VMF は時間発展をうまく捉えており, 非常にタイトな予測分布をもたらしている。

むすび

本研究では DMD の新たなベイズモデルである BDMD-VMF を提案した。BDMD-VMF は少量の本質的なパラメータのみによってモデリングされ, 欠測データの扱いやタイトな予測の点で優れている。今後の主要な研究課題として, 高速なパラメータ推定法の開発やモデルの非線形拡張が挙げられる。本研究は日野英逸教授 (統数研), 庄野逸教授 (電通大) との共同研究である [5]。

参考文献

- Schmid, P. J., "Dynamic mode decomposition of numerical and experimental data," *J. Fluid Mech.* 656, (2010)
- Takeishi, N. et al., "Bayesian Dynamic Mode Decomposition," *IJCAI-17*, (2017)
- Lim, Y. J. et al., "Variational Bayesian Approach to Movie Rating Prediction," *KDD Cup adn Workshop 2007* (2007)
- Chereshnev, R. et al., "HuGaDB: Human Gait Database for Activity Recognition from Wearable Inertial Sensor Networks," *Analysis of Images, Social Networks and Texts*, (2018)
- Kawashima, T. et al., "Bayesian Dynamic Mode Decomposition with Variational Matrix Factorization," *AAAI-21*, (2021)

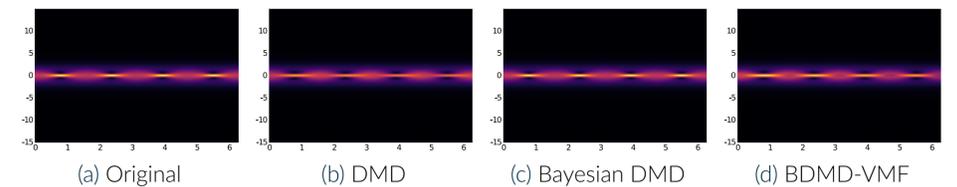


Figure 1. 非線形シュレーディンガー方程式に対する再構成結果 (\$D = 256, T = 256, K = 4\$)。

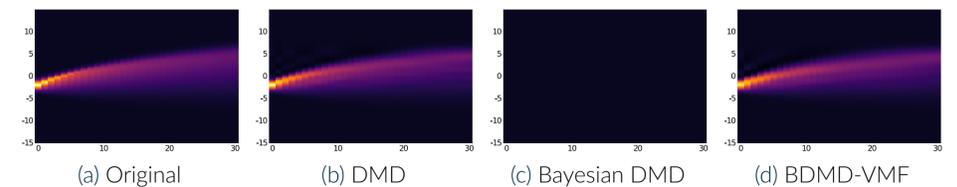


Figure 2. バーガース方程式に対する再構成結果 (\$D = 256, T = 31, K = 7\$)。

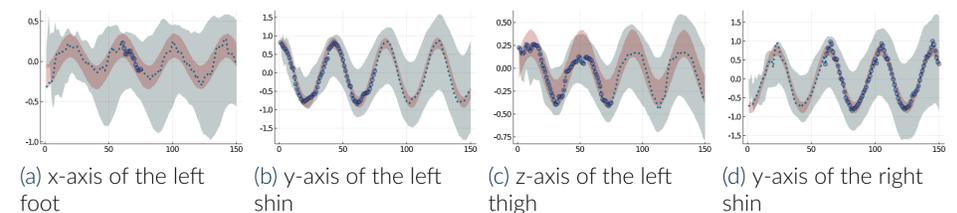


Figure 3. ジャイロセンサデータに対する再構成結果の一部 (\$D = 18, T = 150, K = 2\$)。点線は真のデータ, マーカーは実際に入力データを表す。グレーと赤でハイライトされた領域はそれぞれ Bayesian VAR(2) モデルと BDMD-VMF による 95% 信用区間。