Bayesian Dynamic Mode Decomposition

Naoya Takeishi[§], Yoshinobu Kawahara[†,‡],
Yasuo Tabei[‡], Takehisa Yairi[§]

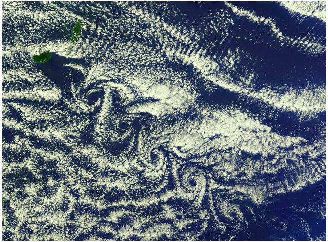
[§]Dept. of Aeronautics & Astronautics, The University of Tokyo
[†]The Institute of Scientific and Industrial Research, Osaka University
[‡]RIKEN Center for Advanced Intelligence Project (AIP)

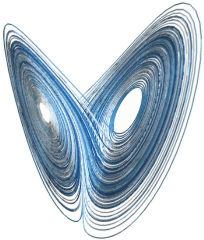
22 August 2017, IJCAI @ Melbourne

Motivation: Analysis of dynamical systems

- Various types of complex phenomena can be described in terms of (nonlinear) **dynamical systems**.

$$\boldsymbol{x}_{t+1} = \boldsymbol{f}(\boldsymbol{x}_t), \quad \boldsymbol{x} \in \mathcal{M} \text{ (state space)}$$





☹ When $\boldsymbol{f}$ is nonlinear, analysis based on trajectories of $\boldsymbol{x}$ is difficult.

Operator-theoretic view of dynamical systems

Definition (Koopman operator [Koopman '31, Mezić '05])

Koopman operator (composition operator) $\mathcal{K}$ represents time-evolution of *observables* (i.e., observation function) $g : \mathcal{M} \to \mathbb{R}$ or $\mathbb{C}$.

$$\mathcal{K}g(\boldsymbol{x}) = g(\boldsymbol{f}(\boldsymbol{x})), \quad g \in \mathcal{F} \text{ (function space)}$$

- $\mathcal{K}$ describes temporal evolution of *function* (infinite-dimensional vector) instead of the finite-dimensional state vector.
- Defining $\mathcal{K}$, we can *lift* the analysis of nonlinear dynamical systems into a linear (but infinite-dimensional) regime!

☺ Since $\mathcal{K}$ is linear, we can analyze dynamics using the spectra of $\mathcal{K}$.

Koopman mode decomposition (KMD)

- Eigenvalues and eigenfunctions of $\mathcal{K}$:

$$\mathcal{K}\,\varphi_i(\boldsymbol{x}) = \lambda_i\,\varphi_i(\boldsymbol{x}) \quad \text{for} \quad i = 1, 2, \dots .$$

- Projection of $g(\boldsymbol{x})$ to $\operatorname{span}\{\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}), \dots\}$ (i.e., transformation to a canonical form). $\Rightarrow$ Coefficients are called *Koopman modes*.

$$g(\boldsymbol{x}) = \sum_{i=1}^{\infty} \varphi_i(\boldsymbol{x})\, v_i$$

- Since φ is eigenfunction,

$$g(\boldsymbol{x}_t) = \sum_{i=1}^{\infty} \lambda_i^t \underbrace{\varphi_i(\boldsymbol{x}_0) v_i}_{w_i}, \quad \text{(KMD)}$$

where $|\lambda_i|$ = decay rate of w_i, $\angle\lambda_i$ = frequency of w_i.

- A numerical realization of KMD is *dynamic mode decomposition* (DMD) [Rowley+ '09, Schmid '10, Tu+ '14].

Dynamic mode decomposition (DMD)

Assumption ($\mathcal{K}$-invariant subspace [Budišić+ '12])

Dataset is generated with a set of observables

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} g_1(\boldsymbol{x}) & g_2(\boldsymbol{x}) & \cdots & g_n(\boldsymbol{x}) \end{bmatrix}^{\mathsf{T}}$$

that spans (approximately) $\mathcal{K}$-invariant subspace.

$\Rightarrow$ Then, KMD can be (approximately) realized by DMD.

Algorithm (DMD [Tu+ '14])

Input time-series $(\boldsymbol{y}_0, \dots, \boldsymbol{y}_m)$ s.t. $\boldsymbol{y}_t = \boldsymbol{g}(\boldsymbol{x}_t)$

Output eigenvalues $\{\lambda\}$, eigenfunctions $\{\varphi\}$, and modes $\{\boldsymbol{w}\}$

1. Estimate a linear model $\boldsymbol{y}_{t+1} \approx \boldsymbol{A}\boldsymbol{y}_t$.
2. On $\boldsymbol{A}$, compute eigenvalues λ_i and right-/left-eigenvectors $\boldsymbol{w}_i$, $\boldsymbol{z}_i^{\mathsf{H}}$.
3. Compute $\varphi_{i,t} = \boldsymbol{z}_i^{\mathsf{H}}\boldsymbol{y}_t$.

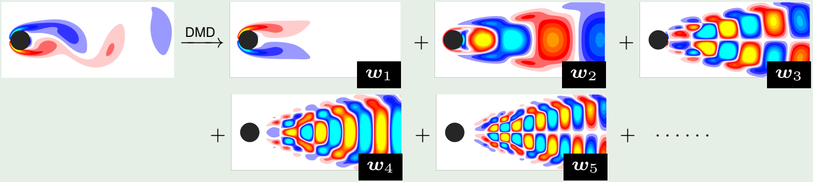
Quasi-periodic modes extraction by KMD/DMD

- Review: KMD/DMD computes the decomposition of time-series into modes $\boldsymbol{w}_i$ that evolve with frequency $\angle\lambda_i$ and decay rate $|\lambda_i|$.
  - $\boldsymbol{w}_i$ is termed *dynamic modes*.

$$\boldsymbol{g}(\boldsymbol{x}_t) \approx \sum_{i=1}^{\infty} \lambda_i^t \, \boldsymbol{w}_i$$

Example (2D fluid flow past a cylinder)

Flow past a cylinder is universal in many natural/engineering situations.



Other applications of KMD/DMD

- Lots of applications in a wide range of domains
  - fluid mechanics [Rowley+ '09, Schmid '10, & many more],
  - neuroscience [Brunton+ '16],
  - image processing [Kutz+ '16, Takeishi+ '17],
  - analysis of power systems [Raak+ '16, Susuki+ '16],
  - epidemiology [Proctor&Eckhoff '15],
  - optimal control [Mauroy&Goncalves '16],
  - finance [Mann&Kutz '16],
  - medical care [Bourantas+ '14],
  - robotics [Berger+ '15], etc.

Issue

- DMD relies on linear modeling $\boldsymbol{g}(\boldsymbol{x}_{t+1}) \approx \boldsymbol{A}\boldsymbol{g}(\boldsymbol{x}_t)$ and eigendecomposition of $\boldsymbol{A}$.
- So it lacks an associated *probabilistic/Bayesian framework*, by which we can
  - consider *observation noise* explicitly,
  - perform a *posterior inference*,
  - consider *DMD extensions* in a unified manner, etc.
- Let's do it!
  - analogously to PCA's formulation as probabilistic/Bayesian PCA [Tipping&Bishop '99, Bishop '99]

Proposed method (1/2): Probabilistic DMD

- Dataset: *snapshot pairs* with observation noise

$$\mathcal{D} = \Big((\boldsymbol{y}_{0,1}, \boldsymbol{y}_{1,1}),\ \dots,\ (\boldsymbol{y}_{0,t}, \boldsymbol{y}_{1,t}),\ \dots,\ (\boldsymbol{y}_{0,m}, \boldsymbol{y}_{1,m}) \Big),$$

where $\boldsymbol{y}_{0,t} = \boldsymbol{g}(\boldsymbol{x}_t) + e_{0,t}$ and $\boldsymbol{y}_{1,t} = \boldsymbol{g}(\boldsymbol{x}_{t+\Delta t}) + e_{1,t}$,

Definition (Generative model of probabilistic DMD)

$$\begin{cases} \boldsymbol{y}_{0,t} \sim \mathcal{CN}\left(\sum_{i=1}^{k} \varphi_{t,i} \boldsymbol{w}_i,\ \sigma^2 I \right) \\ \boldsymbol{y}_{1,t} \sim \mathcal{CN}\left(\sum_{i=1}^{k} \lambda_i \varphi_{t,i} \boldsymbol{w}_i,\ \sigma^2 I \right) \end{cases}$$

$$\varphi_{t,i} \sim \mathcal{CN}(0,\ 1)$$

☺ If $k = n$ and $\sigma^2 \to 0$, the MLE of $(\lambda, \boldsymbol{w})$ coincides with DMD's solution.

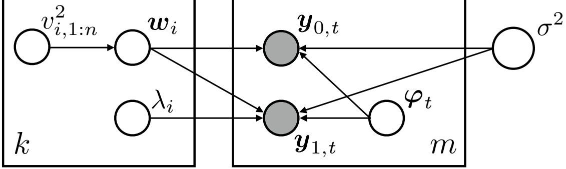
Proposed method (2/2): Bayesian DMD

Definition (Prior on parameters for Bayesian DMD)

$$\boldsymbol{w}_i | v^2_{i,1:n} \sim \mathcal{CN}\left(\mathbf{0},\ \mathrm{diag}\left(v^2_{i,1}, \ldots, v^2_{i,n}\right)\right), \quad v^2_{i,d} \sim \mathrm{InvGamma}\left(\alpha_v,\ \beta_v\right)$$

$$\lambda_i \sim \mathcal{CN}\left(0,\ 1\right)$$

$$\sigma^2 \sim \mathrm{InvGamma}\left(\alpha_\sigma,\ \beta_\sigma\right)$$



☺ For a *posterior inference*, a Gibbs sampler can be constructed easily.

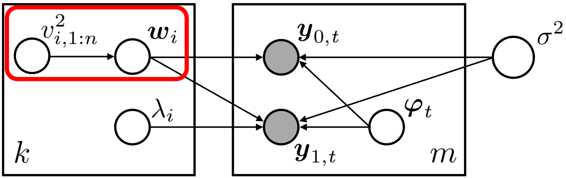
Extension example: Sparse Bayesian DMD

Definition (Prior on parameters for sparse Bayesian DMD)

$$\boldsymbol{w}_i | v_{i,1:n}^2 \sim \mathcal{CN}\left(\mathbf{0},\ \sigma^2 \mathrm{diag}\left(v_{i,1}^2, \ldots, v_{i,n}^2\right)\right), \quad v_{i,d}^2 \sim \mathrm{Exponential}(\gamma_i^2/2)$$

$$\lambda_i \sim \mathcal{CN}\left(0,\ 1\right)$$

$$\sigma^2 \sim \mathrm{InvGamma}\left(\alpha_\sigma,\ \beta_\sigma\right)$$



☺ We can *extend the model* in a unified Bayesian manner.

Numerical example (1)

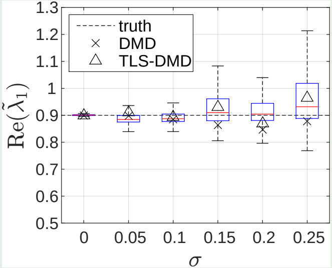
Example (Fixed-point attractor)

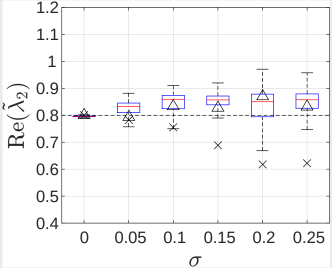
Generate data by

$$\boldsymbol{y}_t = \lambda_1^t \begin{bmatrix}2 & 2\end{bmatrix}^\mathsf{T} + \lambda_2^t \begin{bmatrix}2 & -2\end{bmatrix}^\mathsf{T} + \boldsymbol{e}_t,$$

where e is Gaussian observation noise.

True eigenvalues are $\lambda_1 = 0.9$ and $\lambda_2 = 0.8$.





Numerical example (2)

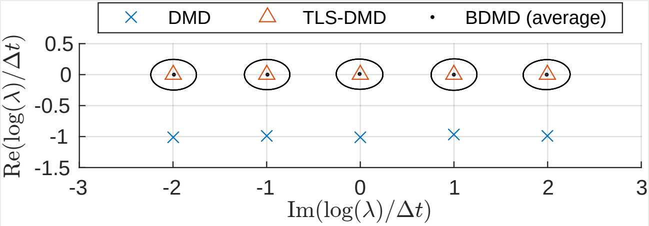
Example (Limit-cycle attractor)

Generate data from Stuart–Landau equation

$$r_{t+1} = r_t + \Delta t(\mu r_t - r_t^3),$$
$$\theta_{t+1} = \theta_t + \Delta t(\gamma - \beta r_t^2),$$

and Gaussian observation noise.

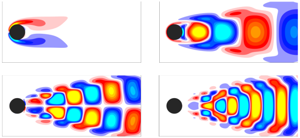
True (continuous-time) eigenvalues lie on the imaginary axis.



Analysis of dynamical systems based on **Koopman operator** is a useful tool.

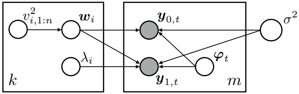
$$\mathcal{K}g(\boldsymbol{x}) = g(\boldsymbol{f}(\boldsymbol{x}))$$

Dynamic mode decomposition (DMD) is a numerical method for Koopman analysis.



In this work, we developed **probabilistic & Bayesian DMDs** to

- consider observation noise,
- infer posterior distribution,
- extend DMD in a unified manner, etc.



Implementation available at
https://github.com/n-takeishi/bayesiandmd