MATH-BIOINF-STATS 547: Mathematics of Data

Due Date: 02/17/2025

Problem Set 2: Dynamic Mode Decomposition

For this problem set, please submit a `.pdf` document with a write-up of your results and observations. We encourage using Overleaf, but the MATLAB Live Editor or other word-processing software is acceptable. We have provided a LaTeX template to help get you started, which is available on Overleaf where you can make a copy of it.

Background

Time-dependent phenomena underlie many areas of scientific research. Recent developments in technology enable efficient collection of time series data. Gaining insight into these data requires application of innovative algorithms. We will introduce a beautiful technique called dynamic mode decomposition (DMD), which was first introduced by Schmid and Sesterhenn in 2008 [1]. DMD is a data-driven and model-free algorithm extracting spatio-temporal patterns in the form of so-called DMD modes and DMD eigenvalues. DMD has been investigated on both practical and theoretical grounds.

First I will give you a simple prototype example motivated by Alan Turing. Here we model dynamics of a single cell with two proteins as an Andronov-Hopf oscillator, which has the form

$$\frac{dx}{dt} = -\left(x^2 + y^2\right) x + ax - by \tag{1a}$$

$$\frac{dy}{dt} = -\left(x^2 + y^2\right) y + ay + bx, \qquad a, b > 0 \tag{1b}$$

where, $x, y \in \mathbb{R}$ can be interpreted as concentrations of the proteins in the cell. To give a graphical illustration, we numerically integrate Equation 1a and 1b. Figure 1 shows the graph of $x(t)$ and $y(t)$ and plots the solutions in a phase plane.

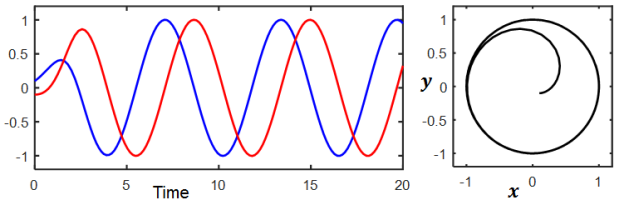

Figure 1: Dynamics of a single cell. Left: Blue ($x(t)$) and red ($y(t)$) represent protein concentrations in a single cell. The initial conditions are $(x(0), y(0) = (0.1, -0.1)$, and $a = b = 1$. Right: The solutions in a phase plane.

Now we extend this formulation to two cells, separated by a membrane, that each have the same two proteins, using Equations 1a and 1b with $a = 1, b = 1$. The two cells are coupled in the following way:

$$\text{Cell 1:} \quad \begin{aligned} \tfrac{dx_1}{dt} &= -\left(x_1^2 + y_1^2\right) x_1 + x_1 - y_1 + c\,(x_2 - x_1) \\ \tfrac{dy_1}{dt} &= -\left(x_1^2 + y_1^2\right) y_1 + y_1 + x_1 + c\,(y_2 - y_1) \end{aligned} \tag{2a}$$

$$\text{Cell 2:} \quad \begin{aligned} \tfrac{dx_2}{dt} &= -\left(x_2^2 + y_2^2\right) x_2 + x_2 - y_2 + c\,(x_1 - x_2) \\ \tfrac{dy_2}{dt} &= -\left(x_2^2 + y_2^2\right) y_2 + y_2 + x_2 + c\,(y_1 - y_2) \end{aligned} \tag{2b}$$

where for protein x, x_1 is the protein concentration for cell 1, and x_2 is the protein concentration for cell 2, and likewise for y_1 and y_2 for cells 1 and 2. This coupling is called Turing-type coupling, where $c > 0$ is the coupling parameter. Now we show the results of a numerical study of System 2 for different values of c

and initial conditions. Figure 2 (bottom) shows that when two cells are coupled, $x_1(t) = x_2(t) = x(t)$ and $y_1(t) = y_2(t) = y(t)$. We then say the proteins are "Beating in Unison," that is two proteins $x(t)$ and $y(t)$ oscillate in phase synchronization.

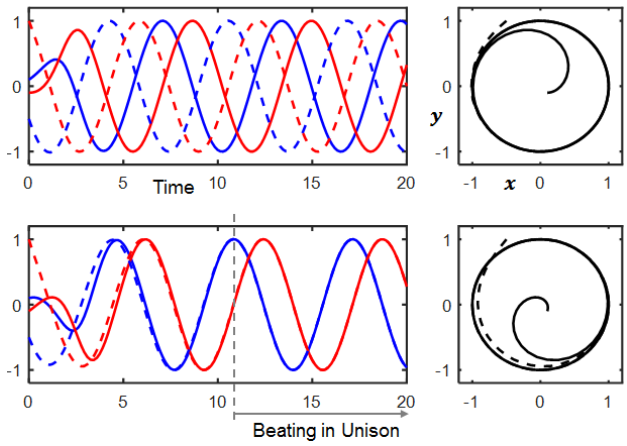

Figure 2: Dynamics of two cells. The initial conditions are $(x_1, y_1, x_2, y_2) = (0.1, -0.1, -0.5, 1)$. Top: $c = 0$, bottom: $c = 0.1$. Red, protein x; blue, protein y; solid line, cell 1; dotted line cell 2. Right panels are phase planes. Solid line is cell 1 (x_1, y_1); dotted line is cell 2 (x_2, y_2).

Figure 3 is an example of two proteins that are not Beating in Unison. This system is a special case where the initial conditions are opposite $(0.1, -0.1, -0.1, 0.1)$. The eigenvalues of the Jacobian matrix at $(0, 0, 0, 0)$ are $1-2c-i, 1-2c+i, 1-i$, and $1+i$. When $c > 0.5$, the eigenvalues $1-2c-i$ and $1-2c+i$ have negative real parts. The corresponding eigenvectors are $(1, -1, -i, 1), (-1, -1, -i, 1), (-1, 1, -i, 1), (1, 1, i, 1)$.

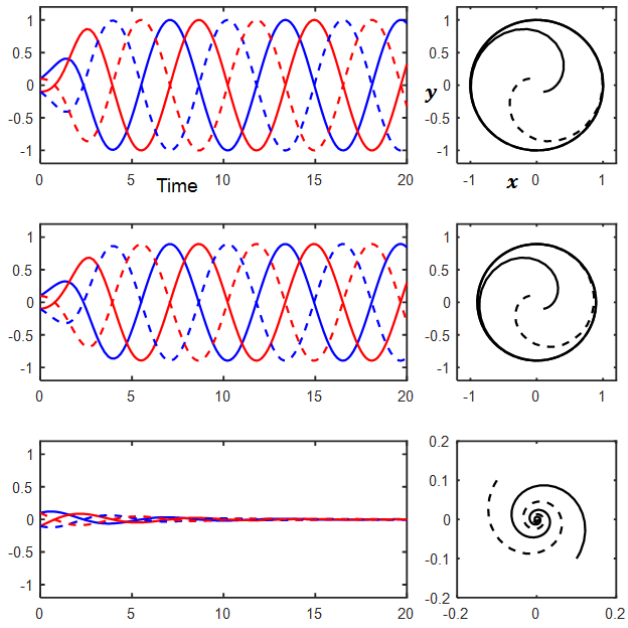

Figure 3: Dynamics of two cells. Initial conditions are $(x_1, y_1, x_2, y_2) = (0.1, -0.1, -0.1, 0.1)$. Top: $c = 0$, middle: $c = 0.1$, bottom: $c = 0.6$. Red, protein x; blue, protein y; solid line, cell 1; dotted line cell 2. Right panels are phase planes. Solid line is cell 1 (x_1, y_1); dotted line is cell 2 (x_2, y_2).

DMD Algorithm

1. Compute the SVD of the first data matrix, $\mathbf{X} = \mathbf{U\Sigma V}^{\mathsf{T}}$.

   Remember that DMD uses a sequence of snapshots

   $$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{m-1}] \in^{n \times m-1}$$

   $$\mathbf{X}' = [\mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_m] \in^{n \times m-1}$$

2. We may now make the substitution into $\mathbf{X}' = \mathbf{AX}$ and write $\mathbf{X}' = \mathbf{AU\Sigma V}^{\mathsf{T}}$.

3. Define $\tilde{\mathbf{A}} \triangleq \mathbf{U}^{\mathsf{T}}\mathbf{AU} = \mathbf{U}^{\mathsf{T}}\mathbf{X}'\mathbf{V\Sigma}^{-1}$.

4. Compute the eigendecomposition of $\tilde{\mathbf{A}}$

   $$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W\Lambda},$$

   where $\mathbf{W}$ is the matrix of eigenvectors, and $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues. Each eigenvalue λ_i is a DMD eigenvalue.

5. Compute the DMD modes,

   $$\mathbf{\Phi} \triangleq \mathbf{X}'\mathbf{V\Sigma}^{-1}\mathbf{W}.$$

   Each column of $\mathbf{\Phi}$ is a DMD mode ϕ corresponding to eigenvalue λ_i.

Note. You may need to install packages from MATLAB to run DMD. MATLAB should prompt you if the packages are missing and direct you on how to install them. These packages should be free to download once you have confirmed that your account is associated with the university.

Exercise 1 (DMD Theory)**.** Building on the DMD note set, derive the following properties of DMD. Given a $n \times m$ data matrix $\mathbf{X}$ with m measurements on n samples, DMD is performed to obtain $\tilde{\mathbf{A}}$ with has eigenvectors Λ and eigenvectors $\mathbf{W}$ and DMD modes Φ. Given a time point x_t, write x_{t+1} in terms of the DMD modes Φ and eigenvalues of $\tilde{\mathbf{A}}$.

Exercise 2 (Synthetic Data)**.** Run the first section of starter code PS2_starter_code.mlx in MATLAB with the following initial conditions and parameters. Then, answer the below questions.

Initial Conditions:

- $(x_1, y_1, x_2, y_2) = (0.1, -0.1, -0.5, 1)$ and $c = 0$
- $(x_1, y_1, x_2, y_2) = (0.1, -0.1, -0.5, 1)$ and $c = 0.1$
- $(x_1, y_1, x_2, y_2) = (0.1, -0.1, -0.1, 0.1)$ and $c = 0.6$

(a) Show the outputs (save figures as images and add to your submission document) and explain what you observe. Are there any relationships between the subplots within each figure?

(b) Compare the three figures. Identify and explain any differences you notice that were caused by manipulating the parameters and initial conditions.

Exercise 3 (Fluid Dynamics)**.** Read Section 2.3 of an example application of DMD to Fluid Dynamics in the DMD Book. The associated data and code of this example have been provided in the problem set MATLAB starter code. Run and follow the MATLAB code to generate similar plots seen in the example (Note: the starter code *should* generate different plots than are in the book). Include your figures and answer the following questions in your write up:

(a) What do the eigenvalues of A indicate about the cylindrical wake flow?

(b) What are the relative advantages and disadvantages between DMD and POD in this case?

Exercise 4 (Neuroscience)**.** Run the Section 4 of PS2_starter_code.mlx in MATLAB. Read Chapter 12.3 of Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems [2], which explains the data and methods of augmenting `X` and `Xaug` prior to computing DMD. More details on the augmentation of data in this example can be found in Chapter 7.

(a) Show the outputs (save figures as images and add to your submission document) and explain what you observe. What are the relationships between the subplots within each figure?

(b) Compare the figures generated using `Xaug` and `X`. Identify and explain any differences that occurred after augmenting the data.

(c) What are the main advantages and disadvantages of this data augmentation process?

Exercise 5 (Mouse Data)**.** For this exercise, you will work with time series mouse endomicroscopy neuronal data. Run the Section 5 of PS2_starter_code.mlx in MATLAB which preforms DMD on the data.

(a) Show the outputs (save figures as images and add to your submission document).

(b) Interpret the results in context.

Exercise 6 (EEG Data)**.** For this exercise, we will work with implementing DMD on data with multiple samples. For this exercise, we will only look at the first patient.

(a) What is the dimension of the data?

(b) What is the rank of each matrix? Why is this information important for DMD?

(c) Preform DMD on both the eyes open and closed data sets, compare the results. Include eigenvalue plots of your linear approximation.

References

[1] Peter Schmid and Joern Sesterhenn. Dynamic Mode Decomposition of numerical and experimental data. In *APS Division of Fluid Dynamics Meeting Abstracts*, volume 61 of *APS Meeting Abstracts*, page MR.007, November 2008.

[2] J Nathan Kutz, Steven L Brunton, Bingni W Brunton, and Joshua L Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.