Coding Project 5: Background Subtraction through Dynamic Mode Decomposition

Tyler Shakibai

Abstract

Dynamic mode decomposition (DMD) is used to isolate and remove the background modes from video footage. We use DMD to obtain a spectrum of frequencies so that the ones corresponding to the background footage can be subtracted out. To perform this task we approximate a low-rank and sparse DMD reconstruction.

Introduction Dynamic mode decomposition (DMD) is a dimensionality reduction method that operates on time-series data to compute the set of modes that correspond to constant oscillation and decay/growth rates, and incorporate temporal dynamics. These modes are approximations of the modes and eigenvalues of the Koopman operator, derived from nonlinear dynamics.

In this report, we will provide an overview of how dynamic mode decomposition works, and how it be applied to the task of background removal. To perform this task we approximate a low-rank and sparse DMD reconstruction of our time-series measurements. DMD is well-suited to problems where the spatial and temporal components are both crucial to the qualitative behavior of the system.

Theoretical Background Dynamic mode decomposition arises from instances where we would like to understand the underlying dynamics but we only have snapshot measurements of the system. Consider some time-series measurements x^n on the spatial domain where we have

$$X = \begin{bmatrix} U(\vec{x}, t_1) & U(\vec{x}, t_2) & \dots & U(\vec{x}, t_M) \end{bmatrix}$$

where M is the number of time snapshots we have of a certain system. We are interested in taking specific subsets of these measurements, which can be written as

$$X_j^k = \begin{bmatrix} U(\vec{x}, t_j) & U(\vec{x}, t_{j+1}) & \dots & U(\vec{x}, t_k) \end{bmatrix}$$

Following this notation, the expression for X_1 becomes

$$X_1^{M-1} = \begin{bmatrix} x_1 & x_2 & \dots & x_{M-1} \end{bmatrix}$$

Using the Koopman operator A, we can redefine X_1 as

$$X_1 = \begin{bmatrix} x_1 & Ax_1 & \dots & A^{M-2}x_1 \end{bmatrix}$$

We can perform the exact same process with $j = 2$ to get X_2, which is the time series for the next times which is defined as

$$X_2 = \begin{bmatrix} x_1 & Ax_1 & \dots & A^{M-1}x_1 \end{bmatrix}$$

Therefore, the next time series is dependent on the previous time series with the following relation

$$X_2^M = AX_1^{M-1} + \vec{r}e_{M-1}^T$$

Results Dynamic mode decomposition was able to obtain a spectrum of frequencies so that the ones which corresponded to the unwanted background footage could be subtracted from the original footage. DMD was used to compute modes which are approximations of the modes and eigenvalues of the Koopman operator, and remove the necessary ones. More specifically, this was done by taking the singular value decomposition followed by finding the eigenvalues of Σ. Using this data, we were able to iterative find the differential equation that best fit the data from our measurements at each realization. Finally, we were able to subtract the isolated information pertaining to the background from our original data, leaving only the foreground footage.

Conclusion Dynamic mode decomposition was used to isolate and remove the background modes from video footage. We used DMD to obtain a spectrum of frequencies so that the ones corresponding to the background footage can be subtracted from our original data. We performed this task by approximating a low-rank and sparse DMD reconstruction.

After using Matlab to watch the modified footage, I can conclude that DMD was fairly successful in isolating the correct modes to be removed from our original video. With the exception of some subtle noise, the resulting footage appear to be entirely composed of the object in the foreground.

Acknowledgment I attended office hours with the course's TAs, Katie and Michael, and heavily utilized the Numpy and Scipy Documentation.