Supplementary Information for

Analyzing Gang Reduction Strategies Through Dynamic Mode Decomposition

Yifan Li, Jiazhong Mei, Aviva Prins, and Qinyi Zeng

Aviva Prins.
E-mail: aviva@avivaprins.com

This PDF file includes:

Supplementary text
Figs. S1 to S2
References for SI reference citations

Supporting Information Text

1. Earth Mover's Distance

The Earth Mover's Distance is a measure that quantifies the distance between two histograms (or distributions in general). Since we have the distributions of risk factors of some participants in the GRYD program from the beginning and from the end of an intervention period, we design a simple preliminary metric to quantify the decay of their risk factors over time.

For two times t_1 and t_2 such that $t_1 < t_2$, we identify participants who have questionnaire responses at both times and extract their corresponding risk scores at t_1 and t_2 as r_1 and r_2. The we define their **decay score** (DS) between t_1 and t_2 as the following:

$$\mathrm{DS}_{12} = \frac{\sum_i (v_{2i} - v_{1i})}{\sum_i -v_{1i}} \cdot 100 \quad [1]$$

The term $\sum_i (v_{2i} - v_{1i})$ gives the total decay in the mass of risk factors from t_1 to t_2, while $\sum_i -v_{1i}$ gives the total amount of decay if every participant's risk score becomes 0 at t_2. As a result, DS_{12} conveys the amount of decay in risk factors during the time period relative to the amount of decay that can possibly happen in the system.

In a system where $\sum_i v_{2i} < \sum_i v_{1i}$ (i.e. a system where the mass transported to the lower risk scores is less than that transported to the higher risk scores), DS will range from 0 to 100 with 0 indicating no decay and 100 indicating the most extreme decay has happened. Thus the larger DS is, the more intense the decay is in the system.

2. Dynamical Systems and the Koopman Operator

In mathematics, it is common to examine the temporal changes using the theories of dynamical systems. A dynamical system is typically equipped with a set of points sitting closely on a manifold $\mathcal{M}$ and a function f which describes the dynamics of the system over time. That is:

$$\frac{dx}{dt} = f(x, t; \mu), \quad [2]$$

where x is a point on $\mathcal{M}$, t stands for time, and μ are the parameters of the system.

If we consider a system in a setting of discrete time, the system is naturally described as:

$$x_{t+1} = f(x_t), \quad [3]$$

for $t = 0, 1, 2, ...$, and x_t stands for a snapshot of the point x at time t.

Examples of dynamical systems include fluid flows, evolution of certain animal species in a certain habitat, neuron networks, and even financial markets. And as we attempt to examine the properties of the unknown dynamics function f, we turn our attention to the space of functions on $\mathcal{M}$: $\{\varphi : \mathcal{M} \to \mathcal{M}\}$. So instead of looking at f which maps between points, we examine scalar functions $S = \{g : \mathcal{M} \to \mathbb{R}\}$, which are called **observables** or **measurements** on $\mathcal{M}$. For instance, if x were a participant in the GRYD program, then a measurement of x, $g(x)$ could be their answer to one of the 56 questions from the questionnaire. The advantage of looking at this space s that we can easily utilize the Koopman Operator Theory.

The theory provides us with the Koopman Operator, denoted by $\mathcal{K}$, which is linear, infinite-dimensional operator that pushes scalar functions on $\mathcal{M}$ through time. In other words, the following equation holds:

$$\mathcal{K} \circ g = g \circ f. \quad [4]$$

We observe that while f might be non-linear, the operator $\mathcal{K}$ is a linear operator acting on scalar functions on $\mathcal{M}$, which can be seen easily from Equation 4:

$$\begin{aligned} \mathcal{K}[(g_1 + g_2)(x)] &= (g_1 + g_2)(f(x)) \\ &= g_1(f(x)) + g_2(f(x)) \\ &= \mathcal{K}(g_1(x)) + \mathcal{K}(g_2(x)), \end{aligned} \quad [5]$$

for all $g_1, g_2 \in S$.

Therefore, the Koopman operator represents the dynamics in the system, which we can analyze using eigenvalues and eigenfunctions of $\mathcal{K}$ (1, 2). However, since $\mathcal{K}$ is infinite-dimensional, it cannot be solved directly in practice. Thus, we rely on Dynamic Mode Decomposition to find eigenvalues and eigenmodes that are approximates of the Koopman eigenvalues and eigenmodes (3).

3. Dynamic Mode Decomposition (DMD)

Dynamic Mode Decomposition (DMD) is a method that studies the growth or decay and periodicity in a dynamical system. It is computationally favored on large datasets because it projects the system onto a lower dimension.

A. Algorithm. For a dynamical system $AX = Y$, where X and Y are known $m \times n$ matrices and A is an unknown $m \times m$ matrix, A can be solved by the Least Square Solution:

$$A = YX^+. \tag{6}$$

However, when the number of rows in X and Y, m, is too large, as in the problems of fluid flows, it is computationally expensive to calculate A directly. In such cases, we project X and Y onto some lower dimensional spaces using the Singular-Value Decomposition (SVD). We compute the rank-reduced SVD of X as:

$$X = U\Sigma V^\top, \tag{7}$$

and then

$$A = YV\Sigma^{-1}U^\top \tag{8}$$

DMD considers a lower rank approximation of the original system

$$\tilde{A} = U^\top AU = U^\top YV\Sigma^{-1}. \tag{9}$$

Then eigenvalues and eigenvectors of A can be approximated by eigenvalues and eigenvectors of $\tilde{A}$:

Suppose λ is an eigenvalue of $\tilde{A}$ with eigenvector v, $\tilde{A}v = \lambda v$, then we have

$$\begin{aligned} A(YV\Sigma^{-1}v) &= YV\Sigma^{-1}U^\top YV\Sigma^{-1}v \\ &= YV\Sigma^{-1}\tilde{A}v \\ &= YV\Sigma^{-1}(\lambda v) \\ &= \lambda(YV\Sigma^{-1}v), \end{aligned} \tag{10}$$

so λ is also an eigenvalue of A with eigenvector $YV\Sigma^{-1}v$.

It turns out that A is actually a finite approximation of the infinite-dimensional Koopman Operator, as explained in (3): Let $\mathcal{K}$ be the Koopman Operator in $L^2(\mathcal{M})$, and its eigenvalues and eigenfunctions be denoted by λ_j and θ_j such that

$$K \circ \theta_j(x) = \lambda_j \theta_j(x), \text{ for } j = 0, 1, 2, ..., \text{ and } x \in \mathcal{M}. \tag{11}$$

For a given scalar function $g \in L^2(\mathcal{M})$, suppose also that $g \in \text{Span}\{\theta_j : j = 0, 1, 2, ...\}$, then

$$g(x) = \sum_{j=0}^{\infty} c_j \theta_j(x), \text{ for some } c_j \in \mathbb{R} \tag{12}$$

Then by Equation 4:

$$g(f(x)) = \mathcal{K}(g(x)) = \mathcal{K}(\sum_{j=0}^{\infty} c_j \theta_j(x)) = \sum_{j=0}^{\infty} c_j(\mathcal{K} \circ \theta_j(x)) = \sum_{j=0}^{\infty} \lambda_j c_j \theta_j(x), \tag{13}$$

And it is easy to see that

$$\mathcal{K}^{(l)}(g(x)) = \sum_{j=0}^{\infty} \lambda_j^l c_j \theta_j(x), \text{ for } l = 1, 2, 3, ..., \tag{14}$$

And for matrices X and Y whose columns are represented by $(x_1, x_2, ..., x_k, ..., x_n)$ and $(y_1, y_2, ..., y_k, ..., y_n)$, respectively, whenever we calculate A by Equation 6 and eigendecompose it to obtain its eigenvalues b_j and eigenvectors ϕ_j, for $j = 0, 1, 2, ...$, we can write

$$x_k = \sum_{j=0}^{n-1} c_{jk} \phi_j, \text{ for some } c_{jk} \in \mathbb{R}, \tag{15}$$

whenever X has full rank.

Comparing Equation 15 to Equation 13, we see that A is a finite approximation of $\mathcal{K}$.

Algorithm 1 Exact DMD

Input: initial matrix X, final matrix Y, rank r
Output: eigenvalues λ, eigenvectors ϕ
Compute rank-r SVD of $X = U_r\Sigma_r V_r^\top$.
Define $\tilde{A} = U_r^\top YV_r\Sigma_r^{-1}$.
Compute eigenvalues and eigenvectors of $\tilde{A}$ such that $\tilde{A}w_i = \lambda_i w_i$, then λ_is are DMD eigenvalues.
DMD eigenvectors (eigenmodes) are $\phi_i = YV_r\Sigma_r^{-1}w_i$, for $i = 1, 2, ..., r$.

B. Interpreting Eigenvalue. DMD eigenvalues provide information of how observables change after A. Suppose v is an eigenvector of A with eigenvalue λ, its change is expressed as $Av = \lambda v$. Let

$$\lambda = \rho e^{i\omega}, \rho \geq 0, \omega \in \mathbb{R}.$$

Then ρ is the magnitude of the λ, which represents growth or decay in terms of how much the change in magnitude of vector v after applied to A. ω represents frequency as it is the change in angle of vector v. Therefore, it is reasonable to instead compute log eigenvalues and analyze their real and imaginary component in terms of growth and frequency.

The system is considered stable if matrix A has an eigenvalue with a magnitude 1, or equivalently a log eigenvalue with a real component of 0. This means that the magnitude of the vector remains unchanged after going through the system. Otherwise, if the real component of log eigenvalue is negative, the system decays; and if the real component of log eigenvalue of positive, the system is chaotic.

C. Selecting Dominant Eigenvalue-Eigenvector Pair. Given an arbitrary point x, it can be expressed in terms of DMD eigenvectors $x = \sum_{i=1}^{r} c_i\phi_i + C$, where C is the remaining part perpendicular to the eigenbasis. Then,

$$Ax = A(\sum_i c_i\phi_i + C) \approx \sum_i c_i\lambda_i\phi_i$$

since AC tends to be small, and $A^k x \approx \sum_i c_i\lambda_i^k\phi_i$. When k is large, the value is dominated by the term $c_1\lambda_1^k\phi_1$ where λ_1 is the largest DMD eigenvalue. Therefore, we can select most dominant eigenvalue-eigenvector pairs according to the magnitude of $\lambda^k|\phi|$. The convention is to choose k to be the time step in the system (4), so in the case of GRYD dataset where we only look at one step forward, we choose $k = 1$ and select dominant eigenvalue-eigenvector pairs by $\lambda|\phi|$.

4. Understanding DMD on Synthetic Data

A. Dynamics Controlled Experiment. We construct synthetic systems that we are sure of their evolutions and examine the most dominant eigenvalue computed by DMD. At each iteration, we take the following steps:

- Take Q_M to be a random sample with replacement of 100 questionnaire responses from dataset that has a risk factor of 3, 4, or 5;
- Take Q_L (or Q_H) to be a random sample with replacement of 100 questionnaire responses from dataset that has a risk factor of 0 (or 9);
- Let $X = Q_M$ and $Y_t = (1-t)Q_M + tQ_L$ (or Q_H);
- Perform DMD on system $S_t : X \to Y_t$ for $t \in [0, 1]$ and record most dominant eigenvalue λ_t;

In doing so, we control the growth/decay of the synthetic system and study how the dominant eigenvalues from DMD change in response. Figure S1 shows values in terms of $\log(\lambda)$ on the complex plane.

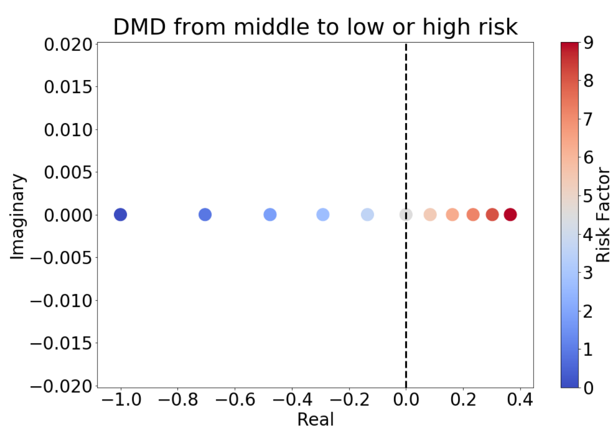
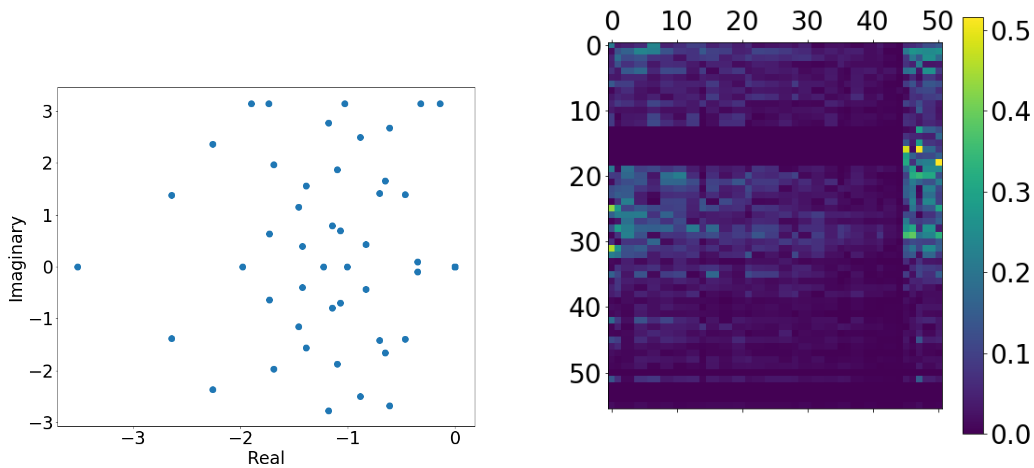


Fig. S1. Positions of log eigenvalue on complex plane for systems that go from middle risk to low or high risk. Blue dots represent systems that decays to a lower risk, and red dots represent systems that grows.

We see that the real component of log eigenvalue is correlated to growth or decay in the system. In particular, when the system is going from mid-risk to zero-risk, the value is -1. As values in Y increase, the growth or decay rate of the system increases. Therefore, the largest eigenvalue from DMD is effective in revealing growth/decay in the system.

B. Segmented Dynamics Experiment. Consider a system where a subset of responses do not change with time. We are interested in how DMD captures such behavior. Theoretically, suppose in system $AX = Y$, row k in the matrix does not change, i.e. $(Ax)_k = x_k \forall x$, then the system should have an eigenvalue $\lambda = 1$ with unit eigenvector $e_k = (0, ..., 0, 1, 0, ..., 0)^T$.

To verify, we construct a synthetic decaying system from dataset similar to previous experiment, and replace row 13-18 in matrix Y to be same as matrix X. Therefore, in this system, responses of question 13-18 remain unchanged between takes of questionnaire. The DMD result of this system is shown in Figure S2.



(a) Log eigenvalues on complex plane.

(b) Eigenvector matrix colored by the magnitude of the entries.

Fig. S2. Eigenvalues and eigenvectors in a decaying system where question 13 to 18 are fixed.

Though it is hard to see from the scatter plot, DMD gives 6 eigenvalues close to 1, which is consistent with the number of questions fixed constant. However, eigenvectors corresponding to these eigenvalues are not unit vectors. Despite having more weights on fixed questions, the eigenvectors also contain non-zero weights on other questions. On the other hand, fixed questions have no weights in eigenvectors with decay.

References

1. Rowley CW, Mezić I, Bagheri S, Schlatter P, Henningson DS (2009) Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics* 641:115–127.
2. Kutz JN, Brunton SL, Brunton BW, Proctor JL (2016) *Dynamic mode decomposition: data-driven modeling of complex systems.* (SIAM) Vol. 149.
3. Tu JH, Rowley CW, Luchtenburg DM, Brunton SL, Kutz JN (2013) On dynamic mode decomposition: theory and applications. *arXiv preprint arXiv:1312.0041.*
4. Proctor JL, Eckhoff PA (2015) Discovering dynamic patterns from infectious disease data using dynamic mode decomposition. *International Health* 7(2):139–145.