A NOTE ON MARKOV NORMALIZED MAGNETIC EIGENMAPS

ALEXANDER CLONINGER

ABSTRACT. We note that building a magnetic Laplacian from the Markov transition matrix, rather than the graph adjacency matrix, yields several benefits for the magnetic eigenmaps algorithm. The two largest benefits are that the embedding becomes more stable as a function of the rotation parameter g, and the principal eigenvector of the magnetic Laplacian now converges to the page rank of the network as a function of diffusion time. We show empirically that this normalization improves the phase and real/imaginary embeddings of the low-frequency eigenvectors of the magnetic Laplacian.

1. INTRODUCTION

We consider the problem of building a diffusion operator on directed networks and graphs with adjacency matrix $W \in \mathbb{R}^{N \times N}$ that parallels Diffusion Maps on undirected networks [1]. A natural approach to this is to build the Magnetic Eigenmaps Laplacian [2],

$$L = D - T \circ W^{(s)},$$

where $(T \circ W^{(s)})_{i,j} = e^{2\pi ig(W_{j,i} - W_{i,j})} \frac{W_{j,i} + W_{i,j}}{2}$ and $D_{i,i} = \sum_j W_{i,j}^{(s)}$. This method assigns a complex rotation to asymmetric links. In this note, we observe that Markov normalization of the adjacency matrix allows us to introduce the notion of diffusion time steps, which is capable of capturing the evolution of the process. Markov normalization of the graph adjacency matrix yields several benefits:

1) normalizing the complex rotation matrix $e^{2\pi ig(W_{j,i} - W_{i,j})}$ by the density of the node, which emphasizes absorbing states and converges to the page rank of the graph for large time [3],

2) stabilizing the phase embedding with respect to choice of the rotation parameter $g$,

3) separating the absorbing states into the first eigenvector and recurrent states into subsequent eigenvectors, and

4) introducing diffusion time, which allows for the study of multi-step neighborhoods.

2. MARKOV PROPERTY AND DIFFUSION TIME

An important aspect of diffusion maps is the ability to separate the statistics of the sampling distribution from the geometry of the underlying manifold [4]. This is accomplished...
A NOTE ON MARKOV NORMALIZED MAGNETIC EIGENMAPS

by normalizing the affinity between points by the overall transition probability,

\[ P = T^{-1}W, \quad \text{where} \quad T_{x,y} = \begin{cases} \sum_y W(x,y), & x = y \\ 0, & x \neq y. \end{cases} \]

This also allows us to consider the properties of the multiple time step diffusion process \( P^t \) for \( t \in \mathbb{R}^+ \). For a symmetric process (i.e. undirected graph), the embedding of this process remains stable. This is because, if the eigendecomposition is \( P = \Phi \Lambda \Phi^* \), then

\[ P^t = (\Phi \Lambda \Phi^*)^t = \Phi \Lambda^t \Phi^*. \]

However, for a non-symmetric graph, this property no longer holds. Thus the eigenvectors of \( P^t \) will exhibit time dependence beyond a simple multiplicity factor.

The markov normalized magnetic Laplacian then takes the form of

\[ L_{i,j}^{(t)} = D_{i,j} - e^{2\pi i (P^t_{i,j} - P^t_{j,i})} \frac{P^t_{i,j} + P^t_{j,i}}{2}, \quad \text{where} \quad D_{i,i} = \frac{1}{2} \sum_j (P^t_{i,j} + P^t_{j,i}). \]

Adjusting the weight matrix to be Markov normalized has several implications on the magnetic Laplacian. In diffusion on a symmetric graph, normalizing by the node density causes the principal eigenvector \( \phi_0 \) to be a constant across all nodes and is thus trivial. However, when the network is not symmetric, the principal eigenvector takes on a much more important role. Namely, the principal eigenvector recovers the degree of asymmetry between nodes. This implies that, as \( t \) increases, the phase of the eigenvectors converges to the page rank of the network.

**Corollary 1.** In the limit as \( t \to \infty \), the principal eigenvector of \( L = \lim_{t \to \infty} L^{(t)} \) (denoted \( \phi \)) converge to

\[ \phi_{i,g}^{(t)} = e^{2\pi i h_i}, \]

where \( h = \text{pagerank}(W) \) with no dampening factor.

**Proof.** From Theorem 1 in [2], we know that the magnetic Laplacian \( L \) has a zero eigenvalue iff \( \exists h \) such that \( a_{i,j} = h_j - h_i \). Furthermore, we know that, in this case, \( \phi_{i,g}^{(t)} = e^{2\pi i h_i} \phi_{i,0}^{(t)} \).

We also know that \( P^t_{i,j} = h \) as \( t \to \infty \), where \( h = \text{pagerank}(W) \) with no dampening factor. That means \( a_{i,j} = P^t_{i,j} - P^t_{j,i} = h_j - h_i \).

Moreover, because \( W \) is row stochastic, the principal eigenvector is the ones vector \( 1 \). Thus, the principal eigenvector of the markov normalized magnetic Laplacian converge to the ones vector rotated into the complex plane by a factor proportional to the page rank of the nodes.

This only applies in the limit as \( t \to \infty \). The embeddings of \( L^{(t)} \) for small \( t \), and even \( t = 1 \), yield valuable properties due to the Markov normalization. We show this empirically in Section 3.
3. Examples

In all examples of the Markov normalized process, we will refer to a value of $g \in [0, 1/2)$ that is on the same scale as for the unnormalized Laplacian. However, due to the weights being normalized to sum to 1, they are actually on a different scale than the unnormalized weights matrix. For that reason, we set $g \leftarrow g/\max(\max(P))$ for the Markov normalized weights so that the degree of rotation is on the same order of magnitude for both methods.

3.1. Three Cluster Example. We begin with the three cluster example from [2]. This example is particularly simple given that the complex rotation of three clusters will, more often than not, keep the clusters separated. However, for small values of $g$, rotation of the unnormalized asymmetric weights matrix yields insufficient separation between the clusters, as shown in Figure 1. In fact, Figure 2 shows the percentage of points that cluster correctly across 100 uniform random small values of $g$ ($g < 0.25$), where clustering was done by k-means with $k = 3$.

![Unnormalized Embedding](image1)

![Unnormalized Phase](image2)

![Unnormalized Phase vs Page Rank](image3)

![Markov Normalized Embedding](image4)

![Markov Normalized Phase](image5)

![Markov Normalized Phase vs Page Rank](image6)

**Figure 1.** Three clusters with $P(\text{inCluster}) = 0.5$, $P(\text{outCluster}) = 0.5$, $P(\text{rotateClockwise}) = 0.9$, and with parameter $g = 0.04$. The markov normalized embeddings are with $t = 1$, though the bottom right page rank plot is for $t = 4$ to allow sufficient time to converge.

By incorporating the notion of diffusion time, we are also able to observe the time dynamics of this process. Figure 7 in the appendix shows the embedding for each $t \in \{1, 2, ..., 9\}$. Notice that after 3 and 4 time steps, the clusters become momentarily blurred. This is because, by that time, most of the energy from cluster 1 will have moved to clusters 2 and 3, etc. However, after that momentary blurring, the process stabilizes and we continue to see the three clusters rotating about each other for larger $t$. This tells us that there is a global mixing time of $t > 1$ for the process to satisfy $P_{i,j} > \epsilon$ for all $i, j$. 
3.2. Circle with Drift. We consider data distributed on a unit circle, with drift in the counterclockwise direction. This is an interesting process as it has a very large mixing time. We define the affinity between points to be

\[
k(x, y) = \begin{cases} 
e^{-\|x-y\|^2/5\sigma^2}, & \angle xy \geq 0 \\ e^{-\|x-y\|^2/\sigma^2}, & \angle xy < 0 \end{cases}
\]

where \(\angle xy\) measures the angle between the points. The affinity matrix is seen in Figure 8 in the appendix.

Figure 3 compares the different normalizations and eigenvectors for the circle with drift. By normalizing the weights matrix, the first eigenvector is once again the trivial eigenvector. Subsequent eigenvectors recover the geometry of the circle.

The higher order eigenvectors also contain useful information, as in diffusion maps. They mimic the sines and cosines of the angle around the circle, as the Fourier functions \(\{e^{inx}\}\) are eigenfunctions of the Laplacian on the circle. We wish to recover these functions independent of the sampling density on the circle. Figure 9 in the appendix shows that the Markov normalized weights matrix gives a closer approximation of the sin multiples on the circle that comes from the unnormalized weights matrix. Note that the data used here is not uniformly sampled across the circle.

3.3. Bow Tie Example. Another issue with not including a diffusion time is the need for \(g\) to implicitly assume the number of clusters. By setting \(g\) to be \(1/k\) for \(k\) clusters as the authors suggest, the complex rotation can lose valuable information about transitions that aren’t \(k\) periodic. We demonstrate this with an example similar to the three cluster example from Section 3.1. We build a three cluster cycle among clusters 1, 2, and 3, and a 5 cluster cycle among clusters 1, 4, 5, 6, and 7. By normalizing the weights, we see in Figure...
that the non-principal eigenvectors capture the geometry of the process. Moreover, we
discount the fact that cluster 1 appears in both cycles, and will end with twice the energy
of the other clusters for large $t$.

We are also able to observe the time dynamics of this process. Figure 10 in the appendix
shows the embedding for each $t \in \{1, 2, ..., 9\}$. Notice that after 7 and 8 time steps, the
clusters become momentarily blurred. This is because, by that time, most of the energy
from one end of the bow tie has moved to the other. However, after that momentary
blurring, the process stabilizes and we continue to see the three clusters rotating about
each other for larger $t$. This tells us that there is a global mixing time of $t > 6$ for the
process to satisfy $P_{i,j} > \epsilon$ for all $i, j$. We can also see these dynamics in the affinity matrix
in Figure 11. And finally, we see in Figure 11 that after a large $t$, the phase of the leading
eigenvector converges to the page rank of the process.

3.4. Hidden Circle. As a final example, we consider a mix of absorbing and recurrent
states. The data we consider is a unit square with drift from left to right via

$$k(x, y) = \begin{cases} 
    e^{-\|x-y\|^2/3\sigma^2}, & x_1 < y_1 \\
    e^{-\|x-y\|^2/\sigma^2}, & x_1 \geq y_1
\end{cases}$$

On top of that, we add an annulus in the middle of the unit square with counterclockwise
flow. The results of the phase and embedding are in Figure 5.
Figure 4. Bow tie clusters with $P(\text{inCluster}) = 0.5$, $P(\text{outCluster}) = 0.5$, $P(\text{rotateClockwise}) = 0.9$, and with $g = 0.04$. For Markov normalization, we take $t = 1$.

Figure 5. $g = 0.24$, $t = 4$
Also, by not normalizing the phase rotation, low density points become separated in the phase space (see Figure 6). This leads to discontinuity in the torus projection of the phase, because the top and bottom edges of the square do not have much energy coming in.

![Unnormalized Phase Torus](image1)

![Markov Normalized Phase Torus](image2)

**Figure 6.** $g = 0.24, t = 1$

4. Conclusion

We note that Markov normalization of the graph adjacency matrix yields several benefits. Normalization of the complex rotation matrix $T$ converges to the page rank of the graph for a large time, as in Section 3.1. The Markov normalization also stabilizes the phase embedding over a larger choice of $g$, as in most examples in Section 3. By emphasizing the absorbing states in the first eigenvector, this normalization separates the recurrent states into subsequent eigenvectors as in Section 3.4. And finally, we can introduce diffusion time into the process and observe the multi-step neighborhood dynamics of the graph, as in most examples in Section 3.

References

[1] Coifman, Ronald R., and Stéphane Lafon. "Diffusion maps." Applied and computational harmonic analysis 21.1 (2006): 5-30.

[2] Fanuel, Michal, et al. Magnetic Eigenmaps for Visualization of Directed Networks. arXiv preprint arXiv:1606.08266 (2016).

[3] Page, Lawrence, et al. The PageRank citation ranking: bringing order to the web. (1999).

[4] Lafon, Stéphane S. Diffusion maps and geometric harmonics. Diss. Yale University, 2004.
Figure 7. Three clusters with $P(\text{inCluster}) = 0.5$, $P(\text{outCluster}) = 0.5$, $P(\text{rotateClockwise}) = 0.9$, and with $g = 1/4$. Evolve cluster from $t \in \{1, ..., 9\}$. 
Figure 8. Affinity matrix for circle with drift

Figure 9. Circle with drift, colored by corresponding real part of real($\phi_i$).
Top: Markov normalized, Bottom: Unnormalized. The eigenvectors correspond to multiples of sinusoids, the eigenvectors of the DFT matrix.
Figure 10. Bow tie clusters with $P(\text{inCluster}) = 0.5$, $P(\text{outCluster}) = 0.5$, $P(\text{rotateClockwise}) = 0.9$, and with $g = 1/4$. Evolve cluster from $t \in \{1, \ldots, 9\}$.

Figure 11. Bow tie clusters affinity at $t = 7$ (left), and bow tie clusters phase($V_0$) at $t = 10$ (right).