ARTICLE IN PRESS
YACHA:1235

Appl. Comput. Harmon. Anal. ••• (••••) •••–•••

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Letter to the Editor

Time coupled diffusion maps

Nicholas F. Marshall [a,*], Matthew J. Hirn [b]

[a] *Department of Mathematics, Yale University, New Haven, CT 06520, USA*
[b] *Department of Computational Mathematics, Science & Engineering and Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA*

A R T I C L E   I N F O

Article history:
Received 30 August 2016
Received in revised form 3 November 2017
Accepted 4 November 2017
Available online xxxx
Communicated by Naoki Saito

Keywords:
Manifold learning
Dimensionality reduction
Diffusion distance
Heat equation
Time-dependent metric

A B S T R A C T

We consider a collection of n points in $\mathbb{R}^d$ measured at m times, which are encoded in an $n \times d \times m$ data tensor. Our objective is to define a single embedding of the n points into Euclidean space which summarizes the geometry as described by the data tensor. In the case of a fixed data set, diffusion maps and related graph Laplacian methods define such an embedding via the eigenfunctions of a diffusion operator constructed on the data. Given a sequence of m measurements of n points, we introduce the notion of time coupled diffusion maps which have natural geometric and probabilistic interpretations. To frame our method in the context of manifold learning, we model evolving data as samples from an underlying manifold with a time-dependent metric, and we describe a connection of our method to the heat equation on such a manifold.

1. Introduction

In many machine learning and signal processing tasks, the observable data is high dimensional, but it lies on a low-dimensional intrinsic manifold. In recent years, several manifold learning methods have emerged which attempt to recover the intrinsic manifold underlying datasets. In particular, graph Laplacian methods have become popular due to their practicality and theoretical guarantees [1–8].

Current graph Laplacian methods implicitly assume a static intrinsic manifold, or equivalently, that the dynamics underlying the data generation process are stationary. For many applications, this stationary assumption is justified, as datasets often consist of a single snapshot of a system, or are recorded over small time windows. However, in the case where data is accumulated over longer periods of time, accounting for changing dynamics may be advantageous. Furthermore, if a system is particularly noisy, combining a large number of snapshots over time may help recover structure hidden in noise. These observations raise the

* Corresponding author.
E-mail addresses: nicholas.marshall@yale.edu (N.F. Marshall), mhirn@msu.edu (M.J. Hirn).

https://doi.org/10.1016/j.acha.2017.11.003
1063-5203/

Please cite this article in press as: N.F. Marshall, M.J. Hirn, Time coupled diffusion maps, Appl. Comput. Harmon. Anal. (2017), https://doi.org/10.1016/j.acha.2017.11.003

following question: how can graph Laplacian methods be extended to account for changing dynamics while maintaining theoretical guarantees?

In this paper, we propose modeling data with changing dynamics by assuming there exists an underlying intrinsic manifold with a time-dependent metric. We will describe the proposed method using the diffusion maps framework: a popular graph Laplacian framework which is robust to non-uniform sampling [4]. We remark that diffusion maps are highly related to other manifold learning methods such as Laplacian eigenmaps and spectral clustering. In fact, if data is uniformly sampled from the underlying manifold, diffusion maps [4] is essentially eigenvalue weighted Laplacian eigenmaps [9].

Although we assume points on the intrinsic manifold are fixed, their geometry, i.e., dependence structure, is allowed change. We can conceptualize samples from a manifold with a time-dependent metric by considering a corresponding point cloud smoothly moving through $\mathbb{R}^d$ produced by isometrically embedding the manifold over time. The evolution of the metric dictates the movement of points, and vice versa. In practice, datasets conforming to this model are commonly encountered, e.g., an RGB video feed consists of a collection of n pixels which move through $\mathbb{R}^3$.

In general, we consider data consisting of a collection of n points in $\mathbb{R}^d$ measured at m times encoded in an $n \times d \times m$ data tensor X. The tensor X can be expressed as a sequence $(X_1, \ldots, X_m)$ of $n \times d$ matrices whose entries correspond across the sequence. Given such as sequence $(X_1, \ldots, X_m)$, the time coupled diffusion map framework introduced in this paper is based on the product operator:

$$\mathbf{P}^{(m)} = \mathbf{P}_m \mathbf{P}_{m-1} \cdots \mathbf{P}_2 \mathbf{P}_1,$$

where each $\mathbf{P}_i$ is a diffusion operator constructed from X_i. We will show that this discrete diffusion process, which is formally defined in the following section, approximates a continuous diffusion process on an assumed underlying manifold with a time-dependent metric. Additionally, we introduce the notion of time coupled diffusion maps, named thus because the time evolution of the data has been coupled to the time evolution of a diffusion process.

1.1. Related works

In the diffusion geometry literature, several techniques have been developed, which also utilize multiple diffusion kernels for a variety of objectives including: iteratively refining the representation of data, facilitating comparison, and combing multiple measurements of a fixed system.

An early example of a multiple kernel method is the denoising algorithm of Szlam, Maggioni, and Coifman [10], which iteratively smooths an image via an anisotropic diffusion process. That is, the algorithm switches between constructing a diffusion kernel on a given data set (in this case an image), and applying the constructed kernel to the data:

$$X_i \to \mathbf{P}_i, \quad X_{i+1} = \mathbf{P}_i X_i$$

where the arrow denotes that $\mathbf{P}_i$ is constructed based on X_i. More recently, in [11] Welp, Wolf, Hirn, and Krishnaswamy introduce an iterative diffusion based construction, which acts to course grain data. From a theoretical perspective, both of these methods can be considered in the context of the time coupled diffusion framework introduced in this paper.

In [12] Wang, Jiang, Wang, Zhou, and Tu introduce the notion of Cross Diffusion as a metric fusion algorithm with applications to image processing. They demonstrate how multiple metrics on a given data set can be combined by considering the iterative cross diffusion

$$\mathbf{P}_1^{(t+1)} = \mathbf{P}_1 \mathbf{P}_2^{(t)} \mathbf{P}_1^T, \quad \text{and} \quad \mathbf{P}_2^{(t+1)} = \mathbf{P}_2 \mathbf{P}_1^{(t)} \mathbf{P}_2^T,$$

Please cite this article in press as: N.F. Marshall, M.J. Hirn, Time coupled diffusion maps, Appl. Comput. Harmon. Anal. (2017), https://doi.org/10.1016/j.acha.2017.11.003