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

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Learning the geometry of common latent variables using alternating-diffusion

Roy R. Lederman [a,*], Ronen Talmon [b]

[a] Department of Mathematics, Yale University, New Haven 06520, CT, United States
[b] Department of Electrical Engineering, Technion (IIT), Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 8 November 2014
Accepted 3 September 2015
Available online xxxx
Communicated by Ronald R. Coifman

Keywords:
Common variable
Alternating-diffusion
Diffusion-maps
Multi-view
Multimodal analysis

ABSTRACT

One of the challenges in data analysis is to distinguish between different sources of variability manifested in data. In this paper, we consider the case of multiple sensors measuring the same physical phenomenon, such that the properties of the physical phenomenon are manifested as a hidden common source of variability (which we would like to extract), while each sensor has its own sensor-specific effects (hidden variables which we would like to suppress); the relations between the measurements and the hidden variables are unknown. We present a data-driven method based on alternating products of diffusion operators and show that it extracts the common source of variability. Moreover, we show that it extracts the common source of variability in a multi-sensor experiment as if it were a standard manifold learning algorithm used to analyze a simple single-sensor experiment, in which the common source of variability is the *only* source of variability.

1. Introduction

Measurement systems typically have many sources of variability. When multiple sensors are used to measure the same physical phenomenon, some sources of variability are related to the actual physical phenomenon, whereas other sources of variability are irrelevant, sensor-specific effects. In this case, extracting the common source of variability and discarding the sensor-specific sources may uncover the essence of the data, separating the relevant information from the irrelevant information.

The motivation for this work arises from applications of exploratory data analysis in areas such as complex biological systems, neural systems and biomedical devices. In these problems, the sensor-related sources of variability are not necessarily restricted to noise and interferences that can often be suppressed by averaging, but also include variables such as the position and orientation of a sensor, environmental effects, channel characteristics, and local activity in different components of the complex system. The different components

* Corresponding author.
E-mail addresses: roy@aya.yale.edu (R.R. Lederman), ronen@ee.technion.ac.il (R. Talmon).

http://dx.doi.org/10.1016/j.acha.2015.09.002
1063-5203/

Please cite this article in press as: R.R. Lederman, R. Talmon, Learning the geometry of common latent variables using alternating-diffusion, Appl. Comput. Harmon. Anal. (2015), http://dx.doi.org/10.1016/j.acha.2015.09.002

of complex systems like these can be modeled as different “sensors” and the interaction between them can be modeled as a common variable; so the term “sensor” used in this paper is interpreted more generally than simply a physical measurement device.

Unsupervised Manifold Learning is a class of nonlinear data-driven methods, e.g. ISOMAP [1], locally linear embedding (LLE) [2], Hessian Maps [3], and Laplacian Eigenmaps [4], often used to extract the sources of variability in given data sets. Of particular interest in the context of this paper is Diffusion Geometry [5–9], a manifold learning framework, in which discrete diffusion processes are constructed on the given data points; these diffusion processes are designed to capture the geometry of the sources of variability. In the case of multiple sensors, despite having more information, adding sensors adds sources of variability, making it more difficult to extract the common source of variability. Various methods have been proposed to analyze data from multiple sensors within the framework of Manifold Learning. One approach is to concatenate the vectors representing the data into one vector [10], but in this case it is not clear how the data from each sensor should be scaled, especially if the sensors are of very different nature. To address this challenge, it has been proposed in [11] to use Diffusion Maps to obtain a low-dimensional “standardized” representation of data from each sensor, and then to concatenate the low dimensional representations. However, these methods aggregate all sources of variability from all sensors, and they neither distinguish the common variable nor discard the sensor-specific variables.

A classic approach designed to extract the common source of variability from two sensors is Canonical Correlation Analysis (CCA) [12], which recovers highly correlated linear projections in linear systems, but has limited applicability to non-linear problems. Kernel CCA (KCCA), the generalization of CCA to the kernel feature space (e.g. [13,14]), treats some aspects of nonlinearity, but it relies on inversion of covariance matrices, an operation that raises statistical and numerical issues in applications. Another related method [15] also assumes certain linearities in the problem.

In the context of supervised learning, alternating conditional expectations (ACE) has been proposed for regression analysis [16], and linear combinations of kernels have been the subject of considerable work on multi-kernel learning (e.g. [17]). In this paper, we consider an unsupervised setting where no examples of the unknown hidden variables are available, and we propose a data-driven method based on an alternating product of diffusion operators. In the field of multi-view problems, there has been ample work based on various manipulation or combination of operators. For example, several approaches for metric-fusion, clustering and classification, have been proposed, which rely on various manipulation of affinity matrices (e.g. [18–21]), Markov and diffusion matrices (e.g. [22,23]), graph Laplacians (e.g. [24,25]) and sets of nearest neighbors (e.g. [26,27]). Tensor products of Markov matrices have also been considered in [28], and products of Markov matrices and their transposes have been presented in [29]. A recent work on products of kernels extends diffusion maps for the case where the same underlying process is observed using several different modalities [30,31]. Additional aspects of multimodal data fusion are discussed in [32].

The main contribution of this paper is: 1) the presentation of an unsupervised method based on alternating diffusion, and 2) a theoretical analysis of alternating diffusion showing that it recovers the common variable while it discards the sensor-specific variables. More specifically, we show that the common source of variability is extracted by this method from multiple sensors as if it were the *only source of variability* in a single sensor, extracted by a manifold learning method.

The analysis of the method distinguishes between two types of objects: *observable* objects, which are quantities that can be approximated based on the measurements (following the standard practice in manifold learning), and *hidden* objects, which are not approximated/accessible directly. We discuss a hidden effective diffusion process based on hidden objects and use it to develop a manifold learning method for extracting the geometry of the latent common variable. While the hidden effective diffusion is merely a formal object that is not accessible directly, our method is based on observables and builds a geometry equivalent to the geometry of the hidden effective diffusion.

Please cite this article in press as: R.R. Lederman, R. Talmon, Learning the geometry of common latent variables using alternating-diffusion, Appl. Comput. Harmon. Anal. (2015), http://dx.doi.org/10.1016/j.acha.2015.09.002