In many fields including economics, collection of time series such as stocks or energy prices are governed by a similar non-linear dynamical process. These time series are often measured hourly, thus, each day can be viewed as a high-dimensional data point. In this paper, we apply a spectral method, which based on anisotropic diffusion kernels to the model high dimensional electricity price data. We demonstrate the proposed method on price data that was collected from several zones. We show that even though the observed output spaces differ by local spatial influences and noise, the common global parameters that drive the underlying process are extracted.

Modeling zonal electricity prices by anisotropic diffusion embeddings

Neta Rabin and Ronald R. Coifman
Research Report YALEU/DCS/TR-1457
Yale University
June 5, 2012

Approved for public release: distribution is unlimited.
Keywords: *Manifold learning, Time series modeling, Electricity prices*

1 Introduction

Many observed high dimensional time series in science and economics are governed by an underlying non-linear dynamical process. A fundamental task for understanding the observed data is to generate a parametrization that describes the non-linear dynamics that govern the process. In this work we employ a spectral analysis framework, which is based on anisotropic diffusion embeddings, to extract a set of stable independent intrinsic parameters from a number of observed datasets that are governed by the same salient modes.

Electricity price curves from zones that are geographically close, are an example of temporal datasets that are governed by the same set of intrinsic parameters. Although the price curves are gathered from a number of zones, there is a unique set of parameters that govern the observed space and is independent of the zone (station). These parameters capture the physical components such as the season and economical factors, which may depend on the fuel prices and the stability of the market.

Electricity prices are typically treated as a one-dimensional time series and modeled by parametric models such as ARIMA processes [5, 1]. A few recent papers proposed the use of data mining techniques that are based on kernel methods for spike and forecasting analysis [6, 7]. Chen et al. [4] explored manifold learning with local linear embedding for modeling electricity prices. In their model, the daily profile, which consists of 24 hourly price values is considered as a high dimensional data point. LLE then organize the high-dimensional observed data into an embedded space by using the first top embedding coordinates.

In this paper, we follow the above setting by treating the observed data as high-dimensional time series, where each point reside in $\mathbb{R}^{24}$ and holds a daily price profile. We apply a spectral method, which was proposed in [11] by Singer and Coifman for finding the non linear independent components that govern the high dimensional data. The main advantage of the proposed method over the common manifold learning techniques is that the obtained embedding coordinates build an inverse mapping of the observed data into the parametric space. This framework was been recently applied for earth structure classification [3], recovering independent parameters of acoustic channels [10], source localization [9] and non-linear tracking [12].

2 Modeling electricity prices with diffusion kernels

This section describes the diffusion maps framework for organizing and embedding high-dimensional time-series into a low dimensional space. Denote the time series of daily electricity profiles that were collected from the same zone over n consecutive days by

$$\Gamma = \{x_1, x_2, \ldots, x_n\}, \quad x_i \in \mathbb{R}^{24}.$$

Let

$$C = [c_1, c_2, \ldots, c_d]$$

be the physical parameters (such as the daily temperature, fuel prices, etc.) that control the observed space. We assume that there is a non linear map between observed data points $x_i \in \Gamma$ and the physical parameters C.

We first review in Section 2.1 the standard diffusion maps framework, which studies the geometry of the observable manifold and uses the spectral decomposition to embed the data into Euclidean space. Next, we describe in Section 2.2 how to construct an anisotropic diffusion kernel, which is computed based on the geometry of the physical parameters' manifold. Applying this kernel in the diffusion maps framework, embeds the data into Euclidian space based on the distances of the controlling parameters C that effect the observed data Γ.

2.1 Diffusion maps

The diffusion maps framework [2] provide a method for constructing coordinates that parameterize a high-dimensional dataset according to its geometry. The associated diffusion distances are a local preserving metric for this data. The intrinsic geometry of the set Γ is studied by constructing a graph $G = (\Gamma, W)$ with a kernel $W \triangleq w(x_i, x_j)$, which serves as a weight function. A Gaussian kernel $W = e^{\frac{-\|x_i - x_j\|^2}{2\epsilon}}$ is a common choice for a weight function. The kernel is symmetric, positive-preserving and positive semi-definite, as required for the construction of diffusion maps. The normalized kernel

$$P \triangleq p(x_i, x_j) = s(x_i)^{-1} w(x_i, x_j), \quad \text{where} \quad s(x_i) = \sum_j w(x_i, x_j). \tag{1}$$

is a Markov transition matrix. The matrix P consists of nonnegative real numbers, where each row is summed to 1, thus it can be viewed as the probability of moving from x_i to x_j in one time step. The eigendecomposition of the transition matrix P

$$p(x_i, x_j) = \sum_{k \geq 0} \lambda_k \psi_k(x_i) \phi_k(x_j). \tag{2}$$

together with the fact that P is conjugate to a symmetric matrix results in a set of n real eigenvalues $\{\lambda_k\}_{k=1}^{n-1}$ and biorthonormal eigenvectors $\{\psi_k\}$ and $\{\phi_k\}$. A fast decay of the spectrum is achieved by an appropriate choice of ϵ, thus, only a small number of terms are required for sufficient accuracy in the sum (2). The family of diffusion maps $\{\Psi(x_i)\}_{i=1}^{n}$, which are defined by

$$\Psi(x_i) = \left(\lambda_1 \psi_1(x_i), \lambda_2 \psi_2(x_i), \lambda_3 \psi_3(x_i), \cdots\right), \tag{3}$$

embed the dataset into a Euclidean space.

This embedding parameterizes the data into a new space, in which the distances between the data points are determined by the geometric structure of the data. Following the definitions in [8, 2], the diffusion distance between two data points x_i and x_j is the weighted L^2 distance

$$D^2(x_i, x_j) = \sum_{l \in \Gamma} \frac{\left(p(x_i, x_l) - p(x_l, x_j)\right)^2}{\phi_0(x_l)}. \tag{4}$$

In this metric, two data points are close to each other if there exists many paths that connect them. The value of $\frac{1}{\phi_0(x_i)}$ depends on the point's density. Substituting Eq. (2) in

Eq. (4) and using the biorthogonality properties, we obtain that the diffusion distance with the right eigenvectors of the transition matrix P is expressed as

$$D^2(x_i, x_j) = \sum_{k \geq 1} \lambda_k (\psi_k(x_i) - \psi_k(x_j))^2. \tag{5}$$

The Euclidean distance between two points in the embedded space is equivalent to the distances between the points as defined by a random walk.

2.2 Anisotropic diffusion kernels

This section describes the steps for constructing an anisotropic diffusion kernel as was introduced in [11]. Local covariance matrices are used to approximate the Euclidean distance in the physical parameters space instead of the distance of the observable space, which is typically modeled in manifold learning. We assume that there is a non linear map between observed data points in Γ and the physical parameters C.

We begin by computing the local covariance matrices around points of the observed data. For each data point $\{x_i\}_{i=l}^{n}$ the local covariance matrix that is associated with x_i is computed by using the local *time cloud* from the previous L samples $\{x_{i-j}\}_{j=1}^{L}$. This covariance matrix is denoted by $\Sigma_i = Cov(x_i)$. The anisotropic diffusion kernel

$$\bar{W}_{kl} = \frac{\pi}{\sqrt{det(Cov(\frac{x_k+x_l}{2}))}} exp\left(-\frac{(x_k - x_l)^T[\Sigma(x_k) + \Sigma(x_l)]^{-1}(x_k - x_l)}{\epsilon}\right) \tag{6}$$

is constructed. Kushnir et al. [3] showed that by using the metric in (6), the Euclidean distance between the physical parameters, i.e. $\|C_k - C_l\|^2$, is approximated by

$$\|C_k - C_l\|^2 \approx (x_k - x_l)^T[\Sigma_k + \Sigma_l]^{-1}(x_k - x_l). \tag{7}$$

The kernel is normalized like described in Section 2.1 The spectral decomposition of $\bar{W}$ embeds the data in terms of its independent physical parameters.

3 Experimental results

In this section, we demonstrate the proposed method on data gathered in four zones: Capital, Dunwod, Millwd and NPX from New York Independent System Operator. The time series are the LBMP (locational based marginal prices) data of the day-ahead market of the four electricity zones from June 1, 2005 to May 31, 2011. Figure 1 displays two weeks of hourly data from the four stations.

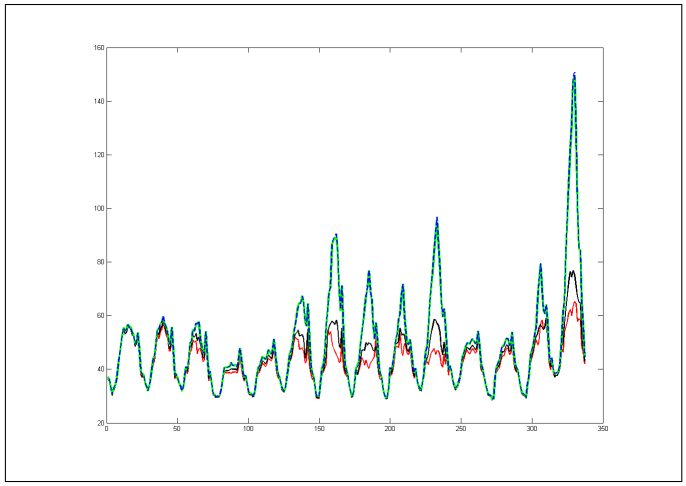

Figure 1: An example of hourly electricity prices gathered from the four zones Capital (red), Millwd (green), NPX (black) and Dunwod (blue) for a period of two weeks.

The application of diffusion maps, as described in Section 2.1 to the high-dimensional price series is presented in Figure 2. The first two diffusion coordinates correlate to the average daily price. The fifth coordinate captures the season.

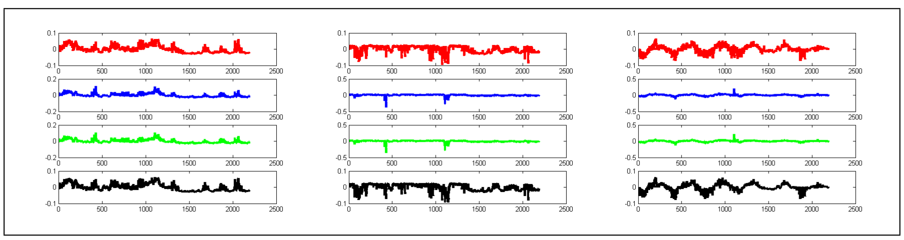

Figure 2: The first, second and fifth diffusion coordinates that embed the prices from Capital zone (red), Dunwod zone (blue) Millwd zone (green) and NPX (black).

The embedding coordinates obtained by applying diffusion maps with the anisotropic diffusion kernels are presented in Fig. 3. These coordinated find a set of consistent controlling parameters which drive the underlying process of the four time series, and are independent of the zone. The first two coordinated may describe some global market behavior. The third embedding coordinate captures the yearly season. The sharp transition in the forth coordinate occurs in the beginning of 2009 and may be related to the financial crisis.

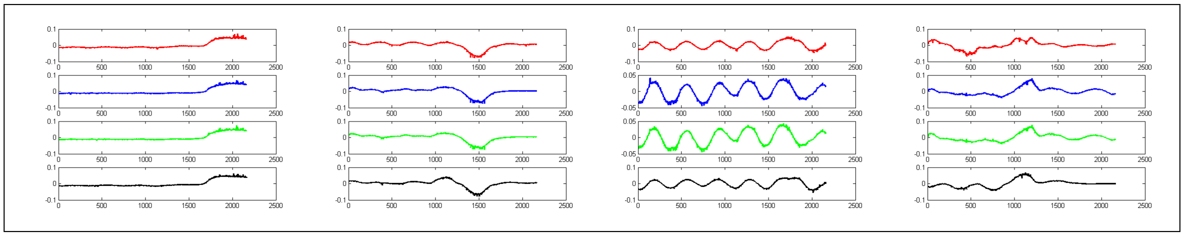

Figure 3: The first four independent embedding coordinates obtained for modeling the data with anisotropic diffusion kernels.

Abnormal spikes are easily detected with this representation. Moreover, application of a standard smoothing filter to the set of embedding coordinates yields a consistent set of salient modes that can be used for analysis tasks such as price forecasting.

References

[1] R. Espnola A. J. Conejo, M. A. Plazas and A. B. Molina. Day-ahead electricity price forecasting using the wavelet transform and arima models. *IEEE Transactions on Power Systems*, 20(2):1035–1042, 2005.

[2] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis: Special issue on Diffusion Maps and Wavelets*, 21:5–30, July 2006.

[3] R. R. Coifman D. Kushnir, A. Haddad. Anisotropic diffusion on sub-manifolds with application to earth structure classification. *Applied and Computational Harmonic Anaysis*, 32(2), 2012.

[4] S. Deng J. Chen and X. Huo. Electricity price curve modeling and forecasting by manifold learning. *IEEE Transactions on Power Systems*, 23, 2008.

[5] F. J. Nogales J. Contreras, R. Espnola and A. J. Conejo. Arima models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, 18(3):1014–1020, 2003.

[6] X. Li J. H. Zhao, Z. Y. Dong and K. P. Wong. A framework for electricity price spike analysis with advanced data mining method. *IEEE Transactions on Power Systems*, 22:376–385, 2007.

[7] X. Li J. H. Zhao, Z. Y. Dong and K. P. Wong. A statistical approach for interval forecasting of the electricity price. *IEEE Transactions on Power Systems*, 23:267–276, 2008.

[8] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis. Diffusion maps, spectral clustering and reaction coordinate of dynamical systems. *Applied and Computational Harmonic Analysis: Special issue on Diffusion Maps and Wavelets*, 21:113–127, July 2006.

[9] I. Cohen R. Talmon and S. Gannot. Supervised source localization using diffusion kernels. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 11 2011.

[10] R. R. Coifman I. Cohen R. Talmon, D. Kushnir and S. Gannot. Parametrization of linear systems using diffusion kernels. *IEEE Trans. Signal Processing*, 60, 3 2012.

[11] A. Singer and R. R. Coifman. Non-linear independent component analysis with diffusion maps. *Applied and Computational Harmonic Analysis*, 25, 2008.

[12] R. Talmon and R. R. Coifman. Differential stochastic sensing: Intrinsic modeling of random time series with applications to nonlinear tracking. *submitted*, 2012.