FRAME BASED KERNEL METHODS FOR AUTOMATIC CLASSIFICATION IN
HYPERSPECTRAL DATA

John Benedetto, Wojciech Czaja, Justin C. Flake, Matthew Hirn

Department of Mathematics
University of Maryland
College Park, MD 20742
jjb@math.umd.edu wojtek@math.umd.edu jcflake@math.umd.edu hirn@math.umd.edu

We propose a new kernel and frame based dimension reducing algorithm by exploiting the synergy between endmembers and kernel based classification schemes. Given a hyperspectral data set \( X = \{x_i\}_{i=1}^N \subseteq \mathbb{R}^D \) consisting of \( N \) pixels in \( D \) dimensions, we propose the following algorithm for processing \( X \): 1.) Landmarking, 2.) Kernel Application, 3.) Out of sample extension, 4.) Endmember selection, 5.) Frame coefficients. Steps 1 and 3 enable the algorithm to run on large data sets. In step 2, we use kernel eigenmap methods to reduce the dimension of the data set \( X \), thus creating a low dimensional data set \( Y = \{y_i\}_{i=1}^N \subseteq \mathbb{R}^d \) that preserves the local geometry of \( X \). \( Y \) consists of \( N \ d \)-dimensional data points, one for each element of \( X \). We assume \( d < D \). Step 4 selects endmembers for the lower dimensional data set \( Y \). Unlike traditional endmember applications in which the the number of endmembers is fewer than the dimension of the data, we select more endmembers than the reduced dimension \( d \). This creates a frame \( \Phi \), for \( \mathbb{R}^d \) by which we can represent the low dimensional data points \( Y \). Frames provide overcomplete representations which gives flexibility in representing mixtures and pure elements. Step 5 computes the frame coefficients of the data points \( Y \) in terms of the endmembers \( \Phi \). There are infinitely many such frame representations - we highlight certain ones that are well suited for classification purposes.

1. LANDMARKING

Our first step is to reduce the complexity of the kernel eigenmap algorithm by selecting a subset of \( X \) on which to compute the kernel. We denote this subset as \( Z = \{z_i\}_{i=1}^n \subseteq X \), where \( n \ll N \). Our current results select the set \( Z \) uniformly at random from the set \( X \). In the future we plan to investigate more systematic ways by which to sample \( X \).

2. KERNEL APPLICATION

Given \( Z \subseteq X \), we construct a kernel for \( Z \). Our results thus far focus on the locally linear embedding (LLE) kernel [2]. The general nature of our framework, though, allows for the use of any kernel eigenmap method, including, e.g., Laplacian eigenmaps [3] or Isomap [4]. We diagonalize the resulting kernel \( K \) and select the \( d \) eigenvectors corresponding to the \( d \) smallest non-zero eigenvalues. Denote the \( j^{\text{th}} \) smallest non-zero eigenvector by \( v_j \), and let the \( i^{\text{th}} \) entry of \( v_j \) be denoted by \( v_j(i) \). The reduced dimension coordinates for the sampled points \( z_i \in Z \) are then given by \( y_i = (v_1(i), v_2(i), \ldots, v_d(i)) \in \mathbb{R}^d \), for all \( i = 1, \ldots, n \).

3. OUT OF SAMPLE EXTENSION

Given the \( n \) low dimensional coordinates \( \{y_i\}_{i=1}^n \) corresponding to the sampled set \( Z = \{z_i\}_{i=1}^n \subseteq X \), we wish to extend these new coordinates to all of \( X \) via an out of sample extension [5]. After a suitable re-indexing of the low dimensional coordinates, we are left with a set \( Y = \{y_i\}_{i=1}^N \subseteq \mathbb{R}^d \), where \( y_i \) is the new low dimensional representation of the original high dimensional data point \( x_i \in X \subseteq \mathbb{R}^D \).

4. ENDMEMBER SELECTION

The fourth step in our algorithm is to select endmembers for the low dimensional space \( Y \subseteq \mathbb{R}^d \). Traditional applications of endmember algorithms are run on the original high dimensional data set \( X \subseteq \mathbb{R}^D \), and if \( s \) denotes the number of endmembers, then \( s < D \). Since we are finding endmembers for the space \( Y \), we propose finding \( s > d \) endmembers, thus creating a frame \( \Phi = \{\varphi_i\}_{i=1}^s \) for \( Y \). Frames arise naturally in dimension reduction, and are in fact a generalization of orthonormal bases. There are many endmember selection algorithms available, e.g., N-FINDR [6], ORASIS [7], and Pixel Purity Index [8]; see also [9] and [10]. The results of this paper employ the Support Vector Data Description (SVDD), see, e.g., [11] algorithm for selecting endmembers. The core idea of SVDD is to obtain a minimal spherical shaped boundary around the data set, which in turn gives a description of the data in terms of a set of support vectors.
5. FRAME COEFFICIENTS

Given a frame $\Phi = \{\varphi_i\}_{i=1}^s$ for $Y$, we shall find a set of coefficients $C = \{c_{i,j}\}_{i,j=1}^{N,s}$ that represents $Y$ in terms of $\Phi$:

$$y_i = \sum_{j=1}^s c_{i,j} \varphi_j \quad \text{for all } i = 1, \ldots, N.$$ 

We propose two separate ways to find $C$. The first is based on the frame operator $S : \mathbb{R}^d \to \mathbb{R}^d$, which is:

$$Sy = \sum_{i=1}^s \langle y, \varphi_i \rangle \varphi_i \quad \text{for all } y \in \mathbb{R}^d.$$ 

For any frame $\Phi$, the frame operator $S$ is invertible, and in fact gives the following representation:

$$y = \sum_{i=1}^s \langle y, S^{-1} \varphi_i \rangle \varphi_i \quad \text{for all } y \in \mathbb{R}^d.$$ 

The coefficients $c_{i,j} = \langle y_i, S^{-1} \varphi_j \rangle, i = 1, \ldots, N, j = 1, \ldots, s$, are called the canonical coefficients and they minimize the $\ell^2$ energy of the coefficient set $C$. An alternative to the canonical coefficient set is to find sparse coefficient representations. Such coefficients are found by minimizing the $\ell^p$ energy of the coefficients, where $0 < p \leq 1$:

$$c_{i,.} = \arg \min_{\tilde{c}} \| \tilde{c} \|_{\ell^p} \quad \text{subject to } y_i = \sum_{j=1}^s \tilde{c}_j \varphi_j.$$ 

6. REFERENCES


