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 [1], Φ, 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 Φ. 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^{th} smallest non-zero eigenvector by v_j, and let the i^{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 in 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 Φ:

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

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

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

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

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

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 ℓ^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 ℓ^p energy of the coefficients, where $0 < p \leq 1$:

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

6. REFERENCES

[1] J. Kovacevic and A. Chebira, "Live beyond bases: The advent of frames," *IEEE Signal Processing Magazine*, 2007.

[2] S. Roweis and L. Saul, "Nonlinear dimension reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[3] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," *Advances in Neural Information Processing Systems*, vol. 14, pp. 585–591, 2002.

[4] J. Tenenbaum, V. de Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[5] Y. Bengio, J.-F. Paiement, P. Vincent, O. Delalleau, N. Le Roux, and M. Ouimet, "Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering," *Advances in Neural Information Processing Systems*, vol. 16, pp. 177–184, 2004.

[6] M. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proceedings of SPIE*, vol. 3753, pp. 266–275, 1999.

[7] J. Bowles, P. Palmadesso, J. Antoniades, M. Baumbeck, and L. Rickard, "Use of filter vectors in hyperspectral data analysis," *Proceedings of SPIE*, vol. 2553, pp. 148–157, 1995.

[8] J. Boardman, F. Kruse, and R. Green, "Mapping target signatures via partial unmixing of aviris data," *Fifth JPL Airborne Earth Science Workshop, volume 1 of JPL Publication 95-1*, pp. 23–26, 1995.

[9] C. Wu and C.-I Chang, "A new simplex growing algorithm for endmember extraction," *IEEE International Geoscience and Remote Sensing Symposium*, July 2005.

[10] D. Gillis, J. Bowles, and M. Winter, "Using endmembers as a coordinate system in hyperspectral imagery," *Proceedings of SPIE*, vol. 4816, pp. 346–354, 2002.

[11] A. Banerjee, P. Burlina, and J. Broadwater, "A machine learning approach for finding hyperspectral endmembers," *IEEE International Geoscience and Remote Sensing Symposium*, vol. IGARSS 2007, pp. 3817–3820, 2007.