Research Statement

Justin D. Marks

December 1, 2017

1 Research Area and Community

Broadly, my research area is in applied mathematics with emphases on matrix analysis, numerical linear algebra, and algebraic geometry. More narrowly, I have developed classification and clustering algorithms on matrix manifolds. Data sets for application of these algorithms are digital photography, hyperspectral imagery, and other large data sets. Relevant journals for submission of articles include SIAM Journal on Matrix Analysis and Applications, Linear Algebra and its Applications, Numerical Algorithms, and Transactions on Geoscience and Remote Sensing. Primarily, my research ability was cultivated within the Pattern Analysis Laboratory at Colorado State University, where we held seminar series on topics such as support vector machines, electroencephalogram data, and application of the l_1-norm for sparsity.

I have developed a professional network of scholars and collaborators within academia and the governmental sector. To disseminate my ideas and connect with other researchers, I have attended and spoken at conferences such as the Gene Golub SIAM Summer School in Italy during June 2010, the NSF/DTRA Algorithms Workshops in Chapel Hill, Boston, San Diego, Boulder, and Washington D.C. during June 2010, June 2011, November 2012, March 2014, and July 2015, respectively, and ICIAM in Vancouver, Canada during July 2011. Participants at these conferences represent Mathematics, Computer Science, Chemistry, Biology, National Laboratories, and the Department of Defense including the Air Force. These conferences led to a collaborative position at MIT Lincoln Laboratory and a postdoctoral position at Air Force Institute of Technology.

2 Published Work

Large data sets, such as collections of digital photographs, can be challenging to correctly cluster and classify. The development of effective algorithms to accomplish these tasks is the primary goal of the Pattern Analysis Lab at Colorado State University [1], of which I was a member. For instance, suppose that we have obtained five images of person A, five images of person B, and five images of person C. Now suppose the shapes in Figure 1 represent the arrangement of these images within some high-dimensional vector-space.

Next, suppose that we aim to classify unlabeled novel images according to the class label of the closest gallery image. The arrangement in Figure 1 suggests that an unlabeled novel image of one of these three people would be difficult to classify accurately if it was located close to the border between two of the clusters. We would benefit from an algorithm that could map the original images in Figure 1 to the arrangement in Figure 2, without human intervention.

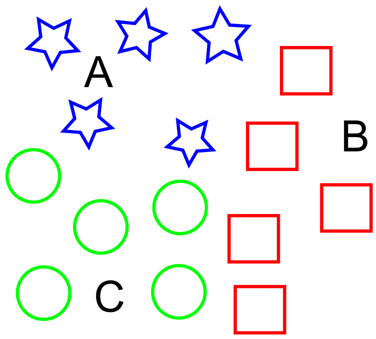

Figure 1: The classification task may have difficulty with this arrangement of class data, since the separation between classes and the density within individual classes is low.

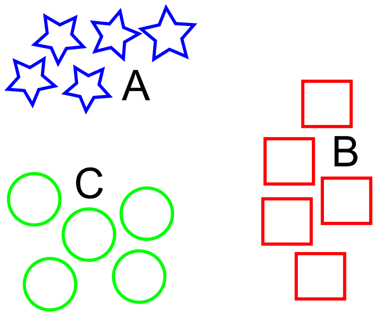

Figure 2: The classification task is facilitated by this alternate arrangement of class data, since the separation between classes and density within individual classes are amplified versus Figure 1.

The effect of this mapping is an amplification of between-class separation and within-class compression. It is likely that the ability to discriminate between classes is improved by the mapping. This is a rough sketch of the motivation and effect of the Discriminative Canonical Correlations algorithm [10], which was the first geometric data analysis algorithm I studied in detail. My master's thesis [12] contains a thorough analysis of the mathematical foundations of the algorithm and various implementations.

The setting for Discriminative Canonical Correlations is the Grassmann manifold. Advancing from my master's work, I pursued a deeper understanding of the geometric framework of the Grassmann manifold, as well as the Stiefel manifold [6]. The tangent space and normal space definitions for these manifolds led to new algorithms designed by my advisors and myself. To motivate these algorithms, let us refer again to our data arrangement example in Figure 1. For many applications, such as statistical analysis or data compression, it is advantageous to compute a mean representative for a cluster of within-class points. In Figure 3 we mark reasonable locations for the means of each of the three classes.

The computation of these means is simple in the case of vector spaces. Given a set of vectors, add the vectors together and then divide by the number of vectors in the set. However, the standard definition of mean used in vector spaces is likely to result in a point not on X if X is not "flat," as in the case of a Grassmann or Stiefel manifold. More sophisticated techniques are required to identify sensible mean representatives in this context, and three novel techniques constitute the core of my doctoral dissertation [13].

The first of our algorithms we name, the *normal mean*. We exploited both the tangent space and the normal space of the Grassmann and Stiefel manifolds in formulating a two-

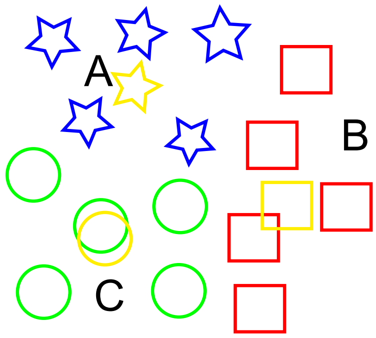

Figure 3: Reasonable locations for means for each class are marked.

phase iterative algorithm that converges to a mean. The normal mean shares algorithmic structure with the Karcher mean [3]. Given a cluster of points $\{X_i\}_{i=1}^N$ on either manifold, we first make a guess for the mean manifold point, μ, e.g., $\mu = X_1$. Then we intersect the normal space at each manifold point X_i with the tangent space to the manifold at μ. This generates a collection of tangent vectors which can be averaged in the naive way to obtain tangent vector δ. Let the thin SVD decomposition of δ be $\delta = U\Sigma V^T$. We map δ to the closest orthonormal matrix, UV^T, which is a point on, or a representative for a point on, the Stiefel or Grassmann manifold, respectively. This point becomes our updated μ iterate. We continue this procedure until the norm of μ is unchanged beyond a determined tolerance, e.g., 10^{-6}.

The *projection mean* is similar in concept to the normal mean. The major modifications are that the tangent projection map is used in place of intersecting the normal and tangent spaces, and the inverse of the tangent projection map is used in place of the closest orthonormal matrix. We note that the Stiefel projection mean algorithm involves an intriguing appeal to numerical algebraic geometry. On the other hand, our third mean is in a separate category from either of these means and deserves a more complete explanation.

Given a collection of subspaces of $\mathbb{R}^n$, $\{S_i\}_{i=1}^N$, perhaps of different dimension $\{d_i\}_{i=1}^N$, we develop a pair of algorithms to compute a *flag mean* of arbitrary dimension. The foundation of each algorithm is the formulation of an optimization problem based on, for each i, the cosine of the principal angle θ_i between subspace S_i and an optimizing line spanned by the unit vector u. The first optimization problem, (1), yields a polynomial system that may be solved using homotopy continuation methods from numerical algebraic geometry:

$$\begin{aligned} &\underset{u}{\text{argmax}} && \sum_{i=1}^{N} \cos\theta_i \\ &\text{subject to} && u^T u = 1, u \in \mathbb{R}^n. \end{aligned} \tag{1}$$

The second optimization problem, using squares of cosines of principal angles, yields an eigenvector problem, thus sidestepping the iterative methods of many manifold means. This second optimization problem is shown to be intimately related to Multi-set Canonical Correlation Analysis [9, 14]. These two types of flag means apply to collections of points on a Grassmann manifold, as an example. However, the domain of application is much more general, a distinct advantage of the flag mean over other manifold means. A flag is a nested sequence of subspaces. Since the result of repeated application of the algorithm with subspace deflation is in fact a flag, the name *flag mean* is appropriate.

With our arsenal of mean algorithms, we made application to a variety of data sets. The Pattern Analysis Lab database of face images, the KTH action database, and a hyperspectral bio-chem image database became the playgrounds for our algorithms. I enjoyed a fruitful collaboration with Dimitris Manolakis at MIT Lincoln Laboratory from November 2011 to June 2012. We combined my geometric data analysis training with nonlinear manifold methods in considering classification and feature extraction schemes for hyperspectral bio-chem data.

From September 2012 to July 2013, I was a postdoctoral researcher at Air Force Institute of Technology (AFIT) at Wright Patterson Air Force Base, during which I began working with Matt Fickus on frame theory. Frame theory involves the analysis of structured linear operators. It is an elegant area, pure in flavor but not without application to practical problems, especially signal processing. Complex analysis and Fourier analysis are essential for this research. We wrote a paper identifying and proving the optimality of an algorithm to complete a frame with given spectrum utilizing frame vectors of prescribed lengths [7]. In addition to frame theory, we continued to examine hyperspectral imagery, with an emphasis on compressed sensing cameras for the collection of data cubes.

3 Ongoing Research Program

3.1 Matrix Manifold Means

While at Bowdoin College from August 2013 to June 2015, at Wesleyan University from September 2015 to June 2016, and now at Gonzaga University, I have continued my collaboration with Michael Kirby and Chris Peterson at Colorado State University. We published a paper on the version of the flag mean which uses the squares of the cosines of the principal angles [5], a paper regarding the version of the flag mean which uses the cosines of the principal angles was [2], and we are preparing to submit a paper on the normal mean and projection mean [11].

A recent publication [8] has cited our unpublished normal mean, and there is now a hole in the literature that we need to fill with a paper. I have conducted the same experiments outlined in this recent publication, and discovered that the paper seems to misrepresent our algorithm. During a research trip to Colorado State University in 2017, my collaborators and I attempted to identify the inverse function of the tangent bundle projection operator for the Grassmann manifold, but it has been elusive. It is possible that this inverse operator does not exist (is not well-defined). This has been an obstacle to including a projection mean in our paper.

I am planning to spend a number of weeks in Colorado during Summer 2018 to develop further research momentum in related projects, such as convergence analysis and comparisons between the means. We hope to write application papers to build upon the first set of papers which developed the mathematical foundations for the algorithms. It is possible that I will eventually write a book about means on matrix manifolds.

3.2 Hyperspectral Imagery

I collaborated with Avishai Ben-David at Edgewood Chemical Biological Center. We published a paper in September 2014 which highlights the utility of geodesic path interpolation between two covariance matrices for background estimation in hyperspectral images [4].

This approach is motivated by the geometry of the Riemannian manifold of covariance matrices. Both Avishai and I are interested in building bridges between manifold geometry and signal processing applications.

A significant open problem arose during my time at MIT Lincoln Laboratory concerning the differences between a statistical approach and my nonlinear geometric approach for detection of a biochemical plume within a hyperspectral image. The results that we obtained were strangely counterintuitive, and Michael Kirby and I remain motivated to demonstrate that a geometric approach to detection is in fact preferred in many circumstances.

3.3 Frame Theory

In 2015, Matt Fickus and I published a paper on optimal frame completions [7], and I hope that we can collaborate further in the years ahead, with emphases on both frame theory and compressed sensing.

I will present one problem in the area of frame theory. This is a problem that has gathered much interest from the signal processing and frame theory community. We seek a lower bound for the number of frame vectors necessary for the recovery of a signal, up to global phase, from a collection of magnitudes of inner products of the frame vectors with the raw signal. We have shown that the lower bound problem is directly related to the following geometry problem:

Problem 1. *In $\mathbb{R}^M$, what is the maximum number N of 2-dimensional subspaces, $\{S_i\}_{i=1}^N$, that can be formed such that there exists another 2-dimensional subspace Q which, for each S_i, contains at least one nonzero vector that is orthogonal to S_i? Equivalently, what is the maximum number N of points $\{X_i\}_{i=1}^N$ on the Grassmann manifold $Gr(M,2)$ that can be selected such that there exists $Y \in Gr(M,2)$ which, for each X_i, forms at least one principal angle of $\frac{\pi}{2}$ with X_i?*

It appeals to me that frame theory problems can, at times, be translated into problems on Grassmannians, a robust mathematical setting which frequently provides illuminating insights.

3.4 Funding

Fortunately, research in the area of geometric data analysis has historically been generously funded. Moreover, geometric algorithm development is increasingly necessary for significant defense applications, and thus I am confident that there will be future support available for advancement of my work.

4 Research Mentoring

From Fall 2014 to Spring 2015, I advised a senior Bowdoin mathematics major in independent study in the area of geometric data analysis. My student read through a graduate-level textbook, *Geometric Data Analysis: An Empirical Approach to Dimensionality Reduction and the Study of Patterns* by Michael Kirby, and completed related problem sets. He simultaneously developed his conceptual understanding of mathematical algorithms and coding abilities in Matlab. We were motivated by electroencephalogram data generated by Bowdoin neuroscience professor Erika Nyhus. We considered geometric data processing algorithms that could contribute to answering a question about causality with regards to similarly structured waves emanating from distinct geographical regions of the brain. This year of independent study was a rich experience for both myself and my student, who remarked that it was the academic highlight of his time at Bowdoin.

Since Fall 2016 I have been a research advisor for Ethan Mahintorabi. We are working on computational mathematics research pertaining to matrix manifold means and geometric data analysis. Ethan is currently a senior Computer Science major with sophisticated coding skills and a moderate background in mathematics courses. I have helped Ethan develop a stronger foundation in linear algebra and subspace-based algorithms. One of Ethan's research projects was to build a classification tree algorithm based on subspaces and to compare to classifiers based on individual instances. He presented his work with a poster at the AMS meeting in Pullman, WA in April 2017, which was his first presentation at a conference. He found this to be a tremendously valuable experience.

I will continue to advise Ethan until the end of Fall 2017. At that point, I will leave for Gonzaga-In-Florence. This academic year we received a McDonald Work Award to support his research. We are working to enhance detection of video piracy, and we are studying how to use Hilbert space-filling curves to vectorize images, as opposed to a more routine raster-scanning (row-wise scanning) approach.

I look forward to further involving students in my research. I will advise students to target problems in geometry, pure linear algebra, statistics, and computational analysis. For example, students could aid me in performing comparative statistical analysis on Grassmann and Stiefel manifolds using each of the mean variants and the corresponding variances. Another project could involve classification algorithms, such as nearest neighbor, paired with data compression based on each of the mean representatives. In addition, students could optimize mean algorithms for computational speedup. As my research interests evolve over the coming years, I expect that corresponding student research will evolve with it. Furthermore, for the sake of networking and exposure to diverse areas of mathematics, I will encourage students to attend and present their work at regional, national, and international conferences.

References

[1] Pattern analysis lab. `http://www.math.colostate.edu/~kirby`. Department of Mathematics, Colorado State University.

[2] Daniel J. Bates, Brent R. Davis, Michael Kirby, Justin Marks, and Chris Peterson. The max-length-vector line of best fit to a set of vector subspaces and an optimization problem over a set of hyperellipsoids. *Numerical Linear Algebra with Applications*, 22(3):453–464, 2015. nla.1965.

[3] Evgeni Begelfor and Michael Werman. Affine invariance revisited. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2*, CVPR '06, pages 2087–2094, Washington, DC, USA, 2006. IEEE Computer Society.

[4] Avishai Ben-David and Justin Marks. Geodesic paths for time-dependent covariance matrices in a riemannian manifold. *Geoscience and Remote Sensing Letters, IEEE*, 11(9):1499–1503, Sept 2014.

[5] Bruce Draper, Michael Kirby, Justin Marks, Tim Marrinan, and Chris Peterson. A flag representation for finite collections of subspaces of mixed dimensions. *Linear Algebra and its Applications*, 451(0):15 – 32, 2014.

[6] Alan Edelman, Tomás Arias, and Steven T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl*, 20:303–353, 1998.

[7] Matthew Fickus, Justin D. Marks, and Miriam J. Poteet. A generalized schur-horn theorem and optimal frame completions. *Applied and Computational Harmonic Analysis*, 2015.

[8] S. Fiori, T. Kaneko, and T. Tanaka. Ieee transactions on signal processing. *Tangent-Bundle Maps on the Grassmann Manifold: Application to Empirical Arithmetic Averaging*, 63(1):155–168, Jan 2015.

[9] Jon R. Kettenring. Canonical analysis of several sets of variables. *Biometrika*, 58(3):433–451, 1971.

[10] Tae-Kyun Kim, Josef Kittler, and Roberto Cipolla. Discriminative learning and recognition of image set classes using canonical correlations. *IEEE*, 29(6):1005–1018, 2007.

[11] Michael Kirby, Justin Marks, and Chris Peterson. Two normal/tangent bundle algorithms for averaging point clouds on grassmann and stiefel manifolds. In preparation.

[12] Justin Marks. Discriminative canonical correlations: An offspring of linear discriminant analysis. Master's thesis, Colorado State University, 2009.

[13] Justin Marks. *Mean Variants on Matrix Manifolds*. PhD thesis, Colorado State University, 2012.

[14] Javier Via, Ignacio Santamaria, and Jesus Perez. A learning algorithm for adaptive canonical correlation analysis of several data sets. *Neural Networks*, 20(1):139 – 152, 2007.