USING RANDOM MATRIX THEORY TO DETERMINE THE NUMBER OF ENDMEMBERS IN A HYPERSPECTRAL IMAGE

K. Cawse[1,2], M. Sears[3], A. Robin[1], S.B. Damelin[1,4], K. Wessels[2], F. van den Bergh[2], R. Mathieu[5]

[1]School of Computational and Applied Maths
University of the Witwatersrand
South Africa

[2]Remote Sensing Research Unit
Meraka Institute, CSIR
South Africa

[3]School of Computer Science
University of the Witwatersrand
South Africa

[4]Department of Mathematical Science
Georgia Southern University
USA

[5]Ecosystems and Earth Observation
Natural Resources and the Environment, CSIR
South Africa

ABSTRACT

Determining the number of spectral endmembers in a hyperspectral image is an important step in the spectral unmixing process, and under- or overestimation of this number may lead to incorrect unmixing for unsupervised methods. In this paper we discuss a new method for determining the number of endmembers, using recent advances in Random Matrix Theory. This method is entirely unsupervised and is computationally cheaper than other existing methods. We apply our method to synthetic images, including a standard test image developed by Chein-I Chang, with good results for Gaussian independent noise.

Index Terms— Hyperspectral Unmixing, Random Matrix Theory, Linear Mixture Model, Virtual Dimension.

1. INTRODUCTION

Determining the number of endmembers in an image is important for the processing of many different types of data, including chemical unmixing [1], extracting speech signals in a noisy band [2], unmixing minerals [3] and unmixing environmental landscapes [4], among many others. This is an important first step in unmixing a hyperspectral image, and this step is often overlooked in image unmixing algorithms [5].

A common model used to unmix hyperspectral images is the linear mixing model, introduced by [6]. This model assumes that each pixel in the image is made up of a linear combination of "endmembers". Mathematically, measurement $\mathbf{x}_i \in \mathbb{R}^p$ for each pixel $i, 1 \leq i \leq p$, where $p \geq 1$ is the number of spectral bands, is represented as:

$$\mathbf{x}_i = \sum_{j=1}^{K} a_{ij}\mathbf{v}_j + \mathbf{n}_i \tag{1}$$

where a_{ij} represents the proportion of endmember $\mathbf{v}_j$ in the mixed pixel $\mathbf{x}_i$, $\mathbf{n}_i$ represents some noise function, and K is the number of endmembers. We assume Gaussian noise following the methods of [1] [5].

The first step in unmixing the image is to determine how many endmembers or constituents are contained in the scene. This is known as the *Virtual Dimension* of the image. An incorrect estimation of this number, K, can seriously detriment the accuracy of many unmixing methods [5], and in practice, algorithms do not perform well when endmembers are included that are not present in the image [7].

There are several existing methods for determining the Virtual Dimension of the image. Most of these methods use the eigenvalues of the observation covariance matrix S, where S is given by

$$S = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T, \tag{2}$$

where N is the number of pixels in the image and $\bar{\mathbf{x}} \in \mathbb{R}^p$ is the mean spectral vector over all the pixels in the image. This matrix is used to distinguish the eigenvalues due to signal and the eigenvalues due to noise. This procedure is not unique to

S.B. Damelin acknowledges support from NSF, University of the Witwatersrand, and Georgia Southern University.

hyperspectral unmixing — it is used to unmix, for instance, signals [2], chemical mixtures [1] or other mixtures [3][4].

In an ideal case, there should be a clear distinction between eigenvalues due to signal and eigenvalues due to noise. However, in a realistic situation, it is extremely difficult to distinguish between a small signal eigenvalue and a large noise eigenvalue. Several methods have been developed to address this problem.

Wu *et al* [5] have provided a summary of methods to determine the number of endmembers, among which are *An Information Criterion* (AIC) [8] and *Minimum Description Length* (MDL) [9] which both determine K by minimising some function depending on the eigenvalues of S, the centered observation covariance matrix. Both of these methods assume that the noise is independent identically distributed (i.i.d.) and Gaussian. These noise assumptions are not appropriate for hyperspectral imagery and are seen as a disadvantage of these methods. Wu *et al.* [5] also discuss the Gerschgorin Radius-Based method [10] (assuming only i.i.d. noise), and Signal Subspace Estimation (SSE) [11] (assuming only Gaussian noise). Next, the paper discusses the Neyman-Pearson Detection Method [12] which assumes white noise with zero mean. All the methods discussed only use the eigenvalues of S or the eigenvalues for the non-centered observation covariance matrix. In all cases, nothing needs to be known about the endmember vectors, which is an advantage over supervised methods.

From real and synthetic experiments described in [5], the authors determined that the best methods were Neyman-Pearson, SSE and Gerschgorin Radius-Based method, and the methods with the strictest noise assumptions (i.e. that the noise is Gaussian and i.i.d.) performed the worst. Several of these methods were very sensitive to user-defined values. The difference between noise and signal eigenvalues can be very subtle.

In the area of chemical unmixing, Kritchman and Nadler [1] have worked with new results in Random Matrix Theory to determine which eigenvalues are due to noise and which are due to signal. The advantage to this method is that there are no parameters that need to be set by the user. Also, it appears to perform better than other methods in finding the threshold between noise and signal eigenvalues. Random Matrix Theory principles that are used are quite recent [13] and are proving to be extremely useful in chemical unmixing [1].

The chemical unmixing method is not directly applicable to hyperspectral imagery, but this method shows promise and we will attempt to adapt it for the hyperspectral case. In this paper, we propose a new approach to determine the threshold between noise and signal eigenvalues, using Random Matrix Theory. In Section 2 we will introduce our model, before showing some experiments in Section 3 and we will then draw some conclusions in Section 4.

2. MODEL FORMULATION

In Random Matrix Theory, research has been done into the first (largest) eigenvalue of a Random Matrix. Since we are assuming Gaussian noise, the first observed eigenvalue due to noise can be thought of as the first eigenvalue in a Random Matrix. According to Johnstone [13], the first eigenvalue of a Random Matrix fulfills the condition

$$\lambda \leq \sigma^2(\mu_{N,p} + s(\alpha)\sigma_{N,p}) \tag{3}$$

where σ is the variance of the Gaussian noise, α is a significance level, $s(\alpha)$ may be found by inverting the Tracy-Widom distribution (in [1] $\alpha = 0.5\%$) and for real valued data,

$$\mu_{N,p} = \frac{1}{N}\left(\sqrt{N-\frac{1}{2}}+\sqrt{p-\frac{1}{2}}\right)^2 \tag{4}$$

$$\sigma_{N,p} = \frac{1}{N}\left(\sqrt{N-\frac{1}{2}}+\sqrt{p-\frac{1}{2}}\right)\times \left(\frac{1}{\sqrt{N-\frac{1}{2}}}+\frac{1}{\sqrt{p-\frac{1}{2}}}\right)^{1/3} \tag{5}$$

Note that these functions do not depend on the number of endmembers, K.

Then, if the eigenvalues, $\{\lambda_j\}_{j=1}^p$, of the non-centered observation covariance matrix are sorted in descending order, so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$, K is defined so that $\lambda_j > \sigma^2(\mu_{N,p} + s(\alpha)\sigma_{N,p})$ for all $j \in \mathbb{Z}, 1 \leq j \leq K$.

Mathematically, the most reasonable approximation to the noise is the average of all the noise eigenvalues (see [1]), given by

$$\sigma^2_{REF} = \frac{1}{p-K}\sum_{j=K+1}^{p}\lambda_j. \tag{6}$$

Kritchman and Nadler [1] found this approximation to be too low for their chemical unmixing application, and so they considered a new noise estimate, given by

$$\sigma_0^2 = \frac{\sigma^2_{REF}}{1-\frac{K}{N}}. \tag{7}$$

However, in hyperspectral cases, N is large in proportion to K and so the extra denominator in σ_0^2 tends to one. Kritchman and Nadler [1] also derived other noise approximations, but their experiments are based on the assumption that N is small. Since for hyperspectral imagery the number of pixels, N, is large, we will consider σ^2_{REF} in our experiments.

Our approach has many advantages for hyperspectral imagery. Firstly, in this case, N is very large in proportion to p, so $\sigma_{N,p}$ in equation (5) becomes very small, and therefore the right hand side of equation (3) is not sensitive to the choice of the confidence interval α. This is in contrast to several methods examined by Wu *et al.* [5], where the user-defined threshold had a large impact on the results. Secondly, this method is computationally efficient when compared to other methods that performed well in [5]. Some of those require the solution to an optimization problem, the computation of eigenvectors, or the computation of eigenvalues for more than one large matrix. The Random Matrix Theory approach only requires the eigenvalues for the observation covariance matrix.

3. EXPERIMENTS

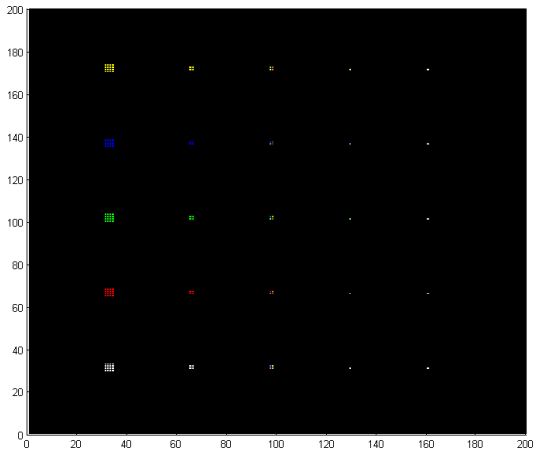

Fig. 1. Synthetic image created according to the instructions in [14].

Chang *et al.* have created a synthetic test dataset to standardise algorithm testing [14]. This dataset is an image of size 200×200 pixels, with pure pixels, mixed pixels and sub-pixel elements inserted into a background (see Figure 1). The image contains five minerals and a background (derived from the five minerals). Wu *et al.* [5] used the dataset to test several reviewed methods for finding the number of endmembers in an image. In this image, there are 5 rows of panels (where each row is associated with a mineral) and 5 columns of panels (where each column represents different size or mixing scenarios). The first column contains 4×4 pure pixels, the second column contains 2×2 pure pixels, the third column contains 2×2 mixed pixels, and the fourth and fifth columns contain sub-pixel elements. For the mixed pixels, every pixel in the panel has a value comprised of half the mineral of that row, and for each of the four pixels, the remaining half comes from one of the four other minerals. For the sub-pixel elements, in the fourth column, the pixel value is half the mineral of that row, and the other half is made up of the background value; for the fifth column, the pixel is 25% mineral and 75% background. Zero-centered, i.i.d, additive Gaussian noise is then added to the image.

We created a first dataset using the following minerals: Alunite; Buddingtonite; Calcite; Kaolinite; and Muscovite (taken from the JPL spectra library [15]). The background is the mean of all five minerals (as in [5]). We also compiled a second synthetic dataset, made up of 5-20 minerals, again chosen from the JPL spectral library [15]. The number of endmembers, K, is randomly selected, and the proportions of each endmember in each pixel is also random, with the only restrictions being the positive and sum-to-one conditions that are enforced on the proportions. Then, by iterating this method, testing may be done on many different "images", since K and the proportions differ for each iteration, and both are randomly selected from a uniform distribution for each iteration.

The random matrix evaluation in (3), with the noise given as in (6), may be rewritten as

$$\begin{aligned} f(p,K) &< 1 \text{ for signal eigenvalues, where} \\ f(p,K) &= \frac{\sum_{j=K+1}^{p} \lambda_j}{\lambda_K(p-K)}(\mu_{N,p} + s(\alpha)\sigma_{N,p}). \end{aligned} \quad (8)$$

In order to test this formula, we compiled a test image as suggested by [14], and a set of synthetic images for various number of endmembers, K, where $5 \leq K \leq 20$. For each K, a hundred different images were generated, with uniformly distributed abundance values. For these test images, we used a noise standard deviation of 0.001, which translates to SNR of 500:1, using the definition in [5]. This SNR is typical of modern airborne sensors.

We also tested our formula on an AVIRIS flight scene collected over Cuprite, Nevada in 1997. This dataset is available online[1]. The image contains 350×350 spatial pixels, with 189 spectral bands. Wu *et al.* [5] determined the number of endmembers in this scene to be between 22 and 28.

The validity of this formula for synthetic data is displayed in Figure 2. Only three synthetic images were displayed for clarity, but the same behavior was observed for all the synthetic images generated. For the synthetic images, the line $f(p,K) = 1$ is a good threshold between signal and noise eigenvalues, as expected from the Random Matrix behavior. But the Cuprite scene behaved differently.

For the synthetic data, it is assumed that the noise is zero-centered Gaussian, i.i.d. and uncorrelated with the signal. In a simple experiment where the variance of the noise in one of the synthetic images is made to depend in some way on the

[1] aviris.jpl.nasa.gov/html/aviris.freedata.html

band number, the synthetic graph approximates the Cuprite graph (as in Figure 2) much more closely, especially in terms of amplitude. By this experiment it appears that our noise assumptions are unrealistic for the Cuprite image. However, Random Matrix Theory can be adjusted for different noise conditions, and further work will focus on how to make these adjustments.

This paper presents a successful application of Random Matrix Theory to hyperspectral imagery in synthetic images.

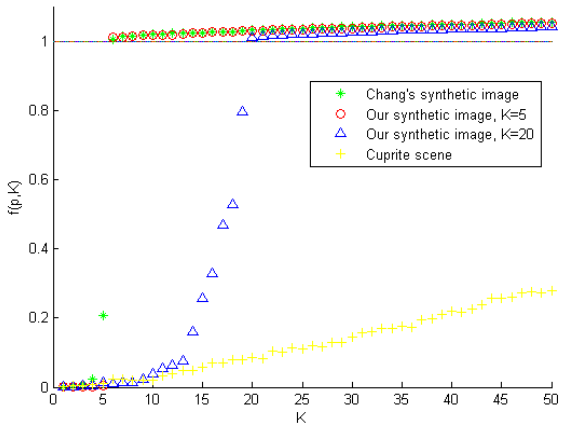

Fig. 2. Using Random Matrix Theory to threshold hyperspectral data, where equation (8) is displayed for synthetic images and a Cuprite scene. For clarity, the synthetic images are shown with low noise (standard deviation = 0.001).

4. CONCLUSION

We have presented a new and innovative method for determining the number of endmembers in a hyperspectral image. Our method uses Random Matrix Theory in order to separate signal and noise eigenvalues from the observation covariance matrix, and it is computationally efficient. This method is not sensitive to user-defined thresholds and performs well on simulated images. The preliminary experiments show that the noise assumptions need to be adjusted for real images, but the Random Matrix Theory does allow for different noise assumptions.

5. REFERENCES

[1] S. Kritchman, B. Nadler, "Determining the number of components in a factor model from limited noisy data," *Chemometrics and Intelligent Laboratory Systems*, vol. 94, pp. 19–32, 2008.

[2] H-M. Park, H-Y. Jung, T-W. Lee, S-Y. Lee, "Subband-based blind signal separation for noisy speech recognition," *Electronic letters*, vol. 35, pp. 2011–2012, 1999.

[3] F.A. Kruse, "Identification and mapping of minerals in drill core using hyperspectral image analysis of infrared reflectance spectra," *International journal of remote sensing*, vol. 17(9), pp. 1623–1632, 1996.

[4] D. Gillis, J. Bowles, E.J. Ientilucci, D.W. Messinger, "A generalised linear mixing model for hyperspectral imagery," *Proc. of SPIE*, vol. 6966, pp. 1–11, 2008.

[5] C-C. Wu, W. Liu, C-I. Chang, "Exploration of methods for estimation of number of endmembers in hyperspectral imagery," *Proc. of SPIE*, vol. 7(43), pp. 1–11, 2006.

[6] H. Horwitz, R. Nalepka, P. Hyde, J. Morgenstern, "Estimating the proportions of objects within a single resolution element of a multispectral scanner," in *Proceedings of the 7th International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan*, 1971, pp. 1307–1320.

[7] D.C. Heinz, C-I. Chang, "Fully Constrained Least Squares Linear Spectral Mixture Analysis Method for Material Quantification in Hyperspectral Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39(3), pp. 529–545, 2001.

[8] H. Akaike, "A new look at the statistical model identification," *IEEE Tran. Automatica Control*, vol. AC-19, pp. 716–723, 1974.

[9] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.

[10] H.T. Wu, J.F. Yang, F.K. Chen, "Source number estimators using transformed Gerschgorin radii," *IEEE Tras. on Signal Process.*, vol. 43(6), pp. 1325–1333, 1995.

[11] J. Dias, J. Nascimento, "Estimation of signal subspace on hyperspectral data," *Proc. of SPIE*, vol. 5982, pp. 191–198, 2005.

[12] J.C. Harsanyi, W. Farrand, C-I. Chang, "Detection of subpixel spectral signatures in hyperspectral image sequences," *ASPRS*, pp. 236–247, 1994.

[13] I.M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *The Annals of Statistics*, vol. 29(2), pp. 295–327, 2001.

[14] Y.C.C. Chang, H. Ren, C-I. Chang, R.S. Rand, "How to design synthetic images to validate and evaluate hyperspectral imaging algorithms," *Proc. of SPIE*, vol. 6966, pp. 69661P–1, 2008.

[15] C.I. Grove, S. Hook, E. Paylor, "Laboratory reflectance spectra of 160 minerals, 0.4 to 2.5 micrometers," *JPL Publications*, vol. 92–2, 1992.