NONLINEAR UNMIXING BY USING NON-EUCLIDEAN METRICS IN A LINEAR UNMIXING CHAIN

Rob Heylen, Paul Scheunders

IMinds-Visionlab, University of Antwerp
Universiteitsplein 1
2610 Wilrijk, Belgium

Anand Rangarajan, Paul Gader

CISE, University of Florida
E301 CSE Building, P.O. Box 116120
Gainesville, Fl 32611, USA

ABSTRACT

In the linear mixing model, many techniques for endmember extraction are based on the assumption that pure pixels exist in the data, and form the extremes of a simplex embedded in the data cloud. These endmembers can then be obtained by geometrical approaches, such as looking for the largest simplex, or by maximal orthogonal subspace projections. Also obtaining the abundances of each pixel with respect to these endmembers can be completely written in geometrical terms. While these geometrical algorithms assume Euclidean geometry, it has been shown that using different metrics can offer certain benefits, such as dealing with nonlinear mixing effects by using geodesic or kernel distances, or dealing with correlations and colored noise by using Mahalanobis metrics. In this paper, we demonstrate how a linear unmixing chain based on maximal orthogonal subspace projections and simplex projection can be written in terms of distance geometry, so that other metrics can be easily employed. This yields a very flexible processing chain: by using other metrics, the same unmixing methodology can be used to deal with a wide range of unmixing models and scenarios. As an example, metrics are provided for dealing with intimate mixtures, nonlinear dimensionality reduction, and colored noise.

1. INTRODUCTION

Most processing chains employed in hyperspectral unmixing employ a sequence of three algorithms, i.e., for estimating the number of endmember, extraction of the endmembers from the data, and inversion of the mixing equation. One often assumes the linear mixing model (LMM), where the observed spectrum is a convex linear combination of endmember spectra. In the LMM, estimation of the number of endmembers often comes down to estimating the dimensionality of the data subspace. Endmember extraction algorithm (EEAs) often assume the presence of pure pixels, and try to identify these as extreme points or as the vertices of a simplex. Inversion algorithms then need to solve a convex optimization problem, where the set of abundances has to be found that minimizes the reconstruction error, subject to convexity constraints.

The LMM assumes that a spectrum $\boldsymbol{x}_n \in \mathbb{R}^D$ is a linear combination of M endmembers $\{\boldsymbol{e}_m\}_{m=1}^M$ and additional noise $\boldsymbol{\eta}_n$, with abundances that are positive and sum to one:

$$\boldsymbol{x}_n = \sum_{m=1}^{M} a_{nm}\boldsymbol{e}_m + \boldsymbol{\eta}_n \qquad (1)$$
$$\forall n, m : a_{nm} \geq 0, \ \sum_m a_{nm} = 1$$

If a data set $\{\boldsymbol{x}_n\}_{n=1}^N$ is given, many EEAs employ geometrical techniques to identify the endmembers, under the assumption that they are present in the data as pure pixels. One approach is to identify maximal projections or distances in subsequent orthogonal projections (orthogonal subspace projection [1], automated target generation process [2], maximum distance algorithm [3], vertex component analysis [4]), or to iteratively identify vertices that maximize the simplex volume (simplex growing algorithm [5]). Since maximizing an orthogonal distance is equivalent to maximizing a simplex volume from a geometrical point of view, all these algorithms perform the same operation at their core. They only differ in terms of their initialization, implementation details, and possible requirement of linear dimensionality reduction.

In this paper, we present an EEA that also belongs to this family, but works in terms of distances between points instead of coordinates (i.e., distance geometry). While this removes the requirement for dimensionality reduction, and hence solves a problem that is often present in alternative EEAs, this is not the main novelty of this paper. The use of distance geometry allows us to explore the effects of different metrics on the unmixing results.

Using different metrics can be beneficial for several reasons. First of all, several nonlinearities can be introduced through the use of a proper metric: Certain nonlinear models can be used to derive a metric (e.g. the Hapke model for intimate mixing [6]), any kernel function naturally leads to a metric, and a geodesic distance matrix will simulate the use of nonlinear dimensionality reduction, but without the large computational overhead created by the linear embedding step. Furthermore, metrics used to deal with correlations or colored

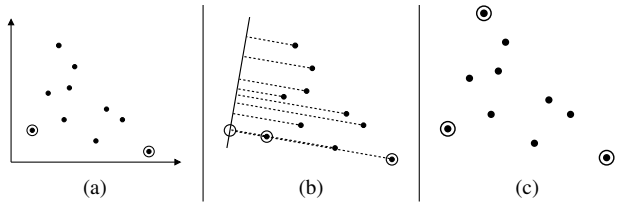


Fig. 1: Illustration of the addition of a third endmember in the MaxD algorithm. (a): The data set, where two endmembers have been found (circles). (b): All points are projected onto the hyperplane orthogonal to the already selected endmembers (indicated with a line). We look for the point with the largest distance from the endmember projection (the circle on the line) (c): The pixel with the largest distance from the endmembers is added as a new endmember. This pixel will also maximize the simplex volume.

noise can be employed as well (e.g., Mahalanobis distance). A similar reasoning was employed in [7], but for the much slower NfindR algorithm, and only for graph-geodesic distances.

2. THE DMAXD ALGORITHM

The algorithm is based on a distance-geometry based version of the MaxD endmember extraction algorithm [3], and is hence called the distance-MaxD or DMaxD algorithm. Consider a spectral space $\mathbb{R}^D$ of D dimensions, and a data set $\{\boldsymbol{x}_n\}_{n=1}^N$ consisting of N points, of which M points are endmembers. The pixel of largest norm is selected as the first endmember:

$$\boldsymbol{e}_1 = \boldsymbol{x}_I, \quad I = \underset{n}{\operatorname{argmax}} \|\boldsymbol{x}_n\|_2 \tag{2}$$

Subsequent endmembers are selected as those with the largest distance from the already obtained endmembers, in the space orthogonal to the endmember plane. See Fig. 1 for an illustration in two dimensions. Note that the space orthogonal to a single point is simply the entire data space.

The orthogonal distance to a hyperplane spanned by several points can be identified in terms of mutual distances with the Cayley-Menger (CM) determinant. Consider the squared distance matrix $\boldsymbol{D}_q = \{d_{ij}\}_{i,j=1}^q$, with $d_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2$ the squared distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. The volume $V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q)$ of the $(q-1)$-simplex spanned by vertices $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_q\}$ can be written as

$$(-1)^{q+1} 2^q (q!)^2 V^2 = \det(\boldsymbol{C}_q) \tag{3}$$

where $\boldsymbol{C}_q$ is the CM matrix of the q points:

$$\boldsymbol{C}_q = \begin{bmatrix} \boldsymbol{D}_q & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \tag{4}$$

$$= \begin{bmatrix} 0 & d_{1,2} & \ldots & d_{1,q-1} & d_{1,q} & 1 \\ \vdots & & & \ddots & & \vdots \\ d_{q,1} & d_{q,2} & \ldots & d_{q,q-1} & 0 & 1 \\ 1 & 1 & \ldots & 1 & 1 & 0 \end{bmatrix} \tag{5}$$

With some algebra (see e.g. [7]), we can use this equation to derive an orthogonal projection operator in terms of mutual distances: The orthogonal distance from $\boldsymbol{x}_q$ to the hyperplane through the points $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{q-1})$ is given by

$$d_\perp(\boldsymbol{x}_q; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{q-1}) = \frac{\boldsymbol{d}_q^{\mathrm{T}} \boldsymbol{C}_{q-1}^{-1} \boldsymbol{d}_q}{2} \tag{6}$$

with $\boldsymbol{C}_{q-1}$ the CM matrix of the set $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{q-1}\}$, and $\boldsymbol{d}_q = (d_{1,q}, \ldots, d_{q-1,q}, 1)$. The DMaxD algorithm then functions as follows: Let I be an index set $I = (i_1, \ldots, i_q)$, $i_j \in [1, \ldots, N]$. Let the notation $\boldsymbol{C}(I)$ be shorthand for $\boldsymbol{C}_q(\boldsymbol{x}_{I_1}, \ldots, \boldsymbol{x}_{I_q})$, which is the Cayley-Menger matrix of the points indexed by I, and let the distance function $D(\boldsymbol{x}, \boldsymbol{y})$ return the quadratic distance between $\boldsymbol{x}$ and $\boldsymbol{y}$.

Algorithm 1: DMaxD

input : Data set $\{\boldsymbol{x}_n\}_{n=1}^N$, number of endmembers M, distance function $D(.,.)$
output: Index set of the endmembers I

```
1  begin
2      I = {argmax_n D(0, x_n)} ;
3      for n = 1, ..., N do
4          d(1, n) = D(x_{I(1)}, x_n)
5      for q = 2, ..., M do
6          P = C(I)^{-1} ;
7          for n = 1, ..., N do
8              v = [d(:, n); 1] ;
9              o(n) = v^T P v ;
10         I = I ∪ {argmax_n o(n)} ;
11         for n = 1, ..., N do
12             d(q, n) = D(x_{I(q)}, x_n);
13     return I ;
```

Line 2 determines the first endmember as the endmember farthest from the origin. Lines 3-4 calculate the distances from this endmember to all other pixels, and store these as a single row in the distance matrix d. Line 5 starts the main loop that will add $M-1$ additional endmembers iteratively. Lines 6-9 calculate the orthogonal distances from each point $\boldsymbol{x}_n$ to the hyperplane through the endmember set. The index corresponding to the highest orthogonal distance is added to the endmember index set I in line 10, and the distance matrix

d, containing the distances from the current endmember set to all other data points, is updated in lines 11-12. Finally, the resulting endmember index set I is returned in line 13.

3. USEFUL METRICS FOR UNMIXING

A metric in spectral space is any function that returns the distance between two points, and obeys certain conditions, such as nonnegativity, symmetry and the triangle inequality. While there is an infinite number of metrics that can be defined, some of them have a certain appeal when considering the unmixing problem. We will consider following metrics:

- The Euclidean distance, yielding the LMM.
- The Mahalanobis distance, which takes the data covariance matrix $\boldsymbol{Z}$ into account, and corresponds to whitening the data:

$$d(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x} - \boldsymbol{y})^T \boldsymbol{Z}^{-1} (\boldsymbol{x} - \boldsymbol{y}) \tag{7}$$

- The single scattering albedo (SSA) distance. Here we use the Hapke model [6] for intimate mineral mixing to convert each reflectance x_d in the vector $\boldsymbol{x}$ to a SSA w_d using the relation

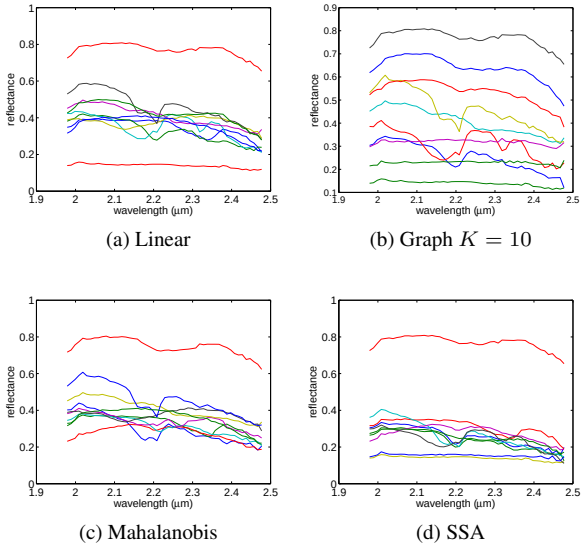
$$x_d = \frac{w_d}{(1 + 2\mu\sqrt{1 - w_d})(1 + 2\mu_0\sqrt{1 - w_d})} \tag{8}$$

  Variables μ and μ_0 are the cosines of the angles with the normal of the incoming and outgoing radiation respectively. The Hapke model assumes that linear mixing happens in the SSA space instead of reflectance space, hence the distances are Euclidean distances between the vectors $\boldsymbol{w}$ instead of $\boldsymbol{x}$:

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{w}_x - \boldsymbol{w}_y\|_2^2 \tag{9}$$

  with $\boldsymbol{w}_x$ and $\boldsymbol{w}_y$ the SSA vectors associated with $\boldsymbol{x}$ and $\boldsymbol{y}$ respectively, given by (8). This approach allows us to incorporate the entire intimate mixing machinery into the proposed metric.

- Graph-geodesic distances over a K-nearest neighbor graph, as in [7]. This distance matrix will approximate the true geodesic distances over the nonlinear data manifold, and allows one to execute geometrical algorithms respecting the manifold geometry. A geodesic distance matrix is generally only a true metric under certain conditions (i.e., zero curvature everywhere), but will yield decent results in practice.



(a) Linear (b) Graph $K = 10$

(c) Mahalanobis (d) SSA

Fig. 2: The endmember spectra retrieved with the DMaxD, for several distance functions. Possible parameters are listed in the subtitles.

type	min	median	mean	max
linear	0.0193	0.0379	0.0368	0.0517
Graph	**0.0177**	**0.0279**	**0.0310**	**0.0506**
Mahalanobis	0.0193	0.0333	0.0349	0.0517
SSA	0.0237	0.0357	0.0389	0.0577

Table 1: The minimum, median, mean and maximum of the spectral angles with the closest library spectra, over the 10 extracted endmembers.

4. EXPERIMENTS

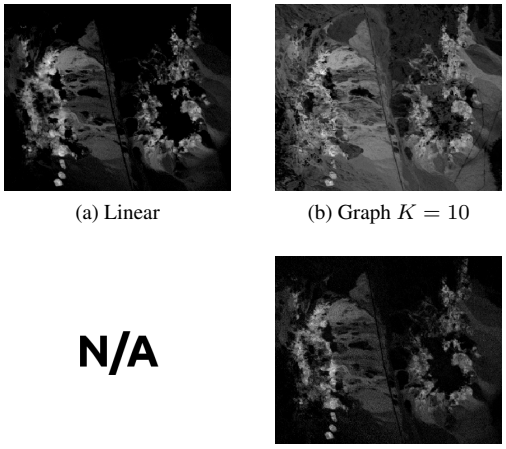
To asses the DMaxD algorithm, we have extracted 10 endmembers from the well-known Cuprite data set using the proposed metrics. The obtained endmember spectra are displayed in Fig. 2, and show that using different metrics will lead to endmember sets that can show substantial differences compared to a linear endmember extraction algorithm.

In order to identify the obtained endmember spectra, each one is compared with the USGS spectral database, and identified with the spectrum of lowest spectral angle. The smaller these angles are, the better the spectral match is, and the more certain that the obtained spectrum can indeed be considered to be the matched database spectrum. These angles, and their statistics, are given in table 1. From this table, it is clear that using other metrics can improve the endmember extraction process. The graph-geodesic metric generally obtains the smallest angles of the metrics considered.

In table 2, we have listed the spectra that were identified by each metric. It is clear that the obtained sets of minerals

Mineral	linear	Graph	Mahal.	SSA
alunite	1	1		1
beryl	1	1		2
chert	2	1	1	1
hydrogros.	1			1
kaolinite	1	1	2	1
mizzonite	1	2		1
mordenite	1		2	1
orthoclase	1		1	
zoisite	1		1	1
other		4	3	1

Table 2: The minerals identified for each metric. For each metric and mineral, we list how many times that mineral was the best fit with a retrieved endmember. The number 0 is not shown for clarity. Minerals that appear only once in a row are grouped together in the "other" category.



(a) Linear (b) Graph $K = 10$

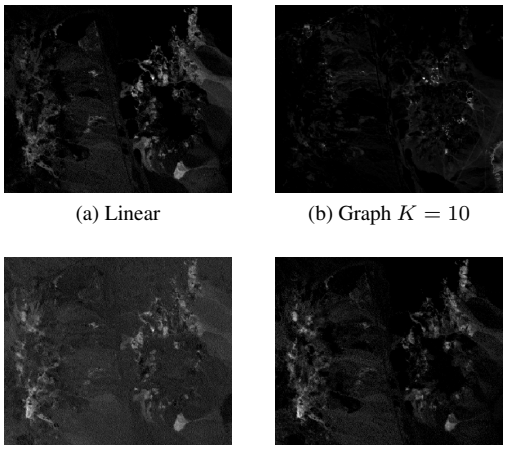
(c) Mahalanobis (d) SSA

Fig. 3: The abundance maps of the alunite endmember.

show a large overlap for most metrics, although some peculiar behavior can be observed as well. For instance, the Mahalanobis distance did not identify an alunite endmember, which is known to be present in large abundances in this data set.

After obtaining the endmembers, the data sets can be unmixed using the distance simplex projection unmixing (DSPU) algorithm. This algorithm is a distance-geometry based fully-constrained unmixing algorithm, and can hence function with different metrics as well [8]. These maps are shown in Fig. 3 and 4.

These preliminary results show that nonlinearity can be introduced into distance-geometry based linear processing chains through the introduction of metrics, and can yield very flexible algorithms, able to incorporate a variety of nonlinear effects and mixing equations.



(a) Linear (b) Graph $K = 10$

(c) Mahalanobis (d) SSA

Fig. 4: The abundance maps of the kaolinite endmember.

5. REFERENCES

[1] J.C. Harsanyi and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, pp. 779–785, 1994.

[2] H. Ren and C.-I Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerospace Elec. Sys.*, vol. 39, no. 4, pp. 1232–1249, October 2003.

[3] J.R. Schott, K. Lee, R. Raqueno, G. Hoffmann, and G. Healey, "A subpixel target detection technique based on the invariance approach," *Proc. AVIRIS Workshop*, 2003.

[4] J.M.P. Nascimento and J.M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, 2005.

[5] C.-I Chang, C.-C. Wu, W.-M. Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2804–2819, 2006.

[6] B. Hapke, "Bidirectional reflectance spectroscopy. 1. Theory," *J. Geophys. Res.*, vol. 86, pp. 3039–3054, 1981.

[7] R. Heylen, D. Burazerović, and P. Scheunders, "Nonlinear spectral unmixing by geodesic simplex volume maximization," *IEEE J. Sel. Top. Sig. Proc.*, vol. 5, no. 3, pp. 534–542, 2011.

[8] R. Heylen and P. Scheunders, "A distance geometric framework for nonlinear hyperspectral unmixing," *IEEE J. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. (under review)*, 2014.