

MINIMUM VOLUME SIMPLEX ANALYSIS: A FAST ALGORITHM TO UNMIX HYPERSPECTRAL DATA

Jun Li and José M. Bioucas-Dias

Instituto de Telecomunicações,
Instituto Superior Técnico, Technical University of Lisbon,
Lisboa, Portugal

ABSTRACT

This paper presents a new method of minimum volume class for hyperspectral unmixing, termed *minimum volume simplex analysis* (MVSA). The underlying mixing model is linear; *i.e.*, the mixed hyperspectral vectors are modeled by a linear mixture of the endmember signatures weighted by the correspondent abundance fractions. MVSA approaches hyperspectral unmixing by fitting a minimum volume simplex to the hyperspectral data, constraining the abundance fractions to belong to the probability simplex. The resulting optimization problem is solved by implementing a sequence of quadratically constrained subproblems. In a final step, the hard constraint on the abundance fractions is replaced with a hinge type loss function to account for outliers and noise.

We illustrate the state-of-the-art performance of the MVSA algorithm in unmixing simulated data sets. We are mainly concerning with the realistic scenario in which the *pure pixel* assumption (*i.e.*, there exists at least one pure pixel per endmember) is not fulfilled. In these conditions, the MVSA yields much better performance than the pure pixel based algorithms.

Index Terms— Hyperspectral unmixing, Minimum volume simplex, Source separation.

1. INTRODUCTION

Hyperspectral unmixing is a source separation problem [1]. Compared with the canonical source separation scenario, the sources in hyperspectral unmixing (*i.e.*, the materials present in the scene) are statistically dependent and combine in a linear or nonlinear fashion. These characteristics, together with the high dimensionality of hyperspectral vectors, place the unmixing of hyperspectral mixtures beyond the reach of most source separation algorithms, thus fostering active research in the field [2].

Given a set of mixed hyperspectral vectors, linear mixture analysis, or linear unmixing, aims at estimating the number of reference materials, also called endmembers, their spectral signatures, and their abundance fractions [1, 2, 3, 4, 5, 6]. The approaches to hyperspectral linear unmixing can be classified as statistical and geometrical based. The former addresses spectral unmixing as an inference problem, often formulated under the Bayesian framework, whereas the latter exploits the fact that the spectral vectors, under the linear mixing model, are in a simplex set whose vertices represent the sought endmembers.

This work was supported by the European Commission Marie Curie training grant MEST-CT-2005-021175. Email: {jun, bioucas}@lx.it.pt

1.1. Statistical approach to spectral unmixing

Modeling the abundance fractions (sources) statistical dependence in hyperspectral unmixing is a central issue in the statistical framework. In [7], the abundance fractions are modeled as mixtures of Dirichlet densities. The resulting algorithm, termed DECA, for dependent component analysis, implements an expectation maximization iterative scheme for the inference of the endmember signatures (mixing matrix) and the density parameters of the abundance fractions.

The inference engine in the Bayesian framework is the posterior density of the entities to be estimated, given the observations. According to the Bayes law, the posterior includes two factors: the observation density, which may account for additive noise, and a prior, which may impose constraints on the endmember matrix (*e.g.*, non-negativity of its elements) and on the abundance fractions (*e.g.*, to be in the probability simplex) and model spectral variability. Works [8, 9] are representative of this line of attack.

1.2. Geometrical approach to spectral unmixing

The geometrical approach exploits the fact that, under the linear mixing model, hyperspectral vectors belong to a simplex set whose vertices correspond to the endmembers. Therefore, finding the endmembers is equivalent to identify the vertices of the referred to simplex.

If there exists at least one pure (*i.e.*, containing just one material) pixel per endmember, then unmixing amounts to find the spectral vectors in the data set corresponding to the vertices of the data simplex. Some popular algorithms taking this assumption are the the N-FINDR [10], the *pixel purity index* (PPI) [11], the *Automated Morphological Endmember Extraction* (AMEE) [12], the *vertex component analysis* (VCA) [4], and the *simplex growing algorithm* (SGA) [13].

If the pure pixel assumption is not fulfilled, what is a more realistic scenario, the unmixing process is a rather challenging task, since the endmembers, or at least some of them, are not in the data set. A possible line of attack, in the vein of the seminal ideas introduced in [6], is to fit a simplex of minimum volume to the data set. Relevant works exploiting this direction are the *non-negative least-correlated component analysis* (nLCA) [14], the *alternating projected subgradients* [15], and the *nonnegative matrix factorization minimum volume transform* (NMF-MVT) [16]. We consider that the NMF-MVT algorithm is representative of the state-of-the-art in the minimum volume simplex fitting approaches.

1.3. Proposed approach

We introduce the *minimum volume simplex analysis* (MVSA) algorithm for unsupervised hyperspectral linear unmixing. As the name suggests, MVSA belongs to the minimum volume class, and thus is able to unmix hyperspectral data sets in which the pure pixel assumption is violated.

Fitting a simplex of minimum volume to hyperspectral data is a hard nonconvex optimization problem, which may end up in a local minimum. To avoid poor quality local minima, a good initialization is of paramount importance. We initialize MVSA with an inflated version of the simplex provided by VCA [2], a pure pixel based algorithm. Although this initialization may be far from the optimum, we have observed that it is systematically in the attraction basin of a good quality local minimum. Furthermore, since VCA yields a simplex defined by spectral vectors existing in the data set, we can discard all the spectral vectors that are inside this simplex, what accelerates the algorithm. Moreover, by carefully choosing the inflating factor, the large majority of constraints related with the abundance source fractions become inactive, what contributes to speeding up the algorithm, as well.

Minimum volume simplex algorithms are very sensitive to outliers. To make MVSA robust to outliers and noise, we run a final step in which the abundance fraction positivity hard constraint is replaced by a hinge type soft constraint. This step, applied after having found the minimum volume simplex, preserves the good quality of local minima.

The paper is organized as follows. Section 2 introduces the core of MVSA algorithm. Section 3 illustrates aspects of the performance of MVSA approach with simulated data, and Section 4 ends the paper by presenting a few concluding remarks.

2. MINIMUM VOLUME SIMPLEX ANALYSIS ALGORITHM (MVSA)

Let $Y \equiv [y_1, \dots, y_N] \in \mathbb{R}^{p \times n}$ be a matrix holding in its columns the spectral vectors $y_i \in \mathbb{R}^p$, for $i = 1, 2, \dots, n$, of a given hyperspectral data set. Although not strictly necessary, we assume in this version of the algorithm that a dimensionality reduction step (see, e.g., [17]) has been applied to the data set and the vectors $y_i \in \mathbb{R}^p$ are represented in the signal subspace spanned by the endmember spectral signatures. Under the linear mixing model, we have

$$\begin{aligned} Y &= MS \\ \text{s.t.} \quad & S \succeq 0, \quad 1_p^T S = 1_n^T, \end{aligned} \quad (1)$$

where $M \equiv [m_1, \dots, m_p] \in \mathbb{R}^{p \times p}$ is the mixing matrix (m_i denotes the i th endmember signature and p is the number of endmembers), and $S \in \mathbb{R}^{p \times n}$ is the abundance matrix containing the fractions ($[S]_{i,j}$ denotes the fraction of material m_i at pixel j). For each pixel, the fractions should be no less than zero, and sum to 1, that is, the fraction vectors belong to the probability simplex. Therefore, the spectral vectors y_i belong, as well, to a simplex set with vertices m_i , for $i = 1, \dots, p$.

Given Y , and inspired by the seminal work [6], we infer matrices M and S by fitting a minimum volume simplex to the data subject to the constraints in (1). This can be achieved by finding the matrix M with minimum volume defined by its columns under the constraints in (1). It can be formulated as the following optimization problem:

$$\begin{aligned} M^* &= \arg \min_M |\det(M)| \\ \text{s.t.} \quad & QY \succeq 0, \quad 1_p^T QY = 1_n^T, \end{aligned} \quad (2)$$

where $Q \equiv M^{-1}$. Since $\det(Q) = 1/\det(M)$, we can replace the problem (2) with the following:

$$\begin{aligned} Q^* &= \arg \max_Q \log |\det(Q)| \\ \text{s.t.} \quad & QY \succeq 0, \quad 1_p^T QY = 1_n^T. \end{aligned} \quad (3)$$

Optimizations (2) and (3) are nonlinear, although the constraints are linear. Problem (2) is non-convex and has many local minima. So, problem (3) is non-concave and has many local maxima. Therefore, there is no hope in finding systematically the global optima of (3). The MVSA algorithm, we introduce below aims at “good” sub-optimal solutions of optimization problem (3).

Our first step is to simplify the set of constraints $1_p^T QY = 1_n^T$ by noting that every spectral vector y in the data set can be written as a linear combination of p linearly independent vectors taken from the data set, say $Y_p = [y_{i_1}, \dots, y_{i_p}]$, where the weights add to one: i.e., $y = Y_p \beta$, where $1_p^T \beta = 1$. It turns out then, the constraint $1_p^T QY = 1_n^T$ is equivalent to $1_p^T QY_p = 1_n^T$ or else to $1_p^T Q = 1_p^T (Y_p)^{-1}$. Defining $q_m = 1_p^T (Y_p)^{-1}$, we get the equality constraint $1_p^T Q = q_m$. Then, the problem (3) simplifies to

$$\begin{aligned} Q^* &= \arg \max_Q \log |\det(Q)| \\ \text{s.t.} \quad & QY \succeq 0, \quad 1_p^T Q = q_m. \end{aligned} \quad (4)$$

We solve the optimization problem (4) by finding the solution of the respective Kuhn-Tucker equations using a sequential quadratic programming (SQP) methods. This methods belongs to the constrained Newton (or quasi-Newton) and guarantee superlinear convergence by accumulating second-order information regarding the Kuhn-Tucker equations [18]. Each quadratic problem builds a quadratic approximation for the Lagrangean function associated to (4). For this reason, we supply the gradient and the Hessian of f in each SQP iteration.

Usually, the hyperspectral data sets are huge and, thus, the above maximization is heavy from the computational point of view. To lighten the MVSA algorithm, we initialize it with the set of endmembers $M \equiv [m_1, \dots, m_p]$ generated by the VCA [2] algorithm. We selected VCA because its is the fastest among the state-of-the-art pure pixel-based methods. Since the output of VCA is a set of p vectors that are in the data set, then we can discard all vectors belonging to the convex set generated by the columns of M . If the number of endmembers is high, it may happen that the initial simplex provided by VCA contains very few pixels inside and, therefore, most are outside, violating the nonnegativity constraints and slowing down the algorithm. In such cases, we expand the initial simplex to increase the number of pixels that are in the convex hull of the identified endmembers, which speeds up the algorithm. The pseudocode for the MVSA method is shown in below. Symbols $g(Q)_{:,j}$ and $g(Q)_{i,:}$ stand for, respectively, the j th column and the i th line of $g(Q)$, the gradient of $f(Q)$.

Algorithm: Minimum Volume Simplex Analysis (MVSA)

Input: $p, Y, (f(Q) \equiv \log |\det(Q)|)$

Output: matrix Q

- 1: $Q_0 := \text{vca}(Y, \text{'Endmembers'}, p)$
- 2: $Q_0 := \text{expand}(M)$;
- 3: $Y := \text{discard}(Y)$; if y is inside the simplex
- 4: Inequality constraint
 $A * Q \geq b, A = Y^T \otimes I_p, b = 0_{pn}$
- 5: Equality constraint
 $Aeq * Q = beq, Aeq = I_p \otimes 1_p^T, beq = q_m^T$

- 6: $g(Q) := -(Q^{-1})^T$, where $g(Q)$ is the gradient of f
- 7: $[H(Q)]_{i,j} := -[g(Q)_{:,j} * g(Q)_{i,:}]$,
where $H(Q)$ is the Hessian matrix of f
- 8: $Q := \text{SQP}(f, Q_0, A, b, Aeq, beq, g, H)$

Based on experimental evidence, we have come to the conclusion that the complexity of the MVSA algorithm is roughly $O(p^3)$, provided that the initial Q is a feasible solution. Otherwise, the complexity depends on the number active constraints. This is the reason why we start the algorithm with VCA, discard the spectral vectors that are inside the inferred initial simplex, and expand it.

3. EXPERIMENTAL RESULTS

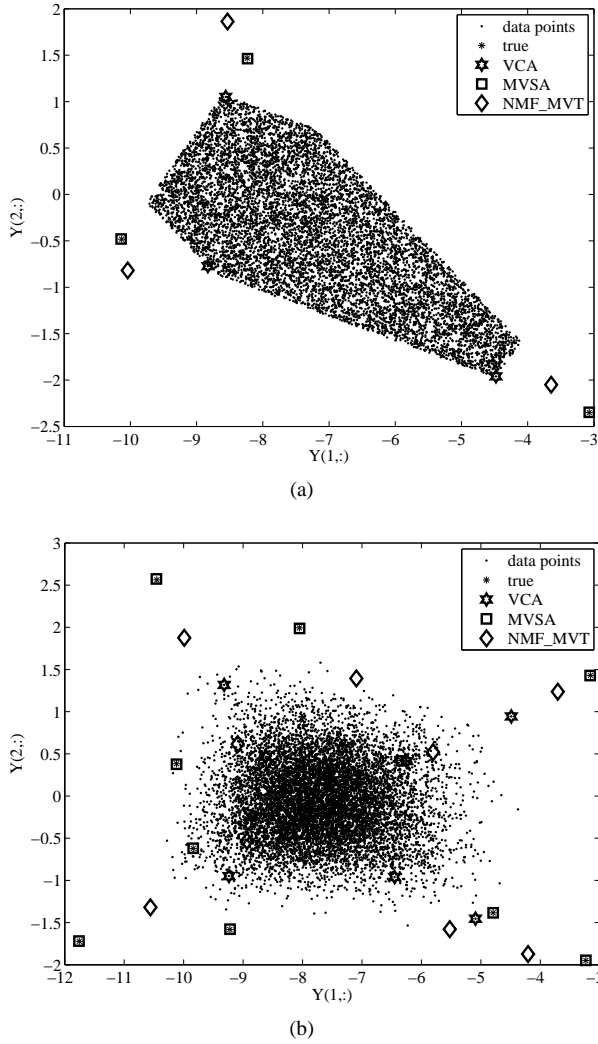


Fig. 1. Unmixing results for (a) $p = 3$ and (b) $p = 10$ number of endmembers for MVSA, MNF-MVT, and VCA algorithms. Dots represent spectral vectors; all other symbols represent inferred endmembers by the unmixing algorithms. Notice the quality of MVSA estimates.

This section presents results obtained by MVSA, VCA, and MNF-MVT unmixing algorithms applied to simulated data sets.

Table 1. Comparison of MVSA and NMF-MVT algorithms for different number of endmembers and sample size $n = 5000$. The time is in seconds and $\|A\|_F$ stands for the Frobenius norm of matrix A .

p	MVSA		NMF-MVT	
	$\ \widehat{M} - M\ _F$	time (sec.)	$\ \widehat{M} - M\ _F$	time (sec.)
3	0.01	4	0.876	153
5	0.04	5	0.785	344
10	0.06	74	5.154	730

Fig. 1 shows a projection on a subspace of the true endmembers, the endmembers inferred by MVSA, VCA, and MNF-MVT, and the spectral vectors. The data set has size $n = 10000$ pixels and a number of endmembers $p = 3$, part a), and $p = 10$, part b). The data is generated according to the linear observation model (1). The abundance fractions are Dirichlet distributed with parameter $\mu_i = 1$, for $i = 1, \dots, p$. The spectral signatures of the endmembers are mineral reflectances, with 224 spectral bands, obtained from a library. To ensure that no pure pixel is present, we discarded all pixels with any abundance fractions larger than 0.8. Notice the high quality of the MVSA estimates in both scenarios: the stars representing the true endmembers are all inside the squares representing the MVSA estimate. The VCA produces the worst estimate, as it was not conceived for data sets failing the pure pixel assumption.

Table 1 shows the times in seconds and the Frobenius norm $\|\widehat{M} - M\|_F$ of the endmember matrix estimates yielded by the MVSA and NMF-MVT algorithms. The algorithms run in a 3.4GHz Pentium 4 PC. MVSA performs much better with respect to both time and error. However, concerning the time complexity, and for the sample size $n = 5000$, the time MVSA takes gets larger than the NMF-MVT time for, roughly, $p > 15$.

3.1. Robustness to outliers and noise

When there are outliers and noise in the data set, we run a final step in which we replace the hard constraint $QY \succeq 0$ with the soft constraint $-1^T \text{hinge}(-QY)1_n$, where $\text{hinge}(x)$ is an element-wise operator that, for each component, yields the negative part of x . The modified optimization problem is

$$Q^* = \arg \max_Q \log |\det(Q)| - \lambda 1^T \text{hinge}(QY)1_n \quad (5)$$

$$\text{s.t.} : \quad 1_p^T Q = q_m,$$

where λ controls the relative weight between the soft constraint and the $\log |\det(Q)|$ term. Notice that, this soft constraint gives zero weight to nonnegative abundance fractions and negative weight to negative abundance fractions. In this way there is slack for the abundance fractions originated in outliers or noise to be negative.

To solve (5), we apply again SQP to the new objective function, but now removing the inequality constraint, *i.e.*,

$$Q := \text{SQP}(f_{soft}, Q_0, Aeq, beq, g, H),$$

where f_{soft} is the new objective function, Q_0 is the output of steps 1 to 8 shown at the end of Section 2, and Aeq, beq, g, H are defined as before.

We applied this robust version of the MVSA algorithm to the data set described above, with $n = 5000$ and $p = 3$, but now introducing additive zero-mean Gaussian noise to the spectral vectors such as the SNR $\equiv \|A\|_F^2 / \|w\|_F^2$ (w denotes the noise cube) was set to 10 dB. The errors $\|\widehat{M} - M\|_F$ of the MVSA and NMF-MVT

estimated endmember matrices were of 0.2 and 1.2, respectively. Fig. 2 shows the results. Notice the good performance of the MVSA algorithm. This are just very preliminar results that, nevertheless, illustrates the potential of this soft constraint tool.

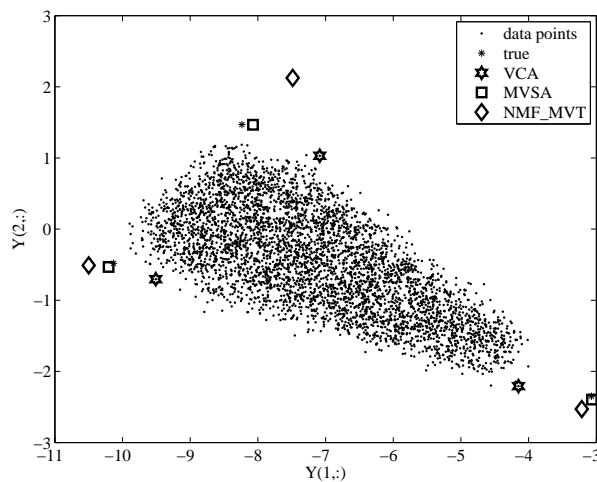


Fig. 2. Noisy scenario. As in Fig. 1 for $n = 5000$, $p = 3$, and $\text{SNR} = 10$ dB.

4. CONCLUSIONS

We have introduced the minimum volume simplex analysis (MVSA) algorithm, a new method to unmix hyperspectral data, under the linear mixing model. MVSA fits a minimum volume simplex to the data set, imposing positivity and sum to one constraints on the abundance fractions. The resulting optimization problem is solved by finding the solution of the respective Kuhn-Tucker equations using a sequential quadratic programming (SQP) method.

A shortcoming of the minimum volume simplex framework is that even a single outlier may force the simplex of minimum volume to be far away from a reasonable solution. To cope with outliers and noise, we have introduced a robust version of the MVSA algorithm. In this version, the positivity hard constraint imposed on the abundance fractions was replaced by a soft constraint of hinge loss type. This formulation seeks for a minimum volume simplex where most abundance fractions are nonnegative allowing, however, some may be negative.

The effectiveness of the new method was illustrated in a limited comparison with the state-of-the-art *non-negative matrix factorization* method [5], where the MVSA method yielded very competitive results.

5. REFERENCES

[1] D.G. Manolakis N. Keshava, J.P. kerekes and G.A. Shaw, "Algorithm taxonomy for hyperspectral unmixing," *Proc. SPIE Vol.4049, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery*, vol. VI, pp. 42, 2000.

[2] J. Nascimento and J. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 898–910, 2005.

[3] R.M. Perez A. Plaza, P. Martinez and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 650–663, 2004.

[4] J. Nascimento and J. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data?," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 175–187, 2005.

[5] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 765–777, 2007.

[6] M. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 542–552, 1994.

[7] J. Nascimento and J. Bioucas-Dias, "Hyperspectral unmixing algorithm via dependent component analysis," *IEEE International Geoscience and Remote Sensing Symposium*, pp. 4033–4036, 2007.

[8] N. Dobigeon, J.-Y. Tourneret, and C.-I. Chang, "Semi-supervised linear spectral unmixing using a hierarchical bayesian model for hyperspectral imagery," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 2684–2695, 2008.

[9] S. Moussaoui, H. Hauksdóttir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Douté, and J. A. Benediksson, "On the decomposition of mars hyperspectral data by ica and bayesian positive source separation," *Neurocomputing*, 2008, accepted.

[10] M. E. Winter, "N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. of the SPIE conference on Imaging Spectrometry V*, vol. 3753, pp. 266–275, 1999.

[11] J. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," in *JPL Pub.93-26, AVIRIS Workshop*, vol. 1, pp. 11–14, 1993.

[12] R. Perez A. Plaza, P. Martinez and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, pp. 2025–2041, 2002.

[13] C.-I. Chang, C.-C. Wu, W. Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 10, pp. 2804–2819, 2006.

[14] Chong-Yung Chi, "Non-negative least-correlated component analysis for separation of dependent sources," *invited talk at the Workshop on Optimization and Signal Processing, The Chinese University Hong Kong, Hong Kong*, 2007.

[15] J.Skaf M. Parente A.Zymnis, S.-J.Kim and S.Boyd, "Hyperspectral image unmixing via alternating projected subgradients," *Proceedings Asilomar Conference*, 2007.

[16] Liming Zhang Xutao Tao, Bin Wang and Jian Qiu Zhang, "A new scheme for decomposition of mixed pixels based on non-negative matrix factorization," *IEEE International Geoscience and Remote Sensing Symposium*, pp. 1759–1762, 2007.

[17] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, 2008.

[18] R. Fletcher, *Practical Methods of Optimization*, John Wiley and Sons, 1987.