Hyperspectral Unmixing based on Constrained Nonnegative Matrix Factorization via Approximate L_0

Tai Gao[1, a], Yang Guo[1, b], Chengzhi Deng[2], Shengqian Wang[2] and Qing Yu[2, c]

[1]School of Jiangxi Science & Technology Normal University, Nanchang, China

[2]School of Nanchang Institute of Technology, Nanchang, China

[a]gaotaihaoa@163.com, [b]accuracygy@gmail.com, [c]qingyu@live.cn

Keywords: Hyperspectral unmixing, AL_0-NMF, sparsity, projected gradient.

Abstract. Hyperspectral unmixing is estimating the endmembers and corresponding abundance fractions in a mixed pixel. In the past decade, NMF have been intensively studied to hyperspectral unmixing. As an important constraint for NMF, sparsity could be modeled making use of the L0 regularize. Unfortunately, the L0 regularize is an N-P hard. In this paper, we uses a novel approximate L0 sparsity constraint (which we name AL0-NMF), we propose a project gradient algorithm for AL0 -NMF. The experimental based on synthetic and real data demonstrate the effectiveness of the propose method.

Introduction

Hyperspectral data is a spectal image cube, containing hundreds of spectral bands and spatial information .Owing to the low spatial resolution of the sensor, there are many mixed pixels in the remote sensing image. Hyperspectral unmixing which decomposes mixed pixels into a collection of constituent spectra, or endmembers, and the corresponding abundance fractions is often used to preprocess hyperspectral data [1].In the past few decades, many hyperspectral unmixing algorithms have been propose under the linear mixing model(LMM) such as vertex component analysis (VCA) [2], independent component analysis (ICA) [3], etc.

As a widely used method of blind source separation (BBS), nonnegative matrix factorization (NMF) [4] can adopted to solve the hyperspectral unmixing. Unfortunately, due to the objective function of NMF is nonconvexity, a lot of local minimum occur. An easy but effective solution to reduce the problem is to introduce further constraints into the NMF algorithm. Now, some researchers used sparsity constraint on basic NMF, like [5]. Regularization method are usually utilized to define the sparsity constraint on the abundance matrix of the endmember. In general, L_1 norm is widely used to instead of L_0 norm. But for L_1 regularization, it will be always a constraint due to the abundances suffer from the sum to one constraint. The L_0 regularization can yield a sparser results, while it is an NP-hard. In this paper, we introduce a smoothed function to approximate the L_0 norm, which can enforce the sparsity of endmember abundances and avoid to solve the NP-hard problem.

In this paper, we introduce a new L_0 regularization into NMF (AL_0-NMF) to enforce the sparsity of abundance matrix. We used projected gradient methods in [6] to ensure convergence. In this approach, the SNC is embedded in the parameter update process. Through synthetic and real hyperspectral data experiments results, we observe our algorithm is state of the art.

The rest of this paper is organized as follows. In Section 2, brief introduces AL_0-NMF model of the hyperspectral unmixing. Section 3 give the algorithm of hyperspectral unmixing Section 4 show the results of synthetic and real hyperspectral data. Sections 5 draws conclusions.

AL_0-NMF Unmixing Model

LMM and NMF. The LMM can be written as follows:

$$X = AS + E \tag{1}$$

Where $X \in R^{L \times N}$ denotes the hyperspectral data, $A \in R^{L \times P}$ denotes the endmember signatures, $S \in R^{P \times N}$ denotes the endmember abundances, $E \in R^{P \times N}$ denotes the additive noise. And L, N, P denotes the number of bands, the number of pixels and the number of endmembers in remote sensing image, respectively.

Since NMF can lead to the part-based linear representations of the nonnegative high-dimensional data, it has received a lot of attention. NMF aims to decompose one nonnegative matrix X into two low-rank nonnegative matrices A and S .So we can solve the NMF problem as an optimization problem by minimizing the Euclidean distance. Cosidering the Abundance Nonnegative Constraint (ANC) and sum to one, the cost function is as follows:

$$f(A,S) = \frac{1}{2}\|X - AS\|_F^2 \quad \sum_{P=1}^{P} S_{pn} = 1, S_{pn} \geq 0, p = 1, 2, ..., P \tag{2}$$

There are numerous optimization algorithms to solving the problem (2), but due to the non-convexity of represent A and S together, it exists a lot of local minimum. In order to reduce the solution set, some additional constraints should be introduced. Sparsity is an intrinsic property of hyperspectral data. In most cases, the mixed pixel is usually the superposition of only a few endmembers, the abundance is localized with a degree of sparseness [7].

For this reason, we used sparse constraints into NMF. The cost function with sparse constraints is as follow:

$$f(A,S) = \frac{1}{2}\|X - AS\|_F^2 + \lambda f(S) \tag{3}$$

AL$_0$-NMF. In the recently, many forms of sparsity constraints have been used, like[8-9].From these paper, we know L$_1$ norm used most due to the L$_1$ norm easy to be solved and L$_0$ norm has the sparsest result, but it is a NP-hard problem. In this paper, an approximate L$_0$ norm we choose [11] and the approximate L$_0$-NMF model (AL$_0$-NMF) is given by

$$f(A,S) = \frac{1}{2}\|X - AS\|_F^2 + \lambda\|S\|_0 \quad \|S\|_0 \approx \sum_{p=1}^{P} \sum_{n=1}^{N} \frac{2}{\pi} \arctan\left(\frac{|s_{pn}|}{\sigma^2}\right) \tag{4}$$

And s_{pn} is the abundance matrix for the row of p at the column of n.

Algorithm for AL0-NMF

Updating Rules. Main text paragraph. We know, the cost function (4) is not convex with respect to A and S together, but they are convex in A or S only. To solve the problem many algorithm has put forward, like multiplicative iterative rules, gradient descent, projected gradient methods [5] and so on. Here, projected gradient algorithm is used on this paper.

To implement the projected gradient algorithm, we need to calculate the partial derivative:

$$\nabla_A f(A,S) = ASS^T - X S^T \quad \nabla_S f(A,S) = A^T AS - A^T X + \frac{2}{\pi \times \sigma^2}\lambda \times sign(S) \times \frac{1}{1 + (\frac{abs(S)}{\sigma^2})^2} \tag{5}$$

In our experiments σ is set to 0.001.
The update rule is
$$A^{k+1} = \max(0, A^k - \alpha^k \nabla_A f(A,S)) \quad S^{k+1} = \max(0, S^k - \beta^k \nabla_S f(A,S)) \tag{6}$$

Where the parameters α^k and β^k are the learning step sizes selected based on the Armijo rule [11].The function max (0, x) is set the negative components to zero to satisfy the ANC.

Implementation Issues. SNC: We adopt the method in [12] to ensure the SNC, which is widely used in hyperspectral unmixing algorithms. We augment the matrix X and matrix A by a row constant denoted by

$$X_f = \begin{bmatrix} X \\ \delta 1_N^T \end{bmatrix} \qquad A_f = \begin{bmatrix} A \\ \delta 1_N^T \end{bmatrix} \tag{7}$$

Where δ is a positive number to control the impact of SNC? In our experiments, we find when we set δ to 20, the abundances almost satisfy the SNC and our experiments has the best result.

Initialization: Random initialization and VCA-FCLS is widely used to initialize matrices A and S. In this paper, VCA-FCLS initialization is employed in experiments to speed up the convergence of the proposed algorithm. While VCA-FCLS to initialization, we should to estimate the number of endmember. Here, we resort Hysime algorithm [13] to estimate the number of endmember.

Stopping criteria: We have two stopping criteria in our experiments. One is the maximum iteration number, which is set to 500 in our experiments. The other is the gradient of the cost function f between the current iteration and the starting should satisfy:

$$\left\| \nabla f(A^i, S^i) \right\| \leq \varepsilon \left\| \nabla f(A^1, S^1) \right\|_F^2 \tag{8}$$

Where ε is set to 10^{-4} in our experiments. Either of the criteria is met, the iteration stop.

Parameter: The parameter λ is relies on the sparseness degrees of the material abundances, which cannot be obtained a prior, we use a possible estimator based on sparseness measure in [5].

$$\lambda = \frac{1}{\sqrt{L}} \sum_l \frac{\sqrt{N} - \|x_l\|_1 / \|x_l\|_2}{\sqrt{N-1}} \tag{9}$$

Where x_l denotes the hyperspectral data at band l.

The AL_0-NMF algorithm is summarized as follows.

Algorithm: AL0-NMF
1: Estimate the number of endmembers P using the HySime algotithm and compute λ by (10).
2: Initialize A and S by VCA-FCLS.
3: while stopping criteria are not met do
a): Update Ak+1 by (7)
b): Augment X and A to get Xf and Af.
c): Updata Sk+1 by (7)
End

Experiments

In this section, a series of experiments are used to test the performance of the proposed AL_0-NMF algorithm. We will compared our algorithm with VCA-FCLS and L_1-NMF.

Two metrics: spectral angle distance (SAD) and root mean square error (RMSE), which is most widely used will used to evaluate the performance of the algorithm. The SAD is used to compare the similarity of the pth true endmember signature A_p and its estimate $\hat{A}_p$. For the RMSE, which is used to compare the similarity of the pth true endmember abundance S_p and its estimate $\hat{S}_p$. For the RMSE, which is used to compare the similarity of the pth true endmember abundance S_p and its estimate $\hat{S}_p$. The SAD and RMSE are defined as

$$SAD_p = arc\cos\left(\frac{A_p^T \hat{A}_p}{\|A_p\|\|\hat{A}_p\|}\right) \quad RMSE_p = (\frac{1}{N}\left|S_p - \hat{S}_p\right|^2)^{\frac{1}{2}} \tag{10}$$

In our experiments, the smaller SAD and RMSE are, the better the algorithm is. In the following experiments, the result is acquired by averaging 30 random tests.

To generate the synthetic endmember signature data, five spectral signatures are chosen from the United States Geological Survey (USGS) digital spectral library. The abundances matrix data generated can be describes as follows: 1) We divide an image of size $r^2 \times r^2 (r \in Z^+)$ into $r \times r$ blocks and each block is initialized with one type of endmember signature; 2) we utilized a $(r+1) \times (r+1)$ low pass filter to generate mixed pixels; 3) set a threshold θ (in this paper we set $\theta = 0.7$), for the pixel whose abundances is larger than θ a mixture composed of only two endmembers will take the place of it; 4) to simulate the noise in (1), zero-mean white Gaussian is added into the mixture data. In this

experiment, we aims to compare the three algorithms VCA-FCLS, L_1-NMF and AL_0-NMF with the different noise levels. The SNR is set to 20 and 50.

Table 1 the results of SAD and RMSE with different unmixing methods

	SNR	VCA-FCLS	L1-NMF	AL0-NMF
SAD	20	0.2057	0.1548	0.0332
	50	0.1782	0.1392	0.0162
RMSE	20	0.1436	0.0760	0.0375
	50	0.1304	0.0601	0.0224

From the table 1, we can see that our algorithm has the best results for both the SAD and the RMSE. That means our algorithm not only robust to the noise corruption, bur also could better represents the sparseness of the abundance matrix than L_1-NMF.
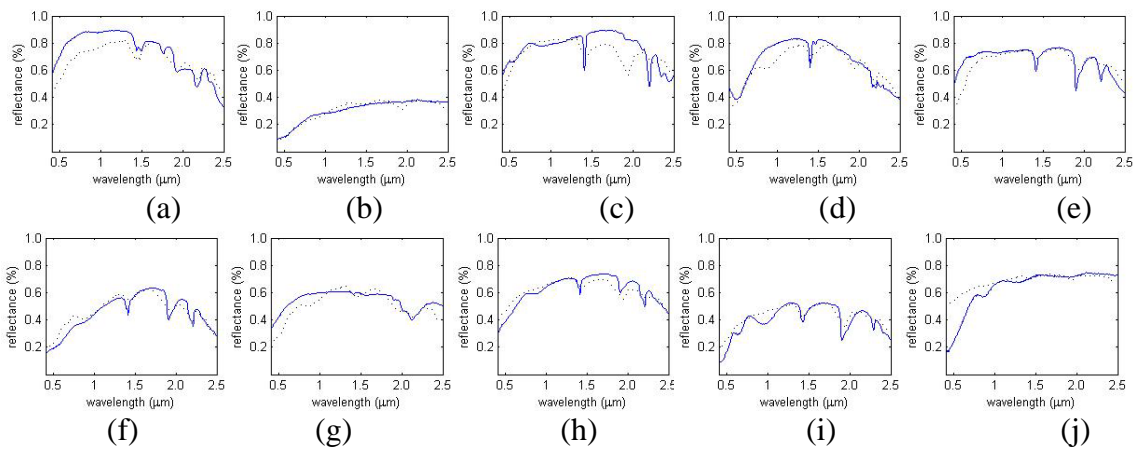

Figure 2. Comparison of USGS library spectra(solid line) with (dotted line) the extracted endmember Comparison of USGS library spectra(solid line) with (dotted line) the extracted endmember signatures using our method.(a) Alunite.(b) Sphene.(c) Muscovite.(d) Dumortierite(e) Montmorillonite.(f) Kaolinite#1.(g) Buddingtonite. (h)Kaolinite#2. (i) Nontronite. (j) Pyrope.

We applied the real hyperspectral data, which is an image acquired by the Airborne Visible /Infrared Imaging spectrometer (AVIRIS) over Cuprite, Nevada, to evaluate the propose method. The cuprite data have been widely used for hyperspectral unmixing research [7]. The size of original image is 250*191. The data has 224 bands that cover the wavelength ranging from 0.4 to 2.5μm. In order to improve the precision, we have remove the low SNR and water vapor absorption bands (1-2, 104-113, 148-167 and 221-224). The other 188 bands are used in our experiment. According to the existing analysis in [2], 14 types of minerals are presented in the image. It is obviously seeing that variants of the same mineral with slightly different spectra can be regarded as the same endmember, so, in the experiments we set p=10. In Fig.2, we show the the comparison between the extracted endmember signatures by our experiments and the corresponding USGS spectra library.

Conclusions

In this paper, we proposed a new approximate L_0-NMF sparsity constrained NMF to solve the hyperspectral unmixing problem. Compare with the other algorithms, our method could make the unmixing results more accurate. Although our method get a better results, we still needs further investigation in other hyperspectral applications.

References

[1]   N. Keshava, "A survey of spectral unmixing algorithms" Lincoln Lab. J., 2003.14, no.1, 55-78.

[2]   J. M. P. Nescient and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmixed hyperspectral data," IEEE Trans. Geosci. Remote Sens., 2005. 43, 4, 898–910.

[3]   J. Wang and C.-I. Chang, "Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery," IEEE Trans. Geosci. Remote Sens., 2006. 44, no. 9, 2601-2616.

[4]   D. D. Lee and H. S. Seung, "Algorithms for nonnegative matrix factorization," MIT Press: Cambridge MA, USA, 2001; 13,556-562.

[5]   P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," Journal of Machine Learning Research, 2004. 5, 1457–1469.

[6]   C. J. Linm, "Projected gradient methods for non-negative matrix factorization," Information and Support Services, 2005. 1, 47–4.

[7]   "Hyperspectral unmixing via L1/2 sparsity-constrained nonnegative matrix factorization," IEEE Trans. Geosci. Remote Sens., 2011. 49, 4282–4297.

[8]   M. D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," IEEE Trans. Geosci. Remote Sens., 2011. 49, 2014–2039.

[9]   M. D. Iordache, A. Plaza, and J. Bioucas-Dias, "Recent developments in sparse hyperspectral unmixing," in Proc. IGARSS, 2010, 1281–1284.

[10]   Z. Guo, T. Wittman, and S. Osher, "L1 unmixing and its application to hyperspectral image enhancement," in Proc. SPIE—Algorithms and Technologies for Multispectral, Hyperspectral, and Ultra spectral Imagery, p. 733 41M, 2009.

[11]   M. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods. New York: Academic, 1982.

[12]   D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," IEEE Trans. Geosci. Remote Sens., 2001. 39, 529–545.

[13]   J. M. B. Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," IEEE Trans. Geosci. Remote Sens., 2008. 46, 2435–2445.