Linear spectral mixture analysis with the Open Leontief Input-Output Model

ZHU Lijiang[1]

National Key Lab on Remote Sensing Information and Image Analysis Technology, Beijing Research Institute of Uranium Geology, P.O. Box 9818, Beijing 100029, China

zljrenwen@hotmail.com

Abstract. Commonly, it requires two constraints imposed on the linear spectral mixture analysis (LSMA).One constraint is the abundance of sum-to-one, which requires the abundance fractions of materials presented in an image pixel to be one and is easy to deal with. The other constraint is that any abundance fractions are always nonnegative and difficult to solve with analytical solution. Most of approaches that provide the solution for the latter problem of LSMA use an optimization or maximization procedure. The results of solution resort to optimization strategies. The Leontief input-output model, of which parameters are very similar to LSMA, is represented by a linear system of equations and the system has a unique nonnegative solution. In this paper, we considered how to determine the parameters of LSMA model, and based on the open Leontief input-output model, we presented a fully constrained linear spectral (FCLS) mixture analysis method for estimating material abundance in spectral mixture pixel. The new FCLS method can not only make the abundance fractions of materials be nonnegative, but also keep them less than one, that always obtained by normalizing procedure in other methods. We also examine a number of approaches, previous FCLS and non-negative matrix factorization (NMF) spectral un-mixing, closely related. A series of computer simulations are conducted to demonstrate the performance of the proposed method in material quantification.

1. Introduction

In recent years, many different algorithms were developed for spectral un-mixing of hyper-spectral imagery. Most spectral un-mixing algorithms think an observed spectrum is linear combination of a limited number of ‘pure’ spectra, where the coefficient of each ‘pure’ spectrum is its abundance in the measured pixel of the imagery scene.

Over time, most algorithms for spectral un-mixing need to perform two steps, the first step is to find ‘pure’ spectrums, which is usually labeled as end member extraction. The other step is to determine the corresponding abundances of each ‘pure’ spectrum in the mixed imagery pixels. The latter is usually labeled as spectral un-mixing, or the spectral mixture analysis. Sometimes the ‘pure’ spectrums of end member can be easily got from the spectral library of JPL, USGS etc. So, our paper focuses only on the latter problem.

[1]This article is supported by the national science and technology support program of China (No.2012BAH27B01)

An approach, which is first proposed by Chang etc, turns the un-mixing problem into a non-negatively constrained least-squares problem that can be solved using the Lawson and Hanson techniques. And this was the first solution obey the ANC and ASC constrains and called FCLS by Chang [1]. Soon after, other algorithms followed that also allowed to find this solution through other approach such as Quadratic programming, hierarchical Bayesian, non-negative matrix factorization (NMF), and so on. Most of the approaches that provide the FCLS solution use an optimization or maximization, or require a random sampler from some distribution like NMF. The result accuracy can be got arbitrarily by modifying the algorithmic parameters, they all do not provide an analytically solution [2].

In this paper, we investigate Linear Mixture Model with the open Leontief Input-output model, which can provide a nonnegative matrix and make solving the solution more easily.

2. Linear Mixture Model

Classical spectral mixture model can be represented as follows:

$$r = \mathbf{M}\alpha + n_0 \tag{2.1}$$

Where n_0 is noise, can be interpreted as a measurement error or some unknown spectral, which proportion in the pixel of same imagery scene is small. r is used to represent spectral reflectance vector of the pixel. Matrix $\mathbf{M} = [\mathbf{m_1}...\mathbf{m_k}]$ is used to represent spectral reflectance of different material in the same imagery scene and α is used to represent the abundance of different material associated with r.

3. The open Leontief input-output model

The open Leontief input-output model, which was originally developed by the economist Wassily Leontief, can be denoted as follows:

$$x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + d_i \tag{3.1}$$

Where $d_j \geq 0$, for $j = 1, \ldots, n$. This leads to the linear system

$$\begin{array}{ccccccccc} (1-a_{11})x_1 & + & (-a_{12})x_2 & + & \ldots & + & (-a_{1n})x_n & = & d_1 \\ (-a_{21})x_1 & + & (1-a_{22})x_2 & + & \ldots & + & (-a_{2n})x_n & = & d_2 \\ & & \vdots & & & & & & \\ (-a_{n1})x_1 & + & (-a_{n2})x_2 & + & \ldots & + & (1-a_{nn})x_n & = & d_n \end{array} \tag{3.2}$$

Where $a_{ij} \geq 0$ for each i and j, $\sum_{i=1}^{n} a_{ij} < 1$.The solution vector X must be nonnegative and unique. We can get this proof in [3].

It is notable that the number of unknowns of open Leontief input-output model is equal to the number of equations. And for the LSMA of hyper-spectral un-mixing, the number of unknowns is always far greater than the number of equations.

4. The proposed LSMA method in this paper

What's the relationship between the model of LSMA and the open Leontief input-output system? Think that r in equation (2.1) represents reflectance of all spectral bands we measured, then

$r = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$, and $0 \leq r_j \leq 1, \dfrac{\sum r_j}{n} < 1$, where n is number of spectral bands we used. So the equation (2.1) can be written in the following form

$$\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} m_{11} & \cdots & m_{k1} \\ & m_{ij} & \\ m_{1n} & \cdots & m_{kn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \tag{4.1}$$

Value $\sum \alpha_i$ is less than one with the ASC constraint.

Because the solution of this equation is not always analytical, so the noise can be omitted from the equation (4.1) in the least squares sense.

$$\begin{pmatrix} m_{11} & \cdots & m_{k1} \\ & m_{ij} & \\ m_{1n} & \cdots & m_{kn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \tag{4.2}$$

Where $\mathbf{m}_i = \begin{pmatrix} m_{i1} \\ \vdots \\ m_{in} \end{pmatrix}$ is represent spectral reflectance of material $\mathbf{m}_i$, $0 \le m_{ij} \le 1$ for any band, $\frac{\sum_j^n m_{ij}}{n} < 1$, the k is usually larger than n. Without any constraints, the least squares estimate of α for model LSMA is given by $\bar{\alpha} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T r$ .Matrix $\mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T$ is called the projection matrix, which project any vector r to the column space of matrix $\mathbf{M}$, the α_j can be negative. If we select some spectral bands to compute the α_j, maybe we can decrease the noise of the model. Typically, the selection makes the number of bands equal to the number of material signature.

As the fact of the reflectance of any broad spectral band is less than 1, equation (4.2) can be rewritten as a linear system:

$$\begin{pmatrix} \frac{m_{11}}{n} & \cdots & \frac{m_{n1}}{n} \\ & \frac{m_{ij}}{n} & \\ \frac{m_{1n}}{n} & \cdots & \frac{m_{nn}}{n} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \frac{\mathbf{M}}{n}\alpha = \frac{r}{n} = \begin{pmatrix} \frac{r_1}{n} \\ \vdots \\ \frac{r_n}{n} \end{pmatrix} \tag{4.3}$$

We denote $\mathbf{A} = \frac{\mathbf{M}}{n}$ for convenient. The entries of coefficient matrix $\mathbf{A}$ have two important properties:

- $\frac{m_{ij}}{n} \ge 0$ for each i and j.
- $\sum_{i=1}^{n} \frac{m_{ij}}{n} < 1$ for each j.

The Open Leontief Input-output model has same properties. That is the critical point to make the solution nonnegative. So the last equation can be written as

$$(I - \mathbf{A})\alpha = \alpha - \frac{r}{n} \tag{4.4}$$

Then solution of equation (4.2) can be given by

$$\alpha = \left(I - \mathbf{A}\right)^{-1}\left(\alpha - \frac{r}{n}\right) \tag{4.5}$$

It is obvious that $\left(I - \mathbf{A}\right)^{-1}$ is nonnegative. An iteration procedure, which is similar to the iteration in [4], is used here to get the solution from equation (4.5), and we use $\bar{\alpha} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T r$ as the initial value of iteration procedure.

5. Computer Simulation and Experiment

In this section, a series of computer simulation and experiment to evaluate performance of the new methods were presented, as comparison, the previous FCLS, which is first proposed by Chang and the NMF method, were also used here. All of three methods are implemented in MATLAB.

Two experiments were designed to demonstrate the performance of the FCLS (Chang), NMF and the new FCLS (Leontief) method in following simulations. 1) When the information of all material spectral signatures is completely known with some noise information was added, 2) when some false information (other material signature, noise and so on) is used.

A set of reflectance spectra selected from ENVI spectral library was used for performance evaluation. This set of reflectance spectra contains five spectra, dray grass, sage brush, black brush, walnut leaf and dark reddish soil. These spectra in the range of 415nm -2500nm were convolved to 15nm spectral bands as shown in Figure 1.

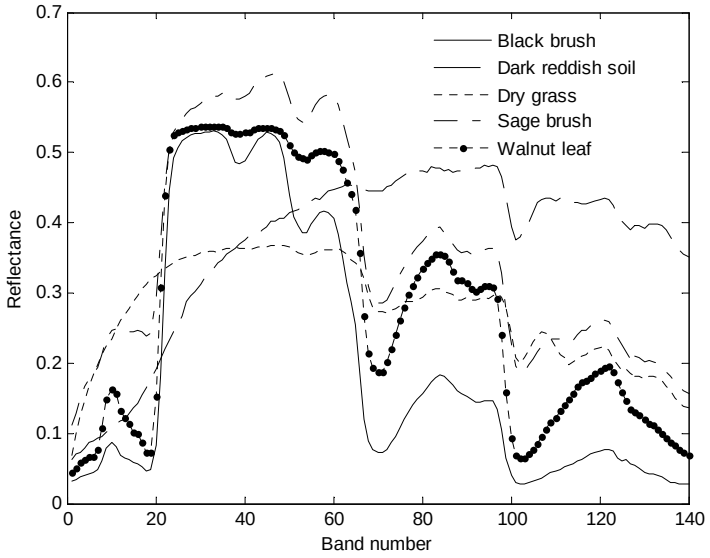

Figure 1. Material reflectance spectra used for mixing pixel simulation

Each simulated data consisted of 100 pixel vector, which were numbered from 1 to 100. First, we create five random data series with MATLAB function ‘rand’, the values of these data series are in range of 0-1 and the length of random data series is 100, which is same to the pixel number of the simulated data. Each series is treated as weight coefficient associated with a material abundance. Then the abundance of different material in the simulation data pixel can be easily got from follow equation:

$$\alpha_i = \frac{\mathbf{w}_i}{k} \tag{5.1}$$

Where $\mathbf{w}_i$ denotes the weight coefficient vector associated with a material abundance, k denotes the number of spectra involved in simulation pixel data. The pixel values of spectral mixing are given by $r = \sum_{1}^{k} \alpha_i \mathbf{m}_i$.

Finally, the White Gaussian noise was added to each pixel vector to achieve a 30:1 SNR, which was defined as average reflectance over all pixels divided by the standard deviation of the noise.

5.1. Experiment : all spectral reflectance is completely known

In this experiment, three reflectance spectra, dry grass, dark reddish soil and walnut leaf, were used to form a reflectance matrix $\mathbf{M} = [\mathbf{m}_1 \quad \mathbf{m}_2 \quad \mathbf{m}_3]$, with their associated abundance fractions given by $\alpha = (\alpha_1 \quad \alpha_2 \quad \alpha_3)^T$ .We denotes the solution as $\overline{\alpha}$,it is obviously that the smaller the difference $\overline{\alpha} - \alpha$,the better the performance. Ideally, the solution $\overline{\alpha}$ is equal to α .

In order to compare the performance of different methods, the same simulation data was also used to evaluate the quantification performance of NMF and FCLS (Chang) methods. All three methods are performed, only results of difference $\overline{\alpha} - \alpha$ in detection of walnut leaf leaves are shown in Figure 2.

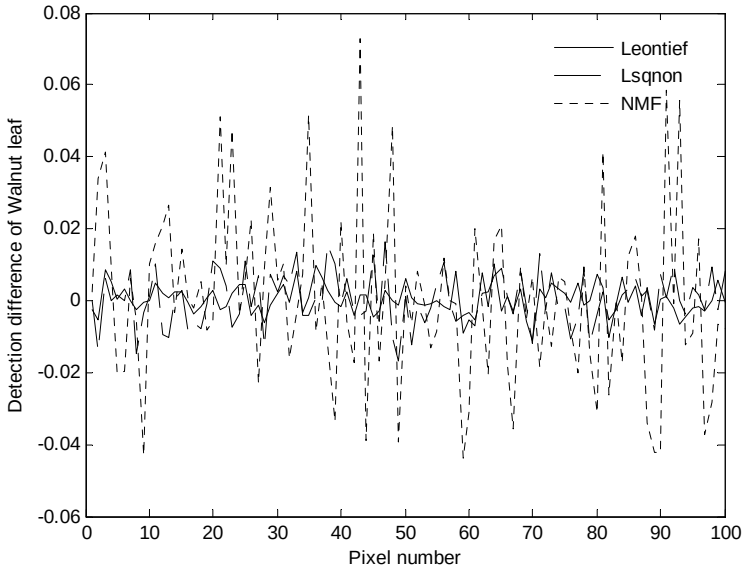

Figure 2. Different $\overline{\alpha} - \alpha$ of walnut leaf quantifications along pixels

The square sums of difference $\overline{\alpha} - \alpha$ are 0.0071 for FCLS (Leontief), 0.0052 for FCLS (Chang and 0.0576 for NMF. Obviously, two FCLS methods have a better performance than NMF methods. The FCLS (Chang) and FCLS (Leontief) performed nearly same. The FCLS (Chang) method has produced best result.

5.2. Experiment: all spectral reflectance are known, but some spectrums are not contained in the simulated data.

In this experiment, five reflectance spectra, dry grass, dark reddish soil, walnut leaf, black brush and sagebrush were used to form a reflectance matrix $\mathbf{M} = [\mathbf{m_1} \quad \mathbf{m_2} \quad \mathbf{m_3} \quad \mathbf{m_4} \quad \mathbf{m_5}]$, with their associated abundance fractions given by $\alpha = (\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5)^T$. To compare different methods, results of difference $\overline{\alpha} - \alpha$ in detection of walnut leaf leaves are computed and shown in Figure 3.

The results of three methods are 0.1077 for FCLS (Leontief), 0.0644 for FCLS (Chang) and 0.0680 for NMF. Unlike Figure 2, it is obviously that the detection performance of FCLS (Leontief) was considerably decreased.

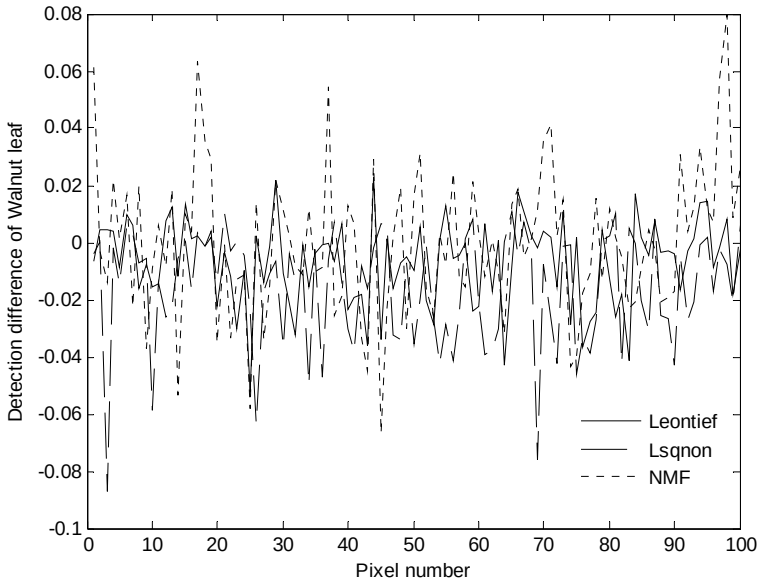

Figure 3. Different $\overline{\alpha} - \alpha$ of walnut leaf quantifications along pixels

But it worth noted that the results of Leontief are more smooth or robust in two experiments.

6. Conclusion

In this article, we have investigated the linear spectral mixture with the Leontief model. After considering how to determine the parameters of LSMA model, we presented a fully constrained linear spectral (FCLS) mixture analysis method for estimating material abundance in spectral mixture pixel. Because of the matrix $\left(I - \mathbf{A}\right)^{-1}$ is nonnegative, the solution of the LSMA model can be kept nonnegative more easily.

Two computer experiments were designed and conducted to demonstrate the performance of the FCLS (Chang), NMF and the new FCLS (Leontief) method. The FCLS (Chang) method has produced best results. The new FCLS (Leontief) method is better than NMF method in first experiments, and worse in second experiments. But the new FCLS (Leontief) method has minor fluctuation in two experiments.

To conclude, the new method seems to be robust or smooth in estimating the abundance fractions of walnut leaf spectrum in computer simulations and experiments. However, the precision of result is unsatisfied so this method needs further investigation to improve the results.

References

[1] Daniel C Heinz, Chein-I Chang 2001 Fully constrained least squares linear spectral mixture analysis method for material quantification in hyper spectral imagery *IEEE Trans. Geosci. Remote Sensing* **39** 529-545

[2] S Jia, Y Qian 2009 Constrained nonnegative matrix factorization for hyper spectral unmixing *IEEE Trans. Geosci. Remote Sensing* **l47** 161-173

[3] Steven J. Leon 2006 *Linear Algebra with Applications* (Beijing: Pearson Education Asia Ltd.) 392-397

[4] C.-I Chang and D. Heinz 2000 Contrained subpixel target detection for remotely sensed imagery *IEEE Trans. Geosci. Remote Sensing* **38** 1144-1159