ESTIMATION OF THE HYPERSPECTRAL TUCKER RANKS

Alexis Huck, Mireille Guillaume

Institut Fresnel, Marseille (France)

ABSTRACT

In hyperspectral image analysis, one often assumes that observed pixel spectra are linear combinations of pure substance spectra. Unmixing a hyperspectral image consists in finding the number of pure substances in the scene, finding their spectral signatures and estimating the abundance fraction of each pure substance spectrum in each spectral pixel. In this paper, we show that the tensor Tucker decomposition could be considered to solve this problem, and a preliminary problem to overcome consists in estimating the 3 required data Tucker ranks, corresponding to the 3 dimensions of the data cube. Then, we propose an optimal method to estimate them.

Index Terms— Hyperspectral, Unmixing, Tensor, Ranks, Non-negative Tucker Decomposition (NTD)

1. INTRODUCTION

Spaceborn and airborn hyperspectral sensors acquire radiance images of a scene in hundreds of narrow and contiguous *spectral bands*, simultaneously. Such data can also be described as a collection of observed spectra contained in the L-dimensional *spectral pixels* of the HyperSpectral Images (HSIs). Thus, hyperspectral data contain two spectral dimensions and one spectral one, and are generally contained in a 3-dimensional array.

1.1. Linear Spectral Mixing Model (LSMM)

It is often relevant to assume that the i^{th} observed spectrum $\mathbf{r}^i$ is a linear mixture of the material spectra present in the respective pixel coverage: $\mathbf{r}^i = \mathbf{A}\mathbf{s}^i + \mathbf{n}^i$, where $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_J]$. The L-dimensional column vectors $\{\mathbf{a}_j\}_{j=1\ldots J}$ contain the sampled spectral signature of the J pure materials constituting the observed scene, also called *endmembers*. The column vector $\mathbf{s}^i = [s_{1i}, \ldots, s_{Ji}]^T$ holds the set of *abundance* fractions and $\mathbf{n}^i$ is an additive noise vector. For a set of mixed spectral pixels, this Linear Spectral Mixing Model (LSMM) [1–3] can be written as follows:

$$\mathbf{R} = \mathbf{A}\mathbf{S} + \mathbf{N} = \mathbf{X} + \mathbf{N} \ . \tag{1}$$

With such notations, the initial 3-dimensional data has been rearranged into a matrix. The I columns of $\mathbf{R}$ contain the spectral pixels and the I columns of $\mathbf{S}$ hold their respective sets of abundance fractions. Thus, the J rows of $\mathbf{S}$ are the *abundance maps* corresponding to the endmembers. At last, $\mathbf{N}$ is the noise matrix. We define:

$$\mathbf{X} = \mathbf{A}\mathbf{S} \tag{2}$$

as the *signal matrix*.

Obviously, the matrices $\mathbf{A}$ and $\mathbf{S}$ must satisfy three physical constraints induced by the LSMM model:

(C1) $\mathbf{A}$ must be non-negative;

(C2) $\mathbf{S}$ must be non-negative;

(C3) The columns of $\mathbf{S}$ must be summed-to-unity.

For practical purposes, $\mathbf{A}$ and $\mathbf{S}$ are generally unknown and one would like to estimate them from $\mathbf{R}$ only, under the LSMM assumption. This problem is called linear *spectral unmixing*. Among various strategies, a non-negative PARAFAC decomposition has shown success in [4] to unmix hyperspectral data. We believe that Non-negative Tucker Decomposition (NTD or HONMF for High Order Non-Negative Matrix Factorization) algorithms [5, 6] could also unmix hyperspectral data, finding a Tucker decomposition of the HSI such that the core tensor and the flattening matrices have only positive entries.

1.2. Tucker Decomposition

Let us consider a multidimensional matrix (or tensor) $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. A Tucker decomposition approximates $\mathcal{X}$ as follows:

$$\mathcal{X}_{i_1,i_2,\ldots,i_N} \approx \mathcal{L}_{i_1,i_2,\ldots,i_N} = \sum_{j_1,j_2,\ldots,j_N} \mathcal{G}_{j_1,j_2,\ldots,j_N} \mathbf{A}^{(1)}_{i_1,j_1} \mathbf{A}^{(2)}_{i_2,j_2} \ldots \mathbf{A}^{(3)}_{i_3,j_3} \tag{3}$$

where $\mathcal{G} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is the *core tensor* and the matrices $\mathbf{A}^{(n)} \in \mathbb{R}^{I_n \times J_n}$ are called the *flattening matrices*. $J_n, \ n = 1, \ldots, N$ are the *Tucker ranks* of the tensor $\mathcal{X}$, which are integers such that $0 < J_n \leq I_n, \ \forall n = 1, \ldots, N$. The Tucker decomposition can equivalently be written as follows:

$$\mathcal{X} \approx \mathcal{L} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \cdots \times_N \mathbf{A}^{(N)} \tag{4}$$

e-mail: alexis.huck@fresnel.fr, mireille.guillaume@fresnel.fr

where the operator $\times_n$ is the n-mode product given by:

$$(\mathcal{Q} \times_n \mathbf{P})_{i_1,i_2,\ldots,j_n,\ldots,i_N} = \sum_{i_n} \mathcal{Q}_{i_1,i_2,\ldots,i_n,\ldots,i_N} \mathbf{P}_{j_n,i_n} \quad (5)$$

One can rearrange the tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ into a matrix $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times I_1 \ldots I_{n-1} I_{n+1} \ldots I_N}$. Thus, (4) can equivalently be unfolded as follows:

$$\begin{aligned} \mathbf{X}_{(n)} \approx \mathbf{L}_{(n)} = & \quad (6) \\ &= \mathbf{A}^{(n)} \mathbf{G}_{(n)} \mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(n+1)} \otimes \mathbf{A}^{(n-1)} \otimes \cdots \otimes \mathbf{A}^{(1)} \\ &= \mathbf{A}^{(n)} \mathbf{S}^{(n)} \end{aligned}$$

where $\otimes$ denotes the Kronecker product. If positivity constraints are imposed on the flattening matrices and the core tensor elements, the approximation (3) is called Non-negative Tucker Decomposition (NTD) [5, 6].

1.3. Link between the LSMM model and the NTD decomposition

A hyperspectral image can be assumed as a 3-dimensional signal. Thus a non-negative hyperspectral tensor signal $\mathcal{X} \in \mathcal{R}_+^{(I_1 \times I_2 \times I_3)}$ is approximated as follows:

$$\mathcal{X} \approx \mathcal{L} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)} \ . \quad (7)$$

From (6), one can equivalently unfold the tensor as follows:

$$\mathbf{X}_{(3)} \approx \mathbf{L}_{(3)} = \mathbf{A}^{(3)} \mathbf{G}_{(3)} \mathbf{A}^{(2)} \otimes \mathbf{A}^{(1)} = \mathbf{A}^{(3)} \mathbf{S}^{(3)} \ . \quad (8)$$

One can observe that the LSMM model (2) and the NTD decomposition (8) can be identified, which shows that an NTD algorithm could be used to perform spectral unmixing. Table 1 summarizes the correspondances between the LSMM model and the NTD decomposition. The existing algorithms

LSMM (2)	NTD (8)	Physical meaning
$\mathbf{A}$	$\mathbf{A}^{(3)}$	Columns: endmember spectra
$\mathbf{S}$	$\mathbf{S}^{(3)}$	Rows: abundance maps
-	I_1	Number of rows of the HSI
-	I_2	Number of columns of the HSI
L	I_3	Number of spectral bands
I	$I_1 I_2$	Number of spectral pixels
-	J_1	-
-	J_2	-
J	J_3	Number of endmembers

Table 1. Correspondance between the LSMM model parameters and the NTD parameters, and physical interpretations

that perform NTD satisfy the non-negativity of $\mathbf{A}^{(3)}$ and $\mathbf{S}^{(3)}$. However, they require the *a priori* knowledge of the hyperspectral Tucker ranks J_1, J_2 and J_3 as input parameters, which are unknown for practical purposes. In the following section, we propose a method to estimate them.

2. ESTIMATION OF THE HYPERSPECTRAL TUCKER RANKS

The proposed approach is a generalisation of the HySime algorithm [2] to the Tucker decomposition. The first step consists in estimating the signal tensor $\mathcal{X}$ from the observed tensor $\mathcal{R}$. The second step finds optimal n-rank ($n = 1, 2, 3$) estimates.

2.1. Estimation of the signal tensor $\mathcal{X}$

To estimate the signal tensor, we consider the 3-mode unfolded hyperspectral image: $\mathbf{R} = \mathbf{R}_{(3)}$. The rows of $\mathbf{R}$ hold the spectral bands. As they represent the same scene in thin and contiguous spectral bands, one expects much redundance among the spectral bands. Let us define $\mathbf{R}_{\backslash l} = \left[r_{1,:}^T, r_{2,:}^T, \ldots, r_{l-1,:}^T, r_{l+1,:}^T, \ldots, r_{L,:}^T\right]^T$, *i.e.* the matrix $\mathbf{R}$ but the l^{th} row. Thus, one can expect that the denoised spectral band $\mathbf{X}_{l,:}$ is obtained by a combination of the remaining spectral bands: $\mathbf{X}_{l,:} \approx \beta_l^T \mathbf{R}_{\backslash l}$. The (L-1)-dimensional vector β_l is a regression vector and its least square estimate is given by: $\hat{\beta}_l = \left(\mathbf{R}_{\backslash l} \mathbf{R}_{\backslash l}^T\right)^{-1} \mathbf{R}_{\backslash l} \mathbf{X}_{l,:}^T$.

2.2. Rank estimations

This section explains how to estimate the rank of the n-mode. First, let fix $n = 1$, 2 or 3.

The n-mode unfolded noise tensor is denoted by the matrix $\mathbf{N}_{(n)} = \mathbf{R}_{(n)} - \mathbf{X}_{(n)}$. We assume the columns of $\mathbf{N}_{(n)}$ are random I_n-dimensional centered vectors. The maximum likelihood estimation of their covariance matrix is given by:

$$\hat{\mathbf{K}}_{\mathbf{n}_{(n)}} = \mathbf{N}_{(n)} \mathbf{N}_{(n)}^T / I_n \ . \quad (9)$$

The signal correlation matrix is estimated as follows:

$$\hat{\mathbf{K}}_{\mathbf{x}_{(n)}} = \mathbf{X}_{(n)} \mathbf{X}_{(n)}^T / I_n \ . \quad (10)$$

The eigen decomposition of $\hat{\mathbf{K}}_{\mathbf{x}_{(n)}}$ can be written as:

$$\hat{\mathbf{K}}_{\mathbf{x}_{(n)}} = \mathbf{E}^{(n)} \Sigma^{(n)} \mathbf{E}^{(n)T} \ , \quad (11)$$

where $\mathbf{E}^{(n)}$ contains the eigenvectors ordered by decreasing magnitude of the respective eigenvalues. Thus $\mathcal{R} = \mathcal{E}_k \oplus \mathcal{E}_k^{\perp}$, where $\mathcal{E}_k$ is the subspace of $\mathcal{R}$ spanned by the first k eigenvectors. The purpose is to find the value J_n of k that best imprisons the signal. As the data are unfolded in the n-mode, I_n is the data dimension and J_n is the signal subspace dimension, that is, the n-rank of the data tensor.

$\mathbf{U}_k^{(n)} = \mathbf{E}_k^{(n)} \mathbf{E}_k^{(n)T}$ is the projection matrix onto $\mathcal{E}_k$ and let us define $\hat{\mathbf{x}}_k^{(n)} = \mathbf{U}_k^{(n)} \mathbf{r}_{(n)}$. It has been shown in [2] that

the Mean Squared Error between $\mathbf{x}^{(n)}$ and $\hat{\mathbf{x}}_k^{(n)}$ is given by:

$$\begin{aligned} \mathrm{MSE}^{(n)}(k|\mathbf{x}_{(n)}) = & \quad (12) \\ = E\left[(\mathbf{x}^{(n)} - \hat{\mathbf{x}}_k^{(n)})^T(\mathbf{x}^{(n)} - \hat{\mathbf{x}}_k^{(n)})|\mathbf{x}\right] & \\ = \mathbf{b}_k^{(n)T}\mathbf{b}_k^{(n)} + Tr\left(\mathbf{U}_k^{(n)}\hat{\mathbf{K}}_{\mathbf{n}_{(n)}}\mathbf{U}_k^{(n)T}\right) \ , & \quad (13) \end{aligned}$$

where $\mathbf{b}_k^{(n)} = \mathbf{x}^{(n)} - \hat{\mathbf{x}}_k^{(n)} = \mathbf{U}_k^{(n)\perp}\mathbf{x}^{(n)}$.

The n-rank J_n is the value of k which minimizes the MSE. Note that the first term in (13) represents the signal that is lost with the projection $\mathbf{U}_k^{(n)}$ and decreases when k increases. The second term in (13) can be related to the amount of noise retained in the signal subspace and increases with k. Thus, the best trade-off between the two terms gives the n-rank. As this method was initially proposed to estimate the number of endmembers in the data, it is now generalized to estimate optimal Tucker n-ranks, n = 1,2 and 3 if the hyperspectral data are assumed as tensors.

3. EXPERIMENTS

3.1. Synthetic data sets

We evaluated the method on two synthetic data sets. The first one is a (30 rows $\times$ 50 columns $\times$ 148 spectral bands)-HSI generated from J=3 endmember spectra. The abundance maps have been generated following Dirichlet probability density functions to ensure the additivity constraint (C3). The second considered synthetic HSI has the same dimensions, but the spectral pixels have been generated from 5 endmembers. We impaired the datasets with a gaussian white noise with different Signal to Noise Ratio (SNR) values. The results of the experiments are presented in Table 2. Remarks about this

SNR	20 dB			30 dB			40 dB		
J_3	$\hat{J}_1$	$\hat{J}_2$	$\hat{J}_3$	$\hat{J}_1$	$\hat{J}_2$	$\hat{J}_3$	$\hat{J}_1$	$\hat{J}_2$	$\hat{J}_3$
3	20	26	3	30	45	3	30	50	3
5	22	27	5	30	49	5	30	50	5

Table 2. Experimental results

experiment.

1.One can note that the 3-rank J_3 is always correctly estimated.

2. The optimal values of J_1 and J_2 (the "spatial" ranks) are more dependent of the endmember number and of the SNR. More specifically, the 1-rank and 2-rank increase when the number of endmembers or the SNR increase. In the case of visible and near infra-red wavelength real HSIs, the noise is generally low and the corresponding scene is more complex than a five endmember-based scene. Thus, we can expect the optimal spatial ranks (1 and 2) to be equal to the HSI spatial dimensions.

Fig. 1. True color image of the considered scene

3.2. Real data set

We performed the Tucker rank estimation on a real-world HSI of a desert scene, from the well known HYperspectral Digital Imagery Collection Experiment (HYDICE). Only 167 out of 210 initial spectral bands have been kept, others suffer from low Signal to Noise Ratios (SNR) due to atmosphere absorption. Thus, the considered HSI dimensions are $150 \times 150 \times 167$. The estimated Tucker Ranks are: $J_1 = 150$, $J_2 = 150$ and $J_3 = 20$. As expected: the spatial ranks are equal to the respective dimensions, and the scene contains 20 pure substances.

4. CONCLUSION

In the framework of Hyperspectral Image (HSI) analysis, the studied HSI is often reshaped into a matrix before processing, and it is consequently hard to consider spatial assumptions in the process. In addition, non-negative tensor decomposition methods are of growing interest and can be appropriate tools for HSI analysis. Especially, we show in this paper how the non-negative Tucker decomposition is expected to unmix hyperspectral data. Nevertheless, the Tucker n-ranks of the data are required input parameters in existing Tucker decomposition algorithms. In this paper, we proposed to estimate the Tucker ranks of the hyperspectral data. The method generalizes to tensors an efficient rank estimator dedicated to matricizing-based analysis algorithms. As the Tucker ranks are required for Tucker-based HSI analysis, it is know possible to adapt the non-negative Tucker decomposition algorithms to HSI analysis and this point will be the subject of further works.

5. REFERENCES

[1] C.I. Chang, *Hyperspectral Imaging : techniques for spectral detection and classification*, Kluwer academic/ Plenium publishers, New york, 2003.

[2] J. M. P. Nascimento and J. M. B. Bioucas Dias, "Hyperspectral subspace identification," *IEEE Transactions On Geoscience and Remote Sensing*, vol. 46, 2008.

[3] C. Brie D. Djafari A.M. Moussaoui, S. Carteret, "Bayesian analysis of spectral mixture data using markov chain monte carlo methods," *Chemometrics and Intelligent Laboratory Systems*, vol. 81, pp. 137–148, 2006.

[4] Qiang Zhang, Han Wang, Robert Plemmons, and V. Paul Pauca, "Spectral unmixing using nonnegative tensor factorization," in *ACM-SE 45: Proceedings of the 45th annual southeast regional conference*, New York, NY, USA, 2007, pp. 531–532, ACM.

[5] M. Mørup, "Sparse higher order non-negative matrix factorization," 2006.

[6] Y. D. Kim and S. Choi, "Nonnegative tucker decomposition," in *Workshop on Component Analysis Methods for Classification, Clustering, Modeling and Estimation Problems in Computer Vision*, 2007, pp. 1–8.