Scale disambiguation of blindly unmixed endmember abundances

E. Salerno
Istituto di Scienza e Tecnologie dell'Informazione, CNR,
Area della Ricerca di Pisa, via G. Moruzzi 1, 56124 Pisa, Italy.

15/09/2006

Abstract

The problem of estimating the scale factors related to blindly separated relative abundance maps is addressed. The nature of the quantities looked for is such that their sum must be everywhere equal to one. Exploiting this property, the problem is solved within the framework of a linear noisy mixture model, with signal-independent and zero-mean noise.

1 Introduction

A common problem in blind source separation is that the unmixed signals are normally affected by scale and permutation ambiguities. Whereas permutation is often unimportant, in some applications a quantitative analysis is essential. This is the case, for example, of the analysis of the endmember configuration in remote-sensed hyperspectral images where the relative endmember abundances have been extracted blindly.

MaxNG, a dependent component analysis separation method proposed in [1], has demonstrated its effectiveness in solving the very problem of spectral unmixing from hyperspectral remote-sensed data. In [2], it is also demonstrated that, since the relative abundances must sum up to one everywhere, the problem of scale ambiguity can be solved in terms of the covariance matrix of the unmixed variables.

In this report, I reformulate this problem within a noisy model, assuming that the noise is signal-independent and has a known covariance structure. The result is similar to the one shown in [2], but contains the covariance matrix of the output noise. Moreover, the residual ambiguities highlighted in that paper are solved by relying again on the only source constraint. The analysis of the pure pixels, which was envisaged as the tool to solve these ambiguities, is considered here as an additional resource to make the result more robust against spectral variability and other nonidealities affecting the practical problem.

In Section 2, I complete the result shown in [2] by including the output noise covariance matrix. The effects of sphering and data reduction are also highlighted. In Section 3, the problem of the residual ambiguities is solved by evaluating the mean

source values and a residual global scaling. These results are discussed in Section 4, where a possible use of pure pixels is hypothesized as a check for the validity of the previous results. For conveninece, in Appendix A, I entirely included a short note where the use of pure pixels was discussed with reference to a noiseless model. The notation used is different from the one adopted in the rest of this report, but the issues raised could well be useful for future developments.

2 Exploiting the source constraint within a noisy model

Let us follow the derivation of *MaxNG* from the usual linear instantaneous model with additive independent noise:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t). \tag{1}$$

Before continuing, it is worth specifying the sizes of the arrays that appear in Equation (1). The data vector $\mathbf{x}$ and the noise vector $\mathbf{n}$ have as many components as the number N of the sensor bands. The source vector $\mathbf{s}$ has size P, the number of hypothesized endmembers in the hyperspectral image. Actually, the number of endmembers is very likely to be unknown, so let us assume that P is larger than the number of endmembers and smaller than the number N of frequency bands. In any case, this will be shown to be unessential in our case. Equation (1) holds true for all the values of t between 1 and the number T of pixels in the hyperspectral image. The $N \times P$ mixing matrix $\mathbf{A}$ and the noise process are assumed pixel-independent. The noise covariance matrix will be denoted by $\mathbf{N}$.

The first step in *MaxNG* is data sphering. The data set used in the separation procedure is

$$\tilde{\mathbf{x}} = \Lambda^{-\frac{1}{2}}\mathbf{V}^T(\mathbf{x} - \bar{\mathbf{x}}) = \Lambda^{-\frac{1}{2}}\mathbf{V}^T\mathbf{A}(\mathbf{s} - \bar{\mathbf{s}}) + \Lambda^{-\frac{1}{2}}\mathbf{V}^T\mathbf{n}, \tag{2}$$

where $\bar{\mathbf{x}}$ and $\bar{\mathbf{s}}$ are the mean data and source vectors, respectively. In hyperspectral imaging, the number of frequency bands is always very large (many tens or hundreds of channels). The sphering of the data described by Equation (2) is useful to both obtaining unit-norm centered data and reducing their dimensionality. The guide to this dimension reduction is the data covariance matrix $\mathbf{R}_{\mathbf{xx}}$.[1] We choose a number $M \geq P$ of effective channels. This choice can be guided by an evaluation of the spectrum of matrix $\mathbf{R}_{\mathbf{xx}}$. The size of the sphered data vector $\tilde{\mathbf{x}}$ is M if Λ in Equation (2) is a diagonal matrix containing the M largest eigenvalues of $\mathbf{R}_{\mathbf{xx}}$ and the $N \times M$ matrix $\mathbf{V}$ contains the M related eigenvectors.

MaxNG finds P vectors of size M, $\mathbf{d}_i$, whose application to the sphered data gives an estimate of the centered source vector $(\mathbf{s} - \bar{\mathbf{s}})$.

$$y_i = \mathbf{d}_i^T\tilde{\mathbf{x}} = \mathbf{d}_i^T\Lambda^{-\frac{1}{2}}\mathbf{V}^T\mathbf{A}(\mathbf{s} - \bar{\mathbf{s}}) + \mathbf{d}_i^T\Lambda^{-\frac{1}{2}}\mathbf{V}^T\mathbf{n}. \tag{3}$$

If we assume that the output variables, y_i, are not contaminated by residual interfering sources, we have

$$y_i = h_i(s_i - \bar{s}_i) + \nu_i, \qquad i = 0, \ldots P-1, \tag{4}$$

[1] The noisy ICA approaches, e.g., [3], require a *quasi-sphericization* based on matrix $(\mathbf{R}_{\mathbf{xx}} - \mathbf{N})$. For dependent component analysis, I did not find any theoretical reason to prefer this strategy. The performances of the two options should be assessed experimentally.

where h_i is the specific scale factor for the i-th endmember abundance and, with no loss of generality, no permutation is assumed. The variables ν_i are the components of the output noise vector, whose covariance matrix is

$$\Sigma = \mathbf{D}^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \mathbf{N} \mathbf{V} \Lambda^{-\frac{1}{2}} \mathbf{D}. \tag{5}$$

where $\mathbf{D}$ is a $P \times M$ matrix whose columns are the vectors $\mathbf{d}_i$ appearing in Equation (3).

When a quantitative analysis on the relative abundances has to be performed, an estimation of the scale factors h_i is needed. To this end, we exploit the property for which the relative endmember abundances must sum up to 1 at each pixel. Let us set $q_i = 1/h_i$. Equation (4) can be rearranged as follows:

$$q_i(y_i - \nu_i) = (s_i - \bar{s}_i), \qquad i = 0, \ldots P-1. \tag{6}$$

Thus, the centered source constraint becomes (see [2])

$$\sum_{i=0}^{P-1} q_i(y_i - \nu_i) = 0. \tag{7}$$

Of course, this is a stochastic constraint for the presence of the output noise samples, and cannot be enforced directly. We can try to minimize the expectation of its square instead:

$$\begin{aligned} E\left\{ \left[\sum_{i=0}^{P-1} q_i(y_i - \nu_i) \right]^2 \right\} &= E\left\{ \textstyle\sum_{j=0}^{P-1} \sum_{i=0}^{P-1} q_i(y_i - \nu_i)(y_j - \nu_j) q_j \right\} = \\ &= \textstyle\sum_{j=0}^{P-1} \sum_{i=0}^{P-1} q_i E\left\{ (y_i - \nu_i)(y_j - \nu_j) \right\} q_j = \\ &= \mathbf{q}^T (\mathbf{R}_{\mathbf{yy}} + \Sigma) \mathbf{q}, \end{aligned} \tag{8}$$

where $\mathbf{q}$ is the P-vector with elements q_i, and the last equality holds true if the output noise and the output variables y_i are uncorrelated. It is clear that the minimum value of (8) is reached when $\mathbf{q}$ is the eigenvector of $(\mathbf{R}_{\mathbf{yy}} + \Sigma)$ related to the eigenvalue with minimum magnitude.

3 Estimating the original sources

By finding the minimum-magnitude eigenvalue of $(\mathbf{R}_{\mathbf{yy}} + \Sigma)$ and the related eigenvector $\mathbf{q}$, the scale factors are determined up to a global coefficient α. Thus, an expression similar to (6) holds true:

$$\alpha q_i(y_i - \nu_i) = (s_i - \bar{s}_i), \qquad i = 0, \ldots P-1, \tag{9}$$

where the factors q_i are the known elements of $\mathbf{q}$. To estimate the original endmember abundances, we still need to evaluate α and the P mean source values $\bar{s}_i$.

From Equations (2) and (3), we have

$$\mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \mathbf{x} - \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}} = \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \mathbf{A}\mathbf{s} - \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \mathbf{A}\bar{\mathbf{s}} + \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \mathbf{n}, \tag{10}$$

from which

$$\mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T (\mathbf{x} - \mathbf{A}\mathbf{s} - \mathbf{n}) - \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}} = -\mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \mathbf{A}\bar{\mathbf{s}}, \tag{11}$$

and, by virtue of Equations (1) and (4),

$$\mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}} = h_i \bar{s}_i. \tag{12}$$

Thus

$$\bar{s}_i = q_i \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}}. \tag{13}$$

The mean source values have thus been found up to a common scaling factor. From Equations (9) and (13),

$$\alpha q_i y_i + \alpha q_i \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}} = s_i + \alpha q_i \nu_i. \tag{14}$$

The form on the left-hand side is thus a noisy estimate of the i-th source up to a coefficient α, which is now the only quantity to be estimated in order to solve our problem. Let us enforce the source constraint again. From Equation (14), we have

$$\alpha \sum_{i=0}^{P-1} \left[q_i y_i + q_i \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}} - q_i \nu_i \right] = 1. \tag{15}$$

Considering that the output noise has zero mean, if we average the expression in brackets over all the pixels, we find

$$\alpha = \frac{1}{E\left\{ \sum_{i=0}^{P-1} \left[q_i y_i + q_i \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}} \right] \right\}}, \tag{16}$$

or, if we introduce again the dependence on the pixel index t,

$$\alpha = \frac{T}{\sum_{i=0}^{P-1} q_i \sum_{t=0}^{T-1} y_i(t) + T \sum_{i=0}^{P-1} q_i \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}}} = \frac{1}{\sum_{i=0}^{P-1} q_i \mathbf{d}_i^T \Lambda^{-\frac{1}{2}} \mathbf{V}^T \bar{\mathbf{x}}}, \tag{17}$$

since the means of the centered estimates y_i are obviously zero. This completes the recovery of the original sources.

4 Discussion

Formally, the developments reported in Sections 2 and 3 solve perfectly the problem of scale disambiguation in blind spectral unmixing. However, if the simplifying assumptions made to derive the *MaxNG* strategy are considered, it is immediate to understand that things may not work so ideally. The basic assumption was to have a space-invariant mixing matrix, which is a very rough approximation of reality, especially in remote sensing. This aspect could be faced by analyzing small terrain patches, but in this case the risk is to lose statistical significance. Moreover, even in ideal situations, it is unlikely to get a separation with zero residual interference, thus the validity of the above relations is compromised.

The only constraint used to estimate the scaling, however, has been the sum of the original relative abundances at each pixel. An effort should be made in order to take further properties into account. In particular, we know that all the source values lie in the range $[0,\ 1]$, and this could be exploited to strengten the estimation of the scale factors.

Let us suppose we have a pixel where all but one abundances are zero. This is referred to as a *pure* pixel. Of course, the only nonzero abundance in a pure pixel must be 1. Ignoring the noise term, having $s_i = 1$ means that the multiplying factor αq_i in the left-hand side of (14) is perfectly determined. Assuming space invariant scale factors, and the availability of at least one pure pixel per endmember, we can also reconstruct the abundance values in all pixels from Equation (14).

For the presence of the output noise, however, an estimate based on a small number of pure pixels is not robust. We should be able to identify a large number of pure pixels and then to perform an average of the corresponding results. This and other issues are discussed in Appendix A, and can be used to complement the results obtained through the derivation shown above. In particular, I would like to stress that pure pixels, if correctly identified, can give us valuable local information to assess the spatial variability of the endmember spectra and, in general, can be assumed as *control points* for all the results obtained via global calculations.

References

[1] C. F. Caiafa, A. N. Proto, "Separation of statistically dependent sources using an L^2-distance non-Gaussianity measure", *Signal Processing*, doi:10.1016/j.sigpro.2006.02.032

[2] C. F. Caiafa, E. Salerno, A. N. Proto, L. Fiumi, "Dependent component analysis as a tool for blind spectral unmixing of remote sensed images", *Proc. EUSIPCO 2006*, Florence, Italy, 4-8 September 2006.

[3] A. Hyvärinen, "Fast independent component analysis for noisy data using Gaussian moments", *Proc. Int. Symp. on Circuits and Syst.*, Orlando. Florida, 1999, pp. V57-V61.

A A memo on endmember estimation from scale-ambiguous separated sources (9 January 2006)

Let us assume the usual linear, instantaneous noiseless data model to hold true:

$$\mathbf{x}(x,y) = \mathbf{A}\mathbf{s}(x,y) \tag{18}$$

Let us also assume that some separation method has estimated the sources, as usual, up to scaling and permutation:

$$\hat{\mathbf{s}}(x,y) = \mathbf{W}\mathbf{x}(x,y) = \mathbf{W}\mathbf{A}\mathbf{s}(x,y) = \mathbf{D}\mathbf{P}\mathbf{s}(x,y) \tag{19}$$

where **D** is a diagonal scaling matrix, in general with distinct scale factors for different sources, and **P** is a permutation matrix, unimportant for our purposes. Note that, if the endmembers are considered as pure elements, then assuming a mixing matrix **A** independent of the pixel (x,y) is a very rough approximation. Thus, we should expect that the scale factors affecting the estimated endmember percentages $\hat{\mathbf{s}}(x,y)$ will also depend on location. A looser interpretation of endmembers could refer to image components having similar frequency spectra. Each of them is represented in a column of matrix **A**. Also in this case, however, and even if the linear model is retained, the mixing matrix will be space-variant, due to both intrinsic spectral variability of natural or man-made elements and the errors normally made in deriving the reflectivity values at different frequencies and locations.

Let us consider a single estimated endmember, that is, a generic element of vector $\hat{\mathbf{s}}$. From Equation (19), we get

$$\hat{s}_i(x,y) = d_k(x,y)s_k(x,y) \tag{20}$$

where index k is a generic element of an unknown permutation of indices $\{i\}$, and the dependence on location of the unknown scale factor d_k has been made explicit. Apart from the latter problem, we know that the original endmembers, $s_i(x,y)$, sum up to 1 in any pixel, for the interpretation we gave them, but this does not help us in their estimation when, as normally happens, factors d_k are different for different k's. Indeed, the following equation

$$\sum_i \hat{s}_i = \sum_i d_i s_i \tag{21}$$

makes the permutation ambiguity unimportant, but does not allow us to normalize the $\hat{s}_i$'s in order to make them sum up to 1, unless the scaling factors are known (possibly in any pixel). This is a basic ambiguity of any blind source separation strategy, when no suitable source priors are available, and, as such, is not solvable when no additional information is introduced.

Let us suppose that, at some location $(\bar{x},\bar{y})$, all the original classes but the k-th are zero. This means that the area in the corresponding pixel is covered by a single element (we are in a *pure* pixel), and it will be $s_k(\bar{x},\bar{y}) = 1$. Equation (21) now becomes

$$\sum_i \hat{s}_i(\bar{x},\bar{y}) = d_k(\bar{x},\bar{y})s_k(\bar{x},\bar{y}) \tag{22}$$

Then, if the separation has been performed correctly as in (19), we will also have

$$\hat{s}_k(\bar{x},\bar{y}) = d_k(\bar{x},\bar{y}) \tag{23}$$

That is, we will be able to evaluate the scaling factor affecting source k at location $(\bar{x},\bar{y})$. If the scaling factors do not vary significantly with position, it will be possible to extend endmember normalization to mixed pixels, provided that at least one pure pixel per endmember can be identified in the separated source maps. In this case, indeed, all the scale factors being known, at each pixel we can write

$$\tilde{s}_i(x,y) = \frac{\hat{s}_i(x,y)/d_i}{\sum_k \hat{s}_k(x,y)/d_k} \tag{24}$$

and

$$\sum_i \tilde{s}_i(x,y) = 1 \tag{25}$$

We are, of course, in an ideal case. Endmember spectral variability is very common and often severe, especially in highly complicated environments. Moreover, a perfect separation as in Equation (19) will not be possible, and, in any case, noise will be present. If we are able to identify a large number of pure pixels per class, however, values $d_k(\bar{x},\bar{y})$ evaluated as in (23) could be taken as indices of the spatial variability of the mixing matrix. Provided that these values are not excessively dispersed, Equation (24) can still be applied to average values, possibly dividing the endmember maps into more or less homogeneous regions. Having a sample of scale factors available can also be useful to perform a statistical analysis of some type.

Identifying a pure pixel in an actual case could be not as easy as described for the above ideal situation. We can expect that, for a good and low-noise separation, a pixel where the sum of all but one endmembers does not exceed a fixed small fraction of the remaining endmember is a good candidate to be a pure pixel. Equations (22)-(25) could thus be also useful for such a kind of pixels and, for reduced scaling factor variance, could also be extended to mixed pixels.

Identifying pure pixels in the source maps could also be useful to estimate the endmember spectra. Note that, in hyperspectral imaging, the size of vector $\mathbf{x}(x,y)$ is normally much larger than the number of endmembers. This also means that matrix $\mathbf{A}$ has many more rows (related to spectral components) than columns (related to endmembers), and that the separation is commonly performed after reducing strongly the dimensionality of the data set, for example, by a PCA preprocessing step. Evaluating a separation matrix $\mathbf{W}$, thus, is not tantamount to estimate the endmember spectra, at least not immediately.

However, if we have a list of pure pixels available, that is, if we know that in a certain location $(\bar{x},\bar{y})$ and for a certain endmember i we have $s_i(\bar{x},\bar{y}) = 1$ and $s_k(\bar{x},\bar{y}) = 0$ for any $k \neq i$, from (18), we can write

$$x_f(\bar{x},\bar{y}) = \mathbf{a}_i \cdot \mathbf{s}(\bar{x},\bar{y}) = a_{i,f} s_i(\bar{x},\bar{y}) = a_{i,f} \tag{26}$$

where $\mathbf{a}_i$ is the i-th row of matrix $\mathbf{A}$ and $a_{i,f}$ is the (i,f)-th element of matrix $\mathbf{A}$, that is, the f-th spectral component of the i-th endmember reflectivity. Also in this case, this is only ideally true, but analyzing the results for a suitable collection of pure pixels could give a useful feedback on the actual endmember spectral variability. For well-behaved environments and data, this could also lead to evaluate sorts of "average endmember spectra".