Bilinear matrix factorization methods and application to unsupervised unmixing of urban hyperspectral images

Yannick Deville[a], Fatima Zohra Benhalouche[a,b], Moussa Sofiane Karoui[a,b,c], Abdelaziz Ouamri[b]

Contact: (a) Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, 14 Av. Edouard Belin, 31400 Toulouse, France
yannick.deville@irap.omp.eu

SFPT-GH conference, Grenoble, France, May 11-13, 2016

ANR: HYperspectral imagery for Environmental urban Planning

Introduction

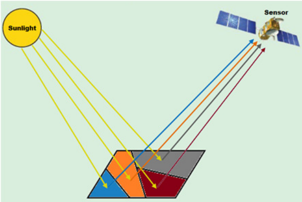
- Blind source separation (BSS)
  ⇒ advanced configurations: nonlinear mixing models
  ⇒ major class: linear-quadratic (LQ), including bilinear
- Bilinear / LQ mixtures:
  theoretical interest + applications:
  - • unmixing of remote sensing data
    • processing of scanned images (show-through effect)
    • analysis of gas sensor array data
  - generic model: truncated polynomial series
    ⇒ approximation of (unknown) model
- Topics of this talk: (1) BMF Methods:
  Bilinear BSS methods based on Matrix Factorization
  + extension to nonnegativity constraints
  (2) unsupervised unmixing of urban hyperspectral images

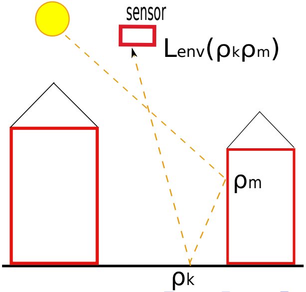
Bilinear mixing model: one sample

- **Scalar form**:

$$x_i(n) = \sum_{j=1}^{M} a_{ij} s_j(n) + \sum_{j=1}^{M-1} \sum_{k=j+1}^{M} b_{ijk} s_j(n) s_k(n) \qquad (1)$$

- Application to **remote sensing [Meganem 2014a]**:
  a_{ij} = **abundance**, $s_j(n)$ = **reflectance** spectrum $\rho_j(\lambda)$,
  single reflection: **linear**, double reflection: **quadratic**





Bilinear mixing model: one sample

- **Scalar form**:

$$x_i(n) = \sum_{j=1}^{M} a_{ij} s_j(n) + \sum_{j=1}^{M-1} \sum_{k=j+1}^{M} b_{ijk} s_j(n) s_k(n) \quad (2)$$

- **First matrix form**:

$$x(n) = As(n) + Bp(n) \quad (3)$$

with column vector $p(n)$: all source products $s_j(n)s_k(n)$

- **Second matrix form**:

$$x(n) = \tilde{A}\tilde{s}(n) \quad (4)$$

with "extended sources" and "extended mixing matrix":

$$\tilde{s}(n) = \begin{bmatrix} s(n) \\ p(n) \end{bmatrix} \text{ and } \tilde{A} = [A \;\; B] \quad (5)$$

Bilinear mixing model: all samples

- **Multi-sample** matrix-form mixing model:

$$x(n) = \tilde{A}\tilde{s}(n) \qquad \Rightarrow \qquad X = \tilde{A}\tilde{S} \tag{6}$$

with

$$\tilde{S} = [\tilde{s}(1), \ldots, \tilde{s}(N)] \tag{7}$$

$$X = [x(1), \ldots, x(N)] \tag{8}$$

BMF methods using a source-constrained structure

Goal of BSS:
provide estimates of source signals,
by using adequately tuned parameters

$\Rightarrow$ **methods**:

- A **standard** approach:
  combine observations according to model which implements class of functions = ***inverse* of class of functions** corresponding to mixing model
  + select parameter values
- Other approach **here**:
  separating system which **models *direct* function**, i.e. mixing function:
  needed for **nonlinear** mixture

... using a source-constrained structure (cont'd)

Mixing and separating data structures:

- **Mixing** function:

$$X = \tilde{A}\tilde{S} \tag{9}$$

  $\Rightarrow$ **variables of separating structure**: matrices C and D, which respectively estimate $\tilde{A}$ and $\tilde{S}$ (possibly up to indeterminacies)

- Rows of $\tilde{S}$ and D: vectors used to **decompose** rows of X
- $\tilde{A}$ and C contain **coefficients** of this decomposition
- **Constraint** on $\tilde{S}$ and therefore D:
  - top M rows of D: **master**, i.e. freely tuned, variables, denoted as d_1 to d_M
  - subsequent rows of D: **slave** variables, updated together with above top rows, so as to contain element-wise **products** $d_j \odot d_k$

… using a source-constrained structure (cont'd)

- Separation principle:
  update C and D so that CD fits X,
  in order to ideally achieve $CD = X$
  $\Rightarrow$ class of methods and separation principle = BMF
- $\Rightarrow$ several adaptation criteria for C and D, e.g:
  1. minimize cost function

$$J_1 = ||X - CD||_F \tag{10}$$

  2. modified approach: see below
- Same as our previous approach for LQ mixtures [Meganem 2014b],
  but here no nonnegativity constraints on sources and mixing coefficients !
  (nor "sum-to-one constraint", thanks to bilinear mixing)

BMF Methods using a doubly-constrained structure

- Matrix C:
  - In above method: **master** variable
  - In following method [Deville 2015]: **slave** variable
    $\Rightarrow$ only master variable: top M rows of D
- New adaptation scheme:
  in each occurence of adaptation **loop for D**,
  slave variable C is set to its optimum value,
  i.e. to its value which **minimizes** $||X - CD||_F$ **wrt** C for considered value of D
  $\Rightarrow$ **least squares** solution:

$$C_{opt} = XD^T(DD^T)^{-1} \tag{11}$$

  $\Rightarrow$ **cost function**:

$$J_2 = ||X(I - D^T(DD^T)^{-1}D)||_F \tag{12}$$

… using a doubly-constrained structure (cont'd)

- Attractive **features**:
  - searched **space** has a much lower dimension
    $\Rightarrow$ computational time, convergence properties
  - J_2 defined by **closed-form** expression
    $\Rightarrow$ gradient-based optimization algorithms

- $\Rightarrow$ various separation **algorithms**:
  - Derivative-free: **Nelder-Mead** method,
    as implemented in fminsearch() Matlab function
    [Deville 2015]
  - **Gradient-based** method $\Rightarrow$ our calculations
    $+$ **nonnegativity** constraint
    [Benhalouche 2016]

Test results

- Pure **spectra**: 8 urban spectra
- **Coefficients**:
  - linear: average classification results over windows
  - quadratic: Fan's model
- Unmixing **methods**:
  - this talk: optimize J_2: 2 versions: (1) gradient, (2) Nelder-Mead, both with nonnegativity constraint
  - linear NMF and extended NMF
  - our previous LQ methods: multiplicative, gradient, gradient-Newton

	SAM (°)	NMSE (%)	SID
Grd-NS-LS-BMF	**2.60**	**15.38**	**1.56**
Nelder-Mead	**2.60**	**15.38**	**1.56**
NMF	99.51	168.04	683.88
Lin-Ext-NMF	16.57	36.87	25.93
Mult-LQNMF	7.85	26.87	4.18
Grd-LQNMF	15.49	43.65	10.85
Grd-New-LQNMF	10.63	29.89	4.32

Separability

- New phenomenon, due to **nonlinearity of mixture**:
  linear independence of sources and some products
  $\Rightarrow$ separability guaranteed for BMF without constraint
  $\neq$ too high indeterminacies in linear BSS
  $\Rightarrow$ e.g. nonnegativity constraint in linear NMF
- Phenomenon due to **separation principle** of BMF:
  select C and D so that CD fits X
- For arbitrary value of top M rows of D:
  row vectors of matrix product CD: combinations of:
  - the M vectors d_1 **to** d_M
  - their products $d_j \odot d_k$

Separability and conditioning

- When each d_j is not collinear to one of the actual source vectors,
  but is a **(bi)linear combination** of the latter vectors: following **property hoped**:
  vector products $d_j \odot d_k$ have "complex form" and are thus outside subspace spanned by actual source vectors and their products,
  i.e. outside subspace spanned by rows of X
  $\Rightarrow$ CD cannot exactly fit X, wathever the value of C
- Conversely, **exact fit $CD = X$ hoped** to be achieved only when D extracts the source signals (up to scaling and permutation)
- Formal **proof** for 2 sources: see [Deville 2015]
- Other issue: **conditioning**

Data

Toy example, related to remote sensing:

- **source** vectors s_1 and s_2:
  reflectance spectra,
  derived from USGS hyperspectral database:
  each source sample is here obtained as average of 200 adjacent samples of an original USGS spectrum
  $\Rightarrow$ source vectors thus reduced to 10 samples
- 10 synthetic but realistic bilinear **mixtures**,
  random, uniform, coefficients: $a_{ij} \in [0, 1]$, $b_{ijk} \in [0, 0.2]$,
  a_{ij} rescaled to sum to one
- 100 Monte-Carlo tests

BSS method, performance criteria

- **Method**: BMF principle, least squares C_{opt}, Nelder-Mead
  no nonnegativity constraint
- **Initialization** of d_1 and d_2:
  s_1 and s_2 + random noise, uniform over $[-0.05, 0.05]$
- Normalized root-mean-square **error for sources**:

$$E_{src} = \frac{\sqrt{\min_{i \neq j \in \{1,2\}} (F_{ij})}}{\sqrt{||s_1||^2 + ||s_2||^2}} \quad (13)$$

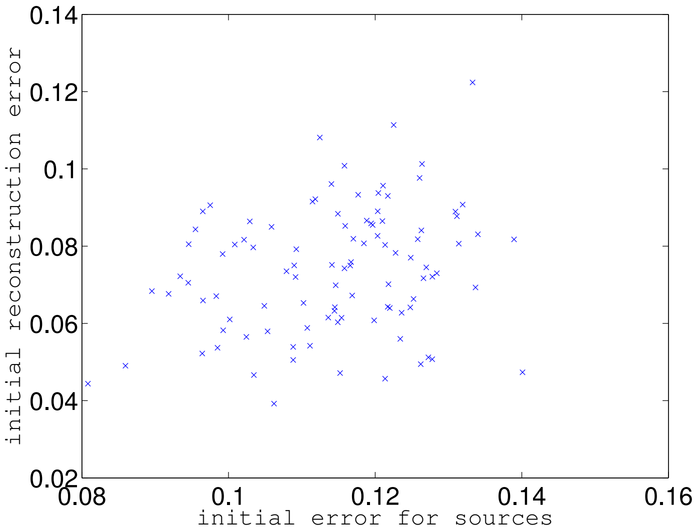
  with F_{ij} equal to:

$$\min_{\epsilon_1 = \pm 1} \left(||s_1 + \epsilon_1 \frac{||s_1||}{||d_i||} d_i||^2 \right) + \min_{\epsilon_2 = \pm 1} \left(||s_2 + \epsilon_2 \frac{||s_2||}{||d_j||} d_j||^2 \right)$$

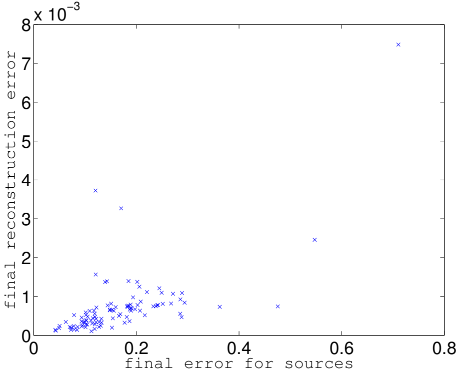
- Normalized **reconstruction error**:

$$E_{recons} = \frac{||X - C_{opt} D||_F}{||X||_F} \quad (14)$$

Scatter plot in (E_{src}, E_{recons}) plane, before BMF



Scatter plot in (E_{src}, E_{recons}) plane, after BMF



- As expected, J_2 strongly decreased
  $\Rightarrow$ **good fit** of CD wrt X
- But **conditioning issue**: source estimates:
  may be significantly different from actual source vectors

Conclusion

- Proposed BMF class of BSS methods is **attractive**:
  1. it initially does not require statistical independence, nonnegativity or sparsity of source signals, but only **linear independence** of sources and some element-wise source products
  2. it does not require knowing analytical form of inverse of mixing model, but only of **direct model**, i.e. mixing model
  3. its separation principle was shown to ensure theoretical **separability** (for 2 sources at this stage)
- Some corresponding practical cost functions and algorithms may lead to **numerical conditioning** issues $\Rightarrow$ avoided with constraints, e.g. **nonnegativity**
- Various extensions: work **in progress**

... questions ?

References

[Meganem 2014a] I. Meganem, P. Déliot, X. Briottet, Y. Deville, S. Hosseini, "Linear-quadratic mixing model for reflectances in urban environments", IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 1, pp. 544-558, Jan. 2014.

[Meganem 2014b] I. Meganem, Y. Deville, S. Hosseini, P. Déliot, X. Briottet, "Linear-quadratic blind source separation Using NMF to unmix urban hyperspectral images", IEEE Transactions on Signal Processing, vol. 62, no. 7, pp. 1822-1833, April 1, 2014.

[Deville 2015] Y. Deville, "Matrix factorization for bilinear blind source separation : methods, separability and conditioning", Proceedings of the 23rd European Signal Processing Conference (EUSIPCO 2015), pp. 1945-1949, Nice, France, Aug. 31 - Sept. 4, 2015.

[Benalouche 2016] F. Z. Benhalouche, Y. Deville, M. S. Karoui, A. Ouamri, "Bilinear matrix factorization using a gradient method for hyperspectral endmember spectra extraction", to appear in the Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2016), Beijing, China, July 10-15, 2016.