Bayesian linear unmixing for spectral mixture analysis
Application to hyperspectral imagery

Nicolas Dobigeon
(Joint work with Prof. J.-Y. Tourneret and other collaborators...)

University of Toulouse
IRIT/INP-ENSEEIHT – TéSA
http://www.enseeiht.fr/~dobigeon

Symposium on Hyperspectral Imaging in Research and Engineering
SHIRE 2014, Sheffield, 10–11 April 2014
Hyperspectral Imagery

*Hyperspectral Images*

- same scene observed at different wavelengths,
Hyperspectral Imagery

Hyperspectral Images

- same scene observed at different wavelengths,

Hyperspectral Cube
Hyperspectral Imagery

*Hyperspectral Images*

- same scene observed at different wavelengths,
- pixel represented by a vector of hundreds of measurements.
Bayesian linear unmixing for spectral mixture analysis

Context

Hyperspectral Imagery

**Hyperspectral Images**

- same scene observed at different wavelengths,
- pixel represented by a vector of hundreds of measurements.

**Hyperspectral Cube**
Spectral Mixture Analysis (SMA)

Linear Mixing Model (LMM): \[ y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \]

Bayesian linear unmixing for spectral mixture analysis

Context

Spectral Mixture Analysis (SMA)

Linear Mixing Model (LMM):

\[ y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \]

- \( L = 825 \) (0.4\( \mu \)m \( \rightarrow \) 2.5\( \mu \)m),
- \( R = 3 \):
  - green grass (solid line),
  - galvanized steel metal (dashed line),
  - bare red brick (dotted line),
- \( a_p = [0.3, 0.6, 0.1]^T \),
- SNR \( \approx 20 \text{dB} \).

Problem

Estimation of \( a_p \) under positivity and additivity constraints and \( m_1, \ldots, m_R \) under positivity constraints.
Bayesian linear unmixing for spectral mixture analysis

Context

Spectral Mixture Analysis (SMA)

Linear Mixing Model (LMM):

\[ y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \]

- \( L = 825 \) \((0.4 \mu m \rightarrow 2.5 \mu m)\),
- \( R = 3 \):
  - green grass (solid line),
  - galvanized steel metal (dashed line),
  - bare red brick (dotted line),
- \( a_p = [0.3, 0.6, 0.1]^T \),
- \( \text{SNR} \approx 20 \text{dB} \).

Problem

Estimation of \( a_p \) under positivity and additivity constraints and \( m_1, \ldots, m_R \) under positivity constraints.
Bayesian linear unmixing for spectral mixture analysis

Context

Spectral Mixture Analysis...

... and applications

\[
\begin{pmatrix}
y_{1,p} \\
\vdots \\
y_{L,p}
\end{pmatrix} =
\begin{pmatrix}
m_1 & \ldots & m_R
\end{pmatrix}
\begin{pmatrix}
a_{1,p} \\
\vdots \\
a_{R,p}
\end{pmatrix} \\
p = 1, \ldots, P
\]

Multi-band imaging

- remote sensing\(^1\), astronomy\(^2\), EELS\(^3\),...
- evolution parameter: pixel index in the image.

Spectrochemical analysis

- Raman\(^4\), NIR\(^5\),...
- evolution parameter: time, temperature, ...

Bioinformatics

- gene expression analysis\(^6\)
- evolution parameter: time, subject, treatment, ...

---

Spectral Mixture Analysis...
... and applications

\[
\begin{pmatrix}
  y_{1,p} \\
  \vdots \\
  y_{L,p}
\end{pmatrix} =
\begin{pmatrix}
  m_1 & \ldots & m_R
\end{pmatrix}
\begin{pmatrix}
  a_{1,p} \\
  \vdots \\
  a_{R,p}
\end{pmatrix}, \quad p = 1, \ldots, P
\]

Multi-band imaging

- remote sensing\(^1\), astronomy\(^2\), EELS\(^3\),...
- evolution parameter: pixel index in the image.

Spectrochemical analysis

- Raman\(^4\), NIR\(^5\),...
- evolution parameter: time, temperature,...

Bioinformatics

- gene expression analysis\(^6\)
- evolution parameter: time, subject, treatment,...

---

Bayesian linear unmixing for spectral mixture analysis

**Context**

Spectral Mixture Analysis...

... and applications

\[
\begin{pmatrix}
  y_{1,p} \\
  \vdots \\
  y_{L,p}
\end{pmatrix}
=
\begin{pmatrix}
  m_1 & \ldots & m_R
\end{pmatrix}
\begin{pmatrix}
  a_{1,p} \\
  \vdots \\
  a_{R,p}
\end{pmatrix}
\]

\( p = 1, \ldots, P \)

**Multi-band imaging**

- remote sensing\(^1\), astronomy\(^2\), EELS\(^3\),...
- evolution parameter: pixel index in the image.

**Spectrochemical analysis**

- Raman\(^4\), NIR\(^5\),...
- evolution parameter: time, temperature,...

**Bioinformatics**

- gene expression analysis\(^6\)
- evolution parameter: time, subject, treatment,...

---

Bayesian linear unmixing for spectral mixture analysis

Context

Spectral Mixture Analysis...
... and applications

\[
\begin{pmatrix}
    y_{1,p} \\
    \vdots \\
    y_{L,p}
\end{pmatrix}
= \begin{pmatrix}
    m_1 & \ldots & m_R
\end{pmatrix}
\begin{pmatrix}
    a_{1,p} \\
    \vdots \\
    a_{R,p}
\end{pmatrix}
\]

\[p = 1, \ldots, P\]

Multi-band imaging

- remote sensing\(^1\), astronomy\(^2\), EELS\(^3\),...
- evolution parameter: pixel index in the image.

Spectrochemical analysis

- Raman\(^4\), NIR\(^5\),...
- evolution parameter: time, temperature, ...

Bioinformatics

- gene expression analysis\(^6\)
- evolution parameter: time, subject, treatment, ...

---

Spectral Mixture Analysis (SMA)

Linear Mixing Model (LMM): \( y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \)

- Supervised case: \( m_1, \ldots, m_R \) are known,
Spectral Mixture Analysis (SMA)

Linear Mixing Model (LMM):

\[ y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \]

- Supervised case: \( m_1, \ldots, m_R \) are known,

- Semi-supervised case: \( m_1, \ldots, m_R \) are partially unknown (\( R \) unknown, the \( m_r \) belong to a fixed spectral library),
Bayesian linear unmixing for spectral mixture analysis

Context

**Spectral Mixture Analysis (SMA)**

Linear Mixing Model (LMM): \( y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \)

- **Supervised case:** \( m_1, \ldots, m_R \) are known,

- **Semi-supervised case:** \( m_1, \ldots, m_R \) are partially unknown (\( R \) unknown, the \( m_r \) belong to a fixed spectral library),

- **Unsupervised case:** \( m_1, \ldots, m_R \) are unknown.
Spectral Mixture Analysis (SMA)

Linear Mixing Model (LMM):

\[ y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p \]

- Supervised case: \( m_1, \ldots, m_R \) are known,

- Semi-supervised case: \( m_1, \ldots, m_R \) are partially unknown (\( R \) unknown, the \( m_r \) belong to a fixed spectral library),

- Unsupervised case: \( m_1, \ldots, m_R \) are unknown.
Outline

Context

Problem formulation

Bayesian modeling
  Likelihood function
  Prior distributions
  Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Outline

Context

Problem formulation

Bayesian modeling
  Likelihood function
  Prior distributions
  Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Spectral unmixing steps

**Preprocessing steps**

Huge data volume to be analyzed

⇒ using data reduction procedures

- Principal component analysis (PCA): projection onto the space spanned by the directions of highest variances,
- Maximum Noise Fraction (MNF): best projection maximizing the SNR,
- ...

Remark: optional steps for some algorithms.

**Estimation steps**

1. **Endmember extraction step** (i.e., identifying the materials),
2. **Inversion step** (abundance estimation),
3. **Joint estimation** of endmembers and abundances.
Endmember extraction algorithms (1)
Exploiting convex geometry principles
Endmember extraction algorithms (1)
Exploiting convex geometry principles

Looking for purest pixels (≈ maximum volume simplex inscribed in the dataset)

(a) Pixel Purity Index (PPI)  
(b) N-FINDR

or projection onto orthogonal subspaces (VCA, ORASIS).

Looking for the minimum volume simplex (inscribing the dataset)
“Minimum Volume Transform” (MVT) algorithm and counterparts.
**Inversion methods (2)**

**Constrained inverse problem**

\[ \text{Minimizing} \quad J(a) = \|y - Ma\|^2 \quad \text{s.t.} \quad \begin{cases} \quad a_r \geq 0, \quad \forall r = 1, \ldots, R \\ \quad \sum_{r=1}^{R} a_r = 1 \end{cases} \]

with \( M = [m_1, \ldots, m_R] \).

- Fully Constrained Least Squares (FCLS) [Heinz et al., 2001],
- Scaled Gradient Methods (SGM) [Theys et al., 2009],
- Interior point primal-dual algorithm [Chouzenoux et al., 2011],
- ...

**Bayesian estimation**

- Choice of a prior distribution ensuring the constraints,
- Difficult statistical estimation → MCMC algorithm.
Joint methods (1)+(2)
Matrix factorization problem

For a given pixel $p$ observed in $L$ spectral bands:

$$y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p$$

$$= Ma_p + n_p$$

Now, consider $P$ pixels:

$$Y = MA + N$$

where

$$Y = [y_1, \ldots, y_P], \quad M = [m_1, \ldots, m_R],$$

$$A = [a_1, \ldots, a_P], \quad N = [n_1, \ldots, n_P].$$

Factorization $Y \approx MA$ formulated as the minimization problem

$$\min_{M,A} D(Y|MA) = \sum_{p} D(y_p|Ma_p) = \sum_{p,\ell} d(y_{\ell,p}|[Ma_p]_{\ell})$$

where $d(a|b)$ is a “distance measure”, e.g., $d(a|b) = \|a - b\|^2$. 
Joint methods (1)+(2)
Matrix factorization problem

For a given pixel \( p \) observed in \( L \) spectral bands:

\[
y_p = \sum_{r=1}^{R} m_r a_{p,r} + n_p = M a_p + n_p
\]

Now, consider \( P \) pixels:

\[
Y = MA + N
\]

where

\[
Y = [y_1, \ldots, y_P], \quad M = [m_1, \ldots, m_R], \\
A = [a_1, \ldots, a_P], \quad N = [n_1, \ldots, n_P].
\]

Factorization \( Y \approx MA \) formulated as the minimization problem

\[
\min_{M,A} D(Y|MA) = \sum_p D(y_p|Ma_p) = \sum_{p,\ell} d(y_{\ell,p}|[Ma_p]_{\ell})
\]

where \( d(a|b) \) is a “distance measure”, e.g., \( d(a|b) = \|a - b\|^2 \).
Bayesian linear unmixing for spectral mixture analysis

Problem formulation

Joint methods (1)+(2)
Matrix factorization problem

Factorization \( Y \approx MA \) is an ill-posed problem!
If \( \{M, A\} \) is a solution, \( \{MP, P^{-1}A\} \) is a solution\(^2\).

\[\Downarrow\]
Additional constraints required!

Factorize \( Y \approx MA \) under positivity and additivity constraints on \( A \)
and positivity constraints on \( M \)

SMA = (constrained) matrix factorization problem
= (constrained) blind source separation problem

\(^2\)For all \( P \) invertible matrix.
Joint methods (1)+(2)
Matrix factorization problem

Factorization $Y \approx MA$ is an ill-posed problem!
If $\{M, A\}$ is a solution, $\{MP, P^{-1}A\}$ is a solution\(^2\).

⇓
Additional constraints required!

Factorize $Y \approx MA$ under positivity and additivity constraints on $A$
and positivity constraints on $M$

SMA = (constrained) matrix factorization problem
= (constrained) blind source separation problem

\(^2\)For all $P$ invertible matrix.
Joint methods (1)+(2)
Matrix factorization problem

Factorization $Y \approx MA$ is an ill-posed problem!
If $\{M, A\}$ is a solution, $\{MP, P^{-1}A\}$ is a solution\(^2\).

$\Downarrow$
Additional constraints required!

Factorize $Y \approx MA$ under positivity and additivity constraints on $A$
and positivity constraints on $M$

$\text{SMA} = (\text{constrained}) \text{ matrix factorization problem}$
$= (\text{constrained}) \text{ blind source separation problem}$

\(^2\)For all $P$ invertible matrix.
Joint methods (1)+(2)
Matrix factorization strategies

1. Principal Component Analysis\(^3\) (PCA) \((Y^T \approx A^T M^T)\)
   - Searching for orthogonal “principal components” (PCs) \(m_r\),
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.

2. Independent Component Analysis (ICA) \((Y^T \approx A^T M^T)\)
   - Maximizing the statistical independence between the sources \(m_r\),
   - Several measures of independence \(\Rightarrow\) several algorithms.
   - Not adapted for hyperspectral images!

3. Nonnegative Matrix Factorization (NMF)
   - Searching for \(M\) et \(A\) with positive entries,
   - Several measures of divergence \(d(\cdot|\cdot)\) \(\Rightarrow\) several algorithms.

4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on \(m_r\) \(\Rightarrow\) positive “sources”
   - Positivity and sum-to-one constraints on \(a_p\)
     \(\Rightarrow\) mixing coefficients = proportions/concentrations/probabilities.

\(^3\)PCA \(\approx\) factor analysis (FA) = empirical orthogonal functions (EOF)
Joint methods (1)+(2)
Matrix factorization strategies

1. Principal Component Analysis\(^3\) (PCA) \((Y^T \approx A^T M^T)\)
   - Searching for orthogonal “principal components” (PCs) \(m_r\),
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.

2. Independent Component Analysis (ICA) \((Y^T \approx A^T M^T)\)
   - Maximizing the statistical independence between the sources \(m_r\),
   - Several measures of independence \(\Rightarrow\) several algorithms.
   - Not adapted for hyperspectral images!

3. Nonnegative Matrix Factorization (NMF)
   - Searching for \(M\) et \(A\) with positive entries,
   - Several measures of divergence \(d(\cdot|\cdot)\) \(\Rightarrow\) several algorithms.

4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on \(m_r\) \(\Rightarrow\) positive “sources”
   - Positivity and sum-to-one constraints on \(a_p\)
   \(\Rightarrow\) mixing coefficients = proportions/concentrations/probabilities.

\(^3\)PCA \(\approx\) factor analysis (FA) = empirical orthogonal functions (EOF)
Joint methods (1)+(2)
Matrix factorization strategies

1. Principal Component Analysis\(^3\) (PCA) \((Y^T \approx A^T M^T)\)
   - Searching for orthogonal “principal components” (PCs) \(m_r\),
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.

2. Independent Component Analysis (ICA) \((Y^T \approx A^T M^T)\)
   - Maximizing the statistical independence between the sources \(m_r\),
   - Several measures of independence ⇒ several algorithms.
   - Not adapted for hyperspectral images!

3. Nonnegative Matrix Factorization (NMF)
   - Searching for \(M\) et \(A\) with positive entries,
   - Several measures of divergence \(d(\cdot|\cdot)\) ⇒ several algorithms.

4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on \(m_r\) ⇒ positive “sources”
   - Positivity and sum-to-one constraints on \(a_p\)
   ⇒ mixing coefficients = proportions/concentrations/probabilities.

\(^3\)PCA ≈ factor analysis (FA) = empirical orthogonal functions (EOF)
Joint methods (1)+(2)
Matrix factorization strategies

1. Principal Component Analysis\(^3\) (PCA) \((Y^T \approx A^T M^T)\)
   - Searching for orthogonal “principal components” (PCs) \(m_r\),
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.

2. Independent Component Analysis (ICA) \((Y^T \approx A^T M^T)\)
   - Maximizing the statistical independence between the sources \(m_r\),
   - Several measures of independence \(\Rightarrow\) several algorithms.
   - Not adapted for hyperspectral images!

3. Nonnegative Matrix Factorization (NMF)
   - Searching for \(M\) et \(A\) with positive entries,
   - Several measures of divergence \(d(\cdot|\cdot) \Rightarrow\) several algorithms.

4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on \(m_r\) \(\Rightarrow\) positive “sources”
   - Positivity and sum-to-one constraints on \(a_p\)
     \(\Rightarrow\) mixing coefficients = proportions/concentrations/probabilities.

\(^3\)PCA \(\approx\) factor analysis (FA) = empirical orthogonal functions (EOF)
Joint method (1)+(2)
Matrix factorization strategies

**NMF-based algorithms**

- Assumption: positivity of the endmember spectra and the abundances,
- Various counterparts, to handle additional constraints (sum-to-one one, minimum volume...)
  Example: MVC-NMF [Miao and Qi, 2007].

**Bayesian estimation**

- Choice of prior ensuring the constraints,
- Difficult statistical estimation $\rightarrow$ MCMC algorithm.
Outline

Context

Problem formulation

Bayesian modeling
  - Likelihood function
  - Prior distributions
  - Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Bayesian inference

Unknown parameters:

- $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_P]$: matrix of the $P$ abundance vectors,
- $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$: matrix of the spectral signatures,
- $\sigma^2$: noise variance,

Unknown parameter vector: $\boldsymbol{\theta} = \{\mathbf{A}, \mathbf{M}, \sigma^2\}$.

Bayes paradigm

Estimation from the posterior distribution $f(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\boldsymbol{\theta})f(\boldsymbol{\theta})$ with:

- Likelihood: $f(\mathbf{Y}|\boldsymbol{\theta})$ (data-fitting term, related to the observation process),
- Parameter prior distribution: $f(\boldsymbol{\theta})$ (prior knowledge on the unknowns).
Computing the Bayesian estimators

**Maximum a posteriori (MAP) estimator**

\[
\hat{\theta}_{\text{MAP}} = \arg \max_{\theta} f(\theta | Y) \\
= \arg \max_{\theta} f(Y | \theta) f(\theta).
\]

**Minimum mean square error (MMSE) estimator**

\[
\hat{\theta}_{\text{MMSE}} = E[\theta | Y] \\
= \int \theta f(\theta | Y) d\theta \\
= \frac{\int \theta f(Y | \theta) f(\theta) d\theta}{\int f(Y | \theta) f(\theta) d\theta}
\]
Bayesian modeling

**Likelihood Function**

The model and the Gaussian property of the noise vector \( n_p \) yield:

\[
f (y_p \mid M, a_p, \sigma^2) = \left( \frac{1}{2\pi\sigma^2} \right)^{ \frac{L}{2} } \exp \left[ - \frac{\| y_p - Ma_p \|^2}{2\sigma^2} \right],
\]

where \( \| \cdot \| \) denotes the standard \( \ell_2 \) norm: \( \| x \|^2 = x^T x \).

Under statistical independence assumption \( n_p \) \( (p = 1, \ldots, P) \):

\[
f (Y \mid M, A, \sigma^2) = \prod_{p=1}^{P} f (y_p \mid M, a_p, \sigma^2),
\]
**Bayesian modeling**

*Endmember prior distribution*

In the observation space:

- For absorbance spectra: exploiting the positivity and the sparsity
  
  ex: spectrochemical analysis [Dobigeon *et al.*, Signal Processing, 2009]
Bayesian modeling

*Endmember prior distribution*

In the observation space:

- For absorbance spectra: exploiting the positivity and the sparsity
  
ex: spectrochemical analysis [Dobigeon *et al*, Signal Processing, 2009]
Endmember prior distribution

In the observation space:

- For absorbance spectra: exploiting the positivity and the sparsity
  ex: spectrochemical analysis [Dobigeon et al, Signal Processing, 2009]

- truncated/half Gaussian distribution $m_{l,r}|\sigma_r^2 \sim \mathcal{N}^+(0, \sigma_r^2)$
Bayesian modeling

Endmember prior distribution

In the observation space:

- For absorbance spectra: exploiting the positivity and the sparsity
  ex: spectrochemical analysis [Dobigeon et al, Signal Processing, 2009]
  - truncated/half Gaussian distribution \( m_{l,r} \mid \sigma_r^2 \sim \mathcal{N}^+ (0, \sigma_r^2) \)
  - exponential distribution \( m_{l,r} \mid \lambda_r \sim \mathcal{E} (\lambda_r) \)
Bayesian modeling

**Endmember prior distribution**

In the observation space:

- For absorbance spectra: exploiting the positivity and the **sparsity**
  
  ex: spectrochemical analysis [Dobigeon *et al*, Signal Processing, 2009]
Endmember prior distribution

In the observation space:

- For absorbance spectra: exploiting the positivity and the sparsity
  ex: spectrochemical analysis [Dobigeon et al, Signal Processing, 2009]

- For reflectance spectra: which feature?
Bayesian linear unmixing for spectral mixture analysis

Bayesian modeling

**Endmember prior distribution**

In the observation space:

- For absorbance spectra: exploiting the positivity and the sparsity
  ex: spectrochemical analysis [Dobigeon *et al*, Signal Processing, 2009]

- For reflectance spectra: which feature?

---

Proposed solution: prior distributions defined on a transformed space!
Bayesian linear unmixing for spectral mixture analysis

Bayesian modeling

Dimensionality reduction

In the noise-free case, the data $\mathbf{Y}$ are represented in a lower-dimensional subset $\mathcal{V}_{R-1}$ of $\mathbb{R}^{R-1}$ without lost of information:

$$\mathcal{V}_{R-1} = \text{span} (\mathbf{v}_1, \ldots, \mathbf{v}_{R-1}).$$
Bayesian linear unmixing for spectral mixture analysis

Bayesian modeling

**Endmember projections**
Projection \( t_r \in \mathbb{R}^{R-1} \) of the endmember spectrum \( m_r \in \mathbb{R}^L \):

\[
t_r = P ( m_r - \bar{y} ) \quad \text{or} \quad m_r = Ut_r + \bar{y},
\]

where \( U = P^\dagger \) and \( \bar{y} \) is the empirical mean of the data.

**Prior for the endmember projections**
Truncated multivariate Gaussian distribution chosen as prior for \( t_r \):

\[
t_r \sim \mathcal{N}_{T_r} ( e_r, s_r^2 I_K ),
\]

where \( T_r \) introduced to ensure positivity of the \( m_{l,r} \) (\( \forall r, r \)):

\[
\{ m_{l,r} \geq 0, \ \forall l = 1, \ldots, L \} \Leftrightarrow \{ t_r \in T_r \}.
\]
Abundance Priors

With the positivity and sum-to-one\(^4\) constraints on \(a_p\), uniform priors chosen for \(a_p\) \((p = 1, \ldots , P)\) on the set \(S\):

\[
S = \{c_p; \|a_p\|_1 = 1 \text{ and } a_p \succeq 0\}.
\]

Variance Prior

Non-informative Jeffreys’ prior:

\[
f(\sigma^2) \propto \frac{1}{\sigma^2}.
\]

\(^4\)Can be relaxed, depending on the target application.
Bayesian linear unmixing for spectral mixture analysis

Bayesian modeling

Posterior distribution of $\theta = \{A, T, \sigma^2\}$

$$f \left( A, T, \sigma^2 \mid Y \right) \propto \prod_{p=1}^{P} 1_{s} (a_p) \times \prod_{r=1}^{R} \exp \left[ -\frac{\|t_r - e_r\|^2}{2s_r^2} \right] 1_{T_r} (t_r) \times \prod_{p=1}^{P} \left[ \left( \frac{1}{\sigma^2} \right)^{\frac{L}{2} + 1} \exp \left( -\frac{\|y_p - (UT + \bar{y}) a_p\|^2}{2\sigma^2} \right) \right].$$

$\rightarrow$ A too complex posterior distribution...

Generation of samples according to $f \left( A, T, \sigma^2 \mid Y \right)$ using MCMC methods.
Outline

Context

Problem formulation

Bayesian modeling
  Likelihood function
  Prior distributions
  Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Gibbs Sampler

Sampling from $f(A|T, \sigma^2, Y)$

$\triangleright$ Sampling from truncated multivariate Gaussian distributions

$$a_p | T, \sigma^2, y_p \sim \mathcal{N}_S(v_p, \Sigma_p).$$

Sampling from $f(T|C, \sigma^2, Y)$

$\triangleright$ Sampling from truncated multivariate Gaussian distributions

$$t_r | T_r, a_r, \sigma^2, Y \sim \mathcal{N}_{T_r}(\tau_r, \Lambda_r).$$

Sampling from $f(\sigma^2|C, T, Y)$

$\triangleright$ Sampling from an inverse-Gamma distribution

$$\sigma^2 | A, T, Y \sim \mathcal{IG} \left( \frac{PL}{2}, \frac{1}{2} \sum_{p=1}^{P} \| y_p - Ma_p \|^2 \right).$$
Approximating the Bayesian estimators

Let $\Theta = \{\theta^{(1)}, \ldots, \theta^{(N_{MC})}\}$ denote a set of $N_{MC}$ samples distributed according to the posterior of interest $f(\theta|Y)$.

**Maximum a posteriori (MAP) estimator**

$$\hat{\theta}_{\text{MAP}} = \arg \max_{\theta} f(\theta|Y)$$

$$\approx \arg \max_{\theta^{(t)} \in \Theta} f(\theta^{(t)}|Y)$$

**Minimum mean square error (MMSE) estimator**

$$\hat{\theta}_{\text{MMSE}} = \mathbb{E}[\theta|Y]$$

$$\approx \frac{1}{N_{MC}} \sum_{t=1}^{N_{MC}} \theta^{(t)}$$
Outline

Context

Problem formulation

Bayesian modeling
  Likelihood function
  Prior distributions
  Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Simulation Results: Synthetic Data

Simulation parameters

- Image: $R = 3$, $L = 413$, and $\text{SNR} = 15\text{dB}$,
Simulation Results: Synthetic Data

Simulation parameters

► Image: $R = 3$, $L = 413$, and $SNR = 15dB$,
Simulation Results: Synthetic Data

Comparisons with standard methods
Posterior distributions of abundances

→ interesting to compute confidence intervals !
Simulation Results: Synthetic Data

Posterior distributions of abundances

→ interesting to compute confidence intervals!
Outline

Context

Problem formulation

Bayesian modeling
  Likelihood function
  Prior distributions
  Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Simulation results: AVIRIS data
[Dobigeon et al., IEEE Trans. SP, 2010]

Simulation parameters

- Image: $50 \times 50$ pixels (Moffett field), $L = 224$ bands.
Simulation results: AVIRIS data
[Dobigeon et al., IEEE Trans. SP, 2010]

Simulation parameters

▶ Image: $50 \times 50$ pixels (Moffett field), $L = 224$ bands.
Simulation results: EELS data
[Dobigeon et al., Ultramicroscopy, 2012]

Simulation parameters

- Image: $64 \times 64$ pixels, $L = 1340$ energy channels.
Simulation results: EELS data

[Dobigeon et al., *Ultramicroscopy*, 2012]

**Simulation parameters**

- Image: $64 \times 64$ pixels, $L = 1340$ energy channels.
Simulation results: EELS data

[Dobigeon et al., Ultramicroscopy, 2012]

Simulation parameters

- Image: $64 \times 64$ pixels, $L = 1340$ energy channels.
Simulation results: Mars data
[Schmidt et al., IEEE Trans. GRS, 2010]

**OMEGA data**

- $L = 184$ spectral bands, $\approx 300 \times 400$ pixels,
- 3 materials: CO$_2$, dust, H$_2$O.
Simulation results: Mars data
[Schmidt et al., IEEE Trans. GRS, 2010]

**OMEGA data**

- \( L = 184 \) spectral bands, \( \approx 300 \times 400 \) pixels,
- 3 materials: CO\(_2\), dust, H\(_2\)O.
Simulation results: spectrometrics
[Dobigeon et al., Signal Processing, 2007]

Chemical mixtures

- $L = 1000$ spectral bands, 10 observed mixtures over time,
- 3 materials: chemical species (reactant A, intermediate C, product D).
Bayesian linear unmixing for spectral mixture analysis
Simulation results on real data

Simulation results: spectrometrics
[Dobigeon et al., Signal Processing, 2007]

Chemical mixtures

- \( L = 1000 \) spectral bands, 10 observed mixtures over time,
- 3 materials: chemical species (reactant A, intermediate C, product D).

(a) Reactant A
(b) Intermediate C
(c) Product D
(d) Mixing coefficients
Outline

Context

Problem formulation

Bayesian modeling
  Likelihood function
  Prior distributions
  Posterior distribution

Markov chain Monte Carlo algorithm

Simulations: synthetic data

Simulation results on real data

Conclusions
Conclusions

Summary

- unsupervised estimation of endmembers and abundances,
- Bayesian model ensuring
  - the positivity and additivity constraints of the abundances,
  - the positivity constraints of the endmember spectra,
- Generation of samples distributed according to the posterior distribution thanks to (hybrid) Gibbs sampler,
- Confidence intervals for the parameter estimates.
Conclusions

Summary

- unsupervised estimation of endmembers and abundances,
- Bayesian model ensuring
  - the positivity and additivity constraints of the abundances,
  - the positivity constraints of the endmember spectra,
- Generation of samples distributed according to the posterior distribution thanks to (hybrid) Gibbs sampler,
- Confidence intervals for the parameter estimates.
Conclusions

**Summary**

- unsupervised estimation of endmembers and abundances,
- Bayesian model ensuring
  - the positivity and additivity constraints of the abundances,
  - the positivity constraints of the endmember spectra,
- Generation of samples distributed according to the posterior distribution thanks to (hybrid) Gibbs sampler,
- Confidence intervals for the parameter estimates.

**Extensions**

- exploiting spatial (or temporal) correlations between neighboring observations,
- estimating the number of components,
- handle nonlinear mixtures (e.g., intimate models, bilinear models,...).
Financial supports
- French Ministry of Defence (MRIS/DGA),
- GdR ISIS network (“Young researcher” project),
- HYPANEMA ANR Project,
- CIMI Lab. of Excellence, Toulouse

Collaborators and/or coauthors
- Jean-Yves Tourneret, IRIT/INP-ENSEEIHT, University of Toulouse,
- Jose M. Bioucas-Dias, Instituto Superior Técnico, Lisbon,
- Steve McLaughlin, Heriot Watt University, Edinburgh,
- Alfred O. Hero, University of Michigan,
- Yoann Altmann, IRIT/INP-ENSEEIHT, University of Toulouse,
- Saïd Moussaoui, IRCCyN, Ecole Centrale de Nantes,
- Nathalie Brun, LPS, University of Paris-Sud,
- Frédéric Schmidt, University of Paris-Sud.
Conclusions

(Subjective) bibliography I


Bayesian linear unmixing for spectral mixture analysis
Application to hyperspectral imagery

Nicolas Dobigeon
(Joint work with Prof. J.-Y. Tourneret and other collaborators...)

University of Toulouse
IRIT/INP-ENSEEIHT – TéSA
http://www.enseeiht.fr/~dobigeon

Symposium on Hyperspectral Imaging in Research and Engineering
SHIRE 2014, Sheffield, 10–11 April 2014