MCMC algorithm for spectral unmixing

Nicolas Dobigeon

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

Atelier Astrostatistique, 8-9 décembre 2011

Hyperspectral Imagery

Hyperspectral Images

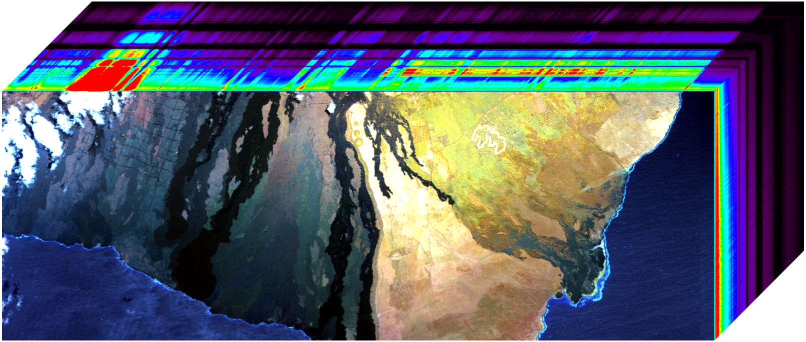
- same scene observed at different wavelengths,

Hyperspectral Imagery

Hyperspectral Images

- same scene observed at different wavelengths,

Hyperspectral Cube



Hyperspectral Imagery

Hyperspectral Images

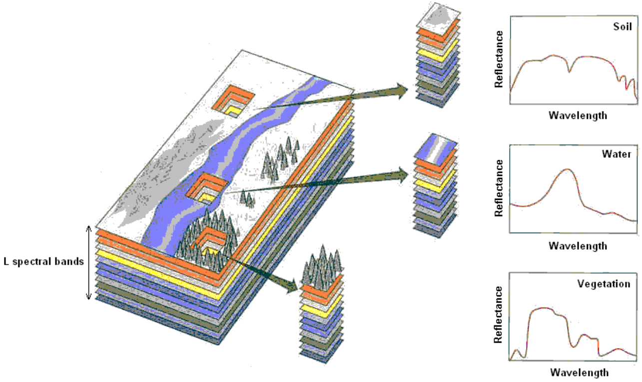
- same scene observed at different wavelengths,
- pixel represented by a vector of hundreds of measurements.

Hyperspectral Imagery

Hyperspectral Images

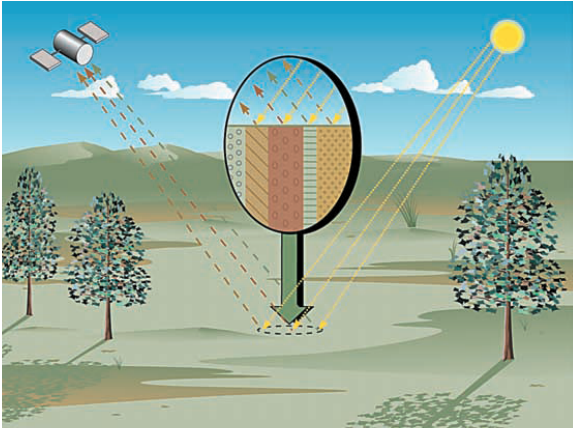
- same scene observed at different wavelengths,
- pixel represented by a vector of hundreds of measurements.

Hyperspectral Cube



Spectral Mixture Analysis

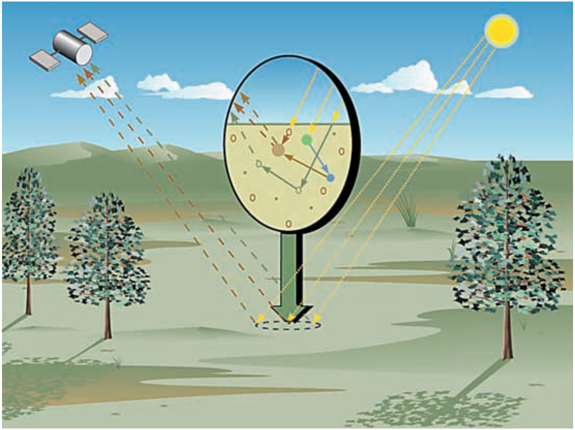
Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$



Reference: IEEE Signal Proc. Magazine, Jan. 2002.

Spectral Mixture Analysis

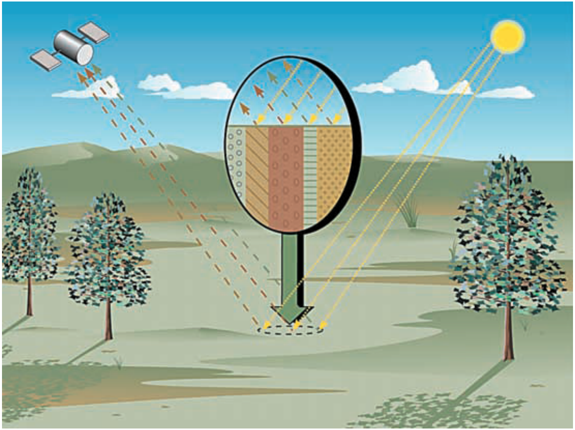
Non-linear Mixing Model: $\mathbf{y}_p = g\left(\{\mathbf{a}_r, \mathbf{m}_r\}_r\right) + \mathbf{n}_p$



Reference: IEEE Signal Proc. Magazine, Jan. 2002.

Spectral Mixture Analysis

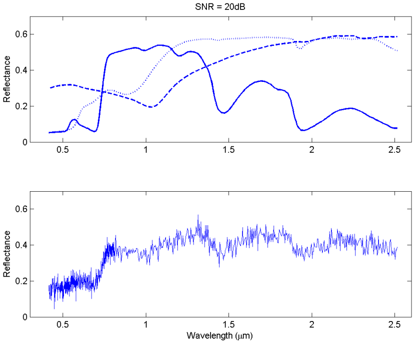
Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$



Reference: IEEE Signal Proc. Magazine, Jan. 2002.

Spectral Mixture Analysis

Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$



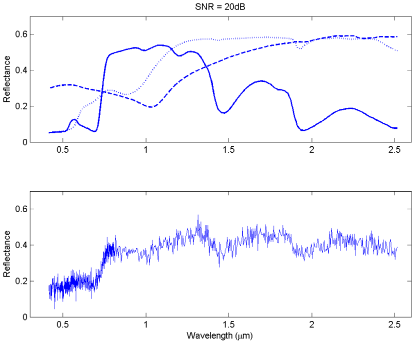
- $L = 825$ ($0.4\mu\text{m} \rightarrow 2.5\mu\text{m}$),
- $R = 3$:
  - green grass (solid line),
  - galvanized steel metal (dashed line),
  - bare red brick (dotted line),
- $\mathbf{a}_p = [0.3, 0.6, 0.1]^T$,
- SNR $\approx$ 20dB.

Problem

Estimation of $\mathbf{a}_p$ under positivity and additivity constraints and $\mathbf{m}_1, \ldots, \mathbf{m}_R$ under positivity constraints.

Spectral Mixture Analysis

Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$



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

Problem

Estimation of $\mathbf{a}_p$ under positivity and additivity constraints and $\mathbf{m}_1, \ldots, \mathbf{m}_R$ under positivity constraints.

Spectral unmixing steps

Preprocessing

Huge data volume $\Rightarrow$ dimensional reduction algorithms

- Principal Component Analysis (PCA): projection into a lower dimensional space spanned by directions of high magnitudes,
- Maximum Noise Fraction (MNF): projection maximizing the SNR,
- ...

Remark: optional step for some algorithms.

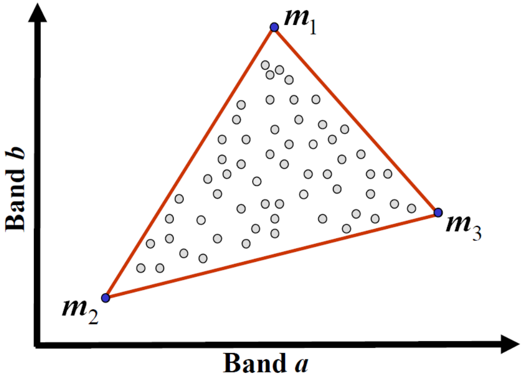
Estimation

(1) **Endmember extraction** step (i.e., estimation of the spectral components),

(2) **inversion** (abundance estimation),

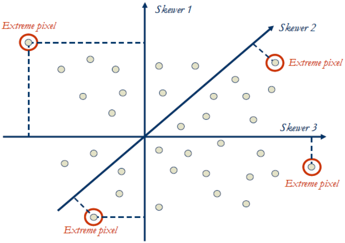
(1+2) **joint** estimation of the endmembers and abundances.

Endmember extraction (1)
Convex geometry

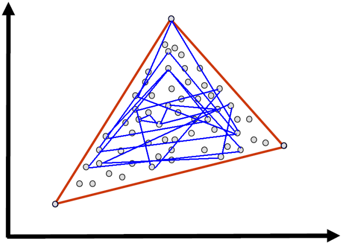


Endmember extraction (1)
Convex geometry

Searching for purest pixels (≈ simplex of maximum volum inscribed in the data)



(a) Pixel Purity Index (PPI)



(b) N-FINDR

or successive projections onto orthogonal subspaces (VCA, ORASIS).

Searching for simplex of minimum volum inscribing the data
"Minimum Volume Transform" (MVT) algorithms and variants.

Inversion (2)
Constrained inverse problem

Constrained optimization

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

with $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$.

- Fully Constrained Least Squares (FCLS) [Heinz *et al.*, 2001],
- Scaled Gradient Methods (SGM) [Theys *et al.*, 2009],
- Primal-dual algorithms [Chouzenoux *et al.*, 2011],
- ...

Spectral Mixture Analysis

Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$

- Supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are known,

Spectral Mixture Analysis

Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$

- Supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are known,
- Semi-supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are partially unknown (R unknown, the $\mathbf{m}_r$ belong to a fixed spectral library),

Spectral Mixture Analysis

Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$

- Supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are known,
- Semi-supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are partially unknown (R unknown, the $\mathbf{m}_r$ belong to a fixed spectral library),
- Unsupervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are unknown.

Spectral Mixture Analysis

Linear Mixing Model (LMM): $\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p$

- Supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are known,
- Semi-supervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are partially unknown (R unknown, the $\mathbf{m}_r$ belong to a fixed spectral library),
- Unsupervised case: $\mathbf{m}_1, \ldots, \mathbf{m}_R$ are unknown.

Outline

Outline

Observation model

For a given pixel p observed in L spectral bands:

$$\begin{aligned} \mathbf{y}_p &= \textstyle\sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p \\ &= \mathbf{M}\mathbf{a}_p + \mathbf{n}_p \end{aligned}$$

Now, consider P pixels:

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \mathbf{N}$$

where

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}_1, \ldots, \mathbf{y}_P], \qquad & \mathbf{M} &= [\mathbf{m}_1, \ldots, \mathbf{m}_R], \\ \mathbf{A} &= [\mathbf{a}_1, \ldots, \mathbf{a}_P], \qquad & \mathbf{N} &= [\mathbf{n}_1, \ldots, \mathbf{n}_P]. \end{aligned}$$

Factorize $\mathbf{Y} \approx \mathbf{M}\mathbf{A}$ under positivity and additivity constraints on $\mathbf{A}$
and positivity constraints on $\mathbf{M}$

Spectral Mixture Analysis = (constrained) matrix factorization
= (constrained) blind source separation

Observation model

For a given pixel p observed in L spectral bands:

$$\begin{aligned} \mathbf{y}_p &= \sum_{r=1}^{R} \mathbf{m}_r a_{p,r} + \mathbf{n}_p \\ &= \mathbf{M}\mathbf{a}_p + \mathbf{n}_p \end{aligned}$$

Now, consider P pixels:

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \mathbf{N}$$

where

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}_1, \ldots, \mathbf{y}_P], & \mathbf{M} &= [\mathbf{m}_1, \ldots, \mathbf{m}_R], \\ \mathbf{A} &= [\mathbf{a}_1, \ldots, \mathbf{a}_P], & \mathbf{N} &= [\mathbf{n}_1, \ldots, \mathbf{n}_P]. \end{aligned}$$

Factorize $\mathbf{Y} \approx \mathbf{M}\mathbf{A}$ under positivity and additivity constraints on $\mathbf{A}$
and positivity constraints on $\mathbf{M}$

Spectral Mixture Analysis = (constrained) matrix factorization
= (constrained) blind source separation

Matrix factorization problem

Factorization $\mathbf{Y} \approx \mathbf{M}\mathbf{A}$ formulated as the minimization problem

$$\min_{\mathbf{M},\mathbf{A}} D(\mathbf{Y}|\mathbf{M}\mathbf{A}) = \sum_{p} D(\mathbf{y}_p|\mathbf{M}\mathbf{a}_p) = \sum_{p,\ell} d\left(y_{\ell,p}|\left[\mathbf{M}\mathbf{a}_p\right]_\ell\right)$$

where $d(a|b)$ is a "distance measure", e.g., $d(a|b) = \|a-b\|^2$.

An ill-posed problem!
If $\{\mathbf{M},\mathbf{A}\}$ is a solution, $\left\{\mathbf{M}\mathbf{P},\mathbf{P}^{-1}\mathbf{A}\right\}$ is a solution[1].

$\Downarrow$

Additional constraints required!

[1] For all $\mathbf{P}$ invertible matrix.

Matrix factorization problem

Factorization $\mathbf{Y} \approx \mathbf{MA}$ formulated as the minimization problem

$$\min_{\mathbf{M},\mathbf{A}} D(\mathbf{Y}|\mathbf{MA}) = \sum_{p} D(\mathbf{y}_p|\mathbf{Ma}_p) = \sum_{p,\ell} d\left(y_{\ell,p}|\left[\mathbf{Ma}_p\right]_{\ell}\right)$$

where $d(a|b)$ is a "distance measure", e.g., $d(a|b) = \|a-b\|^2$.

An ill-posed problem!
If $\{\mathbf{M},\mathbf{A}\}$ is a solution, $\left\{\mathbf{MP},\mathbf{P}^{-1}\mathbf{A}\right\}$ is a solution[1].

$\Downarrow$

Additional constraints required!

[1] For all $\mathbf{P}$ invertible matrix.

Matrix factorization problem

Factorization $\mathbf{Y} \approx \mathbf{MA}$ formulated as the minimization problem

$$\min_{\mathbf{M},\mathbf{A}} D(\mathbf{Y}|\mathbf{MA}) = \sum_{p} D(\mathbf{y}_p|\mathbf{Ma}_p) = \sum_{p,\ell} d\left(y_{\ell,p}|\left[\mathbf{Ma}_p\right]_\ell\right)$$

where $d(a|b)$ is a “distance measure”, e.g., $d(a|b) = \|a-b\|^2$.

An ill-posed problem!
If $\{\mathbf{M}, \mathbf{A}\}$ is a solution, $\left\{\mathbf{MP}, \mathbf{P}^{-1}\mathbf{A}\right\}$ is a solution[1].

$\Downarrow$

Additional constraints required!

[1] For all $\mathbf{P}$ invertible matrix.

Matrix factorization strategies
$\mathbf{Y} \approx \mathbf{M}\mathbf{A} \Leftrightarrow \mathbf{Y}^T \approx \mathbf{A}^T\mathbf{M}^T$

1. Principal Component Analysis(PCA)
   - Searching for orthogonal "principal components" (PCs) $\mathbf{m}_r$,
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.
2. Independent Component Analysis (ICA) (of $\mathbf{Y}^T$)
   - Maximizing the statistical independence between the sources $\mathbf{m}_r$,
   - Several measures of independence $\Rightarrow$ several algorithms.
3. Nonnegative Matrix Factorization (NMF)
   - Searching for $\mathbf{M}$ et $\mathbf{A}$ with positive entries,
   - Several measures of divergence $d(\cdot|\cdot) \Rightarrow$ several algorithms.
4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on $\mathbf{m}_r \Rightarrow$ positive "sources"
   - Positivity and sum-to-one constraints on $\mathbf{a}_p$
     $\Rightarrow$ mixing coefficients = proportions/concentrations/probabilities.

Matrix factorization strategies
$\mathbf{Y} \approx \mathbf{M}\mathbf{A} \Leftrightarrow \mathbf{Y}^T \approx \mathbf{A}^T\mathbf{M}^T$

1. Principal Component Analysis(PCA)
   - Searching for orthogonal "principal components" (PCs) $\mathbf{m}_r$,
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.
2. Independent Component Analysis (ICA) (of $\mathbf{Y}^T$)
   - Maximizing the statistical independence between the sources $\mathbf{m}_r$,
   - Several measures of independence $\Rightarrow$ several algorithms.
3. Nonnegative Matrix Factorization (NMF)
   - Searching for $\mathbf{M}$ et $\mathbf{A}$ with positive entries,
   - Several measures of divergence $d(\cdot|\cdot) \Rightarrow$ several algorithms.
4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on $\mathbf{m}_r \Rightarrow$ positive "sources"
   - Positivity and sum-to-one constraints on $\mathbf{a}_p$
     $\Rightarrow$ mixing coefficients = proportions/concentrations/probabilities.

Matrix factorization strategies
$\mathbf{Y} \approx \mathbf{MA} \Leftrightarrow \mathbf{Y}^T \approx \mathbf{A}^T\mathbf{M}^T$

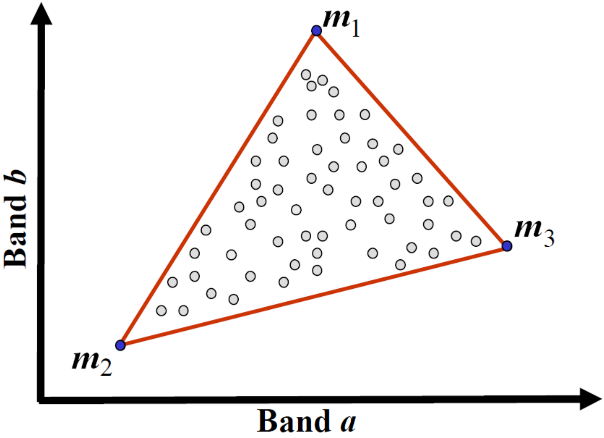
1. Principal Component Analysis(PCA)
   - Searching for orthogonal "principal components" (PCs) $\mathbf{m}_r$,
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.
2. Independent Component Analysis (ICA) (of $\mathbf{Y}^T$)
   - Maximizing the statistical independence between the sources $\mathbf{m}_r$,
   - Several measures of independence $\Rightarrow$ several algorithms.
3. Nonnegative Matrix Factorization (NMF)
   - Searching for $\mathbf{M}$ et $\mathbf{A}$ with positive entries,
   - Several measures of divergence $d(\cdot|\cdot) \Rightarrow$ several algorithms.
4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on $\mathbf{m}_r \Rightarrow$ positive "sources"
   - Positivity and sum-to-one constraints on $\mathbf{a}_p$
     $\Rightarrow$ mixing coefficients = proportions/concentrations/probabilities.

Matrix factorization strategies
$\mathbf{Y} \approx \mathbf{M}\mathbf{A} \Leftrightarrow \mathbf{Y}^T \approx \mathbf{A}^T\mathbf{M}^T$

1. Principal Component Analysis(PCA)
   - Searching for orthogonal “principal components” (PCs) $\mathbf{m}_r$,
   - PCs = directions with maximal variance in the data,
   - Generally used as a dimension reduction procedure.
2. Independent Component Analysis (ICA) (of $\mathbf{Y}^T$)
   - Maximizing the statistical independence between the sources $\mathbf{m}_r$,
   - Several measures of independence $\Rightarrow$ several algorithms.
3. Nonnegative Matrix Factorization (NMF)
   - Searching for $\mathbf{M}$ et $\mathbf{A}$ with positive entries,
   - Several measures of divergence $d\left(\cdot|\cdot\right) \Rightarrow$ several algorithms.
4. (Fully Constrained) Spectral Mixture Analysis (SMA)
   - Positivity constraints on $\mathbf{m}_r \Rightarrow$ positive “sources”
   - Positivity and sum-to-one constraints on $\mathbf{a}_p$
     $\Rightarrow$ mixing coefficients = proportions/concentrations/probabilities.

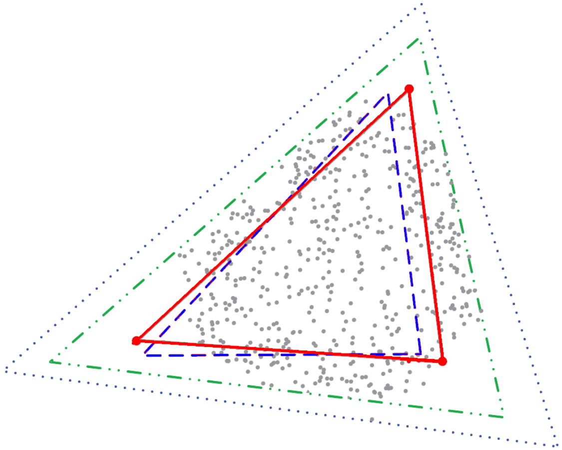
Geometrical formulation of SMA

SMA = looking for a simplex enclosing the data



Geometrical formulation of SMA

In practice: non-unique solution + trade-off noise vs. constraints...



Outline

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,
- σ^2: noise variance,

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

Bayes paradigm: $f(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\boldsymbol{\theta}) f(\boldsymbol{\theta})$ with:

$\longrightarrow$ Likelihood: $f(\mathbf{Y}|\boldsymbol{\theta})$,

$\longrightarrow$ Parameter prior distribution: $f(\boldsymbol{\theta})$.

Hierarchical Bayesian modeling

Likelihood Function

The model and the Gaussian property of the noise vector $\mathbf{n}_p$ yield:

$$f\left(\mathbf{y}_p \middle| \mathbf{M}, \mathbf{a}_p, \sigma^2\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{L}{2}} \exp\left[-\frac{\left\|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p\right\|^2}{2\sigma^2}\right],$$

where $\|\cdot\|$ denotes the standard ℓ_2 norm: $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$.
Assuming prior independence between the $\mathbf{n}_p$'s ($p = 1, \ldots, P$):

$$f\left(\mathbf{Y} \middle| \mathbf{M}, \mathbf{A}, \sigma^2\right) = \prod_{p=1}^{P} f\left(\mathbf{y}_p \middle| \mathbf{M}, \mathbf{a}_p, \sigma^2\right),$$

Hierarchical Bayesian modeling

Endmember prior distribution

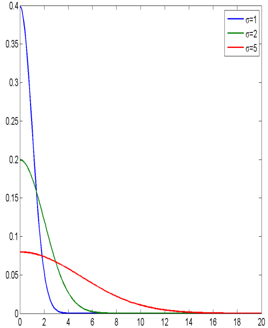
Chosen to ensure positivity constraints on $m_{l,r}$ ($l = 1, \ldots, L$, $r = 1, \ldots, R$).

Hierarchical Bayesian modeling

Endmember prior distribution

Chosen to ensure positivity constraints on $m_{l,r}$ ($l = 1, \ldots, L$, $r = 1, \ldots, R$).
Several choices:

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

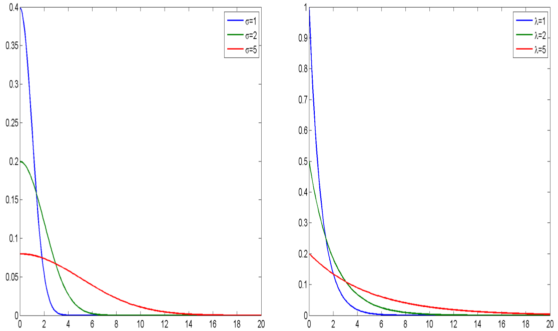


Hierarchical Bayesian modeling

Endmember prior distribution

Chosen to ensure positivity constraints on $m_{l,r}$ ($l = 1, \ldots, L$, $r = 1, \ldots, R$).
Several choices:

- truncated/half Gaussian distribution $m_{l,r}|\sigma_r^2 \sim \mathcal{N}^+ \left(0, \sigma_r^2\right)$
- exponential distribution $m_{l,r}|\lambda_r \sim \mathcal{E}\left(\lambda_r\right)$

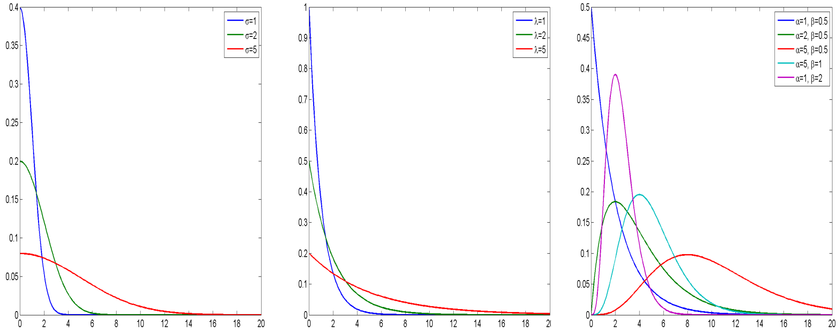


Hierarchical Bayesian modeling

Endmember prior distribution

Chosen to ensure positivity constraints on $m_{l,r}$ ($l = 1, \dots, L$, $r = 1, \dots, R$).
Several choices:

- truncated/half Gaussian distribution $m_{l,r}|\sigma_r^2 \sim \mathcal{N}^+ \left(0, \sigma_r^2\right)$
- exponential distribution $m_{l,r}|\lambda_r \sim \mathcal{E}\left(\lambda_r\right)$
- Gamma distribution $m_{l,r}|\alpha_r, \beta_r \sim \mathcal{G}\left(\alpha_r, \beta_r\right)$



Hierarchical Bayesian modeling

Endmember prior distribution

Gamma distribution with shape parameter α_r and (inverse) scale parameter β_r:

$$f\left(m_{l,r}|\alpha_r,\beta_r\right) = \frac{\beta_r^{\alpha_r}}{\Gamma\left(\alpha_r\right)} m_{l,r}^{\alpha_r-1} \exp\left(-\beta_r m_{l,r}\right) \mathbf{1}_{\mathbb{R}+}\left(m_{l,r}\right)$$

Choice of the hyperparameters? $\rightarrow$ hierarchical Bayesian model...

Endmember hyperparameter prior

Exponential prior for α_r

$$\alpha_r \sim \mathcal{E}(\lambda_{\alpha_r})$$

Conjugate gamma prior for β_r

$$\beta_r \sim \mathcal{G}(\alpha_{\beta_r}, \beta_{\beta_r})$$

where λ_{α_r}, α_{β_r} and β_{β_r} chosen to obtain vague (i.e., flat) hyperpriors.

Hierarchical Bayesian modeling

Endmember prior distribution

Gamma distribution with shape parameter α_r and (inverse) scale parameter β_r:

$$f\left(m_{l,r}|\alpha_r,\beta_r\right) = \frac{\beta_r^{\alpha_r}}{\Gamma\left(\alpha_r\right)} m_{l,r}^{\alpha_r-1} \exp\left(-\beta_r m_{l,r}\right) \mathbf{1}_{\mathbb{R}+}\left(m_{l,r}\right)$$

Choice of the hyperparameters? $\rightarrow$ hierarchical Bayesian model...

Endmember hyperparameter prior

Exponential prior for α_r

$$\alpha_r \sim \mathcal{E}(\lambda_{\alpha_r})$$

Conjugate gamma prior for β_r

$$\beta_r \sim \mathcal{G}(\alpha_{\beta_r}, \beta_{\beta_r})$$

where λ_{α_r}, α_{β_r} and β_{β_r} chosen to obtain vague (i.e., flat) hyperpriors.

Hierarchical Bayesian modeling

Endmember prior distribution

Gamma distribution with shape parameter α_r and (inverse) scale parameter β_r:

$$f\left(m_{l,r}|\alpha_r,\beta_r\right) = \frac{\beta_r^{\alpha_r}}{\Gamma\left(\alpha_r\right)} m_{l,r}^{\alpha_r-1} \exp\left(-\beta_r m_{l,r}\right) \mathbf{1}_{\mathbb{R}+}\left(m_{l,r}\right)$$

Choice of the hyperparameters? $\rightarrow$ hierarchical Bayesian model...

Endmember hyperparameter prior

Exponential prior for α_r

$$\alpha_r \sim \mathcal{E}(\lambda_{\alpha_r})$$

Conjugate gamma prior for β_r

$$\beta_r \sim \mathcal{G}(\alpha_{\beta_r}, \beta_{\beta_r})$$

where λ_{α_r}, α_{β_r} and β_{β_r} chosen to obtain vague (i.e., flat) hyperpriors.

Hierarchical Bayesian modeling

Abundance Priors

With the positivity and additivity constraints on $\mathbf{a}_p$:

$$\mathbf{a}_p = \begin{bmatrix} \mathbf{c}_p \\ a_{p,R} \end{bmatrix} \quad \text{with} \quad \mathbf{c}_p = \begin{bmatrix} a_{p,1} \\ \vdots \\ a_{p,R-1} \end{bmatrix} \quad \text{and} \quad a_{p,R} = 1 - \sum_{r=1}^{R-1} a_{p,r}.$$

Uniform priors chosen for $\mathbf{c}_p$ ($p = 1, \ldots, P$) on the simplex $\mathcal{S}$:

$$\mathcal{S} = \left\{ \mathbf{c}_p; \ \|\mathbf{c}_p\|_1 \leq 1 \text{ and } \mathbf{c}_p \succeq \mathbf{0} \right\}.$$

Variance Prior

Jeffreys' prior:

$$f\left(\sigma^2\right) \sim \frac{1}{\sigma^2}.$$

Hierarchical Bayesian modeling

Posterior distribution of $\boldsymbol{\theta} = \left\{\mathbf{M}, \mathbf{C}, \sigma^2, \boldsymbol{\alpha}\right\}$

$$f\left(\mathbf{M}, \mathbf{C}, \sigma^2, \boldsymbol{\alpha} | \mathbf{Y}\right) \propto f\left(\mathbf{Y} | \mathbf{M}, \mathbf{C}\right) f\left(\mathbf{C}\right) f\left(\mathbf{M} | \boldsymbol{\alpha}, \boldsymbol{\beta}\right) f\left(\boldsymbol{\alpha}\right) f\left(\boldsymbol{\beta}\right) f\left(\sigma^2\right)$$

with $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_R]$ and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_R]$.

$\rightarrow$ A too complex posterior distribution...
Generation of samples according to $f\left(\mathbf{M}, \mathbf{C}, \sigma^2, \boldsymbol{\alpha} | \mathbf{Y}\right)$ using MCMC methods.

Outline

Gibbs Sampler

Sampling from truncated Gaussian distributions $f\left(\mathbf{c}_p \left|\mathbf{M}, \sigma^2, \mathbf{y}_p\right.\right)$

$$\mathbf{c}_p \left|\mathbf{M}, \sigma^2, \mathbf{y}_p \right. \sim \mathcal{N}_{\mathcal{S}}\left(\boldsymbol{\upsilon}_p, \boldsymbol{\Sigma}_p\right)$$

Sampling from $f\left(\alpha_r|\mathbf{C}, \mathbf{M}, \sigma^2, \beta_r, \mathbf{Y}\right)$ using Metropolis-Hastings steps

$$f\left(\alpha_r|\mathbf{C}, \mathbf{M}, \sigma^2, \beta_r, \mathbf{Y}\right) \propto \prod_l^L \left[\frac{\beta_r^{\alpha_r}}{\Gamma\left(\alpha_r\right)} m_{l,r}^{\alpha_r - 1}\right] \exp\left(-\lambda_{\alpha_r}\alpha_r\right) \mathbf{1}_{\mathbb{R}+}\left(\alpha_r\right)$$

Sampling from Gamma distributions $f\left(\beta_r|\mathbf{m}_r, \sigma^2, \alpha_r, \mathbf{Y}\right)$

$$\beta_r|\mathbf{M}, \sigma^2, \boldsymbol{\alpha}, \mathbf{Y} \sim \mathcal{G}\left(1 + L\alpha_r + \alpha_{\beta_r}, \|\mathbf{m}_r\|_1 + \beta_{\beta_r}\right)$$

Sampling from $f\left(m_{l,r}|\mathbf{C}, \sigma^2, \alpha_r, \beta_r, \mathbf{Y}\right)$ using Metropolis-Hastings steps

$$f\left(m_{l,r}|\mathbf{C}, \sigma^2, \alpha_r, \beta_r, \mathbf{Y}\right) \propto m_{l,r}{}^{\alpha_r - 1} \exp\left[-\frac{\left(m_{l,r} - \mu_{l,r}\right)^2}{2\delta_{l,r}^2} - \beta_r m_{l,r}\right] \mathbf{1}_{\mathbb{R}+}\left(m_{l,r}\right)$$

Sampling from an inverse-Gamma distribution $f\left(\sigma^2|\mathbf{C}, \mathbf{T}, \mathbf{Y}\right)$

$$\sigma^2|\mathbf{M}, \mathbf{C}, \mathbf{Y} \sim \mathcal{IG}\left(\frac{PL}{2}, \frac{1}{2}\sum_{p=1}^{P} \|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p\|^2\right)$$

Outline

Simulation Results: Synthetic Data

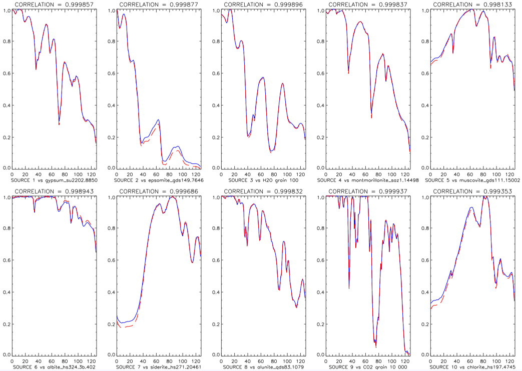
Simulation parameters

- 200×500 pixels,
- $R = 10$ endmembers (H_2O and CO_2 ice spectra + mineral spectra from USGS library),
- $L = 128$ (OMEGA C Channel),
- mixing coefficients drawn uniformly on the admissible set (simplex)

Simulation Results: Synthetic Data

Simulation parameters

- 200×500 pixels,
- $R = 10$ endmembers (H_2O and CO_2 ice spectra + mineral spectra from USGS library),
- $L = 128$ (OMEGA C Channel),
- mixing coefficients drawn uniformly on the admissible set (simplex)



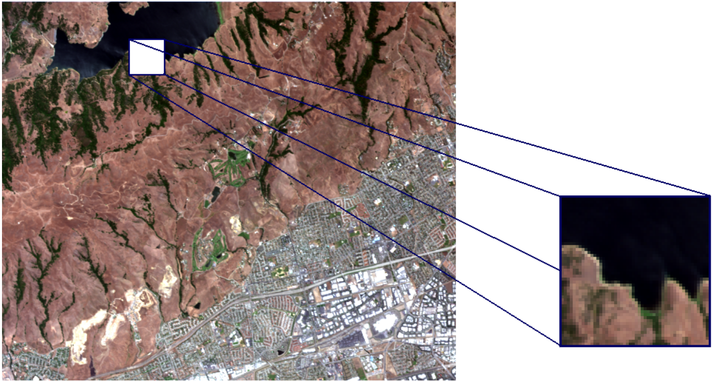
Outline

Simulation results: real data

[Dobigeon *et al*, *IEEE Trans. SP*, 2009]

(remote sensing) AVIRIS data

- Image: 50×50 pixels (Moffett field), $L = 224$ bands,
- 3 materials: vegetation, water, soil.

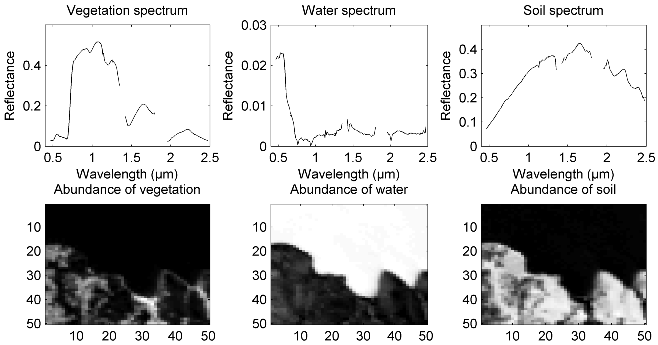


Simulation results: real data

[Dobigeon *et al*, *IEEE Trans. SP*, 2009]

(remote sensing) AVIRIS data

- Image: 50×50 pixels (Moffett field), $L = 224$ bands,
- 3 materials: vegetation, water, soil.



Simulation results: real data

[Schmidt *et al.*, *IEEE Trans. GRS*, 2010]

OMEGA data

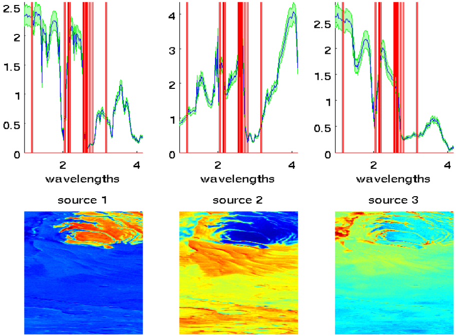
- $L = 184$ spectral bands, $\approx 300 \times 400$ pixels,
- 3 materials: CO_2, dust, H_20.

Simulation results: real 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_20.



Outline

Conclusion

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.

[2] Eches *et al. IEEE Trans. GRS*, 2011.
[3] Dobigeon *et al. IEEE Trans. SP*, 2008.
[4] Iordache *et al.*, *IEEE Trans. GRS*, 2011.
[5] Halimi *et al. IEEE Trans. GRS*, 2011.

Conclusion

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.

Some extensions

- exploiting spatial correlation
  ex: Markovian models[2]
- estimating the number of components (model order selection problem)
  ex: algorithms with jumps[3], "sparse" methods[4]
- nonlinear models (to handle multiple scattering effects or intimate mixtures)
  ex: bilinear models[5]

[2] Eches *et al. IEEE Trans. GRS*, 2011.
[3] Dobigeon *et al. IEEE Trans. SP*, 2008.
[4] Iordache *et al.*, *IEEE Trans. GRS*, 2011.
[5] Halimi *et al. IEEE Trans. GRS*, 2011.

Conclusion

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.

Some extensions

- exploiting spatial correlation
  ex: Markovian models[2]
- estimating the number of components (model order selection problem)
  ex: algorithms with jumps[3], "sparse" methods[4]
- nonlinear models (to handle multiple scattering effects or intimate mixtures)
  ex: bilinear models[5]

[2] Eches *et al. IEEE Trans. GRS*, 2011.
[3] Dobigeon *et al. IEEE Trans. SP*, 2008.
[4] Iordache *et al.*, *IEEE Trans. GRS*, 2011.
[5] Halimi *et al. IEEE Trans. GRS*, 2011.

Acknowledgments

- MRIS/DGA,
- GdR ISIS (Projet Jeunes Chercheurs),
- Alfred O. Hero, University of Michigan,
- Jean-Yves Tourneret, Université de Toulouse,
- Martial Coulon, Université de Toulouse,
- Jérôme Idier, Ecole Centrale de Nantes,
- Eric Le Carpentier, Ecole Centrale de Nantes,
- Saïd Moussaoui, Ecole Centrale de Nantes,
- Frédéric Schmidt, Université Paris 11.

MCMC algorithm for spectral unmixing

Nicolas Dobigeon

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

Atelier Astrostatistique, 8-9 décembre 2011