Nonnegative Matrix and Tensor Factorizations: Models, Algorithms and Applications

Andersen Man Shun Ang
Supervisor: Nicolas Gillis

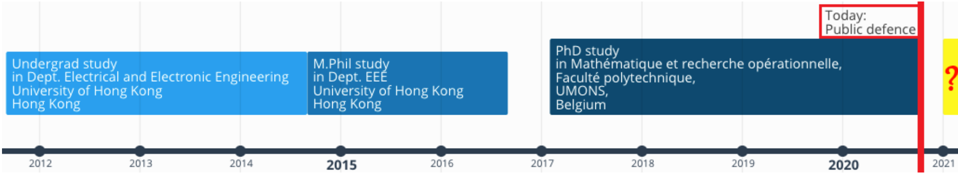
Homepage: angms.science

October 14, 2020

Overview

About myself, and acknowledgments

I spent 9 years learning how to know you don't know.



I acknowledge:

- My boss Nicolas Gillis
- Jury members
  - Xavier Siebert (UMONS), Thierry Dutoit (UMONS), Laurent Jacques (UCL), Francois Glineur (UCL), Lieven De Lathauwer (KUL)
- Colleagues and friends
  Nicolas, Pierre, Hien, Arnaud, Valentin, Maryam, Tim, Francois
  Punit, Jeremy, Junjun, Christophe, Flavia

About this “talk”

- 3 main parts to cover the 8 chapters of the thesis.
  - Ch1 Introduction
  - Ch2 NMF with minimum volume: minvol NMF
  - Ch3 Minvol NMF on hyperspectral unmixing
  - Ch4 Minvol NMF on audio source separation
  - Ch5 NMF with unimodality: NuMF
  - Ch6 Tensor algebra and factorization
  - Ch7 Heuristic extrapolation with restarts
  - Ch8 Conclusion

- Highly compressed.
  $\rightarrow$ details in the thesis

- Only core ideas and eye-catching items.

Overview

For non-NMF people : why NMF ?

- Interpretability

  NMF beats similar tools (PCA, SVD, ICA) due to the interpretability on non-negative data.

- Model correctness

  NMF can find ground truth (under certain conditions).

- Mathematical curiosity

  NMF is related to some serious problems in mathematics.

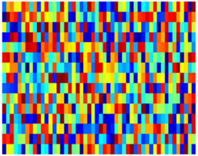
- ~~My boss tell me to do it.~~

12 / 39

(New) It is cool.

Thunder-fast review of NMF

Input



$\mathbf{M} \in \mathbb{R}_{+}^{m \times n}$

and $r \in \mathbb{N}$

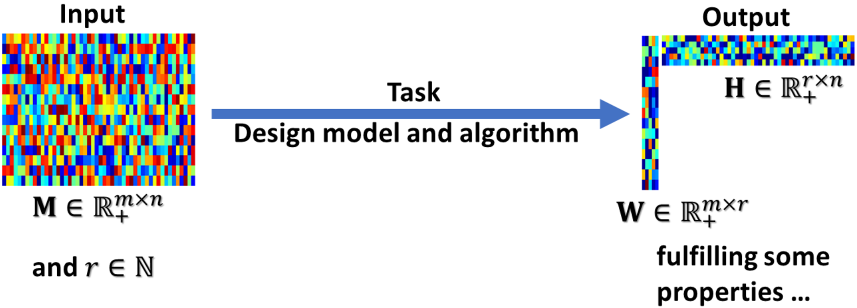
Output

$\mathbf{H} \in \mathbb{R}_{+}^{r \times n}$

$\mathbf{W} \in \mathbb{R}_{+}^{m \times r}$

fulfilling some properties ...

Thunder-fast review of NMF



NMF

- Numerical optimization

$$\underset{\mathbf{W},\mathbf{H}}{\text{minimize}}\ \frac{1}{2}\|\mathbf{M}-\mathbf{W}\mathbf{H}\|_F^2 \ \text{ subject to }\ \mathbf{W}\geq 0 \text{ and } \mathbf{H}\geq 0.$$

NMF

- Numerical optimization

$$\underset{\mathbf{W},\mathbf{H}}{\text{minimize}}\ \frac{1}{2}\|\mathbf{M}-\mathbf{W}\mathbf{H}\|_F^2 \ \text{ subject to } \ \mathbf{W}\geq 0 \text{ and } \mathbf{H}\geq 0.$$

- Regularized model

$$\min_{\mathbf{W},\mathbf{H}} \frac{1}{2}\|\mathbf{M}-\mathbf{W}\mathbf{H}\|_F^2+g(\mathbf{W}) \ \text{ s.t. } \ \mathbf{W}\geq 0, \mathbf{H}\geq 0 \text{ and other constraints.}$$

- Notation: hide $\mathbf{W}\geq 0$, $\mathbf{H}\geq 0$.

The details I skipped

- Popularity of NMF in research
- The nonnegative rank
- How NMF is used in application
- How NMF problems are solved
- HALS
- Solution space of NMF
- NMF is NP-hard
- Separable NMF
- How to solve Separable NMF
- Successive Projection Algorithm

... see chapter 1 of the thesis!

Geometry of $\mathbf{M} = \mathbf{WH}$

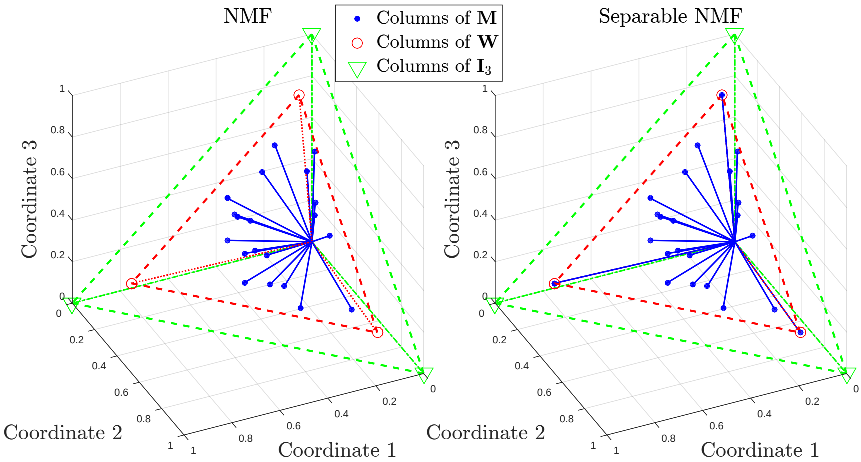


Figure: Nested cone geometry of NMF: blue cone $\subseteq$ red cone $\subseteq$ green cone.

Minimum volume NMF

$$\min_{\mathbf{W},\mathbf{H}} \quad \frac{1}{2}\|\mathbf{M}-\mathbf{W}\mathbf{H}\|_F^2 + \lambda\mathcal{V}(\mathbf{W}) \text{ subject to some constraints.}$$

What is "volume"?

- **Lemma 2.1.1** (p.17 in thesis) If $\mathbf{W} \in \mathbb{R}_+^{m \times r}$ is full rank, then

$$\sqrt{\det(\mathbf{W}^\top \mathbf{W})}$$

  is the volume of $\mathrm{conv}\big([\mathbf{0}\ \mathbf{W}]\big)$ in the column space of $\mathbf{W}$, up to a constant factor.

  Proof idea: Gram-Schmidt orthogonalization and SVD.

- Figure illustration.

What is "volume"?

$$\det(\mathbf{W}^\top \mathbf{W}) = \|\mathbf{w}_1\|_2^2 \cdot \left\|\mathbf{P}_1^\perp \mathbf{w}_2\right\|_2^2 \left\|\mathbf{P}_{1,2}^\perp \mathbf{w}_3\right\|_2^2 \dots \left\|\mathbf{P}_{1,\dots,r-1}^\perp \mathbf{w}_r\right\|_2^2,$$

where $\mathbf{P}_{\mathbf{a}}^\perp = \mathbf{I} - \dfrac{\mathbf{a}\mathbf{a}^\top}{\|\mathbf{a}\|_2^2}$ is the projector on $\text{span}^\perp(\mathbf{a})$,

$$\begin{array}{rclcl}
\mathbf{P}_1^\perp & = & \mathbf{P}_{\mathbf{a}_1}^\perp, & & \mathbf{a}_1 = \mathbf{w}_1 \\
\mathbf{P}_{1,2}^\perp & = & \mathbf{P}_1^\perp \mathbf{P}_{\mathbf{a}_2}^\perp, & & \mathbf{a}_2 = \mathbf{P}_1^\perp \mathbf{w}_2 \\
\mathbf{P}_{1,2,3}^\perp & = & \mathbf{P}_{1,2}^\perp \mathbf{P}_{\mathbf{a}_3}^\perp, & & \mathbf{a}_3 = \mathbf{P}_{1,2}^\perp \mathbf{w}_3 \\
\vdots & & \vdots & & \\
\mathbf{P}_{1,\dots,r-1}^\perp & = & \mathbf{P}_{1,2,\dots,r-2}^\perp \mathbf{P}_{\mathbf{a}_{r-1}}^\perp, & & \mathbf{a}_{r-1} = \mathbf{P}_{1,2,\dots,r-2}^\perp \mathbf{w}_{r-1}.
\end{array}$$

Some volume functions for $\min_{\mathbf{W},\mathbf{H}} \frac{1}{2}\|\mathbf{M} - \mathbf{W}\mathbf{H}\|_F^2 + \lambda \mathcal{V}(\mathbf{W})$

Name	Definition	In $g \circ \boldsymbol{\sigma}(\mathbf{W})$
Determinant (det)	$\mathcal{V}_{\mathrm{det}}(\mathbf{W}) = \det\left(\mathbf{W}^\top \mathbf{W}\right)$	$\prod_{i=1}^{r} \sigma_i^2$
log-determinant (logdet)	$\mathcal{V}_{\mathrm{logdet}}(\mathbf{W}) = \log\det\left(\mathbf{W}^\top \mathbf{W} + \delta \mathbf{I}_r\right)$	$\sum_{i=1}^{r} \log\left(\sigma_i^2 + \delta\right)$
Frobenius norm squared	$\mathcal{V}_{\mathrm{F}}(\mathbf{W}) = \|\mathbf{W}\|_F^2$	$\sum_{i=1}^{r} \sigma_i^2$
Nuclear norm	$\mathcal{V}_*(\mathbf{W}) = \|\mathbf{W}\|_*$	$\sum_{i=1}^{r} \sigma_i$
Smooth Schatten-p norm	$\mathcal{V}_{p,\delta}(\mathbf{W}) = \mathrm{Tr}\left(\mathbf{W}^\top \mathbf{W} + \delta \mathbf{I}_r\right)^{\frac{p}{2}}$	$\sum_{i=1}^{r} (\sigma_i^2 + \delta)^{\frac{p}{2}}$

New theory developed in the thesis: Identifiability of a minvol (N)MF

- **Theorem 2.1.1** Let $\mathbf{M} = \mathbf{W}\mathbf{H}$ where $\mathbf{H} \in \mathbb{R}_+^{r \times n}$ satisfies the SSC, $\mathbf{W} \in \mathbb{R}^{m \times r}$ satisfies $\mathbf{W}^\top \mathbf{1}_m = \mathbf{1}_r$ and $\operatorname{rank}(\mathbf{M}) = r$. Then the (exact) solution of

$$\begin{array}{rll} \underset{\mathbf{W},\mathbf{H}}{\operatorname{argmin}} & \mathcal{V}_{\det}(\mathbf{W}) & \text{\% Det-volume} \\ \text{s.t.} & \mathbf{M} = \mathbf{W}\mathbf{H} & \text{\% Matrix factorization} \\ & \mathbf{H} \geq 0 & \text{\% Nonnegativity} \\ & \mathbf{W}^\top \mathbf{1}_m = \mathbf{1}_r & \text{\% Normalization of col. of } \mathbf{W} \end{array}$$

  is essentially unique.

- Significance: justify the use of minvol NMF in application.

Illustration

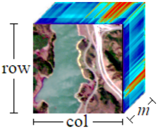
http://angms.science/eg_SNPA_ini.gif

The details I skipped

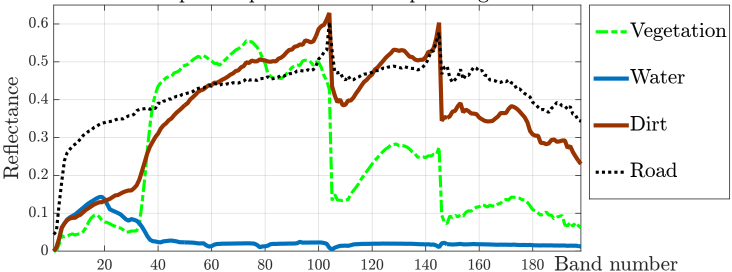
- Details of the formulation of the minvol NMF
- Motivation of minvol NMF
- Details on the volume regularizer
- Comparing the volume regularizer
- Parameter tuning
- Rank deficiency case
- The proof of the identifiability theorem
- How to solve minvol NMF
- Computational cost of the algorithm
- Convergence analysis of the algorithm

… see chapter 2 of the thesis!

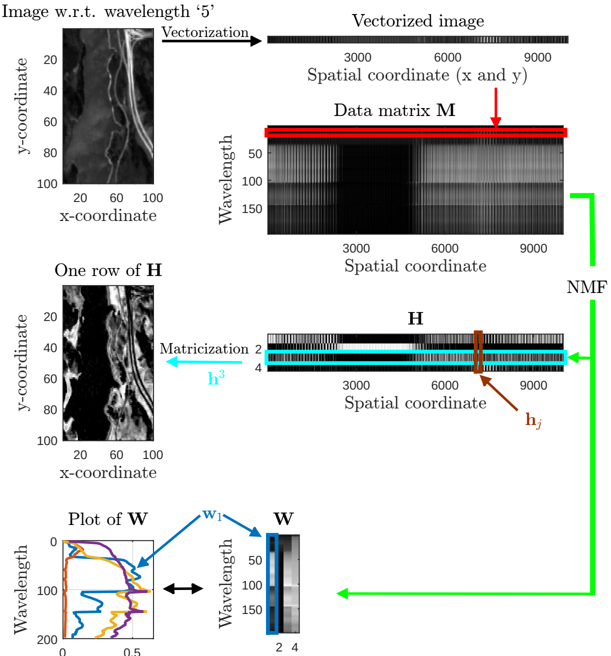
Hyperspectral Unmixing

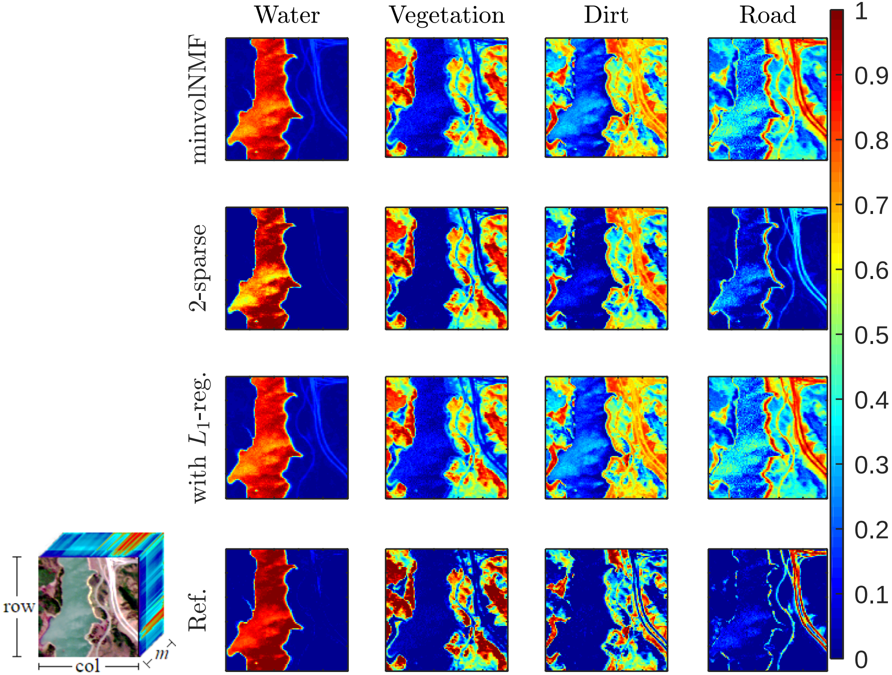


Endmember spectral profiles of the Jasper Ridge HSI data

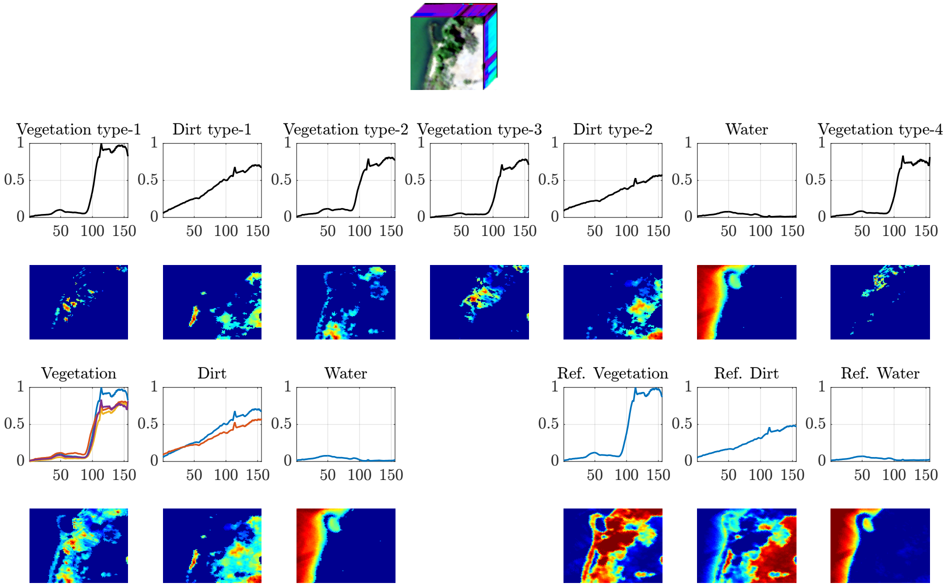


Minvol NMF on Hyperspectral Unmixing





Water
Vegetation
Dirt
Road
minvolNMF
2-sparse
with L_1-reg.
Ref.
row
col
m
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

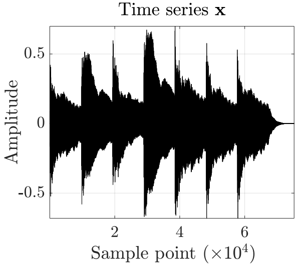

Vegetation type-1
Dirt type-1
Vegetation type-2
Vegetation type-3
Dirt type-2
Water
Vegetation type-4
Vegetation
Dirt
Water
Ref. Vegetation
Ref. Dirt
Ref. Water

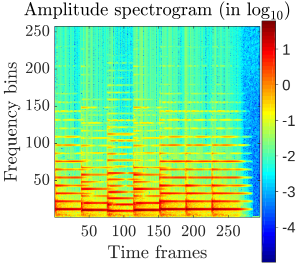
The details I skipped

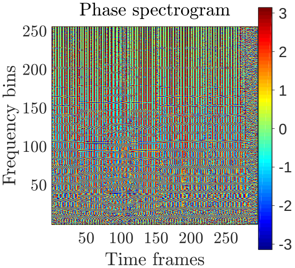
- Details of the hyperspectral image
- Pure-pixel assumption
- Performance metrics for ranking algorithms
- Parameter tuning and experiment setup
- Details of many experimental results
  - minvol NMF $\gg$ `SPA`, `MVC-NMF` and `RVolMin` on synthetic and real datasets.
  - minvol NMF with $\mathcal{V}_{\text{det}}$ and $\mathcal{V}_{\text{logdet}}$ are the top two methods among the tested methods.

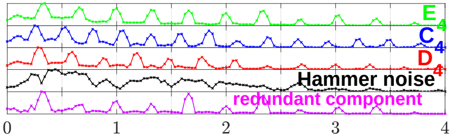
... see chapter 3 of the thesis!

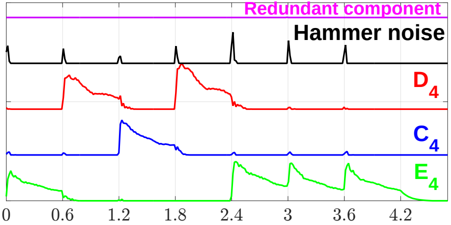
Audio source separation







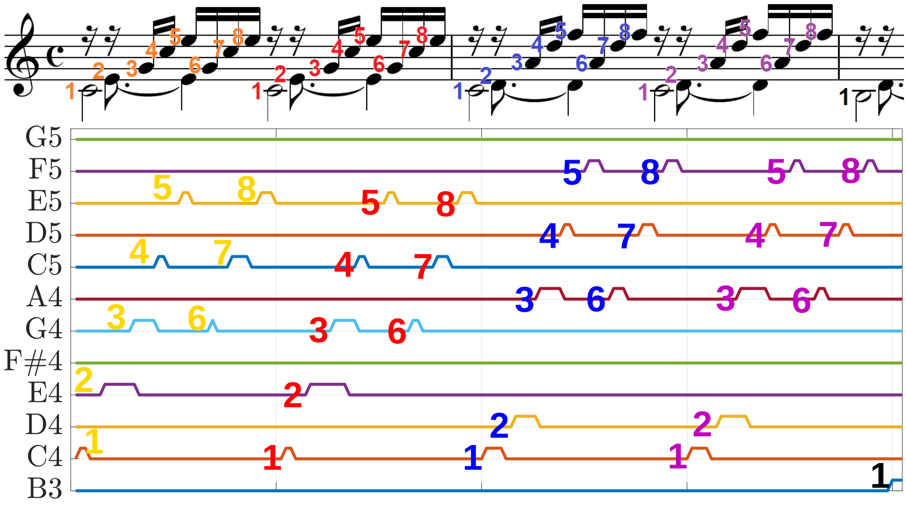




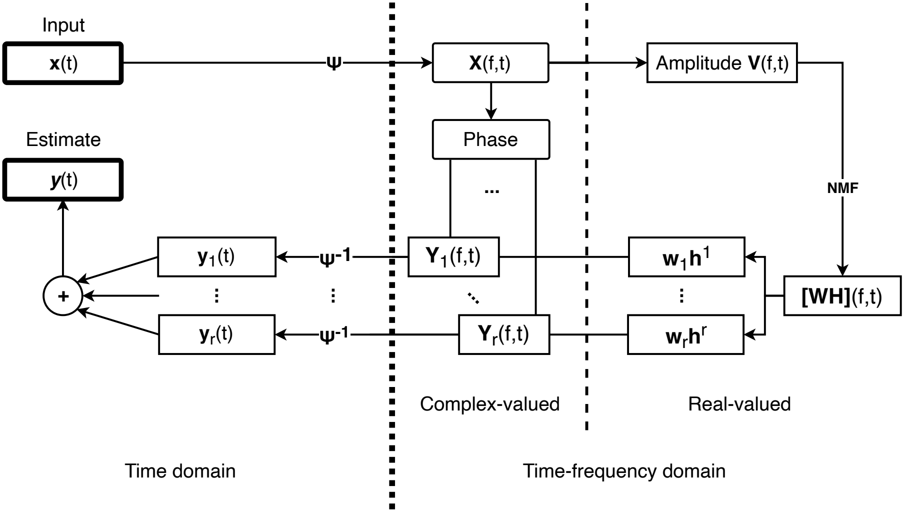
Illustration

https://www.youtube.com/watch?v=1BrpxvpghKQ

Correct identification of the note sequence



How does it work?



And the data fitting term is changed from $\frac{1}{2}\|\mathbf{M} - \mathbf{WH}\|_F^2$ to the β-divergence.

Assumptions on applying minvol NMF on audio data

On using minvol NMF to estimate the source $\mathbf{s}$:

- A1 No time delay in the mixing.
- A2 Linear mixing model.
- A3 Sources are balanced.
- A4 Additive mixture $\mathbf{V} = \sum_i |\mathbf{S}_i|$.
- A5 $|\mathbf{S}_i|$ are well approximated by nonnegative rank-1 matrices.
- A6 Reconstructed components are consistent.
- A7 The data has no outlier.

Violating any one of these assumptions leads to errors, or ... new research opportunities!

On the consistency assumption

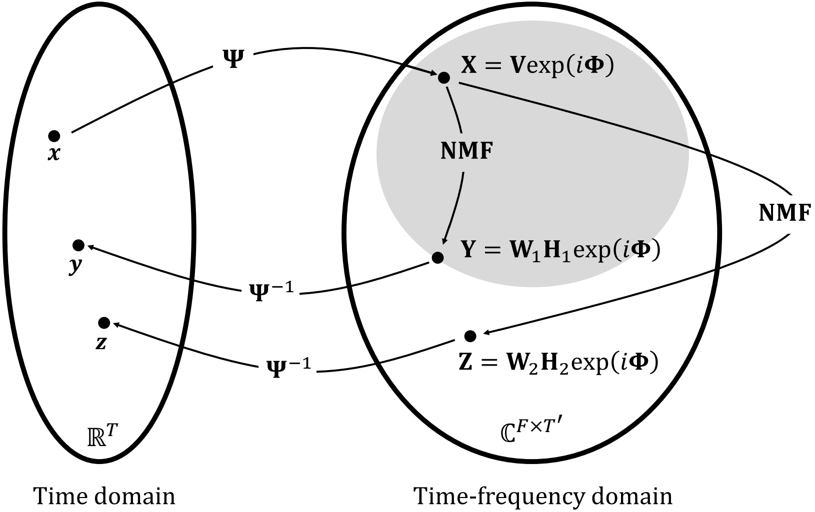


Figure: Illustrating the consistency issue of decomposing a time-frequency matrix. Here the NMF result $\mathbf{W}_1\mathbf{H}_1$ is consistent as the matrix $\mathbf{Y}$ while the result $\mathbf{W}_2\mathbf{H}_2$ is not.

The details I skipped

- Details of the audio source separation process
- Time frequency transform
- β-divergence and β-divergence NMF
- Assumptions on applying NMF on audio data
- Parameter tuning and experiment setup
- Details of experimental results

… see chapter 4 of the thesis!

Overview

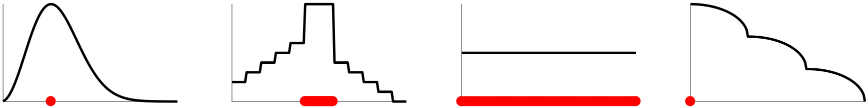
Nonnegative unimodality

- A vector $\mathbf{x}$ is nonnegative unimodal $\exists p \in [m]$ such that

$$0 \leq x_1 \leq x_2 \leq \cdots \leq x_p \text{ and } x_p \geq x_{p+1} \geq \cdots \geq x_m \geq 0.$$

$\mathcal{U}_+^m$: the set of Nu vectors in $\mathbb{R}^m$
$\mathcal{U}_+^{m,p}$: the set of Nu vectors in $\mathcal{U}_+^m$ with tonicity change at p.

- Examples



Characterizing the Nu set

$$\underbrace{\mathbf{x} \in \mathcal{U}_+^m}_{\text{“}\mathbf{x}\text{ is unimodal”}} \iff \exists p \text{ s.t. } \mathbf{x} \in \underbrace{\mathcal{U}_+^{m,p} \cup \mathcal{U}_+^{m,p+1}}_{\text{a convex set}} \iff \underbrace{\left\{ \begin{array}{ccc} 0 & \leq & x_1 \\ x_1 & \leq & x_2 \\ & \vdots & \\ x_{p-1} & \leq & x_p \\ x_{p+1} & \geq & x_{p+2} \\ & \vdots & \\ x_{m-1} & \geq & x_m \\ x_m & \geq & 0 \end{array} \right.}_{\substack{\text{Union of two systems} \\ \text{of monic inequalities.}}}$$

Characterizing the Nu set

$$\left\{\begin{array}{ccc} 0 & \leq & x_1 \\ x_1 & \leq & x_2 \\ & \vdots & \\ x_{p-1} & \leq & x_p \\ x_{p+1} & \geq & x_{p+2} \\ & \vdots & \\ x_{m-1} & \geq & x_m \\ x_m & \geq & 0 \end{array}\right. \iff \mathbf{U}_p \mathbf{x} \geq 0$$

$$\mathbf{U}_p = \left(\begin{array}{c|c} \underbrace{\begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}_{p \times p}}_{\mathbf{D}} & \mathbf{0}_{p \times (m-p)} \\ \hline \mathbf{0}_{(m-p) \times p} & \underbrace{\begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \\ & & & 1 \end{bmatrix}_{(m-p) \times (m-p)}}_{\mathbf{D}} \end{array}\right) \in \{-1, 0, 1\}^{m \times m},$$

NuMF

$$\begin{array}{ll} \min_{\mathbf{W},\mathbf{H}} & \frac{1}{2}\|\mathbf{M}-\mathbf{W}\mathbf{H}\|_F^2 \\ \text{subject to} & \mathbf{U}_{p_j}\mathbf{w}_j \geq 0 \text{ for all } j \in [r], \\ & \mathbf{w}_j^\top \mathbf{1}_m = 1 \text{ for all } j \in [r], \\ & \mathbf{h}^i \geq 0 \text{ for all } i \in [r], \end{array}$$

How to solve it?

- There are r integer unknowns $p_1, \ldots, p_r$
- Idea to solve it: brute force
- Improvement: accelerated projected gradient, peak detection, multi-grid

New theory developed in the thesis: Restriction operator preserves Nu

- **Theorem 5.2.1** Let $\mathbf{x} \in \mathcal{U}_+^m$ and $\mathbf{R} \in \mathbb{R}^{m \times m'}$ with the structure

$$\mathbf{R}(a,b) = \begin{bmatrix} a & b & & & & & \\ & b & a & b & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & b & a & b & \\ & & & & & b & a \end{bmatrix}, a > 0, b > 0, a + 2b = 1,$$

  then $\mathbf{y} = \mathbf{R}\mathbf{x} \in \mathcal{U}_+^{m'}$. Furthermore, if $\mathbf{x} \in \mathcal{U}_+^{m,p}$ where p is even, then $\mathbf{y} \in \mathcal{U}_+^{m,p_y}$ with $p_y \in \left\{ \frac{p}{2} - 1, \frac{p}{2}, \frac{p}{2} + 1 \right\}$.
- Significance: justify the use of Multi-grid as a dimension reduction step in solving NuMF.

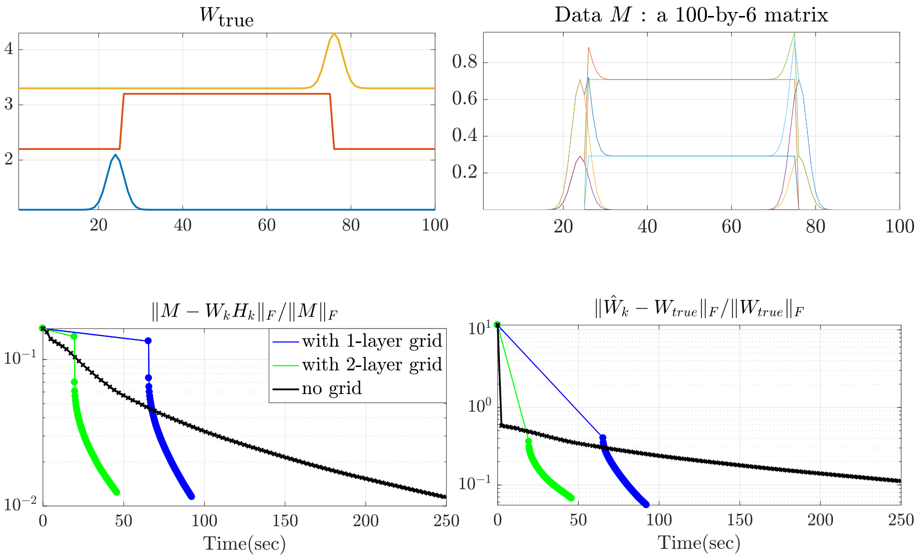
New theory developed in the thesis: 3 theorems related to identifiability of NuMF

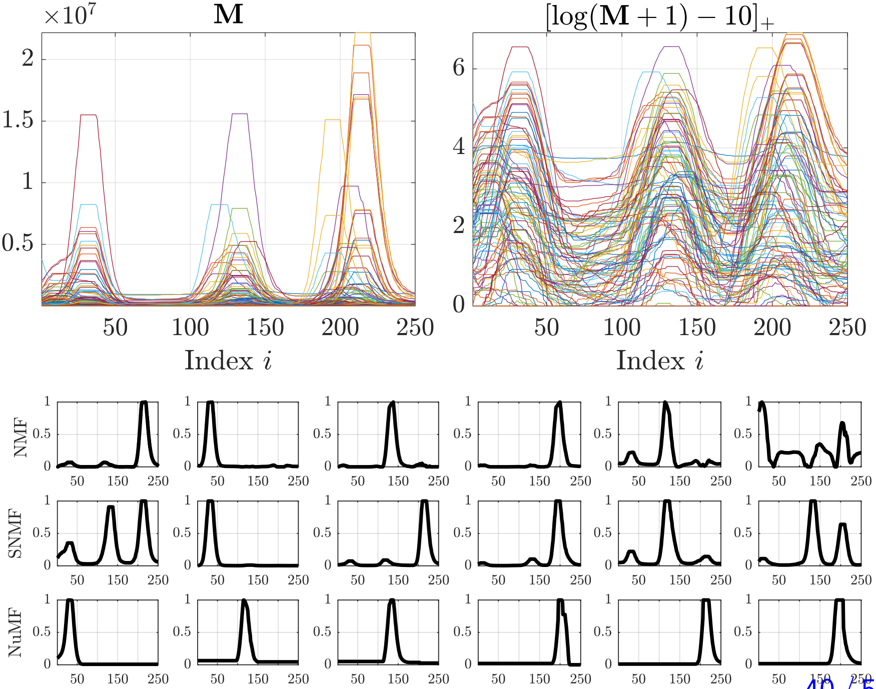
- **Theorem 5.3.1** Informally, if $\mathbf{w}_i$ are Nu and have strictly disjoint support, then the sol. of NuMF is essentially unique.
  - Specialty of the theorem: it works with $n \geq 1$, even if $r > n$.

- **Theorem 5.3.2** Informally, if $\mathbf{w}_i$ are Nu and have ~~strictly disjoint~~ adjacent support, $\mathbf{H}$ satisfies the independent sensing condition, then the (exact) sol. of NuMF is essentially unique.

- **Theorem 5.3.3** Informally, given $\mathbf{x}, \mathbf{y} \in \mathcal{U}_+^m$ s.t. $\mathrm{supp}(\mathbf{x}) \frac{\not\subseteq}{\not\supseteq} \mathrm{supp}(\mathbf{y})$.

  If $\mathbf{x}$, $\mathbf{y}$ can be generated by two non-zero Nu vectors $\mathbf{u}$, $\mathbf{v}$ as $\mathbf{x} = a\mathbf{u} + b\mathbf{v}$ and $\mathbf{y} = c\mathbf{u} + d\mathbf{v}$, then the only possibilities are either $\mathbf{u} = \mathbf{x}$, $\mathbf{v} = \mathbf{y}$ or $\mathbf{u} = \mathbf{y}$, $\mathbf{v} = \mathbf{x}$.


W_{true}
Data M : a 100-by-6 matrix
$\|M - W_k H_k\|_F / \|M\|_F$
$\|\hat{W}_k - W_{true}\|_F / \|W_{true}\|_F$
with 1-layer grid
with 2-layer grid
no grid
Time(sec)
Time(sec)



$\times 10^7$
$\mathbf{M}$
$[\log(\mathbf{M}+1)-10]_+$
Index i
Index i
NMF
SNMF
NuMF

The details I skipped

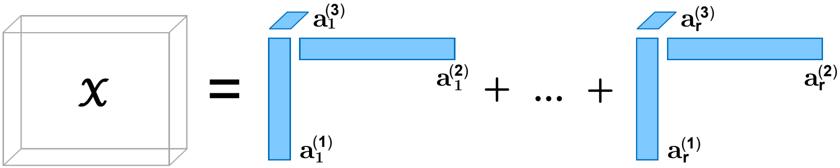
- Details of the accelerated projected gradient
- Details of the peak detection
- Details of multi-grid
- Details of the identifiability results

... see chapter 5 of the thesis!

Overview

HER

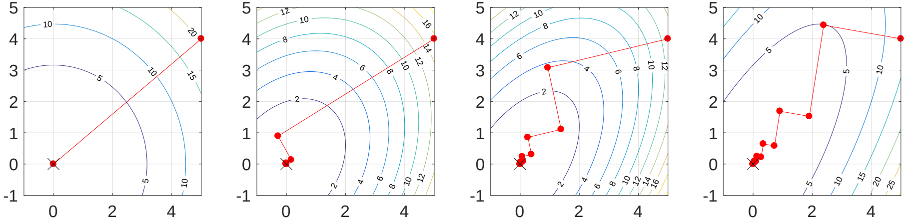
- A framework for accelerating algo. for solving NMF, NTF and CPD.

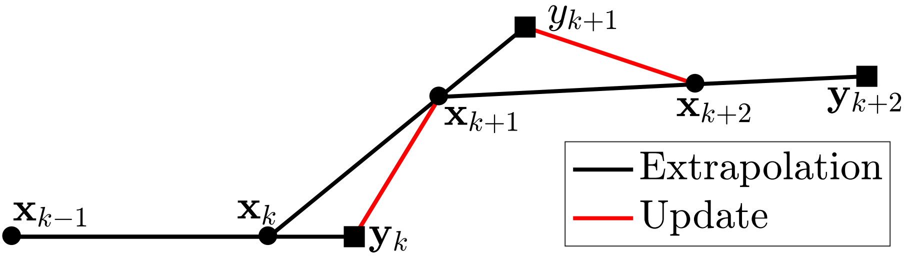


- Key ideas
  - Extrapolation with restart.
  - Cheap computation by making use of computed component in the update.

- Many numerical evidences on the effectiveness of HER.

Extrapolation? Why?





Algorithm 8 HER

Input: a nonnegative N-way tensor

Output: nonnegative factors $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)}$.

Initialization: Choose $\beta_0 \in (0,1)$, $\eta \geq \bar{\gamma} \geq \gamma \geq 1$ and 2 sets of initial factor matrices $\left(\mathbf{A}_0^{(1)}, \ldots, \mathbf{A}_0^{(N)}\right)$ and $\left(\hat{\mathbf{A}}_0^{(1)}, \ldots, \hat{\mathbf{A}}_0^{(N)}\right)$. Set $\bar{\beta}_0 = 1$ and $k = 1$.

for $k = 1, \ldots$ until some criteria is satisfied **do**

  for $i = 1, \ldots, N$ **do**

    Update step Let $\mathbf{A}_k^{(i)}$ be an exact/inexact solution of

$$\min_{\mathbf{A}^{(i)} \geq 0} F\left(\hat{\mathbf{A}}_k^{(1)}, \ldots, \hat{\mathbf{A}}_k^{(i-1)}, \mathbf{A}^{(i)}, \hat{\mathbf{A}}_{k-1}^{(i+1)}, \ldots, \hat{\mathbf{A}}_{k-1}^{(N)}\right). \tag{7.1}$$

    Extrapolation step

$$\hat{\mathbf{A}}_k^{(i)} = \left[\mathbf{A}_k^{(i)} + \beta_{k-1}(\mathbf{A}_k^{(i)} - \mathbf{A}_{k-1}^{(i)})\right]_+. \tag{7.2}$$

  end for

  Compute $\hat{F}_k := F\left(\hat{\mathbf{A}}_k^{(1)}, \hat{\mathbf{A}}_k^{(2)}, \ldots, \hat{\mathbf{A}}_k^{(N-1)}, \mathbf{A}_k^{(N)}\right)$.

  if $\hat{F}_k > \hat{F}_{k-1}$ **then**

    Set $\hat{\mathbf{A}}_k^{(i)} = \mathbf{A}_k^{(i)}$, $\forall i \in [N]$ *% abandon the sequence* $\hat{\mathbf{A}}_k^{(i)}$

    Set $\bar{\beta}_k = \beta_{k-1}$, $\beta_k = \beta_{k-1}/\eta$. *% Update* $\bar{\beta}$*, decrease* β

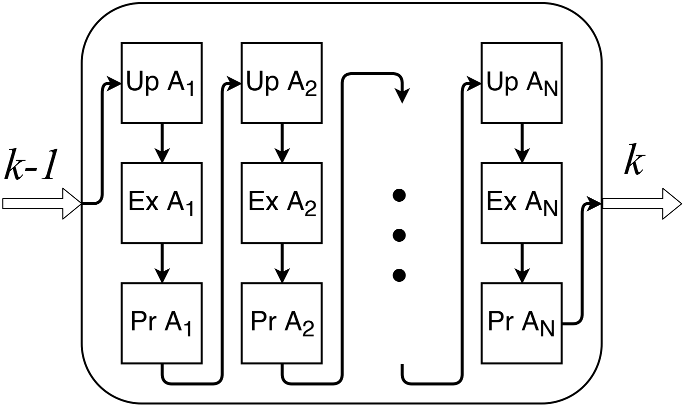
  else

    Set $\mathbf{A}_k^{(i)} = \hat{\mathbf{A}}_k^{(i)}$, $\forall i \in [N]$. *% keep the sequence* $\hat{\mathbf{A}}_k^{(i)}$

    Set $\bar{\beta}_k = \min\{1, \bar{\beta}_{k-1}\bar{\gamma}\}$, $\beta_k = \min\{\bar{\beta}_{k-1}, \beta_{k-1}\gamma\}$. *% Increase* $\bar{\beta}$ *and* β

  end if

end for


k-1
Up A_1
Ex A_1
Pr A_1
Up A_2
Ex A_2
Pr A_2
Up A_N
Ex A_N
Pr A_N
k

See thesis:
Chapter 7, p.113- p.125.

The details I skipped

- The discussion of HER in the original form for solving NMF problem.
- The details of HER on NMF and NTF problems.
- Other similar algorithms.

... see chapter 7 of the thesis!

Overview

Summary of contributions: the modeling aspect

- Minvol NMF
  - It generalizes *Separable NMF*, an important class of NMF.
  - Minvol NMF with $\mathcal{V}_{\text{det}}$ is identifiable under the SSC condition.
  - Minvol NMF is empirically superior than other approaches.
  - Minvol NMF with nuclear norm: new model.

- NuMF
  - Proposed a brute-force heuristic to solve it, with acceleration by APG, peak detection and a dimension reduction step based on a multi-grid method.
  - The multi-grid method preserves unimodality.
  - 3 identifiability theorems of NuMF in 3 special cases.

Summary of contributions: the algorithm aspect

- A generic framework named HER on accelerating both exact and inexact BCD type of algorithms in solving NMF, NTF and CPD.
- Many empirical evidences on the effectiveness of HER.
- The benefits of HER:
  - it can accelerate any BCD algorithm
  - it extrapolates the variable sequence without increasing the per-iteration computational cost of the algorithm
  - the auxiliary extrapolation sequence it produces is always feasible.
- The main shortcomings of HER are: no theoretical convergence guarantee; requires parameter tuning.
- Apart from HER, we provided efficient algorithms for solving minvol NMF and NuMF.

Summary of contributions: the application aspect

- Geography
  Hyperspectral unmixing in remote sensing.

- Music
  Audio blind source separation.

- Chemistry
  Chromatography - mass spectrometry

The end

- PhD Thesis:
  Nonnegative Matrix and Tensor Factorizations: Models, Algorithms and Applications
  Available online: `https://angms.science/doc/PhDThesis.pdf`
  Errata: `https://angms.science/doc/PhDThesisErrata.pdf`
  slide: https://angms.science/doc/PhDPublicDefSlide.pdf

- Works during the PhD period
  green = journal paper, black = conference paper
  1. **A.**, and Gillis, Volume regularized non-negative matrix factorizations, WHISPERS18
  2. **A.**, and Gillis, Algorithms and comparisons of nonnegative matrix factorizations with volume regularization for hyperspectral unmixing, JSTAR, 19
  3. **A.**, and Gillis, Accelerating nonnegative matrix factorization algorithms using extrapolation, Neural Computation, 19
  4. **A.**, Cohen and Gillis, Accelerating approximate nonnegative canonical polyadic decomposition using extrapolation, GRETSI19
  5. Leplat, Gillis, Siebert and **A.**, Séparation aveugle de sources sonores par factorization en matrices positives avec pénalité sur le volume du dictionnaire, GRETSI19
  6. Leplat, **A.** and Gillis, Minimum-volume rank-deficient nonnegative matrix factorizations, ICASSP19
  7. Leplat, Gillis and **A.**, Blind Audio Source Separation with Minimum-Volume Beta-Divergence NMF, TSP, 20
  8. **A.**, Cohen, Le and Gillis, Extrapolated Alternating Algorithms for Approximate Canonical Polyadic Decomposition, ICASSP20
  9. **A.**, Cohen, Gillis and Le, Accelerating Block Coordinate Descent for Nonnegative Tensor Factorization, NLAA, under review, 20
  10. **A.**, Gillis, Vandaele and De Sterck, Nonnegative Unimodal Matrix Factorization, submitted to ICASSP (NEW)