β-NMF and sparsity promoting regularizations for complex mixture unmixing: Application to 2D HSQC NMR.

Afef CHERNI, Sandrine ANTHOINE, Caroline CHAUX
Aix-Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France.

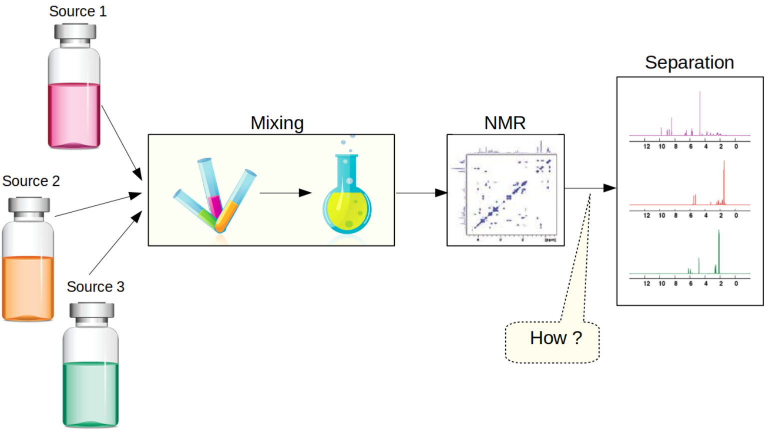
ICASSP 2020, 4-8 May, Barcelone

Outline

Introduction

- HSQC NMR experience
  - **NMR** : Nuclear Magnetic Resonance, a spectroscopy technique used to identify molecules in a given chemical mixture.
  - **2D HSQC** : Heteronuclear Single Quantum Coherence, a NMR experience used to determinate the correlations between a carbon and its attached proteins.
  - **BSS** : Blind Source Separation, an efficient mathematical method used to analyze data which are modeled as the linear combination of elementary sources or components.

Introduction



Explanatory scheme

Introduction

■ Problem statement

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}$$

- $\mathbf{X} = (x_{m,\ell}) \in \mathbb{R}^{M \times L}$: given mixtures
- $\mathbf{S} = (s_{n,\ell}) \in \mathbb{R}^{N \times L}$: unknown sources
- $\mathbf{A} = (a_{m,n}) \in \mathbb{R}^{M \times N}$: unknown mixing matrix
- $\mathbf{N} = (n_{m,\ell}) \in \mathbb{R}^{M \times L}$: acquisition noise

■ Difficulties

✗ Indeterminacies of solutions

$$(\exists \Lambda \in \mathbb{R}^{N \times N}) \quad \text{such that} \quad \mathbf{A}' = \mathbf{A}\Lambda \ \text{ et } \ \mathbf{S}' = \Lambda^{-1}\mathbf{S}$$

where Λ is a diagonal or a permutation matrix.

✗ 2D NMR spectra present a high level of sparsity with a spectral overlap and poor resolution.

Variational formulation

- Regularized approach

$$\underset{\mathbf{A},\,\mathbf{S}}{\text{minimize}}\;\; \Theta(\mathbf{A},\mathbf{S}) := \underbrace{\Phi(\mathbf{A},\mathbf{S})}_{\text{Data fidelity}} + \underbrace{\Psi(\mathbf{A},\mathbf{S})}_{\text{Regularization term}}$$

- Standard choices [Cherni et al., 2019]

$$\underset{\mathbf{A},\mathbf{S}}{\text{minimize}}\;\; \frac{1}{2}\|\mathbf{X}-\mathbf{A}\mathbf{S}\|_F^2 + \lambda_{\mathbf{A}}\Psi_{\mathbf{A}}(\mathbf{A}) + \lambda_{\mathbf{S}}\Psi_{\mathbf{S}}(\mathbf{S})$$

$\lambda_{\mathbf{A}}$ and $\lambda_{\mathbf{S}}$ are regularization parameters.
$\Psi_{\mathbf{A}}$ and $\Psi_{\mathbf{S}}$ are regularization functions.

$$\iota_+(\mathbf{u}) = \begin{cases} 0 & \text{if } u_i \geq 0\ \forall i \\ +\infty & \text{otherwise.} \end{cases}$$

$$\ell_1(\mathbf{u}) = \left(\sum_{i=1}^{L} |u_i|\right)$$

$$\mathrm{Ent}(\mathbf{u}) = \sum_{i=1}^{L} \mathrm{ent}(u_i)$$

$$\mathrm{ent}(u) = \begin{cases} u\log(u) & \text{if } u > 0 \\ 0 & \text{if } u = 0 \\ +\infty & \text{otherwise} \end{cases}$$

Proposed approach

- Generalization

$$\underset{\mathbf{A},\,\mathbf{S}}{\text{minimize}} \;\; \beta\text{-}\mathbf{div}(\mathbf{A},\mathbf{S}) + \lambda_{\mathbf{A}}\Psi_{\mathbf{A}}(\mathbf{A}) + \lambda_{\mathbf{S}}\Psi_{\mathbf{S}}(\mathbf{S})$$

where

$$(\forall\,(\mathbf{u},\mathbf{v}) \in (\mathbb{R}_+^L)^2) \quad \beta\text{-}\mathbf{div}(\mathbf{u},\mathbf{v}) = \sum_{i=1}^{L} \beta\text{-div}(u_i|v_i)$$

and for all $(u,v) \in \mathbb{R}_+^2$

$$\beta\text{-div}(u|v) = \begin{cases} \frac{1}{\beta(\beta-1)}\left(u^\beta + (\beta-1)v^\beta - \beta u v^{\beta-1}\right) & \text{if } \beta \in \mathbb{R} \setminus \{0,1\} \\ \frac{u}{v} - \log\left(\frac{u}{v}\right) - 1 & \text{if } \beta = 0 \\ u\log\left(\frac{u}{v}\right) - u + v & \text{if } \beta = 1 \end{cases}$$

- Frobenius norm is a special case of β-**div** where $\beta = 2$.

MM based multiplicative algorithm

- Generic alternating algorithm [Hunter and Lange, 2000]

For $k = 0, 1, \ldots$

$$\left\lfloor \begin{array}{ll} \mathbf{A}_{k+1} = \underset{\mathbf{A}}{\operatorname{argmin}} \ \beta\text{-}\mathbf{div}(\mathbf{X}, \mathbf{A}\mathbf{S}_k) + \lambda_{\mathbf{A}} \Psi_{\mathbf{A}}(\mathbf{A}) & (\mathbf{I}) \\ \mathbf{S}_{k+1} = \underset{\mathbf{S}}{\operatorname{argmin}} \ \beta\text{-}\mathbf{div}(\mathbf{X}, \mathbf{A}_{k+1}\mathbf{S}) + \lambda_{\mathbf{S}} \Psi_{\mathbf{S}}(\mathbf{S}) & (\mathbf{II}) \end{array} \right.$$

and for $\beta > 2$

(**I**) $\Psi_{\mathbf{A}} = \iota_+$

$$\mathbf{A}_{k+1} = \left(\frac{\left(\mathbf{X} \odot (\mathbf{A}_k\mathbf{S})^{\odot(\beta-2)}\right)\mathbf{S}^T}{(\mathbf{A}_k\mathbf{S})^{\odot(\beta-1)}\mathbf{S}^T} \right)_+^{\odot\frac{1}{\beta-1}} \odot \mathbf{A}_k$$

(**II**)-a) $\Psi_{\mathbf{S}} = \iota_+$

$$\mathbf{S}_{k+1} = \left(\frac{\mathbf{A}^T\left(\mathbf{X} \odot (\mathbf{A}\mathbf{S}_k)^{\odot(\beta-2)}\right)}{\mathbf{A}^T(\mathbf{A}\mathbf{S}_k)^{\odot(\beta-1)}} \right)_+^{\odot\frac{1}{\beta-1}} \odot \mathbf{S}_k$$

(II)-b) $\Psi_{\mathbf{S}} = \ell_1 + \iota_+$

$$\mathbf{S}_{k+1} = \left(\frac{\mathbf{A}^T (\mathbf{X} \odot (\mathbf{A}\mathbf{S}_k)^{\odot(\beta-2)}) - \lambda_{\mathbf{S}}}{\mathbf{A}^T (\mathbf{A}\mathbf{S}_k)^{\odot(\beta-1)}} \right)_+^{\odot \frac{1}{\beta-1}} \odot \mathbf{S}_k$$

(II)-c) $\Psi_{\mathbf{S}} = \mathrm{Ent} + \iota_+$

$$\mathbf{S}_{k+1} = \left(\frac{\gamma}{\alpha} \mathbf{W} \left(\frac{\alpha}{\gamma} \exp(-\frac{\delta}{\gamma}) \right) \right)_+^{\odot \frac{1}{\beta-1}} \odot \mathbf{S}_k$$

where $\odot$ denotes the Hadamard product, $\mathbf{W}$ is the W-Lambert function [Corless et al., 1996] and

$$\begin{cases} \alpha = \mathbf{A}^T (\mathbf{A}\mathbf{S}_k)^{\odot(\beta-1)} \odot \mathbf{S}_k, \\ \gamma = \frac{\lambda_{\mathbf{S}}}{\beta-1} \mathbf{S}_k, \\ \delta = \lambda_{\mathbf{S}} (\mathbf{S}_k + \mathbf{S}_k \odot \log(\mathbf{S}_k)) - \mathbf{A}^T (\mathbf{X} \odot (\mathbf{A}\mathbf{S}_k)^{\odot(\beta-2)}) \odot \mathbf{S}_k. \end{cases}$$

2D HSQC NMR data

- Data

Real case

$\mathbf{X} \in \mathbb{R}^{5\times1024\times2048}$: 5 mixtures

$\mathbf{S} \in \mathbb{R}^{4\times1024\times2048}$: 4 sources

$\mathbf{A} \in \mathbb{R}^{5\times4}$: a mixture matrix

Simulated case

$\mathbf{X}$ is simulated following model

$$\mathbf{X} = \mathbf{AS} + \mathbf{N}$$

$\mathbf{N} \sim \mathcal{N}(0, \sigma^2)$, $\sigma = 2 \times 10^4$

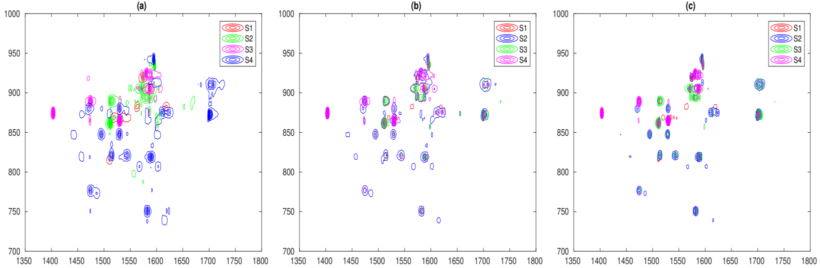
- Performance criteria [Vincent et al., 2006, Moreau et al., 1994]
  - SIR : Source to Interference Ratio
  - SDR : Source to Distortion Ratio
  - SAR : Source to Artifacts Ratio
  - Amari Index

Results on simulated data

Data fidelity term	λ_S	Ψ_S	SDR	SIR	SAR	Amari-index
Squared Frobenius	ι_+		18.073	28.854	18.514	0.0121
	0.1σ	$\ell_1 + \iota_+$	**30.299**	31.475	**39.462**	0.0272
		Ent $+\iota_+$	18.287	36.859	18.354	0.0090
	σ	$\ell_1 + \iota_+$	21.140	21.788	29.872	0.0492
		Ent $+\iota_+$	17.334	**36.909**	17.421	0.0198
	10σ	$\ell_1 + \iota_+$	17.041	25.581	22.104	0.0189
		Ent $+\iota_+$	16.021	30.625	18.216	0.0861
β-divergence	ι_+		**36.711**	40.854	**41.571**	**0.0054**
	0.1σ	$\ell_1 + \iota_+$	36.531	40.853	41.255	**0.0054**
		Ent $+\iota_+$	**36.711**	40.854	41.570	**0.0054**
	σ	$\ell_1 + \iota_+$	32.041	40.868	34.135	**0.0054**
		Ent $+\iota_+$	36.710	40.852	41.570	**0.0054**
	10σ	$\ell_1 + \iota_+$	22.906	**41.140**	23.102	**0.0054**
		Ent $+\iota_+$	36.688	40.851	41.513	**0.0054**

Average criteria obtained with various λ_S for $\beta = 3$.

Results on real data



Pure sources (a), estimated sources using Frobenius norm (b), estimated sources using β-**div** (c).

Data fidelity term	Ψ_S	SDR	SIR	SAR	Amari-index
Squared Frobenius	$\ell_1 + \iota_+$	04.984	13.956	07.951	0.1804
	Ent $+\iota_+$	05.755	**14.434**	08.446	0.1793
β-divergence	$\ell_1 + \iota_+$	**07.240**	11.487	10.574	**0.1610**
	Ent $+\iota_+$	07.220	11.396	**10.632**	0.1657

Average criteria obtained with $\lambda_S = 10\sigma$ for $\beta = 3$.

Conclusion & perspectives

✓ The β-divergence combined with ℓ_1 norm or Ent function ensures the BSS of the 2D HSQC NMR.
✓ In the real case, better SDR and SAR values are obtained using β-divergence. However, a slight deterioration on the SIR values is noticed.

- Optimize the choice of $\lambda_{\mathbf{S}}$.
- Verify the linearity of the model in the context of 2D NMR.

References

Cherni, A., Piersanti, E., Anthoine, S., Chaux, C., Shintu, L., Yemloul, M., and Torrésani, B. (2019).
Challenges in the decomposition of 2d nmr spectra of mixtures of small molecules.
Faraday Discussions.

Corless, R.-M., Gonnet, G.-H., Hare, D.-E., Jeffrey, D.-J., and Knuth, D.-E. (1996).
On the lambert W function.
Advances in Applied Mathematics, 5(1) :329–359.

Hunter, D. R. and Lange, K. (2000).
Quantile regression via an mm algorithm.
Journal of Computational and Graphical Statistics, 9(1) :60–77.

Moreau, E., Macchi, O., and Macchi, O. (1994).
A one stage self-adaptive algorithm for source separation.
In *Proceedings of ICASSP 94. IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, pages III–49. IEEE.

Vincent, E., Gribonval, R., and Févotte, C. (2006).
Performance measurement in blind audio source separation.
IEEE transactions on audio, speech, and language processing, 14(4) :1462–1469.

β-NMF and sparsity promoting regularizations for complex mixture unmixing : Application to 2D HSQC NMR.

Afef CHERNI*, Sandrine ANTHOINE, Caroline CHAUX

Aix-Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France.
Project : BIFROST, http://www.i2m.univ-amu.fr/project/bifrost/
afef.cherni@univ-amu.fr