Log-determinant constrained Non-negative Matrix Factorization

Andersen Man Shun Ang; Nicolas Gillis

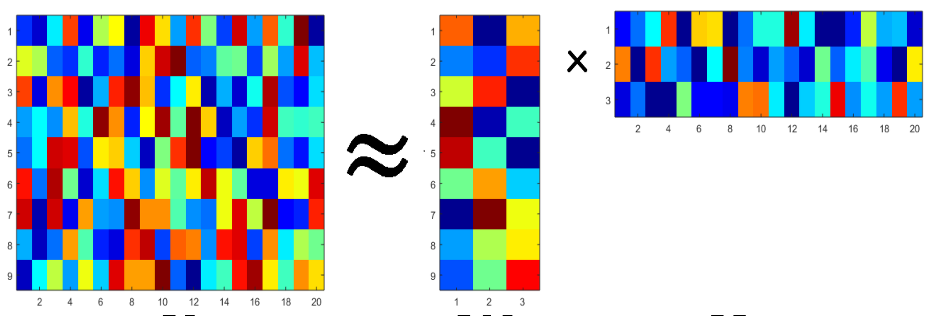
Department of Mathematics and Operational Research, Faculté Polytechnique, Université de Mons

Overview

- Decomposition of non-negative matrices by *Non-negative Matrix Factorization* (**NMF**).
- Regularized by log-determinant which promote "minimum volume".
- Goal : develop fast and flexible algorithm for the log-determinant constrained non-negative matrix factorization.

Introduction to Non-negative Matrix Factorization (NMF)

- **NMF problem** : given data matrix $\mathbf{X} \in \mathbb{R}_{+}^{p\times n}$, find two smaller matrices $\mathbf{W} \in \mathbb{R}_{+}^{p\times r}$, $\mathbf{H} \in \mathbb{R}_{+}^{r\times n}$ such that $\mathbf{X} \approx \mathbf{WH}$, with $r \ll \min\{m,n\}$.

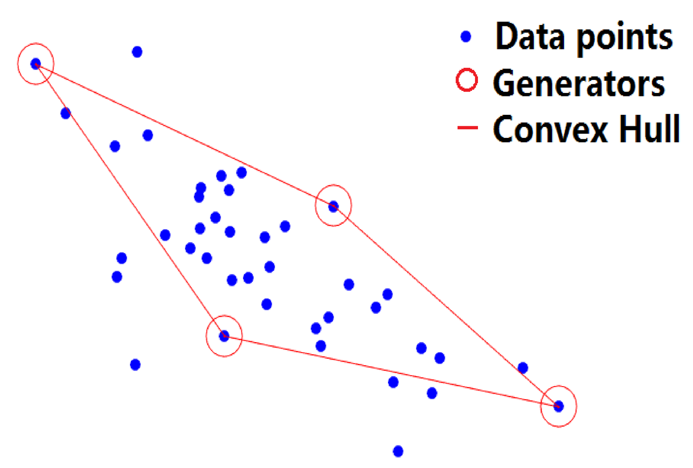


$\boldsymbol{X}$ $\boldsymbol{W}$ $\boldsymbol{H}$

- **The optimization problem**

  $\min_{W,H} f(W,H) = \|\boldsymbol{X} - \boldsymbol{WH}\|_{\phi}^{2}$

  where $\phi \in \{F, 1 \leq p \leq 2\}$.
- **Why NMF** : Interpretation and part-based representation of data. The basis matrix $\boldsymbol{W}$ obtained can be used for data characterization.
- **Solution of NMF is non-unique** → need further constraints on $\boldsymbol{W}$ and/or $\boldsymbol{H}$.
- **Separability condition** : one way to turn NMF from NP-hard to tractable problem. It assumes all data points are spanned by a set of generators within the dataset.



- Algebraically $\mathbf{W} = \boldsymbol{X}_{:A} \in \mathbb{R}_{+}^{p\times|A|}$, where A is a column index set with $|A| = r$.

Log-det constrained NMF

- Determinant ≈ "volume".
- Log-det : log prevents λ_{Max} dominating the volume expression.
- Log-det regularized NMF :

  $F(W,H) = f(W,H) + \beta g(W)$

  $g(W) = \operatorname{logdet}(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W} + \delta\mathbf{I})$
- $\beta \geq 0$ balances the error fitting term and regularization term.
- $\delta \geq 0$ lower bounds the log function.

Problems of the log-det constraint :

- $\operatorname{logdet}(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W} + \delta\mathbf{I})$ is not convex nor concave in $\boldsymbol{W}$.
- Proximal operator for $\operatorname{logdet}(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W} + \delta\mathbf{I})$ cannot be easily computed.

Solving log-det NMF by Coordinate Descent

- Key point : a logdet-trace inequality

  $\operatorname{logdet}(A) \leq \operatorname{Tr}(A - I)$

$\Rightarrow \operatorname{logdet}(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W} + \delta\mathbf{I}) \leq \operatorname{Tr}(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W} + (\delta - 1)I)$

$\Rightarrow F(W,H) \leq f(W,H) + \beta\operatorname{Tr}(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W} + (\delta - 1)I)$

After some algebra, the function F (defined above) on one column w_i of the matrix $\boldsymbol{W}$ is

$$F(w_i) \leq w_i Q_w w_i + p_w^T w_i + c_w$$

where

- $Q_w = (\|h_i\|_2^2 + \beta\,)I_p$
- $p_w = -2h_i^T X_i^T$
- $X_i = X - \sum_{j\neq i} w_j h_j$
- $F(h_i)$: the function F on h_i of the matrix $\boldsymbol{H}$, has a similar expression.

Log-determinant constrained NMF Coordinate Descent Algorithm

INPUT : $\mathbf{X} \in \mathbb{R}_{+}^{p\times n}$, desired r

Initialization of $\boldsymbol{W}$ and $\boldsymbol{H}$

FOR $k = 1$ *to* k_{Max} iteration

    FOR $i = 1$ *to* r iteration

        Update w_i by minimizing $F(w_i)$

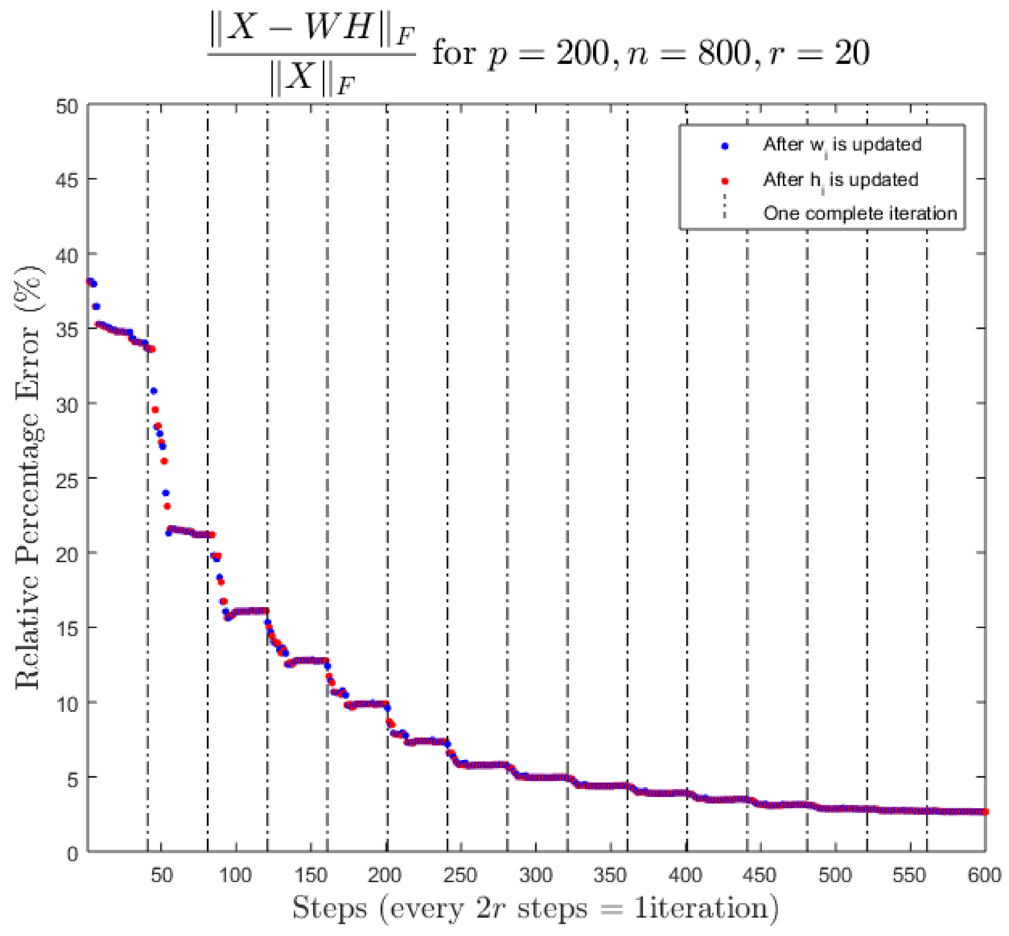
        Update h_i by minimizing $F(h_i)$

    END FOR

END FOR

Simulation

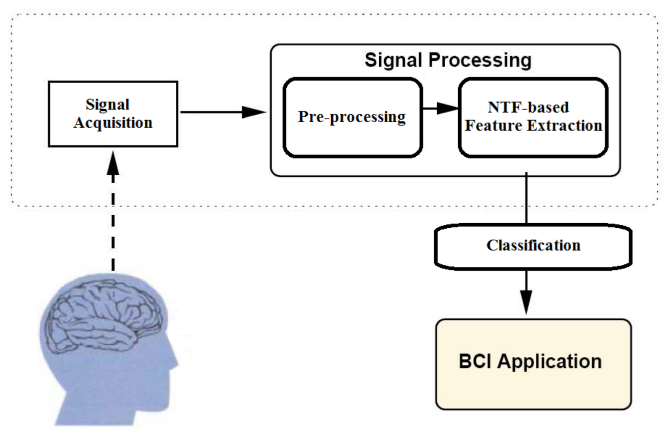
- As $Q_w = (\|h_i\|_2^2 + \beta\,)I_m$ is always non-singular, the quadratic sub-problems always have unique solution.
- The expression $\boldsymbol{WH}$ does allow either $\mathbf{W}$ or $\boldsymbol{H}$ to grow very large. Normalization is needed.
- An simulation example shows the algorithm converges.



Applications and extensions

- **Brain Computer Interface**

The NMF method can be applied on the spectrum data in BCI for feature extraction.



- **Human facial data applications. E.g. compression and recognition.**

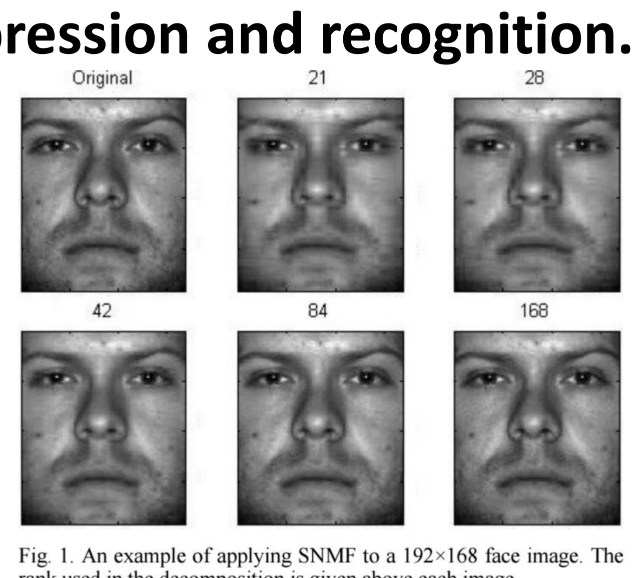
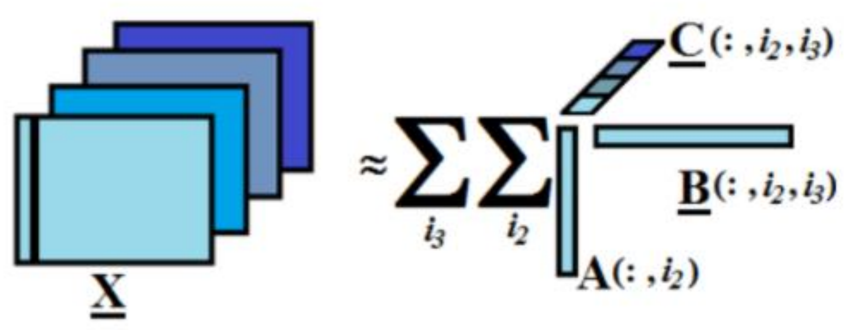


Fig. 1. An example of applying SNMF to a 192×168 face image. The rank used in the decomposition is given above each image.

- **Non-negative Tensor Factorization**



Contact

Andersen Man Shun ANG

Department of Mathematics and Operational Research

Faculté polytechnique, Université de Mons

Email: ManShun.ANG@umons.ac.be

Homepage: angms.science

UMONS Université de Mons

European Research Council
Established by the European Commission