Nonnegative matrix factorization with polynomial signals via hierarchical alternating least squares

Cécile HAUTECOEUR and François GLINEUR: UCLouvain, CORE & ICTEAM/INMA

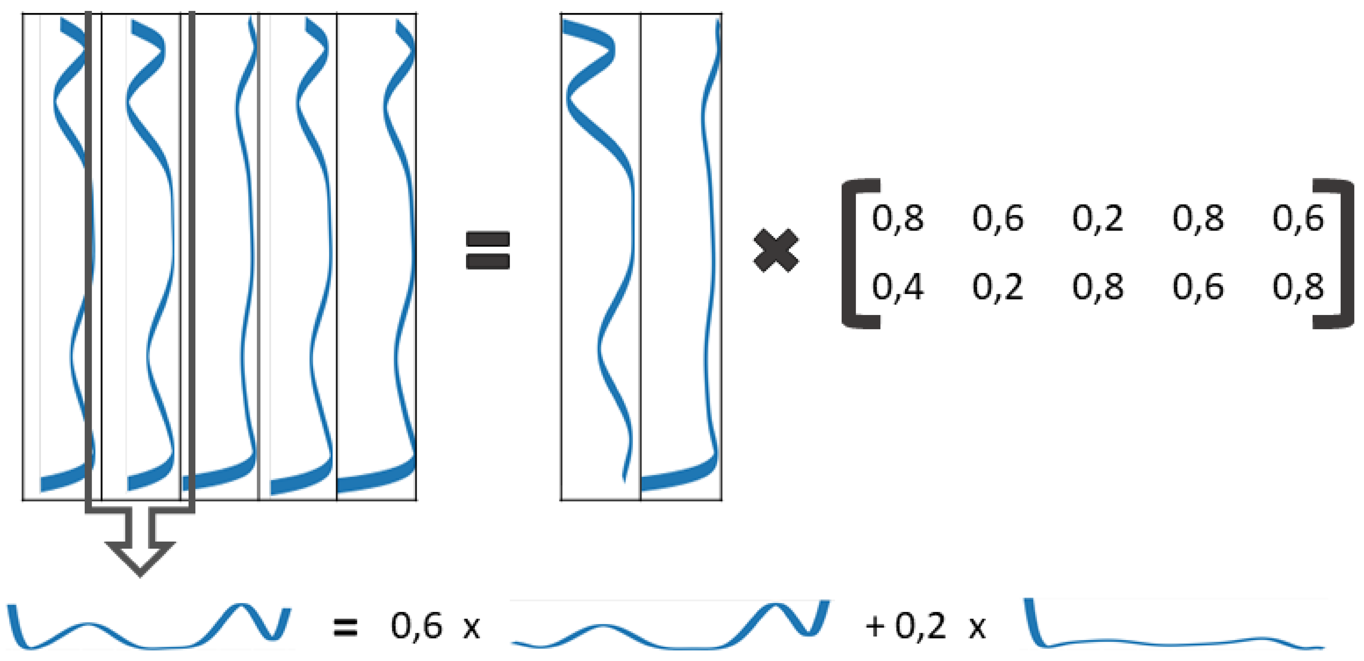
Nonnegative Matrix Factorization (NMF)

Problem: Given nonnegative matrix $Y \in \mathbb{R}_+^{m\times n}$ and integer rank $r > 0$, find nonnegative matrices $A \in \mathbb{R}_+^{m\times r}$ and $X \in \mathbb{R}_+^{r\times n}$ such that $||Y - AX||$ is minimal ($Y \simeq AX$).

Hence NMF expresses each element in given data (Y) as a nonnegative linear combinations of a few well-chosen common basis elements (in X).

NMF with polynomial signals (P-NMF)

A novel extension of usual NMF: The provided data, the n columns of matrix Y, are (observations of) **continuous signals**, $Y = \{y_i(t)\}_{i=1}^n$, over interval $t \in [a,b]$. Columns of matrix A are (observations of) **polynomials** with fixed degreee D, $A = \{a_k(t)\}_{k=1}^r$, and not vectors anymore.



Two possible cost functions:

- **PS-NMF:** Data are discretized over m points $\{\tau_j\}_{j=1}^m$:
$$\min_{A,X} \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_i(\tau_j) - \textstyle\sum_{k=1}^r a_k(\tau_j)x_{k,i}\right)^2.$$
- **PI-NMF:** Data are polynomials with known coefficients:
$$\min_{A,X} \sum_{i=1}^{n}\int_a^b \left(y_i(t) - \textstyle\sum_{k=1}^r a_k(t)x_{k,i}\right)^2 dt.$$

Algorithms

Hierarchical alternating least squares (HALS): The idea of this algorithm is to update alternatively matrix A and matrix X until convergence. Moreover, during the update of A or X, each column is successively updated individually, taking into account updates of previous columns [1].

Usual NMF:

$$a_{:j} \leftarrow \left[\frac{Yx_{j:}^\top - \sum_{k\neq j} a_{:k}x_{k:}\,x_{j:}^\top}{x_{j:}x_{j:}^\top}\right]_+ \quad x_{j:} \leftarrow \left[\frac{a_{:j}^\top Y - \sum_{k\neq j} a_{:j}^\top a_{:k}x_{k:}}{a_{:j}^\top a_{:j}}\right]_+$$

PS-NMF: We consider B the matrix of coefficients of polynomials in A, Π the Vandermonde matrix ($A = \Pi B$) and $\tilde{\Pi} = \Pi^\top\Pi$. Moreover, P is the set of polynomials nonnegatives over interval $[a,b]$. HALS become:

$$b_{:j} \leftarrow \left[\frac{\Pi^\dagger Y x_{j:}^\top - \sum_{k\neq j} b_{:k}x_{k:}\,x_{j:}^\top}{x_{j:}x_{j:}^\top}\right]_P \quad x_{j:} \leftarrow \left[\frac{b_{:j}^\top \Pi^\top Y - \sum_{k\neq j} b_{:j}^\top\tilde{\Pi}b_{:k}x_{k:}}{b_{:j}^\top\tilde{\Pi}b_{:j}}\right]_+$$

PI-NMF: We consider Z the matrix of coefficients of polynomials in Y and $M = \int_a^b \Pi(t)^\top\Pi(t)\,dt$. Surprisingly, updates are still closed form:

$$b_{:j} \leftarrow \left[\frac{Zx_{j:}^\top - \sum_{k\neq j} b_{:k}x_{k:}\,x_{j:}^\top}{x_{j:}x_{j:}^\top}\right]_P \quad x_{j:} \leftarrow \left[\frac{b_{:j}^\top MZ - \sum_{k\neq j} b_{:j}^\top Mb_{:k}x_{k:}}{b_{:j}^\top Mb_{:j}}\right]_+$$

However both PS-NMF and PI-NMF require a new operation $[\cdot]_P$ namely **projection over the set of nonnegative polynomials.**

Projection onto the set of nonnegative polynomials

Problem to solve: Given a polynomial f, find the polynomial g that is both nonnegative and closest to f over interval $[a,b]$.

$$\min_g \int_a^b (g(t) - f(t))^2\,dt \qquad \text{such that } g \in P.$$

Parametrization of nonnegative polynomial over [-1,1]:

$g(t) \geq 0$ for $t \in [-1,1] \Leftrightarrow g(t) = a(t) + (1-t^2)c(t)$ where $a(t), c(t) \geq 0\ \forall t$.

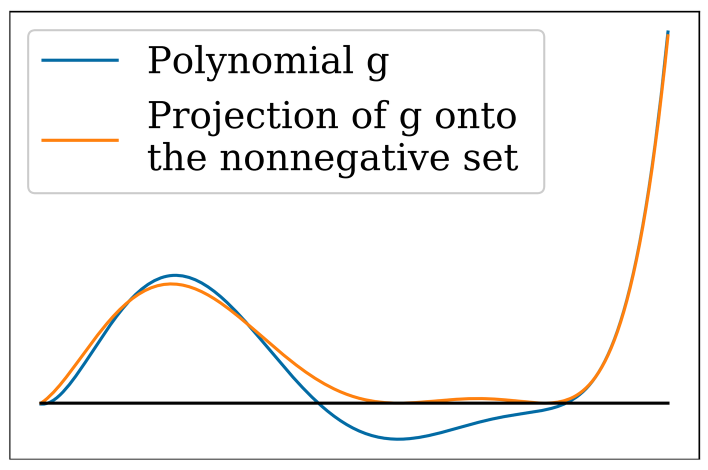
a has degree D (as g) and c has degree $D-2$ (Markov-Lukacs).

Parametrization of nonnegative polynomial over $\mathbb{R}$:

$$g(t) \geq 0\ \forall t \Leftrightarrow g(t) = \textstyle\sum_i h_i(t)^2 \qquad \text{(SOS)}.$$

Moreover, the set of nonnegative polynomials of degree D can be represented using an LMI (linear matrix inequality) involving a positive semidefinite matrix of size $(D/2+1)$.

Algorithm: Using previous information it is possible to define our projection as a semidefinite optimization problem, and solve it with an appropriate solver (such as interior-point MOSEK).



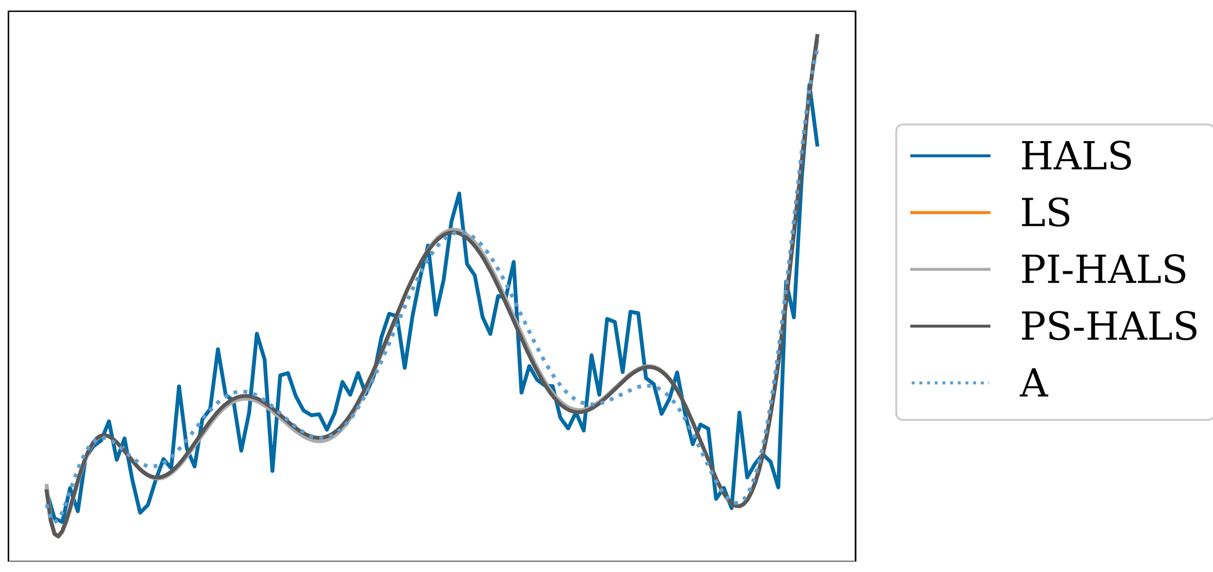
Observations and results

The P-NMF problem has already been considered by Debals and al. in [2]. In this paper, the authors use a least-squares solver in a non-alternative way. We compare our algorithms to usual HALS and the least-square approach, denoted LS.

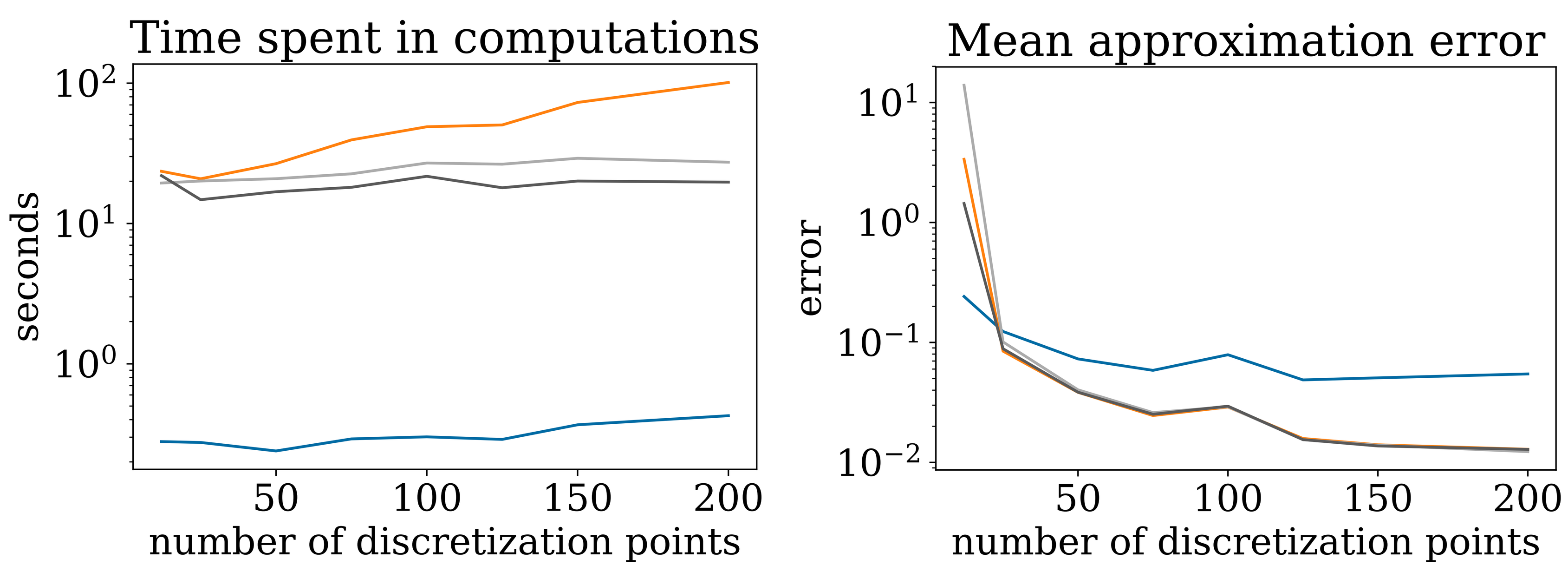
Signals recovered in A:

◇ Less sensitive to noise than HALS.

◇ Recover smoother signals.



Performances:



◇ Error similar to LS. ◇ CPU time increases slowly with problem size.

Further work

◇ Accelerate projection.
◇ Consider other parametrizable signals (such as splines).

References and acknowledgments

[1] Cichocki A., and al. *Hierarchical als algorithms for nonnegative matrix and 3d tensor factorization.* In International Conference on Independent Component Analysis and Signal Separation, pp.169–176, Springer, 2007.

[2] Debals O., and al. *Nonnegative matrix factorization using nonnegative polynomial approximations.* IEEE Signal Processing Letters, 24(7), pp.948–952, 2017.

This work was supported by the Fonds de la Recherche Scientifique – FNRS and the Fonds Wetenschappelijk Onderzoek – Vlaanderen under EOS Project no 30468160.