Hierarchical Nonnegative Tensor Decompositions

by Jamie Haddock
(Harvey Mudd College, Department of Mathematics)
on September 30, 2023,
AWM Research Symposium 2023 “Tensor Methods for data modeling”

https://ieeexplore.ieee.org/document/9022678 (CAMSAP 2019)
joint with M. Gao•, D. Molitor, E. Sadovnik, T. Will•, R. Zhang•, D. Needell

https://ieeexplore.ieee.org/document/9723126 (ACSSC 2021)
joint with Joshua Vendrow•, Deanna Needell

https://ieeexplore.ieee.org/document/9747810 (ICASSP 2022)
joint with Joshua Vendrow•, Deanna Needell

NSF DMS #2211318

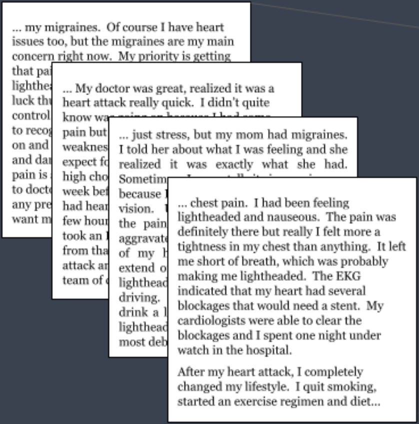
matH of Algorithms, Data & Decisions research group @ HARVEY MUDD COLLEGE

Consider applying for the
Institute for Advanced Study
Women and Mathematics
summer program!

Deadline is in February each year.

Motivation

» Learn trends in high-dimensional data

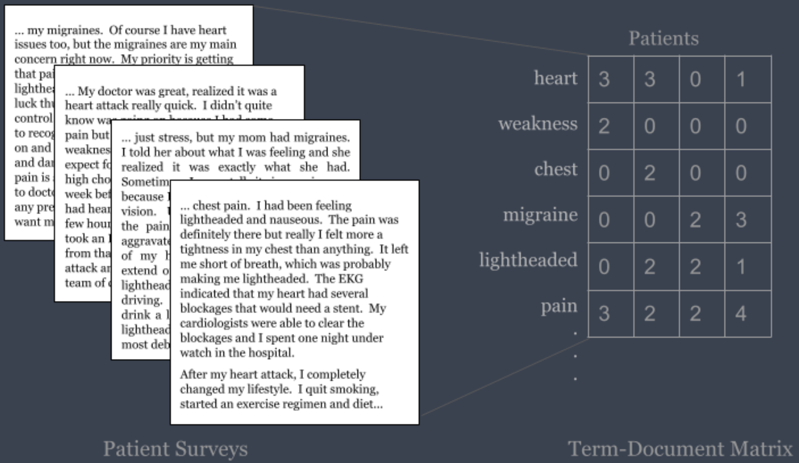


Patient Surveys

	Patients			
heart	3	3	0	1
weakness	2	0	0	0
chest	0	2	0	0
migraine	0	0	2	3
lightheaded	0	2	2	1
pain	3	2	2	4
⋮				

Term-Document Matrix

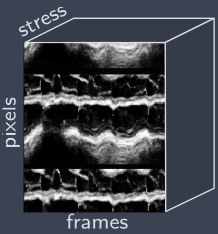
» Learn trends in high-dimensional data



Understand symptom trends and shared patient experiences automatically.

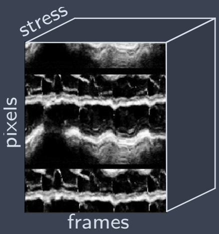
» Learn trends in high-dimensional data

» Learn trends in high-dimensional data



» Learn trends in high-dimensional data

Learn cohesive parts and separate noise in medical image studies.



Can we tell how the resulting parts/topics are related?

Can we tell how the resulting parts/topics are related?

How do we choose the number of topics or parts to learn?

Can we tell how the resulting parts/topics are related?

How do we choose the number of topics or parts to learn?

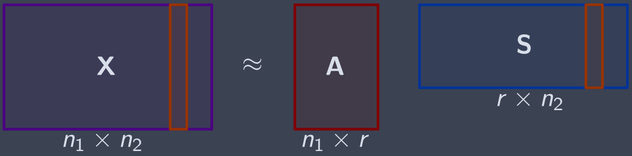
Hierarchical matrix factorization and tensor decomposition topic models!

Introduction

» Nonnegative Matrix Factorization (NMF)

Model: Given nonnegative data $\mathbf{X}$, compute nonnegative $\mathbf{A}$ and $\mathbf{S}$ of lower rank so that

$$\mathbf{X} \approx \mathbf{A}\mathbf{S}.$$

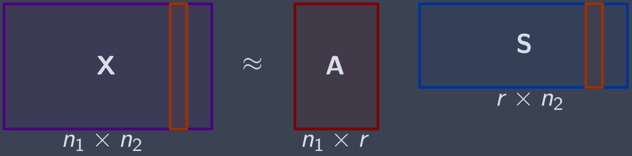


Lee, Daniel D., and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." Nature 401.6755 (1999): 788-791.

» Nonnegative Matrix Factorization (NMF)

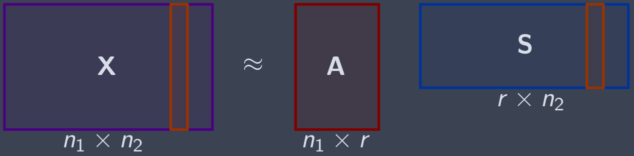
Model: Given nonnegative data $\mathbf{X}$, compute nonnegative $\mathbf{A}$ and $\mathbf{S}$ of lower rank so that

$$\mathbf{X} \approx \mathbf{A}\mathbf{S}.$$



Lee, Daniel D., and H. Sebastian Seung. “Learning the parts of objects by non-negative matrix factorization.” Nature 401.6755 (1999): 788-791.

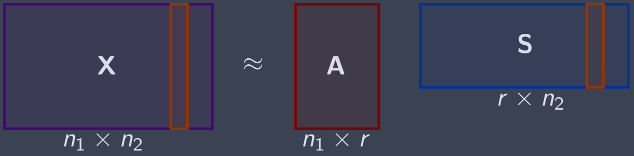
» Nonnegative Matrix Factorization (NMF)



▷ Employed for dimensionality-reduction and topic modeling

Lee, Daniel D., and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." Nature 401.6755 (1999): 788-791.

» Nonnegative Matrix Factorization (NMF)

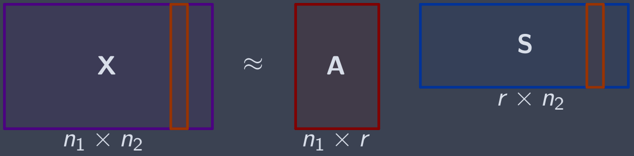


- ▷ Employed for dimensionality-reduction and topic modeling
- ▷ Often formulated as

$$\min_{\mathbf{A}\in\mathbb{R}_{\geq 0}^{n_1\times r},\mathbf{S}\in\mathbb{R}_{\geq 0}^{r\times n_2}} \|\mathbf{X}-\mathbf{AS}\|_F^2 \quad \text{or} \quad \min_{\mathbf{A}\in\mathbb{R}_{\geq 0}^{n_1\times r},\mathbf{S}\in\mathbb{R}_{\geq 0}^{r\times n_2}} D(\mathbf{X}\|\mathbf{AS}).$$ [1]

[1] Lee, Daniel D., and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." Nature 401.6755 (1999): 788-791.

» Nonnegative Matrix Factorization (NMF)



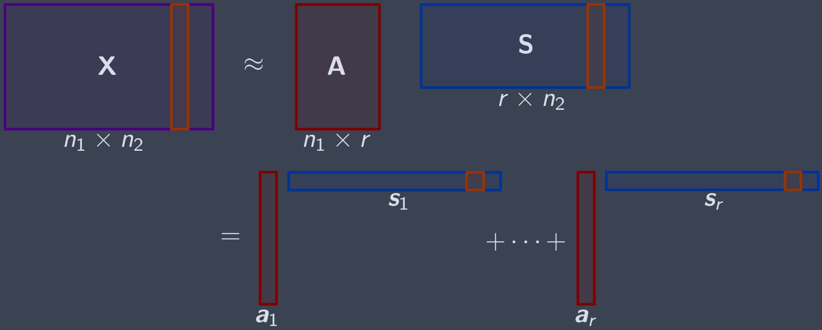
- ▷ Employed for dimensionality-reduction and topic modeling
- ▷ Often formulated as

$$\min_{\mathbf{A}\in\mathbb{R}_{\geq 0}^{n_1\times r},\mathbf{S}\in\mathbb{R}_{\geq 0}^{r\times n_2}} \|\mathbf{X}-\mathbf{AS}\|_F^2 \quad \text{or} \quad \min_{\mathbf{A}\in\mathbb{R}_{\geq 0}^{n_1\times r},\mathbf{S}\in\mathbb{R}_{\geq 0}^{r\times n_2}} D(\mathbf{X}\|\mathbf{AS}).$$[1]

- ▷ non-convex optimization problems

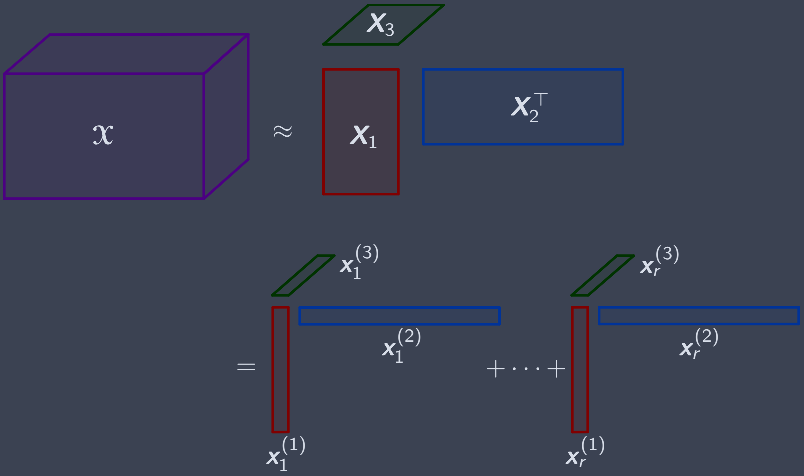
[1] Lee, Daniel D., and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." Nature 401.6755 (1999): 788-791.

» Nonnegative Matrix Factorization (NMF)



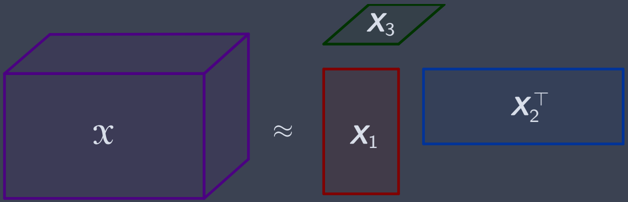
Lee, Daniel D., and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." Nature 401.6755 (1999): 788-791.

» Nonnegative CANDECOMP/PARAFAC (CP) decomposition (NCPD)



Carroll, J. Douglas, and Jih-Jie Chang. "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition." Psychometrika 35.3 (1970): 283-319.
Harshman, Richard A. "Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multimodal factor analysis." (1970): 1-84.

» Nonnegative CANDECOMP/PARAFAC (CP) decomposition (NCPD)



- formulated as $\min_{\boldsymbol{X}_i \geq 0} \|\mathcal{X} - [\![\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_k]\!]\|_F^2$ where

$$[\![\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_k]\!] \equiv \sum_{j=1}^{r} \boldsymbol{x}_j^{(1)} \otimes \boldsymbol{x}_j^{(2)} \otimes \cdots \otimes \boldsymbol{x}_j^{(k)}$$

Carroll, J. Douglas, and Jih-Jie Chang. "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition." Psychometrika 35.3 (1970): 283-319.
Harshman, Richard A. "Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multimodal factor analysis." (1970): 1-84.

» Hierarchical NMF

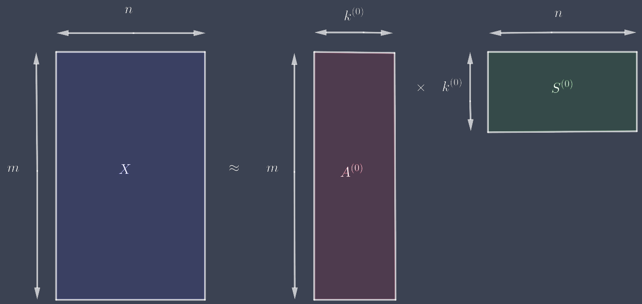
Model: Sequentially factorize

Cichocki, Andrzej, and Rafal Zdunek. "Multilayer nonnegative matrix factorisation." ELECTRONICS LETTERS-IEE 42.16 (2006): 947.

» Hierarchical NMF

Model: Sequentially factorize

$$X \approx A^{(0)} S^{(0)}$$

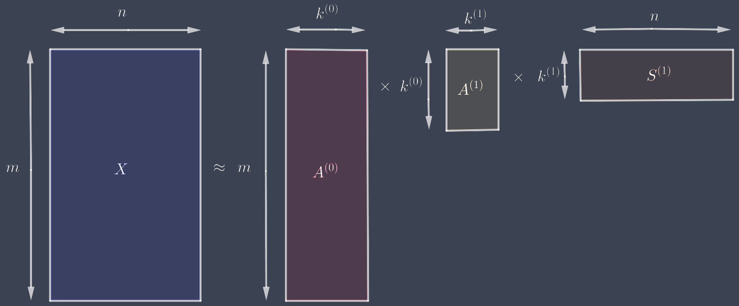


Cichocki, Andrzej, and Rafal Zdunek. "Multilayer nonnegative matrix factorisation." ELECTRONICS LETTERS-IEE 42.16 (2006): 947.

» Hierarchical NMF

Model: Sequentially factorize

$$X \approx A^{(0)} S^{(0)}, S^{(0)} \approx A^{(1)} S^{(1)}$$

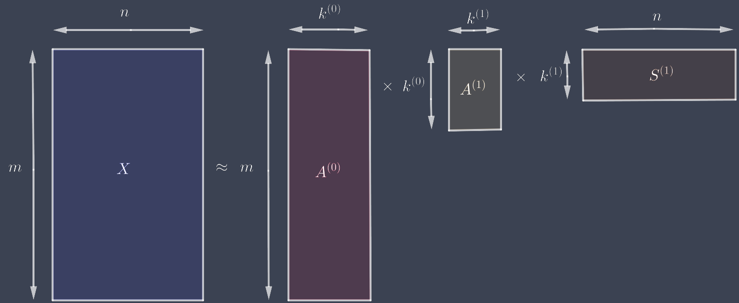


Cichocki, Andrzej, and Rafal Zdunek. "Multilayer nonnegative matrix factorisation." ELECTRONICS LETTERS-IEE 42.16 (2006): 947.

» Hierarchical NMF

Model: Sequentially factorize

$$X \approx A^{(0)}S^{(0)}, S^{(0)} \approx A^{(1)}S^{(1)}, S^{(1)} \approx A^{(2)}S^{(2)}, ..., S^{(\mathcal{L}-1)} \approx A^{(\mathcal{L})}S^{(\mathcal{L})}.$$



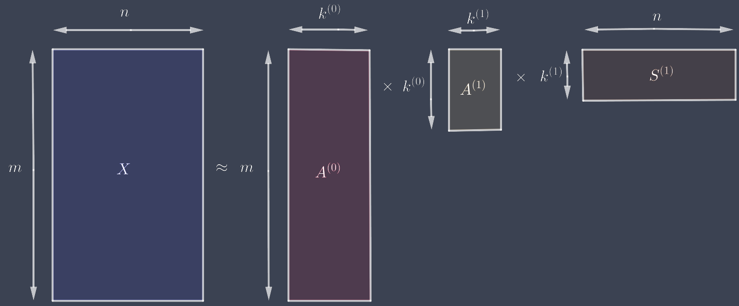
- ▷ $k^{(\ell)}$: supertopics collecting $k^{(\ell-1)}$ subtopics

Cichocki, Andrzej, and Rafal Zdunek. "Multilayer nonnegative matrix factorisation." ELECTRONICS LETTERS-IEE 42.16 (2006): 947.

» Hierarchical NMF

Model: Sequentially factorize

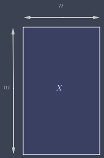
$$X \approx A^{(0)}S^{(0)}, S^{(0)} \approx A^{(1)}S^{(1)}, S^{(1)} \approx A^{(2)}S^{(2)}, ..., S^{(\mathcal{L}-1)} \approx A^{(\mathcal{L})}S^{(\mathcal{L})}.$$



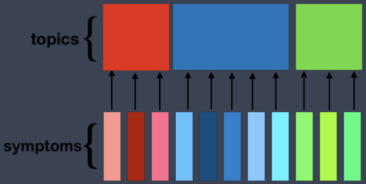
- ▷ $k^{(\ell)}$: supertopics collecting $k^{(\ell-1)}$ subtopics
- ▷ provides relationship between data matrix modes and $k^{(\ell)}$ topics

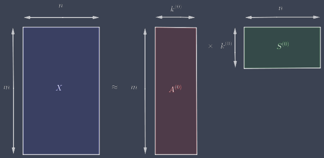
Cichocki, Andrzej, and Rafal Zdunek. "Multilayer nonnegative matrix factorisation." ELECTRONICS LETTERS-IEE 42.16 (2006): 947.

» Hierarchical NMF

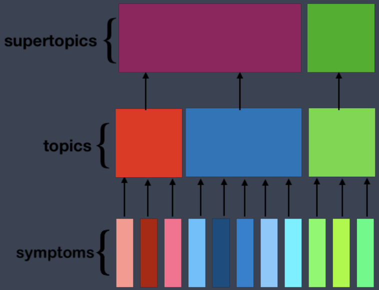


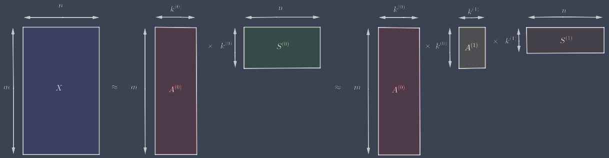
» Hierarchical NMF



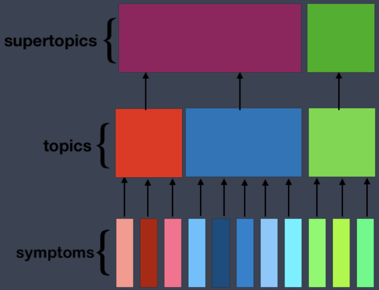


» Hierarchical NMF

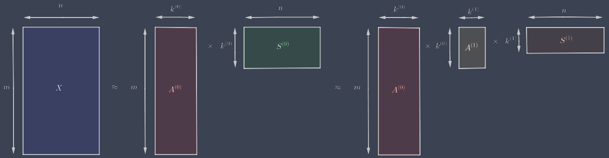




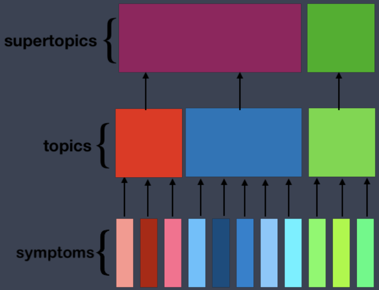
» Hierarchical NMF



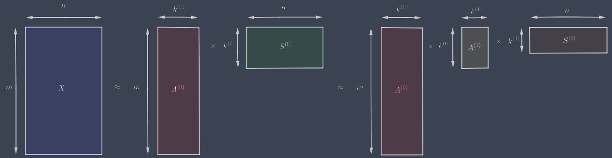
▷ elucidates the hierarchical relationships of learned topics



» Hierarchical NMF



- ▷ elucidates the hierarchical relationships of learned topics
- ▷ no need to choose a fixed model rank (number of topics)



Hierarchical Models

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

- Cichocki, Andrzej, Rafal Zdunek, and Shun-ichi Amari. "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization." International Conference on Independent Component Analysis and Signal Separation. Springer, Berlin, Heidelberg, 2007.

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

- Cichocki, Andrzej, Rafal Zdunek, and Shun-ichi Amari. "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization." International Conference on Independent Component Analysis and Signal Separation. Springer, Berlin, Heidelberg, 2007.

Results depend upon hyperparameter choice (mode).

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- ▷ Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- ▷ Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- ▷ Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

- ▷ Cichocki, Andrzej, Rafal Zdunek, and Shun-ichi Amari. "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization." International Conference on Independent Component Analysis and Signal Separation. Springer, Berlin, Heidelberg, 2007.

Results depend upon hyperparameter choice (mode).

- * Vendrow, Joshua, Jamie Haddock, and Deanna Needell. "Neural nonnegative CP decomposition for hierarchical tensor analysis." 2021 55th Asilomar Conference on Signals, Systems, and Computers. IEEE, 2021.

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- ▷ Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- ▷ Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- ▷ Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

- ▷ Cichocki, Andrzej, Rafal Zdunek, and Shun-ichi Amari. "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization." International Conference on Independent Component Analysis and Signal Separation. Springer, Berlin, Heidelberg, 2007.

Results depend upon hyperparameter choice (mode).

- * Vendrow, Joshua, Jamie Haddock, and Deanna Needell. "Neural nonnegative CP decomposition for hierarchical tensor analysis." 2021 55th Asilomar Conference on Signals, Systems, and Computers. IEEE, 2021.

Not a single hierarchical relationship, good training method.

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- ▷ Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- ▷ Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- ▷ Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

- ▷ Cichocki, Andrzej, Rafal Zdunek, and Shun-ichi Amari. "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization." International Conference on Independent Component Analysis and Signal Separation. Springer, Berlin, Heidelberg, 2007.

Results depend upon hyperparameter choice (mode).

- * Vendrow, Joshua, Jamie Haddock, and Deanna Needell. "Neural nonnegative CP decomposition for hierarchical tensor analysis." 2021 55th Asilomar Conference on Signals, Systems, and Computers. IEEE, 2021.

Not a single hierarchical relationship, good training method.

- * Vendrow, Joshua, Jamie Haddock, and Deanna Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

» Hierarchical Tensor Decompositions

How do we generalize HNMF to a higher-order tensor model?

- ▷ Vasilescu, M. Alex O., and Eric Kim. "Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors." arXiv preprint arXiv:1911.04180 (2019).
- ▷ Song, Le, et al. "Hierarchical tensor decomposition of latent tree graphical models." International Conference on Machine Learning. PMLR, 2013.
- ▷ Grasedyck, Lars. "Hierarchical singular value decomposition of tensors." SIAM Journal on Matrix Analysis and Applications 31.4 (2010): 2029-2054.

- ▷ Cichocki, Andrzej, Rafal Zdunek, and Shun-ichi Amari. "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization." International Conference on Independent Component Analysis and Signal Separation. Springer, Berlin, Heidelberg, 2007.

Results depend upon hyperparameter choice (mode).

- * Vendrow, Joshua, Jamie Haddock, and Deanna Needell. "Neural nonnegative CP decomposition for hierarchical tensor analysis." 2021 55th Asilomar Conference on Signals, Systems, and Computers. IEEE, 2021.

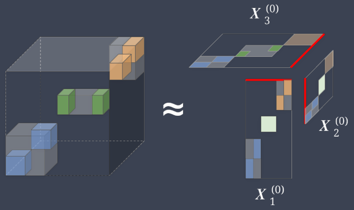
  Not a single hierarchical relationship, good training method.
- * Vendrow, Joshua, Jamie Haddock, and Deanna Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

  ** Single hierarchical relationship, naive training method. **

» Multi-HNTF Model

This model learns

$$\mathcal{X} \approx [\![\boldsymbol{X}_1^{(0)}, \boldsymbol{X}_2^{(0)}, \cdots, \boldsymbol{X}_k^{(0)}]\!]$$



Vendrow, H., Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

» Multi-HNTF Model

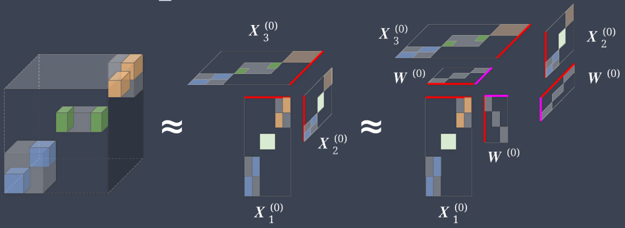
This model learns

$$\mathcal{X} \approx [\![\boldsymbol{X}_1^{(0)}, \boldsymbol{X}_2^{(0)}, \cdots, \boldsymbol{X}_k^{(0)}]\!] \approx [\![\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \cdots, \boldsymbol{X}_k^{(1)}]\!]$$

where

$$\boldsymbol{X}_j^{(\ell+1)} = \boldsymbol{X}_j^{(\ell)} \boldsymbol{W}^{(\ell)},$$

and $\boldsymbol{W}^{(\ell)} \in \mathbb{R}_{\geq 0}^{r^{(\ell-1)} \times r^{(\ell)}}$.



Vendrow, H., Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

» Multi-HNTF Model

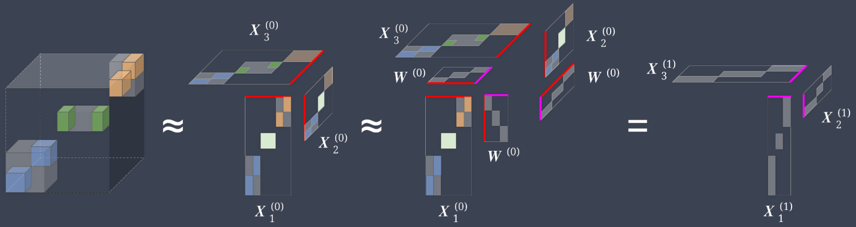
This model learns

$$\mathcal{X} \approx [\![X_1^{(0)}, X_2^{(0)}, \dots, X_k^{(0)}]\!] \approx [\![X_1^{(1)}, X_2^{(1)}, \dots, X_k^{(1)}]\!]$$

where

$$X_j^{(\ell+1)} = X_j^{(\ell)} W^{(\ell)},$$

and $W^{(\ell)} \in \mathbb{R}_{\geq 0}^{r^{(\ell-1)} \times r^{(\ell)}}$.



Vendrow, H., Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

» Multi-HNTF Model

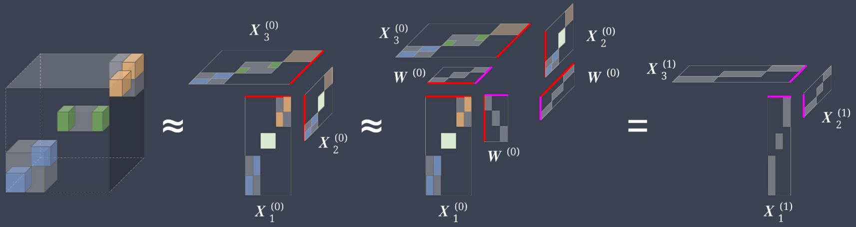
This model learns

$$\mathcal{X} \approx [\![X_1^{(0)}, X_2^{(0)}, \dots, X_k^{(0)}]\!] \approx [\![X_1^{(1)}, X_2^{(1)}, \dots, X_k^{(1)}]\!] \approx \cdots$$
$$\approx [\![X_1^{(\mathcal{L}-1)}, X_2^{(\mathcal{L}-1)}, \dots, X_k^{(\mathcal{L}-1)}]\!]$$

where

$$X_j^{(\ell+1)} = X_j^{(\ell)} W^{(\ell)},$$

and $W^{(\ell)} \in \mathbb{R}_{\geq 0}^{r^{(\ell-1)} \times r^{(\ell)}}$.



Vendrow, H., Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

» Multi-HNTF Model

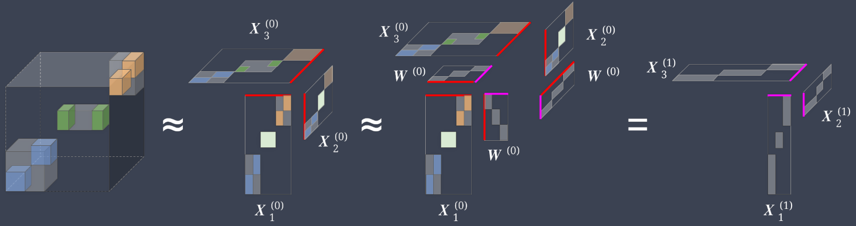
This model learns

$$\mathcal{X} \approx [\![X_1^{(0)}, X_2^{(0)}, \ldots, X_k^{(0)}]\!] \approx [\![X_1^{(1)}, X_2^{(1)}, \ldots, X_k^{(1)}]\!] \approx \cdots$$
$$\approx [\![X_1^{(\mathcal{L}-1)}, X_2^{(\mathcal{L}-1)}, \ldots, X_k^{(\mathcal{L}-1)}]\!]$$

where

$$X_j^{(\ell+1)} = X_j^{(\ell)} W^{(\ell)},$$

and $W^{(\ell)} \in \mathbb{R}_{\geq 0}^{r^{(\ell-1)} \times r^{(\ell)}}$.



A single hierarchical relationship for all modes!

Vendrow, H., Needell. "A Generalized Hierarchical Nonnegative Tensor Decomposition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

» Training Process

procedure $\textsc{Multi-HNTF}(\mathcal{X})$

$\quad \{\boldsymbol{X}_i^{(0)}\}_{i=1}^{k} \leftarrow \mathrm{NCPD}(\mathcal{X}, r_0)$

$\quad$ **for** $\ell = 0 \ldots \mathcal{L}$ **do**

$\quad\quad \boldsymbol{W}^{(\ell)} \leftarrow \mathrm{argmin}_{\boldsymbol{W} \in \mathbb{R}_+^{r_\ell \times r_{\ell+1}}} \|\mathcal{X} - [\![\boldsymbol{X}_1^{(\ell)}\boldsymbol{W}, \ldots, \boldsymbol{X}_k^{(\ell)}\boldsymbol{W}]\!]\|$

$\quad\quad$ **for** $i = 0 \ldots k$ **do**

$\quad\quad\quad \boldsymbol{X}_i^{(\ell+1)} = \boldsymbol{X}_i^{(\ell)}\boldsymbol{W}^{(\ell)}$

» Training Process

procedure MULTI-HNTF($\mathcal{X}$)
  $\{\boldsymbol{X}_i^{(0)}\}_{i=1}^{k} \leftarrow \mathrm{NCPD}(\mathcal{X}, r_0)$
  for $\ell = 0 \ldots \mathcal{L}$ **do**
    $\boldsymbol{W}^{(\ell)} \leftarrow \mathrm{argmin}_{\boldsymbol{W} \in \mathbb{R}_+^{r_\ell \times r_{\ell+1}}} \|\mathcal{X} - [\![\boldsymbol{X}_1^{(\ell)}\boldsymbol{W}, \ldots, \boldsymbol{X}_k^{(\ell)}\boldsymbol{W}]\!]\|$
    for $i = 0 \ldots k$ **do**
      $\boldsymbol{X}_i^{(\ell+1)} = \boldsymbol{X}_i^{(\ell)}\boldsymbol{W}^{(\ell)}$

▷ Can be approximated via NMF method on each mode with averaging of learned $\boldsymbol{W}$ matrix across modes.

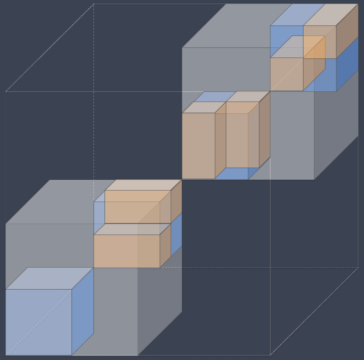
» Training Process

procedure $\textsc{Multi-HNTF}(\mathcal{X})$
    $\{\boldsymbol{X}_i^{(0)}\}_{i=1}^k \leftarrow \mathrm{NCPD}(\mathcal{X}, r_0)$
    for $\ell = 0 \ldots \mathcal{L}$ **do**
        $\boldsymbol{W}^{(\ell)} \leftarrow \operatorname{argmin}_{\boldsymbol{W} \in \mathbb{R}_+^{r_\ell \times r_{\ell+1}}} \|\mathcal{X} - [\![\boldsymbol{X}_1^{(\ell)} \boldsymbol{W}, \ldots, \boldsymbol{X}_k^{(\ell)} \boldsymbol{W}]\!]\|$
        for $i = 0 \ldots k$ **do**
            $\boldsymbol{X}_i^{(\ell+1)} = \boldsymbol{X}_i^{(\ell)} \boldsymbol{W}^{(\ell)}$

- Can be approximated via NMF method on each mode with averaging of learned $\boldsymbol{W}$ matrix across modes.
- Could/should also be trained in a neural network framework.

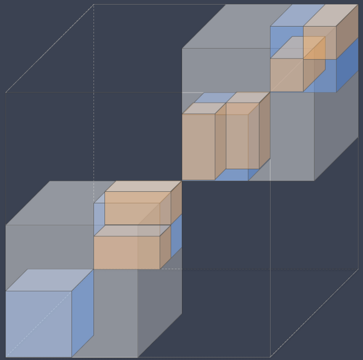
Experiments

» Synthetic Tensor



The table lists relative reconstruction errors on the tensor on the left for models learned with 7-4-2 topic structure. Below, we visualize the Multi-HNTF learned approximations for a synthetic tensor with 7-3 topic structure.

» Synthetic Tensor

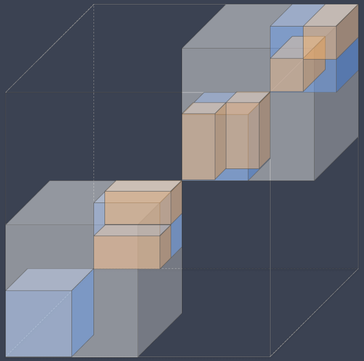


The table lists relative reconstruction errors on the tensor on the left for models learned with 7-4-2 topic structure. Below, we visualize the Multi-HNTF learned approximations for a synthetic tensor with 7-3 topic structure.

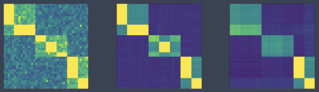
Relative reconstruction error.

Method	$r_0 = 7$	$r_1 = 4$	$r_2 = 2$
Multi-HNTF	0.454	0.548	0.721
Standard HNCPD [Vendrow, et. al.]	0.454	0.612	0.892
HNTF-1 [Cichocki, et. al.]	0.454	0.576	0.781
HNTF-2 [Cichocki, et. al.]	0.454	0.587	0.765
HNTF-3 [Cichocki, et. al.]	0.454	0.560	0.747

» Synthetic Tensor



The table lists relative reconstruction errors on the tensor on the left for models learned with 7-4-2 topic structure. Below, we visualize the Multi-HNTF learned approximations for a synthetic tensor with 7-3 topic structure.



Projections of tensor approximation at each layer of Multi-HNTF.

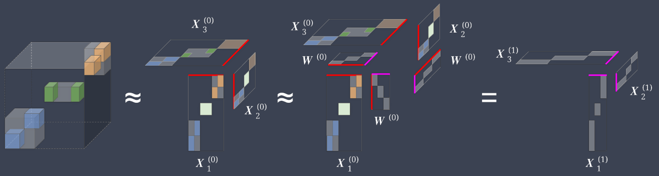
Relative reconstruction error.

Method	$r_0 = 7$	$r_1 = 4$	$r_2 = 2$
Multi-HNTF	0.454	0.548	0.721
Standard HNCPD [Vendrow, et. al.]	0.454	0.612	0.892
HNTF-1 [Cichocki, et. al.]	0.454	0.576	0.781
HNTF-2 [Cichocki, et. al.]	0.454	0.587	0.765
HNTF-3 [Cichocki, et. al.]	0.454	0.560	0.747
Neural HNCPD [Vendrow, et. al.]	0.454	**0.508**	**0.714**

Conclusions

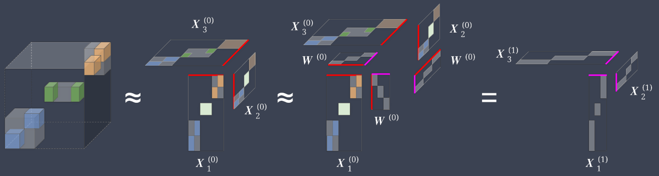
Conclusions

- Multi-HNTF is a hierarchical tensor decomposition model that generalizes hierarchical NMF.



» Conclusions

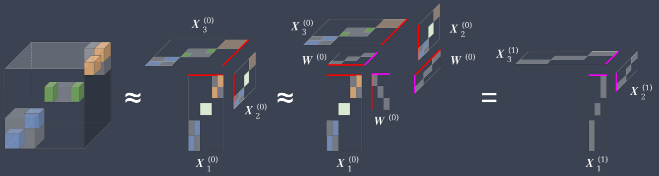
- Multi-HNTF is a hierarchical tensor decomposition model that `generalizes` hierarchical NMF.



- Model can be trained by your `favorite NMF method` with an additional projection step.

» Conclusions

- Multi-HNTF is a hierarchical tensor decomposition model that `generalizes` hierarchical NMF.



- Model can be trained by your `favorite NMF method` with an additional projection step.

- Develop backpropagation framework for Multi-HNTF and first layer NCPD.

» Thanks for listening! Questions?

[1] M. Gao, J. Haddock, D. Molitor, D. Needell, E. Sadovnik, T. Will, and R. Zhang. Neural nonnegative matrix factorization for hierarchical multilayer topic modeling. In Proc. Interational Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2019.

[2] J. Vendrow, J. Haddock, and D. Needell. Neural nonnegative CP decomposition for hierarchical tensor analysis. In Asilomar Conf. on Signals, Systems, Computers (ACSSC), 2021.

[3] J. Vendrow, J. Haddock, and D. Needell. A generalized hierarchical nonnegative tensor decomposition. In Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), 2022.

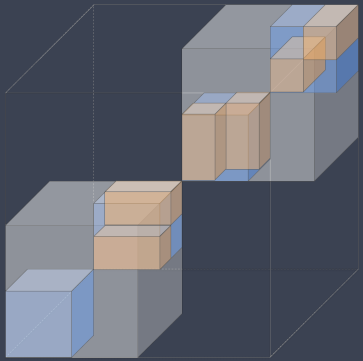
[4] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755):788–791, 1999.

[5] J Douglas Carroll and Jih-Jie Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika, 35(3):283–319, 1970.

[6] Richard A Harshman et al. Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multimodal factor analysis. 1970.

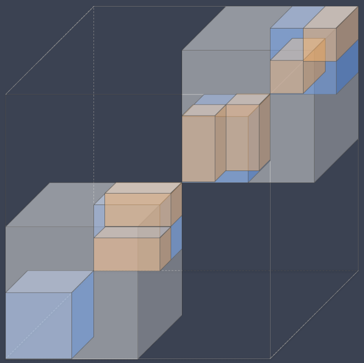
[7] Andrzej Cichocki and Rafal Zdunek. Multilayer nonnegative matrix factorisation. Electronics Letters, 42(16):947–948, 2006.

» Synthetic Tensor



The table lists relative reconstruction errors on the tensor on the left for models learned with 7-4-2 topic structure. Below, we visualize the Multi-HNTF learned approximations for a synthetic tensor with 7-3 topic structure.

» Synthetic Tensor

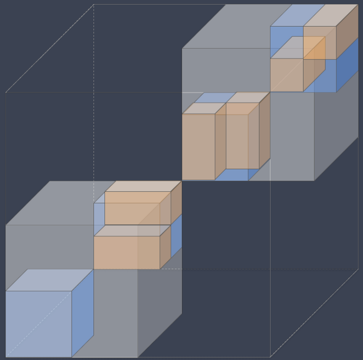


The table lists relative reconstruction errors on the tensor on the left for models learned with 7-4-2 topic structure. Below, we visualize the Multi-HNTF learned approximations for a synthetic tensor with 7-3 topic structure.

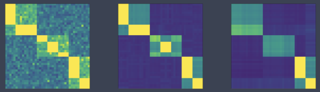
Relative reconstruction error.

Method	$r_0 = 7$	$r_1 = 4$	$r_2 = 2$
Multi-HNTF	0.454	0.548	0.721
Neural HNCPD [Vendrow, et. al.]	0.454	**0.508**	**0.714**
Standard HNCPD [Vendrow, et. al.]	0.454	0.612	0.892
HNTF-1 [Cichocki, et. al.]	0.454	0.576	0.781
HNTF-2 [Cichocki, et. al.]	0.454	0.587	0.765
HNTF-3 [Cichocki, et. al.]	0.454	0.560	0.747

» Synthetic Tensor



The table lists relative reconstruction errors on the tensor on the left for models learned with 7-4-2 topic structure. Below, we visualize the Multi-HNTF learned approximations for a synthetic tensor with 7-3 topic structure.



Projections of tensor approximation at each layer of Multi-HNTF.

Relative reconstruction error.

Method	$r_0 = 7$	$r_1 = 4$	$r_2 = 2$
Multi-HNTF	0.454	0.548	0.721
Neural HNCPD [Vendrow, et. al.]	0.454	**0.508**	**0.714**
Standard HNCPD [Vendrow, et. al.]	0.454	0.612	0.892
HNTF-1 [Cichocki, et. al.]	0.454	0.576	0.781
HNTF-2 [Cichocki, et. al.]	0.454	0.587	0.765
HNTF-3 [Cichocki, et. al.]	0.454	0.560	0.747