Alternating Minimization Algorithms for Graph Regularized Tensor Completion

Yu Guan
Shuyu Dong, Pierre-Antoine Absil, François Glineur

INMA, UCLouvain

Oct 1, 2019

UCLouvain icteam

What is a tensor

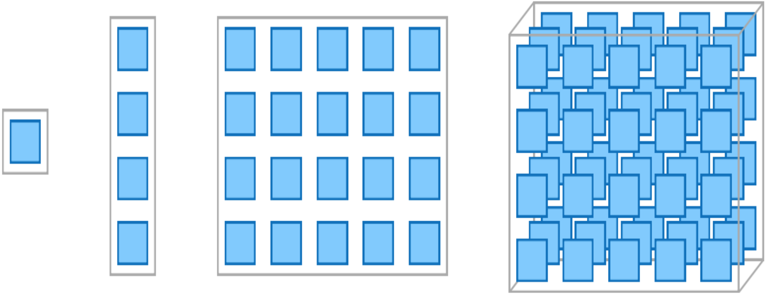


Figure : $x \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^4, X \in \mathbb{R}^{4\times5}, \mathcal{X} \in \mathbb{R}^{4\times5\times3}$

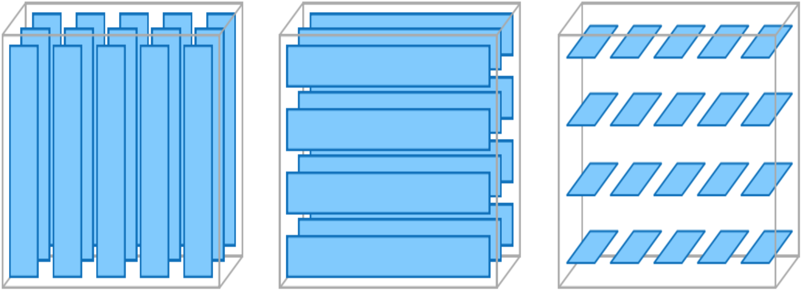


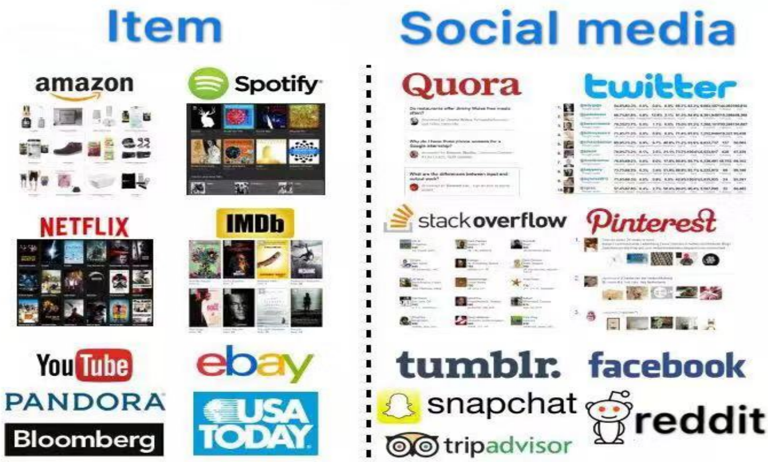
Figure : Column, row, and tube fibers of a order-3 tensor

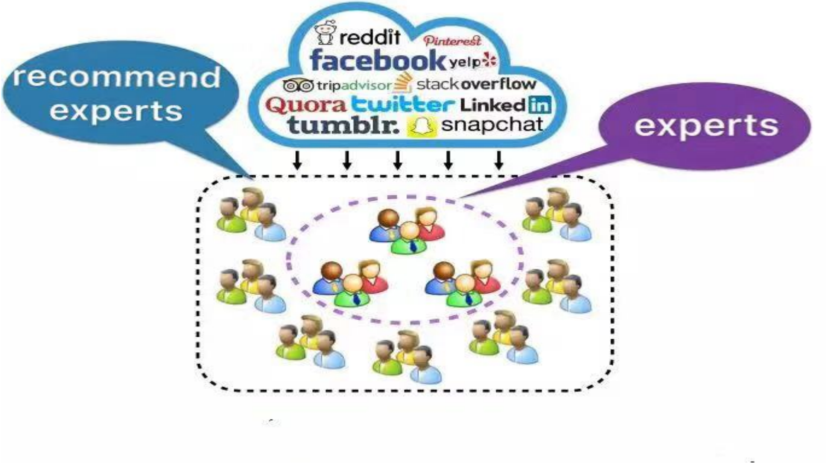
Multiway Data

- Psychometrics: individual $\times$ variable $\times$ time
- Time-series analysis: time $\times$ variable $\times$ lag
- Neuroscience: electrodes $\times$ time $\times$ frequency
- Social networks: users $\times$ keywords $\times$ time
- Facial image: people $\times$ view $\times$ illumination $\times$ expression $\times$ pixels
- Atmospheric science: location $\times$ variable $\times$ time $\times$ observation

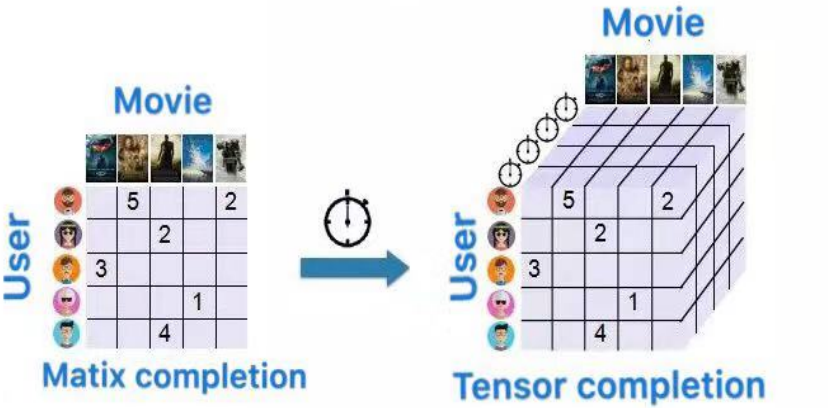
Tensor Completion

Recommender System



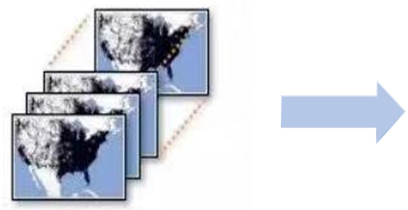

recommend experts
reddit
Pinterest
facebook
yelp
tripadvisor
stackoverflow
Quora
twitter
Linkedin
tumblr.
snapchat
experts

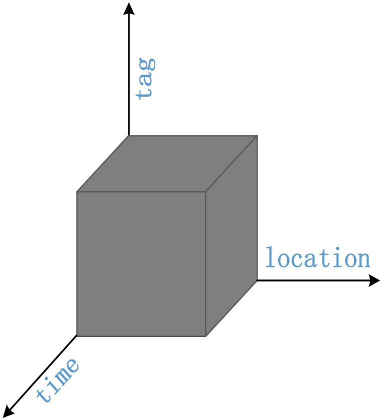
Movie Rating



Spatio-Temporal Data

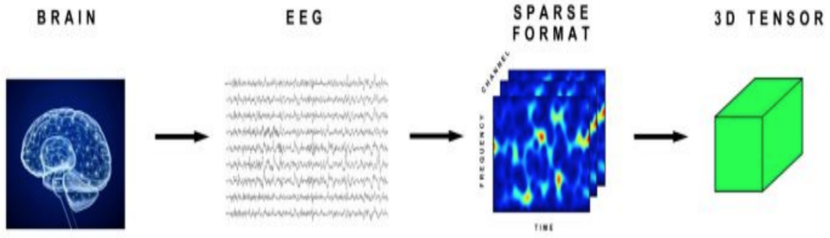
Hashtags





Other Examples

- Image Inpainting
- Video Inpainting
- Link Prediction
- EEG Data
- ...



Why some data is missing

- API restriction
- Error occurs when collecting the data
- Access restriction
- Sampling method
- Some of the data does not exist

Kronecker Product.

The Kronecker product of vectors $\mathbf{u} = [u_r] \in \mathbb{R}^{I_1}$ and $\mathbf{v} = [v_r] \in \mathbb{R}^{I_2}$ results in a vector $\mathbf{u} \otimes \mathbf{v} \in \mathbb{R}^{I_1 I_2}$ defined as

$$\mathbf{u} \otimes \mathbf{v} = \begin{bmatrix} u_1 \mathbf{v} \\ u_2 \mathbf{v} \\ \vdots \\ u_{I_1} \mathbf{v} \end{bmatrix}.$$

Khatri-Rao Product.

The Khatri-Rao product $U \odot V$ of two matrices $U = [u_{\ell,r}] \in \mathbb{R}^{I_1 \times R}$ and $V = [v_{\ell,r}] \in \mathbb{R}^{I_2 \times R}$ is

$$U \odot V = [u_{:,1} \otimes v_{:,1}, \ldots, u_{:,R} \otimes v_{:,R}]$$

Tensor matricization.

Let $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_k}$ is a k-th order tensor, the unfolded matrix is

$$\mathcal{T}_{(\ell)} \in \mathbb{R}^{I_\ell \times (I_1 \ldots I_{\ell-1} I_{\ell+1} \ldots I_k)}$$

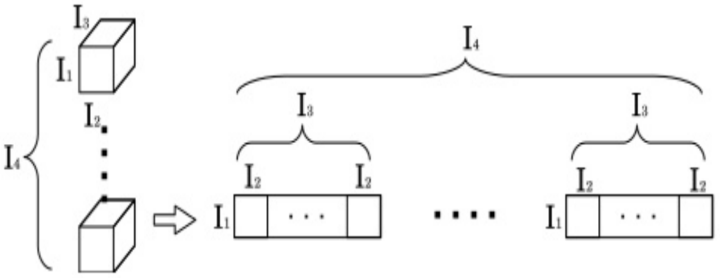
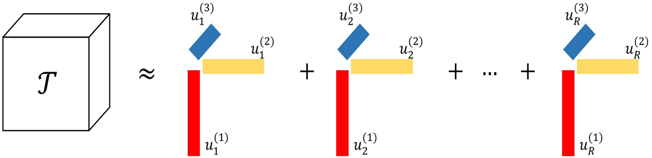


Figure : Unfolding Image

CP Decomposition



The CANDECOMP/PARAFAC(CP) model represents a tensor as a sum of rank-one tensors

$$T = \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \circ \cdots \circ \mathbf{u}_r^{(k)} = [\![U^{(1)}, \ldots, U^{(k)}]\!]$$

Low Rank Tensor Completion (LRTC)

For a given tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times \ldots \times I_k}$, we want to find a low rank tensor $\mathcal{Z}$ by solving the following optimization problem

$$\min_{\mathcal{Z}} \quad \frac{1}{2}\|\mathcal{P}_\Omega(\mathcal{T} - \mathcal{Z})\|_F^2 + \sum_{i=1}^{k} \lambda_i \|\mathcal{Z}_{(i)}\|_*$$

where $\mathcal{P}_\Omega$ is the projection operator that only retains those entries of the tensor that lie in the set Ω.

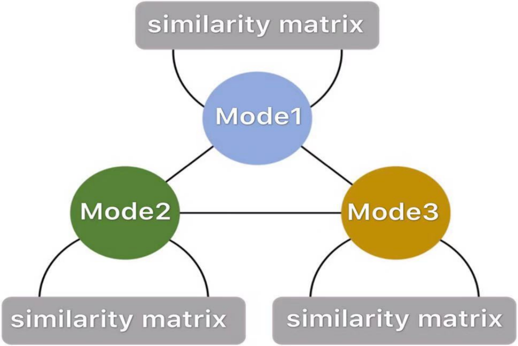
Low Rank Tensor Completion (LRTC)

For a given tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times \cdots \times I_k}$, we want to find a low rank tensor $\mathcal{Z}$ by solving the following optimization problem

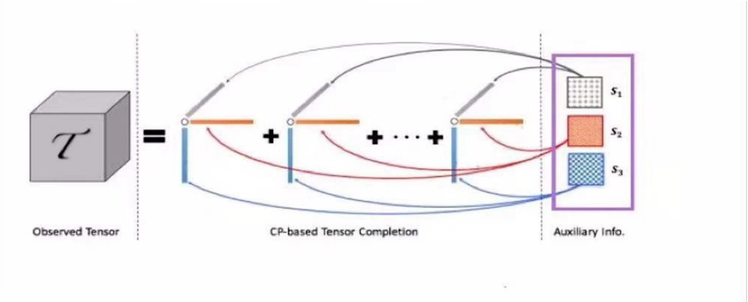
$$\min_{\mathcal{Z}} \quad \frac{1}{2}\|\mathcal{P}_\Omega(\mathcal{T} - \mathcal{Z})\|_F^2 + \sum_{i=1}^{k} \frac{\lambda_i}{2}\{\|U^{(i)}\|_F^2 + \|(U^{(j)})^{\odot_{j \neq i}}\|_F^2\}$$

where $\mathcal{P}_\Omega$ is the projection operator that only retains those entries of the tensor that lie in the set Ω.

Auxiliary Similarity



Our LRTC Model



$$\begin{aligned}\min_{U^{(1)},\dots,U^{(k)}} f(U^{(1)},\dots,U^{(k)}) =& \frac{1}{2}\|\mathcal{P}_\Omega(\mathcal{T} - [\![U^{(1)},\dots,U^{(k)}]\!])\|_F^2 \\ &+ \sum_{i=1}^{k} \frac{\lambda_i}{2}\|U^{(i)}\|_F^2 + \sum_{i=1}^{k} \frac{\lambda_i}{2}\|(U^{(j)})^{\odot_{j\neq i}}\|_F^2 \\ &+ \boxed{\sum_{i=1}^{k} \frac{\lambda_L}{2}\left\langle U^{(i)}U^{(i)\top}, \mathbf{Lap}(U^{(i)})\right\rangle}.\end{aligned}$$

Algorithm 1 Altmin for solving LRTC

Initialization: $U_0^{(1)}, \ldots, U_0^{(k)}$

1: **for** $p = 0, \ldots, k$ **do**
2: **for** $i = 1 \ldots, k$ **do**
3: $U_{p+1}^{(i)} = argminf(U_{p+1}^{(1)}, \ldots, U_{p+1}^{(i-1)}, U^{(i)}, U_p^{(i+1)}, \ldots, U_p^{(k)})$
4: **end for**
5: **end for**

Compared with other algorithms, alternating minimization has several advantages for LRTC problem:

- It is easy to implement as there is no need to tune optimization parameters like step sizes.
- It converges very fast in practice.
- The objective function f is not convex with respect to variables $U^{(1)}, \ldots, U^{(k)}$ together while the subproblem is easy to solve as it has a closed-form solution.
- Obviously, the objective function f is monotone decreasing.

Theorem

The iterates $\{U_p^{(1)}, \ldots, U_p^{(k)}\}$ *generated by Algorithm 1 from any initialization converges globally to a critical point of* f.

Proof.

- f is a KŁ function with $\theta \in [1/2, 1)$
- Each subproblem is strongly convex.
- Notice f is coercive and real analytic, it is guaranteed to produce a bounded sequence.
- f is a $\mathbf{C}^\infty$ function, thus ∇f is Lipschitz continuous on any bounded subset of domain.

□

Numerical Example

- Model advantage
  - Our model
  - Model without graph regularizer
  - Model without any regularizer
- Effectiveness of the proposed methods

- CPU time
- Reconstruction error (RE)$= \frac{\|\mathcal{T} - [\![U^{(1)}, U^{(2)}, U^{(3)}]\!]\|_F}{\|\mathcal{T}\|_F}$, where $\mathcal{T}$ is the ground truth.
- Root mean squared error (RMSE)$= \frac{\|\mathcal{P}_{\Omega^c}(\mathcal{T} - [\![U^{(1)}, U^{(2)}, U^{(3)}]\!])\|_F}{|\mathcal{P}_{\Omega^c}|^{1/2}}$, where $\mathcal{P}_{\Omega^c}$ denotes the unobserved projection and $|\mathcal{P}_{\Omega^c}|$ is the number in $\mathcal{P}_{\Omega^c}$ set.

Synthetic Data

- Randomly generate $U^{(i)} \in \mathbb{R}^{100\times 10}$
- generate the graph Laplacian matrix $\mathbf{Lap}(U^{(1)}) \in \mathbb{R}^{100\times 100}$ of $U^{(1)}$
- $\mathbf{Lap}(U^{(1)}) = V\Lambda V^\top$ by SVD

$\mathcal{T}$ can be generated as follows

$$\mathcal{T}_{(1)} = V\Lambda.^{-2} U^{(1)} (U^3 \odot U^{(2)})^\top,$$

	SR=1%			SR=5%		
Method	GM	WGM	NM	GM	WGM	NM
Altmin-CG	0.2342	0.4395	0.4006*10	$0.6611*10^{-3}$	$0.6356*10^{-3}$	$0.1098*10^{-1}$
Altmin-ADMM	0.2484	0.3644	$0.1492*10^{2}$	$0.6553*10^{-3}$	$0.6375*10^{-3}$	$0.1095*10^{-1}$

TABLE 6.3

RE of the three models: GM 1 (with graph Laplacian), WGM (without graph Laplacian), NM (no regularizer).

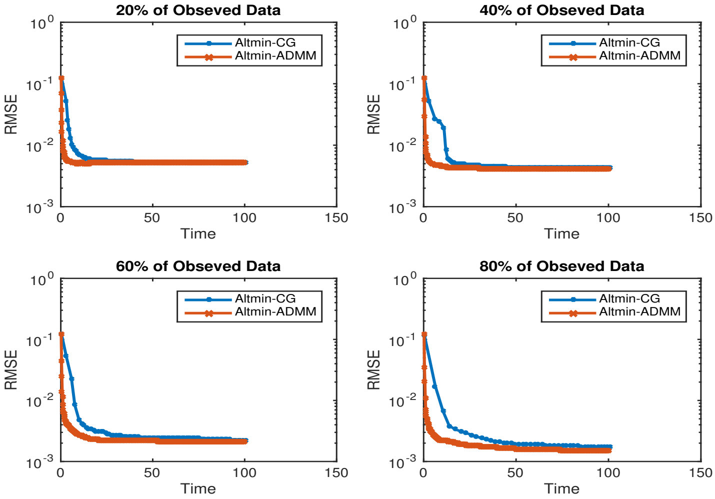
Flow Injection Analysis

- Rank-deficient spectral FIA dataset
- The represented tensor is of size $12(\textit{substances}) \times 100(\textit{wavelengths}) \times 89(\textit{reactiontimes})$.
- For 12 chemical substances, we build the similarity between two substances as the inverse of Euclidean distance between their feature vectors.
- For wavelengths and reaction times, the similar matrices are tridiagonal.

	SR=1%			SR=5%			SR=10%		
Method	GM	WGM	NM	GM	WGM	NM	GM	WGM	NM
Alt-CG	0.6469	0.8795	8.2624	0.0220	0.0326	3.9585	0.0135	0.0166	0.6746
Alt-ADMM	0.5752	0.7678	1.5776	0.0113	0.0113	0.1107	0.0140	0.0171	0.0380

TABLE 6.5

RE of the three models: GM 1 (with graph Laplacian), WGM (without graph Laplacian), NM (no regularizer).


20% of Obseved Data
40% of Obseved Data
60% of Obseved Data
80% of Obseved Data
Altmin-CG
Altmin-ADMM
RMSE
Time
10^{0}
10^{-1}
10^{-2}
10^{-3}
0
50
100
150

Conclusion

- Low-rank tensor completion problem (LRTC) is to find a tensor with the minimum rank, which subjects to the equality constraints given by the observations.
- Graph Laplacian regularizer can help to improve the recovery quality when the missing ratio is high.
- Alternating minimization method is efficient to solve the LRTC problem.
- Our model achieve comparable error rates, while being significantly scalable.

Reference

[1] SONG, QINGQUAN AND GE, HANCHENG AND CAVERLEE, JAMES AND HU, XIA, *Tensor completion algorithms in big data analytics*, ACM Transactions on Knowledge Discovery from Data (TKDD), 13(1) (2019).

Questions?

Thank you!