

Discovery of Latent Factors in High-dimensional Data via Spectral Methods

Furong Huang

University of Maryland

Presented by Xuchen You

2018 SIAM Annual Meeting (AN18)

Machine Learning - Excitements

Success of Supervised Learning



Image classification



Speech recognition



Text processing

Key to Success

- Deep composition of nonlinear units
- Enormous labeled data
- Computation power growth

Machine Learning - Modern Challenges

Automated discovery of features and categories?

Real AI requires **Unsupervised Learning**



Filter bank learning



Feature extraction



Embeddings, Topics

- Summarize key features in data
 - ▶ State-of-the-art: Humans are better than machines
 - ▶ Goal: Intelligent machines that summarize key features in data
- Interpretable modeling and learning of the data
 - ▶ Theoretically guaranteed learning
 - ▶ Extracted features are interpretable

Unsupervised Learning with Big Data

Information Extraction

- High dimension **observation** vs Low dimension **representation**



My Solution: A Unified Tensor Decomposition Framework

App 1: Automated Categorization of Documents

≡ SECTIONS ⌂ HOME 🔎 SEARCH

The New York Times

COLLEGE FOOTBALL

At Florida State, Football Clouds Justice

Now, an examination by The New York Times of police and court records, along with interviews with crime witnesses, has found that, far from an aberration, the treatment of the Winston complaint was in keeping with the way the police on numerous occasions have soft-pedaled allegations of wrongdoing by Seminoles football players. From criminal mischief and motor-vehicle theft to domestic violence, arrests have been avoided, investigations have stalled and players have escaped serious consequences.

In a community whose self-image and economic well-being are so tightly bound to the fortunes of the nation's top-ranked college football team, law enforcement officers are finely attuned to a suspect's football connections. Those ties are cited repeatedly in police reports examined by The Times. What's more, dozens of officers work second jobs directing traffic and providing security at home football games, and many express their devotion to the Seminoles on social media.

IMAGE: AP/WIDEWORLD

On Jan. 10, 2013, a female student at Florida State spotted the man she believed had raped her the previous month. After learning his name, Jameis Winston, she reported him to the Tallahassee police.

In the 21 months since, Florida State officials have said little about how they handled the case, which is no investigated by the federal Department of Justice.

Most recently, university officials suspended Mr. Winston for one game after he stood in a public place on campus and, playing off a running Internet gag, shouted a crude reference to a sex act. In a news conference afterward, his coach, Jimbo Fisher, said, "Our hope and belief is Jameis will learn from this and use better judgment and language and decision making."

As The Times reported last April, the Tallahassee police also failed to aggressively investigate the rape accusation. It did not become public until November, when a Tampa reporter, Matt Baker, acting on a tip, sought records of the police investigation.

Upon learning of Mr. Baker's inquiry, Florida State, having shown little curiosity about the rape accusation, suddenly took a keen interest in the journalist seeking to report it, according to emails obtained by The Times.

"Can you share any details on the requesting source?" David Perry, the university's police chief, asked the Tallahassee police. Several hours later, Mr.

Topics

● Education

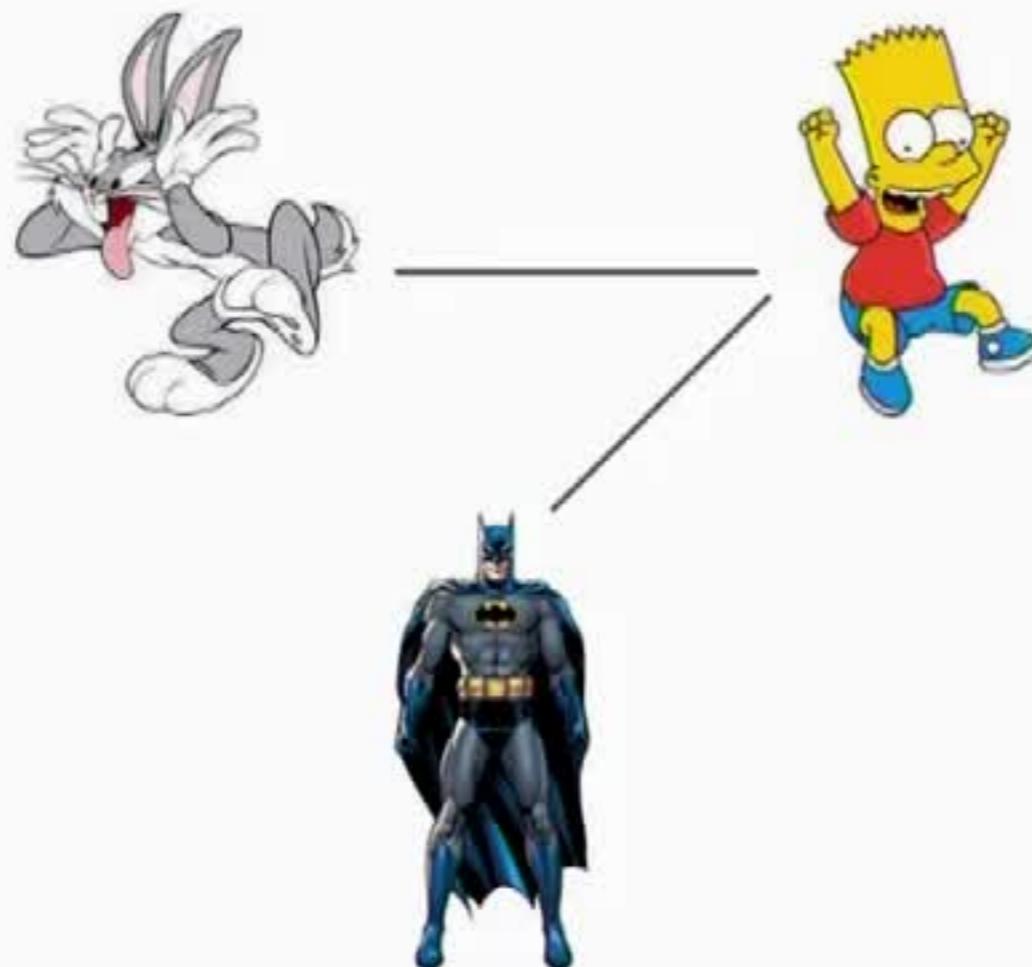
● Crime

● Sports

Document modeling

- Observed: words in document corpus: search logs, emails etc
- Hidden: (mixed) topics: personal interests, professional area etc

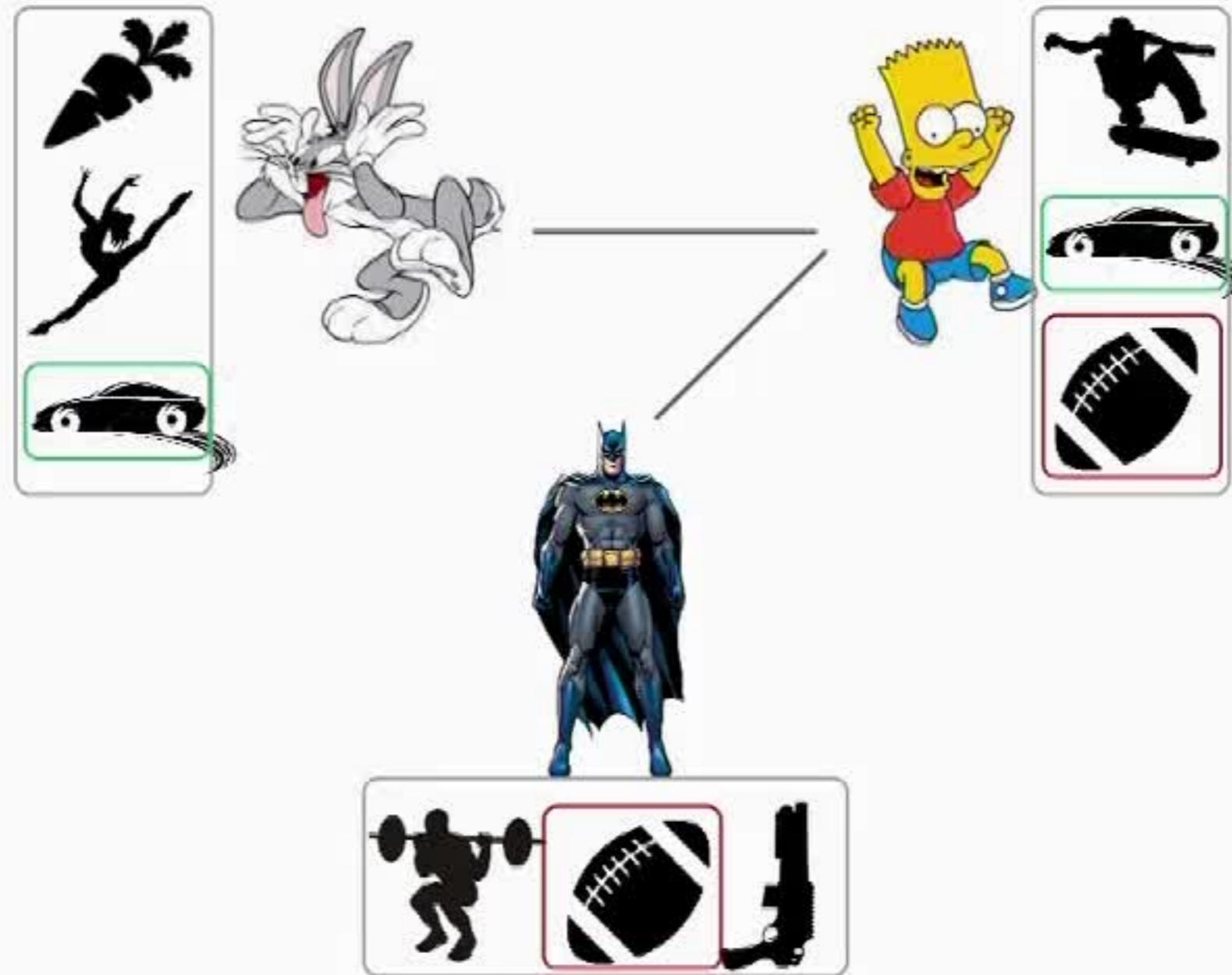
App 2: Community Extraction From Connectivity Graph



Social Networks

- Observed: network of social ties: friendships, transactions etc
- Hidden: (mixed) groups/communities of social actors

App 2: Community Extraction From Connectivity Graph

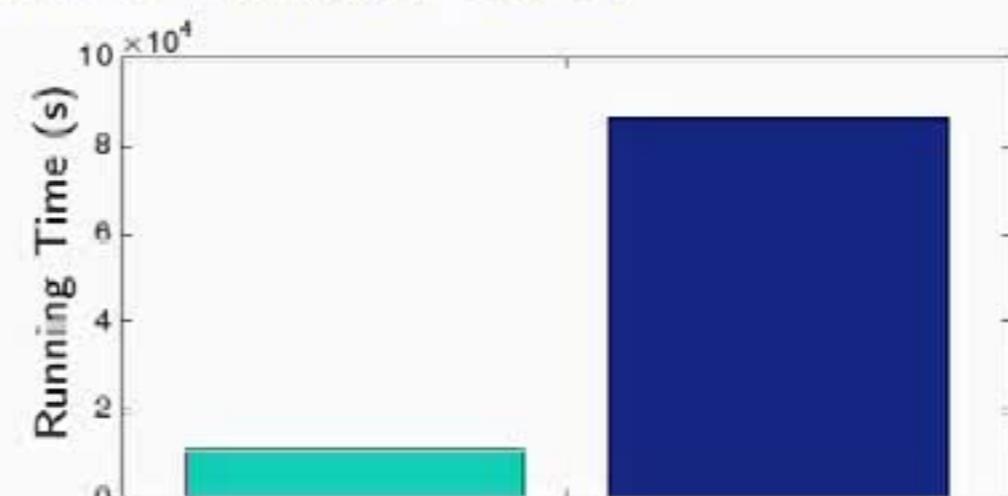
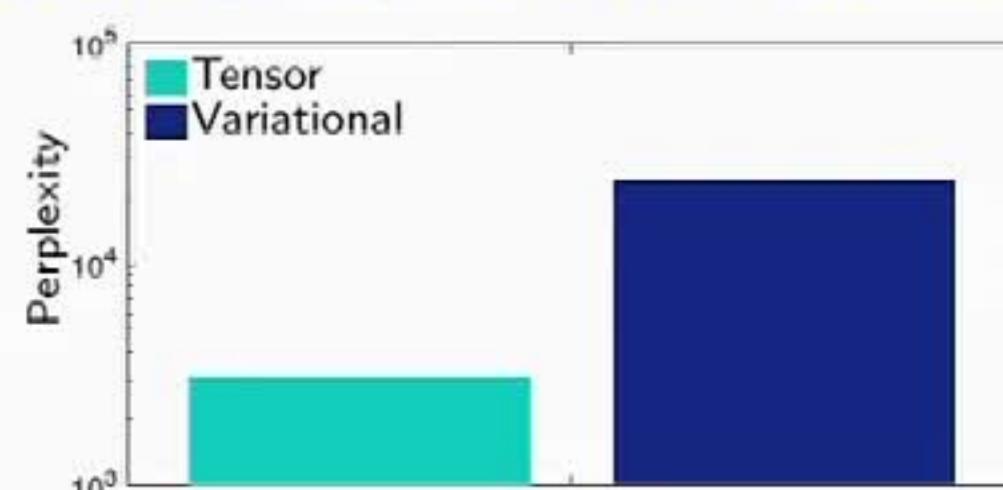


Social Networks

- Observed: network of social ties: friendships, transactions etc
- Hidden: (mixed) groups/communities of social actors

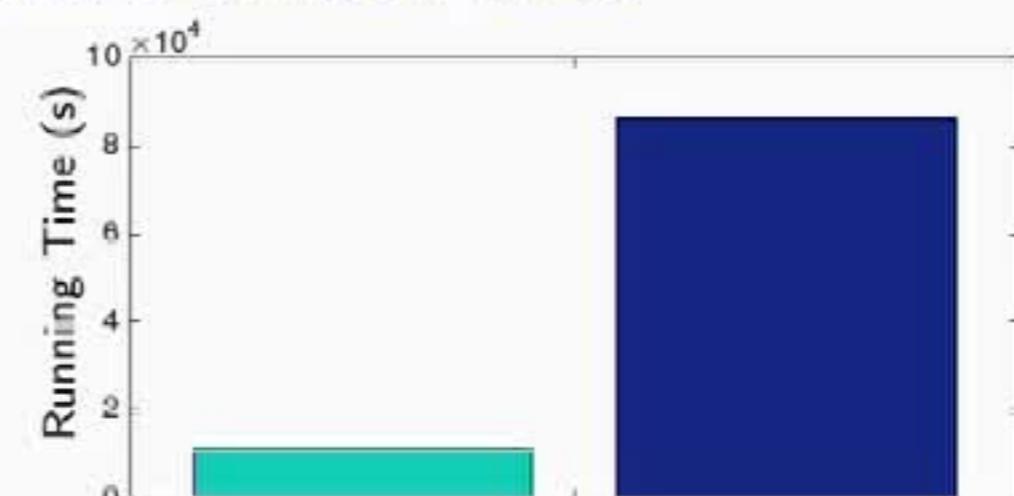
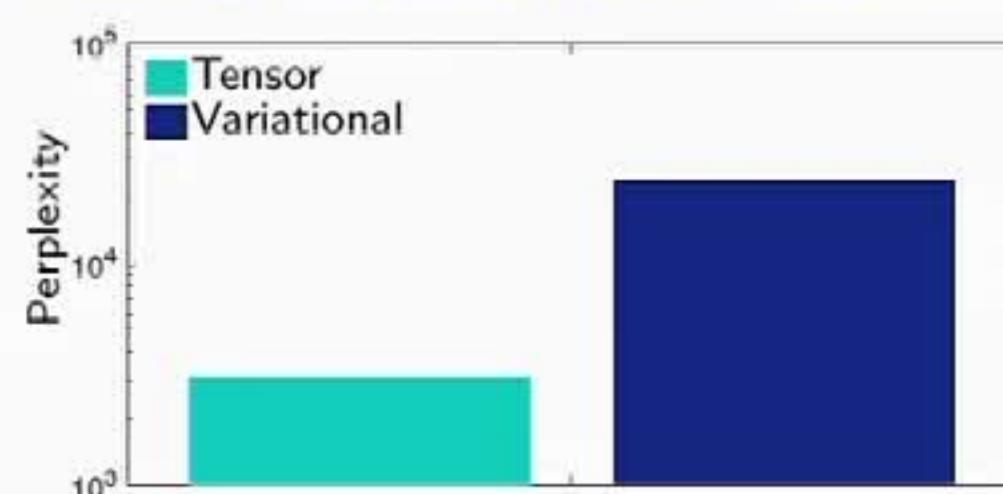
Tensor Methods Compared with Variational Inference

Learning Topics from PubMed on Spark: 8 million docs



Tensor Methods Compared with Variational Inference

Learning Topics from PubMed on Spark: 8 million docs



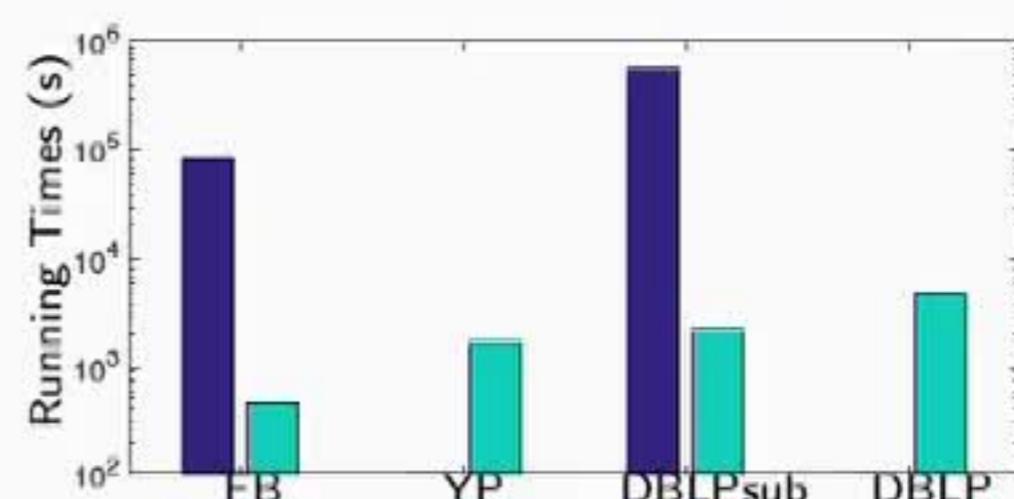
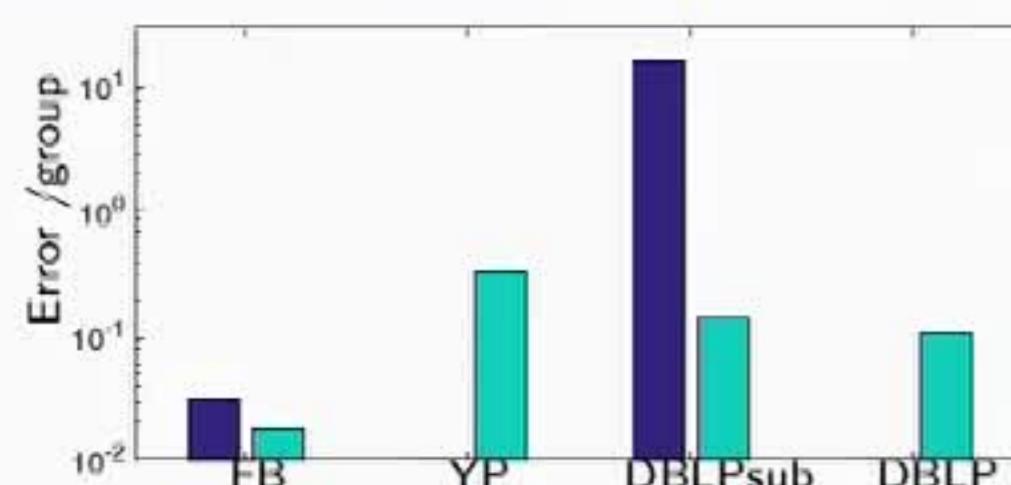
Learning Communities from Graph Connectivity

Facebook: $n \sim 20k$

Yelp: $n \sim 40k$

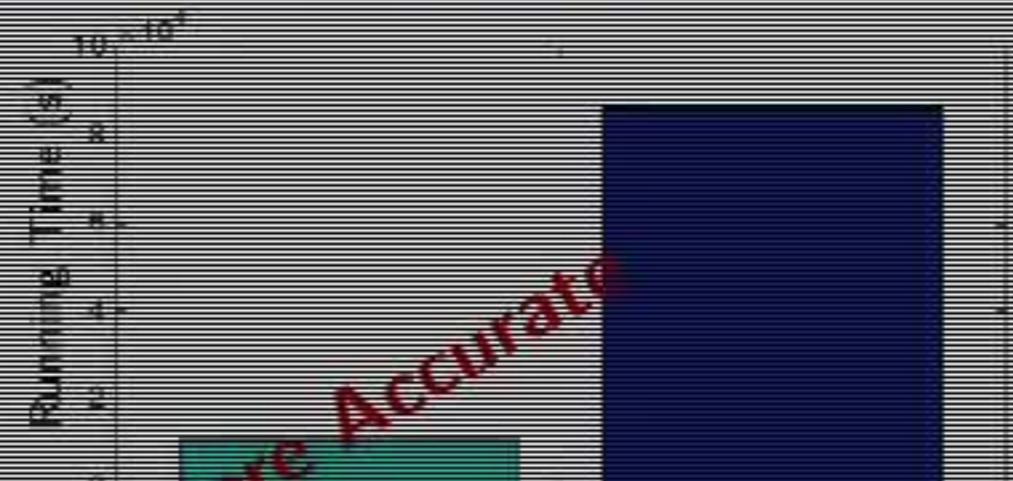
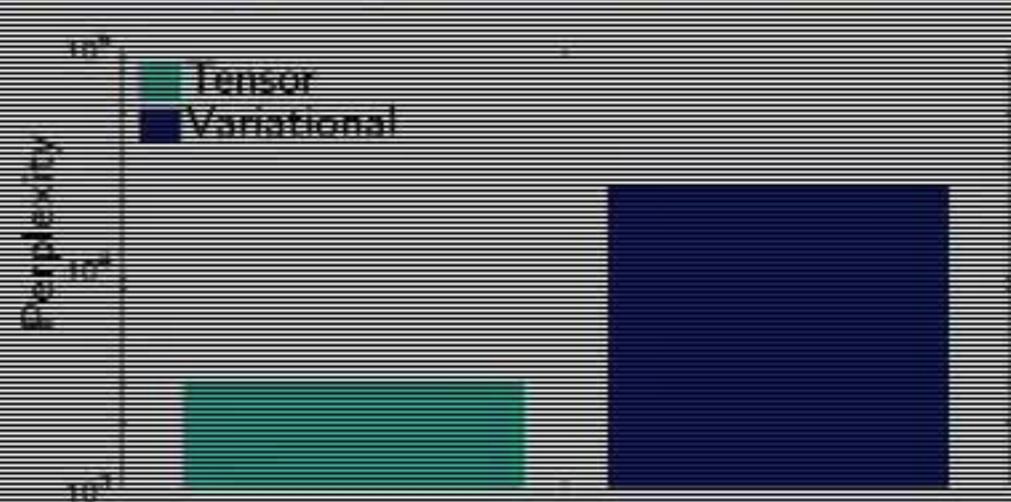
DBLPsub: $n \sim 0.1m$

DBLP: $n \sim 1m$



Tensor Methods Compared with Variational Inference

Learning Topics from PubMed on Spark: 8 million docs



Learning Communities from Graph Connectivity

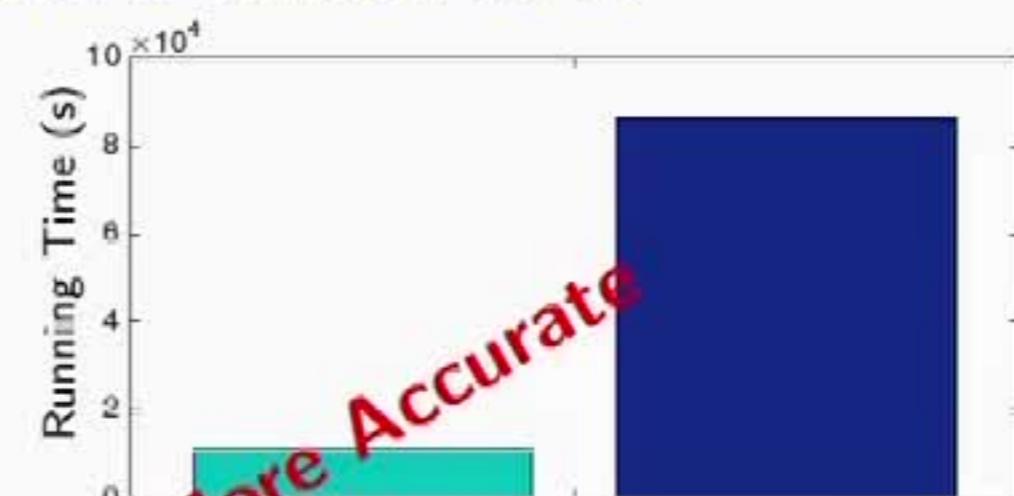
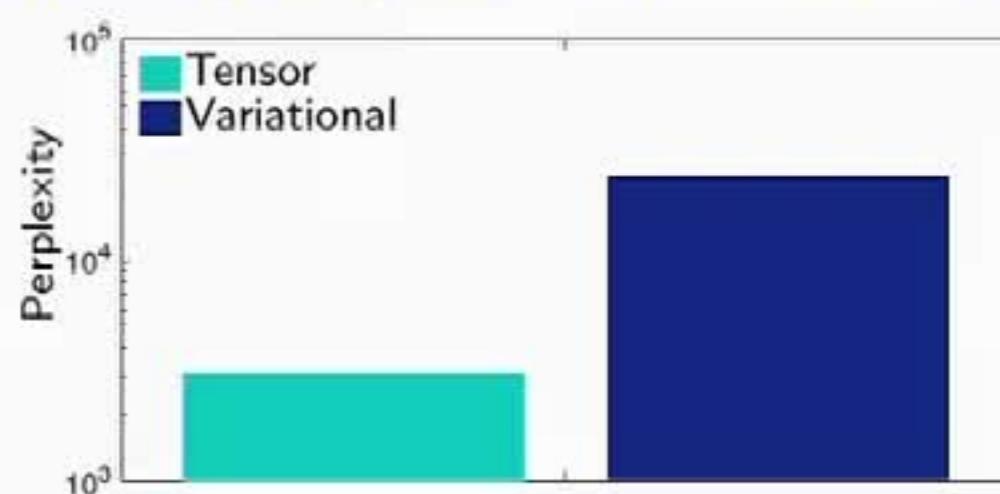
Faster & More Accurate

db: $n \sim 0.1m$

DBLP: $n \sim 1m$

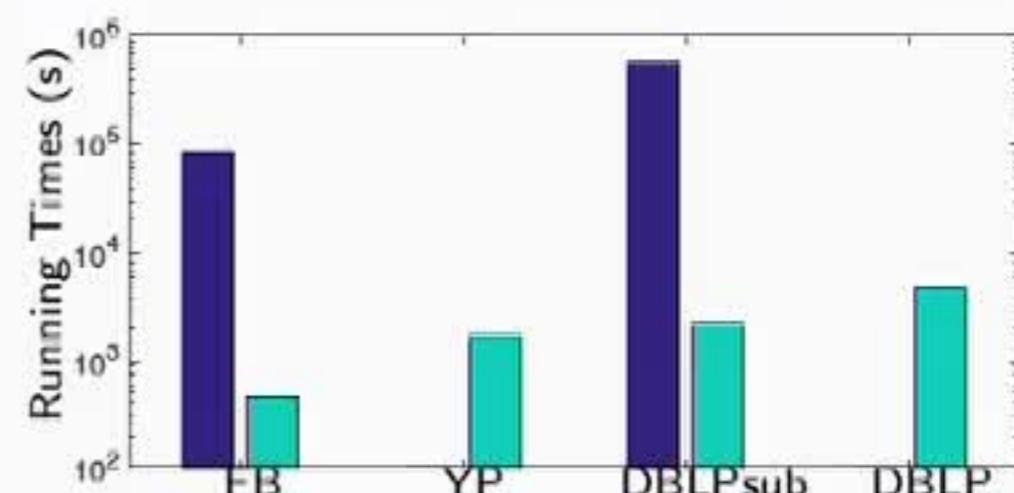
Tensor Methods Compared with Variational Inference

Learning Topics from PubMed on Spark: 8 million docs



Learning Communities from Graph Connectivity

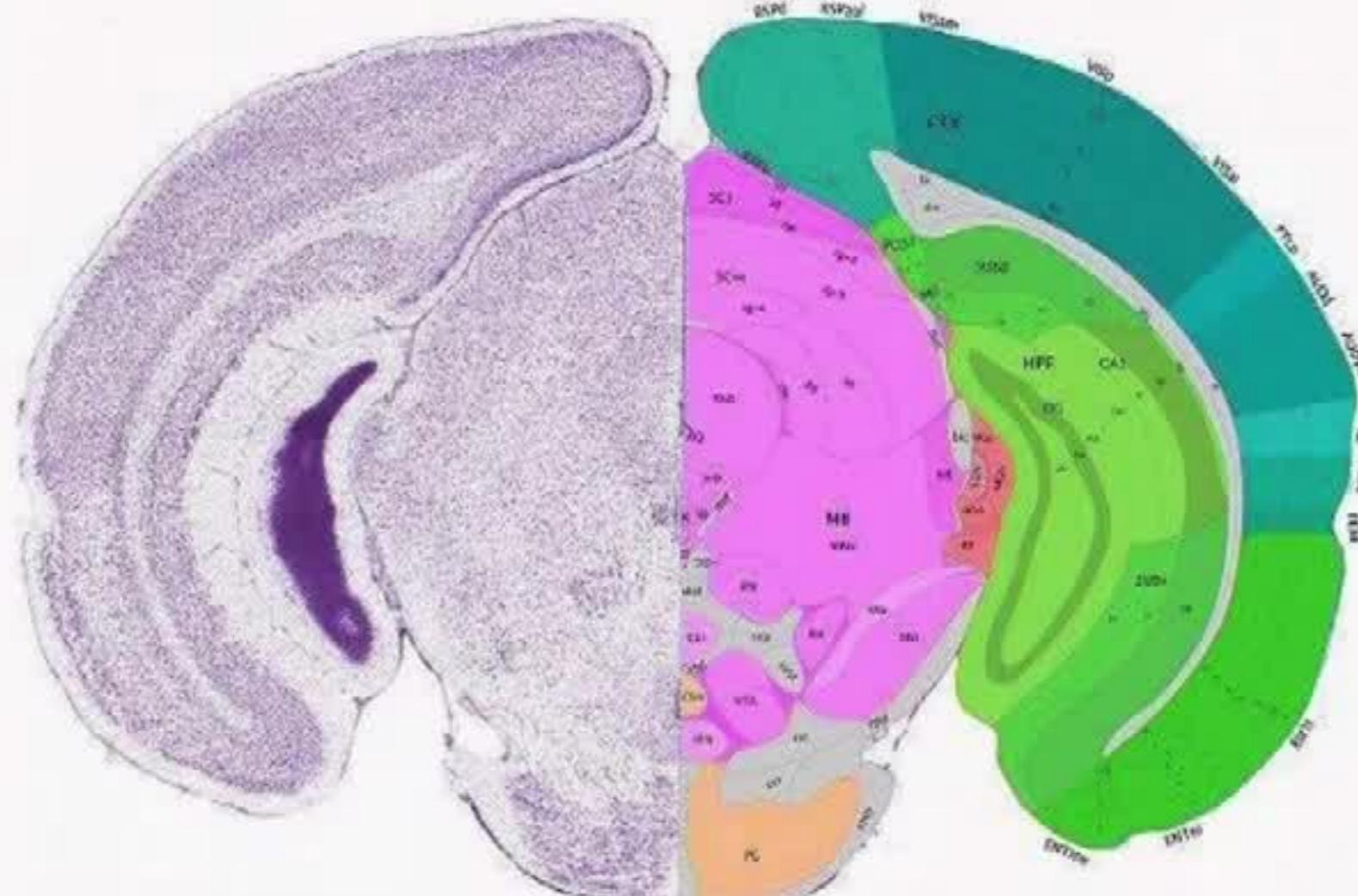
Facebook: $n \sim 20k$ Yelp: $n \sim 40k$ DBLPsub: $n \sim 0.1m$ DBLP: $n \sim 1m$



"Online Tensor Methods for Learning Latent Variable Models", F. Huang, U. Niranjan, M. Hakeem, A. Anandkumar, JMLR14.

"Tensor Methods on Apache Spark", F. Huang, A. Anandkumar, Oct. 2015.

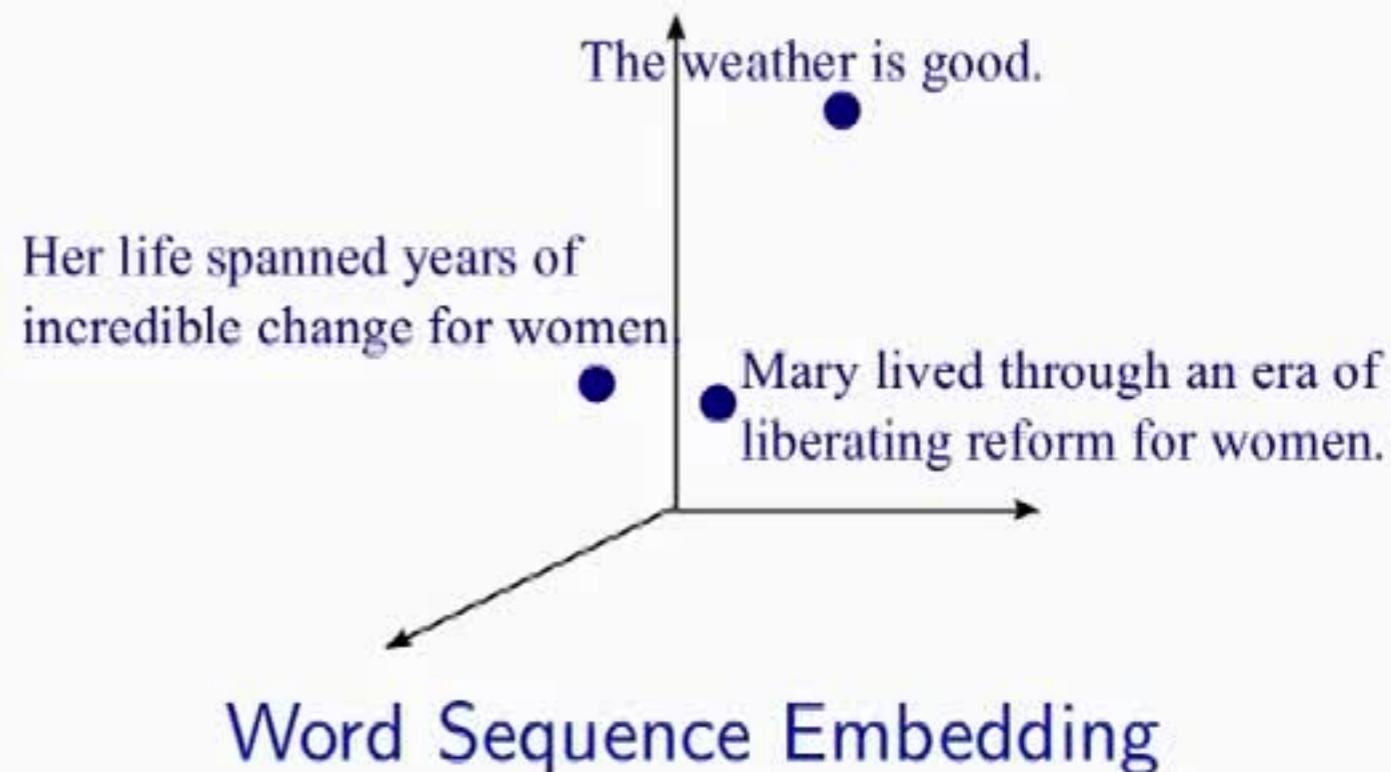
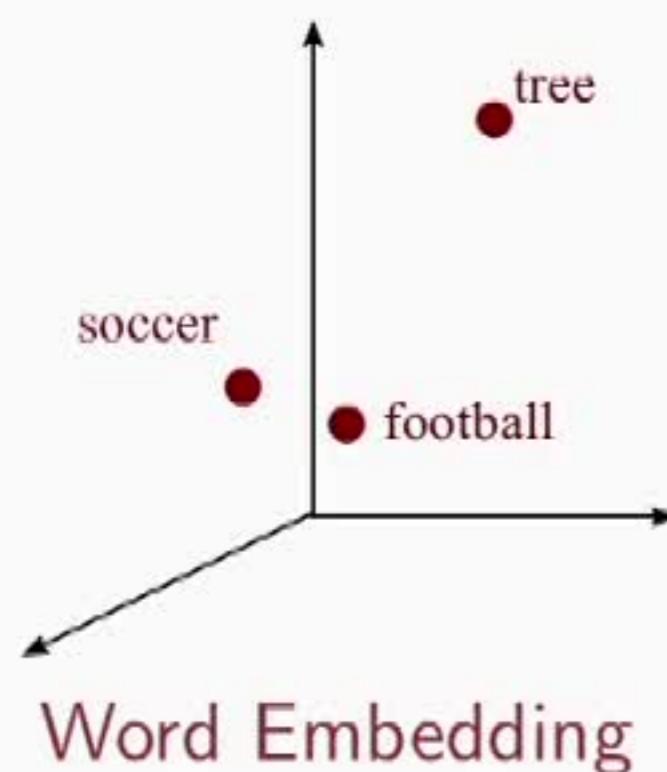
App 3: Cataloging Neuronal Cell Types In the Brain



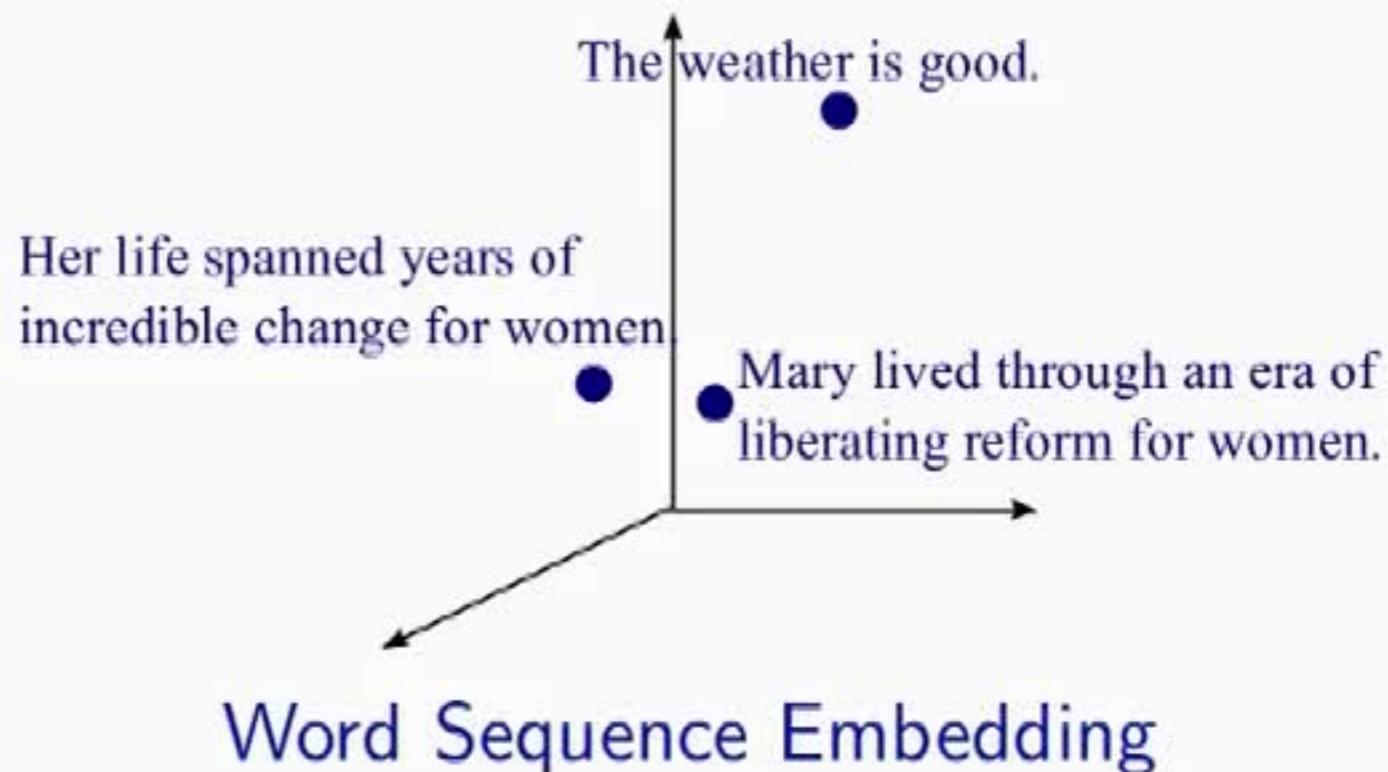
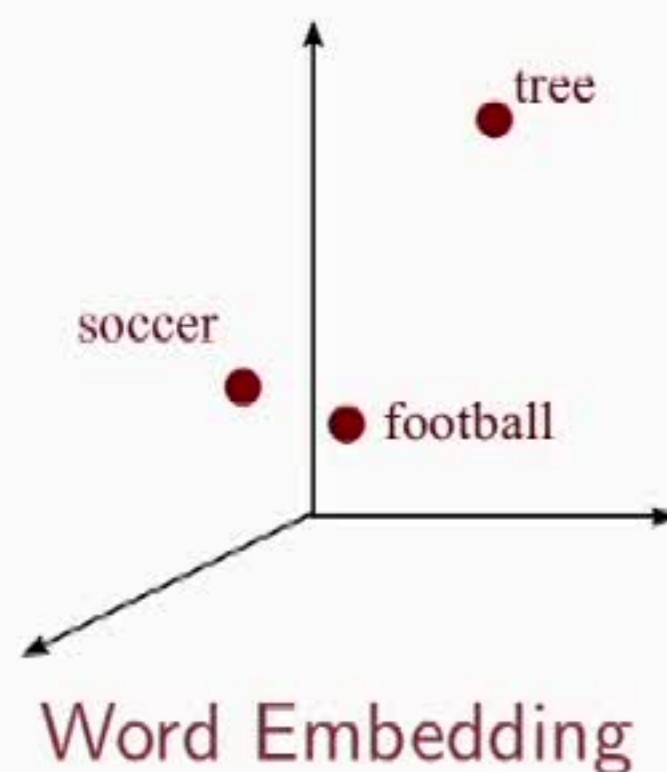
Neuroscience

- Observed: cellular-resolution brain slices
 - Hidden: neuronal cell types

App 4: Word Sequence Embedding Extraction

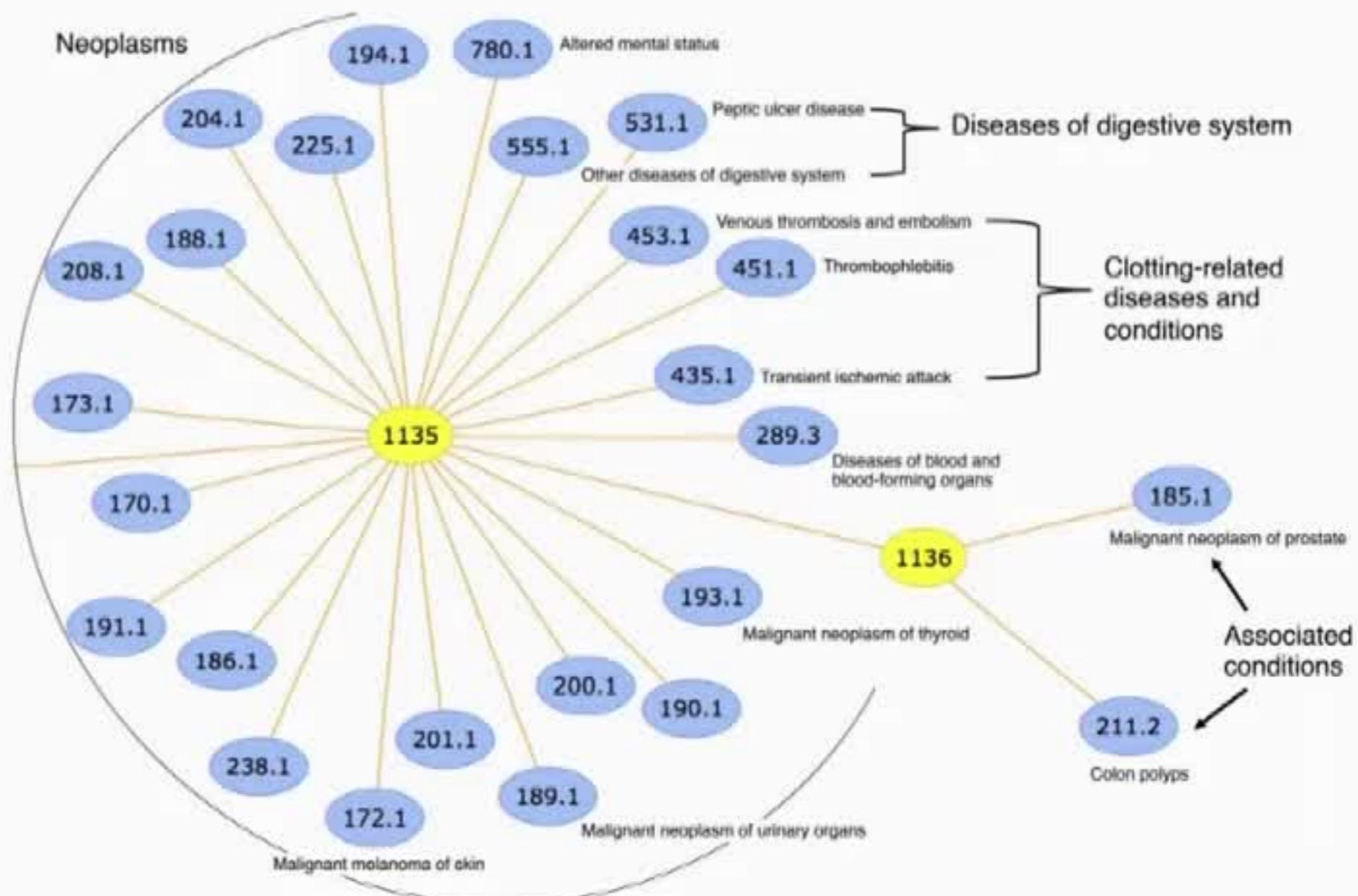


App 4: Word Sequence Embedding Extraction



App 5: Human Disease Hierarchy Discovery

CMS: 1.6 million patients, 168 million diagnostic events, 11 k diseases.



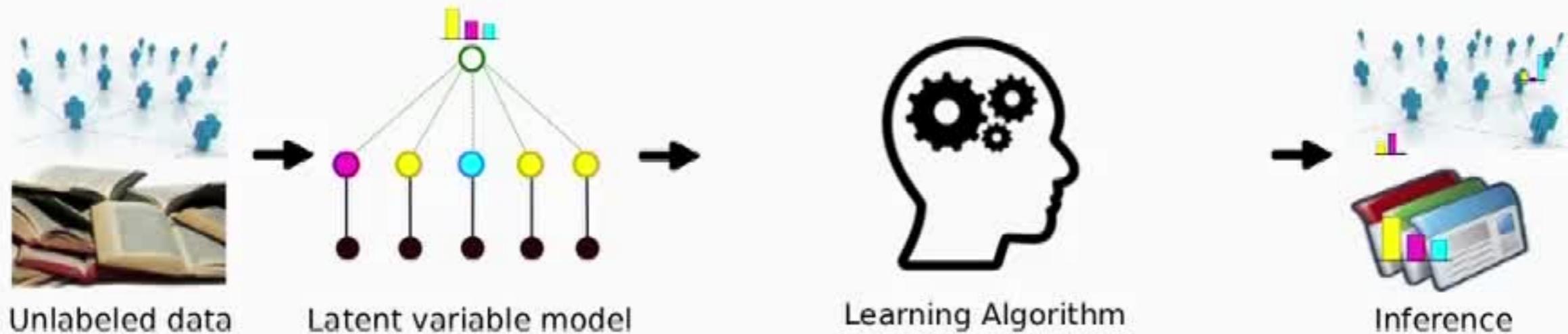
- Observed: co-occurrence of diseases on patients
- Hidden: disease similarity/hierarchy

Involve discovering the **hidden** and **compact** structure
that is embedded in the high-dimensional complex observed data

Challenges in Learning

Basic goal in all mentioned applications

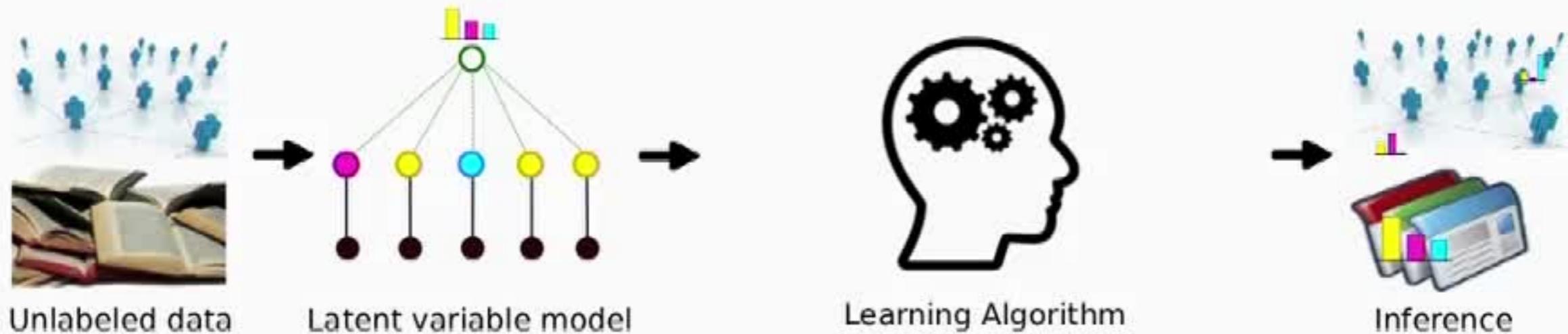
Discover hidden structure in data: **unsupervised** learning.



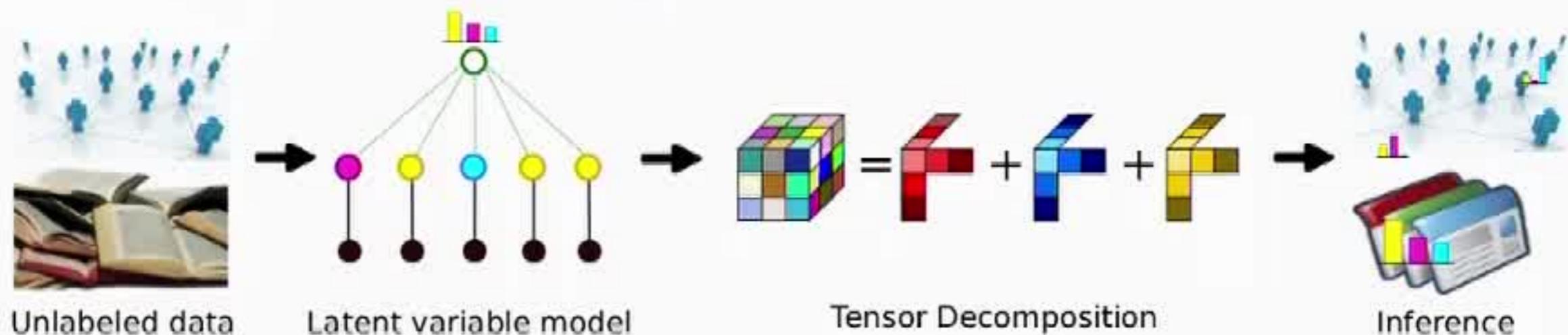
Challenges in Learning

Basic goal in all mentioned applications

Discover hidden structure in data: **unsupervised** learning.



Challenges in Learning – find hidden structure in data



Challenge: Conditions for Identifiability

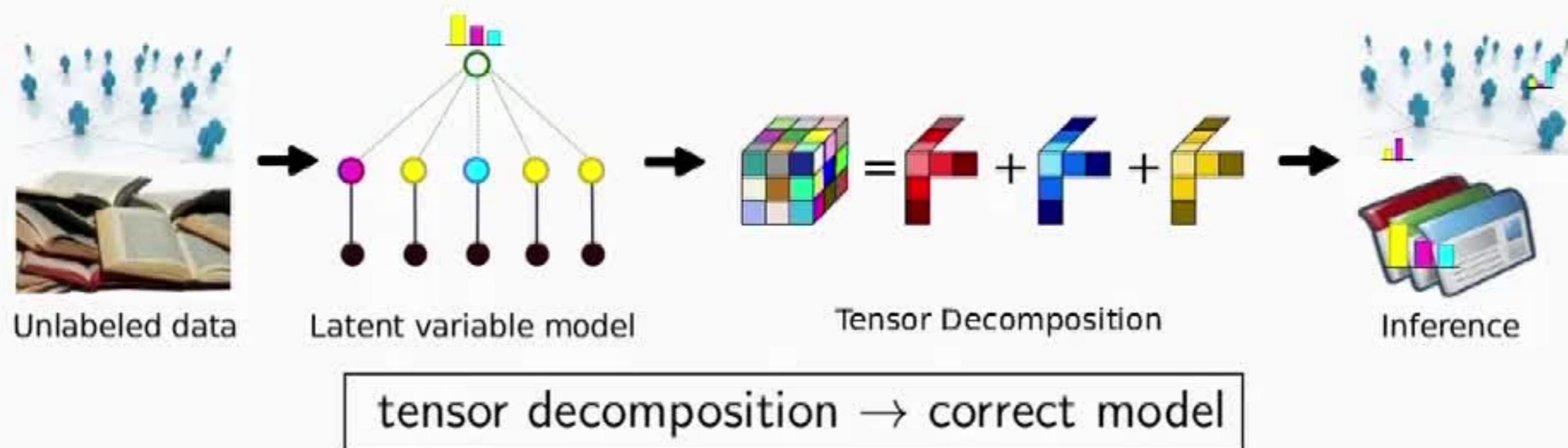
- Whether can model be identified given **infinite computation and data?**
- Are there **tractable algorithms** under identifiability?

Challenge: Efficient Learning of Latent Variable Models

- MCMC: **random sampling**, slow
 - Exponential mixing time
- Likelihood: **non-convex**, not scalable
 - Exponential critical points
- Efficient **computational** and **sample complexities**?

Guaranteed and efficient learning through spectral methods

Unsupervised Learning via Probabilistic Models



Contributions

- Guaranteed **online** algorithm with **global convergence** guarantee
- Highly **scalable**, highly **parallel**, dimensionality reduction
- Tensor library on **CPU/GPU/Spark**
- **Interdisciplinary** applications
- Extension to model with **group invariance**

Outline

1 Introduction

2 Introduction of Method of Moments and Tensor Notations

3 LDA and Community Models

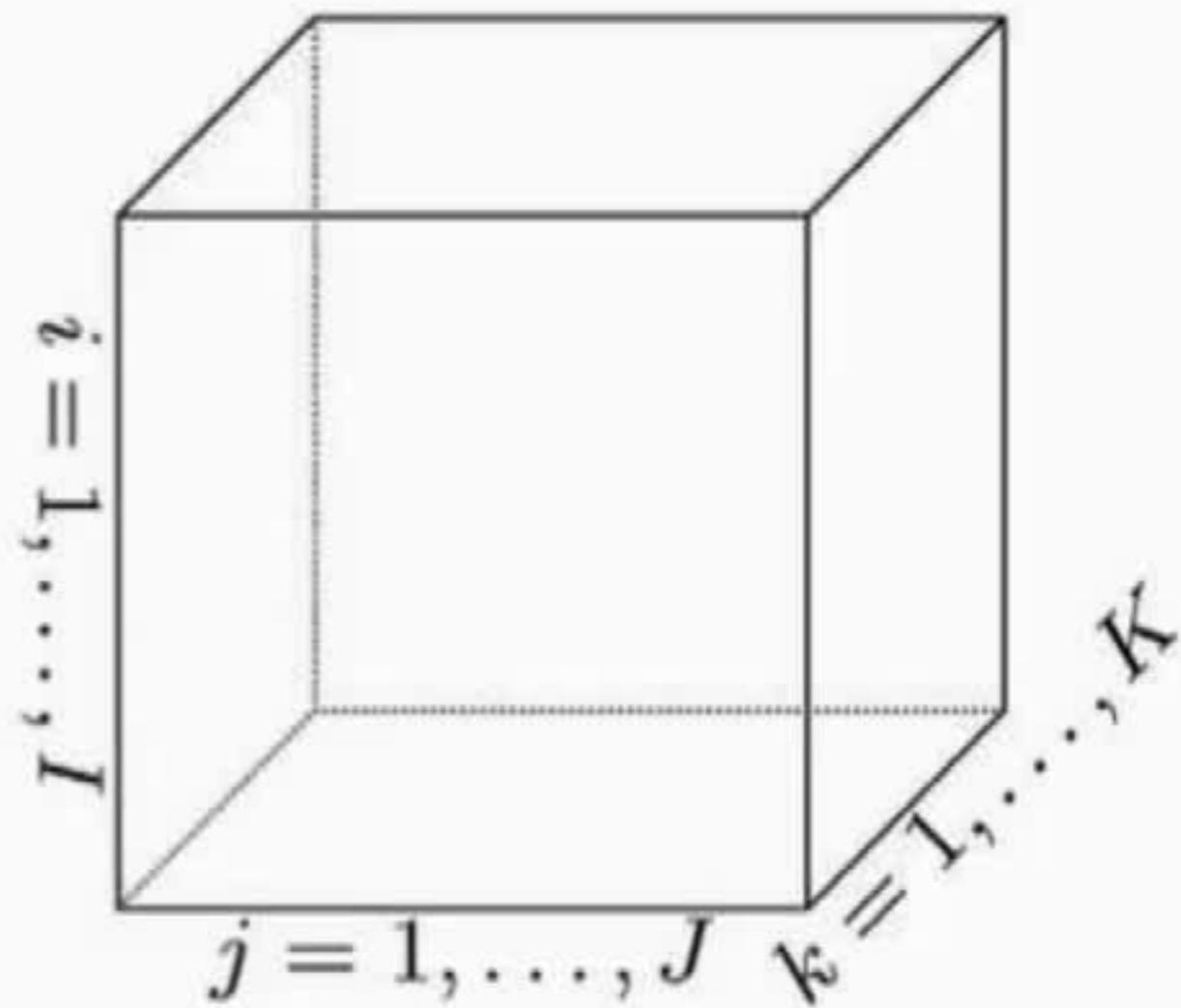
- From Data Aggregates to Model Parameters
- Guaranteed Online Algorithm

4 Conclusion

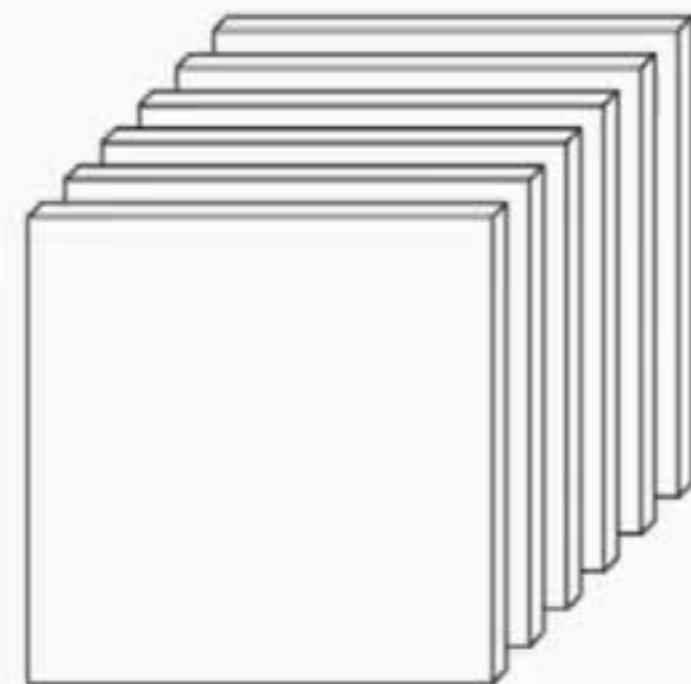
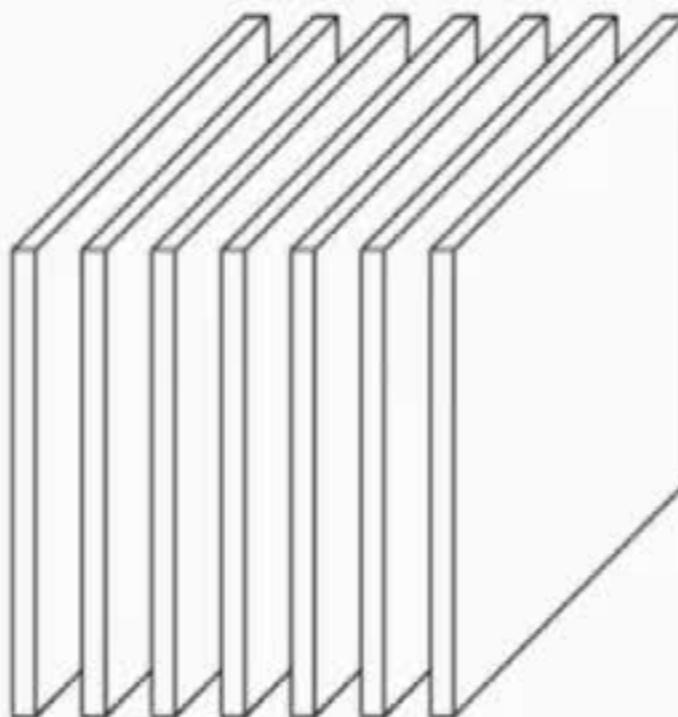
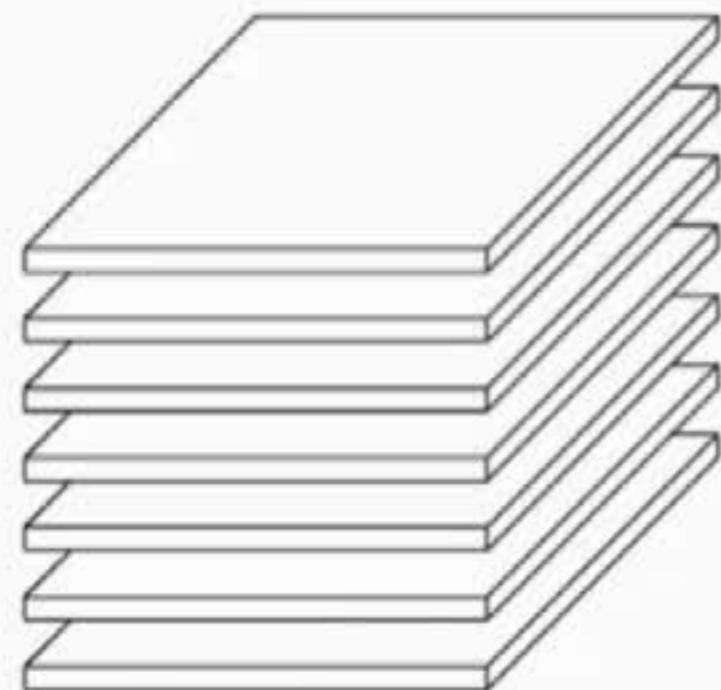
What is a tensor?

Multi-dimensional Array

- Tensor - Higher order matrix
- The number of dimensions is called tensor order.



Slices



- Horizontal slices

- Lateral slices

- Frontal slices

Tensor Product

$$[\mathbf{a} \otimes \mathbf{b}]_{i_1, i_2} = a_{i_1} b_{i_2}$$

The diagram illustrates the tensor product of two vectors. On the left, a 2x4 matrix a is shown with a red column at index i_1 . To its right is an equals sign. Next is a 2x2 vector b with a red row at index i_2 . To the right of b is the expression b_{i_2} .

$$[\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}]_{i_1, i_2, i_3} = a_{i_1} b_{i_2} c_{i_3}$$

The diagram illustrates the tensor product of three tensors. On the left, a 2x2x2 cube a is shown with a red slice at index i_1 . To its right is an equals sign. Next is a 2x2 vector b with a red row at index i_2 . To the right of b is the expression b_{i_2} . Above b is a 2x2 vector c with a red row at index i_3 . To the right of c is the expression c_{i_3} .

- $[\mathbf{a} \otimes \mathbf{b}]_{i_1, i_2} = a_{i_1} b_{i_2}$
- Rank-1 matrix
- $[\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}]_{i_1, i_2, i_3} = a_{i_1} b_{i_2} c_{i_3}$
- Rank-1 tensor

Tensor Product

$$[\mathbf{a} \otimes \mathbf{b}]_{i_1, i_2} = a_{i_1} b_{i_2}$$

The diagram illustrates the tensor product of two vectors. On the left, a 2x4 matrix a is shown with its columns labeled $a_{i_1}, a_{i_2}, a_{i_3}, a_{i_4}$. To its right is an equals sign. Next is a 2x2 vector b with its columns labeled b_{i_1}, b_{i_2} . The result of the multiplication is a 2x2 matrix where each element is the product of the corresponding elements from a and b .

$$[\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}]_{i_1, i_2, i_3} = a_{i_1} b_{i_2} c_{i_3}$$

The diagram illustrates the tensor product of three vectors. On the left, a 2x4x2 cube a is shown with its columns labeled $a_{i_1}, a_{i_2}, a_{i_3}, a_{i_4}$. To its right is an equals sign. Next is a 2x2 vector b with its columns labeled b_{i_1}, b_{i_2} . Then comes another 2x2 vector c with its columns labeled c_{i_1}, c_{i_2} . The result of the multiplication is a 2x2x2 cube where each element is the product of the corresponding elements from a , b , and c .

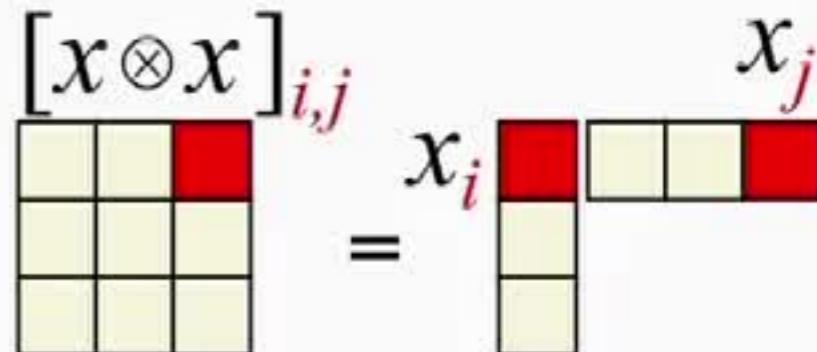
- $[\mathbf{a} \otimes \mathbf{b}]_{i_1, i_2} = a_{i_1} b_{i_2}$
- Rank-1 matrix
- $[\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}]_{i_1, i_2, i_3} = a_{i_1} b_{i_2} c_{i_3}$
- Rank-1 tensor

Tensors in Method of Moments

Matrix: Pair-wise relationship

- Signal or data observed $\mathbf{x} \in \mathbb{R}^d$
- Rank 1 matrix: $[\mathbf{x} \otimes \mathbf{x}]_{i,j} = x_i x_j$
- Aggregated pair-wise relationship

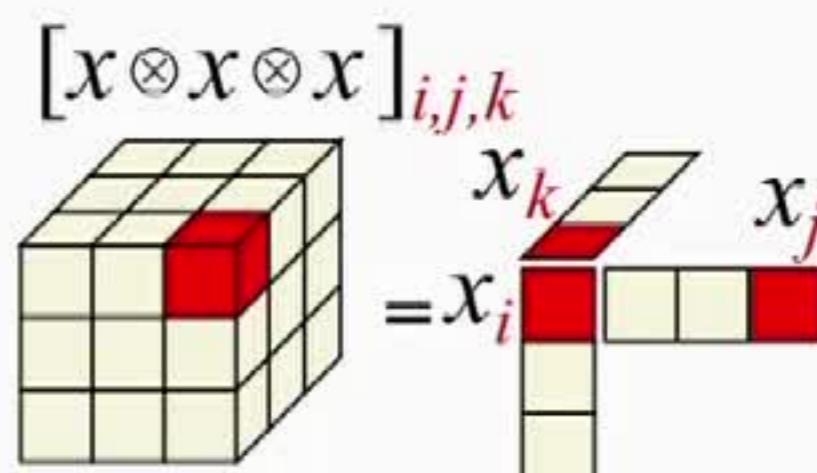
$$\mathbf{M}_2 = \mathbb{E}[\mathbf{x} \otimes \mathbf{x}]$$



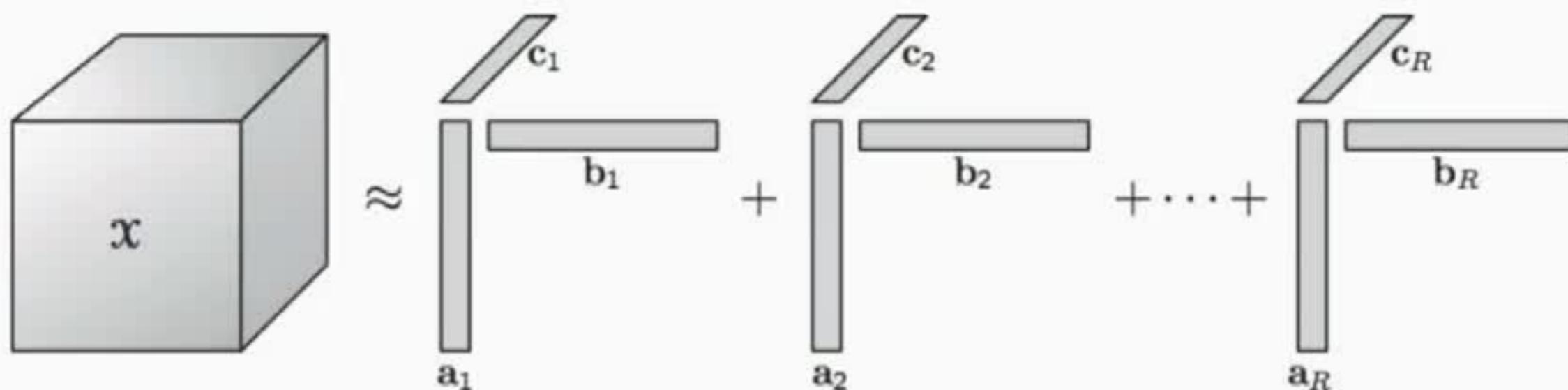
Tensor: Triple-wise relationship or higher

- Signal or data observed $\mathbf{x} \in \mathbb{R}^d$
- Rank 1 tensor:
 $[\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}]_{i,j,k} = x_i x_j x_k$
- Aggregated triple-wise relationship

$$\mathcal{M}_3 = \mathbb{E}[\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}] = \mathbb{E}[\mathbf{x} \otimes 3]$$



CP decomposition

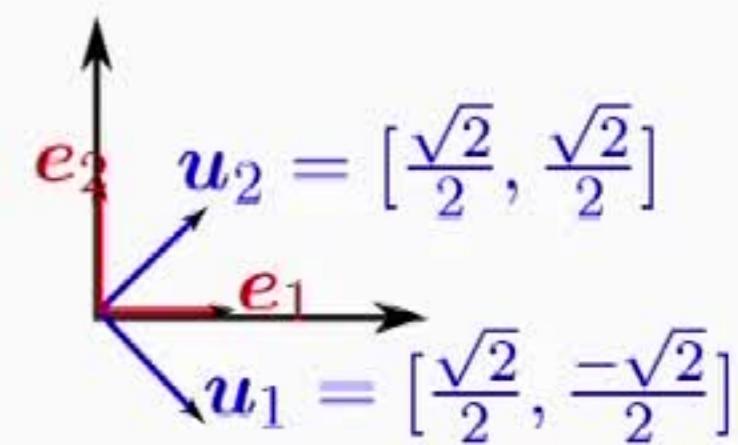


- $\mathcal{X} = \sum_{h=1}^R \mathbf{a}_h \otimes \mathbf{b}_h \otimes \mathbf{c}_h$
- Summation of rank-1 tensors

Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap
 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$
- Unique with eigenvalue gap



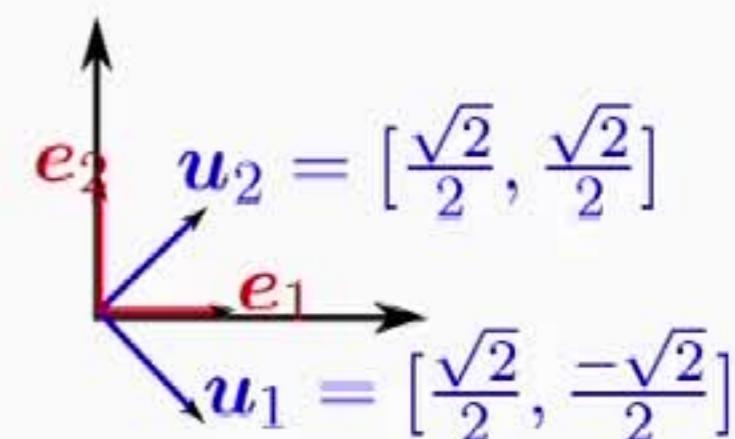
Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$$

- Unique with eigenvalue gap



Tensor Orthogonal Decomposition (Harshman, 1970)

- Unique: eigenvalue gap not needed

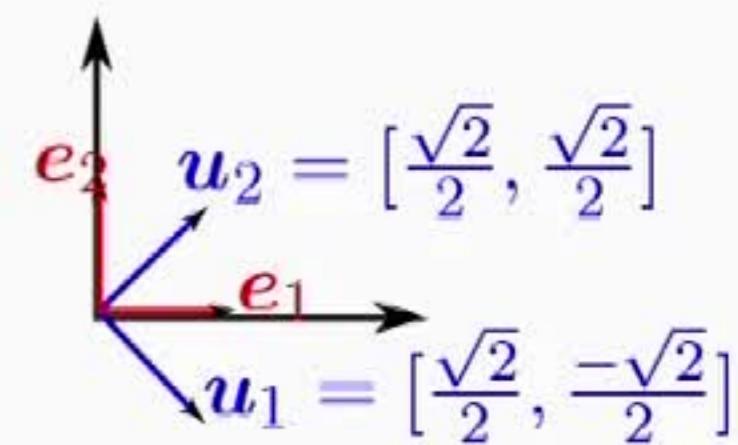
A 3D tensor decomposition diagram. On the left is a large, multi-colored cube representing the original tensor. To its right is an equals sign. Following the equals sign are two tensors with smiley faces, representing the components of the decomposition. A plus sign is placed between the two components. Below the diagram is the mathematical equation:

$$Tensor = \mathbf{u}_1 \otimes \mathbf{u}_1 \otimes \mathbf{u}_1 + \mathbf{u}_2 \otimes \mathbf{u}_2 \otimes \mathbf{u}_2$$

Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap
 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$
- Unique with eigenvalue gap



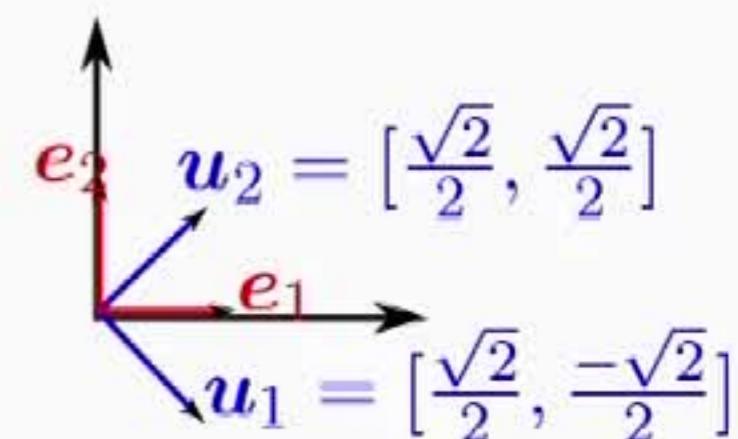
Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap

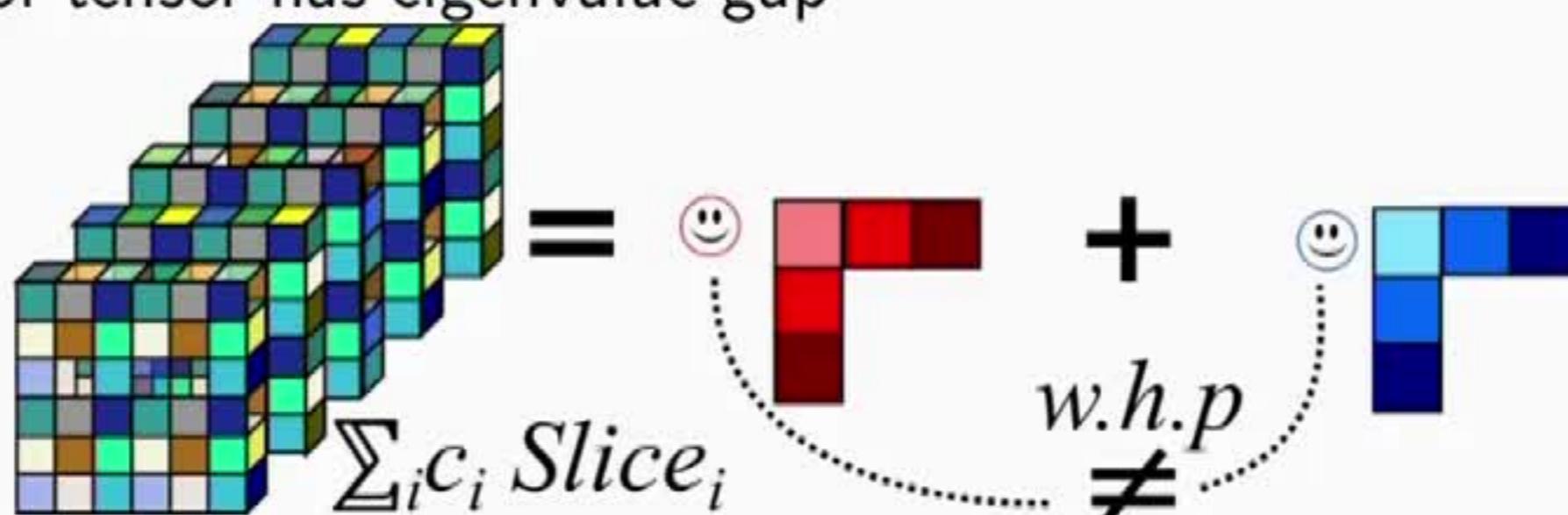
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$$

- Unique with eigenvalue gap



Tensor Orthogonal Decomposition (Harshman, 1970)

- Unique: eigenvalue gap not needed
- Slice of tensor has eigenvalue gap



Outline

1 Introduction

2 Introduction of Method of Moments and Tensor Notations

3 LDA and Community Models

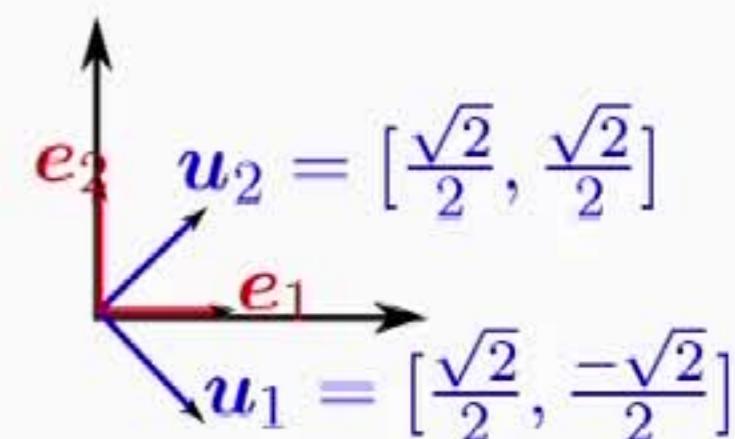
- From Data Aggregates to Model Parameters
- Guaranteed Online Algorithm

4 Conclusion

Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap
- $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$
- Unique with eigenvalue gap



Tensor Orthogonal Decomposition (Harshman, 1970)

- Unique: eigenvalue gap not needed
- Slice of tensor has eigenvalue gap

A 3D tensor is shown as a stack of colored cubes. To its right is an equals sign. Following the equals sign are two smiley face icons, a plus sign, and another two smiley face icons. Below the tensor and slices is the equation:

$$Slice_i = u_1(i) \mathbf{u}_1 \otimes \mathbf{u}_1 + u_2(i) \mathbf{u}_2 \otimes \mathbf{u}_2$$

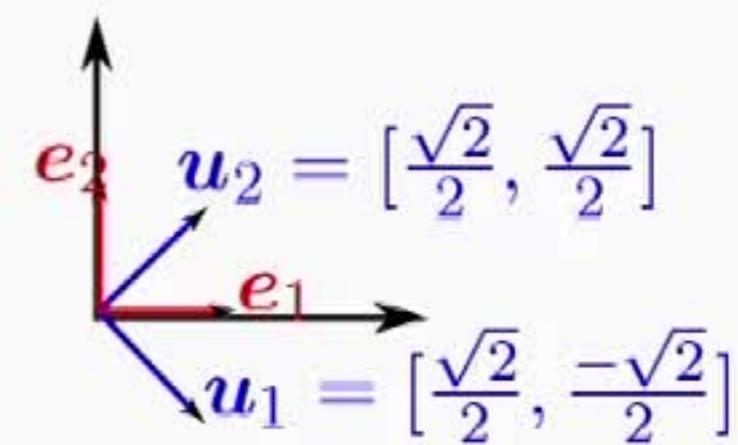
Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$$

- Unique with eigenvalue gap



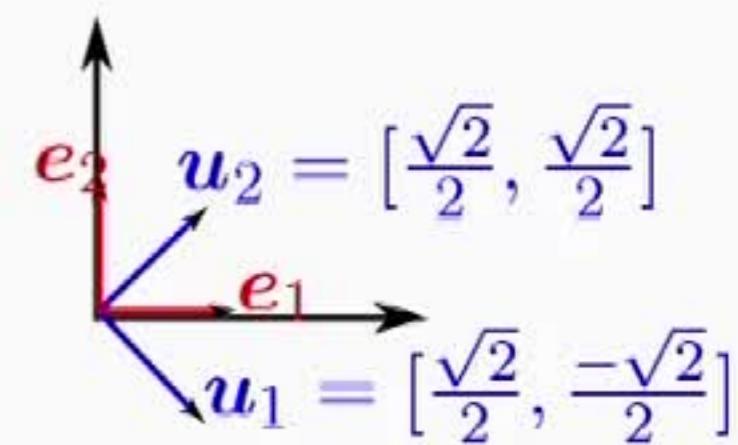
Why are tensors powerful?

Matrix Orthogonal Decomposition

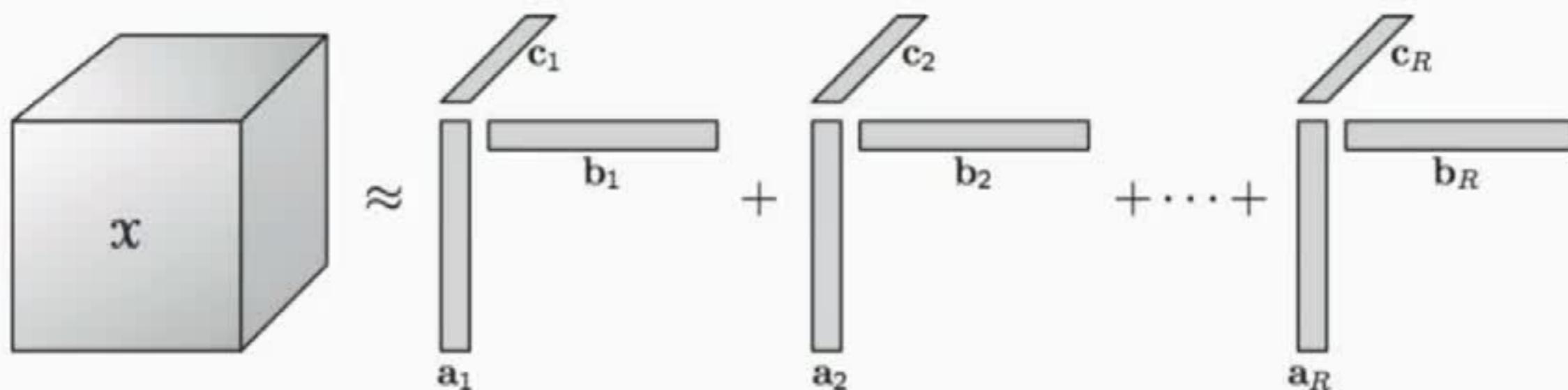
- Not unique without eigenvalue gap

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$$

- Unique with eigenvalue gap



CP decomposition



- $\mathcal{X} = \sum_{h=1}^R \mathbf{a}_h \otimes \mathbf{b}_h \otimes \mathbf{c}_h$
- Summation of rank-1 tensors

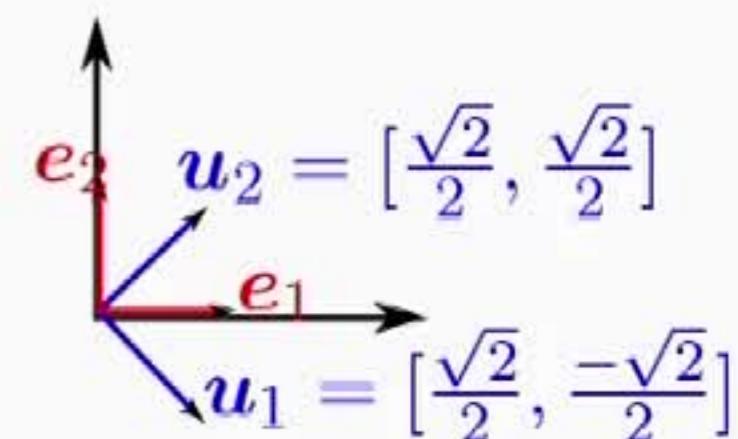
Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$$

- Unique with eigenvalue gap



Tensor Orthogonal Decomposition (Harshman, 1970)

- Unique: eigenvalue gap not needed

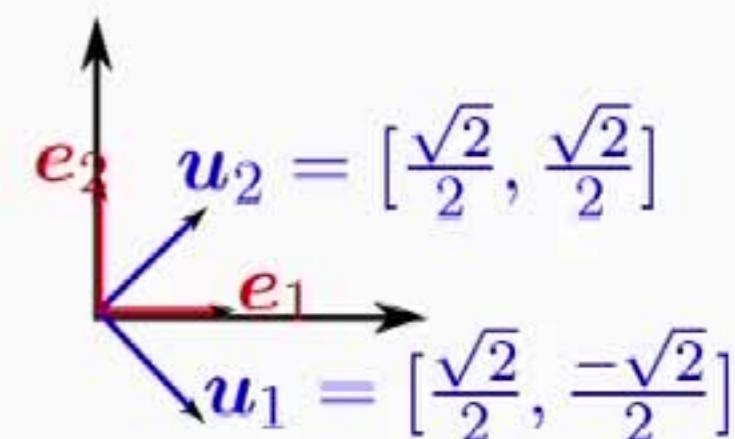
A diagram illustrating tensor decomposition. On the left is a large 3D cube composed of many small colored blocks. To its right is an equals sign. To the right of the equals sign are two smaller 3D tensors, each with a smiley face icon above it. The first smaller tensor has a red color gradient, and the second has a blue color gradient. Below the tensors is the mathematical equation:

$$Tensor = \mathbf{u}_1 \otimes \mathbf{u}_1 \otimes \mathbf{u}_1 + \mathbf{u}_2 \otimes \mathbf{u}_2 \otimes \mathbf{u}_2$$

Why are tensors powerful?

Matrix Orthogonal Decomposition

- Not unique without eigenvalue gap
- $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top = \mathbf{u}_1 \mathbf{u}_1^\top + \mathbf{u}_2 \mathbf{u}_2^\top$
- Unique with eigenvalue gap



Tensor Orthogonal Decomposition (Harshman, 1970)

- Unique: eigenvalue gap not needed
- Slice of tensor has eigenvalue gap

A 3D tensor is shown as a stack of colored cubes. To its right is an equals sign. Following the equals sign are two smiley face icons, a plus sign, and another two smiley face icons. Below the tensor and the decomposition symbols is the equation:

$$Slice_i = u_1(i) \mathbf{u}_1 \otimes \mathbf{u}_1 + u_2(i) \mathbf{u}_2 \otimes \mathbf{u}_2$$

Outline

1 Introduction

2 Introduction of Method of Moments and Tensor Notations

3 LDA and Community Models

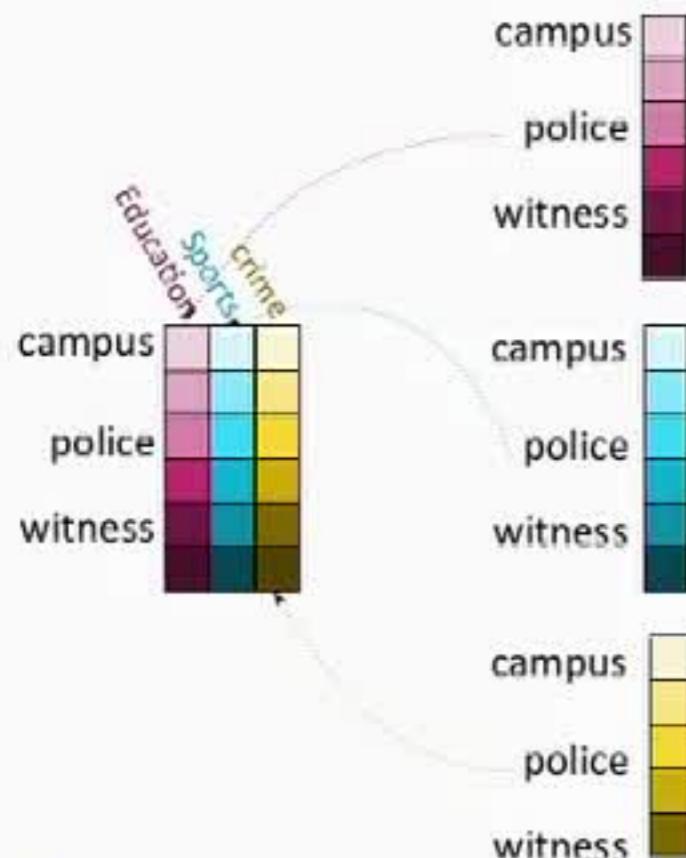
- From Data Aggregates to Model Parameters
- Guaranteed Online Algorithm

4 Conclusion

Probabilistic Topic Models - LDA

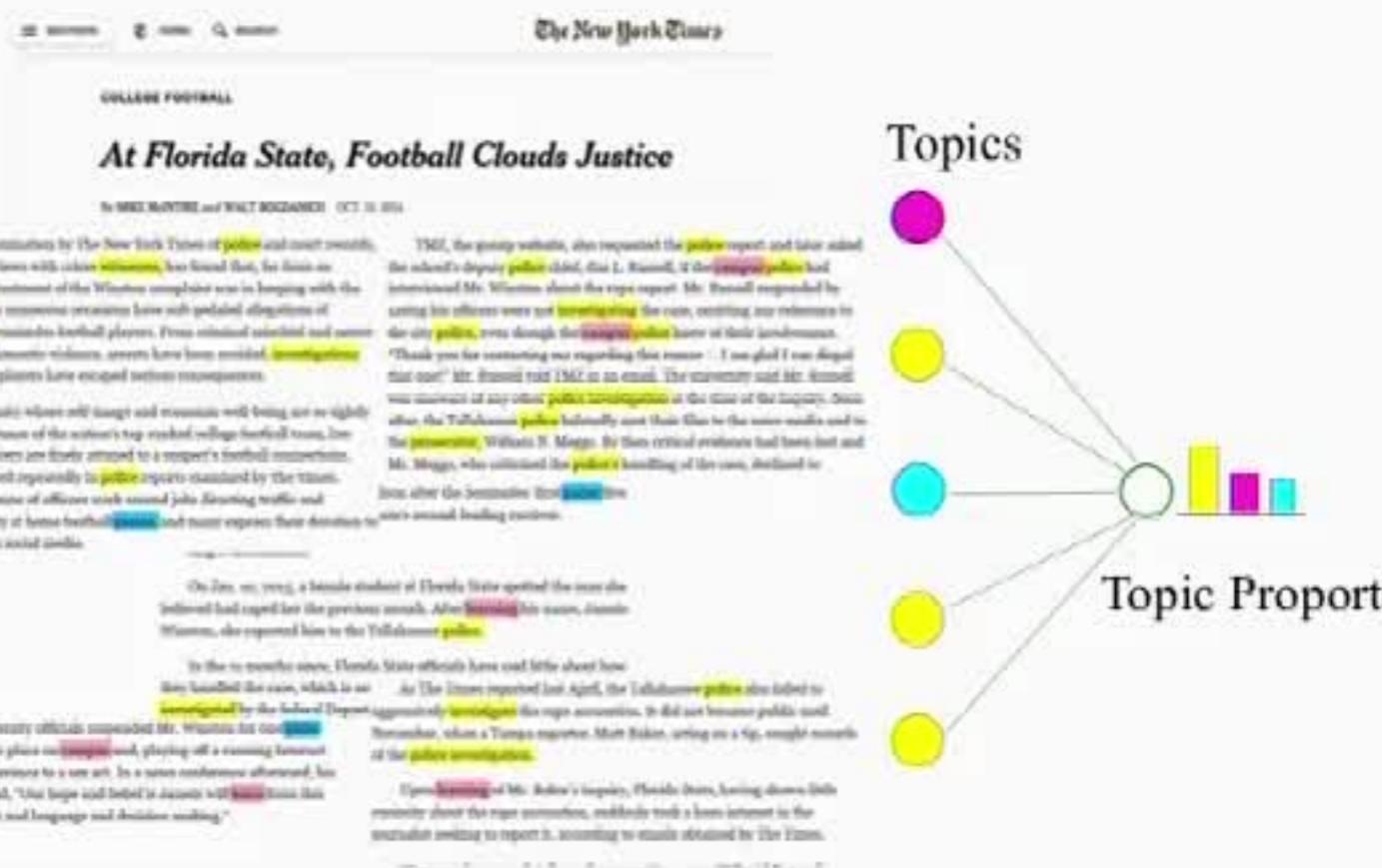
Bag of words

- Infer topics of documents
- Learn hidden process drives the obs.



Generative model

- Topic proportion $\sim \text{Dir}(\alpha)$ for a doc
- Draw a topic, then a word for a token



Goal



- Topic-word matrix

$$\mathbb{P}[\text{word} = e_i | \text{topic} = j]$$

Moments Matching

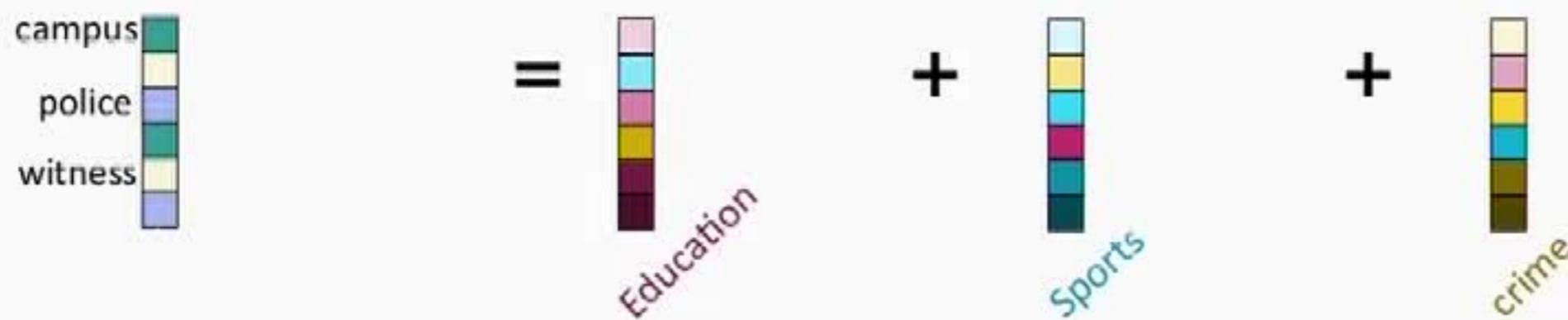
Goal: Linearly independent topic-word table

	campus
	police
	witness

$$\mathbb{E}[\text{word} | \text{topic} = j] = \sum_i \mathbb{P}[\text{word} = e_i | \text{topic} = j] e_i = \text{column } j$$

M_1 : Occurrence Frequency of Words

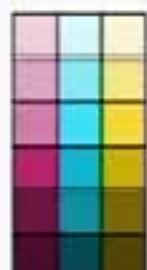
$$\mathbb{E}[\text{word}] = \sum_j \mathbb{E}[\text{word} | \text{topic} = j] \mathbb{P}[\text{topic} = j]$$



No unique decomposition of vectors

Moments Matching

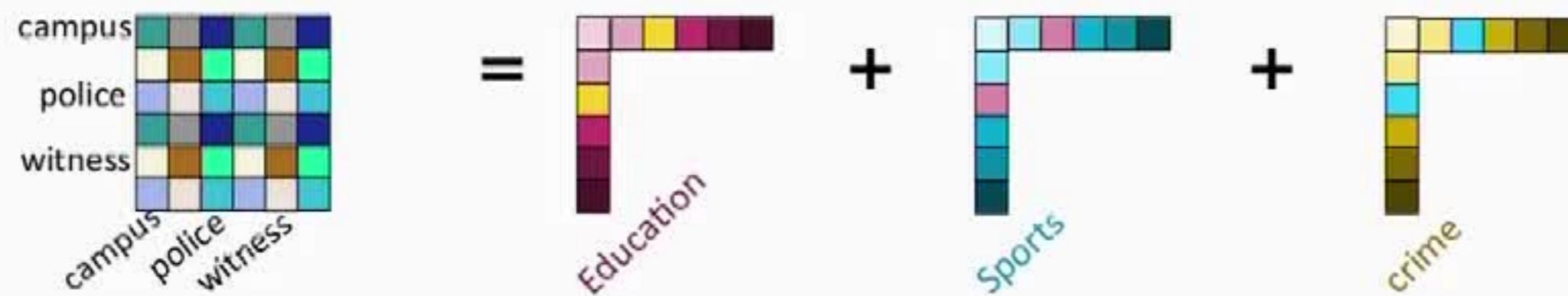
Goal: Linearly independent topic-word table


campus
police
witness

$$\mathbb{E}[\text{word} | \text{topic} = j] = \sum_i \mathbb{P}[\text{word} = e_i | \text{topic} = j] e_i = \text{column } j$$

M_2 : Modified Co-occurrence Frequency of Word Pairs

$$\mathbb{E}[\text{word}_1 \otimes \text{word}_2] = \sum_{j,j'} \mathbb{E}[\text{word}_1 | \text{topic}_1 = j] \otimes \mathbb{E}[\text{word}_2 | \text{topic}_2 = j'] \mathbb{P}[\text{topic}_1 = j, \text{topic}_2 = j']$$

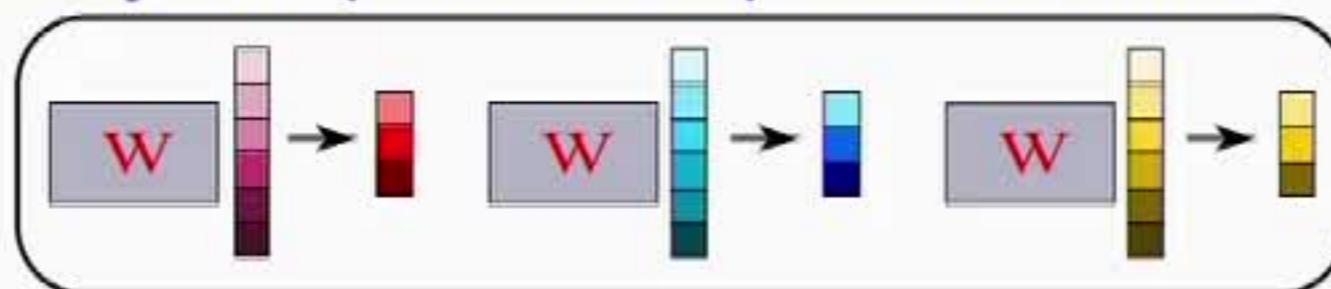


Matrix decomposition recovers subspace, not actual model

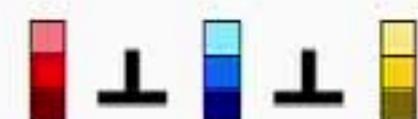
Moments Matching

Goal: Linearly independent topic-word table

Find a W

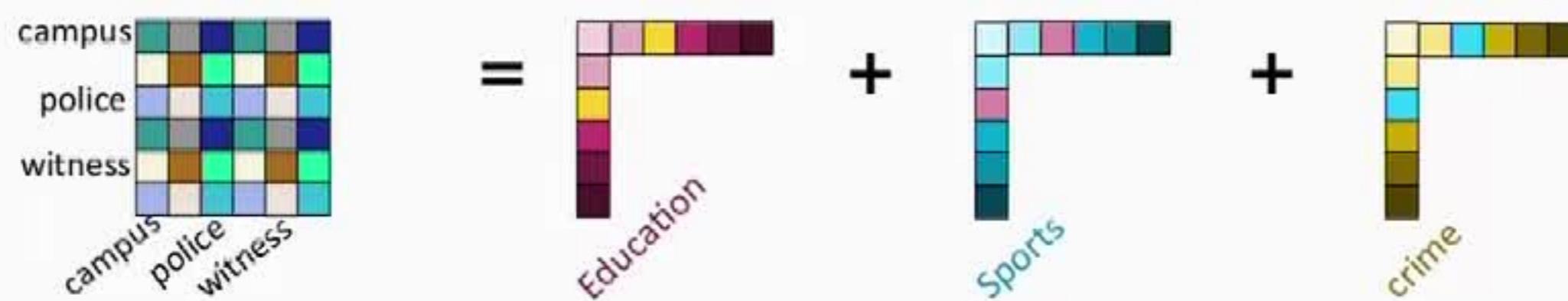


such that



M_2 : Modified Co-occurrence Frequency of Word Pairs

$$\mathbb{E}[\text{word}_1 \otimes \text{word}_2] = \sum_{j,j'} \mathbb{E}[\text{word}_1 | \text{topic}_1 = j] \otimes \mathbb{E}[\text{word}_2 | \text{topic}_2 = j'] \mathbb{P}[\text{topic}_1 = j, \text{topic}_2 = j']$$

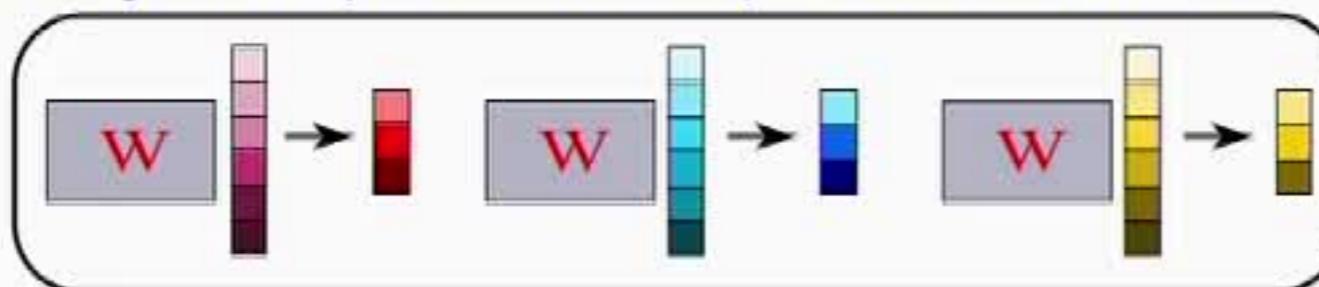


Many such W 's, find one, project data with W

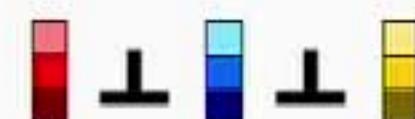
Moments Matching

Goal: Linearly independent topic-word table

Know a W



such that



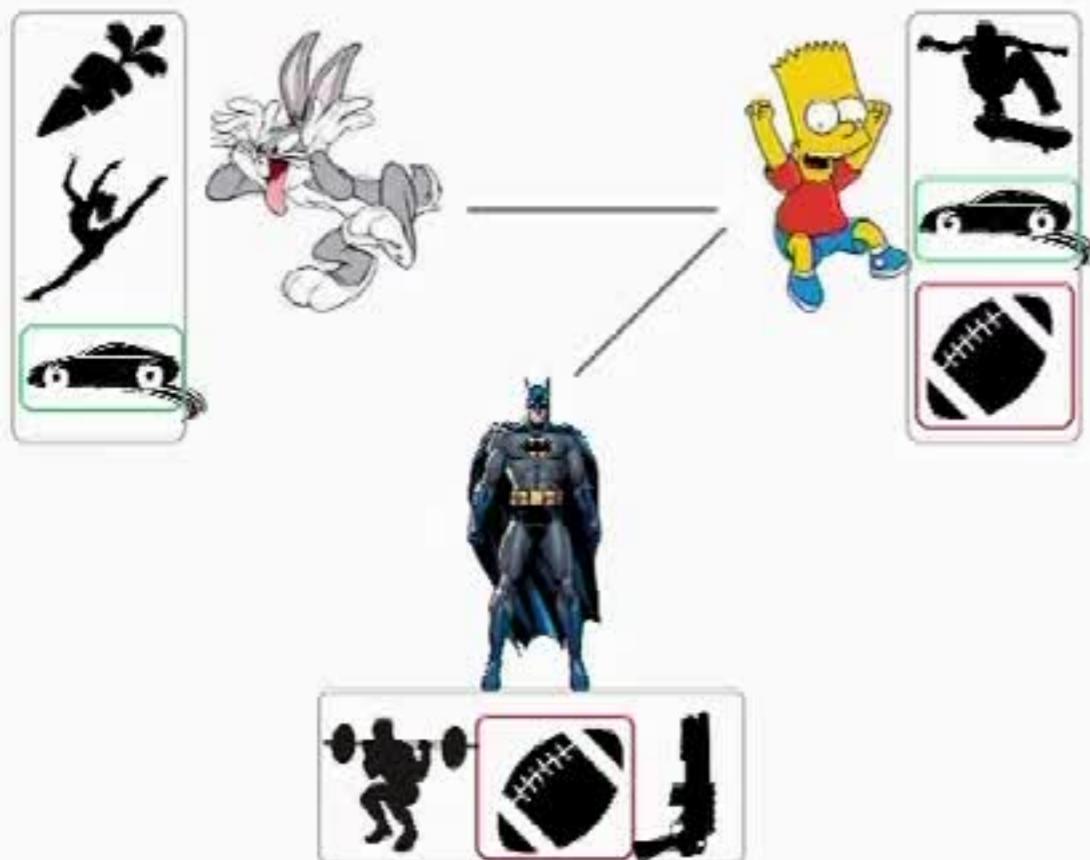
M_3 : Modified Co-occurrence Frequency of Word Triplets

The diagram shows a 3D tensor W represented as a cube with colored blocks. It is shown to be equal to the sum of three smaller tensors f_1 , f_2 , and f_3 . Each f_i is a tensor with a similar 3D structure but with fewer colors, representing the components of the decomposition.

Tensor decomposition uniquely discovers the correct model

Mixed Membership Community Models

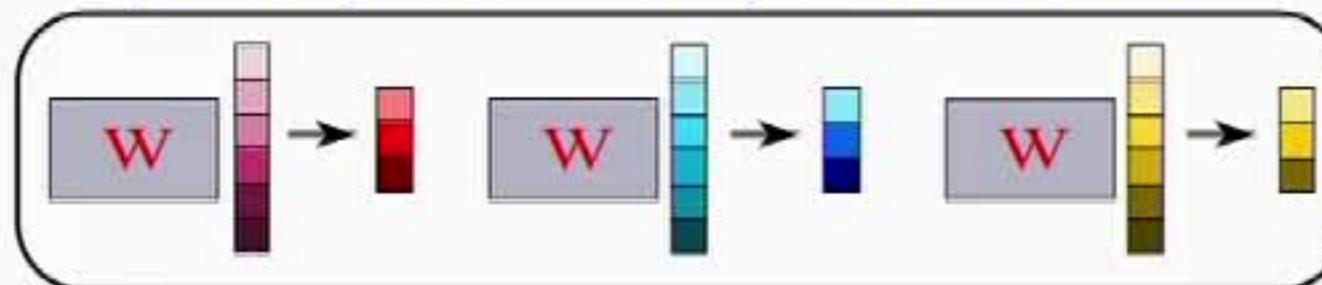
Mixed memberships



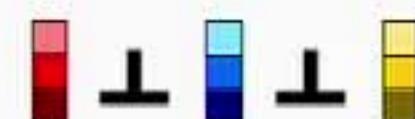
Moments Matching

Goal: Linearly independent topic-word table

Know a W



such that



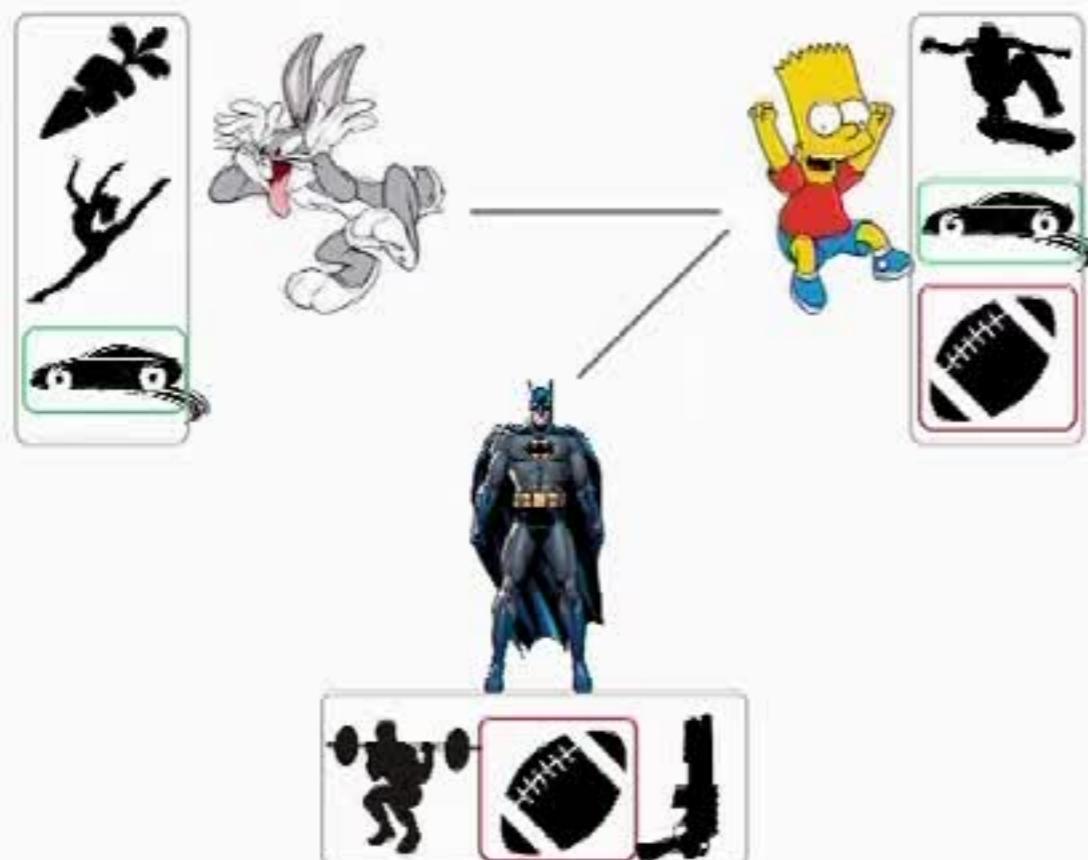
M_3 : Modified Co-occurrence Frequency of Word Triplets

The diagram shows a 3D tensor M_3 represented as a cube with three 'w' labels on its top face. An equals sign follows the cube, followed by three smaller 3D tensors F_1 , F_2 , and F_3 , each with a plus sign between them, indicating they are summed to form the original tensor.

Tensor decomposition uniquely discovers the correct model

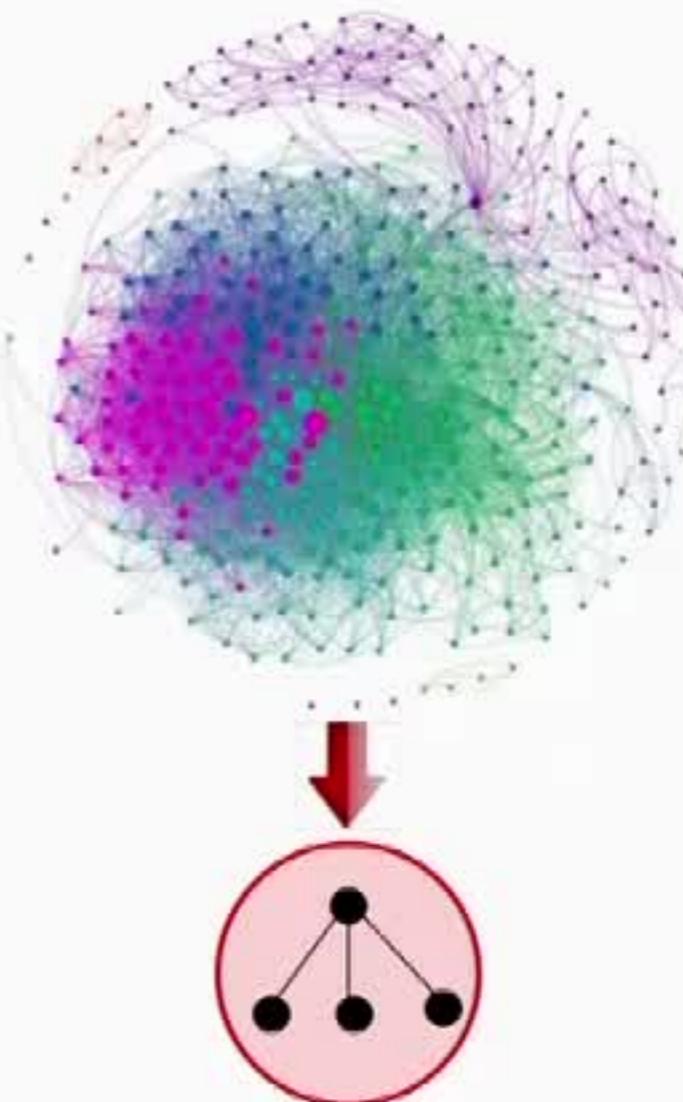
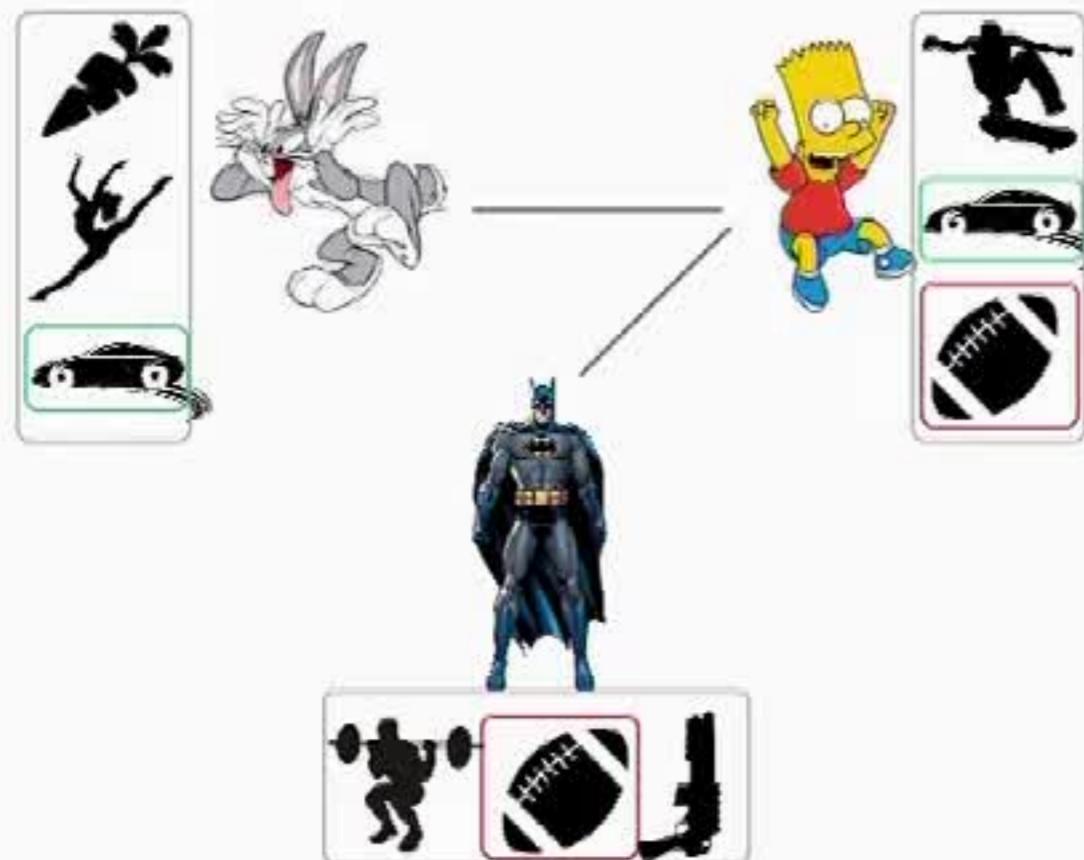
Mixed Membership Community Models

Mixed memberships

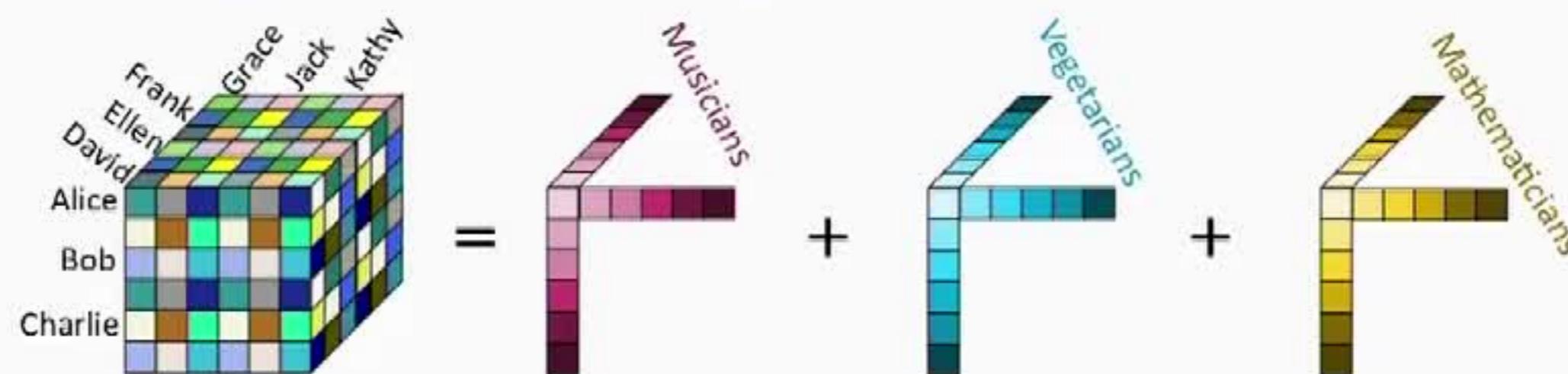


Mixed Membership Community Models

Mixed memberships



What ensures guaranteed learning?



Guaranteed Online Tensor Decomposition

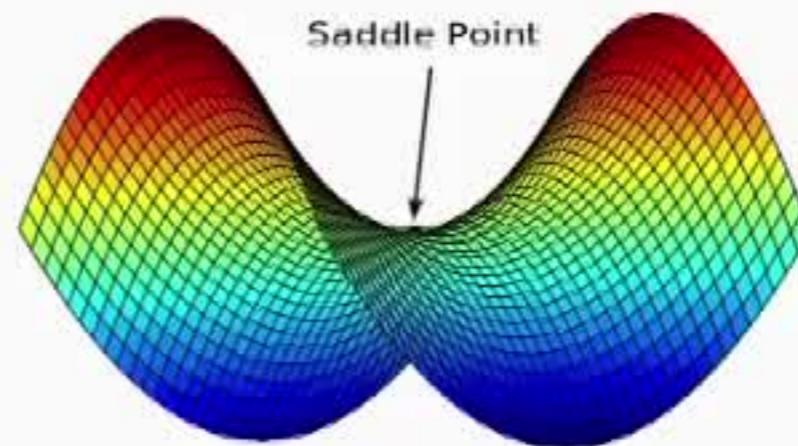
Model is uniquely identifiable! How to identify?

Online Tensor Decomposition

- Tensor $T = \sum_i a_i \otimes a_i \otimes a_i \otimes a_i$, where $\|a_i\| = 1, a_i^\top a_j = 0$
Objective?
- Objective $\min_{\forall i, \|u_i\|^2=1} \sum_{i \neq j} T(u_i, u_i, u_j, u_j)$ **Non-convex!**

Theorem: The proposed objective function has equivalent local optima.

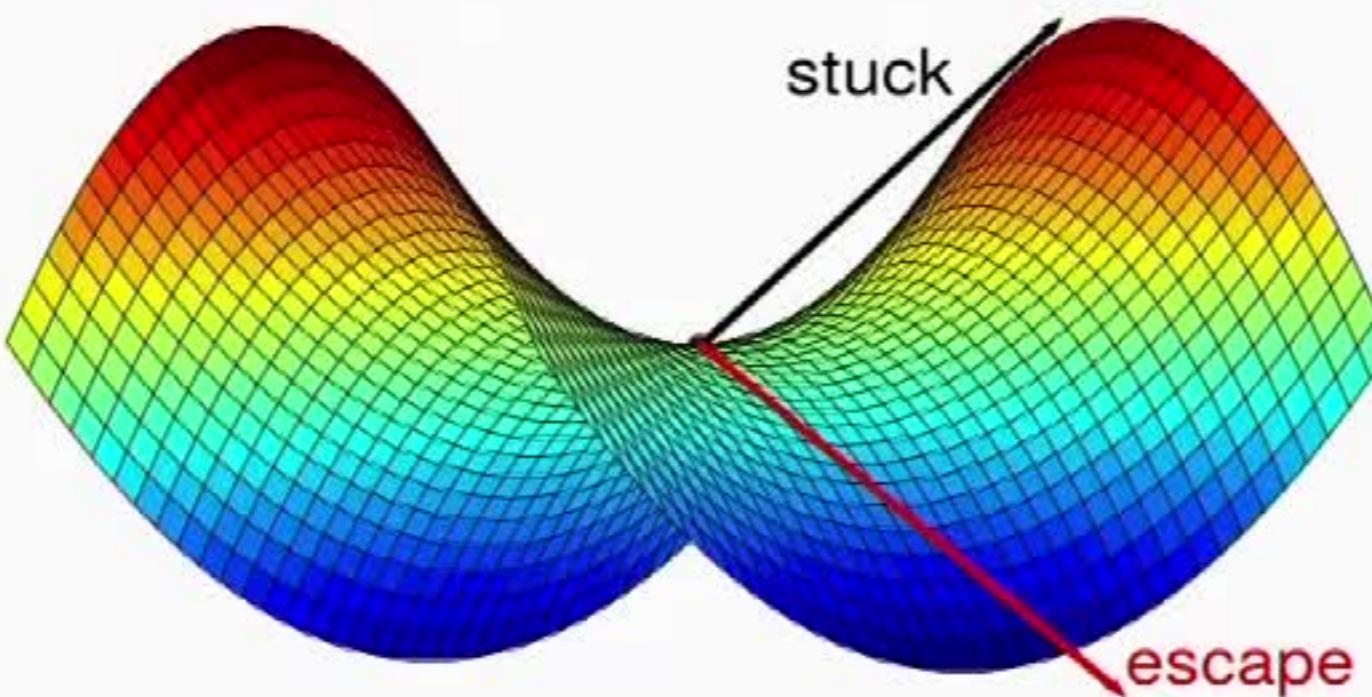
Will SGD work?



Theorem: For smooth, twice-diff fn. with non-degenerate saddle points, **noisy SGD** converges to a local optimum in polynomial steps.

Why could we escape from saddle points?

Stochastic Gradient Descent with Noise



- Saddle point has 0 gradient
- Non-degenerate saddle: Hessian has \pm eigenvalue
- Negative eigenvalue: direction of escape

Noise could help!

Outline

1 Introduction

2 Introduction of Method of Moments and Tensor Notations

3 LDA and Community Models

- From Data Aggregates to Model Parameters
- Guaranteed Online Algorithm

4 Conclusion

Summary

Spectral methods reveal hidden structure

- Text/Image processing
- Social networks
- Neuroscience, healthcare ...



Summary

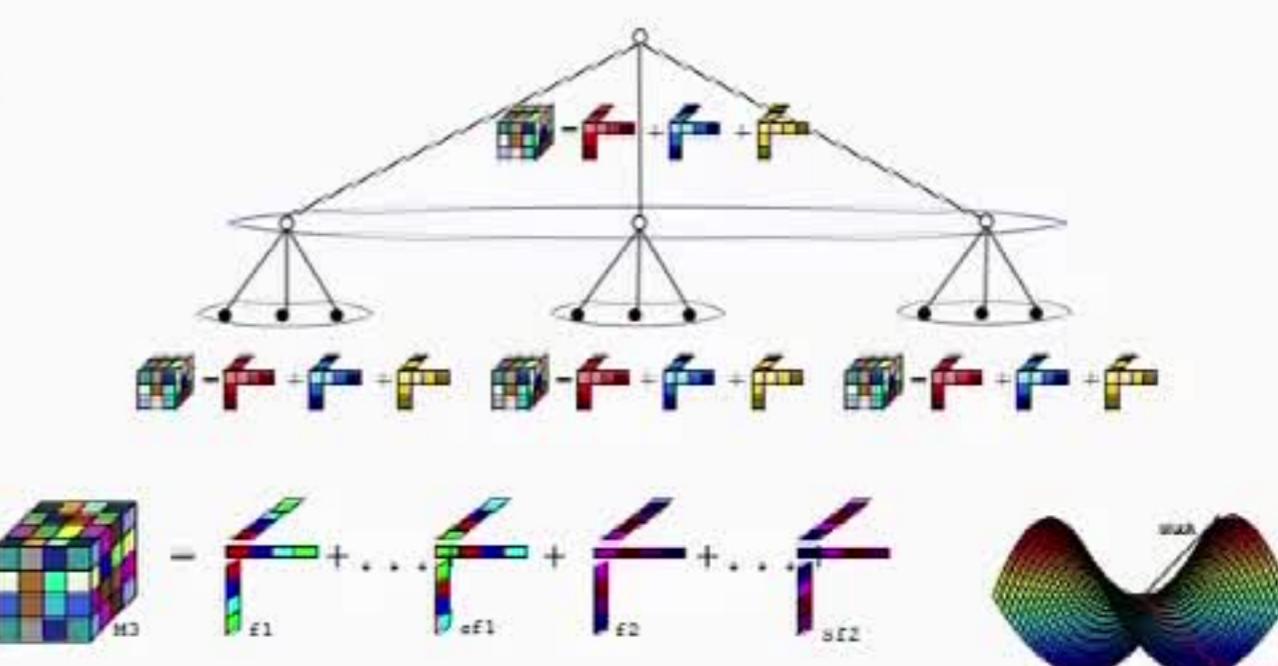
Spectral methods reveal hidden structure

- Text/Image processing
- Social networks
- Neuroscience, healthcare ...



Versatile for latent variable models

- Flat model \rightarrow hierarchical model
- Sparse coding \rightarrow convolutional model
- Efficient, convergence guarantee



A list of papers related to this talk

Topic Model, Community Detection, Feature Learning

- "Online Tensor Methods for Learning Latent Variable Models", by **F. Huang**, U.N. Niranjan, M.U. Hakeem, A. Anandkumar, JMLR 2014.
- "Convolutional Dictionary Learning through Tensor Factorization", by **F. Huang** and A. Anandkumar, conference and workshop proceeding of JMLR, vol.44, Dec 2015.
- "Tensor Methods on Apache Spark", by **F. Huang**, A. Anandkumar, Oct. 2015.

Guaranteed tensor decomposition

- "Escaping From Saddle Points — Online Stochastic Gradient for Tensor Decomposition", by R. Ge, **F. Huang**, C. Jin, Y. Yuan, COLT 2015.

Application in Health Analytics and Neuroscience

- "Scalable Latent TreeModel and its Application to Health Analytics", by **F. Huang**, U.N. Niranjan, I. Perros, R. Chen, J. Sun, A. Anandkumar, NIPS 2015 MLHC workshop.
- "Discovering Neuronal Cell Types and Their Gene Expression Profiles Using a Spatial Point Process Mixture Model", by **F. Huang**, A. Anandkumar, C. Borgs, J. Chayes, E. Fraenkel, M. Hawrylycz, E. Lein, A. Ingrosso, S. Turaga, NIPS 2015 BigNeuro workshop.

Collaborators



Anima Anandkumar
Caltech



Rong Ge
Duke University



Srini Turaga
Janelia Research



Chi Jin
UC Berkeley



Jennifer Chayes
MSR



Christian Borgs
MSR



Ernest Fraenkel
MIT



Yang Yuan
Cornell U



UN Niranjan
UC Irvine