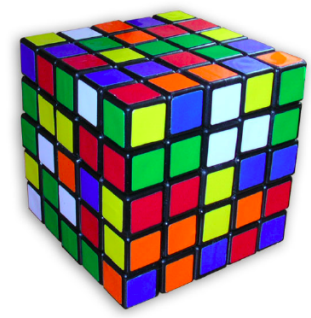
Beyond matrices: tensor decomposition and its application in sciences

Miaoyan Wang

Department Faculty lighting talk

Thanks: NSF DMS 1915978 and OVCRG Grant

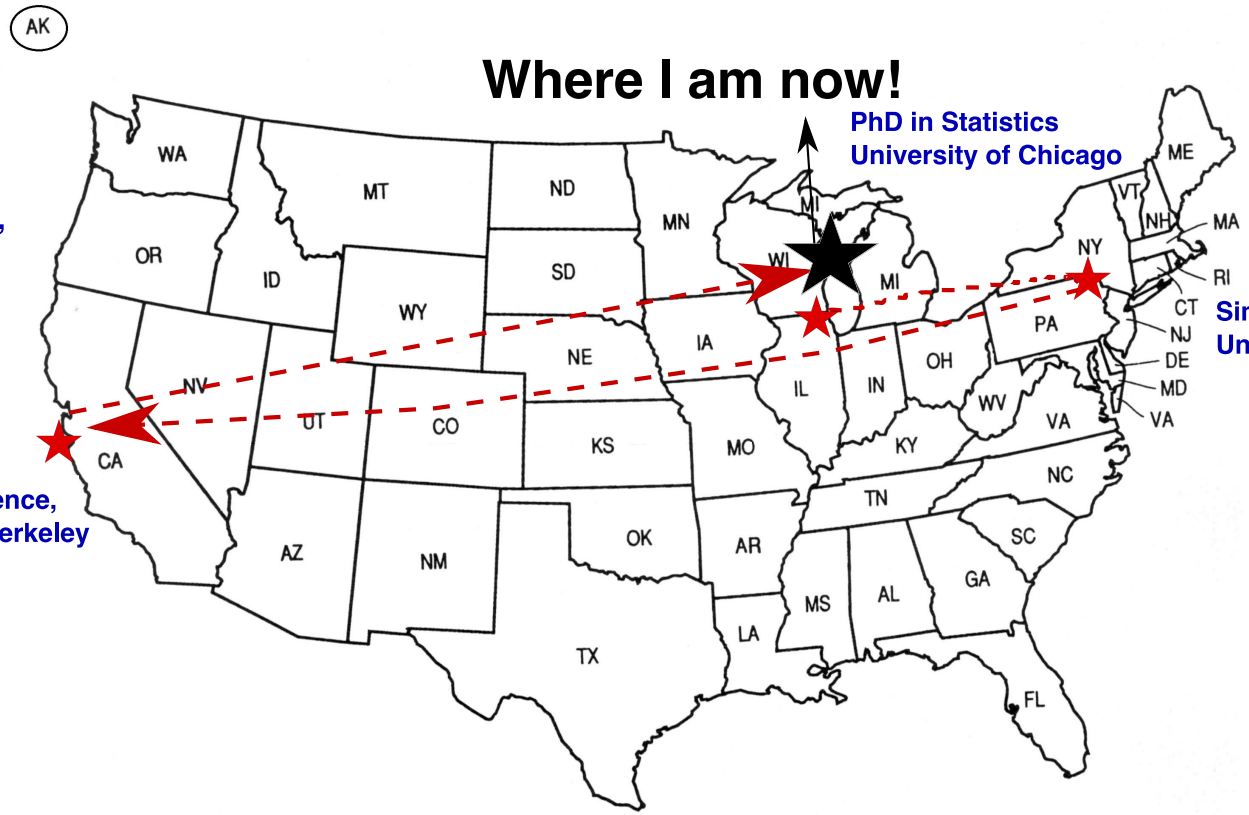
Sep 1, 2019



Who am I?

Originally from China,
B.S. in Mathematics,
Fudan University

Postdoc in Computer Science,
University of California, Berkeley



My research

Statistical machine learning:

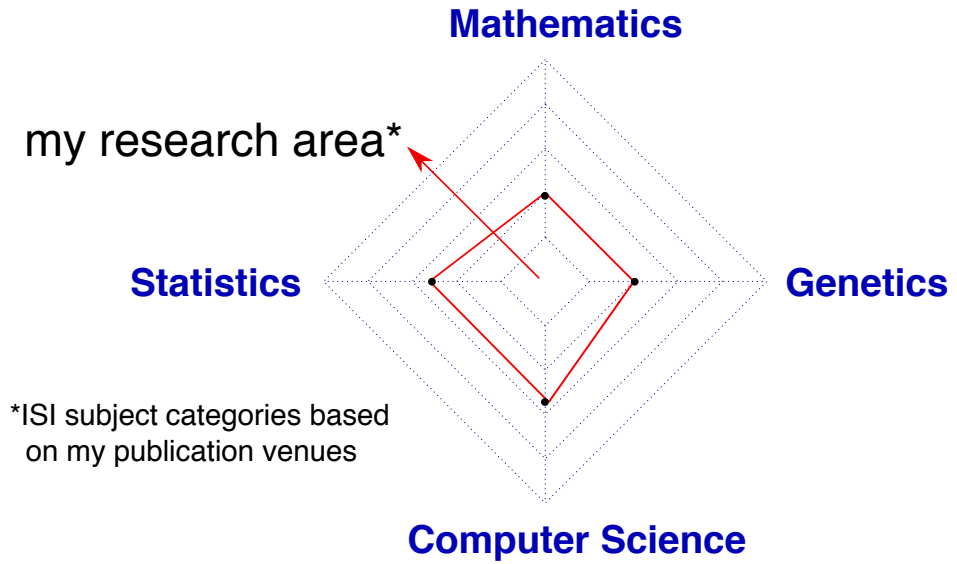
- tensor/matrix decomposition, high-dimensional statistics.

Applied Mathematics:

- numerical algebra, multilinear optimization, combinatorics.

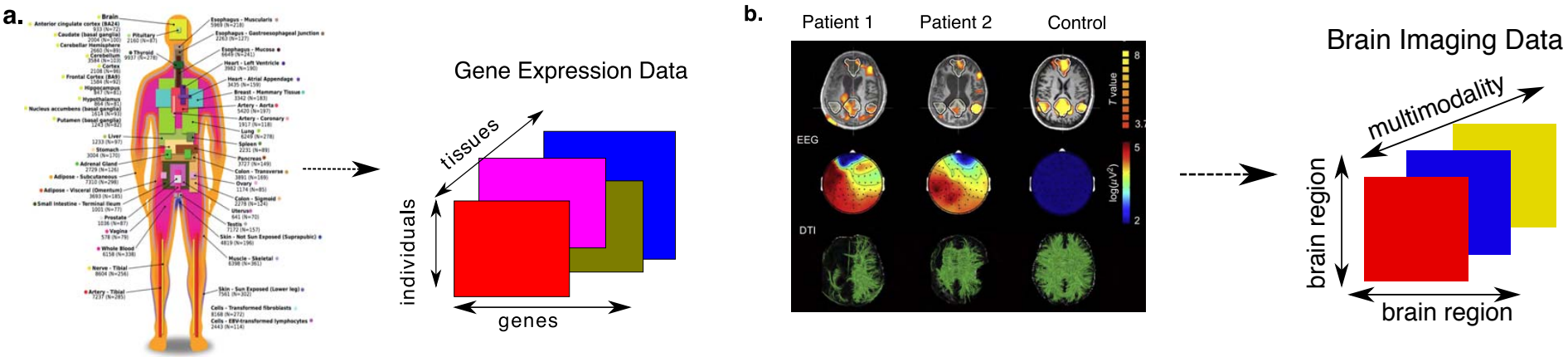
Genetics:

- genetic association studies, gene expression, neuroimaging.



Tensors in science

- Many biomedical datasets come naturally in a multiway form.
- Multi-tissue, multi-individual gene expression measures could be organized as an order-3 tensor $\mathcal{A} = [\![a_{git}]\!] \in \mathbb{R}^{n_G \times n_I \times n_T}$.



Multi-way Clustering in Gene Expression

To identify subsets of genes that are similarly expressed within subsets of individuals and tissues, we seek local blocks in the expression tensor.

Tensors in statistical modeling

"Tensors are the new matrices" that tie together a wide range of areas:

- Longitudinal social network data $\{\boldsymbol{Y}_t : t = 1, ..., n\}$
- Spatio-temporal transcriptome data
- Joint probability table of a set of variables $\mathbb{P}(X_1, X_2, X_3)$
- Higher-order moments in single topic models
- Markov models for the phylogenetic tree $K_{1,3}$

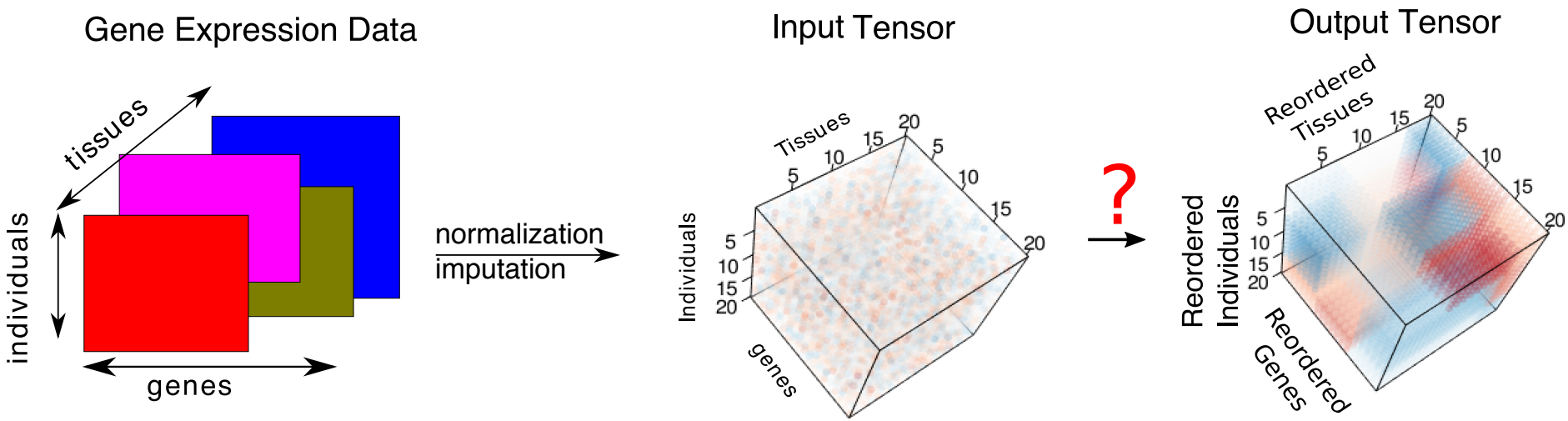
M. Yuan et al 2017, P. Hoff 2015, Montanari-Richard 2014

Anandkumar et al 2014, Mossel et al 2004, P. McCullagh 1987

Why study tensors?

Tensors provide a rich source of

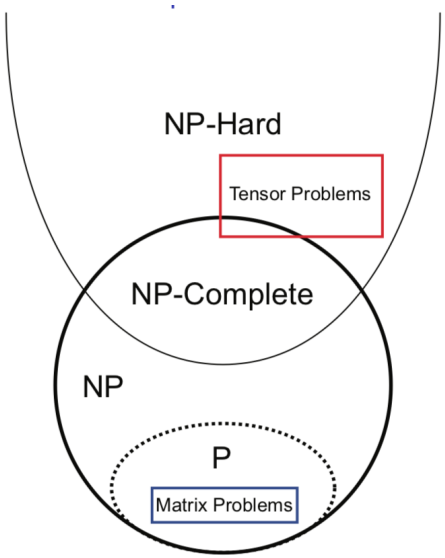
- fundamental problems in data science.
- new tools for long-standing questions.
- huge potentials for new applications.



My research

Prohibitive Computational Complexity

Most higher-order tensor problems are NP-hard [Hillar & Lim, 2013].



Fortunately, the tensors sought in statistical and machine learning applications are often specially structured:

- Low-rankness
- Sparsity
- Non-negativity
- ...

Breaking previous limits

My group is developing a framework of statistical models, efficient algorithms, and fundamental theory to analyze large-dimension large-scale tensor/matrix data.

For potential Ph.D students

Good niche if you are

- comfortable with mathematical statistics, probability theory, and optimization.
- enthusiastic about using your quantitative skills to advance our understanding in sciences.
- actively interested in learning about nearby research areas (which areas are up to you).

As an advisor, I will

- provide you the skills, experiences, and connections that make you excel in your future career.
- be a demanding advisor, probably more demanding than most :)
- not be a good fit for you if you are not broadly interested in attending talks outside your thesis topics, or if you do not like being technical in a serious way.

If the above sounds exciting to you, drop by my office and talk with me!