nexidia

Dale Smith, Ph.D.
Nexidia, Inc.

Tensor Decompositions and Machine Learning

What is a Tensor?

- We're moving from linear algebra to multilinear algebra…

- We know about vectors and matrices (linear transformations) from Linear Algebra. But tensors are not so familiar. A Rubik's Cube can help with visualization…and so can What's a Tensor? via YouTube!

- Think of a hypercube in your data warehouse.

What is a Tensor Decomposition?

nexidia

- Think of a hypercube in your data warehouse – can you do a decomposition into lower-rank objects that reveal hidden features or hierarchies?

- A rank-2 tensor decomposition – the Singular Value Decomposition – computes $A = UDV^T$. Principal Components Analysis approximates the original matrix by reducing the rank.

- For rank higher than 2 there are two types of decompositions: a sum of rank-1 tensors and a decomposition into orthonormal subspaces. Oh, and there is a multilinear PCA as well…

Tensor Decompositions in Machine Learning

nexidia
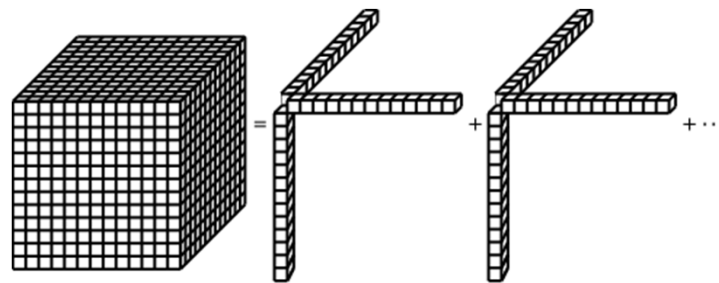
- Feature discovery via hierarchical structures in high-dimensional data sets.

- Recommender systems – similarity measures between users and items.

- Time-evolving network expressed as a 3-tensor.

- Topic modeling to extract hidden topics in a document corpus.

- Multilinear subspace learning – a technique of dimensionality reduction.

- Tensor methods are fast, scalable, and in many cases are embarrassingly parallel.

Tensor Decompositions in Machine Learning

nexidia

- Anima Anandkumar, et. al.: *Fast Detection of Overlapping Communities via Online Tensor Methods*, JMLR 2014 and *Tensor Methods for Large-Scale Machine Learning*, Strata+Hadoop, February 2015.

- Computational issues with estimating latent variable models, e.g. topic models.

  – Local optima.

  – Slow convergence.

  – Markov Chain Monte Carlo can fail.

- Tensor decomposition approaches *guarantee* finding the global optimum.

- Observed speedups of 100x over the Expectation-Maximization and other estimation methods using data from Facebook and Yelp restaurant reviews.

Tensor Decomposition Software

MATLAB	C++	R	Python
Tensor Toolbox	TensorDecomposition4Topic Modeling	rTensor	TensorToolbox
Tensorlab	TH++		PyTensor
			scikit-tensor

A Few References

- Survey papers with excellent visualizations of tensor decompositions

  – *Tensor Decompositions for Signal Processing Applications*, Cichocki, De Lathauwer, et. al., IEEE Signal Processing Magazine, March 2015 (arXiv.org version).

  – *Tensor Decompositions and Applications*, Tamara G. Kolda and Brett W. Bader, SIAM Review, 2009 (Sandia pdf report).

- 2009 Workshop Report: *Future Directions in Tensor-Based Computation and Modeling*.

- Tutorials

  – *Tutorial on Spectral and Tensor Methods for Guaranteed Learning*.

  – *Mining Large Time-evolving Data using Matrix and Tensor Tools*.

  – *Mining and Forecasting of Big Time-Series Data*.

Recent Events

- Blog Posts and Podcasts
  - Anima Anandkumar, *Tensor Methods for Large-Scale Machine Learning*, Strata+Hadoop, February 2015.
  - Ben Lorica, *Let's build open source tensor libraries for data science*, O'Reilly Radar, March 17, 2015.
  - *The tensor renaissance in data science* and associated podcast, O'Reilly Radar, May 7, 2015.
- *Tensor Decompositions in Smart Patient Monitoring*, SIAM News, September 2015.
- SIAM Conference on Applied Linear Algebra, October 26th – 30th, Atlanta.
  - Sessions on Tuesday the 27th.
  - Keynote on Friday the 30th.
- Workshop on Tensor Decompositions and Applications, January 2016.

Takeaways

- Do you have…

    – Time series of matrices or cubes?

    – Multiple aspects: time, location, type?

    – Time-evolving graphs?

    – Primary source data with multiple metadata elements?

- …then you may want to consider tensor decomposition methods to estimate models.