Summary and closing thoughts

Lecture 25
April 25, 2019

Looking back: topics covered

- Intro to randomized algorithms, probabilistic inequalities, hashing, limited independence
- Streaming algorithms: Morris counter, sampling, frequency moments (distinct elements (ℓ_0), F_2, heavy hitters (ℓ_∞))
- Sketching in streaming: Count-Min, Count and applications including sparse recovery
- Dimensionality reduction, JL Lemma, Subspace embeddings
- Similarity estimation, nearest neighbor search via LSH
- Quantiles, sorting, selection
- Graph streaming
- Coresets
- Faster numerical linear algebra: approximate matrix multiplication, regression, and low-rank approximation via sampling and subspace embeddings

Related topics that we did not have time for

- Clustering
- More advanced topics in frequency moments, graph streaming, coresets
- More advanced topics in NLA and Compressed Sensing
- Sparse Fourier Transform
- Large scale graphs and networks
- Lower bounds for streaming, sketching, JL
- Parallel models of computing for large scale data: Map Reduce, GPUs, and related
- Big data aspects of ML (sampling, kernel methods, mixtures of Guassians, matrix completion, stochastic gradient descent, etc)
- Practical aspects

Looking forward

Course evaluation/ICES forms: give me constructive feedback

- Topics: likes, dislikes, others
- Expectations: what were you hoping to learn and what did you learn and whether you want some thing else.
- Evaluation: Homeworks and exam and project. More, less, different, easier, harder, etc.
- Teaching style, pace, slides/notes, etc

Project report: due on Friday, May 3rd on Gradescope

Next time: I will teach on the board (black or white)!

Thanks!