Research Statement

Niranjan Uma Naresh

Big-data analysis and large-scale machine learning have become a central part of many disciplines. My research revolves around elegant modeling of problems of practical interest, understanding the fundamental principles of statistical machine learning, designing scalable algorithms with guarantees, and analyzing their computational and information-theoretic properties rigorously. The application areas that are impacted include domains such as computer vision, natural language processing, speech recognition, quantum physics and computational biology.

Below I summarize the topics I have worked on during my PhD followed by connections and discussions regarding my future research interests.

Probabilistic models: The framework of graphical models provides a unified way to perform inference and answer queries about various real world phenomena by modeling and estimating dependencies between random variables. As an alternative to likelihood-based estimation, I used moment-based methods [2, 3] for learning models such as community-formation, latent tree and topic models. The common theme is that we employ novel stochastic tensor methods to do the estimation efficiently. The theoretical advantage of over other approaches is that moment-based methods enjoy certain statistical and computational guarantees. We also provide state-of-the-art empirical results via implementations on GPUs and CPUs for large-scale machine learning.

Kernel methods: In a collaborative project [5] with the Department of Chemistry at UC Irvine, I used both unsupervised and supervised kernel methods for solving problems in density functional theory. I implemented and investigated the performance of different kernels in approximating the kinetic energy of non-interacting particles in a box.

High-performance computing: Thanks to the easy availability of computational resources, we can now handle large-scale learning problems easily. However, optimizing linear and multilinear algebraic operations is essential for implementing many learning algorithms. To this end, in [8], we use novel extended BLAS kernels to compute tensor contractions. Using thorough benchmarking simulations on the CPU and the GPU, I show significant speedups over existing state-of-the-art libraries for multilinear algebra.

Deep learning and feature extraction: Extraction of meaningful features is essential for good performance in multi-class classification. In [4], the novel feature learning method involves higher-order score functions computed for certain models from data. I also investigated the autoencoders, a popular unsupervised deep learning model, extensively. I evaluated this approach using standard accuracy measures, obtaining competitive classification performance on real-world datasets such as MNIST and Paraphrase data, even with reduced number of training samples.

In [6], we investigate a deterministic approach for extracting geometric features from polygons using local convex hull computations. I developed a novel tree-based data structure for efficient querying and a polynomial time algorithm for matching complementary features. This research suggests that combining geometry-based and content-based information retrieval is likely to confer better performance for various learning tasks.

Non-convex optimization: In [7], I worked on the robust PCA problem. The goal here is to decompose a given matrix as a low-rank plus a sparse matrix. The problem is inherently non-convex and the existing guaranteed solution techniques relied on convex relaxation. We show that an alternating minimization algorithm using non-convex projections yields guaranteed recovery of the low-rank and sparse matrices with better computational and sample complexities. I applied the algorithm to real-world datasets, mainly in the vision domain, for the task of background-foreground separation. We obtain better results with much lesser time and space complexity.

Matrix factorization problems are relatively better understood than tensor problems. While tensor algorithms have nice recovery guarantees, especially in learning probabilistic models, optimization problems
involving them are usually challenging due to the exponential number of saddle points that tensors have. In [1], we study the tensor analogue of the matrix robust PCA problem. The results allow us to tolerate more noise while maintaining faithful recovery.

Current and future research directions

Optimal estimation: In statistical decision theory and Bayesian estimation, one notion of optimality is often characterized by trying to answer the question, “what is the best possible estimation one could achieve in the worst possible scenario?” To this end, I’m interested in characterizing the information-theoretic hardness arising in estimation problems of practical interest including learning latent variable models and time-series models.

Topological data analysis: The study of random graphs is a classical area with far-reaching connections and implications in mathematics and computer science. Generalizing this structure to topological objects such as random complexes gives us better understanding of the shape of data. I’m interested in unraveling connections between topology, probability and optimization while understanding the computational hardness involved. Some applications of practical interest are feature extraction and manifold learning.

Stochastic optimization: Designing scalable algorithms is crucial to extract knowledge by processing massive amounts of data. To this end, I’m interested in iterative numerical methods for stochastic optimization. The advantage of such algorithms is the ability to handle streaming data which leads to efficient time and space complexity per iteration. Despite this gain, a challenge in the online setting is to guarantee fast convergence, which I aim to address.

Deep learning: While neural networks hold the state-of-the-art empirical performance credits and have shown tremendous promise in many domains, much is yet to be understood about them. To this end, I’m interested in understanding their fundamental properties by trying to answer the question, “why deep learning models work the way they work?” In addition, I’m also interested in applying neural networks to push the state-of-the-art results in various problems of practical importance in artificial intelligence.

References


