

Data-Pooling in Optimization

ABSTRACT - Managing large-scale systems often involves simultaneously solving thousands of potentially unrelated stochastic optimization problems, each with limited data. Intuition suggests decoupling these unrelated problems and solving them separately. We propose a novel data-pooling algorithm that disproves this intuition and shows that combining data across problems often outperforms decoupling, even when there is no a priori structure linking the problems, and data are drawn independently. Our approach does not require strong distributional assumptions and applies to constrained, possibly non-convex optimization problems such as vehicle-routing, economic lot-sizing or facility location. We compare and contrast our results to a similar phenomenon in statistics (James-Stein Phenomenon) and show that, unlike the classical statistical setting, the potential benefits of data-pooling, in general, depend strongly on the problem structure, and, in some cases, data-pooling offers no benefit. We prove that as the number of problems grows large, our method learns if pooling is necessary and the optimal amount to pool, even if the expected amount of data per problem is fixed and bounded. We further show that pooling offers significant benefits over decoupling when there are many problems, each of which has a small amount of relevant data. We demonstrate the practical benefits of data-pooling using real data from a chain of retail drug stores in the context of inventory management. Overall, we see data-pooling as a thus far underutilized phenomenon that might be exploited to design more effective algorithms for data-driven optimization.



Dr. Vishal Gupta

Assistant Professor
Dept. of Data Science & Operations
University of Southern California

SPEAKER BIO - Vishal Gupta is an Assistant Professor of Data Sciences and Operations at the USC Marshall School of Business. Vishal earned his B.A. in Mathematics and Philosophy from Yale University, graduating Magna Cum Laude with honors, and completed Part III of the Mathematics Tripos at the University of Cambridge with distinction. After Cambridge, Gupta spent four years working as a “quant” at Barclays Capital focusing on commodities modeling, derivatives pricing and risk management. Realizing how much he missed research working towards a larger mission of impact, he left private industry to complete his Ph.D. at MIT in 2014.

Vishal's research focuses on decision-making in data-scarce environments. His work has been recognized by various awards, including as a finalist in the Pierskalla Best Paper Award, the CHOM Best Paper Award, the Service Science Best Paper Award and the Nicholson Best Student Paper award.

USC Viterbi
School of Engineering
*Daniel J. Epstein Department of
Industrial and Systems Engineering*

TUESDAY, SEPTEMBER 3, 2019

3:30 PM – 4:50 PM

USC ANDRUS GERONTOLOGY CENTER (GER), Room 206