2023 INFORMS Annual Meeting

AI4OPT Schedule of Events and Awards

October 15-18, 2023 || Phoenix, AZ

Search INFORMS Annual Meeting Program Schedule, here.

Find full program, here.

Oct 15 – Sunday

Optimizing Truck Fleet Scheduling for Fuel Deliveries

Time/Location

8 AM - 9:15 AM; CC-ECC 208A

Session

SA75 - Optimization of Logistics Delivery Systems

Authors

Vahid Eghbal Akhlaghi[1], Hongzhao Guan[1], Jason Lu[1], Pascal Van Hentenryck[2], [1]Georgia Institute of Technology, Atlanta, GA, [2]ISyE Georgia Tech, Atlanta, GA, Contact: vahid.eghbal@gatech.edu

Abstract

In this paper, we optimize the delivery schedules for a fleet of trucks with specified availability time windows to meet fuel replenishment needs at various sites. A mixed-integer programming model is proposed, incorporating a two-stage optimization algorithm: firstly, ensuring sufficient deliveries to maintain site operations, and secondly, enhancing driver shift and truck utilization while minimizing total traveled distance. The problem is refined via techniques to make it computationally tractable. A case study in Estonia provides valuable insights, including cost savings, improved customer satisfaction, enhanced business resilience, and reduced fuel consumption and emissions. This approach offers a compelling alternative to current manual drag-and-drop tools, underscoring the potential for automation in fleet scheduling.

Congratulations to AI4OPT members who are this year's winners of the INFORMS Optimization Society Prizes:

Khachiyan Prize: **Renato Monteiro**

- A Journey Through Many Exciting Topics of Continuous Optimization
- Time/Location: 10:45 AM – Noon; CC-North 224A
- Session: Session SB38 - Optimization Society's Award Session I

Farkas Prize: **Alper Atamturk**

- Strengthening MIP Formulations of Hybrid Model Predictive Control
- Time/Location: 10:45 AM – Noon; CC-North 224A
- Session: Session SB38 - Optimization Society's Award Session I

OR/AI: Integrating Operations Research with Artificial Intelligence

Time/Location
12:45 PM – 2 PM; CC-North 226A

Panel Session
SC42 - OR/AI: Integrating Operations Research with Artificial Intelligence

Abstract
OR/AI: Integrating Operations Research with Artificial Intelligence Ahmed Abbasi, University of Notre Dame, Notre Dame, IN This panel session entitled "OR/AI: Integrating Operations Research with Artificial Intelligence" will feature panelists conducting research at the intersection of operations research (OR) and artificial intelligence (AI). The panel will discuss research and funding opportunities at the intersection of OR and AI.

Panelist
Ramayya Krishnan, Carnegie Mellon University, Pittsburgh, PA
Bistra Dilkina, USC, Los Angeles, CA
Nan Zhang, University of Florida, Gainesville, FL
Cole Smith, Syracuse University, Syracuse, NY

K-12 Education Outreach in a Post-COVID World

Time/Location
2:15 PM - 3:30 PM; CC-West 212B

Panel Session
SD81 - K-12 Education Outreach in a Post-COVID World

1. Session Chair: Zihan Zhang, Georgia Tech, Atlanta, GA
   In March 2020, the start of the COVID-19 pandemic brought a sea change to K-12 education in the U.S. Every school district was forced into virtual learning overnight. Now that the pandemic is fading into the rear-view mirror, those who value the success of K-12 students have the opportunity to begin anew K-12 education outreach projects. Successful, sustained programs will simultaneously work within longstanding constraints facing public education in this country and leverage new opportunities resulting from pandemic response. Lessons learned from first-hand experience will be discussed.
2. Panelist:

Afrooz Jalilzadeh, The University of Arizona, Tucson, AZ
Pascal Van Hentenryck, ISyE Georgia Tech, Atlanta, GA
Mary Ogidigben, Pennsylvania State University, State College, PA
Paul Hand, Northeastern University, Boston, MA
Neil Desnoyers, Saint Joseph's University, Philadelphia, PA

Bucketized Active Sampling for Acopf Optimization Proxies

Time/Location
4 PM - 5:15 PM; CC-North 227A

Session
SE45 - Machine Learning and Optimization in Power Systems

Authors
Michael Klamkin, Georgia Tech, Atlanta, GA, Contact: klam@isye.gatech.edu

Abstract
This paper considers optimization proxies for Optimal Power Flow (OPF), i.e., machine-learning models that approximate the input/output relationship of OPF. Their training requires significant data, each instance necessitating the (offline) solving of an OPF. To meet market-clearing requirements, this paper proposes Bucketized Active Sampling (BAS), a novel active learning framework that aims at training the best OPF proxy within a time limit. BAS partitions the input distribution and uses an acquisition function to determine where to sample. Numerical experiments are conducted on large power grids, benchmarking different BAS strategies against classical approaches wherein training data is generated offline. Overall, BAS achieves the same or better accuracy but does so in half the compute time.

End-To-End Feasible Optimization Proxies for Large-Scale Economic Dispatch

Time/Location
4 PM - 5:15 PM; CC-North 227A

Session
Session SE45 - Machine Learning and Optimization in Power Systems

Authors
Wenbo Chen[1], Mathieu Tanneau[1], Pascal Van Hentenryck[2], [1]Georgia Tech, Atlanta, GA, [2]ISyE Georgia Tech, Atlanta, GA, Contact: wchen616@gatech.edu

Abstract
The paper proposes a novel End-to-End Learning and Repair (E2ELR) architecture for training optimization proxies for economic dispatch problems. E2ELR combines deep neural networks with closed-form, differentiable repair layers, thereby integrating learning and feasibility in an end-to-end fashion. E2ELR is also trained with self-supervised learning, removing the need for labeled data and the solving of numerous optimization problems offline. E2ELR is evaluated on industry-size power grids with tens of thousands of buses using an economic dispatch that co-optimizes energy and reserves. The results demonstrate that the self-supervised E2ELR achieves state-of-the-art performance, with optimality gaps that outperform other baselines by at least an order of magnitude.

Multiperiod Stochastic Security-Constrained Economic Dispatch: A Systematic Study

Time/Location
4 PM - 5:15 PM; CC-North 227A

Session
SE45 - Machine Learning and Optimization in Power Systems

Authors
Haoruo Zhao, Mathieu Tanneau, Pascal Van Hentenryck, Georgia Institute of Technology, Atlanta, GA, Contact: hzhao306@gatech.edu

Abstract
US Independent System Operators co-optimize energy and reserve dispatches every five minutes to minimize operating costs while meeting physical and regulatory constraints. However, due to the growing penetration of renewable energy generation, the current single-period, deterministic, security-constrained economic dispatch (SCED) formulations struggle to manage operational uncertainty effectively. This project proposes a multi-period stochastic look-ahead (SLAD) formulation to explicitly consider uncertainty in real-time markets. An accelerated Benders' decomposition is utilized to solve the formulation efficiently. The approach is evaluated on an industry-size transmission grid, and the numerical results demonstrate that SLAD significantly improves reliability and reduces costs, saving approximately 2% on high net load days compared to SCED.

Online Risk Assessment with Optimization Proxies

Time/Location
4 PM - 5:15 PM; CC-North 227A

Session
SE45 - Machine Learning and Optimization in Power Systems

Authors
Mathieu Tanneau[1], Pascal Van Hentenryck[2], [1]Georgia Tech, Atlanta, GA, [2]ISyE Georgia Tech, Atlanta, GA

Abstract
The increased penetration of renewable energy and distributed energy resources is fueling an increase in operational uncertainty in modern power grids.To manage this uncertainty, grid operators must continuously monitor and manage risk. This, in turn, requires tools that can quantify risk in a real-time fashion.
This talk presents how optimization proxies, i.e., machine learning models that approximate the input-output mapping of an optimization problem, offer a scalable tool for real-time risk assessment. Numerical results on an industry-size system will be presented.

Oct 16 – Monday

Bus Line Design for On-Demand Multimodal Transit Systems

Time/Location

8 AM - 9:15 AM; CC-ECC 208A

Session
SE45 - On Bus Transportation System Planning

Authors
Anthony Trasatti[1], Beste Basciftci[2], Pascal Van Hentenryck[1], [1]ISyE Georgia Tech, Atlanta, GA, [2]University of Iowa, Iowa City, IA, Contact: anthony.j.trasatti@gmail.com

Abstract
Many transit agencies have started experimenting with adding on-demand services in attempt to maximize their service quality, accessibility, and coverage within a fixed budget, using shuttles to alleviate their first and last mile issues as part of On-Demand Multimodal Transit Systems (ODMTS). This work presents a novel mixed-integer program (MIP) formulation to incorporate bus line design into the network design problem for On-Demand Multimodal Transit Systems that allow the model to accurately capture wait time and transfer costs in addition to travel time and vehicle costs. To solve large-scale instances, a two-stage reformulation is presented where the first-stage problem decides bus arcs to open and which arcs immediately follow and the second-stage problem decides the multimodal path for each individual trip.

Value-Oriented Loss Function Tuning for Timeseries Forecasting in Energy Markets

Time/Location 8 AM - 9:15 AM; CC-North 126B

Session
MA16. Emerging Topics in Artificial Intelligence: Methods and Applications

Author
Ruben Smets[1], Jean-François Toubeau[1], Mihaly Dolanyi[2], Kenneth Bruninx[3], Erik Delarue[1], [1]Katholieke Universiteit Leuven, Leuven, Belgium; [2]Université de Mons, Mons, Belgium; [3]TU Delft, Delft, Netherlands. Contact: ruben.smets1@kuleuven.be

Abstract
In complex markets with a high amount of uncertainty, such as balancing markets, significant forecast errors are inevitable. Recently, the paradigm of value-oriented forecasting is gaining traction. While these techniques typically require a specific implementation dependent on the application, we propose a generally applicable value-oriented methodology for training time series forecasters. We achieve this by introducing a generalized loss function with a continuous exponent and variability component, and leveraging the maximum downstream value as the selection criterion in the hyperparameter tuning step. The proposed methodology is tested on the optimal control of different types of energy storage systems in the Belgian balancing market. We show that our methodology outperforms traditional benchmarks with 13% to 106% in terms of out-of-sample profit.

Plenary Panel: Harnessing the Data Revolution in Supply Chains

Time/Location
9:45 AM - 10:35 AM; CC-North 301ABC

Panel Session
MP01. Plenary Panel: Harnessing the Data Revolution in Supply Chains

1. "Harnessing the Data Revolution in Supply Chains" by Anne G. Robinson, Kinaxis, Ottawa, ON, Canada, discusses the vast scope and impact of supply chains, which generate substantial amounts of data, representing a treasure trove of insights. While some areas, such as inventory optimization, demand forecasting, production planning, scheduling, and transportation, have well-established techniques, recent technical advances in data storage and computation have opened up new possibilities. The internet of things generates signals throughout the supply chain, and there is a wealth of external data sources, including microeconomic data, weather, consumer demand, and social media sentiment. Machine learning can process these signals to enhance supply chain insights, and algorithmic research offers innovative ways to integrate machine learning and optimization. These changes provide an unparalleled opportunity to leverage the data revolution in supply chain management, particularly in the face of ongoing disruptions that continue to make supply chains a top priority in boardroom discussions.
2. Panelist:
    - Derrick Fournier, Bristol Myers Squibb, Philadelphia, PA
    - **Pascal Van Hentenryck, ISyE Georgia Tech, Atlanta, GA**
    - Kelly Thomas, Worldlocity, Ottawa, ON, Canada
    - Feryal Erhun, University of Cambridge, Cambridge, United Kingdom

---

Optimization Proxies for Continuous and Discrete Optimization

Time/Location
10:45 AM - 12:00 PM; CC-North 227A

Session
MB45 - Machine Learning for Discrete and Global Optimization 1

Author
Pascal Van Hentenryck, ISyE Georgia Tech, Atlanta, GA

Abstract
Optimization proxies aim at replacing an optimization model by the combination of a machine-learning model and a repair step. They have been shown to be particularly useful in real-time applications in a variety of areas including energy systems, supply chains, and transportation systems. This talk will provide an overview of the science, engineering, and applications of optimization proxies.

Optimization-Based Learning for Tactical Load Plan Modification in Trucking Service Networks

Time/Location
12:45 PM - 2:00 PM; CC-North 225A

Session
MC40 - Advances in Optimization for Logistics Network Design

Author

Ritesh Ojha[1], Wenbo Chen[1], Hanyu Zhang[1], Reem Khir[2], Pascal Van Hentenryck[3], Alan Erera[4], [1]Georgia Institute of Technology, Atlanta, GA, [2]Purdue University, West Lafayette, IN, [3]ISyE Georgia Tech, Atlanta, GA, [4]Georgia Tech, Atlanta, GA, Contact: riteshojha8@gatech.edu

Abstract
A critical service network design challenge for package carriers is the so-called load planning problem. Load planning refers to decisions about how many trailers or container loads (perhaps of different types) to plan for dispatch over time between pairs of terminals. Such planned loads are the transportation capacity of the network. Another key component to the consolidation transportation plan are the decisions regarding which package volumes to assign into planned loads. This work considers dynamic optimization of load plans given fixed primary and alternate flow planning decisions. An extensive computational study on real-life instances, provided by our industry partner, including few of their largest terminals in their network, reports an average of 8-12\% potential improvement in trailer capacity utilization at the respective terminals.

Bayesian Risk-Averse Q-Learning with Streaming Data

Time/Location
12:45 PM - 2 PM; CC-North 224A
Session
MC38. Data-driven Decision Making: Robustness, Learning, and Optimization

Authors
Yuhao Wang, Ph.D., supervised by Prof. Enlu Zhou.

Yuhao Wang[1], Enlu Zhou[2], [1]georgia institute of technology, atlanta, GA, [2]ISyE Georgia Tech, Atlanta, GA, Contact: yuhaowang@gatech.edu

Abstract
We consider an off-policy reinforcement learning problem with streaming data, where we need to solve a discounted infinite-horizon Markov decision process (MDP) with an unknown transition model. We adopt a Bayesian risk MDP (BRMDP) formulation, which uses Bayesian posterior to estimate the transition model and imposes a risk functional on rewards at each stage with respect to the posterior distribution to account for the misspecification of the transition model. We develop a multi-stage Bayesian risk-averse Q-learning algorithm by solving BRMDP with dynamically updated Bayesian posterior on the transition model. The proposed algorithm learns a risk-averse policy that is optimal conditional on all the observed data. In particular, the learned policy converges to the true optimal policy if we have an infinite number of observed data for each state-action pair.

Data-Driven Distributionally Robust Vehicle Routing for Heterogeneous Heavy-Duty Trucks

Time/Location
2:15 PM - 3:30 PM; CC-North 221B

Session
MD32 - Data-driven Operation of Emerging Mobility Systems

Authors
On Behalf of **Scott Moura** – Ruiting Wang, University of California Berkeley, Berkeley, CA, Contact: rtwang@berkeley.edu

Abstract
In this work, we study the routing of electric trucks through an application of distributionally robust optimization (DRO) for route planning and dispatch. This approach aims to minimize the cost of operation for the fleet and considers the variability in energy consumption due to uncertain road conditions, traffic, weather, and driving behavior. Furthermore, we augment the distributionally robust energy-minimizing vehicle routing problem by learning the energy efficiency distribution over a horizon. We show that convergence to the true distribution is achieved while learning from samples taken from vehicles in operation on the network. Numerical experiments are conducted to validate this method and to compare it with the deterministic model.

Transforming Healthcare Using Artificial Intelligence

Time/Location
4 PM - 5:15 PM ; CC-North 127C
Session
ME19 - Transforming Healthcare Using Artificial Intelligence
Contact: pollack9@mit.edu

Authors
***Dan Molzahn**'s Ph.D. student **Ryan Piansky** who will attend on Dan's behalf and present work at the intersection of the ethical AI thrust and the energy systems thrust (incorporating fairness in resilient power grid planning problems).*

Swati Gupta1, Alyssa Kody2, Daniel K. Molzahn3, **Ryan Piansky**3, Madeleine Pollack1, 1Massachusetts Institute of Technology, Cambridge, MA, 2Argonne National Laboratory, Lemont, IL, 3Georgia Institute of Technology, Atlanta, GA, Contact: pollack9@mit.edu

Abstract
In this work, we model emergency power shutoffs to mitigate wildfire risk on a synthetic transmission network in Texas. To eliminate both wildfire risk and power loss, one can "underground" a transmission line at a high cost per mile. While considering multiple objectives, we explore how a fixed budget for line undergrounding is distributed across the network and which populations benefit most from the investment. Due to the intersection of network characteristics from the power grid, the spatiotemporal-varying distribution of wildfire risk, and demographic patterns; many vulnerable communities are more likely to experience power loss, according to our model. This necessitates the incorporation of "equity" objectives into the optimization model, and we study their impact on the structure of infrastructure investment decisions.

Network Flows and Minimum Cuts in Ranking, Clustering, Machine Learning, Imaging and Diversity Problems

Time/Location
5:25 PM - 6:15 PM; CC-North 120A

Session
MK02 - Keynote: Network Flows and Minimum Cuts in Ranking, Clustering, Machine Learning, Imaging and Diversity Problems

Author
Dorit Simona Hochbaum, University of California-Berkeley, Berkeley, CA

Abstract
A significant category of integer programming problems, called "monotone" (IPM, Integer Programming Monotone), finds its solution by employing a minimum cut algorithm on an associated graph. Within this category, numerous well-known clustering problems fall under the umbrella of ratio IPM problems. Remarkably, it has been demonstrated that these and the respective ratio problems can all be resolved using a parametric cut procedure, which exhibits the same computational complexity as a single minimum cut procedure. Other applications of IPM include drug ranking, the identification of active neurons in calcium imaging movies, and a wide array of machine learning and classification tasks.
For several problems classified as NP-hard, the incorporation of modeling flexibility transforms them into efficiently solvable IPM problems. This modeling flexibility is exemplified in scenarios such as the threat detection problem, co-segmentation problem, and text summarization problem. In the context of text summarization, for instance, this flexibility is manifested by replacing the concept of minimum similarity with maximum dissimilarity.
A specialized subset within the realm of NP-hard problems is the "budgeted" IPM problems, which incorporate an additional budget constraint. In the case of budgeted IPM problems, it is established that the entire efficient frontier can be generated through the parametric cut procedure. Furthermore, the breakpoints within the efficient frontier are optimal and facilitate the derivation of high-quality solutions for the respective problem. These problems find applications in diverse fields, including facility dispersion, quadratic knapsack, maximum diversity, and text summarization.
The Markov Random Field (MRF) problem originally emerged in the context of machine vision. For convex deviation functions and bilinear separation, the MRF can be efficiently solved through a parametric cut algorithm with a "best possible" runtime. MRF with general convex functions is equivalent to the convex dual of the minimum cost network flow problem, which is also solved efficiently. Applications of the MRF model span various domains, encompassing image segmentation, customer segmentation, credit risk assessment for countries, student paper competitions, semiconductor yield prediction, isotonic regression, general ranking, and recently, the aggregation of voter preference rankings.

Oct 17 – Tuesday

Forecasting and Asset Bundling for Renewable Energy

Time/Location
8 AM - 9:15 AM; CC-North 230

Session
TA52 - Optimization and Machine Learning in Power Systems

Author
Hanyu Zhang[1], Mathieu Tanneau[2], Pascal Van Hentenryck[1], [1]ISyE Georgia Tech, Atlanta, GA, [2]Georgia Tech, Atlanta, GA

Abstract
Renewable energy raises new challenges in forecasting and market-clearing operations for power systems. Their volatility and the sheer number of distributed renewable assets create significant challenges that have not been encountered before. This presentation reviews recent progress in forecasting wind and solar energy sources and in preparing these forecasts for use in stochastic optimization. In particular, the presentation reviews the use of transformer architectures for forecasting large-scale time series arising in renewable energy, as well as novel optimization algorithms for bundling time series.

Path-Based Formulations for the Design of On-Demand Multimodal Transit Systems with Latent Demand Awareness

Time/Location
10:45 AM - 12:00 PM; CC-North 225B

Session
TB41 - Emerging Applications in Transportation and Logistics

Author
Hongzhao Guan[1], Beste Basciftci[2], Pascal Van Hentenryck[3], [1]Georgia Institute of Technology, Atlanta, GA, [2]University of Iowa, Iowa City, IA, [3]ISyE Georgia Tech, Atlanta, GA, Contact: hguan7@gatech.edu

Abstract
This study investigates capturing the latent demand in on-demand multimodal transit systems, namely the ODMTS Design with Adoptions problem (ODMTS-DA) by proposing a new path-based optimization model, called P-Path. This model addresses the computational difficulties from previous work that utilizes bilevel programs. The key idea of P-Path is to enumerate two specific sets of paths which capture the essence of the choice models associated with the adoption behavior of riders. Then ODMTS-DA can be formulated as a single-level MIP model. P-Path is evaluated on two comprehensive case studies. Results show that P-Path solves the mid-size instances in a few minutes, bringing more than two orders of magnitude improvements compared to previous approach. Moreover, the results show that P-Path can solve large-scale ODMTS-DA instances to optimally in a few hours or in a few days.

Portfolio Approximations and Fairness in Combinatorial Optimization

Time/Location
10:45 AM - 12:00 PM; CC-North 126B

Session
TB16 - Fairness in Optimization Models from Operations Management

Author

Swati Gupta[1], **Jai Moondra**[2], Mohit Singh[3], [1]ISyE Georgia Tech, Atlanta, GA, [2]Georgia Institute of Technology, Atlanta, GA, [3]Georgia Institute of Technology, Atlanta, GA, Contact: jmoondra3@gatech.edu

Abstract
Inspired by fairness notions, we look at ordered optimization for discrete optimization problems on real vector spaces (such as all path vectors in a graph), where we are given a weight vector and the cost of a vector in the domain is the dot product of the weight vector with the sorted domain vector. This generalizes the classic min-sum (such as shortest path) and the min-max (such as makespan minimization) objectives. We study solution portfolios for ordered optimization where we ask for a small set of vectors in a fixed domain that are approximately optimal for each ordered optimization problem, generalizing the special case when one domain vector is approximately optimal for all ordered optimization objectives. We show various upper and lower bounds for minimum portfolio sizes for various problems including shortest paths, matchings, routing, and parallel scheduling.

Screening Rules for L0-L2 Regression

Time/Location
12:45 PM - 2:00 PM; CC-North 221C

Session
MC33 - Discrete Optimization for Trustworthy Machine Learning

Author
Anna Deza, Alper Atamturk, IEOR, UC Berkeley, Berkeley, CA, Contact: annadeza@berkeley.edu

Abstract
Logistic regression with a large number of features compared to available labels presents numerous challenges for learning. We present screening rules that safely remove features from the sparse logistic regression with L0-L2 regularization before solving the problem. The screening rules are based on the Fenchel dual of strong conic relaxations of the sparse logistic regression problem. Numerical experiments with real and synthetic data suggest that a high percentage of the features can be effectively and safely removed apriori, leading to substantial speed-up in the computations.

Data-Driven Ranking and Selection

Time/Location
4 PM - 5:15 PM; CC-North 124B

Session
TE12 - Simulation Flash Talks

Authors
Yuhao Wang, Ph.D., supervised by Prof. Enlu Zhou

Enlu Zhou, ISyE Georgia Tech, Atlanta, GA

Abstract
We consider ranking and selection with streaming input data, which arrive in batches, possibly of varying sizes, sequentially over time. We develop a sequential elimination framework for the fixed confidence setting, where the goal is to achieve a specified probability of correct selection (PCS) with as few simulation replications as possible. We also develop procedures based on optimal computing budget allocation for the fixed budget setting, where the goal is to achieve a PCS as high as possible with a given simulation budget.

Oct 18 – Wednesday

Optimization Proxies in the Supply Chain for Semiconductor Manufacturing

Time/Location
8 AM - 9:15 AM; CC-North 224A

Session
WA38 - Semiconductur Industry and Batch Production

Author
Jorge Huertas[1,2], Ashish Nemani[2], Zhenying Zhao[2], Pascal Van Hentenryck[1], [1]ISyE Georgia Tech, Atlanta, GA, [2]Intel, Chandler, AZ, Contact: huertas.ja@gatech.edu

Abstract
In the supply chain for semiconductor manufacturing, many processes at various stages and locations transform the silica wafers into multiple components to finally produce chips. To generate the production plan along this supply chain, a master production scheduler tool executes different optimization models sequentially. Hence, performing sensitivity analyses for planning purposes becomes a challenge, as it is necessary to execute the sequential optimization multiple times to quantify the impact of variations in the input parameters. To overcome this challenge, we combine optimization and machine learning to produce an optimization proxy of the sequential optimization, allowing quick sensitivity analyses for planning purposes.

Path of Solutions for Non-Smooth Fused Lasso

Time/Location
8 AM - 9:15 AM; CC-ECC 211A

Session
WA78 - Optimization for Advacing Machine Learning Methods

Author
Cheng Lu[1], **Tor Nitayanont**[2], Dorit Simona Hochbaum[2], [1]Apple, Cupertino, CA, [2]University of California-Berkeley, Berkeley, CA, Contact: torpong_nitayanont@berkeley.edu

Abstract

In a fused lasso problem on a sequence data, the objective consists of two competing terms: the separation term and the deviation term. The two terms are balanced via a tradeoff parameter. Selecting a value for the tradeoff parameter is critical. However, we often do not know the quality of a certain tradeoff value until we obtain the solution. Here, we propose a minimum cut-based method that solves for all tradeoff values at once. This method eliminates the need to pre-specify tradeoff values since it solves the problem for all values that yield different solutions. Moreover, we prove that, for the L1 deviation, the number of solutions across all tradeoff values is at most 2n^3 for n variables. In fact, experimental results show a linear bound. For a general convex deviation function, the bound is 2n^3(U/ε) where ε is the solution accuracy and U is the range of variables.

Stochastic First-Order Algorithms for Constrained Distributionally Robust Optimization

Time/Location
9:30 AM - 10:45 AM; CC-North 222A

Session
WB34 - First-Order Methods for Challenging Constrained Optimization Problems

Authors
Hyungki Im[1], Paul Grigas[2], [1]UC, Berkeley, Berkeley, CA, [2]UC Berkeley, Berkeley, CA

Abstract
We consider distributionally robust optimization (DRO) problems with multiple expectation constraints. We propose a generic stochastic first-order meta-algorithm, wherein the decision variables and uncertain distribution parameters are each updated separately by applying stochastic first-order methods. We then specialize our results when using particular versions of stochastic mirror descent to update both sets of variables. In this case, we demonstrate that the per iteration cost to update the decision variables is almost independent of the dimension of the distribution parameters. When the ambiguity sets are x^2-divergence sets, we additionally show that the update cost of the distribution parameters is almost independent of the dimension of the distribution parameters.