USC Marshall School of Business

DSO 670: Current Topics in Operations Management
Fall 2020 - Data-Driven Optimization: Theory, Methods and Current Themes

Professor: Vishal Gupta
Meeting Time: Wednesday 9 am – 12pm
Zoom Meeting Link: Prepopulated on Blackboard site
Office: BRI 401 G (See Blackboard for Zoom Details)
Email: guptavis@usc.edu
Office Hours: *By appointment only*

Course Description

This is a Ph.D. level seminar covering modern topics in data-driven optimization and decision-making under uncertainty. The goal is to equip students to pursue their own research in these areas. Hence, we will cover both some fundamental tools ideas in a lecture format, but mostly focus on recent papers and themes in student-led seminar style presentation/discussion. Special emphasis will be placed on modeling, the ways in which optimization settings differ from traditional machine-learning/estimation settings, and possible future areas of research.

Learning Objectives

By the end of the course, student should be able to

- Prove fundamental results about sample average approximation, stochastic gradient descent and other classical approaches and algorithms in data-driven optimization.
- Articulate the placement of research in context of the larger literature on data-driven optimization.
- Critically assess the contribution, strengths and weaknesses of recent research papers in data-driven optimization.
- Formulate their own novel research questions in this domain

Required Materials

There is no required textbook for the course. Lecture notes and recordings will be distributed through Blackboard. Students will be required to read a number of current journal papers and conference proceedings, all available freely via the USC library and posted to Blackboard. I may also post "OPTIONAL" reading material for Blackboard for interested students. After all, your PhD is the time to learn as much as you possibly can.

Prerequisites and/or Recommended Preparation:

There are no formal prerequisites for the course but students should be familiar with basics of mathematical optimization (convexity, duality, optimization algorithms) at the level of ISE 630/630 or equivalent Moreover, students should be very comfortable with basic probability theory (conditional expectation, Chernoff bounds, concentration) Measure theoretic probability is **not** required. Any students concerned about their background ability should reach out to the instructor to discuss their particular situation.

Course Notes:

All meetings for this course will be online over ZOOM. Please log-into class using your USC Zoom account. Note, if you're a Marshall student or have a personal zoom account, you may have multiple log-ins. If you use an account other than your USC zoom account, you will be placed in a "Waiting Room" upon entry. To avoid this, please use your USC zoom account. I will only allow participants in from the waiting room BEFORE class starts, and only participants that are registered or have made special arrangements with me. So if you're late, you won't be able to join. It's better in general to just use your USC account.

All materials for the course will be posted to the course's blackboard site. Please check there regularly. Students interested in auditing the course should reach out to the instructor to discuss and be added to blackboard.

The precise structure of the course will depend in part on how many students ultimately enroll. This will likely be finalized after the first session at which point an updated syllabus will be posted to Blackboard.

Roughly speaking, in the first few weeks of the course, I will be lecturing on some of the big themes in data-driven optimization, introducing fundamental tools used to analyze these methods and proving some of the classical results.

The majority of the course, however, will involve student presentations of papers. (See below for expectations on presentations.) Students will have the opportunity to sign up for papers after our first session.
My goal is that students should present AT MOST 1-2 papers throughout the semester. Clearly, meeting this goal depends on enrollment (hence some of the uncertainty in the syllabus).

All students (not just the presenter) are expected to read and think about papers before the seminar. The presenter will then moderate a discussion of the paper, its strengths, weaknesses and areas for improvement.

I will distribute the precise list of papers in our first session. These papers will roughly span the following topics and often contain two (competing) viewpoints on the same topic. This list is NOT yet final:

- Contextual Stochastic Optimization a.k.a. "Task-based" aka "End-to-End" learning
  Some possible papers:
  - "Smart-Predict-then-Optimize" by Elmachtoub and Grigas
  - "From Predictive to Prescriptive Analytics" by Bertsimas and Kallus
  - "Bootstrap Robust Prescriptive Analytics" by Bertsimas and Van Parys
  - "Melding the Data-Decisions Pipeline: Decision-Focused Learning for Combinatorial Optimization" by Wilder, Dilkina, and Tambe
  - "The Power and Limits of Predictive Approaches to Observational-Data-Driven Optimization" by Bertsimas and Kallus

- Data-Driven Distributionally Robust Optimization
  Some possible papers:
  - "Robust SAA" by Bertsimas, Gupta and Kallus
  - "Data-driven Distributionally Robust Optimization Using the Wasserstein Metric: Performance Guarantees and Tractable Reformulations" by Esfahani and Kuhn
  - "Robust empirical optimization is almost the same as mean–variance optimization" by Gotoh, Kim and Lim

- The Interplay between Optimization and Machine Learning
  Some possible papers:
  - "Testing the Validity of a Demand Model: An Operations Perspective," Omar Besbes, Robert Phillips, Assaf Zeevi
  - "Best Subset Selection via a Modern Optimization Lens" by Bertsimas, King and Mazumder
  - "Extended Comparisons of Best Subset Selection, Forward Stepwise Selection, and the Lasso" by Hastie, Tibshirani and Tibshirani

- “Optimal Classification Trees” by Bertsimas and Dunn
- “Learning Optimal Classification Trees: Strong Max-Flow Formulations” by Aghaei, Gomez and Vayanos
- “Optimization of Tree Ensembles” by Misic
- “Mining Optimal Policies” by Bravo and Yaposhnik

Depending on student interest, I am also happy to incorporate papers on learning in dynamic environments, Online Convex optimization, Causal Inference, and multi-stage optimization. Please email me if there’s a specific topic or paper you’d like to see.

Grading Policies:

There will be several graded deliverables for the class:

- Pre-Class Paper Check-Ins: Before every class in which there is a paper presentation, students will be required to answer a short set of questions on blackboard about the paper. These questions are meant to assess basic understanding of the paper to encourage students to come to class prepared. (In other words, the pre-class questions are EASY if you read the paper.). *Graded as Pass/No-Pass/Not-Complete*

- Post-Class Presentation Feedback: After each session in which a student colleague presents, students will be asked to complete constructive feedback on the presentation to help their colleague improve. Feedback is anonymous and no longer than a paragraph or two. *Graded as Pass/No-Pass/Not-Complete*

- Paper Presentation: Students will sign up for 1 (at most 2) papers to present throughout the semester. Building strong presentation skills is KEY to academic and professional success. To help students build their presentation skills, each presenter is REQUIRED to complete a multi-step process:

  - Before their presentation, the presenting student should do a ``practice-run” of their talk with at least 2 students (at least one of whom should be from our class.) After the practice-run, presenters should take time to debrief with the 2 audience members. What were the best parts of the presentation? What needs work? Are there specific slides or sections that are unclear? What additions should be made, or topics deleted?
    The practice-run and debrief should be completed on Zoom and recorded. You will need to submit the link to the zoom recording on BB as your deliverable. *Graded as Pass/No-Pass/Not-Complete*

  - After you’ve incorporated the feedback from the practice-run, you should schedule a time to meet with me (Vishal) to go through the slides of your talk and sort out any other presentation questions. Keep in mind my schedule is tight. Please plan ahead. *Graded as Pass/No-Pass/Not-Complete*

  - Present your chosen paper in class! While I am generally in favor of whiteboard talks, given the ZOOM nature of the class and the prevalence of slide talks in academia, I will ask that you use PowerPoint or equivalent in your talk. As part of the talk, please be prepared to moderate some discussion on the paper: What does everyone see as its strengths? What is truly novel relative to the literature? What questions remain?

    I will grade your and give written feedback after compiling the other student’s feedback.

- Coding Project: Prototyping and numerically testing algorithms is crucial to research in data-driven optimization. Too many researchers approach these issues from a strictly theoretical perspective and wind up designing algorithms that really don't work in practice.

  As your final project in the course, you can any of the papers we discuss in the course (including one you presented) and code up the method and a series of experiments of that method to better understand its performance. I will provide further details of this requirement in the first week of the course. Some notes:
  - Code can be written in any language (Julia, Python, C++, Matlab, R, etc.) but should be documented and released on GitHub. *Graded as Pass/No-Pass/Not-Complete*
  - The code should be code-reviewed before final submission by at least one other member of the class. *Graded as Pass/No-Pass/Not-Complete*
  - You should create a 5-page report on the numerical experiments you ran with your code explaining the new insights generated. It is my genuine hope that this exploration might encourage you to pursue novel research based on this project.

    I will grade this portion of the assignment and distribute a rubric later in the course.

In addition, participation in discussion and "Zoom-Ettiquette" will contribute to your final grade.

Policy on Group Work

Group discussion is STRONGLY encouraged throughout this class with other students in the class. Throughout your PhD, your peers will always be your best resource. Use them. You may collaborate with other students on ANY of the above deliverables.

However, you MUST always write up your own assignments individually and separately. (Thus, you can talk about a paper together, or even get a peer to read through your report and give you feedback, but you must incorporate that feedback on your own.) Please also list the names of students you collaborated with on the deliverable under your name, with a brief description of their contribution (if you deem it necessary).

For example, on my coding project, I might write:

> Final Project: The Robust-SAA Algorithm under Model Misspecification
> By: Vishal Gupta
>
> Collaborated with: John Snow (helpful discussions throughout project), Sansa Stark (editing final document), Tyrion Lannister (setting up GitHub account)

Grading Breakdown:

Assignment	% of Total Grade
Participation/Discussion	10%
Pre-Class Paper Write-ups	10%
Post-Class Presentation Feedback	10%
Paper Presentation	35%
Coding Project	35%
Total	100%

Synchronous session recording notice

All sessions of the course will be recorded and provided to all students enrolled in the course (and officially auditing non-USC students) via BB. Consequently, it is also important that students respect USC's policy and do NOT share any of the course content outside the course. This includes recordings, lecture notes, or other materials. For clarity, from SCampus:

SCampus Section 11.12(B)

Distribution or use of notes or recordings based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study is a violation of the USC Student Conduct Code. This includes, but is not limited to, providing materials for distribution by services publishing class notes. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relationship to the class, whether obtained in class, via email, on the Internet or via any other media. (SeeSection C.1 Class Notes Policy).

Zoom Etiquette Expectations

Online-learning brings some additional challenges outside the usual environment. I expect students to bring the same curiosity, engagement, and professionalism that they would normally in an in-person class. To that end, students are required

- To ensure they are in a quiet, private place to attend class. They should be able, if called upon, to unmute and participate in discussion.
- To ensure they have adequate technology and internet access to attend class without disturbance. If you have particular technology concerns, please reach out to me ASAP to discuss.
- To be familiar with Zoom: How to share screen, present, mute/unmute audio, raise your hand, ask for a coffee-break, etc.
- To keep their cameras on during class-time and mute their microphones when not speaking
- To participate in breakout rooms and discussion.

In summary, you know how to be a good student and a good citizen. Just do what you know.

ADDITIONAL INFORMATION

USC Statement on Academic Conduct and Support Systems

Explanation - This section, or an enhanced version, is required by the University. You are free to enhance the content as you deem necessary within the structure of the following.

Academic Conduct:

Students are expected to make themselves aware of and abide by the University community's standards of behavior as articulated in the Student Conduct Code. Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Part B, Section 11, "Behavior Violating University Standards" https://policy.usc.edu/scampus-part-b/. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, http://policy.usc.edu/scientific-misconduct.

Support Systems:

It is important to recognize that distance learning is hard. The online platform is difficult, but being isolated, especially in a process as challenging as Ph.D., is also hard. Please look out for one another. If you are feeling overwhelmed, reach out. You may always reach out to me or to your classmates. In other circumstances, you might feel more comfortable reaching out to one of the resources below.

Counseling and Mental Health - (213) 740-9355– 24/7 on call
https://studenthealth.usc.edu/counseling/
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call
suicidepreventionlifeline.org
Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL), press "0" after hours – 24/7 on call
https://studenthealth.usc.edu/sexual-assault/
Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED)- (213) 740-5086 | Title IX – (213) 821-8298
equity.usc.edu, titleix.usc.edu
Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298
https://usc-advocate.symplicity.com/care_report/
Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity |Title IX for appropriate investigation, supportive measures, and response.

The Office of Disability Services and Programs - (213) 740-0776
dsp.usc.edu
Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC is committed to making reasonable accommodations to assist individuals with disabilities in reaching their academic potential. If you have a disability which may impact your performance, attendance, or grades in this course and require accommodations, you must first register with the Office of Disability Services and Programs (www.usc.edu/disability). DSP provides certification for students with disabilities and helps arrange the relevant accommodations. Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to your TA) as early in the semester as possible. DSP is located in GFS (Grace Ford Salvatori Hall) 120 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776. Email: ability@usc.edu.

USC Campus Support and Intervention - (213) 821-4710
https://uscsa.usc.edu/
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC - (213) 740-2101
diversity.usc.edu
Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call
dps.usc.edu, emergency.usc.edu
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call
dps.usc.edu Non-emergency assistance or information.

COURSE CALENDAR

Week	Session Description	Presenter
19-Aug	Introductions and the Data-Driven Optimization Landscape	Vishal
26-Aug	Performance Guarantees for SAA in the Large-Sample Regime	Vishal
2-Sep	More Applications of ULLN and Introduction to Stability	Vishal
9-Sep	Analytics for an Online Retailer: Demand Forecasting and Price Optimization	Student
	A Statistical Learning Approach to Personalization in Revenue Management	Student
16-Sep	The Big Data Newsvendor: Practical Insights from Machine Learning	Student
	From Predictive to Prescriptive Analytics	Student
23-Sep	Optimal Prescriptive Trees	Student
	Recursive Partitioning for Personalization using Observational Data	Student
30-Sep	Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems	Student
	Robust Solutions of Optimization Problems Affected by Uncertain Probabilities	Student
7-Oct	Likelihood robust optimization for data-driven problems	Student
	Smart Predict Then Optimize	Student
14-Oct	Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations	Student
	Distributionally Robust Logistic Regression	Student
21-Oct	From Data to Decisions: Distributionally Robust Optimization is Optimal	Student
	Robust empirical optimization is almost the same as mean–variance optimization	Student
28-Oct	Mining Optimal Policies: A Pattern Recognition Approach to Model Analysis	Student
	Maximizing Intervention Effectiveness	Student
4-Nov	Interpreting Predictive Models for Human-in-the-Loop Analytics	Student
	The Power and Limits of Predictive Approaches to Observational-Data-Driven Optimization	Student

11-Nov	Testing the Validity of a Demand Model: An Operations Perspective	Student
	Small-Data, Large-Scale Linear Optimization with Uncertain Objectives	Student
20-Nov	**Final Coding Project Due**	