Distributionally Robust Optimization

Erick Delage (HEC Montréal),
Daniel Kuhn (Ecole Polytechnique Federale de Laussane),
Karthik Natarajan (Singapore University of Technology and Design),
Wolfram Wiesemann (Imperial College London)

March 4th to 9th, 2018

1 Overview of the Field

A wide variety of real-life decision-making problems, ranging from managerial planning and governmental policy making to finance, engineering, energy and healthcare, can naturally be formulated as *large-scale mathematical optimization problems* which seek to determine values for a set of decision variables that optimize a specified objective function, subject to a number of feasibility constraints. These optimization problems are often affected by substantial *uncertainty* because their parameters are subject to measurement errors or are not yet observable at the planning stage.

Optimization problems under uncertainty have a long history, and they are traditionally solved by methods of stochastic and dynamic programming, both of which have been popularized in the 1950s. Nowadays, the applicability of these methods is challenged for several reasons:

1. **Ambiguity.** In stochastic and dynamic programming, uncertainty is traditionally modeled via probability distributions. However, in many practical decision situations the raw data can be explained by several strikingly different distributions. Naive reliance on a single probabilistic model (e.g., the log-normal distribution underlying the Black Scholes equation of option pricing or the use of the Gaussian copula in the pricing of collateralized debt obligations [24]) can have catastrophic consequences, as has been demonstrated during the recent financial crisis.

2. **"Big Data."** In today's increasingly interconnected world, traditional localized decision problems must be integrated in order to correctly account for all possible synergies and systemic risks. Moreover, through the ongoing proliferation of digital information sources, increasing amounts of decision-relevant data become available (see [15] for a report by McKinsey Global Institute). As a result, modern decision problems have substantially increased in size and often grown beyond the grasp of traditional dynamic and stochastic programming methods.

3. **The Optimizer's Curse.** The solutions of stochastic programming problems parameterized by statistical data tend to display an optimistic bias even if the underlying parameter estimates are unbiased. This phenomenon is referred to as the "optimizer's curse" and can lead to great post-decision disappointment in out-of-sample tests (c.f. [26]).

4. **The Curse of Dimensionality.** In dynamic optimization all future decisions are modeled as contingency plans, that is, as functions mapping observations to actions. In order to solve the emerging functional optimization problems numerically, dynamic and stochastic programming discretize the underlying state space or the probability distribution of the uncertain parameters. In either case, the computation time grows exponentially with the problem size (see [1]), which has been a major impediment to the practical use of classical dynamic and stochastic programming methods.

The new field of *distributionally robust optimization (DRO)* (as popularized with [8]) aims to remedy both the conceptual and the computational shortcomings of classical stochastic and dynamic programming. The central idea is to represent uncertainty through an *ambiguity set*, that is, a family of (possibly infinitely many) probability distributions consistent with the available raw data or prior structural information, and to model the decision-making process as a game against "nature." In this game, the modeler first selects a decision with the goal to maximize expected reward, minimize risk or maximize the probability of constraint satisfaction etc., in response to which "nature" selects a distribution from within the ambiguity set with the goal to inflict maximum harm to the modeler. This setup prompts the modeler to select worst-case optimal decisions that offer performance guarantees valid for all distributions in the ambiguity set.

DRO has several striking benefits. It enables modelers to incorporate information about estimation errors into optimization problems. Therefore, it results in a more realistic account of uncertainty and mitigates the optimizer's curse characteristic for classical stochastic programming. Moreover, surprisingly, DRO problems can often be solved exactly and in polynomial time – in marked contrast to the intractable approximate models obtained via discretizations of stochastic problems tailored to a single nominal distribution. Thus, DRO models have the potential to scale to industrially relevant problem sizes and is already being employed in a number of fields of practice including vehicle routing, fleet management, portfolio selection, revenue management, scheduling, environmental policies, smart grid management, etc.

2 Recent Developments and Open Problems

A distributionally robust optimization model typically takes the form of :

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sup_{F\in\mathcal{D}} \mathbb{E}_F[h(x,\xi)]\,, \tag{1}$$

where $x \in \mathbb{R}^n$ is a vector of decision variables, $\mathcal{X} \subset \mathbb{R}^n$ is the set of implementable decisions, $\xi \in \mathbb{R}^m$ is a random vector for which the distribution F is only known to lie inside an ambiguity set $\mathcal{D}$. Finally, one typically refers to $h(x, \xi)$ as a cost function which depends on both the decisions that are implemented and the realized values for ξ.

One can classify some of the main open questions with the use of the DRO framework in the following categories.

Choice of performance measure in DRO: While the DRO paradigm has to this date been heavily used in its classical form (1), a number of alternative formulations, involving other risk measures than expectation [23], and competing paradigms have recently surfaced (e.g. [5]). This brings the question of how to select the right formulation in any specific context. Perhaps, the pessimistic view is not always the one that should be adopted. Interestingly, Wolfram Wiesemann presented some work establishing an axiomatic motivation for DRO. We also refer to talks summarized in sections 3.1 and 3.2 which shed more light on some of these issues.

Choice of ambiguity sets in DRO: While the community has been very successful, both from a numerical and statistical point of view, employing ambiguity sets that are based on perfect or imperfect moment information (e.g. in [8, 18]), or the notion of "distance" from a reference distribution [2, 27, 10, 4, 17], there are still a number of questions that should be addressed by our community. In particular, one may ask:
- What are the approaches that make the "best" use of data in the design of ambiguity sets?
- What type of algorithms might efficiently handle ambiguity sets that describe more structured random variables such as discrete random variables, dependent variables based on copula information, or even a structure as simple as independence?

Both of these questions were the subject of a number of talks during the workshop. In particular, the first question was the main focus on our second day for which highlights are presented in Section 3.3. One might also be interested by talks from the following presenters regarding the second question: Melvyn Sim, Krzysztof Postek, and Ruiwei Jiang.

Efficiency of solution methods : While, in many cases, it is possible to reformulate DRO problems as tractable mathematical programs (see [11, 19, 30]), there are still a number of DRO models for which high precision solutions are numerically and analytically unattainable. This gives rise to the following questions:
- Can we extend the range of tractable instances?
- Can we design more efficient exact methods or better bounds for approximation algorithms?

In this regard, it is worth examining some of the recent successes achieved in robust optimization (see for instance the talks summarized in Section 3.4). Moreover, a number of the talks summarized in Section 3.5 addressed the second questions.

Distributionally robust optimization in multi-stage settings: While the DRO approach has been very useful in the handling of static optimization problems, its use in multi-stage/dynamic problems is more limited (see e.g. [29, 28, 20] for its use within Markov decision processes). In our opinion, this is due to the fact that the community still has not found satisfying answers to the following questions:
- How can we design exact/approximate algorithms with solution time that scale reasonably with the number of periods?
- Following the work in [22, 13, 3], how can algorithms more efficiently handle discrete delayed decision?
- How can ambiguity set account for conditional independence?
- How severe are the time consistency issues that arise as reported in [25]?

We refer the reader to the presentations of Section 3.6 which shed some light on the first issue.

Applications: We close this section with a brief overview of the areas where distributionally robust optimization as been used: healthcare [16], energy (see [14], portfolio selection [6, 31, 32], statistics [9], vehicule routing [12], etc. A number of new applications were discussed in this workshop and are summarized in Section 3.7.

3 Presentation Highlights

3.1 Emergent Modeling Paradigms

The first session of the workshop had two talks on the general concepts and modeling issues in distributionally robust optimization (DRO) and was conducted by Yinyu Ye (Stanford University). Professor Ye discussed some of the popular modeling approaches for ambiguity sets in distributionally robust optimization using moment, likelihood and Wasserstein ambiguity sets and some of the corresponding tractability results. A second idea that he discussed was that of addressing ambiguity in high dimensions where only the marginal distributions are known but the joint is unknown. Using an approximation to the problem with independent distributions, he showed that for submodular functions, this approach has a good approximation guarantee which is however not valid for supermodular functions. While the complexity of solving the independence model was not discussed in detail in the talk, an interesting question it naturally raises is if the distributionally robust formulation can be directly solved to optimality. He also discussed generalization to online learning using concepts from DRO.

Melvyn Sim (National University of Singapore Business School) proposed new ambiguity sets in his work by building on a scenario wise representation where in each scenario, the uncertain random variables lies in an ambiguity set defined by moments information. Such an ambiguity set can be defined using clustering methods. The model has nice tractable convex optimization representation using linear programming (LP) and second-order cone programming (SOCP) methods and solving the DRO problem can be done in multi-stage problems using decision rules. He also discussed using infinitely constrained ambiguity sets to characterize information such as independence and algorithms to solve it.

Wolfram Wiesemann (Imperial College London) and Professor Erick Delage (HEC Montréal) gave a sequence of two talks on randomized decision making and the value it provides in distributionally robust optimization. Professor Wiesemann proposed a mathematical framework to define ambiguity averse risk measures and developed a representation theorem for it. He discussed conditions under which randomized decisions might be optimal in such a setting and used a simple facility location example to illustrate this. Professor Delage built on this to show the application of this method to distributionally robust facility location problems and developed tractable bounds for this problem while proposing a column generation method to solve this problem. He also proposed that randomized decision-making might be a useful paradigm particularly if the benefits over deterministic decision-making is significant. The idea of randomization in other decision-making problems such as consumer choice is currently under investigation by economists and this stream of work has interesting connections with it, but in an ambiguity averse setting.

Patrick Jaillet (Massachusetts Institute of Technology) discussed an alternative decision-making criterion under uncertainty where instead of maximizing expected utility, the objective is to have a satisfactory solution. He discussed how the classical chance constraint formulation can be generalized by the satisficing model by focusing on the maximum size of the ambiguity set that leads to satisfactory solution. He used a facility location problem to illustrate the model and this leads to interesting questions on how these techniques can be adapted to other decision-making problems.

Marco Campi (University of Brescia) and Simone Garatti (Politecnico di Milano) presented two talks on scenario optimization which appears to be a powerful mathematical tool for controlling the risk that a solution be infeasible. In particular, they demonstrated that in a decision making problem where the constraint set is indexed by scenarios drawn from a distribution, the infeasibility risk is directly related to what they call "complexity", i.e. the minimum number of scenarios needed to identify the optimal solution. While the distribution of scenarios is typically difficult to identify precisely, "complexity" can numerically be evaluated once the optimal solution to the scenario problem is obtained. It is also possible to explore the trade-off between infeasibility risk and expected performance by employing their theory on a version of the decision problem where violations of the constraints are penalized more or less heavily through a tunable penalty parameter. The use of regularization as a way of controlling risk is also discussed.

3.2 The Interplay between DRO and Classical Risk Measures

Jonathan Y. Li (University of Ottawa) reviewed some of the closed form bounds for worst-case conditional value at risk measure under first and second order moment information and showed how it can be generalized to law invariant risk spectral risk measures. His work discusses how for special cases of risk measures, multi-dimensional problems can be reduced to single-dimensional problems. The worst-case distribution in such examples is not two points unlike the classical worst-case conditional value at risk model. A natural question that this work raises is how can these methods be applied to general two stage distributionally robust optimization problems and what are the complexity implications in those cases?

Ruiwei Jiang (University of Michigan) in his talk discussed how DRO chance constraints and conditional value at risk constraints with first and second order moment information can be supplemented with additional structural information such as log-concavity and tail dominance. For the log-concave set of distributions, the results indicate that the DRO constraint with conditional value-at-risk (CVaR) can be formulated in a manner similar to what is currently known but with modified scaling factors for the second order conic terms. In the case of the value-at-risk (VaR) formulation, it is not tight, but it is possible to develop tractable approximations. He discussed applications of the results in optimal power flow problems and appointment scheduling problems.

3.3 The Use of DRO in Data-driven Problems

Dimitris Bertsimas (Massachusetts Institute of Technology) provided a survey of recent developments in the field of data-driven optimization. He showed how one can transform predictive machine learning algorithms to prescriptive ones that are useful for two-stage and multi-stage decision-making problems and characterized the rates of convergence of the resulting algorithms. He also described an approach to make prescriptive approaches immune to overfitting phemonema and to improve the performance of parametric methods via kernel techniques.

Anton Kleywegt (Georgia Institute of Technology) showed how distributionally robust optimization can help to regularize algorithms of statistical learning. This is achieved by minimizing the worst-case expected prediction error across all distributions in a ball, sized with respect to the Wasserstein distance, is centered at the empirical distribution. He identified a broad class of loss functions for which the proposed approach leads to a gradient-norm penalty and discussed important applications in deep learning and discrete choice models.

Andrew Lim (National University of Singapore) proposed a theory for calibrating the ambiguity parameter that typically determines the size of the uncertainty set in robust optimization. He showed that the first-order benefit of injecting robustness into a nominal optimization model is a significant reduction in the variance of the out-of-sample reward, while the corresponding impact on the mean reward is almost an order of magnitude smaller. This observation motivated Andrew to introduce a robust mean-variance frontier, which can be used to tune the ambiguity parameter. He also showed that this frontier can conveniently be approximated using resampling methods like the bootstrap. He then provided evidence that solutions of robust optimization problems whose ambiguity parameters are calibrated to ensure a certain coverage probability may be too conservative out of sample, while tuning the ambiguity parameter in view of the out-of-sample expected reward with no regard for the variance may lack robustness.

Vishal Gupta (University of Southern California) pointed out that practical optimization problems often depend on a huge number of uncertain parameters, for each of which there are only very few historical observations leading to imprecise estimates. He then argued that this large-scale, small-data regime is distinct from the large-sample regime usually studied in statistics. Given a fixed class of candidate policies, he then identified a policy that performs best in this class asymptotically as the number of uncertain parameters tends to infinity (while the number of samples per parameter remains small). He further showed that the loss of optimality of the proposed method relative to the best-in-class policy decays exponentially fast in the number of uncertain parameters for two important policy classes inspired by the empirical Bayes and the regularization literature, respectively.

Karthyek Murthy (Singapore University of Technology and Design) investigated distributionally robust optimization problems where the ambiguity sets are defined via optimal transport distances such as the popular Wasserstein distance. He demonstrated that several widely used machine learning algorithms that employ regularization can be recovered as particular examples of this distributionally robust approach. He also developed a method to calibrate the radius of a Wasserstein ambiguity set by leveraging ideas from empirical likelihood theory in statistics that obviates the need to use brute-force cross-validation techniques. Moreover, he proposed a fast stochastic gradient descent algorithm for solving the resulting optimization problems.

Bart Van Parys (Massachusetts Institute of Technology) studied data-driven optimization problems, where the uncertain parameters that impact the problem's objective function depend on observable contextual information captured by a potentially large number of covariates. He emphasized that a naïve use of the training data (independent samples of the uncertain problem parameters and the covariates) may lead to overfitting. To combat the overfitting, he proposed to leverage ideas from distributionally robust optimization, the statistical bootstrap and Nadaraya-Watson or nearest neighbor density estimation. He showed that the proposed approach leads to tractable convex optimization problems that offers rigorous out-of-sample guarantees.

Henry Lam (Columbia University) discussed new methods to calibrate the uncertainty sets in robust optimization and the Monte Carlo sample sizes in constraint sampling or scenario generation. He proposed strategies to select good parameter values based on data splitting and the validation of their performances in terms of feasibility and optimality. He then showcased the effectiveness of these strategies in relation to the dimension of the underlying optimization model.

Nathan Kallus (Cornell University) proposed a new approach to policy learning from observational data (for example, the transformation of electronic health records to personalized treatment regimes). The task is difficult because only outcomes of the interventions performed are observable, and the distribution of units exposed to one intervention or another differs systematically. Nathan described a new distributionally robust approach to policy learning in the context of personalized medicine that offers strong finite-sample guarantees and preserves the asymptotic optimality and convergence rates of naïve traditional plug-in approaches.

Daniel Kuhn (École Polytechnique Fédérale de Lausanne) outlined an abstract perspective on data-driven stochastic programming whereby one should find a procedure that maps time series data to a near-optimal decision (a prescriptor) and to a prediction of this decision's expected cost under the unknown data-generating distribution (a predictor). He proposed a meta-optimization problem, that is, an optimization problem over optimization problems, that identifies the least conservative predictors and prescriptors subject to constraints on their out-of-sample disappointment. He then showed that the best predictor-prescriptor pair is obtained by solving a distributionally robust optimization problem.

3.4 Computational Lessons Learned from Robust Optimization

Dick den Hertog (Tilburg University) proposed to tackle a robust optimization problem where the cost function $h(x, \xi)$ is convex with respect to both x and ξ. These are generically difficult problems to handle within a classical robust optimization framework, yet have recently been addressed successfully for special instances of DRO models [30, 21]. The proposed approach applies when the uncertainty set is polyhedral and relies on a series of reformulation that converts the problem into a two-stage robust linear program. The resulting problems can subsequently be approximated via classical approximation schemes for these types of robust optimization models. A set of numerical examples provide evidence that this approach often leads to near-optimal solutions.

Grani Hanasusanto (University of Texas at Austin) studied robust quadratically constrained quadratic programs where the uncertain problem parameters contain both continuous and integer components. Grani showed that these problems can be reformulated as copositive programs of polynomial size, which he subsequently approximated via semidefinite programs. Grani demonstrated the superiority of these approximations over the state-of-the-art solution schemes on several problem classes.

Anthony Man-Cho So (Chinese University of Hong Kong) considered a class of robust quadratic optimization problems that arise in various applications in signal processing and wireless communications (such as robust beamforming with cognitive radio constraints). Due to the NP-hardness of this problem class, Anthony developed an approximate solution scheme that uses the so-called epsilon-net technique from functional analysis to offer rigorous approximation guarantees.

3.5 The Use of DRO with Moment Information

A classical form for the DRO problem employs an ambiguity set defined by imposing constraints on moments of the distribution. Typically, these include bounding the mean vector of ξ and its second-order moment matrix $\mathbb{E}[(\xi - \mu)(\xi - \mu)^T)]$. In her talk, Siqian Shen revisited the well-known ambiguity set presented in [8] for the case of a distributionally robust chance constrained (DRCC) optimization problem. She demonstrated that it is possible to reformulate the chance constraint using second-order cone (SOC) constraints instead of linear matrix inequalities (LMI) as was originally proposed. This is interesting since 1) from a numerical perspective, SOC constraints are easier to optimize than LMIs; 2) it somehow ties in together a number of reformulations that are known for these type of constraints and which uses the canonical representation first introduced in [7]. Siqian also presented a branch-and-cut method to improve the solution time in a bin packing problem with these types of constraints. The numerical results were quite promising.

Abdel Lisser (Laboratoire de Recherche en Informatique - University Paris Sud) further presented how the DRCC could be reformulated for a geometric program with uncertainty about the coefficients that multiply each monomial. He also obtained useful reformulation for the cases where the ambiguity set is based on KL-divergence with respect to some reference empirical distribution. While these reformulations were not SOC programs in this case they still could be solved using algorithms that are available for convex optimization problems. Abdel's numerical results involved a shape optimization problem where it appears that there is still important challenges to overcome before large scale problems can be addressed.

Krzysztof Postek (Erasmus University Rotterdam) addresses in his talk the common criticism made against DRO models that they often consider worst-case distribution that are unnatural for the problem at hand. Indeed, it is well known that most moment-based problems have a worst-case distribution that is supported on a handful of scenarios. The proposed remedy consist in using a polynomial function as the density function. In this way, it is possible to control the smoothness of the worst-case density function and prevent it from peaking so significantly at any point of the support set. While a number of numerical difficulties seem to arise when implementing this idea, Krzysztof appeared very resourceful in addressing each of them and will certainly make interesting progress in the months to come.

The talk of Jianqiang Cheng (University of Arizona) proposed methods to improve the computational efficiency in moment-based DRO problems by exploiting principal component analysis (PCA) to reduce the dimensionality of the uncertain vector ξ. This is especially promising in problems that involve covariance information as these problems can take the form of semi-definite programming model where the dimension of ξ has an important effect on solution time. Theoretically, he provides a bound on the size of the approximation error that is introduce through dimensionality reduction which depends on the size of the eigenvalues that are dropped during PCA. The technique is applied on a production-transportation problem where the objective function is a distributionally robust conditional value-at-risk (CVaR) measure. Numerical experiments illustrate the trade-off between the size of the approximation gap and computation time for solving this problem.

Huifu Xu (University of Southampton) presented valuable results concerning the stability of moment-based DRO models. In particular, he considered the case where the moment information is estimated from data and whether the solutions of DRO models constructed based on these estimate converge to a limit solution as more data is used to make the estimation. He also paid special attention to cases where the distribution ambiguity set is designed as decision dependent. Finally, he concluded is presentation with a discussion on the implication of his stability results for a distributionally robust chance constrained problem.

Guzin Bayraksan (Ohio State University) introduced the notion of effective scenarios and ineffective scenarios in the context of a scenario based distributionally robust optimization problem, which is a special case of moment based DRO where the distribution's support is assumed discrete. In particular, she noticed that in some DRO formulations, the optimal solution is not necessarily sensitive to all the scenario with positive probability in a worst-case distribution. For this reason, she considers a scenario to be effective if its removal from the support causes the optimal value of the DRO problem to change, otherwise the scenario is considered ineffective. In some way, these concepts are related to the concept of "complexity" presented by M. Campi and S. Garatti yet the intent here is different. Indeed, Guzin's interest is regarding post-optimization analysis where one can question whether effective scenarios are realistic of perhaps outliers that can be removed from the problem. She also indicates how this information might be helpful from a computational point of view by motivating scenario reduction scheme or provide guidance for improving the effectiveness of decomposition schemes. The applications that illustrates her findings involves a water allocation problem for the Colorado River and where a total variation ambiguity set was used.

3.6 Handling Multi-stage problems with DRO

Sanjay Mehrotra (Northwestern University) proposed decomposition algorithms for two-stage distributionally robust optimization problems with continuous and binary decision variables in both the first and the second stage. The algorithms utilize distribution separation procedures and parametric cuts within a Benders algorithm. The presentation also studied conditions and families of ambiguity set for which the proposed algorithms converge in finite time.

Huan Xu (Georgia Institute of Technology) discussed robust Markov decision process models with parameter uncertainty. His talk discussed how it is possible to learn the uncertainty when a combination of robust and stochastic elements are present in a Markov decision process. He developed an algorithm that combines elements of pessimism and optimism such that it is robust to adversarial uncertainty and optimistic to stochastic uncertainty. This talk discusses new ideas of learning uncertainty sets that might be relevant in other online decision-making problems.

Angelos Georghiou (McGill University) proposed an adaptation of the well-known stochastic dual dynamic programming (SDDP) scheme to multi-stage robust optimization problems. The resulting robust robust dual dynamic programming (RDDP) scheme maintains both inner and outer approximations of the cost to-go for each time stage of the problem. The algorithm converges in finite time, and the presented numerical results show the promise of the proposed scheme.

Vineet Goyal (Columbia University) investigated the performance of affine policies in two-stage robust optimization problems. While it is well-known that their worst-case performance is poor, it has been observed that affine policies perform well in numerical experiments. The presentation has shown that affine policies are a good approximation for two-stage adjustable robust optimization problems with high probability on random instances where the constraint coefficients are generated i.i.d.

Dan Iancu (Stanford University) discussed necessary and sufficient conditions for affine policies to be optimal in multi-stage robust optimization problems. The treatment drew interesting connections with the theory of discrete convexity and global concave envelopes.

Georg Pflug (University of Vienna) in his talk discussed the idea of how distributional robustness might be modeled in multi-stage problems. This is a challenging problem and he proposed an extension of the Wasserstein metric ambiguity set that is relevant for the multi-stage problem using distances between conditional probability measures. The size of the ambiguity sets can be determined from statistical confidence regions. He also discussed applications to the multi-period portfolio optimization using the average value at risk measure under ambiguity.

3.7 Recent Applications of DRO

Tito Homem-de-Mello (Universidad Adolfo Ibáñez) used distributionally robust optimization to model a general class of newsvendor problems where the underlying demand distribution is unknown and the goal is to find an order quantity that minimizes the worst-case expected cost among all distributions within some distance from a given nominal distribution. Due to the specific structure, Tito was able to derive explicit formulas and properties of the optimal solution as a function of the level of robustness, determine the regions of demand that are critical to optimal cost and establish quantitative relationships between the distributionally robust model and the corresponding risk-neutral and classical robust optimization models.

Karthik Natarajan (Singapore University of Technology and Design) also revisited the distributionally robust newsvendor problem but this time exploring conditions for which the optimal order quantity that is returned by DRO can be motived using a risk neutral setting where the distribution is known. While an early result from H. Scarf already established that this is the case for mean-variance models, Karthik presented an analysis that extends, in an asymptotic sense, this equivalence to problems where variance information is replaced by any p-th order moment information. The interesting insight from this range of work is the fact that solutions from DRO models might for some family of problems be explained using a stochastic programming model that uses a heavy-tailed distribution which is not even a member of the original distribution set.

Chung Piaw Teo (National University of Singapore) proposed a new method to integrate probabilistic assessment of disruption risks into the Risk Exposure Index approach for supply chains under disruption. His method measures the resilience of a supply chain by analyzing the Worst-case CVaR (WCVaR) of total lost sales under disruptions. Chung Piaw showed that the optimal strategy in this model can be fully characterized by a conic program, which allows managers to focus their mitigation efforts on critical suppliers and/or installations that will have greater impact on the performance of the supply chain when disrupted.

Selin Damla Ahipasaoglu (Singapore University of Technology and Design) proposed novel classes of equilibria in traffic models with relaxed information assumptions. Selin studied conditions under which these equilibria exist and are unique, and she also provided convex programs for determining these equilibria and developed customized algorithms to calculate the equilibrium flows.

Phebe Vayanos (University of Southern California) studied systems that allocate different types of scarce resources to heterogeneous allocatees based on predetermined priority rules, such as the U.S. deceased-donor kidney allocation system or the public housing program. Phebe proposed to estimate the wait times in such systems via the solution of robust multiclass, multiserver queuing systems. Phebe showcased how the methodology can be applied to the U.S. deceased- donor kidney waitlist.

Jin Qi (Hong Kong University of Science and Technology) studied a finite horizon stochastic inventory routing problem where the supplier determines the replenishment quantities as well as the times and routes to all retailers. The probability distribution governing the uncertain demand of each retailer is assumed to be ambiguous, and the goal is to minimize the risk of the uncertain inventory levels falling outside a pre-specified range. Jin provided algorithms to solve the problem exactly, and she compared the performance of her solutions with several benchmarks to show their benefits.

John Gunnar Carlsson (University of Southern California) considered a distributionally robust version of the Euclidean travelling salesman problem and computed the worst-case spatial demand distribution from within a Wasserstein ball centered at an observed demand distribution. Numerical experiments on a districting problem in multi-vehicle routing confirmed that the proposed approach is useful for decision support tools that divide a territory into service districts for a fleet of vehicles when limited data is available.

4 Outcome of the Meeting

The meeting provided a great opportunity for high quality researchers to meet and discuss some of the most active research directions of this field. Our discussions were particularly enriched by the diversity and range of areas of expertise of the participants. These included experts in optimization, stochastic modeling, game theory, statistics and machine learning, and experts of applications such as financial engineering, vehicle routing, scheduling, health care, to name a few. In the opinion of the organizing committee, there is no doubt that a number of new collaborations have emerged from this event.

References

[1] R. Bellman. *Dynamic Programming*. Courier Dover Publications, 2003.

[2] A. Ben-Tal, D. den Hertog, A. De Waegenaere, B. Melenberg, and G. Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

[3] D. Bertsimas and A. Georghiou. Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Operations Research*, 63(3):610–627, 2015.

[4] D. Bertsimas, V. Gupta, and N. Kallus. Robust sample average approximation. *Mathematical Programming*, Jun 2017.

[5] D. B. Brown and M. Sim. Satisficing measures for analysis of risky positions. *Management Science*, 55(1):71–84, 2009.

[6] G. Calafiore. Ambiguous risk measures and optimal robust portfolios. *SIAM Journal on Optimization*, 18(3):853–877, 2007.

[7] G. Calafiore and L. El Ghaoui. On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications.*, 130(1):1–22, 2006.

[8] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):596–612, 2010.

[9] L. El-Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, October 1997.

[10] R. Gao and A. J. Kleywegt. Distributionally Robust Stochastic Optimization with Wasserstein Distance. *ArXiv e-prints*, April 2016.

[11] J. Goh and M. Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4):902–917, 2010.

[12] C. E. Gounaris, W. Wiesemann, and C. A. Floudas. The robust capacitated vehicle routing problem under demand uncertainty. *Operations Research*, 61(3):677–693, 2013.

[13] Grani A. Hanasusanto, Daniel Kuhn, and Wolfram Wiesemann. K-adaptability in two-stage robust binary programming. *Operations Research*, 63(4):877–891, 2015.

[14] R. Jiang, Y. Guan, and J.-P. Watson. Risk-averse stochastic unit commitment with incomplete information. *IIE Transactions*, 48(9):838–854, 2016.

[15] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Hung Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, May 2011.

[16] F. Meng, J. Qi, M. Zhang, J. Ang, S. Chu, and M. Sim. A robust optimization model for managing elective admission in a public hospital. *Operations Research*, 63(6):1452–1467, 2015.

[17] P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, Jul 2017.

[18] K. Natarajan, D. Pachamanova, and M. Sim. Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Science*, 54(3):573–585, 2008.

[19] K. Natarajan, M. Sim, and J. Uichanco. Tractable robust expected utiltiy and risk models for portfolio optimization. *Mathematical Finance*, 20(4):695–731, 2010.

[20] A. Nilim and L. El Ghaoui. Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, 53(5):780–798, 2005.

[21] K. Postek, A. Ben-Tal, D. den Hertog, and B. Melenberg. Robust optimization with ambiguous stochastic constraints under mean and dispersion information. *Operations Research*, 2018.

[22] K. Postek and D. den Hertog. Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing*, 28(3):553–574, 2016.

[23] K. S. Postek, D. den Hertog, and B. Melenberg. Computationally tractable counterparts of distributionally robust constraints on risk measures. *SIAM Review*, 58(4):603–650, 2016.

[24] F. Salmon. Recipe for disaster: The formula that killed wall street. WIRED Magazine, Magazine-17.03, February 2009.

[25] A. Shapiro and L. XIn. Time inconsistency of optimal policies of distributionally robust inventory models. Submitted draft, September 2017.

[26] J. E. Smith and R. L. Winkler. The optimizers curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006.

[27] Z. Wang, P. W. Glynn, and Y. Ye. Likelihood robust optimization for data-driven problems. *Computational Management Science*, 13(2):241–261, Apr 2016.

[28] W. Wiesemann, D. Kuhn, and B. Rustem. Robust Markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2013.

[29] W. Wiesemann, D. Kuhn, and B. Rustem. Robust Markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2014.

[30] W. Wiesemann, D. Kuhn, and M. Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014.

[31] S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57(5):1155–1168, 2009.

[32] S. Zymler, D. Kuhn, and B. Rustem. Worst-case value at risk of nonlinear portfolios. *Management Science*, 59(1):172–188, 2013.