STOR-i Conference: 9th – 10th January 2020

Titles and Abstracts

Day 1

Analytics for Operations, Scheduling and Pricing in Air Transportation
Alex Jacquillat, MIT
The transportation sector is facing profound challenges in its quest to accommodate growing demand with limited capacity. Analytics can play a catalytic role in enhancing mobility, reliability and profitability. This talk reports results from various analytics projects aimed to support operations, scheduling and pricing practices in air transportation. On the operations side, we present a new integer program for air traffic flow management that explicitly balances flight delays with passenger delays—showing that flight operations can be made more consistent with passengers' itineraries. On the scheduling side, we present a large-scale optimization approach to airport slot allocation—showing that we can derive optimal, or near-optimal solutions at some of the largest schedule-coordinated airports. We also describe a stochastic integer program that optimizes scheduling interventions in networks of airports, while capturing the interdependencies between flight schedules and air traffic flow management operations—showing that the proposed integrated approach outperforms sequential benchmarks that separate scheduling and operations. On the pricing side, we outline a field experiment conducted with a global airline to assess a new competitive pricing strategy—showing that carefully designed adjustments to baseline practices can result in significant increases in revenues and yield. We conclude with a discussion of future research opportunities.

MUMBO: Multi-task Max-value Bayesian Optimisation
Henry Moss, STOR-i PhD Student
We propose MUMBO, the first high-performing yet computationally efficient acquisition function for multi-task Bayesian optimisation. Here, the challenge is to perform efficient optimisation by evaluating low-cost functions somehow related to our true target function, a broad class of problems including the popular task of multi-fidelity optimisation. However, while information-theoretic acquisition functions are known to provide state-of-the-art Bayesian optimisation, existing implementations for multi-task scenarios have prohibitive computational requirements. Previous acquisition functions have therefore been suitable only for problems with both low-dimensional parameter spaces and function query costs sufficiently large to overshadow very significant optimisation overheads. Our approach is scalable, efficient and delivers robust performance across classic optimisation challenges and machine learning hyper-parameter tuning tasks.

Probabilistic Preference Learning via the Mallows rank model: some recent advances
Valeria Vitelli, University of Oslo
Ranking items is crucial for collecting information about preferences in many areas, from marketing to politics. The interest often lies both in producing estimates of the consensus ranking of the items, which is shared among users, and in learning individualized preferences of the users, useful for providing personalized recommendations. In the latter task, it is particularly relevant to have posterior distributions of individual rankings, since these can provide an evaluation of the uncertainty associated to the estimates, and thus they can avoid unnecessarily spamming the users. I will present a statistical model, the Mallows ranking model, which works well in these situations, and which is able of flexibly handling quite different kind of data. The Bayesian paradigm allows a fully probabilistic analysis, and it easily handles missing data and cluster estimation via augmentation procedures. Interestingly, this Bayesian framework has proved to be useful in a variety of situations, from genomic data integration to recommender systems, and I will briefly review some relevant case studies to show the method's potentialities. I will then conclude with some advances we have recently started to work on.

Keywords: Mallows model, Bayesian computing, recommender systems, genomics data integration.

Cuts, Flows and Eigenvalues
Miguel Anjos, University of Edinburgh
Cuts and flows are well-known classical concepts in optimization and operational research. In this talk we will consider two lesser known challenging instances of such problems, namely the max-k-cut problem and the optimal power flow problem, and how eigenvalues can play a role in tackling them.

Valid Inequalities for Mixed-Integer Programs with Fixed Charges on Sets of Variables
Georgia Souli, STOR-i PhD Student
Algorithms and software for integer programming and combinatorial optimisation have improved dramatically over the past couple of decades. A key source of this success is the use of strong valid linear inequalities, also known as cutting planes. One strand of the literature on cutting planes, with numerous applications in Statistics and Operational Research, is concerned with mixed-integer programs that involve fixed charges. A fixed charge is an additional cost that is incurred whenever a certain decision variable takes a positive value. The textbook approach to model such problems tends to perform badly, but several useful families of valid inequalities have been derived. In particular, the so-called “flow cover inequalities” have proven to be so effective that they are now generated by default in both commercial and open-source solvers. In this talk, we consider a more general situation, in which a fixed charge is incurred if and only if at least one of the decision variables in a set takes a positive value. We derive valid inequalities for these problems and show that they generalise and dominate a subclass of the flow cover inequalities for the classical fixed-charge problem. *This is joint work with Adam N. Letchford.

The Use of Shape Constraints for Modeling Time Series of Counts
Richard Davis, Columbia University
For many formulations of models for time series of counts, the specification of a family of probability mass functions relating the observation Yt at time t to a state variable Xt must be explicitly specified. Typical choices are the Poisson and negative binomial distributions. One of the principal goals of this research is to relax this parametric framework and assume that the requisite pmf is a one-parameter exponential family in which the reference distribution is unknown but log-concave. This class of distributions includes many of the commonly used pmfs. The serial dependence in the model is governed by specifying the evolution of the conditional mean process. The particular link function used in the exponential family model depends on the specification of the reference distribution. Using this semi-parametric model formulation, we are able to extend the class of observation-driven models studied in Davis and Liu (2016). In particular, we show there exists a stationary and ergodic solution to the state-space model. In this new semi-parametric framework, we compute and maximize the likelihood function over both the parameters associated with the mean function and the reference measure subject to a concavity constraint. On top of this we can “smooth” the pmf using the Skellam distribution in order to obtain an estimated distribution defined on all the non-negative integers. In general, the smooth version has better performance than existing methods. The estimator of the mean function and the conditional distribution are shown to be consistent and perform well compared to a full parametric model specification. Further limit theory in other situations will be described. The finite sample behavior of the estimators are studied via simulation and empirical examples are provided to illustrate the methodology. This is joint work with Jing Zhang and Thibault Vatter.

Challenges in Real Time Machine Learning
Tom Flowerdew, STOR-i Alumni
A common problem which Featurespace solves for banks, and other financial institutions is to score transactions for their risk of fraud. This problem has a number of challenges:

- Given that a transaction is held while the fraud risk is being calculated, the speed of response is extremely important, meaning that the techniques employed must be deployable in an extremely time-efficient manner;
- Some fraud attacks can be perpetrated in such a way that context from the previous few minutes or seconds is important in order to produce a high-quality score;
- Unlike other Machine Learning problems, the models' decision has an influence on the future data it receives. An example of this might be that a Machine Learning successfully blocks a previously-successful fraud attack vector; the obvious next step for the fraudster would be to change the exact way in which fraud is attempted, or find an entirely new strategy.

This talk will explore some of these challenges in more detail, then explore possible strategies employed during the design of Machine Learning models which allows these problems to be tackled successfully.

Making random things better: Optimisation of Stochastic Systems
Christine Currie, University of Southampton
This talk will take a rather rambling path through the optimisation of stochastic systems, considering diverse applications from the passenger transport industry and healthcare. The focus will be on methods that account for the stochasticity of both inputs and outputs, referring to optimisation-via-simulation methods and robust optimisation. The applications we consider are of complex systems where there are often (but not always) several objectives.

Day 2

Accounting for network effects in credit risk modelling
Veronica Vinciotti, Brunel University
Credit risk models that are traditionally used in industry do not account for possible dependencies in the firms' risk of default. Transaction networks are rarely made available and may not fully account for the firms' dependencies. In this talk, I will firstly explore how transaction data can be used for credit risk modelling and the advantages that this may bring in terms of predictive power. Secondly, I will present a correlated mixed probit model for capturing dependencies that are either not observed or not accounted for by the transaction network. Given the high-dimensionality of the data and the computational challenges of correlated mixed models, we develop an efficient expected-maximisation algorithm for penalised inference. I will show the performance of the methods on a large sample of accounts for small and medium-sized enterprises in the UK.

References:
Vinciotti et al (2019) The effect of interfirm financial transactions on the credit risk of small and medium-sized enterprises. JRSS-A, 182, 4, 1205-1226.
Tosetti and Vinciotti (2019) A computationally efficient correlated mixed probit model for credit risk inference. JRSS-C, 68, 4, 1183-1204.

The Optimization of Explaining Data Science

Dolores Romero Morales, Copenhagen Business School

Data Science models should strike a balance between accuracy and interpretability, providing explanations on the (un)supervised task to the user who needs to interact with the models. Although classification and regression trees have traditionally been considered to be leaders in interpretability, there is a growing literature on enhancing the interpretability of other Data Science methods. In this talk, we will navigate through several Mathematical-Optimization based models in Data Science that illustrate the important role of Mixed-Integer NonLinear Programming in achieving such a balance.

Multi-armed bandit problems with history dependent rewards.

Ciara Pike-Burke, STOR-i Alumni

The multi-armed bandit problem is a common sequential decision making framework where at each time step a player selects an action and receives some reward from selecting that action. The aim is to select actions to maximize the total reward. Commonly it is assumed that the (expected) reward of each action is constant and does not depend on the actions that the player has previously taken. However, in many practical settings this is not realistic. For example in web-advertising, the benefit from showing an advert is likely to depend on the number of times the user has seen it in the past, and in product recommendation the reward of suggesting an item will depend on the time since it was last suggested to the customer. In this talk we will consider several variants of the multi-armed bandit problem where the reward depends on the history of the players actions. For each problem, we will discuss whether learning is possible, and if so provide algorithms that perform well.

Model-based Clustering with Sparse Covariance Matrices

Brendan Murphy, University College Dublin

Finite Gaussian mixture models are widely used for model-based clustering of continuous data. Nevertheless, since the number of model parameters scales quadratically with the number of variables, these models can be easily over-parameterized. For this reason, parsimonious models have been developed via covariance matrix decompositions or assuming local independence. However, these remedies do not allow for direct estimation of sparse covariance matrices nor do they take into account that the structure of association among the variables can vary from one cluster to the other. To this end, we introduce mixtures of Gaussian covariance graph models for model-based clustering with sparse covariance matrices. A penalized likelihood approach is employed for estimation and a general penalty term on the graph configurations can be used to induce different levels of sparsity and incorporate prior knowledge. Model estimation is carried out using a structural-EM algorithm for parameters and graph structure estimation, where two alternative strategies based on a genetic algorithm and an efficient stepwise search are proposed for inference. With this approach, sparse component covariance matrices are directly obtained. The framework results in a parsimonious model-based clustering of the data via a flexible model for the within-group joint distribution of the variables. Extensive simulated data experiments and application to illustrative datasets show that the method attains good classification performance and model quality. This work was completed with Michael Fop and Luca Scrucca.