DISSERTATION PROPOSAL

Bo Yang

"Pathwise Optimization for Merchant Energy Production"

Monday, December 3, 2018
3:00 pm
Tepper Quad 4242

The merchant operations approach formulates the management of energy and commodity conversion assets as real option models in which the market prices of the transformation inputs and outputs evolve stochastically. This representation provides a convenient way to maximize the market value of these assets. We study merchant operations of energy production assets, such as power and natural-gas-processing plants, oil and bio refineries, and ethanol production facilities. The resulting model is a finite-horizon compound switching and timing option, in which the production asset can be periodically operated, suspended, mothballed, reactivated, or abandoned depending on market conditions. The Markov decision process (MDP) formulation of this model is intractable due to intertemporal linkages between decisions and the high dimensionality of the market state variables when using realistic price models. We consider approximate dynamic programming (ADP) approaches that rely on basis-function-based value function approximations to compute both a feasible policy and a lower bound on the optimal policy value of this MDP, as well as a dual (upper) bound on the latter quantity. In particular, we focus on pathwise optimization (PO), because, in theory, it generates the tightest dual bound for a given set of basis functions.

Applying least squares Monte Carlo (LSM), a state-of-the-art ADP approach, in conjunction with information relaxation and duality techniques to realistic merchant ethanol production instances is known to result in sizable (about 13%) average optimality gaps. The first chapter aims at reducing these gaps by extending the PO applicability from optimal stopping to merchant energy production. This extension rests on formulating a novel PO linear program, which is difficult to solve optimally because it is both ill conditioned and large scale. We develop an effective preconditioning approach based on principal component analysis (PCA). We mitigate the dimensionality concern by proposing a provably convergent block coordinate descent (BCD) technique. On the known set of merchant ethanol production instances PO leads to considerably smaller (roughly 7%) average optimality gaps compared to LSM but entails a noticeably larger average computational effort (eleven hours rather than one hour). The optimality gap reduction is almost entirely attributable to the stronger PO-based dual bounds relative to the LSM-based ones. We thus bring to light the near optimality of the LSM-based policy.

Unlike LSM, BCD-based PO is unable to deal with instances that have time horizons longer than twenty-four monthly periods, a practically relevant setting, because of the excessive memory requirement of our BCD method. We conjecture that the insights established in chapter one, in particular the near optimality of the LSM-based policies, continue to hold in these cases. To substantiate this claim, in chapters two and three we extend the applicability of PO to such instances.

Our BCD method exhibits time and space requirements that are both cubic functions of the length of the time horizon and the number of sample paths used to formulate our PO linear program. The second chapter develops an alternating direction method of multipliers (ADMM) technique with time and space complexities that are both quadratic and linear functions of the time horizon and number of samples, respectively, to near optimally solve the dual of the PO linear program. Given a close-to-optimal dual vector, we recover a primal solution by approximately solving via regression a set of conditions inspired by complementary slackness (CS). This approach is both efficient and effective when applied to the known benchmark merchant ethanol production instances used in chapter one, but the regression step cannot deal with the larger instances that we wish to solve. We propose to overcome this obstacle by considering only a subset of the CS-type conditions or adopting incremental regression methods when using all such conditions.

Motivated by the success of aggregation-disaggregation approaches in commercial solvers (e.g., Gurobi), chapter three puts forth such an iterative scheme to directly solve the (primal) PO linear program. The aggregation step combines subsets of the constraints and variables of this model to obtain a relaxed linear program that an off-the-shelf solver can handle. The disaggregation step strengthens this relaxation by efficiently identifying constraint violations in the original linear program, exploiting its structure to split some of the aggregated constraints in the relaxed one. This method provably converges. It can be stopped once the current solution is sufficiently close to the optimal one, which we can check using a bound that we compute when disaggregating constraints, or it leads to good-enough bounds on our MDP optimal policy value. Without aggregation, our approach is equivalent to constraint generation. We have verified both the efficiency and effectiveness of this version of our method on the known benchmark merchant ethanol production instances employed in chapter one. We propose to apply its version that features aggregation and disaggregation to the large size instances to be also used in chapter two.