RTE/Ecole Polytechnique: Internship / Offre de stage

Solving many similar linear programs

Learning to speed up the large-scale simulation of energy systems

- Leo Liberti, CNRS LIX Ecole Polytechnique ⟨leo.liberti@polytechnique.edu⟩
- Antoine Oustry, Thomas Bittar, Manuel Ruiz, RTE

As part of its public service mission, RTE (Réseau de Transport d'Électricité), the French transmission system operator (TSO), conducts prospective studies to evaluate the long-term security of electricity supply, which involves performing intensive probabilistic simulations of the electricity system projected in the future. One of the challenges that arise in this context is that of solving many (from thousands to millions) different linear programs (LP) with similar sizes and, for a given size, with the same matrix structure (positions of zeros/nonzeros in the constraint matrix), but with different values in the nonzero entries.

Even if two LPs share the same sizes and matrix structure, they may have very different optimal solutions: and so, in order to obtain these optimal solutions, one should solve each LP individually with either the simplex algorithm or the interior point method. And yet, because of the high number of LPs to be solved, such methods may take many hours or even days to complete the task, which is too long.

The focus of this internship is to devise and test methods based on a mixture of unsupervised and supervised learning algorithms to speed up the entire solution process. The simplest idea is to store LP solutions (optimal vector and basis) in a hash table indexed by the augmented constraint matrix, and then, for each new LP P that comes along, construct its solution using the optimal basis of the entry in the hash table that has key closest to P. More sophisticated ideas to analyse and test are:

1. form a collection of hash tables from the exact solutions of many sequences of LPs (all having the same size) and treat this as a training set for unsupervised (clustering) and supervised (prediction) learning techniques;
2. use random projections (RP) to compress the key (augmented constraint matrix) and store the optimal solution. This offers some advantages: (i) even LPs with slightly different sizes can refer to the same training set because the random projections can be configured so that they all have the same size; (ii) the projected optima and bases can be used to recover feasible (albeit perhaps sub-optimal) solutions of the original LPs that are not too far from the optima, so a few iterations of interior point method should give almost optimal solutions; (iii) with randomly projected data, the risk of having similar LPs with wildly different optima becomes smaller.

The goal of this internship is to shed light on the ideas sketched above, and, in general, reason about the problem of solving many different LPs with same or similar sizes and structures. Numerical experiments will be carried out on realistic test cases, derived from the prospective modelling of the European electricity system.

Depending on the results and outcomes of this study, a PhD thesis could be considered at the end of the internship to continue the research work.