

Lecture.

A mixed-integer fractional optimization approach to best subset selection

We consider the best subset selection problem in linear regression, i.e., finding a parsimonious subset of the regression variables that provides the best fit to the data according to some predefined criterion. We show that, for a broad range of criteria used in the statistics literature, the best subset selection problem can be modeled as a mixed-integer fractional optimization problem. Then we show how to exploit underlying submodularity in the problem to strengthen the formulations, and propose to tackle the problem by solving a sequence of mixed-integer quadratic optimization problems. The proposed approach can be implemented with off-the-shelf solvers and, in our computations, it outperforms existing alternatives by orders of magnitude.