A fundamental machine learning task is to select amongst a set of features to include in a model. In this module, you will explore this idea in the context of multiple regression, and describe how such feature selection is important for both interpretability and efficiency of forming predictions.

To start, you will examine methods that search over an enumeration of models including different subsets of features. You will analyze both exhaustive search and greedy algorithms. Then, instead of an explicit enumeration, we turn to Lasso regression, which implicitly performs feature selection in a manner akin to ridge regression: A complex model is fit based on a measure of fit to the training data plus a measure of overfitting different than that used in ridge. This lasso method has had impact in numerous applied domains, and the ideas behind the method have fundamentally changed machine learning and statistics. You will also implement a coordinate descent algorithm for fitting a Lasso model.

Coordinate descent is another, general, optimization technique, which is useful in many areas of machine learning.

Less