Title: Two models of double descent for weak features

Abstract:
The "double descent" risk curve was recently proposed to qualitatively describe the out-of-sample prediction accuracy of variably-parameterized machine learning models. This article provides a precise mathematical analysis for the shape of this curve in two simple data models with the least squares/least norm predictor. Specifically, it is shown that the risk peaks when the number of features p is close to the sample size n, but also that the risk decreases towards its minimum as p increases beyond n. This behavior is contrasted with that of "prescient" models that select features in an a priori optimal order.