

Sufficient statistics for online prediction via the Burkholder method

Sasha Rakhlin *

rakhlin@mit.edu

Martingale inequalities and bounds for online prediction methods have one thing in common: they are often proved using a “dynamic programming” approach. In this talk, we discuss a surprising equivalence between existence of online prediction algorithms, validity of certain martingale inequalities, and existence of special Burkholder functions. Building on these connections, we define a notion of a sufficient statistic for online methods and exhibit algorithms that only keep these sufficient statistics in memory. We demonstrate the power of the approach by deriving a new algorithm for matrix completion, as well as several other problems. In particular, concentration inequalities for matrix-valued martingales suggest a sufficient statistic that an online procedure should keep track of; we then find the corresponding Burkholder function and exhibit an efficient algorithm that significantly improves upon the standard adaptive-gradient-descent approach.

This is joint work with Dylan Foster and Karthik Sridharan.

*Center for Statistics, Department of Brain & Cognitive Science, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA