



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

Statistics Colloquium

GUIDO MONTÚFAR

Departments of Mathematics and Statistics
University of California, Los Angeles

The Geometry of Memoryless Stochastic Policy Optimization in Infinite-Horizon POMDPs

MONDAY, April 11, at 4:30 PM, via Zoom

Zoom information will be provided via our weekly email announcement.

ABSTRACT

We consider the problem of finding the best memoryless stochastic policy for an infinite-horizon partially observable Markov decision process (POMDP) with finite state and action spaces with respect to either the discounted or mean reward criterion. We show that the (discounted) state-action frequencies and the expected cumulative reward are rational functions of the policy, whereby the degree is determined by the degree of partial observability. We then describe the optimization problem as a linear optimization problem in the space of feasible state-action frequencies subject to polynomial constraints that we characterize explicitly. This allows us to address the combinatorial and geometric complexity of the optimization problem using recent tools from polynomial optimization. In particular, we estimate the number of critical points and use the polynomial programming description of reward maximization to solve a navigation problem in a grid world. This is work with Johannes Müller.

<https://openreview.net/pdf?id=A05I5IvrdL->