Warren B. Powell

Department of Operations Research and Financial Engineering
Princeton University
http://www.castlelab.princeton.edu/

Optimal learning

- Optimal learning refers broadly to the challenge of efficiently collecting information when observations are “expensive” (depends on the context) and noisy.
- Optimal learning arises in different communities motivated by different problem settings (discrete event simulation, laboratory experiments, field experiments), in both off-line and on-line settings.
- We have been pursuing the generality of a simple idea called the *knowledge gradient* which makes measurements based on the marginal value of an observation. For an early tutorial, see

  W. B. Powell, P. I. Frazier, “Optimal Learning: A Tutorial”

The Knowledge Gradient

- The knowledge gradient (the value of measuring *x*) is given by

$$\nu_x^{KG,n} = E\left\{\max_y F(y, K^{n+1}(x))\right\} - \max_y F(y, K^n)$$

where
$$y = \text{Implementation decision}$$
$$K^n = \text{Our state of knowledge after } n \text{ measurements.}$$
$$F(y, K^n) = \text{Value of making decision } y \text{ given knowledge } K^n.$$
$$x^n = \text{Measurement decision after } n \text{ measurements.}$$
$$W_x^n = \text{Observation resulting from observing } x^n = x$$
$$K^{n+1}(x) = \text{Updated distribution of belief about costs after observing } W_x^{n+1}$$

  - For offline problems, it is the only stationary policy that is myopically and asymptotically optimal.
  - Value of information can be nonconcave. KG(*) algorithm provides a simple strategy to adapt to this situation.

The Knowledge Gradient

- We have been exploring a range of problems:
  - Offline and online problems
  - Increasing complexity of types of measurements
    - Discrete, scalar continuous, multidimensional continuous, subset selection, categorical
  - Increasing generality in the representation of the belief
    - Independent beliefs
    - Correlated beliefs (known covariance function)
    - Parametric models (linear regression)
    - Nonparametric models
    - Learning on graphs, learning parameters of linear programs, value functions for MDP’s, dosage strategies,
- For more information, see

http://www.castlelab.princeton.edu/optimallearning.htm

Other challenges

- How do we simulate intelligence?
  - Often we need to make intelligent decisions within a simulator (optimization).
  - Simulation community tends to focus on tuning myopic policies.
  - Approximate dynamic programming offers a simulation-friendly strategy for optimizing over time, but can introduce difficult machine learning challenges.
  - Recent chapter lists nine types of policies (click here).
  - Does every problem need to be customized? What is the hope of general purpose modeling and algorithmic strategies?