Optimum Follow the Leader Algorithm

Open Problem

Dima Kuzmin

Manfred K. Warmuth

University of California - Santa Cruz, USA

June 28, COLT 2005

Weighted Majority

- For expert setting, n experts, expert i has loss L_i
- Pick an expert with probability $p_i \sim e^{-\eta L_i}$
- Loss bound:

$$L^* + \sqrt{2}\sqrt{L^* R \log n} + O(\log n)$$

[FS97]

- $\sqrt{2}$ factor in some sense optimal [Vov97]

Follow the Perturbed Leader

- New algorithm for choosing an expert [KV05]
- Add random noise ν_i to expert losses and pick expert with smallest perturbed loss
- Same form of the expected loss bound, but $\sqrt{2}$ factor is replaced by 2.
- Can be applied whenever minimum is easy to compute (e.g. online shortest path problem)

WM as FPL

- In expert setting it is possible to implement Weighted Majority as min calculation
- Replace vector of losses with independently drawn exponential random variables with parameters $\lambda_i = e^{-\eta L_i}$
- Pick the expert that has the smallest perturbed loss
- This process implements WMR - $p_i \sim e^{-\eta L_i}$
- Other distributions work

Efficiency

- FPL can be implemented efficiently whenever min calculation is efficient
- Representative problem - online shortest-path
- Straightforward implementation of WM - one weight per path - exponentially many
- FPL only needs one weight for every edge and adds noises to these weights
- However, efficient implementation of WM exists - one weight per edge plus upper bound on path lengths [TW03]
- Question: best possible constant?

Distribution Requirements

- Implementing WM as a min calculation for shortest path problem
- The distribution D of random variable Z parametrized by L has to satisfy the following two properties:
  1. If $Z_1 \sim D(L_1), \ldots, Z_n \sim D(L_n)$ are independent random variables, then
  $$P(\arg\ \min(Z_1, \ldots, Z_n) = i) \sim e^{-L_i}$$
  2. If $Z_1 \sim D(L_1)$ and $Z_2 \sim D(L_2)$ are independent random variables, then
  $$Z_1 + Z_2 \sim D(L_1 + L_2)$$

A Few More Things

- Conditions one appears to be related to closure of distributions under minimum. E.g. all distributions that we know to satisfy 1, also satisfy $\min(Z_1, \ldots, Z_n) \sim D(e^{-L_1} + \ldots + e^{-L_2})$
  So we can ask is there D s.t.
  - $\min(Z_1, Z_2) \sim D(\lambda_1 + \lambda_2)$
  - $Z_1 + Z_2 \sim D(\lambda_1 \lambda_2)$
- A possible problem. Let all edge losses be zero. Then the minimum-path distribution satisfying our conditions would have a parameter related to the number of all paths. Estimating it would solve a #P-complete problem.

Questions

- Does there exist a distribution closed under minimum and addition?
- What is the best constant achievable by FPL-type algorithms?
- What is the class of problems for which WM can be efficiently implemented?