Mechanism Design Notes

Clay Thomas
claytont@princeton.edu

March 14, 2019

1 Regret Minimization and Correlated Equilibrium Concepts

Consider the following online learning problem: There is a set of actions A, $|A| = n$. For each time step $t = 1, \ldots, T$, the decision maker picks a distribution on actions p^t. Afterwards, a cost vector $c^t : A \to [0, 1]$ is selected (in any way - even by an adversary who can see p^t before it picks c^t) and the decision maker "incurs cost" $\mathbb{E}_{a \sim p^t}\left[c^t(a)\right] = \langle p^t, c^t \rangle$. Additionally, the decision maker learns the vector c^t. The goal is to minimize the sum of the costs $\mathbb{E}_{a \sim p^t}\left[c^t(a)\right]$.

$$\sum_{t=1}^{T} \min_a c^t(a)$$

Can we hope to achieve something approximating this? Not quite: an adversary who can first look at p^t can, for example, make all costs very high, except the one with smallest probability in p^t. A more realistic goal is to compare our cost to the following:

$$\min_a \sum_{t=1}^{T} c^t(a)$$

Intuitively, "the cost of the best action which, in hindsight, we could've always taken".

This leads us to define the *regret* of our learning algorithm as

$$\mathbb{E}_{a^t \sim p^t, t=1,\ldots,T}\left[\sum_{t=1}^{T} c^t(a)\right] - \min_a \sum_{t=1}^{T} c^t(a) = \sum_{t=1}^{T} \left\langle p^t, c^t \right\rangle - \min_a \sum_{t=1}^{T} c^t(a)$$

It turns out there is a simple algorithm that achieves low regret.

Parameters: $\epsilon > 0$, T
Initialize $w^1(a) = 1$ for each $a \in A$
for $t = 1, \ldots, T$ **do**
  Choose action $a \in A$ according to distribution p^t proportional to w^t
    i.e. let $p^t(a) = w^t(a)/\Phi^t$ where $\Phi^t = \sum_{a \in A} w^t(a)$
  Observe costs c^t and update weights as follows for each $a \in A$:
    $w^{t+1}(a) = w^t(a)(1 - \epsilon)^{c^t(a)}$.

Theorem 1.1. *Let $OPT = \min_a \sum_{t=1}^{T} c^t(a)$. If we set $\epsilon = \sqrt{\log n / T}$, then after T rounds the total realized cost of the multiplicative weights algorithm is $OPT + O(\sqrt{T \log n})$.*

Proof. The proof follows two observations. First, if the decision maker makes a bad move (i.e. the weight of costly moves is high) then lots of total weight is lost. Second, the best action in hindsight contributes has weight $(1 - \epsilon)^{OPT}$ at time T. Intuitively, this says that the only way for the decision maker to have many bad moves is if every action was very bad.

For the formal proof, start by bounding Φ^{t+1} in terms of Φ^t. In what follows, almost all of the inequalities are standard approximations, which can be proved using Taylor series or other estimations. This is somewhat interesting, and basically just happens because ϵ is so small we can go back and forth between exponentiation and addition really easily. In a sense, this was exactly the property we needed for our multiplicative weights to produce a low additive cost measure. Denote by $\ell^t = \mathbb{E}_{a\sim p^t}\left[c^t(a)\right]$ be the realized cost at time t.

$$\begin{aligned}
\Phi^{t+1} &= \sum_a w^t(a)(1-\epsilon)^{c^t(a)} \\
&\le \sum_a w^t(a)(1-\epsilon c^t(a)) \\
&= \sum_a w^t(a) - \Phi^t \epsilon \sum_a \frac{w^t(a)}{\Phi^t} c^t(a) \\
&= \Phi^t \left(1 - \epsilon \sum_a p^t(a) c^t(a)\right) \\
&= \Phi^t \left(1 - \epsilon \ell^t\right)
\end{aligned}$$

Now we can relate Φ^T to the total loss $L = \sum_{t=1}^T \ell^t$:

$$\begin{aligned}
\Phi^T &= \Phi^0 \prod_{t=1}^T \left(1-\epsilon\ell^t\right) \\
&\le n \prod_{t=1}^T \exp\left(-\epsilon\ell^t\right) \\
&= n \exp\left(-\epsilon \sum_{t=1}^T \ell^t\right) \\
&= n \exp\left(-\epsilon L\right)
\end{aligned}$$

On the other hand, for any action $a \in A$ we have

$$(1-\epsilon)^{OPT} \le w^T(a) < \Phi^T \le n\exp(-\epsilon L)$$

$$OPT(-\epsilon - \epsilon^2/2) \le OPT \log(1-\epsilon) \le \log n - \epsilon L$$

$$\begin{aligned}
L &\le OPT + (\epsilon/2)OPT + \frac{\log n}{\epsilon} \\
&\le OPT + (\sqrt{\log n/T}/2)T + \frac{\log n}{\sqrt{\log n/T}} \\
&= OPT + \frac{3}{2}\sqrt{T\log n}
\end{aligned}$$

□