15-251: Great Theoretical Ideas In Computer Science

Recitation 14 Solutions

Important concepts from lecture

- A **Nash equilibrium** is a choice function from players to strategies such that no one player will benefit from changing their strategy.
- The **social cost** of a given solution (strategy choice function) is the sum of the costs to all players of the resulting game outcome.
- The **price of anarchy** is a metric to compare the social costs of "selfish" play (Nash equilibria) and some form of "cooperation" (the social-cost-optimal solution).
- A **consistent hypothesis** with respect to a set S of labelled points is any hypothesis that labels every point in S correctly.
- A **PAC learning algorithm** is one that, for any $\epsilon, \delta > 0$, any distribution D, and any $m_0(\epsilon, \delta)$ training points distributed according to D, has probability at least $1-\delta$ of arriving at a hypothesis whose error with respect to D is at most ϵ.

The Only Winning Move

Consider the two-player game where each player, without knowledge of the other's choice, chooses a strategy from $S = \{0, 1\}$. Player 1 wins \$100 iff they both choose the same strategy, and Player 2 wins \$100 otherwise.

a) Show that there is no Nash equilibrium.

b) How can we modify S to preserve the nature of the game yet make it so that there will be at least one Nash equilibrium? (Hint: imagine actually playing this game, perhaps repeatedly. How can we make the given model more realistic?)

c) Characterize the set of Nash equilibria after this modification.

a) $(0, 0), (1, 1)$ are not equilibria because player 2 would benefit from changing their strategy. Likewise, $(0, 1), (1, 0)$ are not equilibria because player 1 would benefit from changing their strategy.

b) We can extend S to include probabilistic mixtures of strategies if we let $S' = [0, 1]$ where the strategy r represents playing 1 with probability r and playing 0 with probability $1 - r$.

c) $(0.5, 0.5)$ turns out to be the only Nash equilibrium. With a bit of casework we can verify that neither party benefits from strategy-switching.
To prove uniqueness: AFSOC that some other NE (a, b) exists. If $a > 0.5$, then player 2 would benefit from lowering their strategy to 0. (If player 2's strategy is already 0, then player 1 would benefit from switching to also play 0.) If $a < 0.5$, then player 2 would benefit from raising their strategy to 1. (If player 2's strategy is already 1, then player 1 would benefit from switching to also play 0.) In the last case, where $a = 0.5 \neq b$, player 1 would benefit from switching to play 0 if $b < 0.5$ or 1 if $b > 0.5$.

Alg-chemy

In lecture we saw that, if we're given an algorithm that always finds a consistent hypothesis, we can construct a PAC learning algorithm by finding a hypothesis consistent with some $m_0(\epsilon, \delta)$ points sampled from the input distribution.

Let's try going the other way: given a set S of labelled points, some $\delta > 0$, and a PAC learning algorithm A that may not always output a consistent hypothesis, devise a procedure to find a hypothesis that is consistent with S with probability at least $1 - \delta$.

Given S, let D' be the uniform distribution over points in S. Choose $\epsilon' < \frac{1}{|S|}$ because any hypothesis having error at most ϵ' with respect to D' must actually have error 0. (This is because mislabelling any one example in S will result $\frac{1}{|S|}$ error.
Then we can simply set $\delta' = \delta$, sample $m_0(\epsilon', \delta')$ samples from D', and use our PAC algorithm A. With probability at least $1 - \delta' = 1 - \delta$, the hypothesis output will be consistent with S.

Intersection classes

Let C_1 and C_2 be two concept classes. Define the "intersection class"

$$C = \{c \mid \exists c_1 \in C_1, c_2 \in C_2.\ \forall x \in X.\ c(x) = + \iff c_1(x) = c_2(x) = +\}$$

which is to say that every concept $c \in C$ is the intersection of some $c_1 \in C_1$ and $c_2 \in C_2$. Recall that for any set of examples S and any concept class C', $\pi_{C'}(S)$ is the number of ways of labeling examples in S using concepts from C'. Let $\pi_{C'}(m)$ be the max of $\pi_{C'}(S)$ over all m-sized example sets S, and show that $\pi_C(m) \leq \pi_{C_1}(m) \cdot \pi_{C_2}(m)$.

Take any set X of m points. Let k_1, k_2 be the number of distinct subsets of X labelled $+$ by the concepts in C_1 and C_2 respectively. Note that $k_1 \leq \pi_{C_1}(X) \leq \pi_{C_1}(m)$ and $k_2 \leq \pi_{C_2}(X) \leq \pi_{C_2}(m)$. Now the subsets of X that are labelled $+$ by the concepts in C are formed by intersections of the subsets of X labelled $+$ by the concepts in C_1 and the subsets of X labelled $+$ by the concepts in C_2. So the number of distinct subsets of X labelled $+$ by the concepts in C satsifies $\pi_C(X) \leq k_1 \cdot k_2 \leq \pi_{C_1}(m) \cdot \pi_{C_2}(m)$. Since this holds for all X of size m, we conclude that $\pi_C(m) \leq \pi_{C_1}(m) \cdot \pi_{C_2}(m)$.