TUTORIAL VII

1 Entropy in combinatorics

Our goal is to find a lower bound on the number of comparisons we need to sort a list containing n distinct elements, in the worst case. Let $a_1 < a_2 < \cdots < a_n$ be the sorted elements of our list. There is a unique permutation σ such that the initial list is written $[a_{\sigma(1)}, \ldots, a_{\sigma(n)}]$. Note that when you sort a list by doing comparisons, the exact values of the a_i's do not matter, all that matters is the initial permutation σ of the elements. So in the following, we will associate an initial list with a permutation of $\{1, \ldots, n\}$.

Let X be a random permutation of $\{1, \ldots, n\}$ chosen with uniform distribution. Let A be a sorting algorithm performing t comparisons for all initial permutations (if it needs less for some permutations, just add some useless comparisons). Denote by $Y_1, Y_2, \ldots, Y_t$ the outcome of the comparisons performed by A on the input X (for example, $Y_1 = 0$ if $X[1] > X[2]$, and $Y_1 = 1$ otherwise).

1. Compute $H(X|Y_1, \ldots, Y_n)$.
2. Show that $H(X) \leq t$.
3. Conclude by finding $H(X)$.

2 Algorithmic approach to the channel coding problem

The main objective here is to take an algorithmic approach for the channel coding problem. The input to our algorithmic problem is the specification of a noisy channel W from an input set $\mathcal{X}$ to an output set $\mathcal{Y}$. We would like to send M messages and we ask what is the minimum error probability that we can achieve.

This will be a good opportunity to introduce *submodular* functions which is an interesting property to keep in mind and a rich area of study in optimization and approximation algorithms.

1. (Maximization of submodular functions)
   A function $f : 2^{\mathcal{X}} \to \mathbb{R}_+$ taking as input a subset $S \subseteq X$ that has the following property.

$$f(S \cup T) + f(S \cap T) \leq f(S) + f(T) \,. \tag{1}$$

   It is said to be monotone if $f(S) \leq f(T)$ whenever $S \subseteq T$.

   (a) Show that an equivalent definition for submodular function is that $f(T \cup \{j\}) - f(T) \leq f(S \cup \{j\}) - f(S)$ for any $S \subseteq T$ and any $j \in \mathcal{X} - T$. This can be interpreted as a "diminishing returns" property.

   (b) (Remark: this question is independent of the following questions) Let $Z_1, \ldots, Z_n$ be a family of random variables. For a subset $S \subseteq \{1, \ldots, n\}$, let Z_S be the collection of random variables $\{Z_i\}_{i \in S}$. Show that $f(S) = H(Z_S)$ is a submodular and monotone function.

   (c) Let f be a submodular, monotone and nonnegative function and consider the following optimization problem $\max_{S \subseteq \mathcal{X}, |S| = M} f(S)$. Let S^* of size M be such that $f(S^*) = \max_{S \subseteq \mathcal{X}, |S| = M} f(S)$. Computing such an S^* is computationally hard in general. But there is a natural greedy algorithm for this problem: start with $S_0 = \emptyset$, then choose $S_{i+1} = S_i \cup \arg\max\{f(S_i \cup \{j\}) : j \in \mathcal{X} - S_i\}$. Show that $f(S^*) \leq f(S_i) + M(f(S_{i+1}) - f(S_i))$.

(d) Prove that $f(S^*) - f(S_{i+1}) \leq (1 - \frac{1}{M})(f(S^*) - f(S_i))$.

(e) Conclude that the greedy algorithm gives a constant factor approximation for this problem (and say what the constant is).

2. (Channel coding as a submodular optimization problem) Let $\mathrm{S}(W, M)$ be the largest average success probability of a code for M messages:

$$\mathrm{S}(W, M) = \max_{e,d} \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \mathcal{Y}: d(y)=i} W(y|e(i)) \,, \tag{12}$$

where the maximization is over functions $e : \{1, \ldots, M\} \to \mathcal{X}$ and $d : \mathcal{Y} \to \{1, \ldots, M\}$.

(a) Show that $\mathrm{S}(W, M)$ can be written as maximizing some function f over all subsets of $\mathcal{X}$ of size M. Then show that f is submodular and monotone.

(b) Conclude that it is possible to efficiently (here efficiently means polynomial in the description of the channel W and of M) find a code that achieves a success probability that is at least $(1 - 1/e) \cdot \mathrm{S}(W, M)$.