IITM-CS6100 : Topics in Design and analysis of Algorithms Given on: Feb 27

Problem Set #1 Due on : Mar 14, 23:55 Evaluation Due on : Mar 20

Name: Arabhi Subhash Roll No. : CS17B005

Project Team: Varun(CS17B011), Jaitesh(CS17B021), Krishna(CS17B012)

- Turn in your solutions electronically at the moodle page. The submission should be a pdf file typeset either using LaTeX or any other software that generates pdf. No handwritten solutions are accepted.
- Collaboration is encouraged, but all write-ups must be done individually and independently. For each question, you are required to mention the set of collaborators, if any.
- Submissions will be checked for **plagiarism**. Each case of plagiarism will be reported to the institute disciplinary committee (DISCO).

1. (10 points) Suppose $p_1, \ldots, p_n \in [0,1]^2$. Let $\mathsf{TSP}(p_1, \ldots, p_n)$ denote the smallest cost of a traveling salesman's tour on $p_1, \ldots, p_n$ with respect to Euclidean distance.

   (a) (6 points) For any $n > 0$, show that $\mathsf{TSP}(p_1, \ldots, p_n) \leq c\sqrt{n}$ for some constant $c > 0$.

Solution: Collaborators: Project Team
To prove the upper bound on TSP we need to split $[0,1]^2$ into strips of height $\dfrac{1}{\sqrt{n}}$ (i.e. $\sqrt{n}$ no. of strips). Now consider a journey in which we start from bottom one, traverse left to right, move to above strip, travel in opposite direction and continue this till the last point and then join it with the starting point. The length of optimal journey should be definitely $\leq$ to length of the chosen one. Adding to that, even in our journey the total distance $\leq$ the sum of horizontal and vertical distances travelled (by triangle inequality between successive points and summing up).

Horizontal Distance: As we are moving in one direction (horizontally) in a strip the total horizontal distance travelled in a strip cannot be greater than its length (i.e. 1 unit). Hence the total horizontal distance in the grid cannot be greater than 1×no. of strips (i.e. $1 \times \sqrt{n}$)

Vertical Distance: There are n points in all strips and in the worst case, each of them are consecutively on the top and bottom edges of the strips. So the total distance in all strips can be bounded by $n\times$height of strip (i.e. $n \times \dfrac{1}{\sqrt{n}} = \sqrt{n}$). There is also the travel from one strip to the next one that is no. of strips$\times$2$\times$width of strip (i.e. $\sqrt{n} \times \dfrac{2}{\sqrt{n}} = 2$)

So the TSP is bounded by sum of horizontal distance, vertical distance and distance between first and last points (bounded by $\sqrt{2}$, max. distance between any 2 points).

$$TSP \leq \sqrt{n} + \sqrt{n} + 2 + \sqrt{2} \leq 2 \times \sqrt{n} + 2 + \sqrt{2} \tag{1}$$

As n is natural, n $\geq$ 1, $\sqrt{n} \geq 1$. Hence

$$TSP \leq (4 + \sqrt{2}) \times \sqrt{n} \leq 6 \times \sqrt{n} \tag{2}$$

(b) (4 points) for any $n > 0$, show that there are n points $q_1, \ldots, q_n \in [0,1]^2$ such that $\mathsf{TSP}(p_1, \ldots, p_n) \geq c'\sqrt{n}$ for some constant $c' > 0$.

Solution: Collaborators: Project Team
We know that TSP $\geq$ n $\times$ smallest distance among all pairs of points (As TSP is sum of n distances among those points and each is $\geq$ distance between the nearest pair). Divide the $[0,1]^2$ into grid with small squares of side $\dfrac{1}{\sqrt{n}}$, we get n such squares. Now place n points at the centers of each small squares. The smallest distance between any pair of points here is the distance between centres of adjacent squares (i.e. $\dfrac{1}{\sqrt{n}}$). Now by the inequality described above, for this set of points, TSP $\geq$ n $\times \dfrac{1}{\sqrt{n}} \geq \sqrt{n}$. Therefore these are our n points $(q_1, \ldots, q_n)$ that proves the bound to be tight.

2. (7 points) This exercise to demonstrate the limitations of considering expected running time of an algorithm as a useful measure. For any $n > 0$, describe a function $f : \{0,1\}^n \to \mathbb{N}$ such that

- $\mathbb{E}[f] = \mathbb{E}_{x \in \{0,1\}^n}[f(x)] = n^c$ for some constant c and
- $\mathsf{var}[f] = E[f^2] - (E[f])^2 = \Omega(2^n)$.

I.e., there can be performance measures f which is polynomial in expectation, but variance being exponential. Give formal justification for your answer (i.e., computation of expectation and variance for the function f constructed).

Solution: Let x $= x_1, x_2, \ldots, x_n$ be the inputs to the function where $x_i \in \{0,1\}\ \forall\ i \in \{1, \ldots, n\}$ (lets denote this set as X). Now define f as -

$$f(x_1, \ldots, x_n) = n^c \times 3^{(\Sigma_{i=1}^{n} x_i)} \times (-1)^{n-(\Sigma_{i=1}^{n} x_i)}$$

Computation of E[f] -

$$E[f] = \frac{1}{2^n} \times \Sigma_{x \in X} f(x)$$

This summation can be grouped based on no. of set bits or $\Sigma_{i=1}^{n} x_i$ = t where t in running from $0, 1, 2, \ldots, n$ and each of terms occurring $\binom{n}{t}$ times.

$$E[f] = \frac{1}{2^n} \times n^c \times \Sigma_{t=0}^{n} (3^t \times (-1)^{n-t} \times \binom{n}{t})$$

The summation in RHS of the above equation is expansion of $(x+y)^n$. Hence

$$E[f] = \frac{1}{2^n} \times n^c \times (3-1)^n$$

$$= n^c$$

Computation of var(f), which is equal to $E[f^2]$ - $(E[f])^2$ ($n^{2\times c}$, by above result).

$$E[f^2] = \frac{1}{2^n} \times \Sigma_{x \in X} f^2(x)$$

Performing grouping similar to above

$$E[f^2] = \frac{1}{2^n} \times n^{2\times c} \times \Sigma_{t=0}^{n} (3^t \times (-1)^{n-t})^2 \times \binom{n}{t}$$

$$= \frac{1}{2^n} \times n^{2\times c} \times \Sigma_{t=0}^{n} 9^t \times \binom{n}{t}$$

$$= \frac{1}{2^n} \times n^{2\times c} \times (1+9)^n$$

$$= n^{2\times c} \times 5^n$$

Therefore $var(f) = n^{2\times c} \times (5^n - 1)$. This is equal to $\Omega(2^n)$ because Ω is lower bound on complexity and

$$O(n^{2\times c} \times (5^n - 1)) = O(n^{2\times c} \times 5^n) > O(2^n)$$

Hence we can say that our function satisfies both expectation and variance constraints.

3. (7 points) Obtain profits $p_1, \ldots, p_n$ and weights $w_1, \ldots, w_n$ for the knapsack problem so that $|\mathcal{P}|$ is exponential in n (e.g $2^{\Omega(n)}$). Justify your answer.

Solution: Our solution set is $x_1, \dots, x_n \; \forall x_i \in \{0, 1\}$. The cardinality of the set is 2^n. So as to obtain exponential in n (equal to the order of entire solution set) pareto optimal solutions we need to make sure no solution dominates other. This is satisfied when profits and weights satisfy (assume profits are in ascending order without loss of generality) -

$$\Sigma_{j=1}^{i-1} p_j < p_i \; \forall \; i \in \{2, \dots, n\}$$

$$\Sigma_{j=1}^{i-1} w_j < w_i \; \forall \; i \in \{2, \dots, n\}$$

To be more precise, the case where both the above equations are equal for some values of i is also fine but not the case where one is equal and not other. There are other cases too, like when weights and profits are equal, which guarantee exponential number of pareto optimal solutions but here we are looking at one of the conditions which satisfy the property.

Justification: Consider two solutions to be a and b and let $i_1, i_2, \dots, i_k$ are the indices of the solution (in ascending order) where a and b differ. Without loss of generality assume $a_{i_k} = 1$ and $b_{i_k} = 0$. Now from above conditions -

$$\Sigma_{j=1}^{i_k-1} p_j < p_{i_k}$$

$$\Sigma_{j=1}^{k-1} p_{i_j} < \Sigma_{j=1}^{i_k-1} p_j < p_{i_k}$$

In the above inequality we can see that, P_b (profit of b due to differed bits) $\leq$ LHS (as b's profit is sum of some or all of $p_{i_1}, \dots, pi_{k-1}$) and RHS $\leq P_a$ (profit of a due to differed bits) (as a's profit is P_{i_k} + sum of some or all of $p_{i_1}, \dots, pi_{k-1}$). Therefore we can say that $P_b < P_a$. Same analysis on weights shows that $W_b < W_a$. These inequalities themselves mean that neither a nor b dominate the other. As we did not assume any constraints on a and b we can say that none of elements in solution set dominate other and all are pareto optimal (i.e. $|\mathcal{P}| = 2^n$).

4. (6 points) Read the proof of Lemma 3.4 in the notes by Bodo Manthey (in the google drive shared with the class). The proof assumes that $a = (0, \dots, 0)$ and $b = (\delta, 0, \dots, 0)$. Why is this assumption without loss of generality? Justify your answer.

Solution: The Lemma is set on premise that c is drawn from a Gaussian distribution with standard deviation σ. So even if a and b are not equal to $(0, \dots, 0)$ and $(\delta, 0, \dots, 0)$, we can transform(shift origin) and rotate the axes of the problem to make it happen, this is fine because the transformation and rotation does not effect the Gaussian Distribution(i.e. σ - the standard deviation for each dimension of c wont change).

The transformation involves shifting origin to a and rotating axes such that vector b-a is along x-axis. W.k.t $E(x + c) = E(x) + c$ and $var(x + c) = var(x)$ hence shifting the origin only moves mean but the variance remains same. After rotation, the ith dimension of new c will be $c_i^{'} = c_1 * l_1 + c_2 * l_2 + \ldots + c_d * l_d$. In a case of pure rotation like this one we know that $l_1^2 + l_2^2 + \ldots + l_d^2 = 1$. We also know that $var(x * c) = c^2 * var(x)$ and $var(\Sigma x) = \Sigma var(x)$ - where each x are independent, like in this case where is each dimension of c is an independent gaussian. By combining these results

$$var(c_i^{'}) = var(c_1 * l_1 + c_2 * l_2 + \ldots + c_d * l_d)$$

$$var(c_i^{'}) = var(c_1 * l_1) + var(c_2 * l_2) + \ldots$$

$$= l_1^2 * var(c_1) + l_2^2 * var(c_2) + \ldots$$

As $var(c_j) = \sigma^2 \; \forall j$ given in the lemma

$$var(c_i^{'}) = (l_1^2 + l_2^2 + \ldots) * \sigma^2$$

$$= \sigma^2$$

Hence we can say that gaussian (and there by sampling of c) wont be affected by transformation or rotation of axes.