15-451/651 Algorithms, Spring 2018

Homework #7 **Due: May 1-3 (Tue-Thu), 2018**

This is an oral presentation assignment. There are only 3 problems, each worth (100/3) points. Otherwise same rules as HW4.

The problems in this HW are simple but have a few parts. Also, they cover new concepts, so please start and finish early!

1. **(Roll Over, Gauss!)** Given an $n \times n$ *symmetric* matrix A and an $n \times 1$ vector b, our goal is to solve the equation $Ax = b$ to high accuracy. We will use gradient descent to solve this problem quickly given some assumptions about A; see Lecture 20 for background on gradient descent. (The analysis here is independent of the one from lecture, but you should be comfortable with the ideas there.)

   Recall from linear algebra that every symmetric $d \times d$ matrix A can be written as $V \Lambda V^T$, where V is an $n \times n$ matrix whose columns are the eigenvectors of A, and Λ is a diagonal matrix whose entries are the eigenvalues of A. Recall from Lecture 20 that for $x \in \mathbb{R}^n$, $\|x\|^2 = \sum_{i=1}^n x_i^2$.

   (a) Consider the function $f(x) = \frac{1}{2}\|Ax - b\|^2$. Prove that the gradient $\nabla f(x) = A(Ax - b)$.

   (b) Suppose $x^* = \text{argmin}_x \frac{1}{2}\|Ax - b\|^2$. State why $A^2 x^* = Ab$.

   (c) Suppose we set $x^{(0)} = 0^n$, and

   $$x^{(t+1)} \leftarrow x^{(t)} - \nabla f(x^{(t)}).$$

   Argue for any $i \geq 0$, $A(x^{(i+1)} - x^*) = (I - A^2)(A(x^{(i)} - x^*))$.

   For the next part, you may find the following statements helpful: (1) for a symmetric matrix B and a vector y, $\|By\| \leq \max(|\lambda_{max}|, |\lambda_{min}|) \cdot \|y\|$ where λ_{max} is the maximum eigenvalue of B and λ_{min} is the minimum eigenvalue of B, and (2) for a symmetric matrix B, the eigenvalues of $I - B$ are in the range $[1 - \lambda_{max}, 1 - \lambda_{min}]$. Please try to prove these facts about eigenvalues yourself for practice, though you will not need to prove these to us in the oral presentation. When all eigenvalues of A are in the range $[.9, 1.1]$, A is said to be *well-conditioned*.

   (d) Argue if all eigenvalues of A are in the range $[.9, 1.1]$, then $\|A(x^{(i+1)} - x^*)\| \leq \frac{1}{2}\|A(x^{(i)} - x^*)\|$.

   (e) For a parameter $\epsilon \in (0, 1)$, argue that if $t = O(\log(1/\epsilon))$, then $\|A(x^{(t)} - x^*)\|^2 \leq \epsilon\|b\|^2$.

   (f) Argue that $\|Ax^{(t)} - b\|^2 = \|A(x^{(t)} - x^*)\|^2 + \|Ax^* - b\|^2$. Hint: you may need that for $x, y \in \mathbb{R}^n$, if $\langle x, y \rangle = 0$, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. You may also find part (b) useful. Please show this part even without making any assumption about the eigenvalues.

(g) Assuming that A has m non-zero entries, what is the overall running time of the algorithm for outputting an $x^{(t)}$ for which $\|Ax^{(t)} - b\|^2 \leq \|Ax^* - b\|^2 + \epsilon\|b\|^2$?

2. **(I shall never see a Billboard lovely as a Tree.)** Given a (multi)graph $G = (V, E)$, you want to buy a minimum cost spanning tree on V. Each edge of the graph is owned by a different seller, so there are $m = |E|$ sellers in the system, and you are the only buyer. (This is the mirror image of the usual auction setup with n buyers and 1 seller.)

   The actual cost c_e of each edge is currently unknown to you, and so you have to ask its seller for this information. (Call the reported costs of e the *bid* b_e.) However, the sellers are self-interested and hence may misreport these costs. We need to design a mechanism (which decides the spanning tree to buy, and the prices to pay to each seller) that makes truth-telling a dominant-strategy for the sellers — they should have no incentive to lie, and hence should give bids b_e that equal the true costs c_e.[1]

   (a) Suppose the graph has only two nodes and m parallel links between them. Show the analog of the Vickery auction (which buys the edge e with the least bid, pays that edge's seller the second-lowest bid value, and pays all other sellers zero) is a DSTT mechanism.

   Notice that if there is a single very cheap edge and the rest are costly, everyone is truthful, but still you have to pay a lot more than the actual cost of that edge you bought. But if there a lot of competition—i.e., there are several providers with near-least costs—then you don't overpay by much. In the rest of this problem we will extend this intuition for general graphs.

   Now we have a general graph $G = (V, E)$, with the property that for any partition of the vertices (i.e., a cut), there are at least 2 edges in G crossing this cut.

   (b) Consider the following mechanism. We buy a minimum spanning tree T (which minimizes the sum of the bid values on its edges). The prices: for each edge not in T, we pay that seller zero. For an edge e in T, we look at the (unique) cut induced by the two components of $T - \{e\}$, look at an edge $f \neq e$ crossing this cut with the smallest bid, and pay e's seller the amount b_f. Prove that this mechanism is DSTT.

   This DSTT property means that the sellers have no incentive to lie, and hence we can assume from now on that the bids are equal to the costs. We will now bound the amount the buyer pays, in terms of the structure of the graph.

   (c) **(optional)** Now suppose T is the MST above. Let T' be the min-cost tree that is edge-disjoint from T (assume such a tree exists).[2] Then construct a bipartite graph $H = (E_T, E_{T'})$ whose vertices represent the edges of T and T' respectively, and there is an edge between (e, e') exactly when $T \cup \{e'\} - \{e\}$ is a spanning tree of G. Show that H has a perfect matching. (You may use Hall's theorem from recitation, without having to prove it yourself.)

[1] From now on we will use DSTT to denote *dominant-strategy truth-telling*, or *incentive compatible.*

[2] This just means: delete the edges of T. T' the MST in the resulting graph.

(d) **(not optional)** Use the result of part (c) to show that the sum of payments made to all the sellers by the mechanism in (b) is at most the cost of the minimum-cost spanning tree in G that is edge-disjoint from T

So you may think of the difference in the cost of the MST T', and the cost of T', as a measure of how competitive the market is.

3. **(Oh Spanner, my Spanner!)** We are given an undirected, unweighted graph $G = (V, E_G)$ on n nodes. Say an edge $\{u, v\}$ of G is *heavy* if both of its endpoints u and v have degree greater than $n^{1/3}$ in G, otherwise the edge is *light*.

   We show how to find a 6-*additive spanner* of G, which we denote $H = (V, E_H)$, having $\tilde{O}(n^{4/3})$ edges. Here is the algorithm:

   i. Include all edges in H which are incident to vertices of degree at most $n^{1/3}$ in G.

   ii. Next we sample a set A of $O(n^{2/3} \log n)$ vertices from V uniformly at random. For each vertex $v \in V$, if v is adjacent to one or more vertices in A, then include an edge in H from v to one arbitrary vertex $a \in A$ which is incident to v.

   iii. Moreover, for $i = 0, 1, 2, \ldots, O(\log n)$, we also sample a set B^i of $O(2^{-i} n^{2/3} \log n)$ vertices from V independently and uniformly at random.

   iv. For each i, for each pair of vertices $u \in A$ and $v \in B^i$, if there is at least one shortest path P between u and v in G containing $O(2^i)$ heavy edges, then we choose one such path P and include in our spanner H all heavy edges in this path P.

   To argue H is a 6-additive spanner with the claimed number of edges, answer the questions below:

   (a) How many edges are in H?

   (b) Consider a shortest path P between an arbitrary pair u and v of vertices in G. Suppose there are k heavy edges in P. Argue that there is a set $S(P)$ of $\Omega(kn^{1/3})$ distinct vertices v *adjacent* to P. That is, for every vertex $s \in S(P)$, there should be a vertex u on P such that $\{u, s\}$ is an edge.

   (c) For each pair a, b of vertices in G, fix $P_{a,b}$ to be an arbitrary shortest path between them in G. Argue that with probability at least $1 - 1/\text{poly}(n)$, the following statement holds:

   > For all $a, b \in V$, if the number of heavy edges along $P_{a,b}$ is in the range $[2^{i-1}, 2^i)$, then the set $B^i \cap S(P)$ is non-empty.

   (d) Argue that H is a 6-additive spanner with probability at least $1 - 1/\text{poly}(n)$.

   Hint: For each pair a, b consider the path $P_{a,b}$ and look at the number of heavy edges on it.