CS264: Homework #7

Due by midnight on Thursday, March 2, 2017

Instructions:

(1) Form a group of 1-3 students. You should turn in only one write-up for your entire group. See the course site for submission instructions.

(2) Please type your solutions if possible and feel free to use the LaTeX template provided on the course home page.

(3) All students should complete all of the exercises. Students taking the course for a letter grade should also complete all of the problems.

(4) Write convincingly but not excessively. Exercise solutions rarely need to be more than 1-2 paragraphs. Problem solutions rarely need to be more than a half-page (per part), and can often be shorter.

(5) You may refer to your course notes, and to the textbooks and research papers listed on the course Web page *only*. You cannot refer to textbooks, handouts, or research papers that are not listed on the course home page. (Exception: feel free to use your undergraduate algorithms textbook.) Cite any sources that you use, and make sure that all your words are your own.

(6) If you discuss solution approaches with anyone outside of your team, you must list their names on the front page of your write-up.

(7) Exercises are worth 5 points each. Problem parts are labeled with point values.

(8) No late assignments will be accepted.

Lecture 13 Exercises

Exercise 39

We stated in lecture that sparse approximation is NP-hard for arbitrary matrices $\mathbf{A}$, based on a reduction from Exact Cover by 3-Sets (X3C). In the X3C problem, we are given a set S and a collection $\mathcal{C}$ of 3-element subsets of S. We want to determine whether $\mathcal{C}$ contains an exact cover for S, i.e., a sub-collection $\widehat{\mathcal{C}} \subseteq \mathcal{C}$ such that every element of S occurs exactly once in $\widehat{\mathcal{C}}$. Give a polynomial-time reduction from X3C to sparse approximation.

Lecture 14 Exercises

Exercise 40

The natural formulation for k-means involves centers for each cluster, and measuring the sum of squared distances to cluster centers. In lecture, we introduced a relaxation which takes a suitable average of pairwise squared distances in each cluster. Show that the two formulations are equivalent.

Problems

Problem 26

This problem explores the "Restricted Isometry Property" (RIP), which is another sufficient condition for exact sparse recovery via our ℓ_1-minimization linear program (LP). Formally, fix an integer $s \in \{1, 2, \ldots\}$ for the rest of the problem. We define the *isometry constant* δ_s of a matrix $\mathbf{A}$ as the smallest number such that

$$(1 - \delta_s)||\mathbf{x}||_2^2 \leq ||\mathbf{Ax}||_2^2 \leq (1 + \delta_s)||\mathbf{x}||_2^2 \tag{1}$$

for every s-sparse vector $\mathbf{x}$.

Analogous to our result from lecture for sparseish vectors, we will prove the following theorem:

Theorem 1 *Suppose we have a matrix* $\mathbf{A}$*, with* δ_{2s} *a sufficiently small constant. Let* $\mathbf{x} \in \mathbb{R}^n$ *be s-sparse, and let* $\mathbf{b} = \mathbf{Ax}$*. Then (LP) returns* $\hat{\mathbf{x}} = \mathbf{x}$*.*

(a) (3 points) Prove that for s-sparse vectors $\mathbf{x}, \mathbf{x}'$ supported on disjoint sets,

$$|(\mathbf{Ax}) \cdot (\mathbf{Ax}')| \leq \delta_{2s}||\mathbf{x}||_2 \cdot ||\mathbf{x}'||_2.$$

[Hint: apply RIP to $(\mathbf{x} \pm \mathbf{x}')$.]

(b) (3 points) Let $\mathbf{h} = \hat{\mathbf{x}} - \mathbf{x}$, and recall that $\mathbf{Ah} = 0$. Assume that $\mathbf{h} \neq 0$ (otherwise we're done). Let's decompose $\mathbf{h}$ into a sum of s-sparse vectors $\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_k$. We let $\mathbf{h}_0$ contain the s coefficients of $\mathbf{x}$, and then assign the remaining coefficients, s at a time, to $\mathbf{h}_1, \mathbf{h}_2, \ldots$ in descending order of magnitude. Prove that:

$$\left\| \sum_{j \geq 1} \mathbf{h}_j \right\|_1 \leq \|\mathbf{h}_0\|_1$$

.

[Hint: Use the fact that $\hat{\mathbf{x}}$ is an optimal solution to the LP.]

(c) (9 points) Next, prove that if δ_{2s} is a sufficiently small constant, then $||\mathbf{h}_0||_1 < ||\sum_{j\geq 1} \mathbf{h}_j||_1$. Conclude by deducing Theorem 1.

[Hint: apply RIP to $(\mathbf{h}_0 + \mathbf{h}_1)$, and use part (a) repeatedly.]

Problem 27

The goal of this problem is to extend the compressive sensing result from Lecture #13 to the case where the ground truth vector $\mathbf{z}$ is only *approximately* k-sparse. Let I denote the k coordinates of $\mathbf{z}$ that maximize the ℓ_1 norm $\sum_{i\in I} |z_i|$, and let r denote the residual ℓ_1 norm $\sum_{i\notin I} |z_i|$ on the coordinates outside of I.

Assume that $\mathbf{A}$ is $\frac{1}{4}\sqrt{\frac{n}{k}}$-sparse-ish, as in lecture. Let $\mathbf{w}$ denote the optimal solution to the ℓ_1-minimizing linear program from lecture.

(a) (7 points) Prove that

$$\|\mathbf{z} - \mathbf{w}\|_1 \leq 2\|(\mathbf{z} - \mathbf{w})_I\|_1 + 2r.$$

[This follows from similar maneuvers to the derivation in lecture.]

(b) (3 points) Conclude that the computed vector $\mathbf{w}$ is almost the same as the ground truth vector $\mathbf{z}$, in the sense that $\|\mathbf{z} - \mathbf{w}\|_1 \leq 4r$.

[Use a result from lecture.]

Problem 28

Consider the k-means linear programming relaxation for an instance with two clusters C_1 and C_2 with n points each. Suppose there exist points $x, y \in C_1$, $z \in C_2$ such that $d(x, y) > d(x, z)$. We want to show that the optimal solution to the k-means LP does not correspond to the partition into clusters C_1 and C_2.

(a) (10 points) Prove the claim for the LP considered in lecture, without symmetry constraints:

$$\begin{aligned} \min \quad & \sum_{i,j} z_{ij} \|p_i - p_j\|_2^2 \\ s.t. \quad & z_{ij} \le z_{ii} \quad \forall i, j \\ & \sum_j z_{ij} = 1 \quad \forall i \\ & \sum_i z_{ii} = k \\ & z_{ij} \ge 0 \end{aligned}$$

(b) (10 extra credit points) Prove or disprove the claim for the LP considered in lecture:

$$\begin{aligned} \min \quad & \sum_{i,j} z_{ij} \|p_i - p_j\|_2^2 \\ s.t. \quad & z_{ij} = z_{ji} \quad \forall i, j \\ & z_{ij} \le z_{ii} \quad \forall i, j \\ & \sum_j z_{ij} = 1 \quad \forall i \\ & \sum_i z_{ii} = k \\ & z_{ij} \ge 0 \end{aligned}$$