Homework 1

Out: *Nov 3* Due: *Nov 17*

Instructions:

- Upload your solutions (to the non-extra-credit) as *a single* PDF file (one PDF total) to Gradescope. Please anonymize your submission (do not list your name in the PDF title or in the document itself). If you forget, it's OK.
- If you choose to do extra credit, upload your solution to the extra credits as a single PDF file to Gradescope. Please again anonymize your submission.
- You may discuss ideas for solutions with any classmates, textbooks, the Internet, etc. Please attach a brief "collaboration statement" listing any collaborators at the end of your PDF. **You must write up your solutions individually.**
- For each problem, you should aim to keep your writeup below one page. For some problems, this may be infeasible, and for some problems you may write significantly less than a page. This is not a hard constraint, but part of the assignment is figuring out how to easily convince the grader of correctness, and to do so concisely. "One page" is just a guideline: if your solution is longer because you chose to use figures (or large margins, display math, etc.) that's fine.
- Each problem is worth ten points (even those with multiple subparts).

Problems:

§1 (10 points) This problem explores compressed sensing schemes that work when noise/numerical precision is not an issue. Let $q_1, \ldots, q_n \in \mathbb{R}^n$ be any set of *distinct* numbers. E.g. we could choose $[q_1, \ldots, q_n] = [1, \ldots, n]$. Consider the sensing matrix $A \in \mathbb{R}^{2k \times n}$:

$$A = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ q_1 & q_2 & q_3 & \ldots & q_n \\ (q_1)^2 & (q_2)^2 & (q_3)^2 & \ldots & (q_n)^2 \\ \vdots & \vdots & \vdots & & \vdots \\ (q_1)^{2k-1} & (q_2)^{2k-1} & (q_3)^{2k-1} & \ldots & (q_n)^{2k-1} \end{bmatrix}$$

Show that, if $x \in \mathbb{R}^n$ is a k sparse vector – i.e. $\|x\|_0 \leq k$ – then x can be recovered uniquely given Ax, which is a vector with length $2k$. You don't need to give an efficient algorithm. Just argue that for any given $y \in \mathbb{R}^{2k}$, there is at most one k-sparse x such that $y = Ax$. (Hint: Use that a non-zero degree d polynomial can't have more than d roots.)

§2 In this problem, we will come up with two alternate characterizations of the minimum distance of a binary linear code. Let $E : \mathbb{F}_2^k \to \mathbb{F}_2^n$ be a linear error correcting code that stretches k bits into n bits. Let $g_i = E(e_i)$ be the encoding of the standard basis vectors $e_1, e_2, \ldots, e_k$ (e_i is the vector with all 0s except exactly one 1 in the i-th coordinate) in the k dimensions. Let G be the $k \times n$ matrix with i-th row equal to g_i.

(a) (2 points) Let $C = \mathsf{Span}(g_1, g_2, \ldots, g_k)$ be the linear subspace $\mathbb{F}_2^n$. Prove that every element of C is an encoding of some message.

(b) (3 points) Argue that minimum distance of the code defined by E equals the smallest number of 1s in any non-zero element of C.

(c) (5 points) Prove that if every subset of k columns of G are linearly independent, then, E has minimum distance $d \geq n - k + 1$. (Hint: use the conclusion from part 1 and remember that if every k columns of G are lin independent then every $k \times k$ submatrix of G must be full rank.)

§3 (10 points)

(a) Let M be the transition matrix of a ergodic random walk with mixing time t_0. Let $M' = 1/2(I + M)$ be the "lazy" version of this Markov Chain. Show that the mixing time of M' is at most $10t_0$. It's fine to have any constant (instead of 10) in this bound.

(b) Let M be the transition matrix of a random walk on an undirected d-regular graph G on n vertices that defines an ergodic Markov Chain with stationary distribution π. In the class, we defined the mixing time of this Markov Chain as the smallest integer t_0 such that for every distribution x on the vertices of G, $\|M^{t_0}x - \pi\|_1 \leq 1/4$. Justify this definition by arguing that the distance to stationary distribution shrinks exponentially: i.e., show that after kt_0 steps, $\|M^{kt_0}x - \pi\|_1 \leq 2^{-k}$.

§4 (10 points) Let M be the Markov chain of a 5-regular undirected graph that is connected. Each node has self-loops with probability 1/2. We saw in class that 1 is an eigenvalue with eigenvector $\vec{1}$. Show that every other eigenvalue has magnitude at most $1 - 1/10n^2$. (Hint: check out the proof in the lecture for why a connected graph canot have two eigenvalues that are equal to 1.) What does this imply about the mixing time for a random walk on this graph from an arbitrary starting point?

§5 (Extra credit) (*Sudan's list decoding*) Let $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n) \in F^2$ where $F = GF(q)$ and $q \gg n$. We say that a polynomial $p(x)$ *describes* k of these pairs if $p(a_i) = b_i$ for k values of i. This question concerns an algorithm that recovers p even if $k < n/2$ (in other words, a majority of the values are wrong).

(a) Show that there exists a bivariate polynomial $Q(z, x)$ of degree at most $\lceil\sqrt{n}\rceil + 1$ in z and x such that $Q(b_i, a_i) = 0$ for each $i = 1, \ldots, n$. Show also that there is an efficient (poly(n) time) algorithm to construct such a Q.

(b) Show that if $R(z, x)$ is a bivariate polynomial and $g(x)$ a univariate polynomial then $z - g(x)$ divides $R(z, x)$ iff $R(g(x), x)$ is the 0 polynomial.

(c) Suppose $p(x)$ is a degree d polynomial that describes k of the points. Show that if d is an integer and $k > (d + 1)(\lceil\sqrt{n}\rceil + 1)$ then $z - p(x)$ divides the bivariate polynomial $Q(z, x)$ described in part (a). (Aside: Note that this places an upper bound on the number of such polynomials. Can you improve this upper bound by other methods?)

(There is a randomized polynomial time algorithm due to Berlekamp that factors a bivariate polynomial. Using this we can efficiently recover all the polynomials p of the type described in (c). This completes the description of Sudan's algorithm for *list decoding*.)