ADVANCED ALGORITHM DESIGN

Homework 3

November 17, 2024

Problem 1

This problem explores compressed sensing schemes that work when noise/numerical precision is not an issue. Let $q_1, \cdots, q_n \in \mathbb{R}$ be any set of distinct numbers. E.g. we could choose $q_i = i$. Consider the sensing matrix $A \in \mathbb{R}^{2k \times n}$:

$$A = \begin{bmatrix} 1 & 1 & \cdots & \cdots & 1 \\ q_1 & q_2 & \cdots & \cdots & q_n \\ q_1^2 & q_2^2 & \cdots & \cdots & q_n^2 \\ \vdots & \vdots & & & \vdots \\ q_1^{2k-1} & q_2^{2k-1} & \cdots & \cdots & q_n^{2k-1} \end{bmatrix}.$$

Show that if $x \in \mathbb{R}^n$ is a $k-$sparse vector, that is, $\|x\|_0 \leq k$, 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 $y = Ax$.

Solution

We assume $n \geq 2k$, that is, A is horizontally wide.

WLOG, $q_1 < \cdots < q_n$. For any index set $S = \{i_1, \cdots, i_{2k}\}$ with $1 \leq i_1 < \cdots < i_{2k} \leq n$, we denote by A_S the $2k \times 2k$ matrix formed by taking only the columns $i_1, \cdots, i_{2k}$ from A. This is a Vandermonde matrix with determinant $\det A_S = \prod_{\alpha > \beta}(q_{i_\alpha} - q_{i_\beta}) \neq 0$.

Let $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$ be $k-$sparse vectors such that $A\boldsymbol{x} = A\boldsymbol{z}$. Take $S := \operatorname{supp}(\boldsymbol{x} - \boldsymbol{z})$ so that $|S| \leq 2k$ (WLOG take it to be $2k$ by adding more indices which could be 0 in $\boldsymbol{x} - \boldsymbol{z}$). WLOG say $S = \{1, \cdots, 2k\}$ in an increasing order. Then A_S is invertible by the previous paragraph. Next note that if $\boldsymbol{v} \in \mathbb{R}^{2k}$ then $\boldsymbol{v}\boldsymbol{e}_i^\top$ is the $2k \times 2k$ matrix whose i^{th} column is all $\boldsymbol{v}$ and 0 everywhere else. Take $\boldsymbol{v} := \boldsymbol{x} - \boldsymbol{z}$ now. The next key observation is that $A_S = \sum_{i \in S} A\boldsymbol{e}_i\boldsymbol{e}_i^\top$ and that $\boldsymbol{v}_S = \sum_{j \in S} \boldsymbol{e}_j\boldsymbol{e}_j^\top \boldsymbol{v}$ where $\boldsymbol{v}_S$ is the restriction of $\boldsymbol{v}$ to only the indices in S. Here $\operatorname{supp} \boldsymbol{v} \subseteq S$. Therefore $A_S\boldsymbol{v}_S = \sum_{i \in S} A\boldsymbol{e}_i\boldsymbol{e}_i^\top \boldsymbol{v} = \sum_{i \in S} A\boldsymbol{e}_i v_i = \sum_{i \in [n]} A\boldsymbol{e}_i\boldsymbol{e}_i^\top \boldsymbol{v} = A\boldsymbol{v}$ where the second last equality is because $\boldsymbol{e}_i^\top \boldsymbol{v} = 0$ if $i \notin S$. But $A\boldsymbol{v} = A(\boldsymbol{x} - \boldsymbol{z}) = \boldsymbol{0}$. This means $A_S\boldsymbol{v}_S \implies v_S = 0 \implies \boldsymbol{x}_S = \boldsymbol{z}_S \implies \boldsymbol{x} = \boldsymbol{z}$ where the last implication is because all coordinates of $\boldsymbol{x}, \boldsymbol{z}$ are 0 at indices in $[n] \setminus S$.

Problem 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 $\boldsymbol{g}_i = E(\boldsymbol{e}_i)$ be the encoding of the standard basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_k$ in the k dimensions. Let G be the $k \times n$ matrix with i^{th} row equal to $\boldsymbol{g}_i$.

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

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

(c) 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 (a) and remember that if every k columns of G are linearly independent then every $k \times k$ submatrix of G must be full rank.)

Solution

Assume E is injective.

(a) Let $\boldsymbol{v} \in C$. Then $\boldsymbol{v} = \sum\limits_{i=1}^{k} a_i \boldsymbol{g}_i$ for some scalars $a_i \in \mathbb{F}_2$. So, $\boldsymbol{v} = \sum\limits_{i=1}^{k} a_i E(\boldsymbol{e}_i) = E\left(\sum\limits_{i=1}^{k} a_i \boldsymbol{e}_i\right)$. So $\boldsymbol{v}$ is the encoding of $\sum\limits_{i=1}^{k} a_i \boldsymbol{e}_i$.

(b) Recall the definition of minimum distance: $\Delta = \min\limits_{\boldsymbol{x},\boldsymbol{y} \in \mathbb{F}_2^k, \boldsymbol{x} \neq \boldsymbol{y}} \|E(\boldsymbol{x}) - E(\boldsymbol{y})\|_0$. By linearity of E, $\Delta = \min\limits_{\boldsymbol{x},\boldsymbol{y} \in \mathbb{F}_2^k, \boldsymbol{x} \neq \boldsymbol{y}} \|E(\boldsymbol{x} - \boldsymbol{y})\|_0 = \min\limits_{\boldsymbol{z} \in \mathbb{F}_2^k, \boldsymbol{z} \neq 0} \|E(\boldsymbol{z})\|_0 = \min\limits_{\boldsymbol{z} \in \mathbb{F}_2^k, \boldsymbol{z} \neq 0} \|E(\boldsymbol{z})\|_0 = \min\limits_{\boldsymbol{v} \in E(\mathbb{F}_2^k) = C, \boldsymbol{v} \neq 0} \|\boldsymbol{v}\|_0$.

(c) Every subset of k columns of G is linearly independent. Note that $\boldsymbol{g}_i = G^\top \boldsymbol{e}_i = E(\boldsymbol{e}_i)$. Say $\boldsymbol{a} = E(\boldsymbol{x}) \in C$ has $\geq k$ zero entries, that is, $\|\boldsymbol{a}\|_0 \leq n - k$. WLOG, entries at $S = \{1, \cdots, k\}$ in $\boldsymbol{a}$ are 0. The submatrix G_S of $G^\top$ formed by taking the first k rows has size $k \times k$ and is full rank, thus invertible. Then $\begin{bmatrix} G_S^{-1} & \mathbf{0}_{k \times (n-k)} \end{bmatrix} G_{n \times k}^\top = I_k$ where I_k is the $k \times k$ identity matrix. Therefore, $\boldsymbol{x} = \begin{bmatrix} G_S^{-1} & \mathbf{0}_{k \times (n-k)} \end{bmatrix} G_{n \times k}^\top \boldsymbol{x} = \begin{bmatrix} G_S^{-1} & \mathbf{0}_{k \times (n-k)} \end{bmatrix} \boldsymbol{a} = \mathbf{0}$ where the last equality is true because the last $n - k$ columns of the matrix are 0 are the first k entries of $\boldsymbol{a}$ are 0. Therefore $\boldsymbol{a} = \mathbf{0}$. This means that our assumption, that $\|\boldsymbol{a}\|_0 \leq n - k$ was false, which proves that $\|\boldsymbol{a}\|_0 \geq n - k + 1$ for any nonzero $\boldsymbol{a} \in C$.

Problem 3

(a) Let M be the transition matrix of a ergodic random walk with mixing time t_0. Let $M' = \frac{1}{2}(I + M)$ 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, $\left\|M^{t_0}x - \pi\right\|_1 \leq \frac{1}{4}$. Justify this definition by arguing that the distance to stationary distribution shrinks exponentially: i.e., show that after kt_0 steps, $\left\|M^{kt_0}x - \pi\right\|_1 \leq 2^{-k}$.

Solution

Lemma 1
Let $\boldsymbol{x}, \boldsymbol{y} \in \Delta_{n-1}$ be two probability distributions, that is, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{x} \geq 0, \boldsymbol{y} \geq 0$ and $\|\boldsymbol{x}\|_1 = \|\boldsymbol{y}\|_1 = 1$. Then $\|\boldsymbol{x} - \boldsymbol{y}\|_1 = 2 \sum\limits_{i \in [n]} \mathbf{1}[x_i < y_i] \cdot (y_i - x_i)$.

Proof. Let I denote the set of all $i \in [n]$ for which $x_i = \boldsymbol{e}_i^\top \boldsymbol{x} < \boldsymbol{e}_i^\top \boldsymbol{y} = y_i$. And denote $\boldsymbol{v} := \boldsymbol{x} - \boldsymbol{y}$. Then $\sum_i v_i = 0$. Furthermore $v_i < 0 \iff i \in I$. So $I = \{i \in [n] \mid v_i < 0\}$. Then the sum on the RHS of the given statement is simply $-2 \sum\limits_{i \in S} v_i$. Note that $\|\boldsymbol{x} - \boldsymbol{y}\|_1 = -\sum\limits_{i \in I} v_i + \sum\limits_{i \notin I} v_i = -\sum\limits_{i \in I} v_i + 0 - \sum\limits_{i \in I} v_i$ which is exactly the required quantity. ■

Lemma 2
Let $\boldsymbol{x}, \boldsymbol{y} \in \Delta_{n-1}$ be two probability distributions. For any $i, j \in [n]$, define

$$f(i,j) = \begin{cases} \min\{x_i, y_j\} & \text{if } i = j \\ \frac{2\max\{x_i - y_i, 0\}\max\{y_j - x_j, 0\}}{\|\boldsymbol{x} - \boldsymbol{y}\|_1} & \text{otherwise} \end{cases}.$$

Then $\sum\limits_{i \in [n]} f_t(i,j) = y_j \ \forall\, j \in [n]$, $\sum\limits_{j \in [n]} f_t(i,j) = x_i \ \forall\, i \in [n]$ and $\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_1 = \sum\limits_{i \in [n]} \sum\limits_{j \in [n]} \mathbf{1}_{i \neq j} f(i,j)$. *Essentially this implies that f is a joint distribution with marginals $\boldsymbol{x}$ and $\boldsymbol{y}$.*

Proof. Let $S := \{i \in [n] \mid x_i \geq \boldsymbol{y}_i\}$. This S is simply the complement of I in the proof of lemma 1.

So $\sum\limits_{j \in [n]} f(i,j) = \min\{x_i, y_i\} + 2\max\{x_i - y_i, 0\} \sum\limits_{j \in [n] \setminus \{i\}} \frac{\max\{y_j - x_j, 0\}}{\|\boldsymbol{x} - \boldsymbol{y}\|_1}$.

We will only show $\sum\limits_{j \in [n]} f(i,j) = x_i \ \forall\, i \in [n]$ because the proof for $\sum\limits_{i \in [n]} f(i,j) = y_j \ \forall\, j \in [n]$ is exactly the same.

If $i \in S$, we have

$$\begin{aligned}\sum_{j\in[n]} f(i,j) &= \min\{x_i, y_i\} + 2\max\{x_i - y_i, 0\} \sum_{j\in[n]\setminus\{i\}} \frac{\max\{y_j - x_j, 0\}}{\|\boldsymbol{x}-\boldsymbol{y}\|_1} \\ &= y_i + (x_i - y_i) \sum_{j\in[n]\setminus\{i\}} 2\frac{\max\{y_j - x_j, 0\}}{\|\boldsymbol{x}-\boldsymbol{y}\|_1} = y_i + (x_i - y_i)\cdot 1 = x_i\end{aligned}$$

where the second-last equality follows from lemma 1.

If $i \notin S$, we have

$$\begin{aligned}\sum_{j\in[n]} f(i,j) &= \min\{x_i, y_i\} + 2\max\{x_i - y_i, 0\} \sum_{j\in[n]\setminus\{i\}} \frac{\max\{y_j - x_j, 0\}}{\|\boldsymbol{x}-\boldsymbol{y}\|_1} \\ &= x_i + 2\cdot 0 \cdot \sum_{j\in[n]\setminus\{i\}} \frac{\max\{y_j - x_j, 0\}}{\|\boldsymbol{x}-\boldsymbol{y}\|_1} = x_i.\end{aligned}$$

Finally we show $\|\boldsymbol{x}-\boldsymbol{y}\|_1 = \sum\limits_{i\in[n]}\sum\limits_{j\in[n]} \mathbf{1}_{i\neq j} f(i,j)$. Indeed $\sum\limits_{i\in[n]}\sum\limits_{j\in[n]} \mathbf{1}_{i\neq j} f(i,j) = \sum\limits_{i\in[n]} (x_i - \min\{x_i, y_i\}) = \sum\limits_{i\in S}(x_i - \min\{x_i, y_i\}) + \sum\limits_{i\notin S}(x_i - \min\{x_i, y_i\}) = \sum_{i\in S}(x_i - y_i) \overset{\text{lemma 1}}{=} \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_1$ ■

Corollary 3
Let M be the (symmetric) transition matrix of the random walk on a graph G with n vertices. Define $d(t) := \sup\limits_{\boldsymbol{x},\boldsymbol{y}\in\Delta_{n-1}} \left\|M^t\boldsymbol{x} - M^t\boldsymbol{y}\right\|_1$ for any $t \in \mathbb{N}$. Then $d(s+t) \leq \frac{1}{2}d(s)d(t)$.

Proof. Fix $s, t \in \mathbb{N}$. Let $\boldsymbol{x}, \boldsymbol{y} \in \Delta_{n-1}$. Note that $M^s(\Delta_{n-1}) \subseteq \Delta_{n-1}$. Use the f in the above lemma by replacing $\boldsymbol{x}$ (in the lemma) with $M^s\boldsymbol{x}$ and $\boldsymbol{y}$ with $M^s\boldsymbol{y}$. Note that $\boldsymbol{e}_i^\top M^{t+s}\boldsymbol{x} = \sum\limits_{k=1}^{n} \boldsymbol{e}_i^\top M^t \boldsymbol{e}_k (M^s\boldsymbol{x})_k = \sum\limits_{k=1}^{n} \boldsymbol{e}_i^\top M^s \boldsymbol{e}_k \sum\limits_{j=1}^{n} f(k,j) = \sum\limits_{j\in[n]}\sum\limits_{k\in[n]} f(k,j)\boldsymbol{e}_i^\top M^s\boldsymbol{e}_k$. Similarly, $\boldsymbol{e}_i^\top M^{s+t}\boldsymbol{y} = \sum\limits_{j\in[n]}\sum\limits_{k\in[n]} f(k,j)\boldsymbol{e}_i^\top M^s\boldsymbol{e}_j$. Therefore

$$\begin{aligned}\left\|M^{s+t}(\boldsymbol{x}-\boldsymbol{y})\right\|_1 &= \sum_i \left|\sum_j\sum_k f(k,j)\boldsymbol{e}_i^\top M^s(\boldsymbol{e}_k - \boldsymbol{e}_j)\right| \\ &\leq \sum_{j,k}\sum_i f(k,j)\left|\boldsymbol{e}_i^\top M^s(\boldsymbol{e}_k - \boldsymbol{e}_j)\right| \\ &= \sum_{j,k} f(k,j)\left\|M^s\boldsymbol{e}_k - M^s\boldsymbol{e}_j\right\|_1 \\ &= \sum_{j,k} \mathbf{1}_{j\neq k} f(k,j)\left\|M^s\boldsymbol{e}_k - M^s\boldsymbol{e}_j\right\|_1 \\ &\leq d(s)\sum_{j,k} \mathbf{1}_{j\neq k} f(k,j) \overset{\text{lemma 2}}{=} \frac{1}{2}d(s)d(t).\end{aligned}$$

Since this was for arbitrary $\boldsymbol{x}, \boldsymbol{y} \in \Delta_{n-1}$, taking sup gives the desired result. ■

(b) Let $\boldsymbol{\pi}$ be the stationary distribution. Clearly $\sup\limits_{\boldsymbol{x}\in\Delta_{n-1}} \left\|M^t\boldsymbol{x} - \boldsymbol{\pi}\right\|_1 \leq \sup\limits_{\boldsymbol{x},\boldsymbol{y}\in\Delta_{n-1}} \left\|M^t\boldsymbol{x} - M^t\boldsymbol{y}\right\|_1$ since the constraint $\boldsymbol{y} = \boldsymbol{\pi}$ only makes the feasible set smaller, thus lowering the maximum value. Corollary 3 with induction gives $\sup\limits_{\boldsymbol{x}\in\Delta_{n-1}} \left\|M^{kt_0}\boldsymbol{x} - \boldsymbol{\pi}\right\|_1 \leq d(kt_0) \leq \frac{d(t_0)}{2^k}$ (d is as in corollary 3). But $d(t_0) = \sup\limits_{\boldsymbol{x},\boldsymbol{y}\in\Delta_{n-1}} \left\|M^{t_0}\boldsymbol{x} - M^{t_0}\boldsymbol{y}\right\|_1 \leq \sup\limits_{\boldsymbol{x},\boldsymbol{y}\in\Delta_{n-1}} \left(\left\|M^{t_0}\boldsymbol{x} - \boldsymbol{\pi}\right\|_1 + \left\|M^{t_0}\boldsymbol{y} - \boldsymbol{\pi}\right\|_1\right) < 1$. Therefore $\sup\limits_{\boldsymbol{x}\in\Delta_{n-1}} \left\|M^{kt_0}\boldsymbol{x} - \boldsymbol{\pi}\right\|_1 \leq 2^{-k}$. Note that $d-$regularity was not used.

(a) Assume that the graph is $d-$regular. M is the transition matrix of this random walk. Say its eigenvalues are $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n (> -1)$. $P = \frac{1}{2}(I + M)$ is the lazy version. We want to bound $\left\|P^t(\boldsymbol{x} - \boldsymbol{\pi})\right\|_1$ where $\boldsymbol{\pi}$ is the stationary distribution of M, hence the stationary distribution of P. The eigenvalues of M and P are related as $\lambda_i \leftrightarrow \mu_i := \frac{1+\lambda_i}{2}$. Since M is ergodic, $\mu_2 < 1$ and $\mu_n > 0$. Let $\boldsymbol{x} \in \Delta_{n-1}$ and denote $\boldsymbol{v} := \boldsymbol{x} - \boldsymbol{\pi}$. It's worth noting that $\|M^s\boldsymbol{v}\|_1 \leq \|M^s\|_1 \|\boldsymbol{v}\|_1 = \|\boldsymbol{v}\|_1 \leq \|\boldsymbol{x}\|_1 + \|\boldsymbol{\pi}\|_1 = 2$ where we used that $\|M^s\|_1$ is the maximum of the absolute value column sums which is 1. Take $t := 100t_0$.

$$
\begin{aligned}
\left\|P^t\boldsymbol{v}\right\|_1 &= \left\|\frac{1}{2^t}\sum_{i=0}^{t}\binom{t}{i}M^i\boldsymbol{v}\right\|_1 \\
&\leq \frac{1}{2^t}\sum_{i=0}^{t}\binom{t}{i}\left\|M^i\boldsymbol{v}\right\|_1 \\
&= \frac{1}{2^t}\sum_{i=0}^{t/4}\binom{t}{i}\left\|M^i\boldsymbol{v}\right\|_1 + \frac{1}{2^t}\sum_{25t_0<i\leq t}\binom{t}{i}\left\|M^i\boldsymbol{v}\right\|_1 \\
&\leq 2\sum_{i=0}^{t/4}\binom{t}{i}2^{-t} + \frac{1}{2^t}\sum_{25t_0<i\leq t}\binom{t}{i}\left\|M^i\boldsymbol{v}\right\|_1
\end{aligned}
$$

We use the lower-tail Chernoff bound[1] that if $X_1, \cdots, X_t \in \{0,1\}$ are outcomes of a fair coin toss with $X = \sum_{i=1}^t X_i$ then $\mu = \mathbb{E}\left[X\right] = \frac{t}{2}$ and $p := \sum\limits_{i=0}^{t/4}\binom{t}{i}2^{-i} = \mathbb{P}\left[X \leq \frac{t}{4} = (1 - \frac{1}{2})\mu\right] \leq \exp\left\{-\frac{\mu\cdot(1/2)^2}{2}\right\} = \exp\left\{-\frac{t}{16}\right\} = \exp\left\{-\frac{100t_0}{16}\right\} \overset{[\because t_0\geq 1]}{\leq} \exp\left\{-\frac{100t_0}{16}\right\} \leq e^{-6}$.

Moreover we have (independently) proven in (b) that $\left\|M^{kt_0}\boldsymbol{x} - \boldsymbol{\pi}\right\|_1 \leq 2^{-k}$ whence if $i \geq 25t_0$ then $\left\|M^i(\boldsymbol{x} - \boldsymbol{\pi})\right\|_1 \leq \left\|M^{i-25t_0}\right\|_1 \left\|M^{25t_0}(\boldsymbol{x} - \boldsymbol{\pi})\right\|_1 = 1 \cdot \left\|M^{25t_0}\boldsymbol{x} - \boldsymbol{\pi}\right\|_1 \leq 2^{-25}$

Then we have $\left\|P^t\boldsymbol{v}\right\|_1 \leq 2p + 2^{-t}\sum\limits_{25t_0<i\leq t}\binom{t}{i}\left\|M^i\boldsymbol{v}\right\|_1 \leq 2p + 2^{-t}\sum\limits_{25t_0<i\leq t}\binom{t}{i}2^{-25} = 2p + (1 - p)\cdot 2^{-25} = p(2 - 2^{-25}) + 2^{-25} \leq 2e^{-6} + 2^{-25} < \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$.

(In fact one can improve the above constant 100 to 18 by breaking the sums into $\sum\limits_{i=0}^{t/6}\cdot + \sum\limits_{i>t/6}\cdot$).

[1] $\mathbb{P}\left[X \leq (1-\delta)\mu\right] \leq \exp\left\{-\mu\delta^2/2\right\}\ \forall\, \delta \in (0,1)$

Problem 4

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 **1**. Show that every other eigenvalue has magnitude at most $1-\frac{1}{10n^2}$. What does this imply about the mixing time for a random walk on this graph from an arbitrary starting point?

Solution

Let $\mathcal{L} = 5I - A$ where A is the adjacency matrix of a connected $5-$regular graph (without self loops) $G = ([n], E)$. $\mathcal{L}, A$ have the same eigenvectors. Since A has eigenvalues in $[-5, 5]$ with the second highest eigenvalue < 5 (because connected), the eigenvalues of $\mathcal{L}$ are in $[0, 10]$ where the second smallest eigenvalue (call it λ) is > 0. We will first show that $\lambda \geq \frac{1}{n^2}$.

Let $\boldsymbol{v} \in \mathbb{R}^n$ be an eigenvector of $\mathcal{L}$ with eigenvalue $\lambda > 0$ and normalized so that $\sum_i v_i^2 = \|\boldsymbol{v}\|_2 = 1$ and the entry with the highest magnitude is non-negative (by multiplying $\boldsymbol{v}$ by -1 if necessary). Clearly $\boldsymbol{v}$ satisfies $\sum_i v_i = 0$ because it corresponds to the second lowest eigenvector and the eigenvector for 0 is parallel to $(1, \cdots, 1)$. Thus it must have some positive and some negative entries. Since the norm is 1, the highest entry of $\boldsymbol{v}$, say v_k, must be $\geq \frac{1}{\sqrt{n}}$. Its lowest entry must be negative, say $v_t < 0$. Consider a path $k = x_1 \to x_2 \to \cdots \to x_{r+1} = t$ in G. Then $(v_{x_1} - v_{x_2}) + \cdots + \left(v_{x_r} - v_{x_{r+1}}\right) = v_k - v_t > \frac{1}{\sqrt{n}}$.

Now[2] $\lambda = \boldsymbol{v}^\top \mathcal{L} \boldsymbol{v} = \sum\limits_{\{i,j\} \in E} (v_i - v_j)^2 \geq (v_{x_1} - v_{x_2})^2 + \cdots + \left(v_{x_r} - v_{x_{r+1}}\right)^2 \overset{\text{Cauchy-Schwarz}}{\geq}$ $\frac{1}{r}\left(\sum\limits_{i=1}^{r} (v_{x_i} - v_{x_{i+1}})\right)^2 > \frac{1}{nr} \geq \frac{1}{n^2}$ where the last inequality follows because $r + 1 \leq n$.

The second smallest eigenvalue λ of $\mathcal{L}$ corresponds to the second largest eigenvalue μ_2 of A with the relation that $\lambda = 5 - \mu_2$ so that $\mu_2 = 5 - \lambda$. The random walk described in the question has the transition matrix $\frac{1}{2}\left(I + \frac{1}{5}A\right)$. This matrix has all eigenvalues ≥ 0 and its second largest eigenvalue is $\tilde{\lambda} = \frac{1+\mu_2/5}{2} = \frac{5+\mu_2}{10} = \frac{10-\lambda}{10} = 1 - \frac{\lambda}{10} \leq 1 - \frac{1}{10n^2}$.

[2] $\boldsymbol{v}^\top \mathcal{L} \boldsymbol{v} = 5\sum_i v_i^2 - \sum_i \sum_j \mathbf{1}_{\{i,j\} \in E} \cdot v_i v_j = \frac{1}{2}\sum_i \sum_j \mathbf{1}_{\{i,j\} \in E} \cdot (v_i^2 + v_j^2 - 2v_i v_j) = \sum\limits_{\{i,j\} \in E} (v_i - v_j)^2$

Problem 5

Let $(a_1, b_1), \cdots, (a_n, b_n) \in \mathbb{F}^2$ where $\mathbb{F} = GF(q)$ 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, x such that $Q(b_i, a_i) = 0$ for each $1 \leq i \leq n$. 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 $k > (d + 1)\left(\lceil \sqrt{n} \rceil + 1\right)$ then $z - p(x)$ divides the bivariate polynomial Q described in part (a).

Solution

(a) Take degree $D = \left\lceil \sqrt{2n} \right\rceil$ (I couldn't do $\lceil \sqrt{n} \rceil + 1$). To ensure Q has degree $\leq D$, each monomial $x^i z^j$ should satisfy $j + i \leq D$. Define $Q(z, x) = \sum_{i=0}^{D} \sum_{j=0}^{D-i} c_{ij} x^i z^j$. Treat the c_{ij}'s as the variables and and we try to solve for the simultaneous system of equations $Q(b_l, a_l) = 0$ ($\forall\, 1 \leq l \leq n$) which are all linear in c_{ij}'s. The number of unknown c_{ij} we want to determine is precisely $\sum_{i=0}^{D} (D - i + 1) = (D+1)^2 - \frac{D(D+1)}{2} = \frac{(D+2)(D+1)}{2} > \frac{2n}{2} = n$. Therefore we have more variables than constraints (namely n). So the $\{c_{ij}\}$ admit a nontrivial solution, which can be easily found by Gaussian elimination by forming the required matrix obtained from the equations $\sum_{i=0}^{d} \sum_{j=0}^{d-j} c_{ij} x_l^i z_l^j = 0$ for $1 \leq l \leq n$.

(b) Suppose $z - g(x) | R(x, z)$ in $\mathbb{F}[z, x]$. Then $\exists f(x, z) \in \mathbb{F}[z, x]$ such that $R(z, x) = (z - g(x)) f(z, x)$. Setting $z = g(x)$ gives $R(g(x), x) = 0$.

Suppose $R(g(x), x) = 0$. Recall that $\mathbb{F}[x]$ is an Euclidean domain and so $\mathbb{F}[z, x] \cong (\mathbb{F}[x])[z]$ is a polynomial ring over a Euclidean domain. In simpler words it means that we can divide (with well defined "smaller" remainders) the same way as in $\mathbb{Z}[z]$. The notion of smallness is gives by the degree (in z) of the polynomials. So $\exists q, r \in \mathbb{F}[z, x]$ such that $R(z, x) = (z - g(x)) q(z, x) + r(z, x)$ where either $r = 0$ or $\deg_z(r) = 0$. This simply means that $r \in \mathbb{F}[x]$ and we can write $R(z, x) = (z - g(x)) q(z, x) + r(x)$. Plugging in $z = g(x)$ gives $0 = r(x)$. So $R = (z - g) f$, whence $z - g(x) \mid R(z, x)$.

(c) $\deg p(x) = d$ and define $f(x) := Q(p(x), x)$. Then f has k zeroes (among the first coordinates of the data points). Let's compute the degree of f. Each term $x^i z^j = x^i p(x)^j$ contributes a degree of $i + dj \leq i + d(D - i) = dD - (d - 1)i \leq dD = d\left\lceil \sqrt{2n} \right\rceil$. If $k > d\left\lceil \sqrt{2n} \right\rceil$, then f has more roots than its degree whence f is the zero polynomial (again, I could not do it for $k > (d + 1)\left(\lceil \sqrt{n} \rceil + 1\right)$). By (b), $z - p(x) \mid Q(z, x)$.