Mathematical Tools for CS - Lecture 3

February 20, 2024

Today's Plan

- Random Walk
- Concentration of Measure
- Error Correction Codes

1 Random Walk On the Line

1.1 introduction

This is a stochastic/random process where there is a particle moving, and each step it can move either left or right at probability $\frac{1}{2}$.

Definition. A random variable X is called Radamacher if $Pr(X=1) = Pr(X=-1) = \frac{1}{2}$.

Definition. Let $X_1, X_2, ...$ be independent Radamacher random variables, define $Z_n = \sum_{i=1}^{n} X_i$. The sequence $(Z_1, Z_2, ...)$ is called a random walk on the line (On the line because it's one dimensional).

1.2 Exploting random walks

Question 1: What is $Pr(Z_n = 0)$?
Question 2: How many times will the random walk return to 0?

Theorem 1. For $n \geq 0$, it holds that $Pr(Z_{2n+1} = 0) = 0$. Additionaly, $Pr(Z_{2n} = 0) \sim \frac{1}{\sqrt{\pi n}}$.

Reminder: $f(n) \sim G(n) \iff lim_{n \to \infty} \frac{f(n)}{g(n)} = 1$.

Theorem 2. $Pr(Z_n = 0 \text{ for infinitely many n's}) = 1$

Reminder: Stirling approx: $n! \sim \sqrt{2\pi n} \cdot (\frac{n}{e})^n$

Proof. (Theorem 1) $Z_n = 0$ iff exactly half of $X_1, ..., X_n$ are 1. Hence immidiatly $Pr(Z_{2n+1} = 0) = 0$. Additionaly,

$$\begin{aligned}
Pr(Z_{2n} = 0) &= Pr(\{x \in \{\pm 1\}^{2n} s.t. \sum_{i=1}^{2n} X_i = 0\}) \\
&= \frac{|\{x \in \{\pm 1\}^{2n} s.t. \sum_{i=1}^{2n} X_i = 0\}|}{2^{2n}} \\
&= \frac{\binom{2n}{n}}{2^{2n}} \\
&= \frac{(2n)!}{n!n!2^{2n}} \\
\text{Stirling} &\sim \frac{\sqrt{4\pi n} \cdot (\frac{2n}{e})^{2n}}{2\pi n \cdot (\frac{n}{e})^{2n} \cdot 2^{2n}} \\
&= \frac{(\frac{2n}{e})^{2n}}{\sqrt{\pi n}(\frac{2n}{e})^{2n}} \\
&= \frac{1}{\sqrt{\pi n}}
\end{aligned}$$

□

Proof. (Theorem 2) the number of times we return to zero is $\sum_{n=1}^{\infty} \mathbf{1}_{[\boldsymbol{Z_{2n}=0}]}$. Let's calculate the expectation of that:

$$\begin{aligned}
\mathbb{E}(\sum_{n=1}^{\infty} \mathbf{1}_{[\boldsymbol{Z_{2n}=0}]}) &= \sum_{n=1}^{\infty} \mathbb{E}(1_{[Z_{2n}=0]}) \\
&= \sum_{n=1}^{\infty} Pr(Z_{2n} = 0) \\
\text{up to a constant} &\geq \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \\
&= \infty
\end{aligned}$$

Lemma: Denote $p = Pr(Z_{2n} = 0 \text{ for some } n > 1)$. We claim that $p = 1$ if and only if $\mathbb{E}(\sum_{n=1}^{\infty} \mathbf{1}_{[\boldsymbol{Z_{2n}=0}]}) = \infty$. If we prove this the theorem is proved.

ir $p = 1$ then with probability 1 the random walk will be at 0 infinitely many times. Hence $\mathbb{E}(\sum_{n=1}^{\infty} \mathbf{1}_{[\boldsymbol{Z_{2n}=0}]}) = \infty$.
Otherwise, suppose $p < 1$. In this case, $X = \sum_{n=1}^{\infty} \mathbf{1}_{[\boldsymbol{Z_{2n}=0}]}$ is a geometric $r.v.$ with success probability $1 - p$ (success is **not** going back to 0, if we fail we "try again" until we never go back to 0 again.). As such $\mathbb{E}(X) = \frac{1}{1-p} < \infty$ □

1.3 Random walk of higher dimensions

In 2d, the particle is on a grid, and at each step it can move {right, left} × {up, down} (so it has to move at every axis in each step).

Definition. A random walk in d dimensions is a sequence of random vectors $(Z_1, Z_2, ...)$ s.t. $Z_n = \sum_{i=1}^{n} X_i$ where $(X_1, X_2, ...)$are independent and uniform random vectors in $\{\pm 1\}^d$.

Theorem.
1) $Pr(Z_{2n} = 0) \sim (\frac{1}{\sqrt{\pi n}})^d = (\frac{1}{\pi n})^{\frac{d}{2}}$
2) $Pr(Z_{2n} = 0$ for infinitely many n's$) = 1 \iff d \leq 2$ (WHAT OMG)

Proof. (Sketch, not formal - for part 2)

$$\begin{aligned} Pr(Z_{2n} = 0 \text{ for infinitely many n's}) &\iff \mathbb{E}(\sum_{n=1}^{\infty} 1_{[Z_{2n}=0]}) = \infty \\ &\iff \sum_{n=1}^{\infty} n^{-\frac{d}{2}} = \infty \\ &\iff d \leq 2 \end{aligned}$$

□

2 Concentration of measure

2.1 Philosophy of Concentration of measure

Concentration of measure results show that under various conditions, the probability that the sum of independent r.v.'s deviates from it's expectation is exponentially small.
More generally, instead of sums, we can consider some function of a sequence of random variables. in addition we can sometimes assume they are not independent.

2.2 Random walk concentration

Let $Z_n = \sum_{i=1}^{n}$, where $x_i \in \{\pm 1\}^n$ a random walk on the line. We want to bound $Pr(Z_n \geq a) \leq ?$.
Notice that $\mathbb{E}(Z_n) = 0$ and $Var(Z_n) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} 1 = n.$
This implies via chebychev bound:

$$Pr(Z_n \geq k\sqrt{n}) \leq Pr(|Z_n| \geq k\sqrt{n}) \leq \frac{1}{k^2}$$

but it turns out we can prove a stronger bound:

Claim. $Pr(Z_n \geq a) \leq e^{-\frac{a^2}{2n}}$
in particular, if we take $a = k\sqrt{n}$, then: $r(Z_n \geq a) \leq e^{-\frac{k^2 n}{2n}} = e^{-\frac{k^2}{2}}$. Which is much better than the previous bound.
important fact for the proof:

Fact. $\frac{e^t + e^{-t}}{2} \leq e^{\frac{t^2}{2}}$

Proof. (of the claim) (Many concentration of measure bounds are proved in a similar way). Fix some $t > 0$. $x \to e^{tx}$ is monotone, so therefore:

$$\begin{aligned}
Pr(Z_n \geq a) &= Pr(e^{t\cdot Z_n} \geq e^{t\cdot a})\\
\text{markov inequality } &\leq e^{-ta}\mathbb{E}(e^{tZ_n})\\
\text{def of } Z_n &= e^{-ta}\mathbb{E}(e^{t\sum_{i=1}^{n} X_i})\\
&= e^{-ta}\cdot\mathbb{E}(\prod_{i=1}^{n} e^{tX_i})\\
\text{independence of X's } &= e^{-ta}\prod_{i=1}^{n}\mathbb{E}(e^{tX_i})\\
&= e^{-ta}\prod_{i=1}^{n}\frac{e^{t}+e^{-t}}{2}\\
\text{fact } &\leq e^{-ta}\prod_{i=1}^{n} e^{\frac{t^2}{2}}\\
&= e^{-ta+\frac{1}{2}nt^2}
\end{aligned}$$

now set $t = \frac{a}{n}$ (found by taking derivative and equating 0):

$$\begin{aligned}
e^{-ta+\frac{1}{2}nt^2} &= e^{-\frac{a^2}{n}+\frac{1}{2}n\frac{a^2}{n^2}}\\
&= e^{-\frac{a^2}{n}+\frac{a^2}{2n}}\\
&= e^{-\frac{a^2}{2n}}
\end{aligned}$$

□

2.3 Hoffding inequality.

What we just saw is a special case of (Hoffding's bound).

Theorem. (Hoffding's bound) Let $X_1, ..., X_n$ be independent r.v.'s such that with probability 1, it holds that $X_i \in [a_i, b_i]$. so for any $a \geq 0$:

$$Pr(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n}\mathbb{E}(X_i) + a) \leq e^{-\frac{2a^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}$$

Proof. in the official lecture notes. □

2.4 Error Correction Codes - Hoffding application example

We want to tackle the following scenario. We want to send a sequence and bits, while sending the sequence, some of the bits are modified. We want to be able to send to original sequence despite the modification.
Formally, $X \in \{\pm 1\}^n$ is the original string we want to send. We encode $X \rightarrow C(X) \in \{\pm 1\}^m$. We send $C(X)$. During that, some adversery modifies $\varepsilon \cdot m$ bits $C(X) \rightarrow Z$. We then want to reconstruct $Z \rightarrow D(Z) = X$.
We want to show that we can construct the functions C, D such that we can reconstruct the original message exactly for some ε.

Definition. the Hamming distance between strings $x, y \in \{\pm 1\}^n$ is

$$d(x,y) = |\{i \in [n] : x_i \neq y_i\}|$$

Lemma. The hamming distance d, is a metric - for any $x, y, z \in \{\pm 1\}^n$:

1. (Triangle inequality) $d(x,y) \leq d(x,z) + d(z,y)$
2. $d(x,y) = d(y,x)$
3. $d(x,y) \geq 0$
4. $d(x,y) \geq$ with equlity if and only if $x = y$.

Proof. for triangle inequality.

$$\begin{aligned} d(x,y) &= \sum_{i=1}^{n} \mathbf{1}_{[x_i \neq y_i]} \\ &\leq \sum_{i=1}^{n} \mathbf{1}_{[x_i \neq z_i]} + \mathbf{1}_{[x_i \neq y_i]} \\ &= d(x,z) + d(z,y) \end{aligned}$$

□

Definition. A Code $C : \{\pm 1\}^n \rightarrow \{\pm 1\}^m$ has distance $\varepsilon > 0$ if for any pair $x, y \in \{\pm 1\}^n$ it holds that $d(C(x), C(y)) > 2 \cdot \varepsilon \cdot m$.

Lemma. Suppose $C : \{\pm 1\}^n \rightarrow \{\pm 1\}^m$ has distance $\varepsilon > 0$. Let $x \in \{\pm 1\}^n$ and let $z \in \{\pm 1\}^m$ s.t. $d(C(x), z) \leq \varepsilon \cdot m$. Then

$$x = \arg \min_{y \in \{\pm 1\}^n} d(z, C(y))$$

Proof. Let $y \in \{\pm 1\}^n \setminus \{x\}$. We need to show that $d(z, C(y)) \geq d(z, C(x))$. by triangle inequality:

$$\begin{aligned} d(z,c(y)) &\geq d(c(x),c(y)) - d(z,c(x)) \\ \text{C has distanc } \varepsilon \quad &\geq 2\varepsilon m - d(z,c(x)) \\ \text{assumption} \quad &\geq 2d(C(x),z) - d(z,c(x)) \\ &= d(z,c(x)) \end{aligned}$$

□

Theorem. for $0 \leq \varepsilon \leq \frac{1}{4}$ and $m > \frac{n}{(\frac{1}{2}-2\varepsilon)^2}$ there exists a code $C : \{\pm 1\}^n \to \{\pm 1\}^m$ of distance ε.

Proof. It is enough to show that there are $x_1, ..., x_{2^n} \in \{\pm 1\}^m$ s.t. $d(x_i, x_j) > 2\varepsilon m$. Let $x, y \in \{\pm 1\}^m$ be random vectors. Define $x_i = \begin{cases} 1 & x_i \neq y_i \\ 0 & \text{otherwise} \end{cases}$. so $d(x,y) = \sum_{i=1}^n x_i$. As such $\mathbb{E}(d(x,y)) = \frac{m}{2}$.

$$\begin{aligned} Pr(d(x,y) \leq \varepsilon m) &= Pr(\sum x_i \leq 2\varepsilon m) \\ &= Pr(\sum x_i - \frac{m}{2} \leq (2\varepsilon - \frac{1}{2})m) \\ \text{hoffding (with negative a which is } (2\varepsilon - \tfrac{1}{2})\text{)} \quad &\leq e^{\frac{-(\frac{1}{2}-2\varepsilon)^2 2m^2}{m}} \\ &= e^{-(\frac{1}{2}-2\varepsilon)^2 m} \end{aligned}$$

Let $x_1, ..., x_{2^n} \in \{\pm 1\}^m$ be independent and uniformally distributed random vectors. so:

$$\begin{aligned} Pr(\exists i,j \; s.t. \; d(x_i,x_j) \leq 2\varepsilon m) &= Pr(\bigcup_{(i,j) \in \binom{[2^n]}{2}} d(x_i,x_j) \leq 2\varepsilon m) \\ &\leq \sum_{(i,j) \in \binom{[2^n]}{2}} Pr(d(x_i,x_j) \leq 2\varepsilon m) \\ \text{inequality proved ind last block} \quad &\leq \sum_{(i,j) \in \binom{[2^n]}{2}} e^{-2(\frac{1}{2}-2\varepsilon)^2 m} \\ &\leq 2^{2n} e^{-2(\frac{1}{2}-2\varepsilon)^2 m} \\ \text{assumption} \quad &< 2^{2n} e^{-2(\frac{1}{2}-2\varepsilon)^2 \frac{n}{(\frac{1}{2}-2\varepsilon)^2}} \\ &= 2^{2n} \cdot e^{-2n} \\ &< 1 \end{aligned}$$

□