Lecture 7: Martingales and Concentration

John Sylvester   Nicolás Rivera   Luca Zanetti   Thomas Sauerwald

Lent 2019

UNIVERSITY OF CAMBRIDGE

Outline

Martingales: Definition

A sequence of random variables $Z_0, Z_1, \ldots,$ is a martingale with respect to the sequence $X_0, X_1, \ldots,$ if, for all $n \geq 0$, the following holds:

1. Z_n is a function of $X_0, X_1, \ldots, X_n$
2. $\mathbf{E}[\,|Z_n|\,] < \infty$, and
3. $\mathbf{E}[\,Z_{n+1} | X_0, \ldots, X_n\,] = Z_n$.

We will see later why martingales are useful. For now, just think that being a martingale is good.

Simple Random Walk

- Consider a sequence $X_1, X_2, \ldots,$ of independent random variables with $\mathbf{P}[\, X_i = 1\,] = \mathbf{P}[\, X_i = -1\,] = 1/2$
- For $n \geq 0$ denote by $S_n = X_0 + X_1 + \ldots + X_n$, where $X_0 = k \in \mathbb{Z}$,
- Then $S_0, S_1, \ldots,$ is a martingale with respect to $X_0, X_1, X_2, \ldots,$
- We check the definition
  1. $S_n = \sum_{i=0}^{n} X_i$, i.e. S_n is a function of $X_0, X_1, \ldots, X_n$. Note that $S_0 = X_0 = 0$.
  2. $\mathbf{E}[\,|S_n|\,] \leq \mathbf{E}\left[\, \sum_{i=0}^{n} |X_i|\,\right] \leq n < \infty$
  3. $$\begin{aligned} \mathbf{E}[\, S_{n+1} | X_0, \ldots, X_n\,] &= \mathbf{E}[\, S_n + X_{n+1} | X_0, \ldots, X_n\,] \\ &= S_n + \mathbf{E}[\, X_{n+1} | X_0, \ldots, X_n\,] = S_n \end{aligned}$$

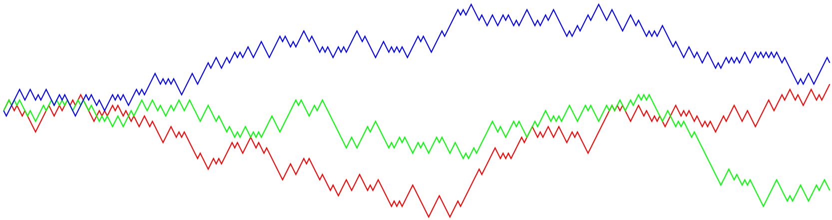
- The usual interpretation is a Gambler who is betting 1 pound each turn, and S_n is the current profit, and X_0 the initial capital.

Simple Random Walk

A bit more interesting is the fact that $W_n = S_n^2 - n$ is also a martingale with respect to $X_0, X_1, \ldots,$

1. $W_n = S_n^2 - n$ is a function of $X_0, X_1, \ldots, X_n$
2. $|S_n^2 - n| \leq (n+k)^2 + n < \infty$
3. $S_{n+1} = S_n + X_{n+1}$ then
4. To save space, we will write $\mathbf{X^m}$ instead of $(X_0, X_1, \ldots, X_m)$, then

$$\begin{aligned}
\mathbf{E}\left[\, S_{n+1}^2 - (n+1) | \mathbf{X_n} \,\right] &= \mathbf{E}\left[\, S_n^2 + 2S_n X_{n+1} + X_{n+1}^2 - (n+1) | \mathbf{X_n} \,\right] \\
&= S_n^2 + 2S_n \mathbf{E}[\, X_{n+1} | \mathbf{X_n} \,] + 1 - (n+1) \\
&= S_n^2 - n
\end{aligned}$$



Biased random Walks

- Consider a sequence $X_1, X_2, \ldots,$ of independent random variables with $\mathbf{P}[\, X_i = 1\,] = p$ and $\mathbf{P}[\, X_i = -1\,] = q$.
- For $n \geq 0$ denote by $S_n = X_0 + X_1 + \ldots + X_n$, where $X_0 = k \in \mathbb{Z}$,
- S_n is not a martingale if $p \neq q$, check it
- Then $Z_n = (q/p)^{S_n}$ is a martingale with respect to $X_0, X_1, X_2, \ldots,$

1. Z_n is a function of S_n which is a function of $X_0, \ldots, X_n$
2. $\mathbf{E}[\,|Z_n|\,] \leq \max\{q/p, 1\}^n < \infty$
3.

$$\begin{aligned}
\mathbf{E}[\, Z_{n+1} | X_0, \ldots, X_n\,] &= (q/p)^{S_n+1} p + (q/p)^{S_n-1} q \\
&= (q/p)^{S_n} \left[(q/p)p + (p/q)q\right] \\
&= Z_n
\end{aligned}$$

Martingales: Definition

A sequence of random variables $Z_0, Z_1, \ldots,$ is a martingale with respect to the sequence $X_0, X_1, \ldots,$ if, for all $n \geq 0$, the following holds:

1. Z_n is a function of $X_0, X_1, \ldots, X_n$
2. $\mathbf{E}[\,|Z_n|\,] < \infty$, and
3. $\mathbf{E}[\,Z_{n+1}|X_0, \ldots, X_n\,] = Z_n$.

Some remarks

- A sequence of random variables $Z_0, Z_1, \ldots,$ is called a martingale when it is a martingale with respect to itself, i.e. $X_i = Z_i$.
- The first index doesn't need to be 0, sometimes it is better to start at 1
- Sometimes we don't define X_0, even though the martingale starts at Z_0
- The index set can be infinity (all natural numbers) or finite
- To save space, we will write $\mathbf{X^m}$ instead of $(X_0, X_1, \ldots, X_m)$

Balls into Bins

Consider m balls assigned uniformly at random into n bins.

Enumerate the balls from 1 to m. Ball i is assigned to a random bin X_i.

Let Z be the number of empty bins (after assigning the balls)

Can we prove concentration here?

Denote $Y_i = 1$ if bin i is empty, 0 otherwise. Then $Z = \sum_{i=1}^{n} Y_i$.

Denote $Z_t = \mathbf{E}\left[Z | \mathbf{X^t} \right]$ and $Z_0 = \mathbf{E}[Z]$. Then Z_t is a martingale wrt $X_1, X_2, \ldots, X_t$

1. $Z_t = f(X_1, \ldots, X_t)$
2. $\mathbf{E}[|Z_t|] \leq n$
3. For $1 \leq t \leq m-1$,

$$\mathbf{E}\left[Z_{t+1} | \mathbf{X^t} \right] = \mathbf{E}\left[\mathbf{E}\left[Z | \mathbf{X^{t+1}} \right] | \mathbf{X^t} \right] = \mathbf{E}\left[Z | \mathbf{X^t} \right]$$

and

$$\mathbf{E}[Z_1] = \mathbf{E}[\mathbf{E}[Z | X_1]] = \mathbf{E}[Z] = Z_0$$

Martingales: Doob Martingale

A Doob martingale refers to a generic construction that is always a martingale. The construction is as follows.

- Let $X_0, \ldots, X_n$ be a sequence of random variables.
- Let Y be another random variable with $\mathbf{E}[\,|Y|\,] < \infty$ (usually Y is a function of $X_0, \ldots, X_n$).
- Define $Z_i = \mathbf{E}\left[\, Y|\mathbf{X^i} \,\right]$.
- $Z_0, \ldots, Z_n$ is a martingale w.r.t $X_0, \ldots, X_n$
  1. Clearly Z_i is a function of $X_0, \ldots, X_i$.
  2. $\mathbf{E}[\,|Z_i|\,] = \mathbf{E}\left[\,|\mathbf{E}\left[\, Y|\mathbf{X^i} \,\right]|\,\right] \leq \mathbf{E}\left[\,\mathbf{E}\left[\,|Y||\mathbf{X^i}\,\right]\,\right] = \mathbf{E}[\,|Y|\,] < \infty$.
  3. $$\mathbf{E}\left[\, Z_{i+1} \,\middle|\, \mathbf{X^i} \,\right] = \mathbf{E}\left[\,\mathbf{E}\left[\, Y \,\middle|\, \mathbf{X^{i+1}} \,\right] \,\middle|\, \mathbf{X^i} \,\right] = \mathbf{E}\left[\, Y \,\middle|\, \mathbf{X^i} \,\right] = Z_i$$
- In most applications, X_0 is undefined/ignored and Z_0 corresponds to $\mathbf{E}[\, Y \,]$ while $Z_n = Y$.

Example of Doob Martingale: Edge Exposure Martingale

- Consider a random graph G sampled from $G_{n,p}$ where the vertex set is $\{1, \ldots, n\}$ and the edge between i, j appears with probability p independent of everything.
- Enumerate all the possible edges from 1 to $m = \binom{n}{2}$. Denote by $X_j = 1$ if edge j appears in G, 0 otherwise.
- Let $F(G)$ be a numerical quantity of G, e.g. number of connected components, number of edges, indicator if G is hamiltonian or not...
- define $Z_i = \mathbf{E}\left[\, F(G) | \mathbf{X^i} \,\right]$ and define $Z_0 = \mathbf{E}[\, F(G) \,]$.
- Z_i is a Doob Martingale wrt to $X_1, \ldots, X_m$, and it is called the edge-exposure martingale
- The interpretation is that instead of computing $F(G)$ by observing G directly, we reveal the edges of G one by one, and estimate $F(G)$ with the given information. With no information the 'best' guess for $F(G)$ is its expectation.

Example of Doob Martingale: Vertex Exposure Martingale

- Similarly, instead of reveal edges one at a time, we can reveal vertices (with the corresponding edges), one at a time.
- Fix the vertices from 1 to n, and let G_i be the subgraph of G induced by the first i vertices.
- let $Z_0 = \mathbf{E}[\, F(G)\,]$ and $Z_i = \mathbf{E}[\, F(G) | G_1, \ldots, G_i\,]$
- this Doob martingale is called the vertex-exposure martingale

Outline

Martingales

Martingale Concentration Inequalities

Examples

Martingales: Azuma-Hoeffding Inequality

Azuma-Hoeffding Inequality

Let $Z_0, \ldots Z_n$ be a martingale wrt $X_0, X_1, \ldots,$ such that

$$a_k \leq Z_k - Z_{k-1} \leq b_k$$

Then, for all $t \geq 0$ and any $k > 0$ it holds

$$\mathbf{P}[\, Z_k - Z_0 \geq t \,] \,, \mathbf{P}[\, Z_k - Z_0 \leq -t \,] \leq \exp\left(- \frac{2t^2}{\sum_{i=1}^{k} (b_i - a_i)^2} \right)$$

Exercise. Check that if Z_0 is deterministic then $\mathbf{E}[\, Z_k \,] = Z_0$.

The proof follows the standard recipe.

1. Let $\lambda > 0$, then

$$\mathbf{P}[\, Z_k - Z_0 \geq t \,] \leq e^{-\lambda t} \mathbf{E}\left[e^{\lambda (Z_k - Z_0)} \right]$$

2. Compute an upper bound for $\mathbf{E}\left[e^{\lambda (Z_k - Z_0)} \right]$
3. Optimise the value of $\lambda > 0$.

We only show that $\mathbf{E}\left[e^{\lambda (Z_k - Z_0)} \right] \leq e^{\sum_{i=1}^{k} \lambda^2 (b_i - a_i)^2 / 8}$, the rest is an **Exercise**.

Martingales: Azuma-Hoeffding Inequality

- Define $Y_i = Z_i - Z_{i-1}$ for $i \geq 1$.
- By martingale properties $\mathbf{E}[\, Y_i | X_0, X_1, \ldots, X_{i-1} \,] = 0$.
- We pretty much follow the same argument used in the Proof of the Hoeffding's Extension Lemma
- By convexity

$$e^{\lambda Y_i} \leq \frac{b_i - Y_i}{b_i - a_i} e^{\lambda a_i} + \frac{Y_i - a_i}{b_i - a_i} e^{\lambda b_i}$$

- Then

$$\mathbf{E}\left[e^{\lambda Y_i} \middle| X_0, \ldots, X_{i-1} \right] \leq \frac{b_i e^{\lambda a_i}}{b_i - a_i} - \frac{a_i e^{\lambda b_i}}{b_i - a_i} \leq \exp\left[\frac{(b_i - a_i)^2 \lambda^2}{8} \right]$$

Exactly as in the Hoeffding's Extension Lemma

- so we have $\mathbf{E}\left[e^{\lambda Y_i} | X_0, \ldots X_{i-1} \right] \leq \exp\left[\frac{(b_i - a_i)^2 \lambda^2}{8}\right]$
- We bound $\mathbf{E}\left[e^{\lambda(X_k - X_0)} \right]$.

$$
\begin{aligned}
\mathbf{E}\left[e^{\lambda(Z_k - Z_0)} \right] &= \mathbf{E}\left[e^{\sum_{i=1}^{k} \lambda Y_i} \right] = \mathbf{E}\left[\prod_{i=1}^{k} e^{\lambda Y_i} \right] \\
&= \mathbf{E}\left[\mathbf{E}\left[\prod_{i=1}^{k} e^{\lambda Y_i} \,\middle|\, X_0, \ldots, X_{k-1} \right] \right] \\
&= \mathbf{E}\left[\prod_{i=1}^{k-1} e^{\lambda Y_i} \mathbf{E}\left[e^{\lambda Y_k} \,\middle|\, X_0, \ldots, X_{k-1} \right] \right] \\
&\leq \mathbf{E}\left[\prod_{i=1}^{k-1} e^{\lambda Y_i} \right] \exp\left[\frac{(b_i - a_i)^2 \lambda^2}{8} \right] \\
&\leq \exp\left[\sum_{i=1}^{k} \frac{(b_i - a_i)^2 \lambda^2}{8} \right]
\end{aligned}
$$

Exercise: Check the previous steps.

- pick $\lambda = \frac{4t}{\sum_{i=1}^{n} (b_i - a_i)^2}$

Method of Bounded Differences

Suppose, we have random variables $X_1, \ldots, X_n$. We want to study the random variable

$$f(X_1, \ldots, X_n)$$

Some examples:

1. $X = X_1 + \ldots + X_n$
2. In balls into bins, X_i indicate where ball i is allocated, and $f(X_1, \ldots, X_m)$ is the number of empty bins
3. X_i indicates if the i-th edge belongs to a random graph G, and $f(X_1, \ldots, X_m)$ represent the number of connected components of G

We can simply prove concentration of X around it means by the so-called Method of Bounded Differences

Method of Bounded Differences

A function f is called Liptchitz of parameter $\mathbf{c} = (c_1, \ldots, c_n)$ if for all i

$$|f(x_1, x_2, \ldots, x_{i-1}, \mathbf{x_i}, x_{i+1}, \ldots, x_n) - f(x_1, x_2, \ldots, x_{i-1}, \mathbf{y_i}, x_{i+1}, \ldots, x_n)| \leq c_i$$

where x_i and y_i are in the domain of the i-th coordinate

McDiarmid's inequality

Let $X_1, \ldots, X_n$ be independent random variables. Let f be Liptchitz of parameter $\mathbf{c} = (c_1, \ldots, c_n)$. Let $X = f(X_1, \ldots, X_n)$. Then

$$\mathbf{P}[\, X - \mathbf{E}[\, X\,] \geq t\,], \mathbf{P}[\, X - \mathbf{E}[\, X\,] \leq -t\,] \leq \exp\left(-\frac{2t^2}{\sum c_i^2}\right)$$

McDiarmid's inequality

Let $X_1, \ldots, X_n$ be independent random variables. Let f be Liptchitz of parameter $\mathbf{c} = (c_1, \ldots, c_n)$. Let $X = f(X_1, \ldots, X_n)$. Then

$$\mathbf{P}[\, X - \mathbf{E}[\, X\,] \geq t\,] , \mathbf{P}[\, X - \mathbf{E}[\, X\,] \leq -t\,] \leq \exp\left(-\frac{2t^2}{\sum c_i^2} \right)$$

In our proof we are going to assume the X_i are discrete random variables. Nevertheless, the result can be proven for continuous random variables. f

Proof: Use that $Z_i = \mathbf{E}[\, f(X_1, \ldots, X_n) | X_1, \ldots, X_i\,]$ with

$\mathbf{E}[\, Z_0\,] = \mathbf{E}[\, f(X_1, \ldots, X_n)\,]$ is a (Doob) Martingale.

We just need bounds for $Z_i - Z_{i-1}$ for all $i \geq 1$.

- Recall that $Z_i - Z_{i-1} = \mathbf{E}[\, f | X_1, \ldots, X_i \,] - \mathbf{E}[\, f | X_1, \ldots, X_{i-1} \,]$
- For $i < j$ write $\mathbf{X_i^j} = (X_i, \ldots, X_j)$. By definition of conditional expectation $Z_i - Z_{i-1}$ equals (exercise)

$$\sum_{\mathbf{x}} f(\mathbf{X_1^{i-1}}, X_i, \mathbf{x}) \mathbf{P}\left[\, \mathbf{X_{i+1}^n} = \mathbf{x} \,\right] - \sum_{(y,\mathbf{x})} f(\mathbf{X_1^{i-1}}, y, \mathbf{x}) \mathbf{P}\left[\, (X_i, \mathbf{X_{i+1}^n}) = (y, \mathbf{x}) \,\right]$$

- Here we use that X_i are independent: $\mathbf{P}[\, (X_i, \mathbf{X_{i+1}^n}) = (y, \mathbf{x}) \,] = \mathbf{P}[\, X_i = y \,] \, \mathbf{P}[\, \mathbf{X_{i+1}^n} = \mathbf{x} \,]$
- Therefore $Z_i - Z_{i-1}$ equals

$$\sum_{\mathbf{x}} \sum_{y} \left[f(\mathbf{X_1^{i-1}}, X_i, \mathbf{x}) - f(\mathbf{X_1^{i-1}}, y, \mathbf{x}) \right] \mathbf{P}[\, X_i = y \,] \, \mathbf{P}\left[\, \mathbf{X_{i+1}^n} = \mathbf{x} \,\right]$$

- Denote $a_i = \inf_{y'} \left[f(\mathbf{X_1^{i-1}}, y', \mathbf{x}) - f(\mathbf{X_1^{i-1}}, y, \mathbf{x}) \right]$ and $b_i = \sup_{z'} \left[f(\mathbf{X_1^{i-1}}, z', \mathbf{x}) - f(\mathbf{X_1^{i-1}}, y, \mathbf{x}) \right]$.
- Note that

$$\left| \left[f(\mathbf{X_1^{i-1}}, z', \mathbf{x}) - f(\mathbf{X_1^{i-1}}, y, \mathbf{x}) \right] - \left[f(\mathbf{X_1^{i-1}}, y', \mathbf{x}) - f(\mathbf{X_1^{i-1}}, y, \mathbf{x}) \right] \right| \leq c_i$$

- Hence $b_i - a_i \leq c_i$

McDiarmid's inequality

Let $X_1, \ldots, X_n$ be independent random variables. Let f be Liptchitz of parameter $\mathbf{c} = (c_1, \ldots, c_n)$. Let $X = f(X_1, \ldots, X_n)$. Then

$$\mathbf{P}[\, X - \mathbf{E}[\, X \,] \geq t \,] , \mathbf{P}[\, X - \mathbf{E}[\, X \,] \leq -t \,] \leq \exp\left(-\frac{2t^2}{\sum c_i^2}\right)$$

Outline

Martingales

Martingale Concentration Inequalities

Examples

Examples: Balls into Bins

Consider m balls assigned uniformly at random into n bins.

Enumerate the balls from 1 to m. Ball i is assigned to a random bin X_i.

Let Z be the number of empty bins (after assigning the balls)

$Z = f(X_1, \ldots, X_n$ and f is Liptchitz with $\mathbf{c} = (1, \ldots, 1)$ (because if we move one ball to another bin, the number of empty bins changes at most in 1)

By the McDiarmid's inequality

$$\mathbf{P}[\,|F - \mathbf{E}[\,F\,]\,| > t\,] \leq 2e^{-2t^2/m}$$

Example: Bin Packing

Consider the Bin Packing problem

1. We are given n items of sizes in the unit interval $[0, 1]$
2. We want to pack those items into the fewest number of unit-capacity bins as possible
3. Suppose that the item sizes X_i are independent random variables in the interval $[0, 1]$
4. let $B = B(X_1, \ldots, X_n)$ the optimal number of bins that suffice to pack the items
5. The Lipschitz conditions holds with $\boldsymbol{c} = (1, \ldots, 1)$, **Why?**
6. Therefore

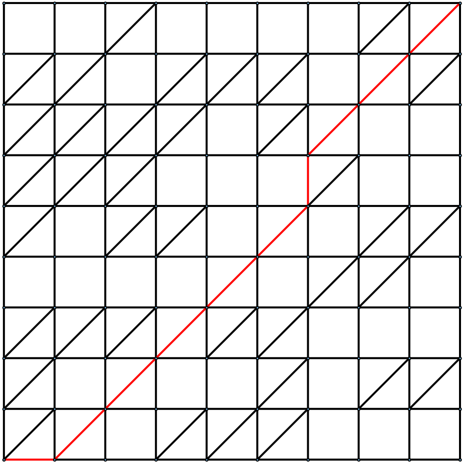
$$\mathbf{P}[\, B - \mathbf{E}[\, B\,] \geq t\,] \,, \mathbf{P}[\, B - \mathbf{E}[\, B\,] \leq -t\,] \leq e^{-2t^2/n}.$$

A random distance problem

Consider an n by n square grid $\{0, 1, \ldots, n\}^2$, where each point is connected to each of its (at most) four neighbours (N, S, E, W). Within each inner square of the grid, we draw a diagonal from *SW* to *NE* with probability p.

We say that $(0, 0)$ is on the bottom left corner and (n, n) in the top right corner.

Can we prove concentration of the shortest path from $(0, 0)$ to (n, n)?



A random distance problem

Can we prove concentration of the shortest path from $(0,0)$ to (n,n)?
Yes! Let Z be the total length of the shortest path.Two options

1. Define $X_{ij} = 1$ if there is a diagonal in square ij, otherwise $X_{ij} = 0$. Then $Z = f(X_{11}, \ldots, X_{nn})$ satisfies the Lipschitz conditions with $\boldsymbol{c} = (2 - \sqrt{2})(1, \ldots, 1)$, **Why?** .
   Then
   $$\mathbf{P}[\,|Z - \mathbf{E}[\,Z\,]\,| \geq t\,] \leq 2\exp\left[\frac{-t^2}{(2-\sqrt{2})^2 n^2}\right]$$
2. Enumerate the columns of squares from 1 to n. Let $Y_i = (X_{1i}, \ldots, X_{ni})$. Then $Z = g(Y_1, \ldots, Y_n)$. g satisfies the Lipschitz conditions with $\boldsymbol{c} = (2 - \sqrt{2})(1, \ldots, 1)$. **Why?**
   Then
   $$\mathbf{P}[\,|Z - \mathbf{E}[\,Z\,]\,| \geq t\,] \leq 2\exp\left[\frac{-t^2}{(2-\sqrt{2})^2 n}\right]$$

   Note the second bound is way more useful than the first one.

Example: Clique Number in Random Graphs

1. Consider a random graph $G = G_{n,p}$ on n vertices where each possible edge appears with probability p independent of each other.
2. Denote by K the clique number of G defined as the size of the largest complete subgraph of G.
3. K is a function of the number of edges of the graph, i.e. $K = K(X_1, \ldots, X_{\binom{n}{2}})$ where X_i represent if the i-th possible edge is in the graph or not.
4. Lipschitz conditions holds with $\boldsymbol{c} = (1, \ldots, 1)$. **Why?**
5. Therefore, for $t > 0$

$$\mathbf{P}[\, K - \mathbf{E}[\, K\,] \geq t\,]\,, \mathbf{P}[\, K - \mathbf{E}[\, K\,] \leq t\,] \leq e^{-2t^2/\binom{n}{2}}.$$

Example: Clique Number in Random Graphs

1. Consider a random graph $G = G_{n,p}$ on n vertices where each possible edge appears with probability p independent of each other.
2. Denote by K the clique number of G defined as the size of the largest complete subgraph of G.
3. Enumerate the vertices from 1 to n
4. Let $X_{i,j} = 1$ if there is a edge between vertices i and j, otherwise $X_{i,j} = 0$
5. Let $Y_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,i-1})$
6. K is a function of the Y_i.
7. Lipschitz conditions holds with $\boldsymbol{c} = (1, \ldots, 1)$. **Why?**
8. Therefore, for $t > 0$

$$\mathbf{P}[\, K - \mathbf{E}[\, K\,] > t\,]\,, \mathbf{P}[\, K - \mathbf{E}[\, K\,] < t\,] \leq e^{-2t^2/n}.$$

Observe this bound is better than the previous one

MaxCut on Random Graphs

We analyse the Max-Cut problems on Random Graphs, i.e. instead of assuming worst case input, we assume a random input.

1. Consider a random graph $G_{n,1/2}$ on vertices $[n] = \{1, \ldots, n\}$ where each possible edge appears with probability $1/2$
2. Let $S \subseteq [n]$. Denote by $E(S : S^c)$ be the set of edges between S and its complement (i.e. the size of the cut given by S).
3. $\mathbf{E}[\,|E(S : S^c)|\,] = \frac{|S|(n-|S|)}{2} \leq n^2/8$
4. Note that $C_S = |E(S : S^c)|$ depends on the possible $|S|(n - |S|)$ edges between S and S^c
5. $C_S = C_S(X_1, \ldots, X_m)$ where $m = |S|(n - |S|)$, where X_i indicates if the i-th edge appears in the cut or not
6. C_S is Lipschitz with $\boldsymbol{c} = (1, \ldots, 1)$
7. Therefore, for $\delta > 0$,

$$\mathbf{P}[\, C_S - \mathbf{E}[\, C_S \,] \geq \delta \mathbf{E}[\, C_S \,]\,] \leq \exp\left(-\frac{2\delta^2 \mathbf{E}[\, C_S \,]^2}{|S|(n - |S|)}\right)$$

8. **Exercise:** Deduce that for any $S \subseteq [n]$,

$$\mathbf{P}\left[C_S \geq \frac{n^2}{8} + \delta \frac{n^2}{4} \right] \leq e^{-\Omega(\delta^2 n^2)}$$

9. By the union bound, we have that

$$\mathbf{P}\left[\exists S : C_S \geq \frac{n^2}{8} + \delta \frac{n^2}{4} \right] \leq 2^n e^{-\Omega(\delta^2 n^2)} = 2^n e^{-\Omega(c^2 n)}$$

10. Recall that $\delta = c/\sqrt{n}$, now we pick c to be large enough, such that $2^n e^{-\Omega(c^2 n)} = 2^{-n}$
11. The main result is:

Theorem

There is a constant c, such that w.h.p. the Max Cut in $G_{n,1/2}$ is at most $n^2/8 + cn^{1/2}$