Lecture 3: "Linearity of Expectation"

Sotiris Nikoletseas
Professor

Computer Engineering and Informatics Department
2023 - 2024

Summary of previous lecture

- Common, underlying concept of all techniques:
  "Non-constructive proof of existence of combinatorial structures that have certain desired properties."
- Method of "positive probability":
  - Construct (by using abstract random experiments) an appropriate probability sample space of combinatorial structures (points $\leftrightarrow$ structures).
  - Prove that the probability of the desired property in this space is positive (i.e. non-zero). $\Rightarrow$ There is at least one combinatorial structure (since there is at least one point) with the desired property.

Summary of this lecture

i. Non-existence proofs using the *Markov Inequality*
ii. Proofs of existence using the *Linearity of Expectation method*

(I) Markov Inequality

Theorem 1

Let X be a non-negative random variable. Then:

$$\forall t > 0: \quad \Pr\{X \geq t\} \leq \frac{E[X]}{t}$$

Proof:

$$E[X] = \sum_{x} x \Pr\{X = x\} \geq \sum_{x \geq t} x \Pr\{X = x\}$$

$$\geq \sum_{x \geq t} t \Pr\{X = x\} = t \sum_{x \geq t} \Pr\{X = x\} = t \cdot \Pr\{X \geq t\}$$

$$\Rightarrow E[X] \geq t \cdot \Pr\{X \geq t\}$$

$$\Rightarrow \Pr\{X \geq t\} \leq \frac{E[X]}{t}$$

(I) Markov Inequality Application

It is actually a (weak) concentration inequality:

$$\Pr\left\{X \geq 2 \cdot E[X]\right\} \leq \frac{1}{2}$$

$$\Pr\left\{X \geq 3 \cdot E[X]\right\} \leq \frac{1}{3}$$

$$\vdots$$

$$\Pr\left\{X \geq k \cdot E[X]\right\} \leq \frac{1}{k}$$

A Basic Theorem

Theorem 2

Let X be a non-negative integer random variable. Then

$$\textit{if } E[X] \to 0 \quad \textit{then} \quad \Pr\{X = 0\} \to 1$$

Proof:
Using Markov's inequality for t=1 we have that:

$$\Pr\{X \geq 1\} \leq E[X]$$

If $E[X] \to 0$ then $\Pr\{X = 0\} \to 1$. □

1. Construct (by using abstract random experiments) an appropriate probability sample space of combinatorial structures.
2. Define the random variable X that corresponds to the number of structures with the desired property.
3. a. Express X as a sum of indicator variables $X = X_1 + X_2 + \cdots X_n$ where

$$X_i = \begin{cases} 1 & \text{the desired property holds} \\ 0 & \text{otherwise} \end{cases}$$

   b. Calculate $E[X]$ using linearity of expectation.
   c. Prove that $E[X] \to 0$ when $X \to \infty$
4. Conclude by using theorem 2 that w.h.p. the r.v. X is limited to 0. Hence, almost certainly there is no structure with the desired property.

Random Graph

Definition 1 (Random Graph)

A random graph is obtained by starting with a set of n isolated vertices and adding successive edges between them at random. In the $G_{n,p}$ model every possible edge occurs independently with probability $0 < p < 1$.

Example - Dominating Set

Definition 2 (Dominating Set)

Given an undirected graph $G = (V, E)$, a dominating set is a subset $S \subseteq V$ of its nodes such that for all nodes $v \in V$, either $v \in S$ or a neighbor u of v is in S.

Remark: The problem of finding a minimum dominating set is NP-hard. We will here address it by employing randomness.

We will show that smaller than logarithmic-size dominating sets do not exist (w.h.p.) in dense random graphs.

Theorem 3

For any $k < \ln n$: $\Pr\{\exists d.s.\ \text{of size } k \text{ in } G_{n,\frac{1}{2}}\} \to 0$

Proof of theorem 3 (1/2)

1. Let G be a graph generated using $G_{n,\frac{1}{2}}$ and S be any fixed set of k vertices of G.
2. Define r.v. X that corresponds to the number of dominating sets of size k.
3. a. $X = \sum_{S,|S|=k} X_S$ where X_S are indicator variables

$$X_S = \begin{cases} 1 & \text{S is d.s.} \\ 0 & \text{otherwise} \end{cases}$$

   b. Calculate $E[X]$: Using linearity of expectation:

$$E[X] = E\left[\sum_{S,|S|=k} X_S\right] = \sum_{S,|S|=k} E[X_S]$$

   - Calculate expectation of indicator variable X_S:

$$E[X_S] = 1 \cdot \Pr\{\text{S is d.s.}\} + 0 \cdot \Pr\{\text{S is not d.s.}\}$$

$$\Rightarrow E[X_S] = \Pr\{\text{S is d.s.}\}$$

Proof of theorem 3 (2/2)

- assume a vertex $v \notin$ d.s.: $\Pr\{\nexists(v,u) : u \in S\} = (\frac{1}{2})^k$

  $\Rightarrow \Pr\{\exists(v,u) : u \in S\} = 1 - (\frac{1}{2})^k$

  $\Rightarrow \Pr\{\forall v \text{ out of S}, \exists(v,u) : u \in S\} = (1 - \frac{1}{2^k})^{n-k}$

  $\Rightarrow E[X_S] = \Pr\{\text{S is d.s.}\} = (1 - \frac{1}{2^k})^{n-k}$

So,

$$E[X] = \sum_{S,|S|=k} E[X_S] = \binom{n}{k}\left(1 - \frac{1}{2^k}\right)^{n-k}$$

c. It holds that $\binom{n}{k} \leq n^k$ and $(1 - \frac{1}{2^k})^{n-k} < e^{-\frac{n-k}{2^k}}$

$$\Rightarrow E[X] \leq n^k e^{-\frac{n-k}{2^k}} \leq e^{\frac{k}{2^k}}\left(e^{k \ln n - \frac{n}{2^k}}\right)$$

If $k \ln n - \frac{n}{2^k} \to -\infty$ then $E[X] \to 0$,

So, if $k < \ln n \Rightarrow E[X] \to 0$

4. Using theorem 2 we prove that almost certainly there are no dominating sets of size $k < \ln n$

1. Construct (by using abstract random experiments) an appropriate probability sample space of combinatorial structures.
2. Define a random variable X that corresponds to the desired quantitative characteristics (e.g. the number or the size of the structures).
3. Express X as a sum of random indicator variables: $X = X_1 + X_2 + \cdots X_n$ where

$$X_i = \begin{cases} 1 & \text{the desired property holds} \\ 0 & \text{otherwise} \end{cases}$$

4. Calculate $E[X_i] = \Pr\{X_i = 1\}$.

5. Linearity of Expectation:

$$E[X] = E\left(\sum_i X_i\right) = \sum_i E[X_i]$$

even when X_i are dependent.

6. Obvious observation:
   - a random variable gets at least one value $\leq E[X]$ and at least one value $\geq E[X]$.
     Proof by contradiction:

     $$\mu = E[X] = \sum_x x \cdot f(x) > \sum_x \mu \cdot f(x) = \mu \sum_x f(x) = \mu$$

   - $\Rightarrow \exists$ at least one point (a structure) in the sample space for which $X \geq E[X]$ and at least one point (a structure) for which $X \leq E[X]$

(II) Linearity of Expectation Method

method's capabilities and limitations

- estimation of expectation suffices.
  - technically easy (indicator random variable $\Leftrightarrow$ probability of property)
  - linearity does not require stochastic independence (generic method)
- is associated with first moment Markov inequality:

$$\Pr\{X \geq t\} \leq \frac{E[X]}{t} \Leftrightarrow \Pr\{X \geq t \cdot E[X]\} \leq \frac{1}{t}$$

- for more powerful results:
  - inequalities with higher moments e.g. Chebyshev's inequality :

$$\Pr\{|X - \mu| \geq \lambda\sigma\} \leq \frac{1}{\lambda^2}$$

  - technical difficulties: linearity of variance generally requires stochastic independence (less generic methods)

Example - Tournament with many Hamiltonian Paths (1/2)

Theorem 4 (Szele, 1943)

For every positive integer n, there exists a tournament on n vertices with at least $n! \cdot 2^{-(n-1)}$ Hamiltonian paths.

Proof:

1. We construct a probability sample space with points corresponding to random tournaments by choosing the direction of each edge at random, equiprobably for the two directions and independently for every edge.
2. We define the r.v. X that corresponds to the number of Hamiltonian Paths.
3. Let σ be a permutation of the vertices of the tournament. We have that $X = \sum_\sigma X_\sigma$ where:

$$X_\sigma = \begin{cases} 1 & \sigma \text{ leads to a Hamiltonian Path} \\ 0 & \text{otherwise} \end{cases}$$

Example - Tournament with many Hamiltonian Paths (2/2)

4. A permutation σ leads to a Hamiltonian Path only if all edges have the same direction.

$$E[X_\sigma] = \Pr\{X_\sigma = 1\} = \left(\frac{1}{2}\right)^{n-1} = 2^{-(n-1)}$$

5. By linearity of Expectation:

$$E[X] = E\left(\sum_\sigma X_\sigma\right) = \sum_\sigma E[X_\sigma]$$

So, we have that:

$$E[X] = n! \cdot 2^{-(n-1)}$$

6. Thus, there must exist at least one tournament on n vertices which has at least $n! \cdot 2^{-(n-1)}$ Hamiltonian paths.

□

Example - Bipartite Subgraphs

Definition 3 (Bipartite Graph)

A bipartite graph is a graph whose vertices can be divided into two disjoint sets V_1 *and* V_2 *such that every edge connects a vertex in* V_1 *to one in* V_2*.*

Theorem 5

Every graph G=(V,E) has a bipartite subgraph with at least $\frac{|E|}{2}$ *edges.*

Proof of theorem 5

1. We construct a random sample space by choosing for every vertex in which set (V_1 or V_2) it belongs at random, equiprobably for the two sets and independently for each vertex. Thus, the points are random "bipartitions" of V.
2. We define the r.v. X that corresponds to the number of "crossing" edges (joining vertices in different parts).
3. Let g be an edge. We have that $X = \sum_{g \in E(G)} X_g$ where:

$$X_g = \begin{cases} 1 & \text{g is crossing} \\ 0 & \text{otherwise} \end{cases}$$

4. $E[X_g] = \Pr\{X_g = 1\} = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = 2 \cdot \frac{1}{4} = \frac{1}{2}$
5. By linearity of Expectation
$E[X] = E\left(\sum_{g \in E(G)} X_g\right) = \sum_{g \in E(G)} E[X_g] = |E| \cdot \frac{1}{2}$
6. Thus, there must exist a bipartite subgraph which has at least $\frac{|E|}{2}$ edges.