67865 - Mathematical Tools In Computer Science

February 20, 2024

Topics:

The course is divided into 3 mini courses.

1. Probability. (Weeks 1-3)
   (a) Rehersal
   (b) Useful Tools In research and CS
   (c) week 4 = break
2. Advanced topics in linear algebra. (Weeks 5-7/8)
3. Optimization. (Week 8/9 - 11)

Technicalities:

- Exercise every week.
- Final grade is the average of the best (n-3) exercises.
- Every week of miluim is excemption from 1 exercise. At least 4 exercises must be submitted.
- There is no exam, and there are no quizzes. :)
- The topics are going to be quite similar to previous years, so notes apply. If Amit does anything different than previous years, he will publish notes.

Part I

Probabilty

1 Rehersal

Today's Topics:

- Probablity spaces
- Conditional Probablity
- Independence
- Random Variables
- Expectation

1.1 Probablity Spaces (Rehersal)

Definition 1. (Probabilty Space/Merhav Histabrut) A probabilty space is a set Ω (is finite or countable - in this course) and a function $Pr : \Omega \to [0, 1]$ s.t. $\sum_{\omega\in\Omega} Pr(\omega) = 1$

Terminology

- $\omega \in \Omega$ is called an elementary/basic event.
- $A \subseteq \Omega$ is called an Event.
- Pr is extended to events by $Pr(A) = \sum_{\omega\in A} P(\omega)$
- Pr is uniform if for any $\omega \in \Omega$ it holds that $Pr(\omega) = \frac{1}{|\Omega|}$

Example. Throwing a die: $\Omega = \{1, 2, ..., 6\}$, $Pr(i) = \frac{1}{6}$
Throwing a pair of dice: $\Omega = \{1, ..., 6\}^2$, $Pr((i, j)) = \frac{1}{36}$

1.2 Conditional Probablity (Rehersal)

Definition 2. (Conditional Probablity) Let $A, B \subseteq \Omega$, s.t. $Pr(A) > 0$. Define the conditional probablity of B given A: $Pr(B|A) = \frac{Pr(A\cap B)}{Pr(A)}$

1.3 Independence (Rehersal)

Definition 3. (Independence) $A, B \subseteq \Omega$ are independent if $Pr(A \cap B) = Pr(A) \cdot Pr(B)$.

Note: If $Pr(A) > 0$, then A, B are independent $\iff P(B|A) = P(B)$

1.4 Random Variables (Rehersal)

Definition 4. (Random Variable) A random variable is a function $X : \Omega \to \mathbb{R}$.

Terminology

- If $Range(X) \subseteq \{0, 1\}$ then X is called Indicator random variable.
- $(X + Y)(\omega) = X(\omega) + Y(\omega)$
- $(X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega)$
- $(aX)(\omega) = aX(\omega)$
- $P(X = a) = P(\{\omega \in \Omega : X(\omega) = a\})$
- $P(X \leq Y) = P(\{\omega \in \Omega : X(\omega) = X(\omega)\})$
- $P(X = 7 : Y = 3) = P(\{\omega \in \Omega : X(\omega) = 7\}|\{\omega \in \Omega : Y(\omega) = 3\})$
- $\vdots$

Fact. If X is an indicator R.V. $\mathbb{E}(X) = Pr(X = 1)$.

Definition 5. (Independent Random Variables) Random Variables X, Y are independent if for any $a, b \in \mathbb{R}$ the events $\{\omega \in \Omega : X(\omega) = a\}$and $\{\omega \in \Omega : Y(\omega) = b\}$ are independent.

1.5 Expectation (Rehersal)

Definition 6. (Expectation) The expectation of a random variable X is $\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega) \cdot Pr(X = \omega)$.

Example.

$\Omega = \{1, 2, 3\} \times \{1, 2, 3\}$. $Pr((i, j)) = \frac{1}{9}$.

$X((i, j)) = i + j$

$\mathbb{E}(X) = \frac{1}{9}(X((1, 1)) + X((1, 2)), ..., X((3, 3))) = 4$

Lemma. (Linearity Of Expectation) Let X, Y be random variables, $a, b \in \mathbb{E}$. then:

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$$

Example. The expected number of fixed points in a random permutation.

- A fixed point in a permuation $\pi : [n] \to [n]$ (1:1 function) is $i \in [n]$ s.t. $\pi(i) = i$.
- $S_n = \{(\pi : [n] \to [n]) : \pi$ pi is a permutation$\}$

$Pr(\pi) = \frac{1}{n!}$
$X(\pi) = |\{i : \pi(i) = i\}|$ (number of fixed points of pi) This is indeed a random variables. $X : S_n \to \mathbb{R}$.
Question: What is the expected value of X.
For every $i \in [n]$, we define $X_i(\pi) = \begin{cases} 1 & \pi(i) = i \\ 0 & otherwise \end{cases}$.
So $X = \sum_{i=1}^{n} X_i$
$\Rightarrow \mathbb{E}(X) \overset{*_1}{=} \sum_{i=1}^{n} \mathbb{E}(X_i) \overset{*_2}{=} \sum_{i=1}^{n} Pr(X_i = 1) = \sum_{i=1}^{n} \frac{1}{n} = 1$
Where $*_1$ is from linearity of expectation and $*_2$ is from X_i's being indicator random variables.

1.6 Inequalities (Rehersal)

1.6.1 Union Bound

Lemma. (Union Bound) If $A, ..., A_k \subseteq \Omega$ then:

$$Pr(\bigcup_{i=1}^{k} A_i) \leq \sum_{i=1}^{k} Pr(A_i)$$

1.6.2 Markov

Theorem. (Markov Inequality) Let $X : \Omega \to \mathbb{R}_+$ (NON NEGATIVE!) be a R.V. then for any $c > 0$:

$$Pr(X \geq c\mathbb{E}(X)) \leq \frac{1}{c} \iff Pr(X \geq c) \leq \frac{\mathbb{E}(X)}{c}$$

Proof.

$$
\begin{aligned}
\mathbb{E}[X] &\overset{def}{=} \sum_{\omega\in\Omega} Pr(\omega)\cdot X(\omega) \\
&\overset{*_1}{\geq} \sum_{\substack{\omega\in\Omega \\ X(\omega)\geq c\mathbb{E}[X]}} Pr(\omega)\cdot X(\omega) \\
&\geq \sum_{\substack{\omega\in\Omega \\ X(\omega)\geq c\mathbb{E}[X]}} Pr(\omega)\cdot c\mathbb{E}[X] \\
&= c\mathbb{E}[X]\cdot \sum_{\substack{\omega\in\Omega \\ X(\omega)\geq c\mathbb{E}[X]}} Pr(\omega) \\
&\overset{def}{=} c\mathbb{E}[X]\cdot Pr(X\geq c\mathbb{E}[X])
\end{aligned}
$$

$*_1$ is because X is non negative.

$$\Rightarrow Pr(X \geq c\mathbb{E}(X)) \leq \frac{1}{c}$$

$\square$

1.6.3 Chebyshev

Definition 7. (Variance and standard deviation) The Variance of R.V. X is:

$$Var(X) = \mathbb{E}(X-\mathbb{E}(X))^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

The standard deviation of X is $\sigma = \sqrt{Var(X)}$.

Fact.

$$
\begin{aligned}
Var(X) &= \mathbb{E}(X-\mathbb{E}(X))^2 \\
&= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) \\
&= \mathbb{E}(X^2) - \mathbb{E}(2X\mathbb{E}(X)) + \mathbb{E}(X)^2 \\
&= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\
&= \mathbb{E}(X^2) - \mathbb{E}(X)^2
\end{aligned}
$$

Fact. $Var(X) \leq \mathbb{E}(X^2)$ ($\leftarrow$ called second moment)

Theorem. (Chebychev inequality) Let X be a R.V. then:

$$Pr\left(|X - \mathbb{E}(X)| \geq c\sqrt{Var(X)}\right) \leq \frac{1}{c^2} \iff Pr\left(|X - \mathbb{E}(X)| \geq c\right) \leq \frac{Var(X)}{c^2}$$

Proof. Denote $Y(\omega) = (X - \mathbb{E}(X))^2$. By markov inequality:

$$Pr(Y \geq c^2\mathbb{E}(Y)) \leq \frac{1}{c^2}$$

$$\begin{aligned} Pr(Y \geq c^2\mathbb{E}(Y)) =& Pr((X - \mathbb{E}(X))^2 \geq c^2 Var(X)) \\ =& Pr(X - \mathbb{E}(X) \geq c\sqrt{Var(X)}) \end{aligned}$$

□

2 Ramsey Theory & Probabalistic Method

2.1 Ramsey Theory

Ramsey theory tries to find the largest possible "island" of "order" in a "chaos".

Definition 8. A Coloring of a graph $G = (V, E)$ is a function $c : E \to \{Red(R), Black(B)\}$.

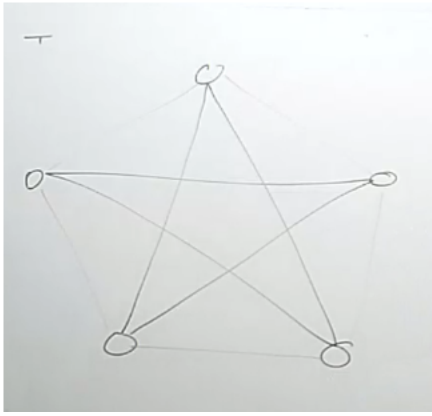
Definition 9. $S \subseteq V$ is Monochromatic w.r.t coloring c, if all edges whose both vertices are in S are in the same colors.

Definition 10. The Ramsey Number R(k) is the minimal number R s.t. for for each coloring of K_R(The complete graph of R vertices) there is a monochromtic set $S \subseteq V$ of size k.

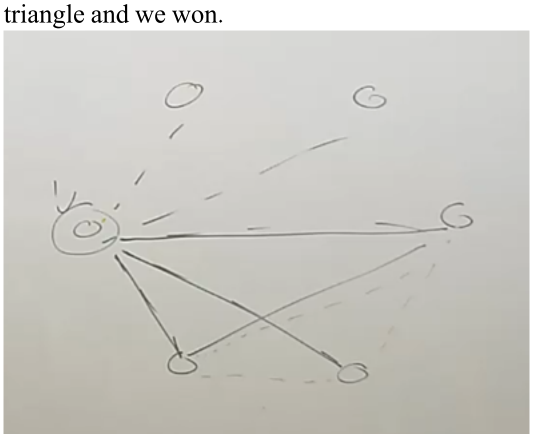
Claim. $R(2) = 2$ (trivial)

Claim. $R(3) = 6$

Proof. first, we can see that $R(3) > 5$ becuase of the following coutner example.



Now we will show that $R(3) \leq 6$. We look at a verticie. 3 of it's edges have to be either black or red. assume they are red, now look at the 3 edges connecting the other vrtices of the red edges. if they are all black we won. if one of them is red there is a red triangle and we won.



□

Claim. $R(4) = 18$ (hard, discovered 1978)

Claim. $R(5)$ = Unknown (between 43 and 48)

Theorem. (Ramsey)

$$2^{\frac{k}{2}-1} < R(k) \leq 2^{2k}$$

Proof. Upper bound in the tirgul: Generalization of $R(3) \leq 6$.

Lower Bound: Using the **probabilstic method**.

Denote $N = 2^{\frac{k}{2}-1}$. Let C be a random coloring of K_N (The complete graph of N vertices). Denote A the event that there is a mono set of size k: $A = \{c : E \to \{R, B\}$: there is a mono set of size k$\}$

is is enough to show that $Pr(A) < 1$.

For $S \subseteq V$s.t. $|S| = k$, Denote $A_S = \{c : S$ is mono w.r.t $c\}$.

As such $A = \bigcup_{S \subseteq V, |S|=k} A_s$.

$Pr(A_s) = 2^{-\binom{k}{2}+1}$ because there a $\binom{k}{2}$ edges formed in S. and they all need to be the same color. the $+1$ is because they can all be either red or black.

By the union bound:

$$
\begin{aligned}
Pr(A) \leq & \sum_{S \subseteq V, |S|=k} Pr(A_s) \\
= & \sum_{S \subseteq V, |S|=k} 2^{-\binom{k}{2}+1} \\
= & \binom{N}{k} \cdot 2^{-\binom{k}{2}+1} \\
\leq & N^k \cdot 2^{-\binom{k}{2}+1} \\
= & 2^{(\frac{k}{2}-1)k} \cdot 2^{-\binom{k}{2}+1} \\
= & 2^{\frac{k^2}{2}-k-\frac{k(k-1)}{2}+1} \\
= & 2^{-\frac{k}{2}+1} \\
\overset{*}{<} & 1
\end{aligned}
$$

$*$ is for $k \geq 3$, for $k = 2, 1$ the case is trivial. $\square$