CMPT 210: Probability and Computing

Lecture 12

Sharan Vaswani

October 17, 2024

Recap - (Basic) Frievald's Algorithm

Q: For $n \times n$ matrices A, B and D, is $D = AB$?

Last class, we proved that:

	Yes	No
$D = AB$	1	0
$D \neq AB$	$< \frac{1}{2}$	$\geq \frac{1}{2}$

Table 1: Probabilities for Basic Frievalds Algorithm

Frievald's Algorithm

- By repeating the *Basic Frievald's Algorithm* m times, we will amplify the probability of success.

The resulting complete Frievald's Algorithm is given by:

1. Run the Basic Frievald's Algorithm for m independent runs.
2. If *any* run of the Basic Frievald's Algorithm outputs "no", output "no".
3. If *all* runs of the Basic Frievald's Algorithm output "yes", output "yes".

	Yes	No
$D = AB$	1	0
$D \neq AB$	$< \frac{1}{2^m}$	$\geq 1 - \frac{1}{2^m}$

Table 2: Probabilities for Frievald's Algorithm

- If $m = 20$, then Frievald's algorithm will make mistake with probability $1/2^{20} \approx 10^{-6}$.

Computational Complexity: $O(mn^2)$

Probability Amplification

- Consider a randomized algorithm $\mathcal{A}$ that is supposed to solve a binary decision problem i.e. it is supposed to answer either Yes or No. It has a one-sided error – (i) if the true answer is Yes, then the algorithm $\mathcal{A}$ correctly outputs Yes with probability 1, but (ii) if the true answer is No, the algorithm $\mathcal{A}$ incorrectly outputs Yes with probability $\leq \frac{1}{2}$.

Let us define a new algorithm $\mathcal{B}$ that runs algorithm $\mathcal{A}$ m times, and if *any* run of $\mathcal{A}$ outputs No, algorithm $\mathcal{B}$ outputs No. If *all* runs of $\mathcal{A}$ output Yes, algorithm $\mathcal{B}$ outputs Yes.

Q: What is the probability that algorithm $\mathcal{B}$ correctly outputs Yes if the true answer is Yes, and correctly outputs No if the true answer is No?

Probability Amplification - Analysis

If A_i denotes run i of Algorithm $\mathcal{A}$, then

$$\Pr[\mathcal{B} \text{ outputs Yes} \mid \text{true answer is Yes }]$$

$$= \Pr[\mathcal{A}_1 \text{ outputs Yes } \cap \mathcal{A}_2 \text{ outputs Yes } \cap \ldots \cap \mathcal{A}_m \text{ outputs Yes} \mid \text{true answer is Yes }]$$

$$= \prod_{i=1}^{m} \Pr[\mathcal{A}_i \text{ outputs Yes} \mid \text{true answer is Yes }] = 1 \qquad \text{(Independence of runs)}$$

$$\Pr[\mathcal{B} \text{ outputs No} \mid \text{true answer is No }]$$

$$= 1 - \Pr[\mathcal{B} \text{ outputs Yes} \mid \text{true answer is No }]$$

$$= 1 - \Pr[\mathcal{A}_1 \text{ outputs Yes } \cap \mathcal{A}_2 \text{ outputs Yes } \cap \ldots \cap \mathcal{A}_m \text{ outputs Yes} \mid \text{true answer is No }]$$

$$= 1 - \prod_{i=1}^{m} \Pr[\mathcal{A}_i \text{ outputs Yes} \mid \text{true answer is No }] \geq 1 - \frac{1}{2^m}.$$

When the true answer is Yes, both $\mathcal{B}$ and $\mathcal{A}$ correctly output Yes. When the true answer is No, $\mathcal{A}$ incorrectly outputs Yes with probability $< \frac{1}{2}$, but $\mathcal{B}$ incorrectly outputs Yes with probability $< \frac{1}{2^m} << \frac{1}{2}$. By repeating the experiment, we have "amplified" the probability of success.

Questions?

Random Variables

Definition: A random "variable" R on a probability space is a total function whose domain is the sample space $\mathcal{S}$. The codomain is usually a subset of the real numbers.

Example: Suppose we toss three independent, unbiased coins. Let C be the number of heads that appear.

$\mathcal{S} = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

C is a total function that maps each outcome in $\mathcal{S}$ to a number as follows: $C(HHH) = 3$, $C(HHT) = C(HTH) = C(THH) = 2$, $C(HTT) = C(THT) = C(TTH) = 1$, $C(TTT) = 0$.

C is a random variable that counts the number of heads in 3 tosses of the coin.

Example: I toss a coin, and define the random variable R which is equal to 1 when I get a heads, and equal to 0 when I get a tails.

Bernoulli random variables: Random variables with the codomain $\{0, 1\}$ are called Bernoulli random variables. E.g. R is a Bernoulli r.v.

Back to throwing dice

Q: Suppose we throw two standard dice one after the other. Let us define R to be the random variable equal to the sum of the dice. What is the domain, range of R?

Ans: $R : \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} \to \mathbb{N} \cap [2,12]$.
$R((4,7)) = 11$, $R((4,1)) = 5$, $R((1,1)) = 2$, $R((6,6)) = 12$.

Q: Three balls are randomly selected from an urn containing 20 balls numbered 1 through 20. The random variable M is the maximal value on the selected balls. What is the domain, range of M? Ans: $M : \{1,2,\ldots,20\} \times \{1,2,\ldots,20\} \times \{1,2,\ldots,20\} \to \{1,2,\ldots,20\}$

Q: In the above example, what is $2 \times M((1,4,6))$? Is M an invertible function? Ans: 12, No since M maps both $\{1,2,5)$ and $(3,4,5)$ to 5.