CMPT 210: Probability and Computing

Lecture 11

Sharan Vaswani

February 9, 2023

Recap - (Basic) Frievald's Algorithm

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

Algorithm:

1. Generate a random n-bit vector x, by making each bit x_i either 0 or 1 *independently* with probability $\frac{1}{2}$. E.g, for $n = 2$, toss a fair coin independently twice with the scheme – H is 0 and T is 1). If we get HT, then set $x = [0\,;\,1]$.
2. Compute $t = Bx$ and $y = At = A(Bx)$ and $z = Dx$.
3. Output "yes" if $y = z$ (all entries need to be equal), else output "no".

Computational complexity: Step 1 can be done in $O(n)$ time. Step 2 requires 3 matrix vector multiplications and can be done in $O(n^2)$ time. Step 3 requires comparing two n-dimensional vectors and can be done in $O(n)$ time. Hence, the total computational complexity is $O(n^2)$.

(Basic) Frievald's Algorithm

Let us analyze the algorithm for general matrix multiplication.

Case (i): If $D = AB$, does the algorithm always output "yes"? Yes! Since $D = AB$, for any vector x, $Dx = ABx$.

Case (ii) If $D \neq AB$, does the algorithm always output "no"?

Claim: For any input matrices A, B, D if $D \neq AB$, then the (Basic) Frievald's algorithm will output "no" with probability $\geq \frac{1}{2}$.

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

Table 1: Probabilities for Basic Frievalds Algorithm

(Basic) Frievald's Algorithm

Proof: If $D \neq AB$, we wish to compute the probability that algorithm outputs "yes" and prove that it less than $\frac{1}{2}$.
Define $E := (AB - D)$ and $r := Ex = (AB - D)x = y - z$. If $D \neq AB$, then $\exists (i, j)$ s.t. $E_{i,j} \neq 0$.

$$
\begin{aligned}
\Pr[\text{Algorithm outputs "yes"}] &= \Pr[y = z] = \Pr[r = \mathbf{0}] \\
&= \Pr[(r_1 = 0) \cap (r_2 = 0) \cap \ldots \cap (r_i = 0) \cap \ldots] \\
&= \Pr[(r_i = 0)] \Pr[(r_1 = 0) \cap (r_2 = 0) \cap \ldots \cap (r_n = 0) | r_i = 0] \\
&\qquad\qquad \text{(By def. of conditional probability)} \\
\implies \Pr[\text{Algorithm outputs "yes"}] &\leq \Pr[r_i = 0] \qquad \text{(Probabilities are in } [0, 1])
\end{aligned}
$$

To complete the proof, on the next slide, we will prove that $\Pr[r_i = 0] \leq \frac{1}{2}$.

(Basic) Frievald's Algorithm

$$r_i = \sum_{k=1}^{n} E_{i,k}x_k = E_{i,j}x_j + \sum_{k \neq j} E_{i,k}x_k = E_{i,j}x_j + \omega \qquad (\omega := \textstyle\sum_{k \neq j} E_{i,k}x_k)$$

$$\Pr[r_i = 0] = \Pr[r_i = 0 | \omega = 0]\Pr[\omega = 0] + \Pr[r_i = 0 | \omega \neq 0]\Pr[\omega \neq 0]$$

(By the law of total probability)

$$\Pr[r_i = 0 | \omega = 0] = \Pr[x_j = 0] = \frac{1}{2} \qquad (\text{Since } E_{i,j} \neq 0 \text{ and } \Pr[x_j = 1] = \tfrac{1}{2})$$

$$\Pr[r_i = 0 | \omega \neq 0] = \Pr[(x_j = 1) \cap E_{i,j} = -\omega] = \Pr[(x_j = 1)]\Pr[E_{i,j} = -\omega | x_j = 1]$$

(By def. of conditional probability)

$$\implies \Pr[r_i = 0 | \omega \neq 0] \leq \Pr[(x_j = 1)] = \frac{1}{2} \qquad (\text{Probabilities are in } [0,1],\ \Pr[x_j = 1] = \tfrac{1}{2})$$

$$\implies \Pr[r_i = 0] \leq \frac{1}{2}\Pr[\omega = 0] + \frac{1}{2}\Pr[\omega \neq 0] = \frac{1}{2}\Pr[\omega = 0] + \frac{1}{2}[1 - \Pr[\omega = 0]] = \frac{1}{2}$$

$(\Pr[E^c] = 1 - \Pr[E])$

$$\implies \Pr[\text{Algorithm outputs ``yes''}] \leq \Pr[r_i = 0] \leq \frac{1}{2}.$$

(Basic) Frievald's Algorithm

Hence, if $D \neq AB$, the Algorithm outputs "yes" with probability $\leq \frac{1}{2} \implies$ the Algorithm outputs "no" with probability $\geq \frac{1}{2}$.

In the worst case, the algorithm can be incorrect half the time! We promised the algorithm would return the correct answer with "high" probability close to 1.

A common trick in randomized algorithms is to have m independent trials of an algorithm and aggregate the answer in some way, reducing the probability of error, thus *amplifying the probability of success*.

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

$$
\begin{aligned}
&\Pr[\mathcal{B} \text{ outputs Yes} \mid \text{true answer is Yes }] \\
&= \Pr[\mathcal{A}_1 \text{ outputs Yes } \cap \mathcal{A}_2 \text{ outputs Yes } \cap \ldots \cap \mathcal{A}_m \text{ outputs Yes} \mid \text{true answer is Yes }] \\
&= \prod_{i=1}^{m} \Pr[\mathcal{A}_i \text{ outputs Yes} \mid \text{true answer is Yes }] = 1 \qquad \text{(Independence of runs)} \\
&\Pr[\mathcal{B} \text{ outputs No} \mid \text{true answer is No }] \\
&= 1 - \Pr[\mathcal{B} \text{ outputs Yes} \mid \text{true answer is No }] \\
&= 1 - \Pr[\mathcal{A}_1 \text{ outputs Yes } \cap \mathcal{A}_2 \text{ outputs Yes } \cap \ldots \cap \mathcal{A}_m \text{ outputs Yes} \mid \text{true answer is No }] \\
&= 1 - \prod_{i=1}^{m} \Pr[\mathcal{A}_i \text{ outputs Yes} \mid \text{true answer is No }] \geq 1 - \frac{1}{2^m}.
\end{aligned}
$$

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.

Random Variables and Events

Indicator Random Variable: An indicator random variable maps every outcome to either 0 or 1.

Example: Suppose we throw two standard dice, and define M to be the random variable that is 1 iff both throws of the dice produce a prime number, else it is 0.

$M : \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} \to \{0, 1\}$. $M((2, 3)) = 1$, $M((3, 6)) = 0$.

An indicator random variable partitions the sample space into those outcomes mapped to 1 and those outcomes mapped to 0.

Example: When throwing two dice, if E is the event that both throws of the dice result in a prime number, then random variable $M = 1$ iff event E happens, else $M = 0$.

The indicator random variable corresponding to an event E is denoted as $\mathcal{I}_E$, meaning that for $\omega \in E$, $\mathcal{I}_E[\omega] = 1$ and for $\omega \notin E$, $\mathcal{I}_E[\omega] = 0$. In the above example, $M = \mathcal{I}_E$ and since $(2, 4) \notin E$, $M((2, 4)) = 0$ and since $(3, 5) \in E$, $M((3, 5)) = 1$.