Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

- Approach 1: Compute $A \cdot B$ and compare its entries with C. Difficulty: The fastest known algorithm for matrix multiplication works in time $O(n^{2.37})$.

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

- Approach 1: Compute $A \cdot B$ and compare its entries with C. Difficulty: The fastest known algorithm for matrix multiplication works in time $O(n^{2.37})$.
- Can we do better?

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

- Approach 1: Compute $A \cdot B$ and compare its entries with C. Difficulty: The fastest known algorithm for matrix multiplication works in time $O(n^{2.37})$.
- Can we do better? Yes, we can.

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

- Approach 1: Compute $A \cdot B$ and compare its entries with C. Difficulty: The fastest known algorithm for matrix multiplication works in time $O(n^{2.37})$.
- Can we do better? Yes, we can.
- Frievalds' randomized algorithm: For an integer parameter k

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

- Approach 1: Compute $A \cdot B$ and compare its entries with C. Difficulty: The fastest known algorithm for matrix multiplication works in time $O(n^{2.37})$.
- Can we do better? Yes, we can.
- Frievalds' randomized algorithm: For an integer parameter k
- Works in time $O(kn^2)$.

Matrix multiplication testing

Observation (From last time)

$P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$ *if* $B_1, \ldots, B_n$ *are pairwise disjoint,* $B_1 \cup \cdots \cup B_n = \Omega$, *and* $P[B_j] > 0$ *for* $j \in [n]$.

Problem

Give three $n \times n$ *matrices* A, B, C, *find an algorithm which quickly determines whether* $A \cdot B = C$.

- Approach 1: Compute $A \cdot B$ and compare its entries with C. Difficulty: The fastest known algorithm for matrix multiplication works in time $O(n^{2.37})$.
- Can we do better? Yes, we can.
- Frievalds' randomized algorithm: For an integer parameter k
- Works in time $O(kn^2)$.
- Answers correctly with probability at least $1 - 2^{-k}$.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0 - 1$ vector r with n coordinates.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0-1$ vector r with n coordinates.
- Compute $v = A(Br) - Cr$.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0-1$ vector r with n coordinates.
- Compute $v = A(Br) - Cr$.
- Answer YES, if $v = 0$, NO otherwise.

Correctness:

- If $AB = C$, then the answer must be YES.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0-1$ vector r with n coordinates.
- Compute $v = A(Br) - Cr$.
- Answer YES, if $v = 0$, NO otherwise.

Correctness:

- If $AB = C$, then the answer must be YES.
- If $AB \neq C$, we will show that the algorithm NO with probability at least $\frac{1}{2}$.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0-1$ vector r with n coordinates.
- Compute $v = A(Br) - Cr$.
- Answer YES, if $v = 0$, NO otherwise.

Correctness:

- If $AB = C$, then the answer must be YES.
- If $AB \neq C$, we will show that the algorithm NO with probability at least $\frac{1}{2}$.
- we generate r k-times independently in each run, and the final answer will be NO if at least one NO appears.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0-1$ vector r with n coordinates.
- Compute $v = A(Br) - Cr$.
- Answer YES, if $v = 0$, NO otherwise.

Correctness:

- If $AB = C$, then the answer must be YES.
- If $AB \neq C$, we will show that the algorithm NO with probability at least $\frac{1}{2}$.
- we generate r k-times independently in each run, and the final answer will be NO if at least one NO appears.
- Then the probability of a mistake is at most 2^{-k}.

Frievalds' algorithm

Algorithm

INPUT: $n \times n$ matrices A, B, C (say over $\mathbb{Q}$).
OUTPUT: Answer whether $AB = C$?
ALGORITHM:

- Pick a random $0-1$ vector r with n coordinates.
- Compute $v = A(Br) - Cr$.
- Answer YES, if $v = 0$, NO otherwise.

Correctness:

- If $AB = C$, then the answer must be YES.
- If $AB \neq C$, we will show that the algorithm NO with probability at least $\frac{1}{2}$.
- we generate r k-times independently in each run, and the final answer will be NO if at least one NO appears.
- Then the probability of a mistake is at most 2^{-k}.

Running time: $O(n^2)$ for single run, $O(kn^2)$ for k runs.

Reliability of Frievalds' algorithm

Claim

If $AB \neq C$, then the algorithm answers NO with probability at least $\frac{1}{2}$.

Proof.

- Let $D := AB - C \neq 0$. If $D = (d_{ij})$, then $\exists i, j$ with $d_{ij} \neq 0$.

Reliability of Frievalds' algorithm

Claim

If $AB \neq C$, then the algorithm answers NO with probability at least $\frac{1}{2}$.

Proof.

- Let $D := AB - C \neq 0$. If $D = (d_{ij})$, then $\exists i, j$ with $d_{ij} \neq 0$.

$$\begin{pmatrix} — & d_{ij} & — \end{pmatrix} \begin{pmatrix} r_1 \\ \\ r_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \\ v_n \end{pmatrix} \qquad v_i = \sum_{k=1}^{n} d_{ik} r_k = d_{ij} r_j + y$$

Reliability of Frievalds' algorithm

Claim

If $AB \neq C$, then the algorithm answers NO with probability at least $\frac{1}{2}$.

Proof.

- Let $D := AB - C \neq 0$. If $D = (d_{ij})$, then $\exists i, j$ with $d_{ij} \neq 0$.

$$\begin{pmatrix} — & d_{ij} & — \end{pmatrix} \begin{pmatrix} r_1 \\ \\ r_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \\ v_n \end{pmatrix} \qquad v_i = \sum_{k=1}^{n} d_{ik} r_k = d_{ij} r_j + y$$

- $P[v_i = 0] = \underbrace{P[v_i = 0 | y = 0]}_{=1/2} P[y = 0] + \underbrace{P[v_i = 0 | y \neq 0]}_{\leq 1/2} P[y \neq 0] \leq \frac{1}{2}(P[y = 0] + P[y \neq 0]) = \frac{1}{2}$. □

2nd moment

Definition

The variance (rozptyl) of a real random variable X is defined as

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

2nd moment

Definition

The variance (rozptyl) of a real random variable X is defined as

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

Remarks:

- Standard deviation (směrodatná odchylka) $\sigma = \sqrt{\mathrm{Var}\, X}$.

2nd moment

Definition

The variance (rozptyl) of a real random variable X is defined as

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

Remarks:

- Standard deviation (směrodatná odchylka) $\sigma = \sqrt{\mathrm{Var}\, X}$.
- It could seem more natural to work with $E[|X - E[X]|]$; however, in reality, it is much more complicated with this expression.

2nd moment

Definition

The variance (rozptyl) of a real random variable X is defined as

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

Remarks:

- Standard deviation (směrodatná odchylka) $\sigma = \sqrt{\mathrm{Var}\, X}$.
- It could seem more natural to work with $E[|X - E[X]|]$; however, in reality, it is much more complicated with this expression.

Definition

The covariance (kovariance) of two random variables X, Y is defined as

$$\mathrm{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

2nd moment

Definition

The variance (rozptyl) of a real random variable X is defined as

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

Remarks:

- Standard deviation (směrodatná odchylka) $\sigma = \sqrt{\mathrm{Var}\, X}$.
- It could seem more natural to work with $E[|X - E[X]|]$; however, in reality, it is much more complicated with this expression.

Definition

The covariance (kovariance) of two random variables X, Y is defined as

$$\mathrm{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

- If X and Y are independent, then $\mathrm{Cov}[X, Y] = 0$.

Variance of a sum of random variables

Lemma

Let $X_1, \ldots, X_n$ be random variables. Then

$$\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{i \neq j} \mathrm{Cov}[X_i, X_j].$$

Variance of a sum of random variables

Lemma

Let $X_1, \ldots, X_n$ be random variables. Then

$$\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{i \neq j} \mathrm{Cov}[X_i, X_j].$$

Proof.

$$\begin{aligned}
\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] &= E\Big[\Big(\sum_{i=1}^{n} X_i\Big)\Big(\sum_{j=1}^{n} X_j\Big)\Big] - E\Big[\Big(\sum_{i=1}^{n} X_i\Big)\Big] E\Big[\Big(\sum_{j=1}^{n} X_j\Big)\Big] \\
&= \sum_{i,j=1}^{n} E[X_i X_j] - \sum_{i,j=1}^{n} E[X_i] E[X_j] \\
&= \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j]
\end{aligned}$$

□

Chebyshev's inequality

Lemma (Chebyshev's inequality)

Let X be a random variable with finite variace. Then

$$P[|X - E[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$$

for every $t > 0$.

Chebyshev's inequality

Lemma (Chebyshev's inequality)

Let X be a random variable with finite variace. Then

$$P[|X - E[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$$

for every $t > 0$.

Proof.

Let $Y = (X - E[X])^2$. Then:

Chebyshev's inequality

Lemma (Chebyshev's inequality)

Let X be a random variable with finite variace. Then

$$P[|X - E[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$$

for every $t > 0$.

Proof.

Let $Y = (X - E[X])^2$. Then:

$$P[|X - E[X]| \geq t] = P[Y \geq t^2] \leq \frac{E[Y]}{t^2} = \frac{\mathrm{Var}[X]}{t^2}. \quad \square$$

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] =$

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] = \frac{2m}{2} = m$.

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] = \frac{2m}{2} = m$.
- $\mathrm{Var}[X] =$

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] = \frac{2m}{2} = m$.
- $\mathrm{Var}[X] = \frac{2m}{4} = \frac{m}{2}$.

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] = \frac{2m}{2} = m$.
- $\mathrm{Var}[X] = \frac{2m}{4} = \frac{m}{2}$.
- Chebyshev: $P[|X - m| < \sqrt{m}] \geq 1 - \frac{m/2}{m} = \frac{1}{2}$.

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] = \frac{2m}{2} = m$.
- $\mathrm{Var}[X] = \frac{2m}{4} = \frac{m}{2}$.
- Chebyshev: $P[|X - m| < \sqrt{m}] \geq 1 - \frac{m/2}{m} = \frac{1}{2}$.
- $P[X = m + k] = \binom{2m}{m+k}\frac{1}{2^{2m}}$. Altogether:

Bounding the middle binomial coefficient

- It is well known that $\binom{2m}{m} = \Theta(\frac{4^m}{\sqrt{m}})$.
- We will show one inequality.

Proposition

For every $m \geq 0$ we have $\binom{2m}{m} \geq \frac{4^m}{4\sqrt{m}+2}$.

Proof.

- Let $X_1, \ldots, X_{2m}$ be independent random $0/1$ variables such that each value is attained with probability $\frac{1}{2}$.
- Let $X = X_1 + \cdots + X_{2m}$.
- $E[X] = \frac{2m}{2} = m$.
- $\mathrm{Var}[X] = \frac{2m}{4} = \frac{m}{2}$.
- Chebyshev: $P[|X - m| < \sqrt{m}] \geq 1 - \frac{m/2}{m} = \frac{1}{2}$.
- $P[X = m + k] = \binom{2m}{m+k}\frac{1}{2^{2m}}$. Altogether:
- $\frac{1}{2} \leq \sum_{|k|<\sqrt{m}} P[X = m + k] \leq (2\sqrt{m} + 1)\binom{2m}{m}\frac{1}{2^{2m}}$. □