Contents

Algorytmika
Ćwiczenia 3

Adrian Herda

2025-06-29

$$\mu = E[X]$$

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

$$\sigma = \mathrm{Std}(X) = \sqrt{\mathrm{Var}(x)} = \sqrt{E[X^2] - E[X]^2}$$

1. Zadanie 12 - losowanie monetą

Rzucamy monetą dwa razy:

- $\Pr(\text{Orzeł}, \text{Reszka}) = p \cdot q$ – Orzeł monety normalnej
- $\Pr(\text{Reszka}, \text{Orzeł}) = q \cdot p$ – Reszka monety normalnej

Oba wyniki są sobie równe, jeśli będziemy ignorować wyniki (Orzeł, Orzeł) oraz (Reszka, Reszka), dostajemy monetę uczciwą.

$$\Pr(\text{Orzeł}) = \Pr(\text{Reszka}) = \frac{pq}{pq + pq} = \frac{1}{2}$$

2. Zadanie 13 - jednostajnie losowane $[0, \cdots, 6]$

1. Mamy generator liczb losowych ze zbioru $\{0, \cdots, 4\}$, niech wynik tego generatora będzie oznaczony literą X.
2. Tworzymy generator liczb losowych ze zbioru $\{1, \cdots, 24\}$

$$Y = 5a + b$$

   gdzie $a, b \in \{0, \cdots, 4\}$ są generowane przez generator z punktu 1. Z tego wynika że:

$$\Pr(Y = y) = \Pr(Y = 5a + b) = \Pr(X = a) \cdot \Pr(X = b) = \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25}$$

3. Zbiór liczb $\{0, \cdots, 20\}$ ma $3 \cdot 7$ liczb co pozwoli nam na dopasowanie równej ilości liczb do zbioru $\{0, \cdots, 6\}$ za pomocą operacji $\bmod 7$.
4. Jeśli $Y > 20$ odrzucamy wynik i ponownie losujemy dwie liczby, powtarzamy dopóki nie wylosujemy $Y \leq 20$, w takim przypadku wylosowaną zmienną liczymy

$$Z = Y \bmod 7$$

   i w ten sposób dostajemy liczbę ze zbioru $\{0, \cdots, 6\}$ z prawdopodobieństwem:

$$\forall_{z \in \{0, \cdots, 6\}} \left(\Pr(Y \bmod 7 = z) = \frac{3}{25} \right) \Longrightarrow \Pr(Z = z) = \frac{1}{7}$$

3. Zadanie 14 - pole koła z Monte Carlo

3.1. Var(X)

Losujemy n punktów $(x_i, y_i) \in [0,1]^2$
Niech:

$$Z_i = \begin{cases} 1 \text{ jeśli } x_i^2 + y_i^2 \leq 1 \\ 0 \text{ w przeciwnym razie} \end{cases}$$

$$\left(\forall_{i \in [n]}\right)\left(\Pr(Z_i = 1) = \frac{\frac{\text{Area}_\bigcirc}{4}}{\Omega} = \frac{\frac{\pi}{4}}{1 \cdot 1} = \frac{\pi}{4} \Rightarrow E[Z_i] = \frac{\pi}{4} \cdot 1 + \left(1 - \frac{\pi}{4}\right) \cdot 0 = \frac{\pi}{4}\right)$$

Mamy do czynienia z rozkładem Bernouliego z prawdopodobieństwem wynoszącym tyle co pole ćwiartki koła czyli $p = \frac{\pi}{4}$. Zatem wiemy od razu że

$$\left(\forall_{i \in [n]}\right)\left(E[Z_i] = p = \frac{\pi}{4}\right)$$

oraz że

$$\left(\forall_{i \in [n]}\right)(\text{Var}(Z_i) = p(1-p))$$

Wtedy pole:

$$A_n = \frac{1}{n} \sum_{i=1}^{n} Z_i$$

Wiemy że wybierane punkty są od siebie niezależne więc $\left(\forall_{i,j \in [n], i \neq j}\right)\left(\text{Cov}\left(Z_i, Z_j\right) = 0\right)$. Ze wzoru na wariancje z mnożnikiem stałym oraz ze wzoru na wariancje sumy zmiennych niezależnych:

$$\sigma_n^2 = \text{Var}(A_n) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^{n} Z_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^{n} Z_i\right) = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}(Z_i) = \frac{1}{n^2} \cdot n \cdot p(1-p) = \frac{p(1-p)}{n}$$

$$\Downarrow$$

$$\sigma_n^2 = \frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n} = \frac{\frac{\pi}{4} \cdot \frac{4-\pi}{4}}{n} = \frac{4\pi - \pi^2}{16n}$$

3.2. Minimalne n potrzebne do dokładności 0.01 z prawdopodobieństwem 0.99

Nierówność Czebyszewa:

$$\Pr(X \geq \varepsilon) \leq \frac{E[X]}{\varepsilon}$$

Z Treści zadania chcemy żeby $\Pr(|A_n - E[A_n]| \geq 0.01) \leq 0.01$

$$\Pr(|A_n - E[A_n]| \geq 0.01) = \Pr\left(|A_n - E[A_n]|^2 \geq 0.01^2\right) \leq \frac{\sigma_n^2}{0.01^2}$$

$$\Downarrow$$

$$0.01 = \frac{\sigma_n^2}{0.01^2} = \frac{4\pi - \pi^2}{16n \cdot 0.01^2}$$

$$\Downarrow$$

$$n = \left\lceil \frac{4\pi - \pi^2}{0.01^3 \cdot 16} \right\rceil \approx \lceil 168547.888329363395938 \rceil = 168548$$

3.3. To samo ale z nierównością Chernoffa

Nierówność Chernoffa:

$$X = \sum_{i=1}^{n} Z_i$$

$$\mu_n = E[X] = \frac{n\pi}{4}$$

$$\Pr(|X - \mu_n| > \varepsilon\mu) \leq 2\exp\left(-\frac{\varepsilon^2}{2+\varepsilon}\mu\right)$$

Z treści zadania chcemy żeby $\Pr(|A_n - E[A_n]| \geq 0.01) \leq 0.01$

$$\Pr(|A_n - E[A_n]| \geq 0.01) = \Pr\left(\frac{1}{n}\ |X - \mu_n| \geq 0.01\right) = \Pr(|X - \mu_n| \geq 0.01n) \leq 2\exp\left(-\frac{\varepsilon^2}{2+\varepsilon}\mu_n\right)$$

$$0.01n = \varepsilon\mu_n = \varepsilon\frac{n\pi}{4} \Rightarrow \varepsilon = \frac{0.04}{\pi}$$

$$\Downarrow$$

$$\frac{0.01}{2} = 0.005 = \exp\left(-\frac{\varepsilon^2}{2+\varepsilon}\mu_n\right) = \exp\left(-\frac{\left(\frac{0.04}{\pi}\right)^2}{2+\frac{0.04}{\pi}} \cdot \frac{n\pi}{4}\right) =$$

$$= \exp\left(-\frac{0.0016 \cdot n\pi}{\pi^2 \cdot \frac{2\pi+0.04}{\pi} \cdot 4}\right) = \exp\left(-\frac{0.0016n}{8\pi + 0.16}\right) = e^{-\frac{0.0016n}{8\pi+0.16}}$$

$$\Downarrow$$

$$-\ln(0.005) = \frac{0.0016n}{8\pi + 0.16} \Rightarrow (\pi + 0.02) \cdot \ln\left(0.005^{-1}\right) = 0.0002n$$

$$\Downarrow$$

$$n = \lceil(\pi + 0.02) \cdot \ln(200) \cdot 5000\rceil \approx \lceil 83755.6063123274 \rceil = 83756$$

4. Zadanie 15 - Standardowe odchylenia przy obliczaniu ćwiartek koła

4.1. Metoda losowania z $[0,1]^2$

Niech:

$$X_i = \begin{cases} 1 \text{ jeśli } x_i^2 + y_i^2 \leq 1 \\ 0 \text{ w przeciwnym wypadku} \end{cases}$$

Wtedy:

$$A_X = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\Downarrow$$

$$\sigma_X^2 = \mathrm{Var}(A_X) = \mathrm{Var}\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\ \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) =$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2} \cdot n \cdot \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right) = \frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n}$$

$$\Downarrow$$

$$\sigma_X = \sqrt{\frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n}} \approx \sqrt{\frac{0.1685}{n}} \approx \frac{0.4105}{\sqrt{n}}$$

4.2. Metoda losowania z $[0, 1]$

Niech:

$$Y_i = \sqrt{1 - x_i^2}$$

$$\Downarrow$$

$$E[Y_i] = \int_0^1 \sqrt{1 - x_i^2} dx = \frac{\pi}{4}$$

$$E[Y_i^2] = \int_0^1 1 - x_i^2 dx = x - \frac{x_i^3}{3}\ |_0^1 = 1 - \frac{1}{3} = \frac{2}{3}$$

$$\sigma_Y^2 = E[Y_i^2] - E[Y_i]^2 = \frac{2}{3} - \left(\frac{\pi}{4}\right)^2 \approx 0.0498$$

Wtedy:

$$A_Y = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

$$\Downarrow$$

$$\sigma_Y^2 = \mathrm{Var}(A_Y) = \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}(Y_i) = \frac{1}{n^2} \sum_{i=1}^{n} \frac{2}{3} - \left(\frac{\pi}{4}\right)^2 = \frac{\frac{2}{3} - \left(\frac{\pi}{4}\right)^2}{n}$$

$$\sigma_Y = \sqrt{\frac{\frac{2}{3} - \left(\frac{\pi}{4}\right)^2}{n}} \approx \sqrt{\frac{0.0498}{n}} \approx \frac{0.2232}{\sqrt{n}}$$

4.3. Wnioski

$$\sigma_X > \sigma_Y$$

Metoda losowania na jednej osi daje lepsze wyniki (o mniejszej wariancji) i zbiega do wartości oczekiwanej szybciej niż metoda losowania z dwóch osi

5. Zadanie 16

Niech:

$$X \in [0,1]$$

$$Y = \begin{cases} X \text{ jeśli } X \in \left[\frac{1}{2}, 1\right] \\ 1 - X \text{ jeśli } X \in \left[0, \frac{1}{2}\right] \end{cases} \Rightarrow Y \in \left[\frac{1}{2}, 1\right]$$

Dystrybuanta Y:

$$F_Y(y) = \Pr(Y \leq y) = \Pr(\max\{X, 1-X\} \leq y) = \Pr(X \in [1-y, y]) \text{ dla } y \in \left[\frac{1}{2}, 1\right]$$

Jako że X jest losowane z rozkładem jednostajnym to:

$$F_Y(y) = \frac{y - (1-y)}{1} = 2y - 1 \text{ dla } y \in \left[\frac{1}{2}, 1\right]$$

Gęstość prawdopodobieństwa:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} 2y - 1 = 2$$

$$\Downarrow$$

Y ma rozkład jednostajny $\Rightarrow E[Y] = 0.75$

$$E[Y^2] = \int_{\frac{1}{2}}^{1} y^2 \cdot f_Y(y) dy = \int_{\frac{1}{2}}^{1} 2y^2 dy = \left[2\frac{y^3}{3}\right]_{\frac{1}{2}}^{1} = \frac{2}{3} - 2 \cdot \frac{1}{8} \cdot \frac{1}{3} = \frac{8}{12} - \frac{1}{12} = \frac{7}{12}$$

$$\Downarrow$$

$$\sigma_Y^2 = E[Y^2] - E[Y]^2 = \frac{7}{12} - \left(\frac{3}{4}\right)^2 = \frac{28}{48} - \frac{27}{48} = \frac{1}{48}$$

6. Zadanie 17

X ma rozkład jednostajny na $[0,1]$ więc:

$$f_X(x) = 1$$

$$E[X] = \frac{1}{2}$$

$$E[X^2] = \int_0^2 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$

$$\sigma_X^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

$$\Downarrow$$

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(X, 1-X) = \mathrm{Cov}(X, -X) = -\mathrm{Cov}(X, X) = -\sigma_X^2 = -\frac{1}{12}$$

Algorytmika
Ćwiczenia 4

May 7, 2025

Adrian Herda
Informatyka Algorytmiczna
Politechnika Wrocławska

7. Zadanie 18

Lemat Schwartza-Zippela:

$$P \in R[x_1, \cdots, x_n]$$

To niezerowy wielomian o stopniu całkowitym $\deg(P) = d \geq 0$ nad domeną R. Jeśli S jest skończonym podzbiorem R to wybierając niezależnie i jednostajnie losowe $r_1, \cdots, r_n$ należące do S^n

$$\Pr[P(r_1, ...c, r_n) = 0] \leq \frac{d}{|S|}$$

7.1. Oszacowanie liczby zer

Niech $f \in k[x_1, \cdots, x_n]$ będzie wielomianem w którym każda zmienna występuję w stopniu co najwyżej d. Załóżmy, że $f \neq 0$, oraz $A \subseteq k$ i $|A| > d$. Chcemy oszacować:

$$|\{a \in A^n : f(a) = 0\}| \leq d \cdot |A|^{n-1}$$

Z Lematu Schwartza-Zippela dla $a \in A^n$:

$$\frac{d}{|A|} \geq \Pr[f(a) = 0] = \frac{|\{a \in A^n : f(a) = 0\}|}{|A^n|}$$

$$\text{mnożymy to przez } |A^n| = |A|^n$$

$$\Downarrow$$

$$|\{a \in A^n : f(a) = 0\}| \leq d \cdot |A|^{n-1}$$

7.2. Prawdopodobieństwo dla $|A| = 2d$

Z Lematu Schwartza-Zippela wynika, że dla $|A| = 2d$ mamy:

$$\Pr[f(\zeta) = 0] \leq \frac{d}{|A|} = \frac{d}{2d} = \frac{1}{2}$$

7.3. Dla Ciągu lsoowych punktów

$|A| = 2d$ oraz $\{\zeta_1, \cdots, \zeta_{40}\} \subseteq A$ to ciąg niezależnych oraz losowych punktów ze zbioru A^n.

$$\forall_{n \in [40]} \Pr[f(\zeta_n) = 0] \leq \frac{1}{2}$$

Jako że ζ są niezależne to:

$$\Pr[f(\zeta_1) = \cdots = f(\zeta_{40}) = 0] = \Pr[f(\zeta_1) = 0] \cdot .... \cdot \Pr[f(\zeta_{40}) = 0] \leq \left(\frac{1}{2}\right)^{40} < 10^{-12}$$

8. Zadanie 21

Rzucamy igłę długości l na płaszczyznę z pionowymi liniami równoległymi w odległościach $k = 1$.

Parametry:

- $t = 1$ - odległość między liniami
- $l \leq t$ - długość igły
- $x \sim U[0, \frac{t}{2}]$ - odległość środka igły od linii
- $\theta \sim U[0, \frac{\pi}{2}]$ - kąt nachylenia igły do poziomu

$$f_X(x) = \begin{cases} \frac{2}{t} : 0 \leq x \leq \frac{t}{2} \\ 0 : \text{gdziekolwiek indziej} \end{cases}$$

$$f_\Theta(\theta) = \begin{cases} \frac{2}{\pi} : 0 \leq \theta \leq \frac{\pi}{2} \\ 0 : \text{gdziekolwiek indziej} \end{cases}$$

$$\Downarrow$$

$$f_{X\Theta}(x, \theta) = f_x(x) \cdot f_\Theta(\theta) = \begin{cases} \frac{4}{\pi t} : 0 \leq x \leq \frac{t}{2}, 0 \leq \theta \leq \frac{\pi}{2} \\ 0 : \text{gdziekolwiek indziej} \end{cases}$$

Igła przecina linię, gdy:

$$x \leq \frac{l}{2} \cos\theta$$

Teraz liczymy prawdopodobieństwo:

$$\Pr[\text{Igła przecina linię}] = \Pr\left[x \leq \frac{l}{2}\cos\theta\right] = \int_{\theta=0}^{\frac{\pi}{2}} \int_{x=0}^{\frac{l}{2}\cos\theta} f_{X\Theta}(x, \theta) dx d\theta =$$

$$\int_{\theta=0}^{\frac{\pi}{2}} \int_{x=0}^{\frac{l}{2}\cos\theta} \frac{4}{\pi t} dx d\theta = \int_0^{\frac{\pi}{2}} \left[\frac{4x}{\pi t}\right]_0^{\frac{l}{2}\cos\theta} d\theta = \frac{4l}{2\pi t} \int_0^{\frac{\pi}{2}} \cos\theta d\theta = \frac{2l}{\pi t} \cdot [\sin\theta]_0^{\frac{\pi}{2}} = 2\frac{l}{\pi t}$$

Algorytmika
Ćwiczenia 5

Adrian Herda

2025-06-29

9. Zadanie 22

Zmienna losowa

$$L \sim \mathrm{Uni}\{0, \cdots, n-1\}$$

to znaczy że:

$$\Pr[L = k] = \frac{1}{n}, \text{ dla } 0 \leq k < n.$$

Zmienna losowa

$$X = \max\{L, n - L - 1\}$$

$$X = \begin{cases} L, & \text{jeśli } L > \frac{n-1}{2} \\ n - L - 1, & \text{jeśli } L \leq \frac{n-1}{2} \end{cases}$$

$$X = \begin{cases} k, & \text{jeśli } k > \frac{n-1}{2} \\ n - k - 1, & \text{jeśli } k \leq \frac{n-1}{2} \end{cases}$$

Dla $n = 2m$ czyli parzystego mamy $\frac{n-1}{2} = m - \frac{1}{2}$ wtedy:

$$\sum_{k=0}^{2m-1} \max\{k, 2m - k - 1\} = \sum_{k=0}^{m-1} (2m - k - 1) + \sum_{k=m}^{2m-1} k =$$

$$\frac{m(2m - 1 + m)}{2} + \frac{m(m + 2m - 1)}{2} = \frac{m(3m - 1) + m(3m - 1)}{2} = 3m^2 - m$$

A więc wartość oczekiwana:

$$E(X) = \frac{1}{n} \sum_{k=0}^{2m-1} \max\{k, 2m - k - 1\} = \frac{1}{2m} \cdot (3m^2 - m) = \frac{3m - 1}{2} = \frac{3}{4}n - \frac{1}{2}$$

Dla $n = 2m + 1$ czyli parzystego mamy $\frac{n-1}{2} = \frac{2m}{2} = m$ wtedy:

$$\sum_{k=0}^{2m} \max\{k, 2m - k\} = \sum_{k=0}^{m} (2m - k) + \sum_{k=m+1}^{2m} k =$$

$$\frac{(m+1)(2m + m)}{2} + \frac{m(m + 1 + 2m)}{2} = \frac{3m^2 + 3m + 3m^2 + m}{2} = \frac{6m^2 + 4m}{2} = 3m^2 + 2m$$

A więc wartość oczekiwana:

$$E(X) = \frac{1}{n} \sum_{k=0}^{2m} \max\{k, 2m - k\} = \frac{m(3m + 2)}{n} = \frac{(n-1) \cdot \left(3 \cdot \left(\frac{n-1}{2}\right) + 2\right)}{2n} =$$

$$\frac{(n-1) \cdot (3n + 1)}{4n} = \frac{3n^2 - 2n - 1}{4n}$$

Algorytmika
Ćwiczenia 6

May 28, 2025

Adrian Herda
Politechnika Wrocławska

10. Zadanie 27

Niech X będzie licznikiem morrisa. Licznik rozpoczyna z wartością 0 i dla każdej operacji inkremetowany jest z prawdopodobieństwem 2^{-X}. Estymator morrisa jest zdefiniowany jako:

$$\hat{n} = 2^{X_n} - 1$$
$$\hat{n} + 1 = 2^{X_n}$$

gdzie X_n jest licznikiem morrisa po n operacjach. Wartość oczekiwana tego estymatora jest równa:

$$E[\hat{n}] = n$$

Dowód przez indukcję:
1. Dla $n = 0$ mamy $E[\hat{n}] = 2^{X_0} - 1 = 2^0 - 1 = 0$.
2. Załóżmy, że dla dowolnego n mamy $E[\hat{n}] = n$.
3. Dla $n + 1$ mamy:

$$\begin{aligned}
E\left[\widehat{n+1}\right] &= E[2^{X_{n+1}} - 1] \\
&= \sum_{i=1}^{\infty} P(X_n = i) \cdot E[2^{X_{n+1}} \mid X_n = i] - 1 \\
&= \sum_{i=1}^{\infty} P(X_n = i) \cdot \left(\underbrace{\frac{1}{2^i} \cdot 2^{i+1}}_{\text{increment}} + \underbrace{\left(1 - \frac{1}{2^i}\right) \cdot 2^i}_{\text{no increment}} \right) - 1 \\
&= \sum_{i=1}^{\infty} P(X_n = i) \cdot (2^i + 1) - 1 \\
&= \sum_{i=1}^{\infty} P(X_n = i)2^i + \sum_{i=1}^{\infty} P(X_n = i) - 1 \\
&= E[2^{X_n}] + 1 - 1 \\
&= E[\hat{n} + 1] = n + 1
\end{aligned}$$

■

10.1. Wariancja

$$
\begin{aligned}
E[2^{2X_n}] &= \sum_{i=1}^{\infty} P(X_n = i) \cdot 2^{2i} \\
&= \sum_{i=1}^{\infty} \left(P(X_{n-1} = i-1) \cdot \frac{1}{2^{i-1}} + P(X_{n-1} = i) \cdot \left(1 - \frac{1}{2^{i-1}}\right)\right) \cdot 2^{2i} \\
&= \sum_{i=1}^{\infty} 2^{i+1} P(X_{n-1} = i-1) + \sum_{i=1}^{\infty} 2^{2i} P(X_{n-1} = i) - \sum_{i=1}^{\infty} 2^{i} P(X_{n-1} = i) \\
&= 4\sum_{i=1}^{\infty} 2^{i-1} P(X_{n-1} = i-1) + E[2^{2X_{n-1}}] + E[2^{X_{n-1}}] \\
&= 4E[2^{X_{n-1}}] + E[2^{2X_{n-1}}] + E[2^{X_{n-1}}] \\
&= 3E[2^{X_{n-1}}] + E[2^{2X_{n-1}}] \\
&= 3n + E[2^{2X_{n-1}}]
\end{aligned}
$$

Ta formuła jest rekurencyjna, więc możemy j rozwin i otrzymamy ciąg arytmetyczny:

$$
E[2^{2X_n}] = 1 + \sum_{i=1}^{n} 3i = 1 + \frac{3}{2}n(n+1)
$$

A więc:

$$
\begin{aligned}
\mathrm{Var}(\hat{n}) &= E\left[(2^{X_n} - 1)^2\right] - E[2^{X_n} - 1]^2 \\
&= E[2^{2X_n} - 2 \cdot 2^{X_n} + 1] - n^2 \\
&= E[2^{2X_n}] + 1 - 2E[2^{X_n}] - n^2 \\
&= 1 + \frac{3}{2}n(n+1) + 1 - (2E[2^{X_n}] - 2) - n^2 - 2 \\
&= \frac{3}{2}n^2 + \frac{3}{2}n - (2E[2^{X_n} - 1]) - n^2 \\
&= \frac{3}{2}n^2 + \frac{3}{2}n - 2n - n^2 \\
&= \frac{1}{2}n^2 - \frac{1}{2}n \\
&= \frac{1}{2}n(n-1)
\end{aligned}
$$

10.2. Błąd standardowy

Błąd standardowy estymatora morrisa jest równy pierwiastkowi z wariancji podzielonej przez liczbę operacji:

$$
\mathrm{SE}[\hat{n}] = \sqrt{\mathrm{Var}\left(\frac{\hat{n}}{n}\right)} = \sqrt{\frac{1}{2}\frac{n(n-1)}{n^2}} = \sqrt{\frac{n-1}{2n}}
$$

11. Zadanie 28

Niech $h(x)$ będzie funkcją hashującą, która dla każdego x zwrca wartość z tablicy haszującej o długości m.

$$P(\text{kolizja}) = P\Big(\exists_{i \neq j}\big(h(x_i) = h(x_j)\big)\Big)$$

11.1. Dowód wzoru

Załóżmy że wstawione zostało już n elementów. To znaczy że pozstało $m - n$ miejsc w tablicy haszującej. Prawdopodobieństwo kolizji przy wstawianiu następnego elementu $x_{\{n+1\}}$ jest równe:

$$P(\text{kolizja}) = \frac{n}{m}$$

Przy wstawainiu kolejnch n elementów prawdopodobieństwo kolizji liczone jesst w następujący sposób:

$$\left.\begin{array}{l} P(\text{brak kolizji}) = \prod_{i=0}^{n-1} \frac{m-i}{m} \approx \exp\left(-\frac{n(n-1)}{2m}\right) \\ \frac{1}{2} \leq P(\text{kolizja}) \implies \frac{1}{2} \geq P(\text{brak kolizji}) \end{array}\right\} \implies \frac{1}{2} \approx \exp\left(-\frac{n(n-1)}{2m}\right)$$

$$\begin{aligned} \ln(2) &\approx \frac{n(n-1)}{2m} \\ n(n-1) &\approx 2\ln(2)m \\ n^2 - n &\approx 2\ln(2)m \end{aligned}$$

Dla dostatecznie dużych n możemy przyjąć, że

$$n^2 - n \approx n^2$$

więc:

$$\begin{aligned} n^2 &\approx 2\ln(2)m \\ n &\approx \sqrt{2\ln(2)m} \end{aligned}$$

■

11.2. Wartość n dla $m = 2^{16}$

$$\begin{aligned} n &\approx \sqrt{2\ln(2)2^{16}} \\ &\approx \sqrt{2 * 0.693147 * 65536} \\ &\approx \sqrt{90852.18725} \\ &\approx 301.417 \end{aligned}$$

12. Zadanie 29

Niech $h_b(x)$ będzie funkcją skrótu zwracającą b-bitowe ciągi. Rozmiar przestrzeni obrazów funkcji $h_b(x)$ dla dowolnego b wynosi:

$$|\{y : h_b(x) = y\}| = 2^b = m$$

Z poprzedniego zadania mamy wzór przybliżający wartość n dla której prawdopodobieństwo kolizji wynosi conajmniej $\frac{1}{2}$:

$$n \approx \sqrt{2m\ln(2)} = \sqrt{2 \cdot 2^b \cdot \ln(2)} = 2^{\frac{b}{2}} \cdot \sqrt{2\ln(2)}$$

12.1. Dla $b = 64$

$$n \approx 2^{\frac{64}{2}} \cdot \sqrt{2\ln(2)} = 2^{32}\sqrt{2\ln(2)} \approx 5.05694 \cdot 10^9$$

12.2. Dla $b = 128$

$$n \approx 2^{\frac{128}{2}} \cdot \sqrt{2\ln(2)} = 2^{64}\sqrt{2\ln(2)} \approx 2.17194 \cdot 10^{19}$$

12.3. Dla $b = 256$

$$n \approx 2^{\frac{256}{2}} \cdot \sqrt{2\ln(2)} = 2^{128}\sqrt{2\ln(2)} \approx 4.00652 \cdot 10^{38}$$

Algorytmika
Ćwiczenia 7

June 10, 2025

Adrian Herda
Informatyka Algorytmiczna
Politechnika Wrocławska

13. Zadanie 33 - wyznacznik Vandermonda

13.1. Treść

Pokazać że

$$V(x_1, ..., x_n) = \prod_{i<j}(x_j - x_i)$$

13.2. Rozwiązanie

Macierz Vandermonda jest zdefiniowana jako:

$$\begin{pmatrix} 1 & x_0 & x_0^2 & ... & x_0^n \\ 1 & x_1 & x_1^2 & ... & x_1^n \\ 1 & x_2 & x_2^2 & ... & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & ... & x_n^n \end{pmatrix}$$

A jej wyznacznik to

$$V(x_1, ..., x_n) = \begin{vmatrix} 1 & x_0 & x_0^2 & ... & x_0^n \\ 1 & x_1 & x_1^2 & ... & x_1^n \\ 1 & x_2 & x_2^2 & ... & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & ... & x_n^n \end{vmatrix}$$

Główna idea dowodu: Jeśli do kolumny macierzy dodamy (lub od niej odejmiemy) inną kolumnę pomnożoną przez pewien skalar, to wyznacznik macierzy nie zmienia się.

A więc w każdej kolumnie oprócz pierwszej odejmujemy poprzednią pomnożoną przez x_0. To daje nam macierz:

$$V = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 1 & x_1 - x_0 & x_1(x_1 - x_0) & ... & x_1^{n-1}(x_1 - x_0) \\ 1 & x_2 - x_0 & x_2(x_2 - x_0) & ... & x_2^{n-1}(x_2 - x_0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_0 & x_n(x_n - x_0) & ... & x_n^{n-1}(x_n - x_0) \end{pmatrix}$$

Teraz wykonując rozwinięcie Laplace'a względem pierwszego wiersza otrzymujemy $\det(V) = \det(B)$ gdzie:

$$B = \begin{pmatrix} x_1 - x_0 & x_1(x_1 - x_0) & \dots & x_1^{n-1}(x_1 - x_0) \\ x_2 - x_0 & x_2(x_2 - x_0) & \dots & x_2^{n-1}(x_2 - x_0) \\ \vdots & \vdots & \ddots & \vdots \\ x_n - x_0 & x_n(x_n - x_0) & \dots & x_n^{n-1}(x_n - x_0) \end{pmatrix}$$

Jako że wszystkie wartości na i-tym wierszu mają współczynnik w postaci $x_{i+1} - x_0$ możemy je wyciągnąć przed macierz i otrzymać równość:

$$\begin{aligned} \det(V) &= (x_1 - x_0)(x_2 - x_0)\cdots(x_n - x_0) \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} \\ &= \prod_{1 \le i \le n} (x_i - x_0) \det(V') \end{aligned}$$

gdzie V' jest macierzą Vandermonda dla $x_1, \dots, x_n$. powtarzając ten proces na coraz mniejszych macierzach Vandermonda otrzymujemy produkt:

$$\begin{aligned} \det(V) &= \prod_{0 < i \le n} (x_i - x_0) \cdot \prod_{1 < i \le n} (x_i - x_1) \cdot \dots \cdot \prod_{n-1 < i \le n} (x_i - x_{n-1}) \\ &= \prod_{0 \le j < n} \left(\prod_{j < i \le n} (x_i - x_j) \right) \\ &= \prod_{0 \le j < i \le n} (x_i - x_j) && \text{Zamieniając notacje } i - j \\ &= \prod_{0 \le i < j \le n} (x_j - x_i) \end{aligned}$$

■

14. Zadanie 36 - maksymalne sparowanie dzięki specyfikacji

14.1. Treść

Należy udowodnić że jeśli w grafie $G = (V, E)$ zachodzą warunki specyfikacji S albo *single* to zbiór $M = \{(p, \mathrm{pref}_p) : \mathrm{pref}_p \neq \mathrm{NULL}\}$ jest sparowaniem maksymalnym.

$$S = (\forall_{p \in V})(\mathrm{married}(p) \lor \mathrm{single}(p))$$

14.2. Dowód

Dowód będzie nie wprost

Załóżmy przeciwnie, że zbiór M nie jest maksymalnym sparowaniem, to znaczy że istnieje M' taki że $M \subset M'$.

1. M jest poprawnym sparowaniem. Z definicji *married*:

$$\mathrm{married}(p) \equiv \mathrm{pref}_p = q \in N(p) \land \mathrm{pref}_q = p \in N(q)$$

Zatem jeśli $\text{pref}_q \neq \text{NULL}$ to żeby dodać (p, pref_p) do M, musimy mieć:
- $\text{pref}_p = q$
- $\text{pref}_q = p$

A więc jako że każdy wierzchołek może być albo *married* albo *single* to żaden wierzchołek nie może być częścią obu par.

2. Załóżmy że M nie jest maksymalne

Jako że M nie jest maksymalne to znaczy że istnieje para $(p, q) \in M' \setminus M$ którą można by dodać do M. Oznacza to, że:
- $\text{pref}_p = \text{NULL}$,
- $\text{pref}_q = \text{NULL}$,
- $p \in N(q) \wedge q \in N(p)$,

To oznacza że oba wierzchołki są *free* co zaprzecza warunkom specyfikacji.

Zatem M musi być sparowaniem maksymalnym.

■

15. Zadanie 37 - specyfikacja kończy jakąkolwiek pracę w algorytmie

15.1. Treść

Należy udowodnić że jeśli zachodzą warunki specyfikacji S to konfiguracja jest ostateczna (żaden krok algorytmu nie zmieni konfiguracji)

$$S = (\forall_{p \in V})(\text{married}(p) \vee \text{single}(p))$$

15.1.1. Algorytm

```
do forever
  if pref_p == NULL && (exists q in N(p))(pref_q == p)
    pref_p ← q
  end if
  if pref_p == NULL
    && (forall q in N(p))(pref_q != p)
    && (exists q in N(p))(pref_q == NULL)
    pref_p ← q
  end if
  if pref_p == q && pref_q != p && pref_q != NULL
    pref_p ← NULL
  end if
end do
```

15.2. Dowód

Jeśli zachodzą warunki specyfikacji S to znaczy że nie ma żadnych wierzchołków w stanach *free*, *wait* oraz *chain*.

1. Pierwsza klauzula `if` sprawdza czy istnieje wierzchołek p taki że $\text{pref}_p = \text{NULL}$ oraz istnieje wierzchołek $q \in N(p)$ taki że $\text{pref}_q = p$ a to znaczyłoby że wierzchołek q musi być w stanie *wait*. Jako że nie ma już takich wierzchołków dzięki warunkom specyfikacji, ta klauzula nie może zostać wykonana.
2. Druga klauzula `if` sprawdza istnienie wierzchołka p takiego że $\text{pref}_p = \text{NULL}$ oraz takiego wierzchołka $q \in N(p)$ że $\text{pref}_q = \text{NULL}$. Taka klauzula spełniona byłaby tylko gdyby $\text{free}(p) \wedge \text{free}(q)$. Jako że nie ma już takich wierzchołków dzięki warunkom specyfikacji, ta klauzula nie może zostać wykonana.
3. Trzecia klauzula `if` sprawdza istnienie wierzchołka p takiego że $\text{pref}_p = q$ oraz takiego wierzchołka $q \in N(p)$ że $\text{pref}_q \neq p \wedge \text{pref}_q \neq \text{NULL}$ to znaczy że $\text{pref}_q = r \wedge r \neq p$. To znaczyłoby że raka klauzula byłaby spełniona tylko dla wierzchołka $\text{chain}(p)$. Jako że nie ma już takich wierzchołków dzięki warunkom specyfikacji, ta klauzula nie może zostać wykonana.

Z punktów powyższych klauzul wynika że algorytm nie może wykonać żadnego kroku, a więc konfiguracja jest ostateczna.

■