אלגוריתמים 1: הרצאה 11 ־ חסמי צ'רנוף

20 בינואר 2026

גיא יער־און

בהרצאה זו נלמד כלי הסתברותי מתמטי, ובהמשך נדון כיצד לממש כלים אלו באלגוריתמים.

0.1 אי שוויון מרקוב וצ'בישב

נניח כי אנחנו מטילים מטבע כמות מסוימת של פעמים ונרצה להבטיח כי בסיכוי גבוה (של לפחות $1-\frac{1}{n^c}$) נבטיח שיצא עץ לפחות פעם אחת. הסיכוי שלא יצא עץ אף פעם הינו $\frac{1}{2^k}$. נראה כי אם $k=clog(n)$ אזי בהסתברות לפחות $1-\frac{1}{n^c}$ נקבל עץ.

מה אם נרצה $log(n)$ פעמים שיצא עץ בסיכוי גבוה פולינומית? נטיל $clog^2(n)$ פעמים, בכל $clog(n)$ הטלות הסיכוי לקבלת עץ הוא $1-\frac{1}{n^c}$. הסיכוי שקיימת סדרה של אחת מ$clog(n)$ ההטלות שבה לא יצא עץ $\leq \frac{log(n)}{n^c} < \frac{1}{n^{c-1}}$. נבחין כי במחשבה הזו מגיעים ל$\Omega(log^2(n))$ הטלות. צ'רנוף יראה שמספיקים $\Theta(logn)$ הטלות ־ זו החזקה של צ'רנוף שנרצה להשתמש בה.

אי שוויון מרקוב: אם X מ"מ אי שלילי, ויהי $k>0$ אזי $Pr[X>k]<\frac{\mathbb{E}[X]}{k}$.
נבחין כי התלות בk היא לינארית במכנה של ההסתברות. בפרט, $Pr[X>k\mu]<\frac{1}{k}$.

אי שוויון צ'בישב: אם X מ"מ (לא בהכרח אי שלילי), ונסמן $\mu=\mathbb{E}[X]$ וכן $\sigma^2=Var[X]=\mathbb{E}[(X-\mu)^2]$. וכן $t>0$ אז,

$$Pr[|X-\mu|\geq t]\leq\frac{\sigma^2}{t^2}$$

נבחין כי זה גורר $Pr[X>k\mu]=Pr[X-\mu>(k-1)\mu]<_* Pr[|X-\mu|\geq(k-1)\mu]\leq\frac{\sigma^2}{(k-1)^2\mu^2}$
שכן $*$ נכון כי בהוספת ערך מוחלט נוספו הסתברויות (אפשרויות לערכים) שלא היו קודם, מס' שהיו שליליים קודם כעת נכנסים.

כעת נבחין, כי הקשר בין $(k-1)$ ל$(k-1)^2$ במכנה הוא כעת ריבועי. יותר טוב ממרקוב. (כל עוד μ,σ מתנהגים "יפה") ־ אם לא יתנהגו יפה כלומר לא יהיו קרובים יחסית זה לזה, יתכן שעדיף להשתמש במרקוב.

ההוכחה של צ'בישב:

$$Pr[|X-\mu|>t]=Pr[(X-\mu)^2>t^2]\leq\frac{\mathbb{E}[(X-\mu)^2]}{t^2}=\frac{Var[X]}{t^2}=\frac{\sigma^2}{t^2}$$

כעת נבחין כי לא חייבים להעלות רק בריבוע, אלא כל העלאה של פונקציה **מונוטונית** שתהפוך את המשתנה לאי שלילי בהכרח (לכן העלאה בשלישית לא הייתה עובדת).

טענה: $2\ell - moment$: יהי X מ"מ כך ש$\mathbb{E}[X] = \mu$. אזי לכל $t > 0$ ולכל $\ell \geq 1$ שלם מתקיים:

$$Pr[|X-\mu| > t] < \frac{\mathbb{E}[(X-\mu)^{2\ell}]}{t^{2\ell}}$$

הוכחה:

$$Pr[|X-\mu| > t] = Pr[(X-\mu)^{2\ell} > t^{2\ell}] \leq \frac{\mathbb{E}[(X-\mu)^{2\ell}]}{t^{2\ell}}$$

נניח כי הטלנו מטבע $8log(n)$ פעמים. מה הסיכוי שיצאו לפחות $log(n)$ פעמים עץ?
נסמן בX את מס' הפעמים שיצא עץ.

לפי אי שוויון מרקוב: נעזר במאורע המשלים (יהיו $8log(n) - X$ הטלות של פלי, נראה כי הסיכוי לכך שיצאו לכל היותר $logn$ הטלות של עץ (המשלים) שווה לסיכוי שיצאו לפחות $7log(n)$ פעמים פלי)

$$Pr[8log(n) - X > 7logn] < \frac{\mathbb{E}[8log(n) - X]}{7log(n)} = \frac{8log(n) - \mathbb{E}[X]}{7log(n)} =$$

נבחין $\mathbb{E}[X] = \frac{1}{2} \times 8log(n) = 4log(n)$ [חצי מהפעמים יצא עץ כי מטבע מאוזן].

$$= \frac{8log(n) - 4log(n)}{7log(n)} = \frac{4}{7}$$

לפי אי שוויון צ'בישב:

$$Pr[8log(n) - X > 7logn] = Pr[(8logn - X) - 4log(n) > 3log(n)]$$

$$\leq Pr[|4log(n) - X| > 3log(n)] = Pr[|X - 4log(n)| > 3log(n)] \leq_* \frac{2log(n)}{(3logn)^2} = \frac{2}{9log(n)}$$

שכן $*$ נכון כי נבחין שהשונות הינה $8logn = 2log(n) \times \frac{1}{2} \times \frac{1}{2}$. נבחין כי השתפרנו כעת, עוד לא פולינומי אך השתפרנו.

לפי $2\ell - moment$ נקבל:

$$Pr[8logn - X > 7log(n)] \leq Pr[|4log(n) - X| > 3log(n)] <$$

$$< \frac{\mathbb{E}[(4log(n) - X)^{2\ell}]}{(3logn)^{2\ell}}$$

מבלי לפתוח את המתמטיקה (**לא פתחנו זאת בהרצאה**), נבחין כי נקבל כאן לבסוף לכל ℓ משהו מהצורה של $\frac{1}{(logn)^{\ell}}$ (בתוספת קבועים). כלומר:

$$\frac{\mathbb{E}[(4log(n) - X)^{2\ell}]}{(3logn)^{2\ell}} \approx \frac{1}{c \times log^{\ell}(n)}$$

נבחר: $\ell = log_{log(n)}(n) = \frac{log(n)}{log(log(n))}$ ונקבל סיכוי גבוה־ זה די מכוער. המטרה תהיה להשתמש בצ'רנוף כמובן.

0.2 חסמי צ'רנוף: הגדרה

הנחות יסוד: נניח כי X הוא סכום של משתנים מקריים אינדיקטורים (כל אחד מהם יכול להתפלג שונה!) **בלתי תלויים** (מאוד חשוב!).
נסמן:

$$X = \sum_i x_i$$

הרעיון מבוסס על השיטה שבה עבדנו קודם. השתמשנו בהוכחות קודם בפונקציה מונוטונית אי שלילית. איזו פונקציה מונוטונית אי שלילית אנחנו מכירים? $f(x) = e^x$. בואו ננסה ־

$$Pr[X \geq a] =_{\forall t>0} Pr[e^{tX} \geq e^{ta}] \leq_{markov} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$$

נרצה לחשב את $\mathbb{E}[e^{tX}]$.

$$\mathbb{E}[e^{tX}] = \mathbb{E}[e^{t\sum x_i}] = \mathbb{E}[\prod e^{tx_i}] =_* \prod \mathbb{E}[e^{tx_i}]$$

המעבר ה$*$ חוקי כיוון שהמשתנים המקריים ב"ת ולכן גם e^{tx_i}.

אי שוויון צ'רנוף: עבור $0 < \delta < 1$ מתקיים: $Pr[X > (1+\delta)\mu] \leq (\frac{e^{\delta}}{(1+\delta)^{1+\delta}})^{\mu}$.
כמובן שזה נראה רע ומפחיד, ולכן ישנן שתי גרסאות פשוטות יותר:
.1

$$Pr[X > (1+\delta)\mu] < e^{\frac{-\mu\delta^2}{3}}$$

.2

$$Pr[X < (1-\delta)\mu] < e^{\frac{-\mu\delta^2}{2}}$$

0.3 הוכחת אי שוויון צ'רנוף

הערה. צבי אמר שאנו לא נדרשים לדעת את ההוכחה בע"פ, אך העקרונות שלה כן חשובים ולכן היא גם מוצגת כאן.

באשר $a = (1+\delta)\mu$ (ממה שפיתחנו קודם) נקבל:

$$Pr[X \geq (1+\delta)\mu] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}}$$

נרצה לחשב את $\mathbb{E}[e^{tX}] = \prod \mathbb{E}[e^{tx_i}]$. כלומר נחשב את $\mathbb{E}[e^{tx_i}]$.
נסמן: (נבחין הפונקציות לא בהכרח שוות הסתברות)

$$X_i = \left\{ \begin{array}{cc} 1 & p_i \\ 0 & 1-p_i \end{array} \right\}$$

$$\mathbb{E}[e^{tx_i}] = p_i e^t + (1-p_i)e^0 = 1 + p_i(e^t - 1) \leq_* e^{p_i(e^t-1)}$$

כאן $*$ נעזרנו באי השוויון $1 + x \leq e^x$.
לכן:

$$\mathbb{E}[e^{tX}] = \prod \mathbb{E}[e^{tx_i}] \leq \prod e^{p_i(e^t-1)} = e^{(e^t-1)\times\sum p_i} =_{**} e^{(e^t-1)\times\mu}$$

$**$ שכן $\mathbb{E}[x_i] = p_i$. ומלינאריות התוחלת: $\mathbb{E}[X] = \sum \mathbb{E}[x_i] = \sum p_i$.
כעת נחזור לאי השוויון מעלה:

$$Pr[X \geq (1+\delta)\mu] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} \leq \frac{e^{(e^t-1)\times\mu}}{e^{t(1+\delta)\mu}} = (e^{(e^t-1)-t(1+\delta)})^\mu$$

נבחר t על מנת להקטין את ההסתברות. נגדיר: $\phi(t) = (e^t - 1) - t(1+\delta)$

$$\phi'(t) = e^t - (1+\delta) = 0 \Longrightarrow t = ln(1+\delta)$$

(זה אכן מינימום ע"י גזירה נוספת). כעת: $t = ln(1+\delta)$

$$e^{\mu(e^{ln(1+\delta)}-1-ln(1+\delta)(1+\delta))} =_*$$

$$(\frac{e^{1+\delta-1}}{e^{(1+\delta)ln(1+\delta)}})^\mu = (\frac{e^\delta}{(1+\delta)^{1+\delta}})^\mu$$

$*$ שכן $e^{ln(1+\delta)} = 1+\delta$. כנדרש.

כעת נחזור לדוגמה מקודם. $Y = 8logn - X$. כמובן $\mathbb{E}[Y] = 4log(n)$ כפי שראינו. נרצה לחשב חסם ל:

$$Pr[Y > 7log(n)] = Pr[Y > (1+\frac{3}{4}\mathbb{E}[Y]] < e^{\frac{-\mu\delta^2}{3}} = e^{\frac{-4log(n)(\frac{3}{4})^2}{3}} = e^{-\frac{3}{4}log(n)} = \frac{1}{n^\alpha}$$

עבור $\alpha > 1$ קבוע.

קבועים: נתבונן כי לבסוף עברנו אל $\frac{1}{n^\alpha}$ אך עלינו לוודא זאת (!) שאכן $\alpha > 1$ [אם זה לא היה עובד עם $8log(n)$ ננסה $10log(n)$ ואם לא מספיק אז ננסה $20log(n)$ זה עדיין O של]. במבחן: עלינו בצד לוודא זאת שאכן $\alpha > 1$.

מסקנה: עבור $log(n)$ הטלות עץ מספיק להטיל $8log(n)$ פעמים.

מסקנה: בעוד מרקוב נותן תלות לינארית, צ'בישב ריבועית, צ'רנוף נותן תלות לוגריתמית (ובהסתברות קבועה של לפחות $1 - \frac{1}{n^\alpha}$). כעת נקבל עם צ'רנוף:

$$Pr[X > (k-1)\mu] \le (\frac{e^{k-1}}{(k-1)^{k-1}})^\mu$$

שהרי זו כבר תלות בחזקת $k-1$.

0.4 מיון מהיר $(Quick\ Sort)$

ניתחנו בהרצאה 10 את תוחלת מס' ההשוואות וראינו כי תוחלת מס' ההשוואות = זמן הריצה היא $O(nlogn)$. כעת, ננתח את החסם עליון על זמן הריצה בסיכוי גבוה.

נתבונן בעץ הרקורסיה של המיון . בתחילה בידינו n איברים שנכנסו אל עץ הרקורסיה. כל קודקוד, בחר מסלול משלו.
נתבונן בעלה בעץ הרקורסיה. נסמנו $a \in A$.
מה הסיכוי שהעומק d_a של a הוא "גדול"? (גדול ממש מ $10ln(n)$). נוכיח בשלב ראשון שהסיכוי לכך קטן פולינומית.

איטרציה טובה היא איטרציה שבה כל קריאה רקורסיבית היא מגודל לכל היותר $\frac{3}{4}$ מגודל הקלט המקומי. הסיכוי לאיטרציה טובה הוא בדיוק $\frac{1}{2}$.
כמה איטרציות טובות ישנן? נבחין כי אם מס' האיטרציות הטובות $< log_{\frac{4}{3}}(n)$ אזי בהכרח נשארו 0 איברים.

מה ההסתברות שב$10ln(n)$ איטרציות שבהן a משתתף, לא היו $log_{\frac{4}{3}}(n)$ איטרציות טובות? [שכן אם היו לפחות $log_{\frac{4}{3}}(n)$ אזי סיימנו. לכן אנחנו מחשבים את המאורע המשלים].
כמובן שזה צועק: צ'רנוף.

תוחלת מס' האיטרציות הטובות הינו $\frac{1}{2} \times 10ln(n) = 5ln(n)$. נחשב את הסיכוי שאנחנו רחוקים מהתוחלת. נסמן X כמשתנה מקרי של מס' האיטרציות הטובות (שהוא סכום של משתני אינדיקטור ב"ת שכן כל איטרציה ב"ת באיטרציה אחרת) עבור עלה כלשהו, נרצה לחשב את:

$$Pr[X < log_{\frac{4}{3}}(n)]$$

כלומר: ההסתברות כי מס' האיטרציות הטובות היה קטן מערך זה. נרצה לחלץ את δ. כיצד?
נזכר בחוק הלגוריתמים לשינוי בסיס: $log_b(a) = \frac{ln(a)}{ln(b)}$:

$$log_{\frac{4}{3}}(n) = (1-\delta) \times 5ln(n) \Longrightarrow 1-\delta = \frac{log_{\frac{4}{3}}(n)}{5ln(n)} = \frac{\frac{ln(n)}{ln(\frac{4}{3})}}{5ln(n)} = \frac{1}{5ln(\frac{4}{3})}$$

$$\Longrightarrow \delta = 1 - \frac{1}{5ln(\frac{4}{3})} \approx 0.3$$

נבחין כי אכן $0 < \delta < 1$ וזה מתאים לתנאי צ'רנוף. לכן:

$$Pr[X < log_{\frac{4}{3}}(n)] = Pr[X < (1-\delta)\mu] \leq e^{\frac{-\mu\delta^2}{2}}$$

$$= e^{-\frac{5ln(n)\delta^2}{2}} = e^{ln(n^{\frac{-5\delta^2}{2}})} = n^{\frac{-5\delta^2}{2}} = \frac{1}{n^{\frac{5\delta^2}{2}}} = \frac{1}{n^{\alpha}}$$

עבור $\alpha > 1$ (הערה, בדוגמה שלקחנו זה לא קורה. היה צריך לבחור $20ln(n)$ למשל. בכל מקרה מדובר בO של).

מכאן, שבסיכוי גבוה של לפחות $1 - \frac{1}{n^{\alpha}}$ העומק של האיבר הנ"ל בעץ הרקורסיה הוא לוגריתמי (יהיו $log_{\frac{4}{3}}(n)$ איטרציות טובות ונסיים איתו).

כעת נשאל מהו הסיכוי שקיים עלה עמוק כלשהו, מה הוכחנו עבור איבר ספציפי? הראינו כי עבור $a \in A$ כלשהו מתקיים:

$$Pr[depth(a) > 10ln(n)] \leq \frac{1}{n^{\alpha}}$$

כלומר: הסתברותו להתקע בעומק גדול מדי די קטנה. מכאן עבור כל האיברים ־ יש לכל היותר n איברים (עלים פוטנציאלים). נרצה לדעת מה ההסתברות שקיים לפחות איבר אחד שהענף שלו ארוך מדי. כלומר:

$$Pr[\exists a : depth(a) > 10ln(n)] \leq \sum_{i=1}^{n} Pr[depth(a_i) > 10ln(n)] \leq \frac{n}{n^{\alpha}} = \frac{1}{n^{\alpha-1}}$$

ולכן,

$$Pr[\forall a : depth(a) < 10ln(n)] = \overline{Pr[\exists a : depth(a) > 10ln(n)]} \geq 1 - \frac{1}{n^{\alpha-1}}$$

ישנה תלות בין האיטרציות. לכן בשלב זה השתמשנו בחסם האיחוד ולא בצ'רנוף. כלומר: ההסתברות שכל העץ יהיה בעומק לוגריתמי היא לפחות $1 - \frac{1}{n^{\alpha-1}}$ עבור $\alpha > 1$ קבוע כלשהו. כאשר עומק העץ הוא $O(logn)$, בכל רמה מתבצעת לכל היותר $O(n)$ עבודה - השוואות (בעת $partiton$) אזי זמן הריצה הכולל של האלגוריתם (בהסתברות גבוהה) יהיה:

$$O(n) \times O(logn) = O(nlogn)$$

מה הוכחנו כאן? זמן הריצה בוורסט קייס - הוא $O(nlogn)$ - בסיכוי גבוה.