

תרגול מס' 13 – אלגוריתמים הסתברותיים

עד עתה דיברנו רק על אלגוריתמים דטרמיניסטיים. בהינתן קלט קבוע לאלגוריתם דטרמיניסטי, כל הרצה שלו תחזיר את אותו הפלט. נשים לב כי גם באלגוריתמים דטרמיניסטיים, בהם יש בחירה שרירותית כחלק מהאלגוריתם (לדוגמא קרוסקל שבוחר שרירותית בין צלעות קלות ביותר), הבחירה אינה הסתברותית. הוכחת נכונות של אלגוריתם דטרמיניסטי דורשת שעבור כל קלט האלגוריתם יתן תוצאה נכונה. אלגוריתם הסתברותי הינו אלגוריתם שבו חלק הפעולות הינן הגרלת ערכים למשתנים (לפעמים נאמר שהאלגוריתם "מטיל מטבעות"), ופעולת האלגוריתם תלויה בתוצאת ההגרלה. על כן, ייתכן ששתי הרצות של אלגוריתם הסתברותי על אותו קלט יסתיימו בפלט שונה. כשנדבר על נכונות אלגוריתם הסתברותי, נתייחס להסתברות שהאלגוריתם יחזיר תשובה נכונה על קלט כלשהו. היתרון של האלגוריתמים ההסתברותיים שנראה הוא שהם יעילים יותר מהאלגוריתם הדטרמיניסטי לפתרון אותה הבעיה.

בעיית Matrix Identification:

מופיע: 3 מטריצות מסדר $n \times n$, A, B, C , מעל השדה $\{0,1\}$,

צ"ל: האם $AB \neq C$?

שימו לב! הבעיה לעיל הינה בעיית הכרעה.

פתרון טריוויאלי: לבצע הכפלת מטריצות ולהשוות. זמן ריצה: $O(n^3)$.

היות ואנו מעוניינים בהכרעה בלבד - האם $AB \neq C$, נראה כי העלות לעיל היא יותר מהנדרש. נראה כי ניתן לשפר את זמן הריצה באופן משמעותי אם נסכים לטעות חד-צדדית זניחה.

מה הכוונה טעות חד צדדית כאן?

אם $AB = C$, תמיד נענה "שווים".

אם $AB \neq C$ נענה "שוניים" בהסתברות גבוהה מאד.

הערה: יתכן ש- $AB \neq C$ אבל נשיב "שווים" (כלומר נחשוב ש $AB=C$), אך בהסתברות נמוכה מאד (ולכן נטעה בהסתברות זניחה).

במה תלוי זמן הריצה של האלגוריתם שנייצר?

- בגודל הקלט (כמובן).

- במידת הטעות לה אנו שואפים.

במקרה שלנו נקבל זמן ריצה של $O(k * n^2)$ עבור אלגוריתם בעל טעות חד צדדית בהסתברות $1/(2^k)$. (עבור $k = 70$ זה יוצא טעות נדירה מאד!!)

שלבי התכנון:

בשלב הראשון נייצר אלגוריתם הרץ ב- $O(n^2)$, וטועה (חד – צדדית) בהסתברות $1/2$.

בשלב השני נתאר אלגוריתם הטועה בהסתברות $1/(2^k)$.

שלב ראשון: אלגוריתם אקראי עם הסתברות לטעות חז-כיוונית 1/2

הסבר כללי:

האלגוריתם מבוסס על העובדה הבאה:

אם $AB=C$, אזי לכל וקטור v באורך n מתקיים $AB\vec{v} = C\vec{v}$.

אם נמצא וקטור שעבורו $AB\vec{v} \neq C\vec{v}$ נוכל להיות בטוחים כי $AB \neq C$. לכן אנו נגריל וקטור, ואז נבדוק האם הוא מקיים את השוויון. בהמשך נראה כי קיימת שיטה יעילה לבדיקת השוויון וכי כמות הוקטורים שמובילים לתוצאה נכונה היא גדולה (כלומר ההסתברות להגריל וקטור כזה היא גבוהה).

עלינו לברר האם $AB \neq C$.

לשם כך די שנבדוק האם המטריצה $D = AB - C$ היא מטריצת ה-0 או שאינה כזו. מטריצת ה-0 מקיימת את התכונה הבאה:

אם $D=0$, אזי לכל וקטור v באורך n , מתקיים $D\vec{v} = \vec{0}$.

האלגוריתם:

קלט: מטריצות מסדר $n \times n$, A, B, C ו-1

פלט: אם $AB=C$: יחזור "שווים" בהסתברות 1.

אם $AB \neq C$: יחזור "שונים" בהסתברות $\leq 1/2$.

1. הגרל (בהסתברות אחידה) וקטור בוליאני $\vec{v} \in \{0,1\}^n$.
2. אם $AB\vec{v} \neq C\vec{v}$ החזר "שונים", אחרת החזר "שווים".

ניתוח זמן:

1. $O(n)$ (נהוג להניח שהגרלת ביט דורשת $O(1)$)

2. נחלק לפעולות:

a. $O(n^2) \leftarrow C\vec{v}$ - הכפלת $(n \times 1) \cdot (n \times n)$

b. $O(n^2) \leftarrow B\vec{v}$ - הכפלת $(n \times 1) \cdot (n \times n)$

c. $O(n^2) \leftarrow A(B\vec{v})$ - הכפלת $(n \times 1) \cdot (n \times n)$

d. $O(n) \leftarrow A(B\vec{v}) - C\vec{v}$ - חיסור שני וקטורים

3. $O(n)$ - השוואה בין שני ווקטורים.

סה"כ - $O(n^2)$ כנדרש

ניתוח ההסתברות של MatrixID

תזכורת: $D = AB - C$

אם $AB = C$, אזי $D = 0$, ולכן לכל וקטור \vec{v} : $D\vec{v} = \vec{0}$, אזי תמיד נענה "שווים".

אם $AB \neq C$, אזי D היא איננה מטריצת ה-0. נראה שבמקרה זה בהסתברות גדולה שווה לחצי נענה "שונים".

הניתוח הבא יראה כי הסיכוי לטעות במקרה בו $AB \neq C$ הוא $\geq \frac{1}{2}$:

הגדרות: עבור $D \neq \vec{0}$

וקטור "שקרן" הוא וקטור בוליאני " $\vec{v} \in \{0,1\}$ " המקיים $D\vec{v} = \vec{0}$. המשמעות היא שניסינו להגריל וקטור שבאמצעותו נוכל להוכיח כי $AB \neq C$ אך יצא לנו וקטור שעבורו $AB\vec{v} = C\vec{v}$.
 וקטור "הגון" הוא וקטור " $\vec{v} \in \{0,1\}$ " המקיים $D\vec{v} \neq \vec{0}$ (הוא למעשה **עד** לכך ש- $AB \neq C$).

אבחנה:

אם $AB \neq C$, אזי ההסתברות לטעות שקולה להסתברות שנגריל וקטור שקרן.

אבחנה:

באופן דומה, הסתברות לא לטעות שווה להסתברות להגריל וקטור "הגון".

טענה:

תהינה A, B, C מטריצות $n \times n$ בוליאניות כך ש $AB \neq C$. אזי עבור v ווקטור שונה מ-0 הנבחר אקראית בצורה אחידה מ " $\{0,1\}^n$ " מתקיים כי: $\Pr[D\vec{v} = \vec{0}] \leq \frac{1}{2}$ (זו ההסתברות לטעות)

הוכחה:

נראה כי אם $D = (AB - C)$ היא אינה מטריצת ה-0 אזי מספר הוקטורים השקרנים קטן שווה למספר הוקטורים ה"הגונים", ולכן צפיפות הוקטורים הרצויים ("הגונים") היא גבוהה מאוד (לפחות 50%). היות וידוע כי $AB \neq C$ אזי קיים איבר ב D אשר שונה מ 0. נניח כי $D_{ij} \neq 0$.
 תהי \mathbf{d} השורה ב D_{ij} מופיע. נשים לב כי אם $dv \neq 0$ אזי $Dv \neq 0$. נראה אם כך, כי כמות הווקטורים המקיימים תכונה זו הינה 2^{n-1} . כעת, נבחר ווקטור $v = (v_1, v_2, \dots, v_n)$.

$$(1) \quad d_{i1}v_1 + \dots + d_{ij}v_j + \dots + d_{in}v_n = 0$$

נשים לב כי אם נהפוך את הערך של v_i אז נקבל וקטור חדש שהוא "הגון", כלומר על כל וקטור "שקרן" יש לנו וקטור "הגון", כמו כן פעולת ההפיכה הזאת היא פעולה חח"ע ועל כן כמות הווקטורים "ההגונים" גדולה שווה לכמות הווקטורים "השקרנים".

Number of lying vectors ≤ Number of fair vectors

סה"כ:

$$\begin{aligned} & \Pr[\text{The algorithm fails}] \\ &= \frac{\text{Number of lying vectors}}{\text{Number of lying vectors} + \text{Number of fair vectors}} \\ &\leq \frac{\text{Number of lying vectors}}{2 * \text{Number of lying vectors}} = \frac{1}{2} \end{aligned}$$

מש"ל.

שלב שני: שיפור ההסתברות

הוכחנו כי:

1. אם $AB=C$ נזהה זאת בהסתברות 1.
2. אחרת, אם $AB \neq C$ נזהה זאת בהסתברות לפחות $1/2$.

$1/2$ איננה הסתברות שמספקת אותנו. כיצד ניתן לשפרה?

זרד 1:

למשל נריץ 3 פעמים את האלגוריתם באופן בת"ל.
אם בכל 3 הריצות נקבל false נענה false. (במשמעות של "כנראה לא שונים")
אחרת נענה true (במשמעות של "שונה בוודאות" – כי מצאנו עד לשונות).

ומה הסיכוי לטעות?

1. אם $AB = C$, כל הריצות יענו "לא שונים" ולכן נענה "לא שונים" ונצדק.
2. אם $AB \neq C$, ההסתברות שכל הריצות יטעו היא לכל היותר $\frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{8}$. אנו נטעה (ונגיד "לא שונים") רק אם כל הריצות יטעו.

במקרה כללי של k ריצות:

אם בכל k הריצות נקבל "שוים" נענה שווים.
אחרת נענה "שונים".

ומה הסיכוי לטעות?

1. אם $AB = C$, כל הרצות יחזירו "שוים" ולכן נענה "שוים" בצדק.
2. אם $AB \neq C$, ההסתברות שבכל הריצות נקבל "שוים" לכל היותר $\frac{1}{2^k}$ ולכן נטעה רק

בהסתברות לכל היותר $\frac{1}{2^k}$.

זרד 2:

נגריל את הווקטורים מקבוצה גדולה יותר $\{0,1,\dots,j\}$ וכעת ההסתברות לטעות תקטן ל $1/j$. אך כעת ייצוג המספרים ידרוש $\log j$ ביטים ופעולות המכפלה ידרשו יותר זמן.

דוגמא:

נניח שנתונות המטריצות הבאות:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad C = \begin{pmatrix} 3 & 6 & 99 \\ 3 & 6 & 99 \\ 3 & 6 & 99 \end{pmatrix}$$

ונניח שהוגרל הווקטור (השקרון) הבא:

$$v = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

נשים לב ש:

$$AB = \begin{pmatrix} 6 & 12 & 18 \\ 6 & 12 & 18 \\ 6 & 12 & 18 \end{pmatrix}$$

כשנצבע את ההכפלה $A(Bv)$ נקבל:

$$A(Bv) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} * \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} * \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$A(Bv) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} * \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$$

$$A(Bv) = \begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix}$$

מצד שני כשנצבע את ההכפלה Cv נקבל:

$$Cv = \begin{pmatrix} 3 & 6 & 99 \\ 3 & 6 & 99 \\ 3 & 6 & 99 \end{pmatrix} * \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix}$$

ולכן נסיק בטעות ש- $AB=C$, אך מנגד קיים $v' = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ הוא לא וקטור שקרון.