

# שימושי אלגברה לינארית במדעי המחשב

## מבחן בית

בבחינה זו נתמקד בעיות חישוביות המשלבות גרפים, והילוכים מקריים. הבחינה תשלב שאלות תיאורטיות ושאלות מעשיות אותן תצטרכו לפתור תוך כדי שימוש בתכנית מחשב שאותה תממשו.

לאורך כל סעיפי הבחינה לכל גרף  $G = (V, E)$ , נסמן:  
 $V = \{1, \dots, n\}$  ו-  $E \subseteq V \times V$   
כלומר, תיתכנה קשתות עצמיות וקשת תיוצג על ידי זוג סדור  $(i, j) \in V \times V$   
(אם הגרף אינו מכוון אזי מתקיים:  $(i, j) \in E \iff (j, i) \in E$ ).

בבחינה נתמקד באלגוריתם ה-  $PageRank$  המתואר להלן בקצרה.

קלט: גרף מכוון  $G$ .

פלט: וקטור הסתברות  $\vec{v} \in \mathbb{R}^n$  המייצג דירוג של קודקודי הגרף.

פרמטרים נוספים:  $N, t, \epsilon$ .

מבנה נתונים: מערך  $d$  בגודל  $n$  המאותחל לאפסים.

### תיאור האלגוריתם:

האלגוריתם יריץ  $t$  איטרציות, כאשר בכל איטרציה נבצע:

נבצע הילוך מקרי באורך  $N$  צעדים, כאשר בכל צעד בהילוך:  
בסיכוי  $1 - \epsilon$  נלך לשכן מקרי של הקודקוד בו אנחנו מבקרים (אם אין לו שכנים נלך לקודקוד מקרי בגרף)  
ובסיכוי  $\epsilon$  נלך לקודקוד מקרי.  
יהי  $i$  הקודקוד אליו הגענו לאחר  $N$  צעדים, נעלה את ערכו של  $d[i]$  ב- 1.  
בתום ריצת האלגוריתם, נחלק את ערכי  $d$  ב-  $t$  ונחזיר וקטור שערכיו הם ערכי המערך  $d$ .

שאלה 1: הראו שהוקטור המתקבל אכן וקטור הסתברות.

בבחינה, נבחן את פלט האלגוריתם על קלטים שונים ונבחן את השפעתם של הפרמטרים  $N, t, \epsilon$  על קלטים אלו.

שאלה 2: הסבירו מדוע אם נבחר  $1/\epsilon \ll N$ , הפרמטר  $\epsilon$  חסר חשיבות מעשית.

שאלה 3: נסתכל על איטרציה בודדת (הילוך מקרי באורך  $N$  צעדים).  
מה תוחלת מספר הצעדים בהם נבחר ללכת לקודקוד מקרי בגרף (ולא לשכן מקרי)?

שאלה 4: הסבירו מדוע התפלגויות הקודקדים לאחר הילוך בגודל  $1/\epsilon$  צעדים ולאחר  $2/\epsilon$  צעדים תהינה דומות (הדרכה: מה קורה לאחר שהלכנו לקודקוד מקרי ולא לשכן מקרי).

לאור מה שהוכחתם בסעיפים האחרונים, אם לא צוין אחרת, נגדיר  $N = 2/\epsilon$  (ובכך נקטין את מספר הפרמטרים בדיון).

### משפחה ראשונה בדיון – גרפים מקריים:

בחלק זה נסתכל על גרפים המוגדרים מתוך ההתפלגות הבאה.

נקבע  $0 < p < 1$  ולכל זוג קודקודים:  $(i, j)$  בגרף, נוסיף את הקשת  $(i, j)$  בסיכוי  $p$ .

#### שאלה 6:

עבור גרפים המוגרלים באופן שהוגדר לעיל, כיצד אתם מצפים שתיראה התפלגות הקודקודים בהילוך? האם ישנה תלות בפרמטרים  $p, \epsilon$ ?

#### שאלה 7:

הגרילו גרף כנ"ל המכיל  $2^{10}$  קודקודים, והגדירו את הפרמטר  $p = 1/2^6$ .  
מה יצאה הדרגה הממוצעת של קודקוד בגרף שהגרלתם?

ענת ננסה להבין כיצד יש לבחור את הפרמטרים  $N, t$  עבור הגרף שהוגרל.  
היות וזהו גרף מקרי, נוותר בסעיף זה על הפרמטר  $\epsilon$  כלומר, נקבע אותו ל-0.  
עבור ערך כלשהו של  $N$ , נריץ את האלגוריתם שלנו באופן הבא:

1. נריץ אותו עם פרמטר:  $t = 2$ .

2. נריץ אותו עם פרמטר:  $t = 4$ .

3. נריץ אותו עם פרמטר:  $t = 8$ .

...

מתי נעצור את הריצה? כאשר המרחק בין הרצות עוקבות היה קטן.  
וביתר פירוט: יהי  $v$  הוקטור שהתקבל ברצה עם פרמטר  $t = 2^i$   
ו-  $v'$  הוקטור שהתקבל עבור  $t = 2^{i-1}$  אם:

$$\|v - v'\| < 1/2^8 \text{ נעצור.}$$

#### שאלה 8:

1. הריצו את התהליך שתואר לעיל על הגרף שהגרלתם עם פרמטר  $N = 2^6$ .  
מה היה ערכו של  $t$  כאשר האלגוריתם עצר? האם הוקטור שהתקבל קרוב לוקטור המצופה?

2. חזרו על הסעיף הקודם עבור הפרמטר  $N = 2^7$ . עד כמה השתנה ערכו של  $t$ ?

הערה: אם עבור  $t = 2^{14}$ , המרחק בין הוקטורים שהתקבלו לא קטן דיו, עצרו את התהליך והודיעו שלא קיבלתם תשובה מספקת (ציינו מהו המרחק בין הוקטורים שהתקבלו).  
הערה 2: גם אם עבור  $t = 2^{14}$  עדיין לא התקבלה התשובה הרצויה נסו להגדיל את ערכו של  $t$  ולראות מתי היא מתקבלת.  
אם לתכנית שלכם לוקח יותר מ-10 דקות לחשב את התוצאה עצרו אותה.

### משפחה שניה בדיון – גרפים מקריים עם חטוטר:

בחלק זה נסתכל על גרפים המוגדרים מתוך ההתפלגות הבאה.

ראשית, נבנה גרף המכיל  $2^{10}$  קודקודים על פי ההתפלגות שתוארה לעיל (עבור הפרמטר  $p = 1/2^6$ , לכל זוג קודקודים  $(i, j)$  בגרף, נוסף את הקשת  $(i, j)$  בסיכוי  $p$ ).

לגרף שהתקבל נוסף  $2^6 = \frac{2^{10}}{2^4}$  קודקודים המחוברים זה לזה בגרף טבעת. נחבר את אחד מקודקודי הגרף שהגרלנו אל אחד מקודקודי גרף הטבעת. שימו לב: על הקודקודים הנוספים ניתן להסתכל כעל ניסיון הטיה של תוצאת הדירוג: אם מישו יבחר להוסיף מספר כמות כזו קודקודים (בבעיית דירוג דפי האינטרנט: על ידי הוספת דפים ולינקים כפי שתוארו), אזי הוא עשוי להעלות את הדירוג שלו (גם אם בפועל מבנה הרשת, מלבד הקודקודים שהוא בחר להוסיף, לא השתנה).

#### שאלה 9:

הריצו את הניסוי שהרצנו בשאלה 8 עבור הגרף החדש. האם הקודקודים החדשים (אלו שנוספו בגרף הטבעת) מדורגים גבוה יותר מיתר הקודקודים בגרף? אם כן, בכמה?

#### שאלה 10:

כעת הריצו את הניסוי שהרצנו בשאלה 8, אך כעת הפרמטר  $\epsilon$  אינו 0, והוא יקבל את הערכים בתחום:  $\frac{1}{2}, \dots, \frac{1}{2^{10}}$  (לכל ערך של  $\epsilon$  הריצו את הניסוי) הגדירו:  $N = 2/\epsilon$ . את הפרמטר  $t$  הריצו כפי שתואר בשאלה 8 (כלומר, הגדילו אותו עד אשר הוקטורים המתקבלים אינם שונים בהרבה, כפי שתואר לעיל). עבור אלו ערכים של  $\epsilon$  התקבל דירוג שלא נתן כמעט משקל יתר לקודקודים שהתווספו? האם השתנתה התפלגות הקודקודים בגרף המקורי?

### משפחה שלישית בדיון – מודל אחר גרפים מקריים:

בחלק זה נסתכל על גרפים המוגדרים מתוך ההתפלגות הבאה.

לכל קודקוד  $i$  תותאם הסתברות  $p_i$ .

לכל זוג קודקודים  $(i, j)$  בגרף נוסף את הקשת  $(i, j)$  בהסתברות  $p_j$ .

#### שאלה 11:

הגרילו גרף כנ"ל המכיל  $2^{10}$  קודקודים, והגדירו את הפרמטר  $p_i = 1/\log_2(i + 1)$ . עבור גרפים המוגרלים באופן שהוגדר לעיל, כיצד אתם מצפים שתיראה התפלגות הקודקודים בהילוך?

#### שאלה 12:

חזרו על שאלה 8 עבור הגרף שהגרלתם.

#### שאלה 13:

חזרו על שאלות 9, 10, כאשר הגרף שאליו אתם מוסיפים את הקודקודים המחוברים בגרף הטבעת, הוא הגרף אותו הגרלתם בשאלה 11. כיצד השתנו תשובותיכם? הסבירו את השינוי.  
האם יש הבדל בתוצאה אם חיברנו את גרף השרשרת לקודקוד מספר 1? או לקודקוד מספר  $n$ ?  
הסבירו מדוע.

בהצלחה! ☺