

## מבוא:

אנו מפרידים בין האינטרקציה החברתית באמצעות המטריצה  $W$ , לבין מצב המחלה אצל כל פרט ע"י הווקטור  $v$ .

- כרגע לפחות, המטריצה  $W$  מתארת את ההסתברות המותנית כך ש:

$$W[i, j] = \text{Prob}(i \text{ infected from } j \mid j \text{ is infectious})$$

וניתן לבנות את  $W$  ע"י שילוב מעגלים חברתיים שונים:  $W = W^{(work)} + W^{(family)} + W^{(random)}$ , כשלכל סוג מעגל חברתי ערכי הדבקה אחרים, בין צמתיים (קרי, אנשים) המחוברים לאותו מעגל.

העמודה ה- $j$  של  $W$  מתארת את מי הפרט ה- $j$  עלול להדביק.

השורה ה- $i$  של  $W$  מתארת מיהם מקורות ההדבקה האפשריים עבור הפרט ה- $i$ .

- הווקטור  $v$  מתאר את ההסתברות שכל פרט הינו מדבק ביום הנוכחי:

$$v[j] = \text{Prob}(j \text{ is infectious})$$

כך שהמכפלה מתארת את הסתברות ההדבקה:

$$p_{i,j} = W[i, j]v[j] = \text{Prob}(i \text{ infected from } j \& j \text{ is infectious})$$

ערכו של הווקטור  $v$  נקבע על סמך ההדבקה שחלה במהלך כל יום, ועל פי מודל מרקובי (הממומש עם מכונת מצבים) המתאר את התפתחות המחלה אצל כל פרט.

בנייר זה נעסוק בעיקר באינטראקציה החברתית המשפיעה על הסתברות ההדבקה, והמתוארת ע"י  $W$ , ופחות נתייחס למודל המרקובי המתאר את התפתחות המחלה.

## פעולות שאנו רוצים לבצע על $W$ :

א. יצירת מודל של הקשרים החברתיים (גנרציה של המטריצות  $W^{(work)}$ ,  $W^{(family)}$ ,  $W^{(random)}$ , ... המרכיבות את  $W$ ).

ב. נירמול נכון של המטריצה  $W$ . (כפי שנראה מיד, שלב זה מכיל פעולות של סיכום איברי המטריצה וכפל המטריצה בסקלר)

ג. חישוב הסתברות הדבקה של כל פרט באוכלוסייה (החישוב ידרוש פעולה בודדת מסוג  $\log(1-W)$ , על האיברים שבתוך המטריצה. וכן פעולות מכפלה בין מטריצה לווקטור).

ד. מדיניות גורפת של ניתוק מעגלים חברתיים (שימוש בתת קבוצה של המטריצות המרכיבות את  $W$ )  
ה. מדיניות ברמת הפרט של ניתוק מעגלים חברתיים (שקול להכפלת עמודה במטריצה בערך אפס).

נתאר כעת את הפרטים של כל סעיף, כפי שהם ממומשים היום, ועל אופני מימוש אלטרנטיביים עתידיים:

## אלגוריתמיקה ומימושים אלטרנטיביים

### א. מודל קשרים חברתיים:

המימוש הנוכחי למספר הקשרים החברתיים וחוזקם, הוא כיום דטרמיניסטי. (למשל משפחות הינן בגודל קבוע של 5 פרטים, והסתברות ההדבקה בין כל זוג פרטים של אותה המשפחה הינו קבוע). בהמשך, נוכל להרחיב זאת למודל סטוכסטי ומורכב יותר. תתי המטריצות כדוגמת  $W^{(work)}$ ,  $W^{(family)}$ ,  $W^{(random)}$  מחושבות פעם אחת בתחילת הריצה ונשמרות בצד לשימושים עתידיים של ניתוק קשרים חברתיים.  $W$  הינה כאמור סיכום כלל הקשרים החברתיים (סיכום כלל המטריצות).

### ב. נירמול:

פרמטר מרכזי בספרות המתאר את קצב ההדבקה הינו ה reproduction number -  $R_0$  כאשר במודל הפשוט ביותר:

$$R_0 = \beta d$$

עם  $\beta$  מתאר את מספר המודבקים מדי יום ע"י פרט חולה אחד,  $d$  הינו מספר ימי המחלה הכולל של אותו פרט מדבק. במודלים מעט מורכבים יותר (Kartiel, 2013), לוקחים בחשבון את הסתברות ההדבקה של החולה לכל

$$R_0 = \beta \sum_{\tau=1}^d P_{\tau} \quad \text{אורך ימי המחלה } P_{\tau}, \tau \in \{1, 2, \dots, d\} \text{ ואז מקדם ההדבקה הינו:}$$

אצלנו, המטריצה  $W$  מקודדת למעשה את מספר המפגשים החברתיים (ומשקללת אותם לפי הסתברויות ההדבקה בכל מפגש), כך שאצלנו  $\beta$  של אדם מסוים למעשה מיוצג ע"י הסכום לאורך עמודה של  $W$  השייכת לאותו פרט. לשם קבלת אומדן ממוצע לכל האוכלוסייה, ערכו של  $\beta$  יקבע מהערך הממוצע של

$$\text{סכומי העמודות: } \frac{1}{N} \sum_{j=1}^N W[:, j] \quad (\text{שהינו גם: } \frac{1}{N} \sum_{i,j} W[i, j], \text{ כאשר } N \text{ הינו גודל האוכלוסייה שבסימולציה}).$$

בהינתן  $R_0$  ופילוג הסתברויות ההדבקה לאורך ימי המחלה  $P_{\tau}, \tau \in \{1, 2, \dots, d\}$ , נוכל לכן לנרמל את המטריצה כדלקמן:

$$1. \quad \text{חשב } \beta \text{ רצוי: } \beta = \frac{R_0}{\sum_{\tau=1}^d P_{\tau}}$$

$$2. \quad \text{חשב } \hat{\beta} \text{ מצוי: } \hat{\beta} = \frac{1}{N} \sum_{i,j} W[i, j]$$

$$3. \quad \text{הכפל את } W \text{ פי } \hat{\beta} / \beta$$

פקטור נירמול זה יש להפעיל על כל תתי המטריצות (שכן הן ישמשו אותנו בהמשך בעת הפעלת סוגים שונים של מדיניות), כך שבסיום שלב זה יהיו בידנו מטריצות מהצורה:

$$\frac{\beta}{\hat{\beta}} W^{(work)}, \frac{\beta}{\hat{\beta}} W^{(family)}, \frac{\beta}{\hat{\beta}} W^{(random)}$$

וכן המטריצה הכוללת:

$$W = \frac{\beta}{\hat{\beta}} (W^{(work)} + W^{(family)} + W^{(random)})$$

מעתה נתייחס לתתי המטריצות וכן ל  $W$ , יחד עם פקטור הנרמול בפנים, ולא נציין עוד את הגורם  $\beta / \hat{\beta}$  במפורש.

### ג. חישוב הסתברות הדבקה:

גודל זה יש לחשב לפי:

$$p(i \text{ is infected}) = 1 - \prod_{j=1}^N (1 - p_{i,j})$$

כאשר, כזכור,  $p_{i,j} = W[i, j]v[j]$

מאחר שהכפלת מטריצה בווקטור הינה אוסף פעולות הכפלה וסיכום ואינה מבצעת את החישוב הנדרש, **כיום** אנו מבצעים חישוב זה בדרך הבאה:

1. שימוש בווקטור  $v$  בינארי בלבד (כלומר איברי 1 לפרטים באוכלוסייה שהגדרנו אותם להיות מדבקים (ב 100%) באיטרציה הנוכחית של האלגוריתם, ואיברי 0 בווקטור לפרטים באוכלוסייה שהינם לחלוטין אינם מדבקים). ההחלטה אם הפרט  $j$ -י עובר ל  $v[j]=1$ , תלויה בהסתברות ההדבקה שלו באותו יום ומוגרלת בהתאם.

2. חישוב  $\tilde{W} = \ln(1 - W)$  (כשהכוונה כאן הפעלת חישוב זה איבר-איבר על כל איברי המטריצה). (אך אין טעם כמובן לחשב זאת על איברים שלא נרשמו במפורש במטריצה הדלילה. ההסתברות שם היא אפס, וממופה לאפס גם במטריצה  $\tilde{W}$ ). נעיר כי חישוב זה יש לבצע רק פעם אחת. (למשל בעת גנרציית המטריצות, ולאחר שלב הנורמליזציה).

3. חישוב ההכפלה:  $tmp = \tilde{W}v$

4. חישוב ההסתברות:  $rslt = 1 - \exp(tmp)$  בכל איברי ווקטור התוצאה.

### 1 אלגוריתם

נשים לב כי חישוב זה בפועל מבצע:

$$tmp[i] = \sum_j \ln(1 - W[i, j]) v[j]$$

$$rslt[i] = 1 - \exp(tmp[i]) = 1 - e^{\sum_j \ln(1 - W[i, j]) v[j]} = 1 - \prod_j (1 - W[i, j])^{v[j]}$$

מאחר שכיום אנו עובדים עם  $v$  בינארי, ההסתברות המשלימה  $1 - p_{i,j} = 1 - W[i, j]v[j]$  מחושבת כאן

נכון, שכן עבור  $v[j] = 0$  וגם עבור  $v[j] = 1$  אכן מתקיים:

$$(1 - W[i, j])^{v[j]} = 1 - W[i, j]v[j]$$

כמו כן, בעת הפעלת מדיניות גלובאלית כדוגמת הימנעות מללכת לעבודה, כיום אנו מחשבים את  $W$  בשתי צורות: עם וללא תרומת המטריצה  $W^{(work)}$ . לאחר מכן, אנו מחשבים את  $\tilde{W} = \ln(1 - W)$  על 2 התצורות הנ"ל של  $W$  ומפעילים אותן לסירוגין, בהתאם למדיניות של אותו יום.

החיסרון בגישה זו: ברגע שיהיו לנו קומבינציות רבות של מדיניות שאולי נרצה לממש, נצטרך לחשב הרבה צורות שונות ל  $W$ .

לעומת זאת, כהרחבה עתידית פשוטה, אפשר לבצע את החישוב הבא:

1. לחשב את לוג המשלים על כל תתי המטריצות, קרי, לחשב:

$$\begin{aligned}\tilde{W}^{(family)} &= \ln(1 - W^{(family)}), \\ \tilde{W}^{(work)} &= \ln(1 - W^{(work)}), \\ \tilde{W}^{(random)} &= \ln(1 - W^{(random)})\end{aligned}$$

#### 1 Equation

2. החישוב יתבצע על פי המדיניות, ישר על תוצרים אלה.

#### אלגוריתם 2

לשם הדוגמא, נניח כי דגלים בוליאניים קובעים אילו מעגלים חברתיים הינם פעילים ואילו לא. החישוב יהיה אז:

$$\begin{aligned}\text{tmp} = & \\ & \text{bool\_family\_allowed} \cdot \tilde{W}^{(family)}_v + \\ & \text{bool\_work\_allowed} \cdot \tilde{W}^{(work)}_v \\ & \text{bool\_random\_wandering\_allowed} \cdot \tilde{W}^{(random)}_v\end{aligned}$$

#### 2 Equation

וההמשך כרגיל:

$$rslt = 1 - \exp(\text{tmp})$$

ואכן תצורה זו שקולה לחישוב המבוקש:

$$\begin{aligned}
& \text{tmp}[i] = \\
& \text{bool\_family\_allowed} \cdot \sum_j \ln(1 - W^{(\text{family})}[i, j]) v[j] + \\
& \text{bool\_work\_allowed} \cdot \sum_j \ln(1 - W^{(\text{work})}[i, j]) v[j] + \dots \\
& \Rightarrow \\
& \text{rslt}[i] = 1 - \exp(\text{tmp}[i]) = \\
& 1 - \left( \prod_j (1 - W^{(\text{family})}[i, j])^{v[j]} \right)^{\text{bool\_family\_allowed}} \left( \prod_j (1 - W^{(\text{work})}[i, j])^{v[j]} \right)^{\text{bool\_work\_allowed}} \dots
\end{aligned}$$

כאופציה נוספת לחישוב ההסתברות, טל הציע את האלגוריתם הבא (אחרי פרשנות ותרגום שלי. טל – אם לא דייקתי, תרגיש חופשי לתקן):

1. חשב את היחס בין המיפוי הלוגריתמי של ההסתברויות במעגלים החברתיים השונים לזו של מעגל חברתי אחד מסוים. לשם הדוגמא, אם נניח כי ההסתברות להידבק ממישהו אקראי הינה  $p_\alpha$ , וההסתברות להידבק ממישהו במעגל המשפחה הינה  $p_\beta$ , ולהידבק ממישהו בעבודה הינה  $p_\gamma$ , אנו

$$\text{נחשב את } x_\beta = \frac{\ln(1 - p_\beta)}{\ln(1 - p_\alpha)} \text{ וכן את } x_\gamma = \frac{\ln(1 - p_\gamma)}{\ln(1 - p_\alpha)}$$

2. בכל מקום במטריצה המקורית  $W$ , שם היה רשום  $p_\alpha$ , נציב 1. במקומות בהם הופיעו ערכי  $p_\beta$ , נציב

$$x_\beta \text{ ובאופן דומה ערכי } p_\gamma \text{ יוחלפו ב } x_\gamma. \text{ נקרא למטריצה המתקבלת } \tilde{W}.$$

נשים לב כי למעשה יש כאן בניית מטריצה המתקבלת ע"י המיפוי:

$$\tilde{W}[i, j] = f(W[i, j]) = \frac{\ln(1 - W[i, j])}{\ln(1 - p_\alpha)}$$

הסתברות מפגש אקראי, בדוגמא שלנו).

3. חשב את המכפלה  $\tilde{W}v$ .  $\text{tmp} = \tilde{W}v$  (הווקטור  $v$  הינו בינארי גם באלגוריתם של טל).

נשים לב כי כל ערך תוצאה בווקטור  $\text{tmp}$  מתאר למעשה את כמות הפעמים בהם סיכמנו תרומות מסוג

$$\ln(1 - p_\alpha). \text{ לשם הדוגמא, אם יצא } \text{tmp}[i] = c \text{ המשמעות היא כי בשורה ה-} i \text{ של } \tilde{W} \text{ היו איברים}$$

כך ש:

$$c = \sum_j \tilde{W}[i, j] v[j] =$$

$$1 \cdot (\text{number of random infectious contacts}) +$$

$$x_\beta \cdot (\text{number intra-family infectious contacts}) +$$

$$x_\gamma \cdot (\text{number intra-work infectious contacts})$$

$$4. \text{ חשב את } rslt[i] = 1 - (1 - p_\alpha)^c$$

### אלגוריתם 3

תוצאה זו נכונה שכן:

$$rslt[i] = 1 - (1 - p_\alpha)^c = 1 - (1 - p_\alpha)^{Nr + x_\beta Nf + x_\gamma Nw}$$

כאשר סימנו ע"י  $Nr, Nf, Nw$  את מספר המפגשים המדבקים בעבודה, במשפחה, ומפגשים רנדומליים, בהתאמה.

אם נפרט יותר מדוע האלגוריתם של טל נכון:

$$\begin{aligned} rslt[i] &= 1 - (1 - p_\alpha)^{Nr} (1 - p_\alpha)^{x_\beta Nf} (1 - p_\alpha)^{x_\gamma Nw} \\ &= 1 - (1 - p_\alpha)^{Nr} (1 - p_\alpha)^{\frac{\ln(1 - p_\beta)}{\ln(1 - p_\alpha)} Nf} (1 - p_\alpha)^{\frac{\ln(1 - p_\gamma)}{\ln(1 - p_\alpha)} Nw} \\ &= 1 - (1 - p_\alpha)^{Nr} (1 - p_\alpha)^{\log_{(1 - p_\alpha)}(1 - p_\beta) Nf} (1 - p_\alpha)^{\log_{(1 - p_\alpha)}(1 - p_\gamma) Nw} \\ &= 1 - (1 - p_\alpha)^{Nr} (1 - p_\beta)^{Nf} (1 - p_\gamma)^{Nw} \end{aligned}$$

לדעתי, לאחר שמסבירים את האלגוריתם שטל הציע בצורה זו, רואים 2 דברים:

1. לא באמת חוסכים כאן חישובי log. הם עדיין מתרחשים פעם אחת בעת בניית המטריצות. (ובכל זאת, מה שכן אפשר לקחת מפה הוא, שאם יש לנו סט קטן וסופי של הסתברויות הבונות את  $W$ , עדיף כנראה לחשב  $\log(1-p)$  עליהן, ואז להציב במטריצה הגדולה תוצרים אלה, במקום לבצע חישוב  $\log(1-W)$  על כל האיברים של  $W$ . עשוי לחסוך זמן חישוב).
2. בגלל שהכל מבוטא ע"פ "הסתברות בסיס" כלשהיא,  $p_\alpha$  בדוגמא זו, זה כן חוסך לנו את שלב האקספוננט בסוף. כלומר בעוד שבאלגוריתם הנוכחי שלנו (אלגוריתם 1) צריך בסוף לחשב  $rslt = 1 - \exp(tmp)$  על ווקטור התוצאה tmp, הרי שבצורה שטל מציע, מחשבים  $rslt[i] = 1 - (1 - p_\alpha)^c$  (או ברישום ווקטורי:  $rslt = 1 - (1 - p_\alpha)^{tmp}$ ). אולי זה יהיה באמת מהיר יותר? אפשר לבדוק.

### ד. מדיניות גורפת של ניתוק מעגלים חברתיים:

זו דרישה די פשוטה; כאמור מתבקש אז להשתמש ב subset של המטריצות שמרכיבות את  $W$ . הראיתי גם איך מכלילים זאת למישור המיפוי עם ה log. (אלגוריתם 2). בגרסא עתידית אפשר לשקול גם שימוש בניתוק "רך" של מעגלים חברתיים. ניתן לעשות זאת ע"י בניית צרוף בסגנון:

$$W = \sum_k \alpha_k W^{(k)} \text{ כאשר } W^{(k)} \text{ מתאר את הקשרים החברתיים מסוג } k, \text{ ו } \alpha_k \text{ מקיימים } 0 \leq \alpha_k \leq 1.$$

כך ש  $\alpha_k$  קרוב ל 1 משמעו כי המעגל החברתי כמעט ואינו נפרם, ועבור  $\alpha_k$  קרוב לאפס המשמעות היא כי הציבור מציית והסתברויות המפגש במעגל זה צונחות קרוב לאפס.

ה. מדיניות ניתוק מעגלים ברמת הפרט:

- פעולה זו מתרחשת למשל כאשר פרט מבין כי הינו חולה ולכן מציב את עצמו בריחוק מהמעגלים החברתיים שלו.
- אם הוא מתנתק מכל קשריו החברתיים (הוא לא פוגש אחרים ואחרים לא פוגשים אותו), הדבר שקול להצבת אפס לאורך השורה והעמודה המתאימה של  $W$ .
  - אם הפרט מתנתק רק מחלק מקשריו החברתיים, הדבר שקול לאיפוס השורה והעמודה במטריצה המתאימה של אותו מעגל חברתי.
- להבנתי כיום יש לנו קושי חישובי לעשות עדכוני שורה ועמודה במטריצה, תוך כדי ריצה.

הפתרון היחיד שאני חושב עליו בשלב זה (מעבר ללכתוב class חדש שיתאים לנו ככפפה ליד) הוא להשתמש במספר עותקים של  $v$  ואיפוס האיבר המתאים ב  $v$  בעת ביצוע ההכפלה במטריצות.

כלומר כמו ב Equation 1 נחשב תחילה את המיפוי הלוגריתמי:

$$\begin{aligned}\tilde{W}^{(family)} &= \ln(1 - W^{(family)}), \\ \tilde{W}^{(work)} &= \ln(1 - W^{(work)}), \\ \tilde{W}^{(random)} &= \ln(1 - W^{(random)})\end{aligned}$$

ואז נבצע חישוב, בדומה ל Equation 2, אך עם מספר עותקים של  $v$ :

$$tmp = \tilde{W}^{(family)} v_{copy1} + \tilde{W}^{(work)} v_{copy2} + \tilde{W}^{(random)} v_{copy3}$$

(אני מתעלם כאן מהמדיניות הגורפת, שעדיין יכולה לבוא לידי ביטוי עם דגלים בוליאניים, כמו ב Equation 2).

אם למשל פרט מספר  $j$  מחליט למנוע מעצמו את מעגל העבודה, אזי נציב  $v_{copy2}[j] = 0$  לאיטרציה זו.

## Bibliography

Kartiel, G. (2013). STOCHASTIC DISCRETE-TIME AGE-OF-INFECTION EPIDEMIC MODELS. *International Journal of Biomathematics*.