בס"ד

מבוא להסתברות וסטטיסטיקה 88-165

תרגילי בית 12 - שרשראות מרקוב - פתרונות

1. שלושה כדורים שחורים ושלושה כדורים לבנים נמצאים בשני כדים, עם שלושה כדורים בכל כד. אומרים שהמערכת נמצאת במצב i אם יש i כדורים לבנים בכד הראשון, $i = 0, 1, 2, 3$. בכל צעד מעבירים באקראי כדור אחד מהכד הראשון לשני, וכדור אחד מהשני לראשון. אם X_n מציין את המצב של המערכת בזמן n, מצא את מטריצת המעבר של שרשרת המרקוב X_n. מצא גם את ההתפלגות הקבועה.

חישוב מטריצת המעבר הוא תרגיל פשוט בקומבינטוריקה. התשובה:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} & 0 \\ 0 & \frac{4}{9} & \frac{4}{9} & \frac{1}{9} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

כדי למצוא התפלגויות קבועות פותרים את המשוואה $wP = w$ (w וקטור שורה). הפתרון היחיד (מנורמל כך שסכום הרכיבים הוא 1):

$$w = \begin{pmatrix} \frac{1}{20} & \frac{9}{20} & \frac{9}{20} & \frac{1}{20} \end{pmatrix}$$

---

2. בסדרה של ניסויי ברנולי בלתי תלויים עם סיכוי להצלחה p, אומרים שהמערכת במצב 0 אם שני הניסויים האחרונים נתנו הצלחות, במצב 1 אם הניסוי האחרון נתן הצלחה אבל הניסוי אחד לפני אחרון נתן כשלון, במצב 2 אם הניסוי האחרון נתן כשלון אבל הניסוי אחד לפני אחרון נתן הצלחה, ובמצב 3 אם שני הניסויים האחרונים נתנו כשלונות. מצא את מטריצת המעבר P של שרשרת מרקוב זה. הוכח שכל המטריצות P^n עבור n גדול מ-1 הן שוות, והסבר את המשמעות.

אם לשרשרת מרקוב יש מטריצת מעבר P, וקיימים $N > 1$ ומספרים Π_j כך ש-

$$(P^N)_{ij} = \Pi_j \ ,$$

הוכח שלכל $n > N$

$$(P^n)_{ij} = \Pi_j \ .$$

$$P = \begin{pmatrix} p & 0 & 1-p & 0 \\ p & 0 & 1-p & 0 \\ 0 & p & 0 & 1-p \\ 0 & p & 0 & 1-p \end{pmatrix}$$

מוצאים ש-

$$P^2 = P^3 = \ldots = \begin{pmatrix} p^2 & p(1-p) & p(1-p) & (1-p)^2 \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \end{pmatrix}$$

המשמעות - בסדרה של ניסוי ברנולי ב"תים אין השפעה של זוג של ניסויים על כל הניסויים הבאים אחריו. לכן, לא משנה באיזה מצב אני מתחיל בו, אחרי שני שלבים או יותר יש הסתברות p^2 להיות במצב 0, הסתברות $p(1-p)$ להיות במצב 1 וכו'.

אם נתון ש-

$$(P^N)_{ij} = \Pi_j$$

אזי

$$\begin{aligned}(P^{N+1})_{ij} &= (P \cdot P^N)_{ij} \\ &= \sum_k P_{ik}(P^N)_{kj} \\ &= \sum_k P_{ik}\Pi_j \\ &= \Pi_j\end{aligned}$$

היות והסכום של כל האיברים בשורה של מטריצת המעבר שווה 1 ($\sum_k P_{ik} = 1$). ובהמשך באנדוקציה, לכל $n \geq N$, $(P^n)_{ij} = \Pi_j$.

---

3. אם בשרשרת מרקוב מסמנים את הסתברות המעבר $\mathbf{P}(X_{n+1} = j | X_n = i)$ על ידי P_{ij}, ואת הסתברות ההתחלתית $\mathbf{P}(X_0 = i)$ על ידי a_i, מצא

(א) $\mathbf{P}(X_1 = X_2)$

(ב) $\mathbf{P}(X_1 = i | X_2 = j)$

(ג) $\mathbf{P}(X_3 = X_0 | X_1 = X_2)$

(א)

$$\begin{aligned}\mathbf{P}(X_1 = X_2) &= \sum_i \sum_j \mathbf{P}(X_2 = X_1 = i, X_0 = j) \\ &= \sum_i \sum_j a_j P_{ji} P_{ii}\end{aligned}$$

(בשורה האחרונה רשמתי את ההסתברות להתחיל במצב j, כפול ההסת' לעבור מ- j ל- i, כפול ההסת' לעבור מ- i ל- i.) ניתן גם לרשום את התשובה בצורה $\sum_i (aP)_i P_{ii}$ - אם ווקטור השורה a נותן את ההתפלגות של מצבים בצעד התחלתי, aP נותן את ההתפלגות בצעד הבא, aP^2 בצעד אחרי זה וכו'.

(ב)

$$\begin{aligned}\mathbf{P}(X_1 = i | X_2 = j) &= \frac{\mathbf{P}(X_1 = i, X_2 = j)}{\mathbf{P}(X_2 = j)} \\ &= \frac{P_{ij}(aP)_i}{(aP^2)_j}\end{aligned}$$

השתמשתי בהערה מהסעיף הקודם. שים לב ש- $(aP^2)_j = \sum_i (aP)_i P_{ij}$ ולכן המונה קטן או שווה למכנה.

(ג)

$$\begin{aligned}\mathbf{P}(X_3 = X_0 | X_1 = X_2) &= \frac{\mathbf{P}(X_3 = X_0, X_1 = X_2)}{\mathbf{P}(X_1 = X_2)} \\ &= \frac{\sum_i \sum_j \mathbf{P}(X_3 = X_0 = i, X_1 = X_2 = j)}{\sum_i \sum_j \mathbf{P}(X_0 = i, X_1 = X_2 = j)} \\ &= \frac{\sum_i \sum_j a_i P_{ij} P_{jj} P_{ji}}{\sum_i \sum_j a_i P_{ij} P_{jj}}\end{aligned}$$

---

4. לשרשרת מרקוב עם מטריצת המעבר

$$\begin{pmatrix} \frac{1}{5} & \frac{4}{5} & 0 & 0 & 0 & 0 \\ \frac{3}{5} & \frac{2}{5} & 0 & 0 & 0 & 0 \\ \frac{1}{10} & \frac{1}{10} & \frac{3}{10} & \frac{3}{10} & \frac{1}{10} & \frac{1}{10} \\ 0 & \frac{1}{10} & \frac{2}{5} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{10} & \frac{7}{10} \\ 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix}$$

(א) מצא את המחלקות המקושרות.

(ב) מצא את כל ההתפלגויות הקבועות.

(א) אומרים ששני מצבים מתקשרים זה עם זה אם ניתן להגיע מהמצב הראשון לשני ולשני מהראשון (כולל אחרי 0 צעדים, כך שכל מצב מתקשר עם עצמו). זה יחס שקילות המחלק את המצבים למחלקות שקילות. יש פה 3 מחלקות מקושרות: המצבים 1 ו-2, המצבים 5 ו-6, והמצבים 3 ו-4.

(ב) פורמלית יש לנו לפתור את המשוואה $wP = w$. יש פתרון אחד לכל מחלקה מתמשכת: למחלקה של המצבים $1, 2$ ההתפלגות היא

$$w_1 = \begin{pmatrix} \frac{3}{7} & \frac{4}{7} & 0 & 0 & 0 & 0 \end{pmatrix}$$

למחלקה של המצבים $5, 6$ ההתפלגות היא

$$w_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{7}{9} \end{pmatrix}$$

ניתן לבדוק שהפתרון הכללי של $wP = w$ הוא צירוף לינארי של שני פתרונות אלה, כלומר

$$w = pw_1 + (1-p)w_2 \ , \qquad 0 \le p \le 1$$

---

5. כתוב את מטריצת המעבר למשחק "נחשים וסולמות" למטה (משחקים עם קוביה אחת ומסיימים את המשחק כאשר מגיעים ל-30 או עוברים אותו). מצא את התוחלת של הזמן עד לסיום המשחק.

מטריצת המעבר היא $\frac{1}{6}$ כפול

$$
\begin{matrix}
0&1&0&1&0&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&1&0&1&1&2&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&1&0&1&1&2&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&1&2&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&1&1&1&1&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&1&1&1&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&1&1&0&1&1&1&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&1&1&1&1&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&1&1&1&1&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&1&0&0&0&0&0&0&1&1&1&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&1&0&0&0&0&0&0&0&1&1&0&1&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&1&0&0&1&0&1&0&0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&1&0&0&1&0&1&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&0&1&0&0&0&1&1&0&0&0&0&0&1&0\\
0&0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&1&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&0&1&1&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&1&1&1\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&1&1&2\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&3\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&4\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&5\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&6\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&6
\end{matrix}
$$

כדי למצוא את הזמנים הממוצעים עד שמגיעים למצב הסופג (30) עושים את הפעולות הבאות במטלב:

```
Q=P(1:29,1:29);
N=inv(eye(29)-Q);
t=sum(N,2)
```

זה מייצר ווקטור שנותן את הזמנים הממוצעים עד שמגיעים ל-30 כאשר מתחילים בריבוע

$1, 2, \ldots, 29$:

9.3976
9.2924
9.8441
9.6419
8.6862
8.7446
8.7884
8.6616
8.5665
8.4291
9.5481
9.0947
9.0509
7.9008
7.9959
7.6046
6.7724
6.3741
5.4786
5.7998
5.9383
5.2567
4.6486
3.9845
3.4153
2.9274
1.3611
1.1667
1.0000

הערה 1: ניתן למחוק את המצבים $3, 5, 11, 17, 19, 20, 21, 27$ מהמשחק היות והם ראשי נחשים או זנבי סולמות ולא ניתן להגיע להם. זה בא לידי ביטוי במטריצת המעבר שיש 8 עמודים של אפסים. אם מוחקים עמודים אלה ואת השורות המתאימות זה לא משפיע על החישוב. לבדוק את זה במטלב יש לעשות

```
ro=[1,2,4,6:10,12:16,18,22:26,28:30];

Pres=P(ro,ro);

Q=Pres(1:21,1:21);

N=inv(eye(21)-Q);

t=sum(N,2)
```

מקבלים אותן התשובות - בלי, כמובן, זמנים אם מתחלים באחד המצבים הנמחקים.

הערה 2: בדרך כלל משחקים גרסא של המשחק שצריכים להגיע ממש ל-30 ואם למשל עומדים בריבוע 28 וזורקים 5, אזי הולכים שני צעדים קדימה ל- 30 ו- 3 צעדים אחורה ל-

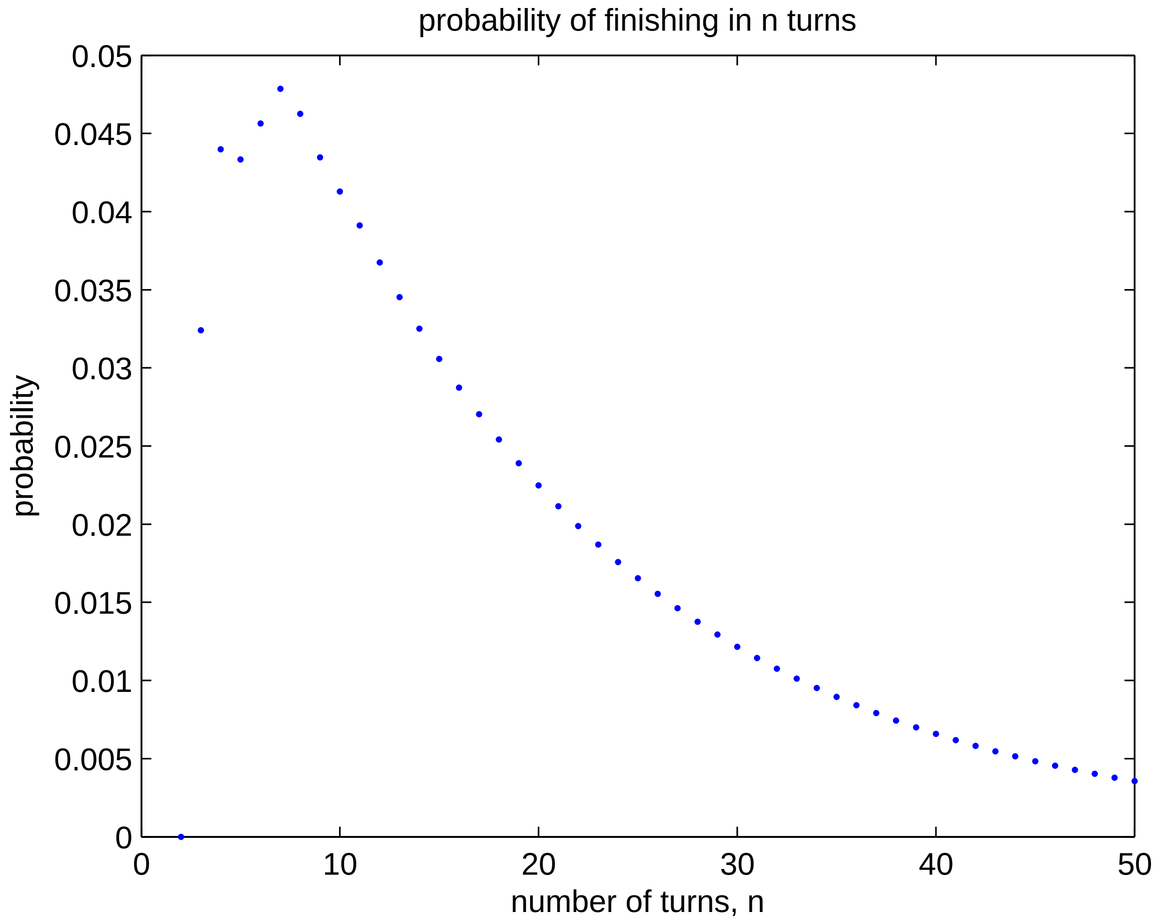
27 - ואז יורדים בנחש ל 1. לפי גרסה זו של המשחק יש לשנות את שורות 25 עד 29 ל-

1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	1
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	1
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1

המשחק מתמשך באופן משמעותי. התוחלות של הזמנים הם עכשיו

19.8333
19.7413
20.3578
20.1648
19.1194
19.2060
19.2709
19.1889
19.1274
19.0067
20.2147
19.7252
19.6603
18.6969
18.7584
18.2827
17.4335
16.7882
15.7292
15.7292
15.8468
15.4282
15.0694
12.9167
12.9167
12.9167
12.9167
12.9167
12.9167

מספרים אלה מחזקים את הדעה שזה מאוד מעצבן שצריך להגיע ממש לריבוע האחרון! אבל יכול להיות שהתוחלת כאן היא לא מדד רלוונטי כל כך. מה שקובע במשחק זה הוא כמה פעמים נפגשים בנחש הגדול מ-27 עד 1. באיור מופיע ההתפלגות של מספר התורים להגיע ל- 30 (כאשר מתחילים ב- 1). רואים שזה רחוק מהתפלגות נורמלית. סטיית התקן הוא בערך 16.3.


probability of finishing in n turns
probability
number of turns, n
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
0.045
0.05
0
10
20
30
40
50