# תהליכים סטוכסטים

**הגדרות:**

- תהליך סטוכסטי (בדיד, סופי) - סדרה של משתנים מקריים $X_0, X_1, ...$ מעל קבוצה סופית של מצבים $S$
- נאמר שלתהליך יש את <u>תכונת מרקוב</u> אם לכל $i > 0$ המשתנה המקרי $X_t|X_{t-1}$ בלתי תלוי בכל המשתנים $X_0, ..., X_{t-2}$ כלומר בכל "רגע" מה יהיה מצב הבא בתהליך תלוי אך ורק במצב שקדם לא (כאילו לא זוכרים את המצבים הקודים)
- <u>שרשרת מרקוב</u> $(Markov\ Chain)$ הינה תהליך סטוכסטי בעל תכונת מרקוב כך שקיימים $\{p_{ij}|i,j \in S\}$ כך שלכל $t > 0$ $i,j \in S$ מתקיים

$$Pr\left[X_t = j|X_{t-1} = i\right] = p_{ij}$$

כלומר ישנה לכל שני מצבים $i,j \in S$ ההסתברות למעבר ממצב $i$ למצב $j$ היא קבוע $(p_{i,j})$ ולא תלויה בזמן שבו היא מתרחשתכל שרשרת מרקוב ניתן להגדיר על ידי מטריצת מעברים

**דוגמה:**

עבור הילוך מקרי בגרף $G = \langle V, E\rangle$ המצבים הם הקודקודים $S = V$

ומטריצת המעברים תהיה $P = \{p_{u,v}\}_{u,v \in V}$ כאשר

$$p_{u,v} = \begin{cases} 0 & (u,v) \notin \mathrm{E} \\ \frac{1}{deg(u)} & (u,v) \in E \end{cases}$$

כל סדרה מוגדרת היטב על ידי התפלגות המצב הראשון $X_0$ ומטריצת המעברים (למעשה נראה עוד מעט שמההתפלגות הראשונה, של ההסתברות להמצאות בכל מצב, ההתפלגות אחרי $n$ צעדים, תתקבל ע"י הכפלת $X_0$ $n$ פעמים במטריצה)

באופן פורמלי לכל סדרת מצבים $\sigma_0, ..., \sigma_n$ נוכל לשאול מה ההסתברות שזה המסלול שנעבור ב $n$ צעדים ונקבל

$$Pr\left[X_0 = \sigma_0, ..., X_n = \sigma_n\right] = Pr\left[X_0 = \sigma_0\right] \cdot \prod_{t=1}^{n} p_{\sigma_{t-1},\sigma_t}$$

**הגדרה:** שרשרת מרקוב המוגדרת על ידי מטריצת המעברים $P$ היא אי־פריקה אם לכל $i,j\in S$ יש מסלול עם הסתברות חיובית מ$i$ ל$j$ לפי $P$

בדגומה של הילוך מקרי בגרף ההגדרה הזאת שקולה לדרישה שהגרף יהיה קשיר

**נשים לב:** אם נסכום את ההסתברויות עבור מצב מסוים, בהינתן כל מצב אפשרי קודם, מהגדרה של הסתברות נקבל כי לכל $i$ מתקיים

$$\sum_{j\in S}p_{i,j}=1$$

**במילים** אחרות, מאחר והכפלה בוקטור אחדות בעצם מחזיר את הסכום בכל שורה, נקבל

$$P\cdot\begin{bmatrix}1\\ .\\ .\\ .\\ 1\end{bmatrix}=\begin{bmatrix}1\\ .\\ .\\ .\\ 1\end{bmatrix}\Rightarrow(P-I)\begin{bmatrix}1\\ .\\ .\\ .\\ 1\end{bmatrix}=0$$

**למה:** $rank\left(P-I\right)=n-1$ כאשר $n=|S|$ והשרשרת־מרקוב היא אי־פריקה

**טענת עזר:** לכל $x=\left(x_1,...,x_n\right)\in\mathbb{R}^n$ אם $\left(P-I\right)x=0$ אזי $x_1=x_2=...=x_n$

כלומר לא רק שוקטור האחדות מאפס את $P-I$ כמו שראינו הוא גם היחיד (עד כדי כפל בסקלר)

אם נוכיח את טענת העזר נוכל להוכיח את הלמה באופן הבא – $rank\left(P-I\right)=n-$ $dim\left(ker\left(P-I\right)\right)=n-1\Leftarrow dim\left(ker\left(P-I\right)\right)=1$

**טריק שימושי:** פונקציות הרמוניות

תהי $f:\{1,...,n\}\to\mathbb{R}$ ונתון כי $\forall t\in\{1,..,n\}:\ f(t)=\frac{f(t+1)-f(t)}{2}$

וכמו כן מספרים לנו שבנקודה $k$ הפונקציה מקבל מקסימום כלומר $argmax(f)=k$

אזי בהכרח $f\left(k+1\right)=f\left(k\right)=f\left(k-1\right)$

**הסבר:** מאחר והערך בנקודה $k$ הוא הממוצע של שני הערכים השכנים לו אם הוא מקסימלי אזי אם אחד מהם קטן ממנו השני היה צריך להיות גדול ממנו כדי לאזן (שהרי ממוצע של שני מספרים נמצא בין שניהם) מאחר ונתון שאין אחד גדול ממנו בהכרח הם לא קטנים ממנו אלא שווים לו (הם לא יכולים להיות גדולים ממנו, כמו שאמרנו, בגלל שהוא מקסימלי)

הטענה נכונה באופן מקביל גם למינימום של $f$

**מסקנה:** בקצוות הקטע הפונקציה מקבלת מקסימום או מינימום

**הוכחת הטענה:** יהי $x=\begin{bmatrix}x_1\\ .\\ .\\ .\\ x_n\end{bmatrix}$ כך ש $(P-I)\,x=0$ או במילים אחרות $Px=x$

נקבל

$$\forall i\ x_i=(p_1,...,p_n)\cdot\begin{bmatrix}x_1\\ .\\ .\\ .\\ x_n\end{bmatrix}=\sum_j p_{ij}x_j$$

כלומר $x_i$ הוא "ממוצע משוקלל" של השכנים של $i$ (של הערכים $\{x_j|p_{ij}>0\}$)

ומכאן - בהינתן $i_0\in S$ כך ש $x_{i_0}=\underset{j}{max}\,(x_j)$

לכל $k$כך ש $p_{ij}>0$ מתקיים $x_j=x_{i_0}\underset{i}{max}\,(x_i)$

כי אחרת אם קיים "שכן" של $i_0$ , כלומר מצב $j$ שההסתברות למעבר מ$i_0$ אליו היא חיובית, כך ש $x_j<x_{i_0}$ ואז כדי "לאזן" את הממוצע המשוקלל שיהיה שווה ל$x_{i_0}$ היה חייב להיות שכן אחר של $i_0$ שמקיים $x_{j'}>x_{i_0}$ בסתירה למקסימליות של $x_{i_0}$

כיוון שהשרשרת היא אי־פריקה לכל מצב $j\in S$ קיים מסלול בעל הסתברות חיובית להתרחשות מ $i_0$ עד אליו ולכן מאותו הטיעון היינו מקבלים באינקודציה שהערך של כל השכנים של $x_{i_0}$ שווה לערך של $max\,(x_i)=x_{i_o}$ והערך של כל השכנים שלהם גם שווה לזה וכן הלאה עד ל$x_j$, כלומר קיבלנו ש $\forall j\in S\ x_j=x_{i_0}\,(=max\,(x_i))$ ובסה"כ נקבל

$$x_1=...=x_n$$

**נשים לב:** אם בזמן $t$ ההתפלגות של המשתנה המקרי $X_t$ נתונה על ידי $q=(q_1,...,q_n)$ כאשר $Pr\,[X_t=i]=q_i$

אזי בזמן $t+1$ ההתפלגות של $X_{t+1}$ היא $qP$

**הסבר:** נסמן $qP=(q'_1,...,q'_n)$ כאשר[1]

$$q'_i=q\cdot\begin{bmatrix}p_{1j}\\ .\\ .\\ .\\ p_{nj}\end{bmatrix}=\sum_{i\in S}q_ip_{ij}=\sum_{i\in S}Pr\left[X_{t+1}=j|X_t=i\right]\overset{*}{=}Pr\left[X_{t+1}=j\right]$$

**הגדרה:** התפלגות $\pi=(\pi_1,...,\pi_n)$ נקראת התפלגות מקובעת אם

$$\pi P=\pi$$

כלומר אם בזמן $t$ ההתפלגות הנוכחית היא $\pi$ אזי גם בזמן $t+1$ היא תישאר $\pi$ (ובעצם גם בזמן הבא וכן הלאה, ההתפלגות בעצם מתקבעת מכאן והלאה)

דוגמה: בהילוך מקרי בגרף קשיר $G=\langle V,E\rangle$ ההתפלגות המקובעת היא

$$\pi_v=\frac{deg\left(v\right)}{2|E|}$$

למה זה נכון?

1. זאת התפלגות -

$$\sum_v\pi_v=\frac{\sum_v deg\left(v\right)}{2|E|}=1$$

כי כזכור סכום הדרגות בגרף שווה ל $2|E|$

2. היא מקובעת -

$$Pr\left[X_{t+1}=v|X_t\sim\pi\right]=\sum_u\pi_up_{uv}=\sum_{u\in\Gamma(v)}\frac{deg\left(u\right)}{2|E|}\cdot\frac{1}{deg\left(u\right)}$$

$$=\frac{1}{2|E|}\sum_{u\in\Gamma(v)}1=\frac{1}{2|E|}deg\left(v\right)=\pi_v$$

כלומר גם $X_{t+1}$ מתפלג בהתפלגות $\pi$

[1] המעבר האחרון (*) נובע מנוסחאת ההסתברות השלמה