תהליכים סטוכסטים 6.10.15

הגדרות:

- תהליך סטוכסטי (בדיד, סופי) - סדרה של משתנים מקריים $X_0, X_1, ...$ מעל קבוצה סופית של מצבים S
- נאמר שלתהליך יש את תכונת מרקוב 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 הקודמים)
- שרשרת מרקוב $(Markov\ Chain)$ 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 מi לj 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 1: $rank\left(P-I\right)=n-1$ **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 – $rank\left(P-I\right)=n-1\Leftarrow dim\left(ker\left(P-I\right)\right)=1$ $dim\left(ker\left(P-I\right)\right)=n-1$

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 kכך 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 מi_0 אליו היא חיובית, כך ש $x_j<x_{i_0}$ ואז כדי "לאזן" את הממוצע המשוקלל שיהיה שווה לx_{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 לx_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]

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

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

$$\pi P=\pi$$

כלומר אם בזמן t ההתפלגות הנוכחית היא π אזי גם בזמן $t+1$ היא תישאר π (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 -

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

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

כלומר גם X_{t+1} מתפלג בהתפלגות π

למה 2: אם P אי־פריקה קיימת התפלגרות מקובעת π וכן

1. $\forall i\in S\ \pi_i>0$
2. ההתפלגות π היא ההתפלגות המקובעת היחידה

הוכחה: בלמה 1 ראינו כי $rank\left(P-I\right)^t=rank\left(P-I\right)=n-1$

ומכאן

$$dim\left(ker\left(P-I\right)^t\right)=1$$

כלומר קיים וקטור $v\neq 0$ כך ש־

$$\exists v\neq 0\ \ \left(P-I\right)^t v=0$$

ומאחר ובמטריצה ריבועית דרגת השורות שווה לדרגת העמודות ולכן גם המימד של הגרעין שלהם שווה ולכן

$$v^t\left(P-I\right)\Rightarrow v^tP=v^t$$

והוקטור יחיד עד כדי מכפלה בסקלר

בפרט קיים וקטור יחיד (ממש) x כך ש $xP=x$ ובנוסף הוא מייצג התפלגות כלומר $\sum_i x_i=1$ (יש דרך אחת לנרמל כראוי את הסכום על ידי כפל בסקלר)

עדיין נותר להראות ש $\forall i\ x_i>0$

נחלק את המצבים לפי x לשתי קבוצות

$$S^+=\{i|x_i>0\}\qquad S^{\leq 0}=\{i|x_i\leq 0\}$$

ונתבונן בסכום $\sum_{j\in S^{\leq 0}}x_j$

נשים לב ש $xP=x$ או במילים אחרות (כאשר P_j מסמל את העמודה ה j)

$$\forall j\ x_j=xP_j=\sum_{i\in S}x_iP_{ji}$$

ולכן נקבל

$$\sum_{j\in S^{\leq 0}}x_j=\sum_{j\in S^{\leq 0}}\left(\sum_{i\in S}x_iP_{ij}\right)=\sum_{j\in S^{\leq 0}}\left(\sum_{i\in S^{\leq 0}}x_iP_{ij}+\sum_{i\in S^+}x_iP_{ij}\right)$$

---

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

$$= \left(\sum_{j \in S} \sum_{i \in S^{\leq 0}} x_i P_{ij} - \sum_{j \in S^+} \sum_{i \in S^{\leq 0}} x_i P_{ij} \right) + \sum_{j \in S^{\leq 0}} \sum_{i \in S^+} x_i P_{ij}$$

נזכור שכל שורה וכל עמודה של P מייצגת התפלגות ולכן

$$\forall i \sum_{i \in S^{\leq 0}} P_{j \in S ij} = 1$$

נציב ונקבל

$$\sum_{i \in S^{\leq 0}} x_i - \sum_{j \in S^+} \sum_{i \in S^{\leq 0}} x_i P_{ij} + \sum_{j \in S^{\leq 0}} \sum_{i \in S^+} x_i P_{ij}$$

ובסה"כ קיבלנו את השיוויון

$$\sum_{j \in S^{\leq 0}} x_j = \sum_{i \in S^{\leq 0}} x_i - \sum_{j \in S^+} \sum_{i \in S^{\leq 0}} x_i P_{ij} + \sum_{j \in S^{\leq 0}} \sum_{i \in S^+} x_i P_{ij}$$

נצנצם ונעביר אגפים ונקבל

$$(*) \sum_{j \in S^+} \sum_{i \in S^{\leq 0}} x_i P_{ij} = \sum_{j \in S^{\leq 0}} \sum_{i \in S^+} x_i P_{ij}$$

נשים לב ש P_{ij} תמיד אי־שלילי ובנוסף מהאופן בו הגדרנו את הקבוצות מתקיים

$$\forall i \in S^+ \ x_i > 0, \ \forall i \in S^{\leq 0} \ x_i \leq 0$$

ולכן ב(*) צד ימין של השיוויון בהכרח אי־שלילי ולעומת זאת צד שמאל אי־חיובי, כי הסימן נקבע על ידי ה x_i והם תלויים במקור של הi (j לא משחק תפקיד בקביעת הסימן) ולכן בהכרח כל המחוברים בשני הצדדים שווה ל - 0

ובפרט

$$(**) \ \forall i \in S^+, j \in S^{\leq 0} : \ x_i P_{ij} = 0$$

מצד שני

$$\forall i \sum_i x_i = 1$$

ולכן בהכרח קיים i כך ש $x_{i_0} > 0$ כלומר קיים i_0 כך ש $i_0 \in S^+$

ועבור i זה מתקיים $i_0 > 0$ ולכן אם קיים $j_0 \in S^{\leq 0}$ ובנוסף מתקיים $P_{ij} > 0$ נקבל שעבור i_0, j_0

$$x_{i_0} P_{i_0 j_0} > 0$$

בסתירה ל(**)

מסקנה: לכל j "שכן" של i_0, כלומר שקיימת הסתברות חיובית למעבר מi_0 ל־j_0 (כלומר $P_{i_0j_0} > 0$) בהכרח

$$j_0 \in S^+$$

ובסה"כ נקבל שלכל $i \in S^+$ כל "שכן" j גם הוא ב S^+ ומכאן, בגלל האי־פריקות, נמשיך באינדוקציה לשכנים של j ולשכנים שלהם וכן הלאה עד שנגיע לכל המצבים (כאמור, בגלל האי־פריקות) ונקבל שכולם ב S^+ או במילים אחרות

$$\forall i \in S \; x_i > 0$$

■

השלב הבא יהיה להראות שבשרשראות מרקוב אי־פריקות תמיד נתכנס להתפלגות π המקובעת. לשם כך נצטרך לסלק מצב בעייתי מסויים - כאשר יש מחזוריות בשרשרת. במצב שבו יש מחזוריות קבוע נקבל שבהינתן מצד התחלתי (אם לצורך הדוגמה נגדיר שההתפלגות ההתחלתית נותנת הסתברות 1 למצב נתון ולשאר 0) בכל שלב לאחר מכן נקבל מחזוריות של ההתפלגויות (למשל בדוגמה נקבל שבכל שלב נוכל ממש להגיד בדיוק איפה אנחנו אמורים להיות בתוך במחזור).

הגדרה: שרשרת מרקוב היא <u>ארגודית</u> אם היא אי־פריקה ובנוסף (התנאים הבאים שקולים זה לזה):

1. **אי־מחזורית.** כלומר $GCD\left(\{|c|\,|c \;-\; circle\ with\ positive\ probability\}\right) = 1$[2]
2. **קיים** n **כך שלכל** $i, j \in S$ **ו** $t > n$ **:** $Pr\left[x_t = j | x_0 = i\right] > 0$
3. **לכל** $i \in S$ **קיים** $n > 0$ **כך שלכל** $j \in S$ **:** $Pr\left[x_n = j | x_0 = i\right] > 0$

הערה: לא נראה את ההוכחה לשקילות ההגדרות

משפט: תהי $X_0, X_1, ...$ שרשרת מרקוב ארגודית שבכל נקודת זמן t מתפלגת $X_t \sim q^{(t)}$

אזי

$$q^{(t)} \xrightarrow{t\to\infty} \pi$$

[2] או במילים - "נאפשר" מעגלים חיוביים אבל לא באופן כזה שכל המעגלים יהיו מאורך שהוא כפולה של מספר קבוע. נניח אם כל המעגלים מאורך שהוא כפולה של 3 נקבל שיש מחזוריות מאורך 3 בשרשרת ולכן היא לא ארגודית.