מבוא להסתברות

© ארזים

25 במאי 2016

1 שרשראות מרקוב

הגדרה 1.1 נתונים:

1. מרחב מצבים סופי S.
2. התפלגות התחלתית μ על הקבוצה S.
3. מטריצת מעבר $P \in M_{|S|}(\mathbb{R})$. נתייחס לאינדקסים של השורות והעמודות כאיברי הקבוצה S. כלומר, לכל $x, y \in S$, ההסתברות לעבור מהמצב X למצב Y תסומן $P(X,Y)$, ותקיים

$$\begin{aligned} P(X,Y) &\geq 0 \\ \sum_{Y \in S} P(X,Y) &= 1 \end{aligned}$$

שרשרת המרקוב עד זמן n היא סדרת המשתנים המקריים $X_0, \ldots, X_n$ המקבלים ערכים מתוך S, המקיימים שלכל $x_0, \ldots, x_n \in S$ מתקיים

$$\mathbb{P}(X_0 = x, \ldots, X_n = x_n) = \mu(x_0) \cdot P(x_0, x_1) \cdot P(x_1, x_2) \cdot \cdots \cdot P(x_{n-1}, x_n)$$

הערה 1.2

1. לכל נתונים S, μ, P קיימת שרשרת מרקוב.
2. יש עקביות בין ערכי n שונים. אם $X_0, \ldots, X_n$ שרשרת המרקוב עד זמן n וכן $Y_0, \ldots, Y_{n+1}$ שרשרת המרקוב עד זמן $n+1$, אזי ההתפלגות (המשותפת) של $Y_0, \ldots, Y_n$ זהה לשל $X_0, \ldots, X_n$.
בשרשרת מרקוב, העבר בלתי תלוי בעתיד בהנתן ההווה.

סימון כדי להדגיש באיזו התפלגות התחלתית μ אנו עובדים כרגע, נסמן $\mathbb{P}_\mu$. כמו כן, עבור מצב $X \in S$ נסמן $\mathbb{P}_X$ עבור $\mathbb{P}_\mu$ שבו $\mu(X) = 1$. בהתאמה, נסמן גם $\mathbb{E}_\mu$, $\mathbb{E}_X$.

1.1 תכונת מרקוב

יהי $k \geq 1$. תהי $V \subseteq S^k$. יהיו $X_k, X_{k+1}, \ldots, X_n \in S$. אזי

$$\begin{aligned}&\mathbb{P}_\mu\left(X_{k+1}=x_{k+1}, X_{k+2}=x_{k+2}, \ldots, X_n=x_n \mid X_k=x_k,\left(X_0, \ldots, X_{k-1}\right) \in V\right)= \\ &=\mathbb{P}_{X_k}\left(X_1=x_{k+1}, X_2=x_{k+2}, \ldots, X_{n-k}=x_n\right)\end{aligned}$$

הוכחה: לכל התפלגות התחלתית μ, כל עוד $\mathbb{P}_\mu\left(X_k=x_k,\left(X_0, \ldots, X_{k-1}\right) \in V\right)>0$. נסמן בתור ν את ההתפלגות שמקיימת $\nu\left(x_k\right)=1$. כעת,

$$\begin{aligned}
&\mathbb{P}_\mu\left(X_{k+1}=x_{k+1}, X_{k+2}=x_{k+2}, \ldots, X_n=x_n \mid X_k=x_k,\left(X_0, \ldots, X_{k-1}\right) \in V\right) = \\
&= \frac{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k, X_{k+1}=x_{k+1}, X_{k+2}=x_{k+2}, \ldots, X_n=x_n\right)}{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)} = \\
&= \frac{\sum_{\left(x_0, \ldots, x_{k-1}\right) \in V} \mathbb{P}_\mu\left(X_0=x_0, \ldots, X_k=x_k, \ldots, X_n=x_n\right)}{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)} = \\
&= \frac{\sum_{\left(x_0, \ldots, x_{k-1}\right) \in V} \mu\left(x_0\right) P\left(x_0, x_1\right) \cdot \cdots \cdot P\left(x_{k-1}, x_k\right) \cdot P\left(x_k, x_{k+1}\right) \cdot \cdots \cdot P\left(x_{n-1}, x_n\right)}{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)} = \\
&= \frac{\nu\left(x_k\right) \cdot P\left(x_k, x_{k+1}\right) \cdot \cdots \cdot P\left(x_{n-1}, x_n\right) \cdot \sum_{\left(x_0, \ldots, x_{k-1}\right) \in V} \mu\left(x_0\right) P\left(x_0, x_1\right) \cdot \cdots \cdot P\left(x_{k-1}, x_k\right)}{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)} = \\
&= \frac{\nu\left(x_k\right) \cdot P\left(x_k, x_{k+1}\right) \cdot \cdots \cdot P\left(x_{n-1}, x_n\right) \cdot \sum_{\left(x_0, \ldots, x_{k-1}\right) \in V} \mathbb{P}_\mu\left(X_0=x_0, \ldots, X_k=x_k\right)}{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)} = \\
&= \frac{\nu\left(x_k\right) \cdot P\left(x_k, x_{k+1}\right) \cdot \cdots \cdot P\left(x_{n-1}, x_n\right) \cdot \mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)}{\mathbb{P}_\mu\left(\left(X_0, \ldots, X_{k-1}\right) \in V, X_k=x_k\right)} = \\
&= \nu\left(x_k\right) \cdot P\left(x_k, x_{k+1}\right) \cdot \cdots \cdot P\left(x_{n-1}, x_n\right) = \\
&= \mathbb{P}_{x_k}\left(X_0=x_k, X_1=x_{k+1}, \ldots, X_{n-k}=x_n\right) = \\
&= \mathbb{P}_{x_k}\left(X_1=x_{k+1}, \ldots, X_{n-k}=x_k\right)
\end{aligned}$$

■

1.2 קידום התפלגויות

למה 1.3 תהי μ התפלגות התחלתית כלשהי (שנתונה כווקטור שורה). נסמן בתור μ_t את ההתפלגות לאחר t צעדים, כלומר ההתפלגות של X_t, גם היא בווקטור שורה. אזי מתקיים $\mu_t=\mu \cdot P^t$, כאשר הכפל הוא כפל מטריציוני.

הוכחה: נוכיח באינדוקציה על t. עבור $t=0$, הטענה ברורה. כעת, נניח שהוכחנו עבור $t-1$, ונוכיח עבור t. יהי $y \in S$, ונחשב

$$\begin{aligned}
\mu_t(y) &= \mathbb{P}_\mu\left(X_t=y\right)=\sum_{\substack{x \in S \\ \mathbb{P}_\mu\left(X_{t-1}=x\right)>0}} \mathbb{P}_\mu\left(X_{t-1}=x, X_t=y\right)= \\
&= \sum_{\substack{x \in S \\ \mathbb{P}_\mu\left(X_{t-1}=x\right)>0}} \mathbb{P}_\mu\left(X_{t-1}=x\right) \cdot \mathbb{P}_\mu\left(X_t=y \mid X_{t-1}=x\right)= \\
&= \sum_{\substack{x \in S \\ \mu_{t-1}(x)>0}} \mu_{t-1}(x) \cdot P(x, y)=\sum_{x \in S} \mu_{t-1}(x) \cdot P(x, y)= \\
&= \left(\mu_{t-1} \cdot P\right)_y=\left(\mu \cdot P^t\right)_y
\end{aligned}$$

■

הגדרה 1.4 (התפלגות סטציונרית) התפלגות Π על S (כווקטור שורה) נקראת סטציונרית (עמידה) אם $\Pi \cdot P = \Pi$. כלומר, לכל $x \in S$,

$$\mathbb{P}_{\Pi}\left(X_1 = x\right) = \Pi\left(x\right)$$