

# מבוא להסתברות

© ארזים

25 במאי 2016

## 1 שרשראות מרקוב

הגדרה 1.1 נתונים:

1. מרחב מצבים סופי  $S$ .
2. התפלגות התחלתית  $\mu$  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

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

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

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

**הערה 1.2** 1. לכל נתונים  $S, \mu, P$  קיימת שרשרת מרקוב.

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

**סימון** כדי להדגיש באיזו התפלגות התחלתית  $\mu$  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  $\mathbb{E}_X, \mathbb{E}_\mu$ .

## 1.1 תכונת מרקוב

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

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

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

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

■

## 1.2 קידום התפלגויות

**למה 1.3** תהי  $\mu$  התפלגות התחלתית כלשהי (שנתונה כווקטור שורה). נסמן בתור  $\mu_t$  את ההתפלגות לאחר  $t$  צעדים, כלומר ההתפלגות של  $X_t$ , גם היא כווקטור 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

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

■

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

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