Poglavlje 10

Markovljevi lanci unatrag

Markovljevo svojstvo kaže da su budućnost i prošlost uvjetno nezavisni uz danu sadašnjost. Po toj definiciji su prošlost i budućnost simetrične što sugerira da Markovljev lanac promatramo u vremenu koje teče unatrag. S druge strane, po teoremu o graničnoj distribuciji jasno je da distribucija koncentrirana u danom stanju teži prema stacionarnoj distribuciji kada vrijeme teče unaprijed. To pokazuje da potpunu simetriju u vremenu nećemo moći dobiti osim u slučaju da je i početna distribucija lanca stacionarna.

Neka je $X = (X_n : n \geq 0)$ Markovljev lanac sa prostorom stanja S, prijelaznom matricom P i stacionarnom distribucijom π takvom da je $\pi_i > 0$ za sve $i \in S$. Definiramo matricu $\widehat{P} = (\widehat{p}_{ij} : i, j \in S)$ formulom

$$\pi_i \widehat{p}_{ij} = \pi_j p_{ji} \,. \tag{10.1}$$

Budući da je

$$\sum_{j \in S} \widehat{p}_{ij} = \frac{1}{\pi_i} \sum_{j \in S} \pi_j p_{ji} = 1 \,,$$

za sve $i \in S$, vidimo da je $\widehat{P}$ stohastička matrica. Nadalje vrijedi

$$\sum_{j \in S} \pi_j \widehat{p}_{ji} = \sum_{j \in S} \pi_i p_{ij} = \pi_i \sum_{j \in S} p_{ij} = \pi_i \,, \tag{10.2}$$

što znači da je π stacionarna distribucija i za prijelaznu matricu $\widehat{P}$.

Pretpostavimo da je π početna distribucija lanca X. Tada je $\mathbb{P}(X_n = i) = \pi_i$ za sve $n \geq 0$ i sve $i \in S$. Slijedi da je

$$\mathbb{P}(X_n = j \,|\, X_{n+1} = i) = \frac{\mathbb{P}(X_{n+1} = i \,|\, X_n = j)\mathbb{P}(X_n = j)}{\mathbb{P}(X_{n+1} = i)} = \frac{p_{ji}\pi_j}{\pi_i} = \widehat{p}_{ij} \,.$$

Teorem 10.1 *Neka je $X = (X_n : n \geq 0)$ (π, P)-ireducibilan Markovljev lanac gdje je π stacionarna distribucija. Za $N \in \mathbb{N}$ i $0 \leq n \leq N$ definiramo $Y_n = X_{N-n}$. Tada je $Y = (Y_n : 0 \leq n \leq N)$ $(\pi, \widehat{P})$-ireducibilan Markovljev lanac kojem je π također stacionarna distribucija.*

Dokaz: Pokazujemo da je uz početnu distribuciju π, Y Markovljev lanac s prijelaznom matricom $\widehat{P}$. Vrijedi

$$\begin{aligned} & \mathbb{P}(Y_0 = i_0, Y_1 = i_1, \dots, Y_N = i_N) = \mathbb{P}(X_0 = i_N, X_1 = i_{N-1}, \dots, X_N = i_0) \\ = \ & \pi_{i_N} p_{i_N i_{N-1}} \dots p_{i_1 i_0} = \widehat{p}_{i_{N-1} i_N} \pi_{i_{N-1}} p_{i_{N-1} i_{N-2}} \dots p_{i_1 i_0} = \cdots \\ = \ & \widehat{p}_{i_{N-1} i_N} \widehat{p}_{i_{N-2} i_{N-1}} \dots \widehat{p}_{i_0 i_1} \pi_{i_0} = \pi_{i_0} \widehat{p}_{i_0 i_1} \dots \widehat{p}_{i_{N-1} i_N} . \end{aligned}$$

Iz Teorema 1.5 slijedi da je Y Markovljev s prijelaznom matricom $\widehat{P}$. Relacija (10.2) pokazuje da je π stacionarna za Y. Preostaje pokazati da je Y ireducibilan. Za dane $i, j \in S$, po Propoziciji 3.2 postoji niz stanja $i_0 = i, i_1, \dots, i_{n-1}, i_n$ takav da je $p_{i_0 i_1} \dots p_{i_{n-1} i_n} > 0$. Slijedi

$$\widehat{p}_{i_n i_{n-1}} \dots \widehat{p}_{i_1 i_0} = \pi_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n} / \pi_{i_n} > 0 \, ,$$

pa po istoj Propoziciji 3.2 imamo $j \longrightarrow i$. $\square$

Za stohastičku matricu P i mjeru λ kažemo da su u *detaljnoj ravnoteži* ako vrijedi

$$\lambda_i p_{ij} = \lambda_j p_{ji} \quad \text{za sve } i, j \in S \, . \tag{10.3}$$

Pretpostavimo da u stanju i imamo masu veličine λ_i, a u stanju j masu λ_j. Redistribuiramo li te mase pomoću prijelazne matrice P, iz stanja i u stanje j odlazi masa $\lambda_i p_{ij}$, dok iz stanja j u stanje i dolazi jednaka masa $\lambda_j p_{ji}$. To pokazuje da je λ invarijantna mjera za P. Formalno,

Lema 10.2 *Ako su P i λ u detaljnoj ravnoteži, tada je λ invarijantna mjera za P.*

Dokaz: Računamo

$$\sum_{i \in S} \lambda_i p_{ij} = \sum_{i \in S} \lambda_j p_{ji} = \lambda_i \, ,$$

zbog (10.3) i činjenice da je P stohastička matrica. $\square$

Neka je λ distribucija na S. Za ireducibilan (λ, P)-Markovljev lanac $X = (X_n : n \geq 0)$ kažemo da je *reverzibilan*, ako je za sve $N \geq 1$, $(X_{N-n} : 0 \leq n \leq N)$ ponovno (λ, P)-Markovljev lanac.

Teorem 10.3 *Pretpostavimo da je $X = (X_n : n \geq 0)$ (λ, P)-Markovljev lanac koji je ireducibilan. Tada je ekvivalentno:*

(a) X je reverzibilan;

(b) P i λ su u detaljnoj ravnoteži.

Dokaz: Pretpostavimo da vrijedi (a), t.j., X je reverzibilan. Tada je za $N = 1$ proces $Y = (Y_0, Y_1) = (X_1, X_0)$ (λ, P)-Markovljev, što znači da je za $i, j \in S$,

$$\mathbb{P}(X_1 = j, X_0 = i) = \mathbb{P}(Y_0 = j, Y_1 = i) = \lambda_j p_{ji} \, .$$

S druge strane, zbog X (λ, P)-Markovljev, imamo $\mathbb{P}(X_0 = i, X_1 = j) = \lambda_i p_{ij}$. Dakle, P i λ su detaljnoj ravnoteži. Obratno, ako vrijedi (b), tada je

$$\widehat{p}_{ij} = \frac{p_{ji}\lambda_j}{\lambda_i} = p_{ij} \, ,$$

odnosno $\widehat{P} = P$. Zbog Leme 10.2 λ je stacionarna distribucija. Po Teoremu 10.1 je za svaki $N \geq 1$ slučajni proces $(X_{N-n} : 0 \leq n \leq N)$ (λ, P)-Markovljev lanac. □

Tipičan primjer reverzibilnih Markovljevih lanaca su slučajne šetnje na grafovima.

Primjer 10.4 Neka je G povezan, lokalno konačan graf. Preciznije, graf je uređen par $G = (V, E)$ vrhova iz skupa V i bridova iz skupa E. Na skup bridova gledamo kao na podskup Kartezijevog produkta $V \times V$ sa svojstvom $(i, j) \in E$ ako i samo ako je $(j, i) \in E$ (ako brid povezuje vrhove i i j, onda povezuje i vrhove j i i). Povezanost grafa znači da za svaka dva vrha $i, j \in V$ postoji konačan niz bridova $(i, i_1), (i_1, i_2), \ldots, (i_n, j)$ koji povezuju ta dva vrha. Lokalna konačnost znači da iz svakog vrha izlazi najviše konačno mnogo bridova.

Pretpostavimo da je svakom bridu $e \in E$ pridružen strogo pozitivan broj c_e koji zovemo *provodljivost* brida. Dakle, $c : E \to (0, \infty)$. Ako je $e = (i, j) = (j, i)$, pišemo $c_e = c_{ij} = c_{ji}$. Ako $(i, j) \notin E$ stavit ćemo $c_{ij} = 0$. Za vrh $i \in V$ definiramo *kapacitet* tog vrha kao $c_i = \sum_{k \in V} c_{ik}$. Slučajna šetnja na G je Markovljev lanac $X = (X_n : n \geq 0)$ sa skupom stanja V i prijelaznom matricom $P = (p_{ij} : i, j \in V)$ gdje je

$$p_{ij} = \frac{c_{ij}}{c_i} \, .$$

Drugim riječima, šetnja iz vrha i prelazi u jedan od susjednih vrhova s vjerojatnošću proporcionalnom provodljivosti odgovarajućeg brida. Zbog pretpostavke o povezanosti grafa slijedi da je šetnja X ireducibilna. Stavimo $c = \sum_{i \in V} c_i$, te definiramo $\pi = (\pi_i : i \in V)$ sa

$$\pi_i = \frac{c_i}{c} \, .$$

Tada je

$$\pi_i p_{ij} = \frac{c_i}{c} \frac{c_{ij}}{c_i} = \frac{c_{ij}}{c} = \frac{c_{ji}}{c} = \frac{c_j}{c} \frac{c_{ji}}{c_j} = \pi_j p_{ji} \, ,$$

odnosno P i π su u detaljnoj ravnoteži. To pokazuje da je π stacionarna distribucija od X, te da je X reverzibilan.

Primjer 10.5 (Model difuzije Paula i Tatiane Ehrenfest) Sljedeći pojednostavljen model difuzije kroz poroznu membranu predložili su 1907. godine austrijski fizičari Tatiana i Paul Ehrenfest. Svrha modela bila je statističko-mehanički opis izmjene topline između dva sustava različitih temperatura. Predloženi model značajno je pomogao u razumijevanju termodinamičke ireverzibilnosti.

Imamo N čestica od kojih se svaka može nalaziti ili u odjeljku A ili u odjeljku B. Pretposavimo da je u trenutku $n \geq 0$ broj čestica u odjeljku A jednak $X_n = i$, $i = 0, 1, \dots, N$. Na slučajan način izabire se čestica (sve čestice su jednako vjerojatne), te se iz odjeljka u kojem jeste premjesti u drugi odjeljak. To znači da je stanje X_{n+1} jednako $i - 1$ (odabrana čestica bila je u odjeljku A) s vjerojatnosti i/N, ili $i + 1$ (odabrana čestica bila je u odjeljku B) s vjerojatnosti $(N - i)/N$. Odmah se vidi da je $X = (X_n : n \geq 0)$ Markovljev lanac sa skupom stanja $S = \{0, 1, \dots, N\}$ i prijelaznim vjerojatnostima $p_{01} = 1$, $p_{N\,N-1} = 1$, te

$$p_{i\,i-1} = \frac{i}{N}\,, \quad p_{i\,i+1} = \frac{N-i}{N}\,, \quad i = 1, 2, \dots, N-1\,.$$

Očito je da sva stanja komuniciraju, što znači da je lanac X ireducibilan. Po Propoziciji 6.10, sva stanja su povratna. Izračunajmo stacionarnu distribuciju π lanca X. Tražimo mjeru $\lambda = (\lambda_i : i \in S)$ koja zadovoljava uvjet detaljne ravnoteže. Dakle, mora vrijediti:

$$\lambda_i p_{i\,i+1} = \lambda_{i+1} p_{i+1\,i}\,, \quad i = 0, 1, \dots, N-1\,,$$

odnosno

$$\lambda_i \frac{N-i}{N} = \lambda_{i+1} \frac{i+1}{N}\,, \quad i = 0, 1, \dots, N-1\,,$$

Iz gornje relacije slijedi

$$\lambda_{i+1} = \lambda_i \frac{N-i}{i+1}\,, \quad i = 0, 1, \dots, N-1\,.$$

Indukcijom se jednostavno pokaže da je

$$\lambda_i = \lambda_0 \binom{N}{i}\,, \quad i = 1, 2, \dots, N\,.$$

Iz $\sum_{i=0}^{N} \lambda_i = \lambda_0 \sum_{i=0}^{N} \binom{N}{i} = \lambda_0 2^N$ slijedi da je sa

$$\pi_i = \frac{1}{2^N} \binom{N}{i}\,, \quad i = 0, 1, \dots, N\,,$$

dana stacionarna distribucija.

Pretpostavimo da lanac kreće iz stanja $i = 0$ (t.j., sve čestice su u odjeljku B). Kako će izgledati tipično ponašanje lanca? Budući da je na početku vrlo mala vjerojatnost da se izabrana čestica nalazi u odjeljku A, očekujemo početni transfer čestica iz odjeljka B u odjeljak A. Nadalje, razumno je očekivati da će se broj čestica u oba odjeljka ujednačiti (barem u smislu distribucija). Pogledajmo što nam kažu rezultati o graničnoj distribuciji. Budući da lanac X ima period 2, možemo primjeniti Teorem 8.12 po kojem je,

$$\lim_{n\to\infty} p_{0j}^{(2n)} = \begin{cases} \frac{1}{2^{N-1}}\binom{N}{j}, & j \text{ paran} \\ 0, & j \text{ neparan}. \end{cases}$$

Iz svojstva binomnih koeficijenata vidimo da će za velike n s velikom vjerojatnosti broj čestica u oba odjeljka biti podjednak. Dakle, ukoliko zanemarimo probleme koji nastaju zbog periodičnosti, Markovljev lanac X po distribuciji konvergira prema ravnotežnom stanju π. S druge strane, zbog povratnosti će se lanac s vjerojatnosti 1 vratiti u stanje $i = 0$, odnosno, odjeljak A biti će ponovno prazan. U svakodnevnom životu to znači da predloženi model predviđa da će se kockica šećera rastopljena u čaju s vjerojatnosti 1 ponovno kristalizirati u istu kockicu šećera. Model difuzije Paula i Tatiane Ehrenfest objašnjava termodinamičku ireverzibilnost, odnosno činjenicu da nikad ne vidimo ponovnu kristalizaciju kocke šećera.

Zaista, može se pokazati da povratnost nije opaziva za stanja koja su daleko od $M = N/2$, N paran. Na primjer, prosječnan (očekivani) broj koraka potrebnih za doći iz stanja M u stanje 0 ($\mathbb{E}_M T_0$) jednak je

$$\frac{1}{2M}\, 2^{2M}(1 + O(M)), \tag{10.4}$$

dok je prosječan (očekivani) broj koraka za doći iz stanja 0 u stanje M ($\mathbb{E}_0 T_M$) jednak

$$M + M \log M + O(1). \tag{10.5}$$

Ukoliko je $M = 10^6$, a jedinica matematičkog vremena jednaka 10^{-6} sekundi (čestice se premještaju svakih 10^{-6} sekundi), tada je vrijeme potrebno za doći iz praznog odjeljka A do stanja u kojem oba odjeljka imaju jednak broj čestica reda veličine nekoliko sekundi ($10^{-6}(10^6 + 10^6 \log 10^6)$). S druge strane, vrijeme potrebno da bi se odjeljak A ispraznio ako na početku oba odjeljka imaju isti broj čestica je reda veličine

$$\frac{1}{2\cdot 10^{12}}\, 2^{2\cdot 10^6} \sim 4.9 \times 10^{602047}$$

sekundi. Starost svemira procjenjena je na 1.37×10^{10} godina, što iznosi 4.320432×10^{17} sekundi.

Mi nećemo izvesti formule (10.4) i (10.5), već ćemo izračunati očekivana vremena povratka u stanje 0, odnosno u stanje $M = N/2$. Iz Teorema 7.14 i Stirlingove formule dobivamo

$$\begin{aligned} \mathbb{E}_0(T_0) &= \frac{1}{\pi_0} = 2^{2M}, \\ \mathbb{E}_M(T_M) &= \frac{1}{\pi_M} = \frac{2^{2M}}{\binom{2M}{M}} \sim 2^{2M}\frac{(\sqrt{2\pi M}e^{-M}M^M)^2}{\sqrt{2\pi 2M}e^{-2M}(2M)^{2M}} = \sqrt{\pi M}. \end{aligned}$$

Gornja očekivanja sličnog su reda kao i očekivanja u formulama (10.4) i (10.5).