REJTETT MARKOV MODELLEK III.

Gerencsér László   Michaletzky György   Molnár-Sáska Gábor
Vágó Zsuzsanna

(Oktatási segédlet)

Budapest, Eötvös Loránd Tudományegyetem
Valószínűségelméleti és Statisztika Tanszék

2002.

1. ML-becslés konzisztenciája

Legyen $\mathcal{X} = \{1, \ldots, m\}$ egy véges állapottér és (X_n) egy Markov-lánc $\mathcal{X}$-en. Legyen (Y_n) a kiolvasási folyamat, ahol $\Theta : \mathcal{X} \longrightarrow \mathbb{R}^p$ paraméterezés esetén a kiolvasás sűrűségfüggvénye $f_{\Theta_x}(y)$.

Példa. Legyen (X_n) egy független azonos eloszlású valószínűségi változókból álló folyamat és $\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1$ a tér egy diszjunkt felbontása. Valamint legyen

$$\Theta_x = \begin{cases} \Theta_0 & \text{ha } x \in \mathcal{X}_0 \\ \Theta_1 & \text{ha } x \in \mathcal{X}_1. \end{cases}$$

Ekkor Y_n eloszlása nyilván egy keverék eloszlás lesz, és feladatunk ebből a keverék eloszlásból megbecsülni Θ_0-t, Θ_1-t és a keverési arányt.

A továbbiakban kimondunk két feltételt, melyeket a fejezet további részében mindig felteszünk.

1.1. Feltétel. *Az (X_n) Markov-lánc irreducibilis és aperiodikus. Legyen a lánc átmenetvalószínűség mátrixa $A = (a_{jk})$, ahol létezik n, hogy $A^n > 0$.*

1.2. Feltétel. *Tegyük fel, hogy*

$$\int \log^+ f(y, \Theta') f(y, \Theta) dy < \infty.$$

Az (X_n, Y_n) folyamat paraméterei az A átmenetvalószínűség mátrix és a $\Theta_1, \ldots, \Theta_m$ kiolvasási paraméterek. Legyen $\Phi = (a_{jk}, \Theta_i)$, ahol $i, j, k = 1, \ldots, m$. A kérdés tehát az, hogy mit mondhatunk $y_1, \ldots, y_n$ megfigyelések mellett a paraméter maximumlikelihood becsléséről, azaz tudunk-e konvergenciát a

$$\frac{1}{n} \log p(y_1, \ldots, y_n, \Theta)$$

kifejezésnek, ha $n \longrightarrow \infty$.

1.1. Megjegyzés. *Ha $\Phi = \Phi_0$ a valódi paraméter, valamint Y_n véges értékkészletű ergodikus folyamat, akkor a Shannon-McMillan-Breiman tétel szerint*

$$-\frac{1}{n} \log p(y_1, \ldots, y_n, \Phi_0) \longrightarrow H$$

1-valószínűséggel, ahol $H(\Phi_0) = \lim_{n\to\infty}(-\frac{1}{n}E\log p(y_1,\dots,y_n,\Phi_0))$ *az entrópia.*

1.1. Lemma. *Ha* (X_n) *ergodikus folyamat, akkor* (Y_n) *is ergodikus lesz.*

Most tegyük fel, hogy (X_n) stacionárius folyamat, és ezzel együtt $P(X_0 = i) > 0$ teljesül $\forall i$-re. Tekintsük a következő azonosságot:

$$p(y_1,\dots,y_{s+t}|X_1 = j,\Phi) =$$

$$= p(y_1,\dots,y_s|X_1 = j,\Phi)p(y_{s+1},\dots,y_{s+t}|y_1,\dots,y_s,X_1 = j,\Phi) \tag{1}$$

A $p(y_1,\dots,y_{s+t}|X_1 = j,\Phi)$ folyamat nem stacionárius, tehát próbáljuk stacionarizálni. Legyen

$$q_0(y_1,\dots,y_t) = \max_j p(y_1,\dots,y_t|X_1 = j,\Phi).$$

Ekkor mivel

$$p(y_1,\dots,y_t,\Phi) = \sum_{j=1}^{m} p(y_1,\dots,y_t|X_1 = j,\Phi)P(X_1 = j),$$

nyilván

$$q_0(y_1,\dots,y_t)\min_j P(X_1 = j) \le p(y_1,\dots,y_t,\Phi) \le q_0(y_1,\dots,y_t),$$

azaz $\frac{1}{n}\log p(y_1,\dots,y_n,\Phi)$ pontosan akkor konvergens, ha $\frac{1}{n}\log q_0(y_1,\dots,y_n)$ konvergens.

$q_0(y_1,\dots,y_t)$ fenti definíciója alapján

$$p(y_{s+1},\dots,y_{s+t}|X_1,\dots,X_{s+1},Y_1,\dots,Y_s,\Phi) = p(y_{s+1},\dots,y_{s+t}|X_{s+1},\Phi) \le q_0(y_{s+1},\dots,y_{s+t})$$

fennáll minden ω-ra, azaz

$$p(y_{s+1},\dots,y_{s+t}|y_1,\dots,y_s,X_1 = j,\Phi) \le q_0(y_{s+1},\dots,y_{s+t}),$$

így (1) alapján kapjuk, hogy

$$q_0(y_1,\dots,y_{t+s}) \le q_0(y_1,\dots,y_s)q_0(y_{s+1},\dots,y_{s+t}). \tag{2}$$

Legyen $W_{st} = \log q_0(y_{s+1},\dots,y_t)$. Mivel feltételezésünk szerint (Y_n) ergodikus folyamat, így van olyan T mértéktartó transzformáció $(\Omega,\mathcal{F},P)$-n, hogy $y_t(\omega) =$

$y_0(T^t\omega)$ és T-nek csak 0 vagy 1 mértékű invariáns részhalmazai vannak. Így az előző definícióval együtt érvényes, hogy

$$W_{st}(\omega) = W_{0(t-s)}(T^s\omega).$$

Továbbiakban legyen $g_n(\omega) = W_{0n}(\omega)$, és így az eddigiek alapján g_n szubadditív folyamat. Alkalmazzuk a szubadditív ergod-tételt. Ehhez persze még az kell, hogy $Eg_1^+ < \infty$. Azonban

$$g_1 = \log q_0(y_1) = \max_j \log p(y_1|X_1 = j, \Phi),$$

és így

$$g_1^+ = \max_j \log^+ p(y_1|X_1 = j, \Phi).$$

A 1.2 feltétel szerint

$$E_{\Phi_0} \log^+ p(y_1|X_1 = j, \Phi) < \infty,$$

így nyilván a fenti véges maximum is integrálható. Tehát a szubadditív ergod-tétel alapján kapjuk, hogy a

$$\lim_n \frac{1}{n} E_{\Phi_0} W_{0n} = -H(\Phi_0, \Phi)$$

határérték létezik, valamint

$$\lim_n \frac{1}{n} W_{0n} = -H(\Phi_0, \Phi)$$

1-valószínűséggel.

Kérdés, hogy ha X_n indítása nem stacionárius, akkor is létezik-e határérték.

1.1. Divergencia reprezentációja

A további célunk az, hogy megmutassuk, hogy $p(X_n = j|Y_{n-1}, \ldots, Y_1)$ stacionáriussá válik.

Tekintsük az (X_n, Y_n) folyamatot a 1.1 és 1.2 feltételek mellett, valamint legyen az X_n indítása stacionárius. Nyilvánvalóan mindkét folyamat ergodikus lesz.

Egy valós értékű X_n ergodikus folyamat reprezentálható az $\Omega' = \{(x_n) : \ -\infty < n < \infty, x_n \in R\}$ sorozatok terén, a következőképpen. Legyen

$$\mathcal{F}_a^b = \sigma(X_n : \ a \le n \le b)$$

$$\mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}^n_{-\infty}$$
$$\mathcal{F}^{\infty} = \mathcal{F}^{\infty}_{-\infty}.$$

Továbbá legyen T a következő $T : \mathcal{F}^{\infty} \longrightarrow \mathcal{F}^{\infty}$ mértéktartó transzformáció: az

$$A = \{(X_{n_i}) \in \overline{A}, n_i \in I\} \quad \text{eseményre}$$

$$TA = \{(X_{n_{i+1}}) \in \overline{A}, n_i \in I\},$$

ahol $\overline{A}$ megfelelő Borel-halmaz. És az $\omega' \in \Omega'$-re fennáll, hogy $(T\omega')_n = \omega'_{n+1}$, $X_n(\omega') = x_n$, $X_n(\omega) = X_0(T^{-n}\omega')$.

Tekintsük az

$$\mathcal{Y} = \{(y_1, y_2, \dots)\}$$

kiolvasási sorozatok terét, ami metrikus tér. Ezen a téren legyen a folyamat eloszlásmértéke P_{Φ_0}.

Jelölje továbbá

$$\mathbb{S}^m = \{u : \ u \in \mathbb{R}^m, u^i \geq 0, \sum_{i=1}^m u^i = 1\}$$

az m-dimenziós szimplexet. Ezekből az elemekből képzett sorozatok tere legyen

$$\mathcal{U} = \{(u_1, u_2 \dots) : \ u_i \in \mathbb{S}^m\}$$

.

Csak speciális $u \in \mathcal{U}$ sorozatokat fogunk tekinteni, melyeket a Baum-egyenlet generál, azaz

$$u^k_{n+1} = \pi(\sum_{j=1}^m a_{jk} f(y_n, \Theta_{0j}) u^j_n),$$

ahol π a normalizálás, és az indítás

$$u^k_1 = P_{\Phi_0}(X_1 = k)$$

stacionárius, vagyis gyakorlatilag az

$$u^k_n = P(X_n = k | y_{n-1}, \dots, y_1) \quad (n = 1, 2, \dots, \quad k = 1, \dots, m)$$

mennyiségek feltételes valószínűségeket jelentenek.

Általános Φ paraméter mellett is felírhatjuk a Baum-egyenletet:

$$u_{n+1}^k = \pi(\sum_{j=1}^m a_{jk}(\Phi) f(y_n, \Theta_j(\Phi)) u_n^j), \tag{3}$$

ahol

$$u_1^k = P_\Phi(X_1 = k).$$

Az (3) összefüggés generál egy $\mathcal{Y} \longrightarrow \mathcal{U}$ leképezést. Tekintsük az

$$\Omega = \mathcal{Y} \times \mathcal{U}$$

teret, illetve az

$$\mathcal{Y} \longrightarrow \Omega$$

injekciót. Így a P mérték átvihető az Ω térre, melyet jelöljünk P'-vel. Legyen az $y \longrightarrow (y, u)$ leképezés gráfja Ω_{00}. (ez függ persze a Φ paramétertől).

Legyen

$$Y_n : (y_1, y_2, \dots) \longrightarrow y_n$$

az n. koordináta kiolvasása, illetve

$$U_n : (u_1, u_2, \dots) \longrightarrow u_n,$$

mely egy valószínűségi eloszlást generál.

Keressünk olyan $\widetilde{P}$ mértéket Ω-n, melyre $\widetilde{P}$ marginálisa az $\mathcal{Y}$ téren éppen P és az (Y_n, U_n) folyamat stacionárius $\widetilde{P}$-ra.

Tekintsük a $T : \Omega \longrightarrow \Omega$ shift-operátort, melyre

$$(T(y, u))_n = (y_{n+1}, u_{n+1}).$$

P' képe T^k alatt legyen $P'T^{-k}$. Vegyük ezen mértékek

$$\widetilde{P}^n = \frac{1}{n} \sum_{k=1}^n P'T^{-k}$$

számtani átlagát.

1.1. Állítás. *A $\widetilde{P}^n$ mértékcsalád feszes.*

Bizonyítás. A $\widetilde{P}^n$ marginálisa az $\mathcal{Y}$ térre nyilvánvalóan P, hiszen a P shiftre nézve invariáns, így $P'T^{-k}$ marginálisa mindig P lesz. Vegyünk egy $A \subset \mathcal{Y}$ kompakt halmazt, melyre $P(A) > 1-\epsilon$. Ekkor egyrészt az $A \times \mathcal{U}$ halmaz kompakt, hiszen két kompakt halmaz szorzata, másfelől

$$\widetilde{P}^n(A \times \mathcal{U}) = P(A) > 1-\epsilon,$$

azaz a Prohorov tétel szerint a $\widetilde{P}^n$ sorozatnak van gyengén konvergens részsorozata, azaz $\exists l_k$ részsorozat, hogy

$$\widetilde{P}^{l_k} \Rightarrow \widetilde{P}$$

1.2. Állítás. *A $\widetilde{P}$ mérték T-invariáns.*

Bizonyítás. A fenti (l_k) sorozatra fennáll, hogy

$$\widetilde{P}^{l_k} = \frac{1}{l_k}\sum_{i=1}^{l_k} P'T^{-i},$$

$$\widetilde{P}^{l_k}T^{-1} = \frac{1}{l_k}\sum_{i=1}^{l_k} P'T^{-i-1}.$$

Mivel a két összeg mindössze 1-1 tagban különbözik, továbbá $P'T^i$ valószínűségi eloszlást jelöl, ezért

$$\widetilde{P}^{l_k} - \widetilde{P}^{l_k}T^{-1} \Rightarrow 0$$

ha $l_k \longrightarrow \infty$, amiből következik, hogy $\widetilde{P}$ T-invariáns mérték.

Tehát kaptuk a következő következményt.

1.1. Következmény. *Az (Y_n, U_n) folyamat stacionárius a $\widetilde{P}$ mérték alatt.*

Felmerül a kérdés, hogy mi lesz $\widetilde{P}$ tartója, azaz melyik az a legszűkebb F zárt halmaz, amelyre $\widetilde{P}(F) = 1$. Ezt úgy fogjuk jelölni, hogy

$$\mathrm{supp}(\widetilde{P}) = F$$

Legyen

$$\Omega_0 = \{(y,u) \in \Omega,\ \text{amelyekre (3) teljesül tetszőleges } u_1 \text{ indítás mellett}\}.$$

1.3. Állítás.

$$\operatorname{supp}(\widetilde{P}) \subset \Omega_0,$$

azaz $\widetilde{P}$ tartóján mindig teljesül a rekurzió.

Bizonyítás. Nyilván Ω_0 zárt halmaz, és

$$\operatorname{supp}(P'T^{-k}) \subset \Omega_0$$

teljesül minden k-ra. Tehát

$$P'T^{-k}(\Omega_0) = 1 \ \forall k - \text{ra},$$

és így

$$\widetilde{P}^n(\Omega_0) = 1$$

is fennáll. Ismert, hogy ha $\widetilde{P}^{l_k} \Rightarrow \widetilde{P}$, akkor minden F zárt halmazra fennáll, hogy

$$\widetilde{P}(F) \geq \limsup \widetilde{P}^{l_k}(F),$$

és így $F = \Omega_0$-t tekintve kapjuk, hogy

$$\widetilde{P}(\Omega_0) = 1, \text{ azaz}$$

$$\operatorname{supp}(\widetilde{P}) \subset \Omega_0.$$

Most térjünk vissza $-\log p(y_1, \ldots, y_n, \Phi)$ kérdésére. Azt láttuk, hogy ennek n-edrésze stacionárius indítás mellett konvergál. A következő állítás a várhatóértékének kiszámításáról szól.

1.4. Állítás.

$$\frac{1}{n} E_P - \log p(y_1, \ldots, y_n, \Phi) = E_{\widetilde{P}} - \log \sum_{j=1}^{m} f(y_1, \Theta_j) u_1^j,$$

azaz a kifejezés csak az első koordinátától függ.

Bizonyítás. A korábbiakhoz hasonlóan írhatjuk, hogy

$$-\log p(y_1, \ldots, y_n, \Phi) = -\sum_{t=1}^{n} \log p(y_t | y_{t-1}, \ldots, y_1, \Phi) =$$

$$-\sum_{t=1}^{n}\log(\sum_{j=1}^{m} f(y_t,\Theta_j)P(X_t=j|y_{t-1},\ldots,y_1,\Phi)) = -\sum_{t=1}^{n}\log f(y_t,\Theta_j)u_t^j,$$

ahol u_t^j-t (3) összefüggés definiálja és $u_1^j = P(X_1 = j|\Phi)$ a kezdeti feltétel.

Mindkét oldalt szorozzuk $\frac{1}{n}$-nel és integráljunk P' szerint. A bal oldal csak y-tól függ, így a bal oldalra kapjuk, hogy

$$\frac{1}{n}E_P \log p(y_1,\ldots,y_n,\Phi).$$

A jobb oldal egy tagjára

$$\frac{1}{n}E_{P'}\log\sum_{j=1}^{m} f(y_t,\Theta_j)u_t^j = \frac{1}{n}E_{PT^{-t}}\log\sum_{j=1}^{m} f(y_1,\Theta_j)u_1^j,$$

melyet összegezve t-re kapjuk a bizonyítandó állítás jobb oldalát.

Fejezetünk egyik fő tétele a következő.

1.1. Tétel. *Az 1.1 és 1.2 feltételek mellett fennáll, hogy*

$$H(\Phi_0,\Phi) = E_{\widetilde{P}}(-\log\sum_{j=1}^{m} f(Y_1,\Theta_j)U_1^j)$$

Bizonyítás. Először belátjuk, hogy a bal oldal legalább akkora, mint a jobb oldal. Tekintsük az

$$A = \{\omega :\ \omega\in\Omega,\ -\log\sum_{j=1}^{m} f(Y_1,\Theta_j)U_1^j \geq 0\}$$

halmazt. Fatou lemma alapján

$$\int_A -\log\sum_{j=1}^{m} f(y_1,\Theta_j)u_1^j d\widetilde{P} \leq \limsup_k \int_A -\log\sum_{j=1}^{m} f(y_1,\Theta_j)u_1^j d\widetilde{P}^{l_k} \tag{4}$$

Most tekintsük az $A^C = \Omega\backslash A$ komplementerhalmazt. Belátjuk, hogy itt $-\log\sum_{j=1}^{m} f(Y_1,\Theta_j)U_1^j$ $\widetilde{P}^{l_k}$ szerint egyenletesen integrálható. Ehhez tekintsük a

$$-\log\sum_{j=1}^{m} f(y_1,\Theta_j)u_1^j \geq -\log\max_j f(y_1,\Theta_j) \geq -\max_j \log^+ f(y_1,\Theta_j)$$

egyenlőtlenséget és vegyünk egy olyan $B \subset \mathcal{Y}$ halmazt, melyre

$$\int\limits_{B} \max_j \log^+ f(y_1, \Theta_j) dP < \epsilon.$$

Ekkor $B \times \mathcal{U}$-n

$$\int\limits_{B \times \mathcal{U}} \max_j \log^+ f(y_1, \Theta_j) d\widetilde{P}^{l_k} < \epsilon,$$

és így

$$\int\limits_{B \times \mathcal{U}} \log \sum_{j=1}^{m} f(y_1, \Theta_j) u_1^j d\widetilde{P}^{l_k} < \epsilon$$

Tehát az A^C halmazon

$$\int\limits_{A^C} -\log \sum_{j=1}^{m} f(y_1, \Theta_j) u_1^j d\widetilde{P} = \lim_k \int\limits_{A^C} -\log \sum_{j=1}^{m} f(y_1, \Theta_j) u_1^j d\widetilde{P}^{l_k} \tag{5}$$

Egybevetve a (4) egyenlőtlenséget és a (5) egyenlőséget, kapjuk, hogy

$$\int\limits_{\Omega} -\log \sum_{j=1}^{m} f(y_1, \Theta_j) u_1^j d\widetilde{P} \leq \limsup_k \int\limits_{\Omega} -\log \sum_{j=1}^{m} f(y_1, \Theta_j) u_1^j d\widetilde{P}^{l_k}$$

Az utóbbi kifejezés 1.4 miatt nem más, mint

$$\limsup_k \frac{1}{l_k} E_P(-\log p(y_1, \ldots, y_{l_k}, \Phi)) = H(\Phi_0, \Phi)$$

Ezzel az egyik irányú egyenlőtlenséget bebizonyítottuk.

Most tekintsük a másik irányt. Ehhez azt használjuk, hogy ha ξ_n olyan stacionárius sorozat, hogy $E\xi_n^+ < \infty$, akkor

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} \xi_k = \eta$$

1-valószínűséggel valamilyen η-val, melyre $E\eta = E\xi_1$ (a várhatóérték akár $-\infty$ is lehet.

Először is belátjuk, hogy

$$E_{\widetilde{P}} \log^+ \sum_{j=1}^{k} f(Y_1, \theta_j) U_1^j \infty.$$

Ehhez

$$E_{\widetilde{P}} \log^+ \sum_{j=1}^{k} f(Y_1, \theta_j) U_1^j \le E_{\widetilde{P}} \log^+ \max_j f(Y_1, \Theta_j) = E_P \max_j \log^+ f(Y_1, \Theta_j) < \infty.$$

Tehát a fenti ergod-tétel alapján

$$\frac{1}{n} \log \sum_{j=1}^{m} f(Y_1, \Theta_j) U_1^j) \longrightarrow Z$$

1-valószínűséggel $\widetilde{P}$-szerint, és

$$E_{\widetilde{P}} Z = E_{\widetilde{P}}(- \log \sum_{j=1}^{m} f(Y_1, \Theta_j) U_1^j).$$

Tehát elegendő látni, hogy

$$E_{\widetilde{P}} Z \ge H(\Phi_0, \Phi).$$

Az eddigiekben láttuk, hogy ha az Ω térben tekintjük az

$$U_{n+1}^j = \pi(\sum_{i=1}^{m} \alpha_{ij} f(Y_n, \Theta_j) U_n^i), \quad U_1^i = P(X_1 = i)$$

rekurziót, ami nem más, mint a Baum-egyenlet, akkor létezik olyan (Y_n', U_n') stacionárius folyamat $\widetilde{P}$ alatt, mely kielégíti ezt a rekurziót, és vetülete az első koordinátára eloszlásban éppen megegyezik az (Y_n) folyamattal. Ehhez azt láttuk, hogy a $P'T^{-l}$ mértékek gyengén konvergálnak. Most tekintsük az alábbi gyengébb állítást.

1.5. Állítás. *U_l' eloszlása P' alatt gyengén konvergens.*

Bizonyítás. U_l' eloszlása P' alatt megegyezik U_l eloszlásával P alatt, ahol a stacionaritás miatt

$$U_l^j = P(X_l = j | Y_{l-1}, \ldots, Y_1) = P(X_1 = j | Y_0, \ldots, Y_{-l+2})$$

A martingálkonvergencia miatt ez utóbbi konvergens P-majdnem mindenütt, ami bizonyítja állításunkat. Legyen $\mathcal{L}(U_l') \Longrightarrow Q_0$ gyengén, valamint Q az az eloszlás, mely $\widetilde{P}$ marginálisa U_1'-re.

1.6. Állítás. $Q = Q_0$.

Bizonyítás. Legyen B folytonossági halmaz Q-ra és Q_0-ra nézve egyaránt. Mivel $\widetilde{P}_{l_k} \Longrightarrow \widetilde{P}$, így

$$Q(B) = \widetilde{P}(U_1' \in B) = \lim_{k\to\infty} \widetilde{P}_{l_k}(B) = \lim_{k\to\infty} \frac{1}{l_k} \sum_{i=0}^{l_k-1} P'T^{-i}(U_1' \in B) =$$

$$\lim_{k\to\infty} \frac{1}{l_k} \sum_{i=0}^{l_k-1} P'(U_{i+1}' \in B) \longrightarrow Q_0(B),$$

hiszen B a Q_0-nak is folytonossági halmaza. Ha két eloszlás a folyonossági halmazokon megegyezik, akkor mindenütt megegyezik, így állításunkat bebizonyítottuk. Most tekintsük az együttes eloszlásokat.

1.7. Állítás. *Ha B folytonossági halmaza Q-nak és $f(y|\Theta_j)$ folytonos y-ban, akkor $A \in \mathcal{B}(\mathcal{Y})$-ra*

$$\widetilde{P}(Y_1' \in A, U_1' \in B) = \int_B \int_A \sum_j f(y_1|\Theta_j) u_1^j dy_1 dQ(u_1).$$

2. Irreducibilitás

Legyen X_n egy Markov-lánc az $(\mathcal{X}, \mathcal{B})$ absztrakt mérhető állapottéren. Először egy olyan lemmát mondunk ki, mely a későbbi bizonyítások során hasznos segédeszköz lesz:

2.1. Lemma. *(Chapman-Kolmogorov egyenlőség)*

$$P^n(x, A) \geq \int_{\mathcal{X}} P^m(x, dy) P^{n-m}(y, A) d\phi(y)$$

for $x \in \mathcal{X}$ and $A \in \mathcal{B}(\mathcal{X})$.

Legyen $A \in \mathcal{B}$-re

$$\tau(A) = \min\{n \geq 1 : \ X_n \in A\}, \quad \text{és}$$
$$L(x, A) = P_x(\tau(A) < \infty).$$

2.1. Definíció. *X_n ϕ-irreducibilis, ha létezik az $(\mathcal{X}, \mathcal{B})$ téren egy ϕ $(\phi \neq 0)$ σ-véges mérték, melyre $\phi(A) > 0$ esetén $L(x, A) > 0$ teljesül $\forall x \in \mathcal{X}$-re .*

Kimondunk három, a fenti definícióval ekvivalens állítást.

2.2. Lemma. *X_n pontosan akkor ϕ-irreducibilis, ha az alábbi állítások közül bármelyike fennáll:*

1. *$\phi(A) > 0$ esetén $\sum\limits_{n=1}^{\infty} P^n(x, A) > 0 \quad \forall x$-re teljesül.*
2. *$\phi(A) > 0$ esetén $\forall x$-re létezik $n \geq 1$, hogy $P^n(x, A) > 0$ teljesül.*
3. *Legyen $K_{\frac{1}{2}}(x, A) = \frac{1}{2} \sum\limits_{n=0}^{\infty} P^n(x, A) 2^{-n}$. Ekkor $\phi(A) > 0$ esetén $K_{\frac{1}{2}}(x, A) > 0$ $\forall x$-re teljesül.*

2.1. Megjegyzés. *Legyen a fenti lemma 1. pontjában lévő összeg $U(x, A)$, azaz*

$$U(x, A) = \sum_{n=1}^{\infty} P^n(x, A).$$

Ha η_A jelöli az A halmazbeli előfordulások számát, akkor

$$U(x, A) = \sum_{n=1}^{\infty} E\chi_A(X_n) = E(\sum_{n=1}^{\infty} \chi_A(X_n)) = E\eta_A.$$

Bizonyítás. Az egyetlen dolog ami igazán bizonyításra szorul, az az, hogy ha $L(x, A) > 0$ minden $x \in A^C$-re, akkor ez teljesül minden $x \in \mathcal{X}$-re is, azaz a $K_{\frac{1}{2}}$-ben szereplő 0-dik tag nincs befolyással az irreducibilitásra. Ez viszont

$$L(x, A) = P(x, A) + \int\limits_{A^C} P(x, dy) L(y, A)$$

miatt nyilvánvaló.

2.1. Maximális irreducibilitási mértékek

2.1. Tétel. *Tegyük fel, hogy X_n ϕ-irreducibilis Markov-lánc. Ekkor létezik egy olyan ψ valószínűségi mérték $\mathcal{B}(\mathcal{X})$-n, melyre*

1. *X_n ψ-irreducibilis,*
2. *Ha X_n ϕ'-irreducibilis, akkor $\phi' \ll \psi$,*
3. *Ha $\psi(A) = 0$, akkor $\psi(\{y : \ L(y, A) > 0\}) = 0$,*
4. *ψ ekvivalens ψ'-vel, ahol*

$$\psi'(A) = \int_{\mathcal{X}} \phi(dy) K_{\frac{1}{2}}(y, A)$$

Bizonyítás. Mivel ϕ egy ς-véges mérték, így normálással elérhetünk egy olyan mértéket, melyre a lánc továbbra is irreducibilis, és $\phi(\mathcal{X}) = 1$. Tekintsük a ψ' mértéket. Nyilvánvalóan ψ' valószínűségi mérték lesz $\mathcal{B}(\mathcal{X})$-n. Mutassuk meg, hogy teljesülnek rá az $(i) - (iii)$ tulajdonságok. Tekintsük a

$$\overline{A}(k) = \{y : \ \sum_{n=1}^{k} P^n(y, A) > \frac{1}{k}\}$$

halmazokat.

(i): Ha $\psi'(A) > 0$, akkor ψ' definíciója alapján létezik olyan k, melyre $\phi(\overline{A}(k)) > 0$, hiszen az $\overline{A}(k)$ halmazok monoton növőleg tartanak az

$$\{y : \ \sum_{n \geq 1} P^n(y, A) > 0\}$$

halmazhoz. Ez azt jelenti, hogy a ϕ-irreducibilitás miatt rögzített x mellett létezik olyan m, melyre $P^m(x, \overline{A}(k)) > 0$. Ekkor használva a Chapman-Kolmogorov egyenlőséget, kapjuk, hogy

$$\sum_{n=1}^{k} P^{m+n}(x, A) = \int_{\mathcal{X}} P^m(x, dy)(\sum_{n=1}^{k} P^n(y, A)) \geq$$

$$\geq \frac{1}{k} P^m(x, \overline{A}(k)) > 0,$$

mely igazolja a ψ'-irreducibilitást.

(ii): Legyen ϕ' olyan, hogy az X_n lánc egyben ϕ'-irreducibilis. Tehát ha $\phi'(A) > 0$, akkor $\sum_n P^n(y, A) > 0$ minden y-ra, és így $\psi'(A) > 0$, azaz valóban $\phi' \ll \psi'$. Másfelől tegyük fel, hogy a lánc ψ'-irreducibilis és $\phi' \ll \psi'$. Ha $\phi'(A) > 0$, akkor $\psi'(A) > 0$ is teljesül, és így a ψ'-irreducibilitás miatt $K_{\frac{1}{2}}(x, A) > 0$ is fennáll minden $x \in \mathcal{X}$-re, azaz X_n valóban ϕ'-irreducibilis.

(iii): Mivel

$$\int_{\mathcal{X}} \psi'(dy) P^m(y, A) 2^{-m} = \int_{\mathcal{X}} \phi(dy) \sum_n P^{m+n}(y, A) 2^{-(n+m+1)} \leq \psi'(A),$$

így a (iii) tulajdonság is nyilvánvaló.

Megjegyezzük, hogy (ii) alapján a maximális irreducibilis mértékek ekvivalensek.

Legyen

$$\mathcal{B}^+ = \{A \in \mathcal{B} : \ \psi(A) > 0\},$$

ahol ψ a fenti tételben szereplő maximális irreducibilitási mérték.

2.2. Definíció. *Egy $A \in \mathcal{B}$ halmaz teli, ha $\psi(A^C) = 0$, illetve egy $A \in \mathcal{B}$ halmaz elnyelő, ha $P(x, A) = 1$, azaz ha $P^n(x, A) = 1 \quad \forall n \geq 1$ és $\forall x$-re.*

2.1. Állítás. *Ha egy halmaz elnyelő, akkor teli.*

Bizonyítás. Indirekten tegyük fel, hogy A halmaz elnyelő, de nem teli, azaz $\psi(A^C) > 0$. Ekkor $\sum_{n=1}^{\infty} P^n(x, A^C) > 0$, azaz $\exists n \geq 1$, melyre $P^n(x, A^C) > 0$, és így nyilván $P^n(x, A) < 1$, ami ellentmond A elnyelőségének.

2.2. Állítás. *Ha A egy teli halmaz, akkor ő tartalmaz egy pozitív ψ-mértékű elnyelő halmazt.*

Bizonyítás. Legyen

$$B = \{y : \sum_{n=0}^{\infty} P^n(y, A^C) = 0\}.$$

Ekkor nyilván $B \subset A$, és így $\psi(A^C) = 0$ miatt $\psi(B) > 0$ is teljesül.

Tegyük fel, hogy B nem elnyelő halmaz, azaz létezik y, hogy $P(y, B^C) > 0$. Ekkor

$$\sum_{n=0}^{\infty} P^{n+1}(y, A^C) \geq \int_{B^C} P(y, dz) \sum_{n=0}^{\infty} P^n(z, A^C) > 0,$$

hiszen $\sum_{n=0}^{\infty} P^n(z, A^C) > 0$, és így ellentmondást kaptunk.

2.2. Kis halmazok

Legyen X_n továbbra is egy ϕ-irreducibilis Markov-lánc.

2.3. Definíció. *Egy $C \in \mathcal{B}(\mathcal{X})$ halmazt kis halmaznak hívunk, ha $\exists m \geq 1$, hogy $\forall B$-re és $\forall x \in C$-re fennáll, hogy*

$$P^m(x, B) \geq \nu_m(B),$$

ahol ν_m egy nem triviális mérték.

2.2. Tétel. *Ha X_n egy ϕ-irreducibilis Markov-lánc, $\mathcal{B}(\mathcal{X})$ végesen generált σ-algebra és $A \in \mathcal{B}^+$, akkor létezik olyan $C \subset A$ kis halmaz, melyre $\phi(C) > 0$.*

Tekintsük P^n felbontását ϕ szerint, azaz

$$P^n(x, B) = \int_B p^n(x, y) d\phi(y) + P^n_{\perp}(x, B).$$

A 2.2 tételt több lemmán keresztül fogjuk bizonyítani.

2.3. Lemma. *(Chapman-Kolmogorov egyenlőtlenség)*

$$p^{n+m}(x,y) \geq \int\limits_{\mathcal{X}} p^n(x,z)p^m(z,y)d\phi(z)$$

ϕ-majdnem minden y-ra.

Bizonyítás.

$$P^{n+m}(x,B) = \int\limits_{\mathcal{X}} P^n(x,dz)P^m(z,B) \geq$$

$$\int\limits_{z\in\mathcal{X}} \int\limits_{y\in B} p^n(x,dz)p^m(z,y)d\phi(y) =$$

$$\int\limits_{B} (\int\limits_{\mathcal{X}} p^n(x,dz)p^m(z,y))d\phi(y) \geq$$

$$\int\limits_{B} (\int\limits_{\mathcal{X}} p^n(x,z)d\phi(z)p^m(z,y))d\phi(y)$$

Az integrandusokat összehasonlítva kapjuk, hogy

$$p^{n+m}(x,y) \geq \int\limits_{\mathcal{X}} p^n(x,z)p^m(z,y)d\phi(z)$$

ϕ-majdnem minden y-ra. A fentihez hasonlóan nyilván az is igaz, hogy

$$p^{n+m}(x,y) \geq \int\limits_{\mathcal{X}} p^n(x,dz)p^m(z,y)$$

ϕ-majdnem minden y-ra.

Legyen rögzített x-re

$$N(x) = \{y : \sum_{n=1}^{\infty} P^n(x,y) = 0\}.$$

2.4. Lemma.

$$\phi(N(x)) = 0 \quad \forall x \in \mathcal{X} - re.$$

Bizonyítás. Tegyük fel, hogy valamely y-ra fennáll, hogy $\phi(N(y)) > 0$. Ekkor

$$P^n(y, N(y)) = \int_{N(y)} p^n(y,z)d\phi(z) + P^n_{\perp}(y, N(y)).$$

Továbbá

$$P^n(y, N(y) \cup A) = \int_{N(y)\cup A} p^n(y,z)d\phi(z) = 0.$$

Azonban $\phi(N(y) \cup A) > 0$, ami ellentmond azzal, hogy ϕ irreducibilis mérték.

2.5. Lemma. *Feltéve, hogy $\mathcal{B}(\mathcal{X})$ végesen generált, $p^n(x,y)$ választható úgy, hogy mérhető függvény legyen.*

Bizonyítás. Legyenek $\mathcal{B}_i \subset \mathcal{B}_{i+1}$ olyan véges halmazalgebrák, melyekre

$$\sigma(\bigcup \mathcal{B}_i) = \mathcal{B}(\mathcal{X}).$$

Definiáljuk minden i-re

$$\phi_i(x,y) = \begin{cases} \frac{P^n(x, B_i(y))}{\phi(B_i(y))} & \text{ha } \phi(B_i(y)) > 0 \\ 0 & \text{ha } \phi(B_i(y)) = 0 \end{cases},$$

ahol $B_i(y)$ jelöli a $\mathcal{B}_i$ véges halmazalgebrának azt a tagját, melynek y eleme. Minden rögzített x-re

$$\phi_i(x,y) \longrightarrow p^n(x,y) \tag{6}$$

ϕ-majdnem mindenütt, és a $\phi_i(x,y)$ függvények mérhetőek (x,y)-ban, hiszen tetszőleges B Borel-halmazra

$$\{(x,y): \ \phi_i(x,y) \in B\} = \bigcup_k \{(x,y): \ (\phi_i(x,y) \in B) \cap (y \in B_i^k)\}.$$

Legyen $N = \{(x,y): \ \lim \phi_i(x,y) \text{ nem létezik}\}$, ami (6) miatt egy mérhető, nullmértékű halmaz. Tehát kapjuk, hogy $p^n(x,y)$ mérhető függvénynek választható.

2.6. Lemma. *Ha X_n egy ϕ-irreducibilis Markov-lánc, $\mathcal{B}(\mathcal{X})$ végesen generált σ-algebra és $A \in \mathcal{B}^+$, akkor létezik olyan $C \subset A$ halmaz, $M > 1$ és $\delta > 0$, hogy $\phi(C) > 0$ és minden $x, y \in C$ párra*

$$p^m(x,y) > \delta.$$

Bizonyítás. $A \in \mathcal{B}^+$ miatt

$$\sum_{n=1}^{\infty} p^n(x,y) > 0$$

ϕ-majdnem minden $y \in A$-ra és $x \in A$-ra, és így léteznek n, m egész számok, hogy

$$\int_A \int_A \int_A p^n(x,y)p^m(y,z)\phi(dx)\phi(dy)\phi(dz) > 0.$$

Legyen

$$A_n(\eta) = \{(x,y) \in A \times A : \ p^n(x,y) \geq \eta\}.$$

Válasszuk az $\eta > 0$ számot olyan kicsinek, hogy

$$\phi \times \phi \times \phi(\{(x,y,z) \in A \times A \times A : \ (x,y) \in A_n(\eta), (y,z) \in A_m(\eta)\}) > 0$$

teljesüljön. A bizonyítás további részében η-t rögzítettnek tekintjük, és az indexekben el fogjuk hagyni.

Tetszőleges x, y-ra legyen $B_i(x,y) = B_i(x) \times B_i(y)$. A 2.4, 2.5 lemmákhoz hasonlóan léteznek $N_k \subset \mathcal{X} \times \mathcal{X}$ ϕ^2-nullmértékű halmazok, hogy minden k-ra és $(x,y) \in A_k \setminus N_k$-ra

$$\lim_{i \to \infty} \frac{\phi^2(A_k \cap B_i(x,y))}{\phi^2(B_i(x,y))} = 1 \tag{7}$$

Válasszunk egy fix (u,v,w) hármast a

$$\{(x,y,z) : \ (x,y) \in A_n \setminus N_n, \ (y,z) \in A_m \setminus N_m\}$$

halmazból. (7) egyenlőség alapján elég nagy j-re

$$\phi^2(A_n \cap B_j(u,v)) \geq \frac{3}{4}\phi^2(B_j(u,v)), \tag{8}$$

$$\phi^2(A_m \cap B_j(v,w)) \geq \frac{3}{4}\phi^2(B_j(v,w)). \tag{9}$$

Legyen

$$A_n(x) = \{y \in A : \ (x,y) \in A_n\}, \text{ illetve}$$
$$A_m^*(z) = \{y \in A : \ (y,z) \in A_m\}.$$

Végül definiáljuk az

$$E_n = \{x \in B_j(u):\ \phi(A_n(x) \cap B_j(v)) \geq \frac{3}{4}\phi(B_j(v))\},\ \text{illetve}$$

$$D_m = \{z \in B_j(w):\ \phi(A_m^*(z) \cap B_j(v))\}$$

halmazokat. (8) és (9) alapján $\phi(E_n) > 0$ és $\phi(D_m) > 0$. Ezek alapján tetszőleges $(x, z) \in E_n \times D_m$ párra

$$\phi(A_n(x) \cap A_m^*(z)) \geq \frac{1}{2}\phi(B_j(v)) > 0.$$

Továbbá 2.3 lemma szerint

$$p^{n+m}(x, z) \geq \int\limits_{A_n(x) \cap A_m^*(z)} p^n(x, y)p^m(y, z)\phi(dy) \geq$$

$$\geq \eta^2\phi(A_n(x) \cap A_m^*(z)) \geq \frac{\eta^2}{2}\phi(B_j(v)).$$

Mivel $\phi(E_n) > 0$, a ϕ-irreducibilitás miatt létezik olyan k egész szám és $C \subset D_m$ halmaz, hogy $P^k(x, E_n) > 0$ minden $x \in C$-re. Legyen $\frac{\eta^2}{2}\phi(B_j(v)) \geq \delta_1 > 0$, illetve $P^k(x, E_n) \geq \delta_2 > 0$.

Az eddigiek alapján így

$$p^{k+n+m}(x, z) \geq \int\limits_{E_n} P^k(x, dy)p^{n+m}(y, z) \geq \delta_1\delta_2,$$

ami $\delta = \delta_1\delta_2$ és $M = k + n + m$ mellett bizonyítja a lemma állítását.

Ezekután a 2.2 tétel bizonyítása már nyilvánvaló.

Bizonyítás. Tulajdonképpen azt látjuk be, hogy minden $A \in \mathcal{B}^+$ halmazhoz létezik $C \subset A$ kis halmaz, melyre $\phi(C) > 0$. 2.6 lemma szerint azt tudjuk, hogy létezik $n > 1$, $\delta > 0$ és $C \subset A$ halmaz, melyre $p^n(x, y) \geq \delta$ minden $x, y \in C$ esetén. Legyen

$$\nu_n(B) = \frac{\phi(B \cap C}{\phi(C)}.$$

Ekkor

$$P^n(x, B) \geq P^n(x, B \cap C) \geq \int\limits_{B \cap C} p^n(x, y)\phi(dy) \geq \delta\phi(B \cap C) \geq \delta\nu_n(B),$$

azaz a tétel állítását bebizonyítottuk.

2.3. Invariáns mérték

Legyen X_n továbbra is egy Markov-lánc az $(\mathcal{X}, \mathcal{B}(\mathcal{X})$ mérhető alapstruktúrán.

2.4. Definíció. *(Bhattacharya-Waymire) Egy $A_0 \in \mathcal{B}(\mathcal{X})$ halmazra teljesül a lokális Doeblin-feltétel, ha $\exists n$, hogy*

$$P^n(x, B) \geq \delta \nu_n(B)$$

fennáll $\forall x \in A_0$ és $\forall B \in \mathcal{B}(\mathcal{X})$-re, ahol $\delta > 0$ és ν_n egy valószínűségi mérték.

Megjegyezzük, hogy a lokális Doeblin-feltétel ekvivalens azzal, hogy a szóbanforgó halmaz kis halmaz, a Doeblin-feltétel pedig azt jelenti, hogy az egész tér kis halmaz.

Legyen

$\Gamma_C = \{\alpha : \ \alpha : \mathcal{X} \longrightarrow \mathcal{X}$ Borel-mérhető leképezés, hogy $\exists y \in A_0$, hogy $\alpha x = y \quad \forall x \in A_0$−ra$\}$.

Ekkor hasonlóan a Rejtett Markov Modellek I.-ben leírtakhoz,

2.7. Lemma. *X_n-re pontosan akkor teljesül a lokális Doeblin feltétel $m = 1$-re, ha $\exists (\alpha_n)$ független azonos eloszlású reprezentációja a láncnak, hogy*

$$P(\alpha_n \in \Gamma_C) \geq \delta.$$

A fenti feltételhez hasonló a lokális minorizálási feltétel:

2.5. Definíció. *Az $A_0 \in \mathcal{B}(\mathcal{X})$ halmazra teljesül a lokális minorizálási feltétel, ha minden $x \in A_0$-ra és minden $B \subset A_0$ Borel-mérhető halmazra fennáll, hogy*

$$P(x, B) \geq \delta \nu(B),$$

ahol ν valószínűségi mérték az A_0 halmazon és $\delta > 0$.

2.6. Definíció. *Az A_0 halmazt rekurrensnek nevezzük, ha $P(\tau_{A_0} < \infty) = 1$ minden $x \in A_0$-ra, ahol $\tau_{A_0} = \min\{k \geq 1 : \ X_k \in A_0\}$*

2.3. Tétel. *Tegyük fel, hogy az* (X_n) *Markov-láncra teljesül a lokális Doeblin-feltétel* $m = 1$*-re az* A_0 *rekurrens halmazon, valamint*

$$\sup_{x \in A_0} E(\tau_{A_0}(x)) < \infty.$$

Ekkor létezik invariáns mértéke a Markov-láncnak.

Bizonyítás. Vezessük be a következő definíciókat és jelöléseket:

$$\tau_{A_0}^0 = 0,$$

$$\tau_{A_0}^1 = \tau_{A_0},$$

$$\tau_{A_0}^{n+1} = \min\{k > \tau_{A_0}^n : \ X_k \in A_0\}.$$

Továbbá legyen

$$\widetilde{X}_n = X_{\tau_{A_0}^n}.$$

Ekkor nyilván $\widetilde{X}_n$ is Markov lánc lesz (lsd Orey), és belátjuk, hogy teljesül rá a Doeblin-feltétel. Legyen α_n az eredeti X_n Markov lánc reprezentációja, azaz α_n független, azonos eloszlású leképezés sorozat, melyre

$$X_n = \alpha_n X_{n-1}.$$

Definiáljuk a β_1 leképezést:

$$\beta_1 y = \alpha_{\tau_{A_0(y)}^1} \dots \alpha_2 \alpha_1 y, \quad y \in A_0,$$

és legyen

$$\beta_1 \widetilde{X}_0 = \widetilde{X}_1$$

Tekintsük a β_n független, azonos eloszlású sorozatot (ugyanolyan eloszlással, mint β_1). Így egy reprezentációját kapjuk az $\widetilde{X}_n$ Markov-láncnak, azaz $\beta_n \widetilde{X}_{n-1} = \widetilde{X}_n$. Mivel

$$P(\beta_1 \in \Gamma_c) \geq P(\alpha_1 \in \Gamma_c) \geq \delta,$$

így az $\widetilde{X}_n$ sorozatra valóban teljesül a Doeblin-feltétel. Ekkor a Rejtett Markov Modellek I. 14.1 Tétele szerint az $\widetilde{X}_n$ láncnak egyértelműen létezik invariáns eloszlása, melyet jelöljünk π_{A_0}-lal.

Tekintsük az $\widetilde{X}_n$ Markov-lánc átmenetvalószínűségeit. Először tetszőleges $B \subset \mathcal{X}$ Borel halmazra bevezetünk egyfajta tabu valószínűséget:

$$P_{A_0}(x, B) = \sum_{n=1}^{\infty} P_x(X_n \in B, \ X_k \in A^C \text{ ha } 1 \leq k < n),$$

ami tulajdonképpen azt jelenti, hogy átlagosan hányszor voltunk B-ben, mielőtt A_0-ba visszatértünk volna. Könnyű meggondolni, hogy A_0-ra megszorítva $P_{A_0}(x,\cdot)$ valószínűségi mérték lesz A_0-on, és ez éppen az $\widetilde{X}_n$ lánc átmenetvalószínűsége.

Definiáljunk egy π mértéket: tetszőleges $B \subset \mathcal{X}$ Borel halmazra legyen

$$\pi(B) = c \int\limits_{A_0} P_{A_0}(y,B)\pi_{A_0}(dy),$$

ahol $c = \frac{1}{E_{\pi_{A_0}}(\tau_{A_0})}$. Belátjuk, hogy az így definiált π mérték invariáns valószínűségi mérték lesz. Először belátjuk, hogy

$$\int\limits_{\mathcal{X}} P(y,B)\pi(dy) = \pi(B).$$

$$\int\limits_{\mathcal{X}} P(y,B)\pi(dy) = \int\limits_{\mathcal{X}} P(y,B)c\int\limits_{A_0} P_{A_0}(x,dy)\pi_{A_0}(dx) =$$

$$= c\int\limits_{\mathcal{X}}\int\limits_{A_0} P(y,B)\sum_{n=1}^{\infty} P_x(X_n \in dy,\ X_k \in A_0^C \text{ ha} 1 \le k < n)\pi_{A_0}(dx) =$$

$$= c\int\limits_{A_0}\sum_{n=1}^{\infty} P_x(X_{n+1} \in B,\ X_k \in A_0^C \text{ ha} 1 \le k < n)\pi_{A_0}(dx) =$$

$$= c\int\limits_{A_0}\{(P_{A_0}(x,B) - P(x,B)) + \sum_{n=1}^{\infty} P_x(X_{n+1} \in B,\ X_n \in A,\ X_k \in A_0^C \text{ ha} 1 \le k < n)\}\pi_{A_0}(dx) =$$

$$= \pi(B) - c\int\limits_{A_0} P(x,B)\pi_{A_0}(dx) + c\int\limits_{A_0}\int\limits_{A_0} P_{A_0}(x,dy)P(y,B)\pi_{A_0}(dx).$$

Felhasználva, hogy π_{A_0} invariáns mértéke az $\widetilde{X}_n$ láncnak,

$$\int\limits_{A_0} P_{A_0}(x,dy)\pi_{A_0}(dx) = \pi_{A_0}(dy),$$

és így kapjuk, hogy π valóban invariáns mérték lesz $\mathcal{X}$-n. Hátra van még, hogy valószínűségi mérték lesz. Definíció szerint

$$\pi(\mathcal{X}) = c\int\limits_{A_0} P_{A_0}(y,\mathcal{X})\pi_{A_0}(dy).$$

Mivel

$$P_{A_0}(y,\mathcal{X}) = E(\tau_{A_0}(y)),\ \text{és}$$
$$\int\limits_{A_0} E(\tau_{A_0}(y))\pi_{A_0}(dy) = E_{\pi_{A_0}}(\tau_{A_0}),$$

így c definíciójából következik, hogy a fent definiált π valószínűségi mérték lesz.

Most definiáljuk a többlépéses visszatérési valószínűségeket:

$$P^1_{A_0}(x,B) = P_{A_0}(x,B),$$

és rekurzióval $n \geq 1$-re

$$P^{n+1}_{A_0}(x,B) = \int\limits_{A_0} P_{A_0}(y,B)P^n_{A_0}(x,dy). \tag{10}$$

Ha $B \subset A_0$, akkor az $\widetilde{X}_n$ láncra vonatkozó Doeblin-feltétel alapján tudjuk, hogy

$$|P^n_{A_0}(x,B) - \pi_{A_0}(B)| \leq (1-\delta)^{[n/N]}. \tag{11}$$

Ezt terjesszük ki, most az eredeti Markov-láncra.

2.3. Állítás. *Tegyük fel, hogy az X_n Markov-láncra teljesül a lokális Doeblin-feltétel m=1-re az A_0 rekurrens halmazon, valamint*

$$\sup_{x\in A_0} E(\tau_{A_0}(x)) < \infty.$$

Ekkor tetszőleges $B \in \mathcal{B}(\mathcal{X})$ halmazra

$$|cP^{n+1}_{A_0}(x,B) - \pi(B)| \leq c'(1-\delta)^{[n/N]}$$

egyenletesen x-ben és B-ben, ahol $c' > 0$ konstans.

Bizonyítás. Felhasználva a (10) rekurziót,

$$|cP^{n+1}_{A_0}(x,B) - \pi(B)| = |c\int\limits_{A_0} P_{A_0}(y,B)(P^n_{A_0}(x,dy) - \pi_{A_0}(dy))| \tag{12}$$

Az állítás feltétele szerint

$$P_{A_0}(y,B) \leq P_{A_0}(y,S) = E(\tau_{A_0}(y)) \leq \sup_{y\in A_0} E(\tau_{A_0}(y)) < \infty.$$

Legyen $c'' = \sup_{y \in A_0} E(\tau_{A_0}(y))$. Ekkor (12) felülről becsülhető

$$cc'' \int_{A_0} (P^n_{A_0}(x, dy) - \pi_{A_0}(dy) \leq 2cc'' \sup_{B \subset A_0} |P^n_{A_0}(x, B) - \pi_{A_0}(B)|,$$

mely (11) miatt felülről becsülhető a bizonyítandó kifejezéssel.

Az előző állítást felhasználva Cesaro értelemben vett konvergenciát bizonyítunk az eredeti Markov-lánc átmenetvalószínűségeire.

2.4. Állítás. *Ismét tegyük fel, hogy az X_n Markov-láncra teljesül a lokális Doeblin-feltétel m=1-re az A_0 rekurrens halmazon, valamint*

$$\sup_{x \in A_0} E(\tau_{A_0}(x)) < \infty.$$

Ekkor

$$\sup_{x \in \mathcal{X},\ B \in \mathcal{B}(\mathcal{X})} |\frac{1}{N} \sum_{n=1}^{N} P^n(x, B) - \pi(B)| \longrightarrow 0$$

Bizonyítás. Legyen

$$\tau_m \leq N < \tau_{m+1}. \tag{13}$$

Először belátjuk, hogy

$$\frac{1}{N} \sum_{n=1}^{N} \chi_{(X_n \in B)} - \sum_{\tau_1 + 1}^{\tau_m} \chi_{(X_n \in B)} \longrightarrow 0 \quad \text{majdnem mindenütt} \tag{14}$$

Felhasználva, hogy $|\chi_{X_n \in B}| \leq 1$ és hogy a két összeg csak az elejében és a végében különbözik,

$$|\frac{1}{N} \sum_{n=1}^{N} \chi_{X_n \in B} - \sum_{\tau_1 + 1}^{\tau_m} \chi_{X_n \in B}| \leq \frac{\tau_1}{N} + \frac{\tau_{m+1} - \tau_m}{N} \leq \frac{\tau_1}{N} + \frac{\tau_{m+1} - \tau_m}{\tau_m},$$

hiszen (13) teljesül. Az összeg első tagja nyilvánvalóan 0-hoz tart. Tekintsük a második tagot. Vegyük észre, hogy

$$\frac{\tau_m}{m} \longrightarrow E_{\pi_{A_0}}(\tau_{A_0}) = \frac{1}{c}.$$

Így

$$\frac{\tau_{m+1}-\tau_m}{\tau_m}=\frac{\frac{m+1}{m}\frac{\tau_{m+1}}{m+1}-\frac{\tau_m}{m}}{\frac{\tau_m}{m}}$$

kifejezés számlálója 0-hoz, míg nevezője $\frac{1}{c}$-hez tart, azaz (14)-t beláttuk. Legyen

$$Z_k=\sum_{\tau_k+1}^{\tau_{k+1}}\chi_{(X_n\in B)}$$

Ekkor (14)-ban a második összeg

$$\frac{1}{N}\sum_{k=1}^{m-1}Z_k=\frac{m}{N}\frac{1}{m}\sum_{k=1}^{m-1}Z_k$$

alakban írható.

Mivel

$$|\frac{m}{N}-\frac{m}{\tau_m}|=|\frac{1}{N}\sum_{n=1}^{N}\chi_{X_n\in A_0}-\frac{1}{\tau_m}\sum_{n=\tau_1+1}^{\tau_m}\chi_{X_n\in A_0}|\leq\frac{\tau_1}{N}+\frac{\tau_{m+1}-\tau_m}{\tau_m}\longrightarrow 0$$

és $\frac{m}{\tau_m}\longrightarrow c$, így kapjuk, hogy

$$\frac{1}{N}\sum_{n=1}^{N}\chi_{X_n\in B}-c\frac{1}{m}\sum_{k=1}^{m-1}Z_k\longrightarrow 0\quad\text{majdnem mindenütt.}$$

Várhatóértéket képezve, mivel

$$E_{x_{\tau_1}}Z_k=E_{x_{\tau_1}}\sum_{n=\tau_k+1}^{\tau_{k+1}}\chi_{X_n\in B}=P_{A_0}^{k+1}(x_{\tau_1},B)$$

és 2.3 állítás alapján

$$|cP_{A_0}^{k+1}(x_{\tau_1},B)-\pi(B)|\longrightarrow 0,$$

Z_k-k számtani átlagára is teljesül a konvergencia.

Most térjünk rá a π invariáns eloszlás egyértelműségére. Ehhez először tekintsünk egy segédállítást.

2.5. Állítás. *Tegyük fel, hogy π egy invariáns eloszlás, A_0 egy rekurrens halmaz, melyre*

$$\sup_{x\in A_0} E(\tau_{A_0}(x)) < \infty.$$

Ekkor A_0-n teljesül, hogy

$$\pi = c\pi_{A_0}$$

valamilyen c konstanssal.

Az állítás következménye:

2.1. Következmény. *Tetszőleges $B \in \mathcal{B}(\mathcal{X})$-re*

$$\pi(B) \geq c\pi_{A_0}(B).$$

Bizonyítás.

$$\pi(B) = \int_{\mathcal{X}} P(x,B)\pi(dx) \geq \int_{A_0} P(x,B)\pi(dx) =$$

$$c\int_{A_0} P(x,B)\pi_{A_0}(dx) = c\pi_{A_0}(B).$$

2.4. Tétel. *Tegyük fel, hogy az X_n Markov-láncra teljesül a lokális Doeblin-feltétel m=1-re az A_0 rekurrens halmazon, valamint*

$$\sup_{x\in A_0} E(\tau_{A_0}(x)) < \infty.$$

Ekkor a Markov-lánc invariáns eloszlása egyértelmű.

Bizonyítás. Legyen π a lánc egy invariáns eloszlása. A folyamat ergodicitásából következik, hogy minden korlátos, mérhető f függvényre

$$\frac{1}{n}\sum_{j=1}^{n} f(X_j) \longrightarrow \int\limits_{\mathcal{X}} f d\pi$$

P_π-majdnem mindenütt, ahol P_π jelöli az $(X_1, \dots)$ sorozat π által indukált mértékét.

Legyen $f = \chi_{A_0}$. Ekkor

$$\frac{1}{n}\sum_{j=1}^{n} \chi_{X_n \in A_0} \longrightarrow \int\limits_{\mathcal{X}} \chi_{A_0} d\pi = \pi(A_0) \quad P_\pi - \text{majdnem mindenütt.} \tag{15}$$

A $N_n = \#\{j : \; X_j \in A_0\}$ jelölés mellett, így

$$\lim_{n\to\infty} \frac{N_n}{n} = \pi(A_0) \quad P_\pi - \text{majdnem mindenütt.}$$

Legyen

$$Z_m = \sum_{j=\tau_{m-1}+1}^{\tau_m} f(X_j)$$

(15) alapján mivel

$$\frac{1}{n}\sum_{j=1}^{n} f(X_j) - \frac{1}{n}\sum_{j=1}^{\tau_{N_n}} f(X_j) \longrightarrow 0 \quad P_\pi - \text{majdnem mindenütt,}$$

így $\frac{N_n}{n} \to \pi(A_0)$ miatt

$$\frac{1}{N_n}\sum_{m=1}^{N_n} Z_m \longrightarrow \frac{1}{\pi(A_0)} \int\limits_{\mathcal{X}} f d\pi \quad P_\pi - \text{majdnem mindenütt.}$$

A kivételes halmaz legyen N, azaz $P_\pi(N) = 0$. Másfelől Z_m egy stacionárius folyamat $P_{\pi_{A_0}}$ alatt, így

$$\frac{1}{N}\sum_{m=1}^{N} Z_m \longrightarrow E_{\pi_{A_0}}(Z_1) \quad P_{\pi_{A_0}} - \text{majdnem mindenütt.}$$

A kivételes halmaz legyen itt N_0, azaz $P_{\pi_{A_0}}(N_0) = 0$. A segédállításunk 2.1 következménye szerint $\pi > \pi_{A_0}$, így $P_\pi > P_{\pi_{A_0}}$, azaz $P_{\pi_{A_0}}(N \cup N_0) = 0$. Tehát kaptuk, hogy

$$E_{\pi_{A_0}}(Z_1) = \frac{1}{\pi(A_0)} \int\limits_{\mathcal{X}} f d\pi \tag{16}$$

(16)-t alkalmazva $f = 1$ esetén kapjuk, hogy

$$E_{\pi_{A_0}}(\tau_{A_0}) = \frac{1}{\pi(A_0)},$$

és így

$$\int\limits_{\mathcal{X}} f d\pi = \frac{E_{\pi_{A_0}}(Z_1)}{E_{\pi_{A_0}}(\tau_{A_0})}.$$

Mivel az A_0 kis halmazon a π_{A_0} mérték egyértelmű, ,így a tétel állítását bebizonyítottuk.