Diszkrét idejű Markov láncok vizsgálata.

Tekintsünk egy diszkrét idejű $X_0, X_1, \dots$ Markov láncot $P(j,k) = P(X_{n+1} = E_k | X_n = j)$, $n = 1, 2, \dots$, átmenetvalószínűségekkel egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amely bizonyos $E_1, E_2, \dots$ értékeket vesz fel. Elsősorban a következő kérdésekre vagyunk kiváncsiak.

a) Milyen $P(j,k)$ átmenetvalószínűségek esetében mondhatjuk azt, hogy a Markov lánc az E_j állapotból kiindulva 1 valószínűséggel (végtelen sokszor) visszatér az E_j állapotba? (Mint látni fogjuk, ha a Markov lánc 1 valószínűséggel visszatér az E_j állapotba, akkor az is igaz, hogy 1 valószínűséggel végtelen sokszor visszatér oda.)

b) Tekintsük egy $X_0, X_1, \dots$ Markov lánc E_j állapotát, és a $P(n,j,j) = P(X_n = E_j | X_0 = E_j)$ n-lépéses átmenetvalószínűségeket. Létezik-e a $\lim\limits_{n\to\infty} P(n,j,j)$ határérték? Ha létezik meg tudjuk-e adni ezt a határértéket viszonylag egyszerű módon? Léteznek-e és jellemezhetőek-e a $\lim\limits_{n\to\infty} P(n,j,k) = \lim\limits_{n\to\infty} P(X_n = E_k | X_0 = E_j)$ határértékek?

c) Mondhatjuk-e, hogy amennyiben E_j és E_k két olyan állapot, amelyekre teljesül az a tulajdonság, hogy a Markov lánc pozitív valószínűséggel jut el az az E_j állapotból az E_k állapotba, illetve az E_k állapotból az E_j állapotba (alkalmas számú lépésben) akkor a Markov láncot az E_j illetve E_k állapotból elindítva egyszerre igaz vagy nem igaz az, hogy a Markov lánc végtelen sokszor visszatér illetve csak véges sok alkalmommal tér vissza oda; egyszerre léteznek vagy nem léteznek pozitív $\lim\limits_{n\to\infty} P(n,j,j) > 0$ és $\lim\limits_{n\to\infty} P(n,k,k) > 0$ határértékek? Általánosabban, fel tudjuk-e osztani a Markov lánc állapotterét annak alapján, hogy mely állapotból mely állapotba lehet eljutni pozitív valószínűséggel természetes módon osztályokra úgy, hogy az egy osztályban levő elemeknek sok fontos hasonló tulajdonsága van?

E kérdések tárgyalása előtt érdemes bevezetni néhány fogalmat és mennyiséget.

Markov lánc rekurrens és tranziens állapotának fogalma. *Legyen $X_0, X_1, \dots$ egy Markov lánc egy $\mathcal{E} = \{E_1, E_2, \dots\}$ állapottéren. Azt mondjuk, hogy a Markov lánc E_j állapota* rekurrens, *ha a Markov láncot az E_j állapotból indítva, azaz ha $P(X_0 = E_j) = 1$, a Markov lánc 1 valószínűséggel visszatér valamikor az E_j állapotba. Ez azt jelenti, hogy*

$$P\left(\bigcup_{n=1}^{\infty} \{\omega \colon X_n(\omega) = E_j\}\right) = 1.$$

A Markov lánc E_j állapota tranziens, *ha a Markov láncot az E_j állapotból indítva, az 1-nél kisebb valószínűséggel tér vissza valamikor az E_j állapotba, azaz*

$$P\left(\bigcup_{n=1}^{\infty} \{\omega \colon X_n(\omega) = E_j\}\right) < 1.$$

Markov lánc egy állapotának a periódusa. *Legyen $X_0, X_1, \dots$ egy Markov lánc egy $\mathcal{E} = \{E_1, E_2, \dots\}$, állapottéren, és jelölje $P(n,j,k) = P(X_n = E_k | X_0 = E_j)$, az n-lépéses átmenetvalószínűségeket. Vezessük be, az $\mathcal{A}(j) = \{n \colon P(n,j,j) > 0\}$, halmazokat, azaz azon n indexek halmazát, amelyekre a $P(n,j,j)$ átmenetvalószínűség szigorúan pozitív. Azt mondjuk, hogy a Markov lánc E_j állapotának periódusa l, ha az $\mathcal{A}(j)$ halmazban szereplő számok legnagyobb közös osztója l. Ha ez a legnagyobb közös osztó 1, akkor a Markov lánc E_j állapotát aperiódikusnak nevezzük. (Ha a $\mathcal{A}(j)$ halmaz üres, azaz a Markov lánc 1 valószínűséggel soha nem tér vissza az E_j állapotba, akkor nem definiáljuk az E_j halmaz periódusát.*

(Egyszerű) feladat. Egy a d-dimenziós tér egész koordiátájú pontjaiból álló rácson történő bolyongás periódusa 2.

Legyen $X_0, X_1, \dots$ Markov lánc egy $\mathcal{E} = \{E_1, E_2, \dots\}$, állapottéren, és jelölje $P(n,j,k) = P(X_n = E_k | X_0 = E_j)$, az n-lépéses átmenetvalószínűségeket. Vezessük be az alábbi mennyiségeket:

$$f_j(n) = P(X_n = E_j, X_m \neq E_j, \text{ ha } 1 \le m < n | X_0 = E_j), \quad n = 1, 2, \dots, \tag{1}$$

azaz $f_j(n)$ annak a valószínűsége, hogy az E_j állapotból indított Markov lánc n lépés múlva tér vissza először az E_j állapotba. Nyílván az E_j állapot akkor és csak akkor rekurrens, ha $\sum\limits_{n=1}^{\infty} f_j(n) = 1$. Ha az E_j állapot rekurrens, akkor vezessük be a

$$\mu_j = \sum_{n=1}^{\infty} n f_j(n) \tag{2}$$

mennyiségeket is. A μ_j mennyiség egyenlő az E_j állapotba való első visszatérés idejének a várható értékével. Mint majd látni fogjuk, a μ_j mennyiség szoros kapcsolatban van a (létező) $\lim\limits_{n \to \infty} P(n,j,j)$ határértékkel.

Vezessük be a

$$P_j(x) = \sum_{n=0}^{\infty} P(n,j,j) x^n, \quad F_j(x) = \sum_{n=0}^{\infty} f_j(n) x^n \tag{3}$$

hatványsorokat, ahol $P(0,j,j) = 1$, $f_j(0) = 0$ definició szerint. Ezek a hatványsorok konvergálnak $|x| < 1$ esetében. Vegyük észre, hogy $P(n,j,j) = f_j(n)P(0,j,j) + f_j(n-1)P(1,j,j) + f_j(2)P(2,j,j) + \cdots + f_j(1)P(n-1,j,j)$ minden $n = 1, 2, \dots$ számra. Innen, illetve a $P(0,j,j) = 1$ relációból következik, hogy

$$F_j(x) P_j(x) = P_j(x) - 1. \tag{4}$$

A (4) azonosság segítségével be fogjuk látni a következő tételt.

Tétel Markov láncok rekurrens és tranziens állapotainak jellemzéséről. *Legyen $X_0, X_1, \dots,$ Markov lánc, $E_1, E_2, \dots$ állapotokkal, és jelölje $P(n,j,k) = P(X_n =$*

$E_k|X_0 = E_j)$ a Markov lánc átmenetvalószínűségeit. A Markov lánc E_j állapota rekurrens, ha

$$\sum_{n=1}^{\infty} P(n,j,j) = \infty,$$

tranziens, ha

$$\sum_{n=1}^{\infty} P(n,j,j) < \infty.$$

Ha az E_j állapot rekurrens, akkor a Markov lánc 1 valószínűséggel végtelen sokszor tér vissza az E_j állapotba. Ha az E_j állapot tranziens, akkor a Markov lánc 1 valószínűséggel csak véges sokszor tér vissza az E_j állapotba.

Bizonyítás: A (4) reláció, illetve annak a ténynek az alapján, hogy $F_j(x)$ és $P_j(x)$ nem negatív együtthatós hatványsorok, kapjuk, hogy minden $N \ge 1$ számra

$$\sum_{n=1}^{N} f_j(n) \le \lim_{x\to 1} F_j(x) = 1 - \lim_{x\to 1} \frac{1}{P_j(x)} \le \sum_{n=1}^{\infty} f_j(n).$$

Ha $\sum\limits_{n=1}^{\infty} P(n,j,j) = K < \infty$, akkor $\sum\limits_{n=1}^{N} f_j(n) \le 1 - \lim\limits_{x\to 1} \frac{1}{P_j(x)} \le 1 - \frac{1}{K}$ minden $N = 1, 2, \dots,$ számra, ezért $\sum\limits_{n=1}^{\infty} f_j(n) < 1$, és az E_j állapot tranziens. Ha Ha $\sum\limits_{n=1}^{\infty} P(n,j,j) = \infty$, akkor $1 \ge \sum\limits_{n=1}^{\infty} f_j(n) \ge 1 - \lim\limits_{x\to 1} \frac{1}{P_j(x)} = 1$, tehát $\sum\limits_{n=1}^{\infty} f_j(n) = 1$, és az E_j állapot rekurrens.

A tétel bizonyításának befejezéséhez elegendő belátni, hogy mivel $F_j = \sum\limits_{n=1}^{\infty} f_j(n)$ annak a valószínűsége, hogy az E_j állapotból induló Markov lánc legalább 1-szer visszatér az E_j állapotba, F_j^k annak a valószínűsége, hogy a Markov lánc legalább k-szor visszatér az E_j állapotba. Ezt az állítást k szerinti teljes indukcióval fogjuk belátni.

Az indukciós feltevés $k = 1$ esetben érvényes. Jelölje $f_j^{(k)}(n)$ annak a valószínűségét, hogy a Markov lánc az n időpontban tér vissza a k-ik alkalommal az E_j állapotba. Az, hogy az indukciós feltevés érvényes k-ra azt jelenti, hogy $\sum\limits_{n=1}^{\infty} f_j^{(k)}(n) = F_j^k$. Annak a valószínűsége, hogy a Markov lánc legalább $k+1$-szer visszatér az E_j állapotba kifejezhető, mint $\sum\limits_{n=1}^{\infty} \sum\limits_{m=1}^{\infty} f_j^{(k)}(n) f_j(m) = \left(\sum\limits_{n=1}^{\infty} f_j^{(k)}(n)\right) \left(\sum\limits_{m=1}^{\infty} f_j(m)\right) = F_j^{k+1}$. A tétel bizonyítását befejeztük.

Érdemes megfogalmazni a fenti eredmény következő Pólya tétel néven ismert híres következményét.

Pólya György tétele véletlen bolyongások visszatéréséről. *Tekintsük a véletlen bolyongást a d-dimenziós egész rácson, azaz tekintsünk egy olyan $X_0, X_1, \dots,$ Markov* *láncot a d-dimenziós tér egész koordinátájú pontjain, amelyre $P(X_0 = (0, \dots, 0))$, az $X_{n+1} - X_n$ valószínűségi változók függetlenek, $P(X_{n+1} - X_n = e_j) = P(X_{n+1} - X_n = -e_j) = \frac{1}{2d}$, $j = 1, \dots, d$, $n = 0, 1, \dots,$ ahol e_j azt a d-dimenziós vektort jelöli, amelynek j-ik koordinátája 1, és összes többi koordinátája nulla. A véletlen bolyongás $d = 1$ és $d = 2$ dimenzióban 1 valószínűséggel végtelen sokszor visszatér az origóba, $d \geq 3$ dimenzióban 1 valószínűséggel csak véges sokszor tér vissza oda.*

A Pólya tétel bizonyítása. Az előző tétel alapján elég belátni azt, hogy $\sum\limits_{n=1}^{\infty} P(n, 0, 0) = \infty$ egy d-dimenziós véletlen bolyongás átmenet valószínűségeire, ha $d = 1$ vagy $d = 2$, és $\sum\limits_{n=1}^{\infty} P(n, 0, 0) < \infty$, ha $d \geq 3$. Ezen relációk megmutatásához elég belátni, hogy $P(2n, 0, 0) \sim Kn^{-d/2}$ alkalmas $K = K(d) > 0$ együtthatóval minden $n = 1, 2, \dots,$ paraméterre és $d = 1, 2, \dots$ dimenzióra. (Jegyezzük meg, hogy $P(2n+1, 0, 0) = 0$, azaz páratlan sok lépésben nem térhetünk vissza az origóba). Viszont ismert az úgynevezett lokális centrális határeloszlástétel, amely jelen esetben azt fejezi ki, kissé felületesen megfogalmazva, hogy a $P(X_n = k)$ valószínűségek úgy viselkednek, mint ahogy azt a centrális határeloszlástétel sugallja.

Felírhatjuk az $X_n = \sum\limits_{j=1}^{n-1} (X_j - X_{j-1})$ relációt, ahol az összegben független és (ismert) egyforma eloszlású valószínűségi változók szerepelnek. Ez lehetővé teszi, hogy a centrális határeloszlástétel bizonyításához hasonlóan, (valójában egyszerűbben), jó aszimptotikus relációt írjunk fel annak valószínűségére, hogy az X_n valószínűségi változó adott értéket vesz fel. Most csak a számunkra a jelen feladatban érdekes formulát látjuk be. Nevezetesen azt, hogy amennyiben $Y_k = X_k - X_{k-1}$, $k = 1, \dots, 2n$ független d-dimenziós térbeli értékeket felvevő valószínűségi változók, $P(Y_k = e_j) = P(Y_k = -e_j) = \frac{1}{2d}$, $j = 1, \dots, d$, $k = 1, \dots, 2n$, akkor

$$\lim_{n\to\infty} n^{-d/2} P(X_{2n} = 0) \lim_{n\to\infty} n^{-d/2} P(Y_1 + \cdots + Y_{2n} = 0) = \frac{2d^{d/2}}{(2\pi)^{d/2}}, \tag{5}$$

Innen következik Pólya György tétele.

Az (5) reláció bizonyításának részleteit elhagyom, Csak rövid magyarázatot adok arra, honnan lehet látni, hogy egy ilyen aszimptotikus formula érvényes, illetve, mi a bizonyítás alapgondolata. A részletek kidolgozása szorgalmi feladat.

Az (5) képlet bizonyos értelemben a centrális határeloszlástétel lokális alakjának tekinthető. A centrális határeloszlástétel bizonyítása azon múlik, hogy az X_{2n} véletlen független összeg karakterisztikus függvényére, illetve annak normalizáltjára jó aszimptotikus becslést tudunk adni. Jelen esetben a karakteresztikus függvény egy olyan (többváltozós) Fourier sor, amelynek tagjai, $c_n(k_1, \dots, k_d) e^{i(k_1 x_1 + \cdots + k_d x_d)}$ alakú függvények, ahol $k_1, \dots, k_d$ egész számok, és $k_1 + \cdots + k_d$ páros szám. Ezért felhasználva, hogy a Fourier sor tagjaiban szereplő függvények ortogonalitásak a $K = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times [-\pi, \pi]^{d-1}$ d-dimenziós téglatesten, (mely állítást külön igazolni kell,) a $c_n(0, \dots, 0)$ Fourier együtthatót ki lehet fejezni a Fourier sor integráljának a segítségével a K téglatesten. Ez, mivel a Fourier sor értékére jó aszimptotikus formulánk van, lehetővé teszi az (5) képlet igazolását. Ennek a számolásnak a fő lépése annak megmutatása, hogy a tekintendő integrál lényegében az origó egy kis környezetébe van koncentrálva, ahol az integrandusra jó aszimptotikus formulát lehet adni.

Magát a végeredményt előre megsejthetjük. A keresett valószínűség közelítőleg egyenlő a megfelelő kovarianciájú, 0 várható értékű normális sűrűségfüggvény integráljával a K téglatesten. Ráadásul, mivel a kovariancia mátrix (nagy n index esetén) nagy, ezért kis hibát követünk el, ha a K téglatest helyett az egész téren integrálunk.

Rátérek a b) kérdés tárgyalására, annak vizsgálatára, hogy mennyivel egyenlő egy Markov lánc átmeneteinek $\lim\limits_{n\to\infty} P(n,j,j)$ határértéke (feltéve, hogy ez a határérték létezik), illetve az így kapott eredményből milyen következtetéseket tudunk levonni a $P(n,j,k)$ átmenetvalószínűségek aszimptotikus viselkedésére nagy n paraméter esetén. A vizsgálat elején csak aperiódikus E_j állapotokat vizsgálunk. Ha ezek viselkedését jól le tudjuk írni, akkor az általános eset vizsgálata viszonylag egyszerűen visszavezethető erre.

A vizsgálat kulcslépése a valószínűségszámítás egyik érdekes eredményének az úgynevezett felújítási tételnek az alkalmazása. Ezt az eredményt itt nem bizonyítom, csak elmagyarázom, hogy szemléletesen nagyon természetes. (Az érdeklődők megtalálhatják ennek az eredménynek egy lehetséges bizonyítását William Feller: Bevezetés a Valószínűségszámításba című könyvében a XII. fejezet 11. pontjában. Az ismertetett eredmény megegyezik az ott tárgyalt 3. tétellel.)

Felújítási tétel. *Legyenek $Y_1, Y_2, \dots,$ független, egyforma eloszlású valószínűségi változók, amelyek (szigorúan) pozitív egész értékeket vesznek fel. Jelölje $S_n = \sum\limits_{j=1}^{n} Y_j$, $n = 1, 2, \dots,$ a tekintett valószínűségi változók részletösszegeit, és $\mathcal{A} = \{m\colon P(Y_1 = m) > 0\}$, azaz azon egész számok halmazát, melyeket az Y_1 valószínűségi változó pozitív valószínűséggel vesz fel. Tegyük fel, hogy az $\mathcal{A}$ halmazban szereplő számok legnagyobb közös osztója 1. Jelölje továbbá $\mu = \sum\limits_{j=1}^{\infty} jP(Y_1 = j)$ az Y_1 valószínűségi változó várható értékét. Ekkor a következő aszimptotikus formulát kaphatjuk annak valószínűségére, hogy valamelyik S_m részletösszeg felvesz egy előre rögzített nagy n értéket:*

$$\lim_{n\to\infty} P\left(\omega\colon \bigcup_{m=1}^{\infty} S_m(\omega) = n\right) = \frac{1}{\mu}. \tag{6}$$

A (6) reláció érvényes mind $\mu < \infty$, mind $\mu = \infty$ esetben. Utóbbi esetben ez a képlet a $\frac{1}{\infty} = 0$ jelöléssel érvényes.

A felújítási tétel segítségével be fogjuk látni a következő tételt.

Tétel Markov láncok átmenetvalószínűségeinek aszimptotikus viselkedéséről. *Legyen $X_0, X_1, \dots$ olyan Markov lánc $P(n,i,j) = P(X_n = E_j | X_0 = E_i)$ átmenetvalószínűségekkel, amelyben az E_j állapot rekurrens és aperiódikus. Ekkor teljesül a*

$$\lim_{n\to\infty} P(n,j,j) = \frac{1}{\mu_j} \tag{7}$$

azonosság, ahol a μ_j mennyiség a (2) formulában van definiálva.

Először elmagyarázom a felújítási tétel szemléletes tartalmát, majd azt, hogy hogyan lehet ennek segítségével az ezt követő tételt belátni.

Természetes feltenni, hogy annak valószínűsége, hogy a független pozitív egész értékeket felvevő valószínűségi változók $S_1, S_2, \dots$, összegei különböző nagy n értékekre körülbelül ugyanolyan valószínűséggel veszik fel valamilyen indexre az n értéket. (Ahhoz, hogy ez a tulajdonság teljesüljön fel kellett tenni, hogy a tételben definiált $\mathcal{A}$ halmazban szereplő számok legnagyobb közös osztója 1.) Ez azt sugallja, hogy ha megjelöljük azokat a (véletlen) pontokat, amelyeket az $S_1(\omega), S_2(\omega), \dots$ részletösszegek meglátogatnak, akkor a megjelölt pontok halmazának az összes pozitív egész számok halmazában van valamilyen sűrűsége, és a (6) képlet baloldalán szereplő kifejezésnek ez a sűrűség a (létező) limesze. Viszont ezt a sűrűséget könnyen kiszámolhatjuk. Ugyanis a nagy számok törvénye alapján $\frac{S_n}{n}$ értéke tart a μ számhoz 1 valószínűséggel, ha $n \to \infty$. Ez azt jelenti, hogy nagy n indexre van egy olyan $[1, A_n]$ intervallum, amelyre $A_n \sim n\mu$, és az $[1, A_n]$ intervallumban pontosan n megjelölt pont van. Ez azt jelenti, hogy a megjelölt pontok sűrűsége $\frac{1}{\mu}$.

A fenti heurisztikus gondolatmenet mutatja, miért hihető, hogy a felújítási tétel igaz. Másrészt, mint látni fogjuk, abban az esetben, ha egy olyan az E_j állapotból induló Markov láncot tekintünk, amelynek E_j állapota rekurrens és aperiódikus, akkor az E_j állapot egymást követő látogatásai között eltelt időszakaszok hosszai olyan független, egyforma eloszlású valószínűségi változók, amelyek teljesítik a felújítási tétel feltételeit $\mu = \mu_j$ várható értékkel. Ezen észrevétel segítségével megkapjuk a (7) formula bizonyítását. A részletek kidolgozása érdekében érdemes bevezetni és bebizonyítani diszkrét idejű Markov láncokra az úgynevezett erős Markov tulajdonságot. Ez egyszerű, de fontos tulajdonság.

Erős Markov tulajdonság definiciója diszkrét idejű Markov láncokra. *Legyen $X_0, X_1, \dots$ (stacionárius) Markov lánc. Azt mondjuk, hogy a Markov lánc teljesíti az erős Markov tulajdonságot, ha a Markov lánc tetszőleges olyan τ megállási szabályára, amelyre $P(\tau(\omega) < \infty) = 1$*

$$\begin{aligned} &P(X_{\tau+1} = E_{u_1}, X_{\tau+2} = E_{u_2}, \dots, X_{\tau+j} = E_{u_j} | \tau = n, X_0 = E_{v_0}, \dots, X_\tau = E_{v_n}) \\ &\qquad = P(X_1 = E_{u_1}, X_2 = E_{u_2}, \dots, X_j = E_{u_j} | X_0 = E_{v_n}) \end{aligned} \tag{8}$$

minden olyan n, $u_1, \dots, u_n$ és $v_1, \dots v_k$ számokra, amelyekre a (8) képlet baloldalán szereplő feltétel nem nulla valószínűségű esemény.

A Markov tulajdonság (stacionárius) Markov láncokra azt mondja ki, hogy rögzítve egy n időpontot, és egy trajektóriát a $[0, n]$ intervallumon egy Markov folyamatnak az n időpont utáni viselkedése, feltéve, hogy a $[0, n]$ időintervallumban az előírt trajektóriát járta be ugyanolyan, mint egy olyan Markov folyamaté, amely a 0 időpontban ennek a trajektóriának a végpontjából indul. Az erős Markov tulajdonság ezt a tulajdonságot fogalmazza meg abban az általánosabb esetben, amikor a Markov folyamatnak nem egy determinisztikus, hanem egy véletlen megállási szabály által definiált időpontja utáni viselkedését kívánjuk leírni. Bár ezt a fogalmat csak diszkrét idejű Markov láncokra fogalmaztuk meg, az erős Markov tulajdonság fogalmát lehet (sőt érdemes) definiálni általános Markov folyamatokra is. Számunkra az erős Markov tulajdonságnak az alább definiált τ_n, $n = 1, 2, \dots$, megállási szabályokra vett alkalmazása lesz elsősorban érdekes.

Legyen $X_0, X_1, \dots$, stacionáris Markov lánc, $P(X_0 = E_j) = 1$, és legyen E_j a Markov lánc egy rekurrens állapota. Legyen

$$\tau_1 = \tau_1(E_j) = \min\{k\colon k > 0,\ X_k = E_j\}. \tag{9}$$

Definiáljuk ezután a τ_n, $n = 1, 2, \dots$, megállási szabályokat az n szám szerinti teljes indukcióval a következő módon. Ha τ_n-et már definiáltuk, akkor legyen

$$\tau_{n+1} = \tau_{n+1}(E_j) = \min\{k\colon k > \tau_n,\ X_k = E_j\}. \tag{9$'$}$$

Szavakkal megfogalmazva, τ_n az E_j állapotba való n-ik visszatérés időpontja. Belátjuk a következő lemmát.

Lemma. *Egy $X_0, X_1, \dots$, Markov lánc teljesíti az erős Markov tulajdonságot. Speciálisan, ha E_j a Markov lánc rekurrens állapota, és $P(X_0 = E_j) = 0$, akkor a (9) és (9$'$) formulákban definiált τ_n, $n = 1, 2, \dots$, megállási szabályokra az $Y_1 = \tau_1$, $Y_k = \tau_k - \tau_{k-1}$, $k = 2, 3, \dots$, valószínűségi változók függetlenek és egyforma eloszlásúak, $P(Y_k = N) = f_j(N)$, ahol $f_j(N)$ a (1) képletben van definiálva.*

A Lemma bizonyítása. Mivel τ megállási szabály, a (8) formula baloldalán szereplő $\{\tau = n, X_0 = E_{v_0}, \dots, X_\tau = E_{v_n}\}$ feltételben definiált esemény részeként szereplő $\{\tau = n\}$ esemény be vagy be nem következtét az $\{X_0 = E_{v_0}, \dots, X_n = E_{v_n}\}$ esemény determinisztikusan meghatározza. Ezért ez a feltétel vagy az üres vagy az $\{X_0 = E_{v_0}, \dots, X_n = E_{v_n}\}$ esemény. Az utóbbi esetben a (8) képlet baloldalán szereplő kifejezés a $P(X_{n+1} = E_{u_1}, X_{n+2} = E_{u_2}, \dots, X_{n+j} = E_{u_j} | X_0 = E_{v_0}, \dots, X_n = E_{v_n})$ feltételes valószínűséggel egyenlő. Ez viszont a (stacionárius) Markov tulajdonság miatt egyenlő a jobboldali kifejezéssel.

A Lemma második állításának bizonyításához elég megmutatni, hogy minden n pozitív, $N_1, N_2, \dots, N_n$ és N nem-negatív egész számokból álló számrendszerre

$$P(\tau_{n+1} - \tau_n = N | \tau_1 = N_1, \tau_2 - \tau_1 = N_2, \dots, \tau_n - \tau_{n-1} = N_n) = P(\tau_1 = N | X_0 = E_j),$$

illetve észrevenni azt, hogy $P(\tau_1 = N | X_0 = E_j) = f_j(N)$ definició szerint. Másrészt $\{\tau_{n+1} - \tau_n = N\}$ eseményt felírhatjuk, mint $\{X_{\tau_n+1} = E_{u_1}, \dots, X_{\tau_n+N} = E_{u_N}\}$ alakú diszjunkt események unióját, és hasonlóan a $\{\tau_1 = N\}$ esemény felírható, mint $\{X_1 = E_{u_1}, \dots, X_N = E_{u_N}\}$ alakú diszjunkt események uniója. Ezért a bizonyítandó azonosság igazolásához elég megmutatni azt, hogy

$$\begin{aligned} &P(X_{\tau_n+1} = E_{u_1}, \dots, X_{\tau_n+N} = E_{u_N} | \tau_1 = N_1, \tau_2 - \tau_1 = N_2, \dots, \tau_n - \tau_{n-1} = N_n) \\ &\quad = P(X_1 = E_{u_1}, \dots, X_{\tau_n+N} = E_{u_N} | X_0 = E_j) \end{aligned}$$

minden lehetséges $E_{u_1}, \dots, E_{u_N}$ állapotokra.

Viszont az utolsó bizonyítandó azonosság baloldalán szereplő $\{\tau_1 = N_1, \tau_2 - \tau_1 = N_2, \dots, \tau_n - \tau_{n-1} = N_n\}$ feltétel felbontható $\{\tau_n = N_1 + \cdots + N_n, X_0 = E_{u_0}, X_1 = E_{u_1}, \dots, X_{N_1+\cdots+N_n} = E_j\}$ alakú diszjunkt események uniójára. (Fontos, hogy az utolsó egyenlet jobboldalán az E_j állapot szerepel.) Ezért elég megmutatni, hogy a bizonyítani kívánt azonosság érvényes, ha az azonosság baloldalán szereplő feltételes valószínűségben szereplő feltételt ilyen alakú feltétellel helyettesítjük. Ez viszont a (8) formulában szereplő erős Markov tulajdonság következménye. A lemmát bebizonyítottuk.

A Markov láncok átmenetvalószínűségeinek aszimptotikus viselkedéséről szóló tétel egyszerű következménye a fenti lemmának és a felújítási tételnek. Valóban, tekintsük a lemma második állításában szereplő τ_n megállási szabályokat. Ezek felírhatóak $\tau_n = \sum\limits_{k=1}^{n} Y_k$ alakban, ahol $Y_1, Y_2, \dots$ független egyforma eloszlású valószínűségi változók, és az az esemény, hogy az E_j pontból induló bolyongás az n-ik időpontban visszatér az E_j állapotba azzal az eseménnyel egyezik meg, hogy az $\tau_1(\omega), \tau_2(\omega), \dots,$ visszatérési idők valamelyike felveszi az n értéket. Mivel ezeket a valószínűségi változókat előállítottuk, mint független, egyforma eloszlású $Y_1, Y_2, \dots$ valószínűségi változók részletösszegeit, azt kell meggondolnunk, hogy ezekre alkalmazható a felújítási tétel, és az az általunk megfogalmazott eredményt adja.

Valóban, megadtuk az Y_1 valószínűségi változó eloszlását, és abból látszik, hogy Y_1 várható értéke a (2) formulában definiált μ_j szám. Másrészt azon pozitív egész halmazának, amelyeket az Y_1 valószínűségi változó pozitív valószínűséggel vesz fel, (azaz a $\mathcal{A}$ halmazban szereplő számoknak) a legnagyobb közös osztója 1 az E_j állapot aperiódikus tulajdonsága miatt. Valóban, ha ezen számok mindegyike osztható volna valamely $d \ge 2$ számmal, akkor a τ_n valószínűségi változók értékei 1 valószínűséggel oszthatók lennének d-vel minden $n = 1, 2, \dots$ számra, és ez azt jelntené, hogy az E_j állapot periódusa osztható d-vel. Így a felújítási tétel megadja a kívánt állítást.

Bevezetem a következő definiciókat.

Markov lánc rekurrens null állapotának a definiciója. *Legyen $X_0, X_1, \dots$ Markov lánc egy $\mathcal{E} = \{E_1, E_2, \dots\}$ állapottéren. Azt mondjuk, hogy a Markov lánc egy E_j állapota rekurrens null állapot, ha az E_j állapotból indított Markov lánc egy valószínűséggel visszatér az E_j állapotba, de a visszatérés idejének a várható értéke végtelen.*

Markov lánc pozitív állapotának a definiciója. *Legyen $X_0, X_1, \dots$ Markov lánc egy $\mathcal{E} = \{E_1, E_2, \dots\}$ állapottéren. Azt mondjuk, hogy a Markov lánc egy E_j állapota pozitív állapot, ha az E_j állapotból indított Markov lánc egy valószínűséggel visszatér az E_j állapotba, és a visszatérés idejének a várható értéke véges. Egy pozitív aperiódikus állapotot ergodikusnak nevezünk.*

Feladat: Láttuk martingálok segítségével, hogy egy a nulla pontból kiinduló 1 dimenziós bolyongás 1 valószínűséggel eljut az 1 pontba, de az eljutás idejének várható értéke végtelen. Mutassuk meg, hogy ez az állítás következik a Markov láncokról tanult eredményekből is.

Segítség: Szimmetria okok miatt a 0-ból az 1-be eljutás idejének a várható értéke megegyezik az 1 vagy -1-ből a nullába való eljutás várható értékével. Másrészt a nullába való visszatérés idejének a várható értéke 1 plusz 1/2-szer az 1-ből nullába való jutás idejének a várható értéke plusz 1/2-szer a -1-ből a nullába jutás idejének a várható értéke.

A következő tételben megfogalmazott eredmény részben összefoglaló jellegű. Ebben felsoroljuk azokat a már bizonyított eredményeket, amelyek a $P(n,j,k)$ átmenetvalószínűségek segítségével jellemzik a tranziens, rekkurens null és pozitív állapotokat. Ezenkívül nagy n idő esetén, aszimptotikus formulát is adunk a $P(n,i,j)$ átmenetvalószínűségekre, azaz annak a valószínűségére, hogy az E_j állapotba jutunk az (attól esetleg különböző) E_i állapotból. Ahhoz, hogy ezeket az eredményeket megkapjuk, először bevezetek néhány jelölést, és megadok néhány egyszerű, de hasznos formulát.

Vezessük be az

$$f_{i,j}(n) = P(X_n = E_j, X_m \neq E_j, \text{ ha } 1 \le m < n | X_0 = E_i), \quad n = 1, 2, \dots, \tag{10}$$

és

$$F(i,j) = \sum_{n=1}^{\infty} f_{i,j}(n) \quad n = 1, 2, \dots, \tag{10'}$$

mennyiségeket. Az $f_{i,j}(n)$ szám annak a valószínűsége, hogy az E_i állapotból elinduló Markov lánc az n időpontban jut először az E_j állapotba, és $F_{i,j}$ annak a valószínűsége, hogy az E_i állapotból elindított Markov lánc valamikor eljut az E_j állapotba.

Igaz a következő azonosság:

$$P(n,i,j) = \sum_{l=1}^{n} f_{i,j}(l) P(n-l,j,j), \tag{11}$$

mert $f_{i,j}(l)P(n-l,j,j)$ annak a valószínűsége, hogy az E_i állapotból kiinduló Markov lánc az l-ik lépésben veszi fel először az E_j értéket, $1 \le l \le n$, és ezután $n-l$ lépésben az E_j állapotból visszajut az E_j állapotba. Most megfogalmazom a következő eredményt.

Tétel egy Markov lánc egy állapotának jellemzéséről. *Legyen $X_0, X_1, \dots$ Markov lánc valamely $\mathcal{E} = \{E_1, E_2, \dots\}$ állapottéren $P(n,i,j) = P(X_n = E_j | X_0 = E_i)$ átmenetvalószínűségekkel.*

a) Az E_j állapot akkor és csak akkor tranzitív, ha

$$\sum_{n=1}^{\infty} P(n,j,j) < \infty.$$

Ebben az esetben

$$\sum_{n=1}^{\infty} P(n,i,j) < \infty \quad \textit{minden } E_i \textit{ állapotra.}$$

b) Az E_j állapot akkor és csak akkor rekurrens null állapot, ha

$$\sum_{n=1}^{\infty} P(n,j,j) = \infty \quad \textit{és} \quad \lim_{n\to\infty} P(n,j,j) = 0.$$

Ebben az esetben

$$\lim_{n\to\infty} P(n,i,j) = 0 \quad \textit{minden } E_i \textit{ állapotra.}$$

c) Egy aperiódikus E_j állapot akkor és csak akkor ergodikus, (azaz akkor és csak akkor teljesül a $\lim\limits_{n\to\infty} P(n,j,j) = \mu_j > 0$ reláció alkalmas $\mu_j > 0$ számmal, ha $\mu_j = \sum\limits_{n=1}^{\infty} n f_j(n) < \infty$, ahol az $f_j(n)$ mennyiség az (1) képlettel van megadva. A két képletben szereplő μ_j szám megegyezik. Ha az E_j állapot ergodikus, akkor

$$\lim_{n\to\infty} P(n,i,j) = F(i,j)\mu_j \quad \textit{minden } E_i \textit{ állapotra.} \tag{12}$$

Az ebben a képletben szereplő $F(i,j)$ számot a $(10')$ képletben definiáltuk.

Markov lánc egy állapotának jellemzéséről szóló tétel bizonyítása. Az a) rész első állítását a Markov láncok rekurrens és tranziens állapotainak jellemzéséről szóló tétel tartalmazza. Az a) rész második állítása következik az a) rész első állításából és a (11) formulából, mert ebből következik, hogy

$$\sum_{n=1}^{\infty} P(n,i,j) = \sum_{n=1}^{\infty}\sum_{l=1}^{n} f_{i,j}(l)P(n-l,j,j) = \left(\sum_{l=1}^{\infty} f_{i,j}(l)\right)\left(\sum_{m=1}^{\infty} P(m,j,j)\right) < \infty.$$

A b) rész első állítása következik a Markov láncok rekurrens és tranziens állapotainak és a Markov láncok átmenetvalószínűségeinek aszimptotikus viselkedéséről szóló tételekből. Ez utóbbi eredmény csak aperiódikus (azaz 1 periódusú) állapotokról szól. Viszont, ha az X_j állapot periódusa $d \ge 2$, akkor tekinthetjük az $\bar{X}_j = X_{dj}$, $j = 0,1,\dots$, $P(\bar{X}_j = 0) = 1$, Markov láncot. Ennek periódusa 1, ezért erre alkalmazva a Markov láncok átmenetvalószínűségeinek aszimptotikus viselkedéséről szóló eredményt kapjuk, hogy $\lim\limits_{n\to\infty} P(X_{nd} = E_j) = 0$. Másrészt $P(X_{nd+r} = E_j) = 0$, ha $0 < r < d$. Ezért a b) rész első állítása igaz az általános esetben.

A b) rész második állítása következik annak első feléből és a (11) formulából, mert

$$P(n,i,j) \le \sup_{m \ge \frac{n}{2}} P(m,j,j) \sum_{l=1}^{n/2} f_{i,j}(l) + \sum_{l=n/2}^{\infty} f_{i,j}(l),$$

és a b) eset feltételei mellett mind a két összeg nagyon kicsi nagy n indexre.

A c) résznek is csak a második állítása új. Az elsőt tartalmazza a Markov láncok átmenetvalószínűségeinek aszimptotikus viselkedéséről szóló tétel. A hiányzó részt az első rész és a (11) formula segítségével láthatjuk be hasonlóan a b) rész indoklásához. Valóban

$$P(n,i,j)=\sum_{l=1}^{n/2} f_{i,j}(l)P(n-l,j,j)+\sum_{l=n/2+1}^{n} f_{i,j}(l)P(n-l,j,j)=\Sigma_1(n)+\Sigma_2(n).$$

Viszont

$$\lim_{n\to\infty}\Sigma_1(n)=\lim_{n\to\infty}\sum_{l=1}^{n/2} f_{i,j}(l)\mu_j=F(i,j)\mu(j),$$

és

$$0\le\limsup_{n\to\infty}\Sigma_2(n)\le\lim_{n\to\infty}\sum_{l=n/2}^{\infty} f_{i,j}(l)=0,$$

ahonnan következik az állítás.

Megadom egy Markov lánc határátmenetvalószínűségeinek egy más tipusú jellemzését is. Ennek alapgondolata a következő. Ha megadjuk a Markov lánc egy olyan kezdeti (úgynevezett stacionárius) eloszlását, amely az idő során nem változik, akkor bizhatunk abban, hogy ennek eloszlásai megegyeznek az előző tételben megadott határértékekkel. Íly módon ki tudjuk számolni az $\frac{1}{\mu_j}$ határértékeket anélkül, hogy a közvetlenül nehezen kiszámítható várható értékeket kellene meghatároznunk. Ehelyett egy (esetleg végtelen) lineáris egyenletrendszert kell megoldanunk. E módszer tárgyalásában szükségünk van az alábbi definicióra.

Markov lánc stacionárius eloszlásáának a definiciója. *Legyen $X_0,X_1,\dots,$ Markov lánc valamely $\mathcal{E}=\{E_1,E_2,\dots\}$ állapottéren $P(i,j)=P(X_1=E_j|X_0=E_i)$ egy lépéses átmenetvalószínűségekkel. Azt mondjuk, hogy egy $u_1\ge 0, u_2\ge 0,\dots,$ nem negatív számokból álló sorozat a Markov lánc stacionárius eloszlását adja meg, ha ez a sorozat teljesíti a*

$$\sum_{j:E_j\in\mathcal{E}} u_j=1 \tag{13}$$

$$u_j=\sum_{i:E_i\in\mathcal{E}} u_iP(i,j)\quad \textit{minden } j=1,2,\dots \textit{ számra.} \tag{14}$$

egyenleteket.

A (13) és (14) formuláknak természetes szemléletes tartalma van. Ha olyan Markov láncot tekintünk, amelynek 0 időpontbeli eloszlása $P(X_0=E_j)=u_j$ minden $j=1,2,\dots$ számra, akkor a (13) képlet alapján a Markov lánc 1 időpontbeli eloszlása $P(X_1=E_j)=\sum\limits_{i:E_i\in\mathcal{E}} P(X_0=E_i)P(X_1=E_j|X_0=E_i)=\sum\limits_{i:E_i\in\mathcal{E}} u_iP(i,j)=u_j$ minden $j=1,2,\dots$ számra. Ez azt jelenti, hogy a Markov láncnak ugyanaz a $P(X_1=E_j)=u_j$,

$j = 1, 2, \ldots$, az eloszlása az $n = 1$ időpontban. De ekkor indukcióval kapjuk, hogy ugyanez a Markov lánc eloszlása minden $n = 1, 2, \ldots$ időpontban.

A korábbi eredmények alapján természetes azt várni, hogy bizonyos nem túl megszorító feltételek teljesülése esetén a Markov láncnak egyetlen stacionárius eloszlása van, amelyet az $u_j = \frac{1}{\mu_j}$ képlettel adhatunk meg. Azt várjuk, hogy ez a számsorozat kielégíti a (13) és (14) képleteket, valamint a Markov lánc tetszőleges kezdeti eloszlás esetén tart a stacionárius eloszláshoz, ha az idő tart a végtelenhez. Ilyen tipusú eredményt fogunk bizonyítani.

Csak abban az esetben várhatjuk, hogy egy Markov láncnak létezik az előbb leírt stacionárius eloszlása, ha annak állapotai rekurrens, pozitív aperiódikus állapotok. Ez a feltétel biztosítja, hogy az $u_j = \frac{1}{\mu_j}$ számok pozitívak. Annak érdekében, hogy lássuk azt, hogy a később megfogalmazott eredmény jól használható, először áttekintjük egy Markov lánc állapotterének szerkezetét. Megmutatjuk, hogy az állapotteret természetes módon fel lehet bontani alkalmas osztályok uniójaként úgy, hogy az egy osztályban levő állapotok egyszerre tranziens, null vagy pozitív rekurrens állapotok, és mindegyiküknek ugyanannyi a periódusa. Az állapottér alkalmas felbontásának megtalálásában a következő eredmény játszik kulcsfontosságú szerepet.

Tétel Markov lánc állapotainak tulajdonságairól. *Legyen $X_0, X_1, \ldots$ Markov lánc valamely $\mathcal{E} = \{E_1, E_2, \ldots\}$ állapottéren $P(n, i, j) = P(X_n = j | X_0 = i)$ átmenetvalószínűségekkel. Legyen E_j rekurrens állapota a Markov láncnak, és definiáljuk az állapottér következő az E_j állapottól függő $\mathcal{C}(j) \subset \mathcal{E}$ részhalmazát. $E_k \in \mathcal{C}(j)$ akkor és csak akkor, ha létezik olyan r pozitív egész szám, hogy $P(r, j, k) > 0$, azaz az E_j állapotból pozitív valószínűséggel el lehet jutni az E_k állapotba.*

A $\mathcal{C}(j)$ halmaz minden eleme rekurrens állapot. Ha $E_k \in \mathcal{C}(j)$, $E_l \in \mathcal{C}(j)$ akkor $F(k, l) = 1$ a $(10')$ képletben definiált $F(\cdot, \cdot)$ mennyiséggel, azaz az E_k állapotból kiinduló Markov lánc 1 valószínűséggel eléri az E_l állapotot. A $\mathcal{C}(j)$ halmazban levő állapotok mindegyike egyszerre null vagy pozitív rekurrens állapot, és mindegyikük periódusa megegyezik. Továbbá minden $E_k \in \mathcal{C}(j)$ (rekurrens) állapotra az E_k állapottól függő $\mathcal{C}(k) \subset \mathcal{E}$ halmazra $\mathcal{C}(k) = \mathcal{C}(j)$.

A tétel bizonyítása. Először azt mutatjuk meg, hogy ha $P(r, j, k) > 0$ valamely $r \geq 1$ számra, azaz az E_j állapotból indított Markov lánc pozitív valószínűséggel eljut az E_k állapotba, akkor $F(k, j) = 1$, azaz az E_k állapotból indított Markov lánc 1 valószínűséggel eljut az E_j állapotba. Valóban, ha ez nem lenne igaz, akkor az $1 - F(k, j) > 0$ és $P(n, k, j) > 0$ relációk miatt pozitív lenne annak a valószínűsége, hogy az E_j állapotból induló Markov lánc az n időpont után nem tér vissza az E_j állapotba, mert az n időpontban az E_k állapotba jut, és onnan nem lép soha az E_j állapotba. Ez viszont ellentmond annak a ténynek, hogy az E_j állapot rekurrens, ezért a Markov lánc 1 valószínűséggel végtelen sokszor visszatér oda.

Mivel $k \in C_j$ esetén léteznek olyan r és $\bar{r}$ számok, amelyekre $P(r, j, k) > 0$ és $P(\bar{r}, k, j)$ ezért $P(r + \bar{r} + n, k, k) \geq P(\bar{r}, k, j) P(n, j, j) P(r, j, k)$. Ebből az egyenlőtlenségből, illetve abból a tényből, hogy a E_j állapot rekurrens tulajdonsága ekvivalens azzal, hogy $\sum\limits_{n=1}^{\infty} P(n,j,j)=\infty$ következik, hogy $\sum\limits_{n=1}^{\infty} P(n,k,k)=\infty$, és az E_k állapot szintén rekurrens. Továbbá, ha az E_j állapot vagy E_k állapot egyikéből egy E_l pozitív valószínűséggel elérhető, akkor ez az E_l állapot a E_j és E_l állapotok másikából is pozitív valószínűséggel elérhető, (esetleg először azt az állapotot látogatva meg, ahonnan tudjuk, hogy az E_l halmaz meglátogatható. Ez azt jelenti, hogy $\mathcal{C}(j)=\mathcal{C}(k)$, ha $E_k\in\mathcal{C}(j)$. Ez speciálisan azt a következményt is maga után vonja, hogy $E_k,E_l\in\mathcal{C}(j)$ esetén $F(k,l)=1$. Valóban (a paraméterek más szereposztásában) láttuk, hogy ez a reláció következik abból, hogy $E_k\in\mathcal{C}(l)$.

Megmutatom, hogy abban az esetben, ha létezik egy olyan $E_k\in\mathcal{C}(j)$, amelyik rekurrens null állapot, akkor E_j is rekurrens null állapot. Innen következik, hogy $\mathcal{C}(j)$ osztályban levő állapotok egyidejűleg null vagy pozitív állapot. Az említett tulajdonság azért érvényes, mert $P(r+\bar r+n,k,k)\ge P(\bar r,k,j)P(n,j,j)P(r,j,k)$ alkalmas $P(\bar r,k,j)>0$ és $P(r,j,k)>0$ számokkal, ezért a $\lim\limits_{n\to\infty}P(n,k,k)=0$ relációból következik, hogy $\lim\limits_{n\to\infty}P(n,j,j)=0$.

A tétel bizonyításának befejezéséhez elegendő azt megmutatni, hogy ha $E_k,E_l\in\mathcal{C}(j)$, és az $\mathcal{A}(k)=\{n\colon P(n,k,k)>0\}$ halmaz elemei oszthatók egy d számmal, akkor az $A(l)=\{n\colon P(n,l,l)>0\}$ halmaz elemei szintén oszthatók ezzel a d számmal. Innen következik, hogy a $\mathcal{C}(j)$ halmaz elemei ugyanolyan periódusú állapotok. A fent megfogalmazott állítás igazolása érdekében vegyük észre, hogy léteznek olyan $r_1>0$ és $r_2>0$, amelyekre $P(r_1,k,l)>0$ és $P(r_2,l,k)>0$. Továbbá $r_1+r_2\in\mathcal{A}(k)$, mert a Markov lánc pozitív valószínűséggel visszajuthat r_1+r_2 lépésben az E_k állapotból az E_k állapotba, r_1 lépésben az E_l majd további r_2 lépésben az E_k állapotba jutva. Ezért r_1+r_2 osztható a d számmal. Továbbá, ha $n\in\mathcal{A}(l)$, azaz $P(n,l,l)>0$ valamely n számra, akkor $P(n+r_1+r_2,k,k)\ge P(r_1,k,l)P(n,l,l)P(r_2,l,k)>0$, ezért $n+r_1+r_2$ osztható a d számmal. A fentiekből következik, hogy n osztható d-vel, és ezt kellett belátnunk.

Bevezetem a következő két definiciót.

Zárt osztályok definiciója egy Markov láncok állapotterében. *Legyen adva egy $X_0,X_1,\dots$ Markov lánc valamely $\mathcal{E}=\{E_1,E_2,\dots\}$ állapotterén $P(j,k)$ 1 lépéses átmenetvalószínűségekkel. Azt mondjuk, hogy egy $\mathcal{C}\subset\mathcal{E}$ halmaz az állapottér egy zárt osztályát alkotja, ha minden $E_j\in\mathcal{C}$ állapotra $\sum\limits_{k\in\mathcal{C}}P(j,k)=1$, azaz, ha a Markov lánc egy $\mathcal{C}$ halmazbeli állapotból indul, akkor egy valószínűséggel a következő lépésben is egy ilyen állapotban marad.*

Irreducibilis, zárt osztályok definiciója egy Markov lánc állapotterében. *Azt mondjuk, hogy egy Markov lánc állapotterének egy zárt osztálya irreducibilis, ha önmagán és az üres halmazon kívül nincs más olyan részhalmaza, amely zárt osztály. Egy Markov láncot irreducibilisnak hívunk, ha a teljes állapottér irreducibilis (zárt) osztály.*

Bebizonyítom az előbbi tétel segítségével az irreduciblis zárt osztályok néhány tulajdonságát.

Tétel Markov lánc irreducibilis zárt osztályainak tulajdonságairól. *Legyen $\mathcal{C}$ egy $X_0, X_1, \dots$, $P(n,j,k) = P(X_n = E_k | X_0 = E_j)$ átmenetvalószínűségekkel rendelkező Markov lánc $\mathcal{E} = \{E_1, E_2, \dots\}$ állapotterének irreducibilis, zárt osztálya. Ekkor minden $E_k \in \mathcal{C}$ és $E_l \in \mathcal{C}$, állapotokra az E_l állapot elérhető az E_k állapotból, azaz létezik olyan r pozitív egész szám, amelyre $P(r,k,l) > 0$. Egy irreducibilis osztály minden elemének ugyanaz a periódusa, és egy irreucibilis osztály minden eleme egyidejűleg, tranziens, null-rekurrens vagy pozitív rekurrens állapot. Ha E_j a Markov lánc rekurrens állapota, akkor az előző tételben definiált $\mathcal{C}(j)$ halmaz, amely azokból az állapotokból áll, amelyeket az E_j állapotból pozitív valószínűséggel el lehet érni, irreducibilis, zárt osztály. Minden olyan zárt, irreducibilis osztály, amely tartalmaz egy E_j rekurrens állapotot megegyezik egy ilyen $\mathcal{C}(j)$ osztállyal.*

A tétel bizonyítása. Adva egy $\mathcal{C}$ zárt osztály és egy $E_k \in \mathcal{C}$ állapot definiáljuk azt a $\mathcal{C}^{(k)} \subset \mathcal{C}$ halmazt, amely azokból az $E_p \in \mathcal{C}$ állapotokból áll, amelyekre $P(n,k,p) > 0$ valamely n számra. Azt állítom, hogy $\mathcal{C}^{(k)}$ is zárt osztály. Innen következik a Tétel első állítása, mert ha a $\mathcal{C}$ halmaz valamelyik állapota nem érhető el az E_k állapotból, azaz ha létezik olyan E_l állapot, amelyre $P(n,k,l) = 0$ minden pozitív egész n számra, akkor $\mathcal{C}^{(k)}$ a $\mathcal{C}$ halmaz olyan nem üres, valódi részhalmaza, amely zárt osztály. Ez azt jelenti, hogy $\mathcal{C}$ nem irreducibilis.

Világos, hogy $\sum\limits_{l: E_l \in \mathcal{C}^{(k)}} P(n,k,l) = 1$ minden $n = 1, 2, \dots$ számra, mert $\mathcal{E} \setminus \mathcal{C}^{(k)}$ csak olyan E_l állapotokat tartalmaz, amelyekre $P(n,k,l) = 0$ minden $n = 1, 2, \dots$ számra. Ha $E_l \in \mathcal{C}^{(k)}$, akkor a $\sum\limits_{p: E_p \in \mathcal{C}^{(k)}} P(n,l,p) = 1$ relációnak is teljesülnie kell minden $n = 1, 2, \dots$ számra, amiből következik, hogy $\mathcal{C}^{(k)}$ zárt osztály. A bizonyítandó állítás igaz, mert ellenkező esetben lenne olyan $n \ge 1$ szám, amelyre $\sum\limits_{p: E_p \in \mathcal{C}^{(k)}} P(n,l,p) < 1$. Ekkor létezik olyan r szám, amelyre $P(r,k,l) > 0$, és $\sum\limits_{\bar{l}: E_{\bar{l}} \in \mathcal{C}^{(k)}} P(n+r,k,\bar{l}) \le \sum\limits_{\bar{l}: E_{\bar{l}} \in \mathcal{E} \setminus \{E_l\}} P(r,k,\bar{l}) + P(r,k,l) \sum\limits_{p: E_p \in \mathcal{C}^{(k)}} P(n,l,p) = 1 - P(r,k,l) + P(r,k,l) \sum\limits_{p: E_p \in \mathcal{C}^{(k)}} P(n,l,p) < 1$, és ez ellentmondás.

Az irreducibilis osztályok már bizonyított tulajdonságából következik az előző tétel bizonyításának befejezéséhez hasonlóan, hogy egy ilyen osztály minden elemének ugyanannyi a periódusa. Az előző tételben láttuk, hogy egy rekkurrens E_j állapotra a $\mathcal{C}(j)$ halmaz az állapottér egy zárt osztálya. Továbbá, ha $\bar{\mathcal{C}} \subset \mathcal{C}(j)$ nem üres zárt osztály, akkor létezik egy $E_k \in \bar{\mathcal{C}}$ elem, és ezért $\mathcal{C}(j) = \mathcal{C}(k) \subset \bar{\mathcal{C}}$. Ezért $\mathcal{C}(j)$ irreducibilis zárt osztály. Továbbá, ha egy irreducibilis zárt osztály tartalmaz egy rekurrens E_j állapotot, akkor tartalmazza a zárt $\mathcal{C}(j)$ osztályt is, ezért megegyezik vele. Innen következik, hogy egy irreducibilis zárt osztálynak vagy mindegyike eleme tranziens állapot vagy megegyezik valamelyik $\mathcal{C}(j)$ halmazzal, ahol E_j rekurrens állapot. Egy ilyen osztály minden eleme rekurrens null vagy rekurrens pozitív állapot. A tétel bizonyítását befejeztük.

Megfogalmazom a fenti eredmények alábbi következményét.

Következmény. *Legyen $X_0, X_1, \dots$ Markov lánc $P(n,j,k) = P(X_n = E_k | X_0 =$*

E_j) átmenetvalószínűségekkel egy $\mathcal{E} = \{E_1, E_2 \dots\}$ állapottéren. Bontsuk fel az $\mathcal{E}$ állapotteret az $\mathcal{E} = \mathcal{R} \cup \mathcal{T}$ képlettel az $\mathcal{R}$ rekurrens és $\mathcal{T}$ tranziens állapotokból álló halmazok uniójaként. A $\mathcal{R}$ halmaz felbontható $\mathcal{C}(j)$ diszjunkt irreducibilis, zárt osztályok uniójaként. Ha $E_k, E_l \in \mathcal{C}(j)$ az $\mathcal{R}$ halmaz egy irreducibilis, zárt $\mathcal{C}(j)$ osztályára, akkor $F(k,l) = 1$ a $(10')$ képletben definiált $F(\cdot,\cdot)$ függvénnyel.

Bizonyítás. A megfogalmazott eredmények következnek a Markov lánc irreducibilis zárt osztályainak és egy Markov lánc állapotainak tulajdonságairól szóló tételek eredményeiből. Azt kell még meggondolnunk, hogy ha két az utóbbi tétel megfogalmazásában definiált $\mathcal{C}(j)$ és $\mathcal{C}(k)$ halmaz nem diszjunkt, akkor $\mathcal{C}(j) = \mathcal{C}(k)$. De ebben az esetben létezik egy $E_l \in \mathcal{C}(j) \cap \mathcal{C}(k)$ állapot, és $\mathcal{C}(j) = \mathcal{C}(l) = \mathcal{C}(k)$.

A fent megfogalmazott következményben egy Markov lánc rekurrens állapotainak halmazát felbontottuk irreducibilis, zárt halmazok uniójaként. A tranziens állapotok halmazának nincs ilyen egyértelmű leírása. Ezt mutatja a következő példa.

Példa. *Legyen $X_1, X_2, \dots$ bolyongás a d-dimenziós téren. Ez irreducibilis Markov lánc (azaz egyetlen irreducibilis osztályból áll), melynek elemei 2 periódikusak. Ha $d = 1$ vagy $d = 2$ akkor ennek az osztálynak az elemei null periódikusak, ha $d \ge 3$ akkor tranziensek.*

Definiáljuk a következő Markov láncot a két-dimenziós tér egész koordinátájú rácspontjain: Ha a Markov lánc valamely (j,k), $j \ne 0$ pontban van, akkor a következő lépésban egyforma $\frac{1}{4}$ valószínűséggel lép a $(j-1,k)$, $(j+1,k)$, $(j,k+1)$, $(j,k-1)$ szomszéd pontok valamelyikébe. Ha a Markov lánc valamely $(0,k)$ alakú pontban van akkor a következő lépésben $\frac{1}{2}$ valószínűséggel lép a $(0,k+1)$ vagy a $(0,k-1)$ pontba. Ennek a Markov láncnak a $(0,k)$, $k = 0 \pm 1, \pm 2, \dots$ alakú pontok a rekurrens állapotai, amelyek irreducibilis zárt osztályt alkotnak. Ennek elemei null-rekurrens, 2 periódusú állapotok. A (j,k), $j \ne 0$ alakú pontok tranziensek. Egy ilyen állapotból kiinduló Markov lánc 1 valószínűséggel eljut a rekurrens állapotokból álló halmazba, ezért nincs olyan irreducibilis zárt osztály, amely ezeket tartalmazza.

Feladat: Bizonyítsuk be a fenti példa állítását.

A következő állítást fogjuk bebizonyítani.

Tétel irreducibilis Markov láncok stacionárius eloszlásáról. *Legyen $X_0, X_1, \dots$ irreducibilis, aperiódikus Markov lánc $P(n,j,k) = P(X_n = E_k | X_0 = E_j)$ átmenetvalószínűségekkel egy $\mathcal{E} = \{E_1, E_2, \dots\}$ állapottéren.*

a) Ha a Markov lánc állapotai pozitív rekurrensek, akkor az $u_j = \lim\limits_{n\to\infty} P(n,i,j) > 0$ határértékek léteznek, $u_j = \frac{1}{\mu_j}$, ahol a μ_j számok a (2) formulában vannak definiálva. Az u_j számok, $E_j \in \mathcal{E}$, a Markov lánc stacionárius eloszlását definiálják, azaz teljesítik a (13) és (14) formulákat.

b) Megfordítva, ha a Markov láncnak van u_j, $j \in \mathcal{E}$, stacionárius eloszlása, akkor a Markov lánc ergodikus, azaz pozitív rekurrens állapotokkal rendelkezik, (és aperiódikus).

A Markov lánc (egyetlen) stacionárius eloszlása teljesíti az $u_j = \frac{1}{\mu_j}$, $E_j \in \mathcal{E}_j$, azonosságot.

A tétel bizonyítása. Tekintsük először azt az esetet, amikor az a) rész feltételei teljesülnek. Az egy Markov lánc egy állapotának jellemzéséről szóló tételben bizonyított (12) formulából és a Markov lánc állapotainak tulajdonságairól szóló tételből következik, hogy $\lim\limits_{n\to\infty} P(n,i,j) = F(i,j)\frac{1}{\mu_j} = \frac{1}{\mu_j}$. Azt kell még belátnunk, hogy az $u_j = \frac{1}{\mu_j}$, $E_j \in \mathcal{E}$, számok teljesítik a (13) és (14) azonosságot. Először azt mutatom meg, hogy

$$u_j = \sum_{k\colon E_k\in\mathcal{E}} u_k P(m,k,j) \quad \text{minden } m = 1,2,\dots, \text{ számra.} \tag{15}$$

Ennek érdekében írjuk fel a

$$P(n+m,i,j) = \sum_{k\colon E_k\in\mathcal{E}} P(n,i,k)P(m,k,j)$$

és próbáljunk limeszt venni mind a két oldalon felhasználva, hogy $\lim\limits_{n\to\infty} P(n+m,i,j) = u_j$ és $\lim\limits_{n\to\infty} P(n,i,k) = u_k$. Mivel a jobb oldalon szereplő $P(m,k,j)$ együtthatók összege (rögzített j-re és a k változó szerint összegezve) lehet divergens is, egyelőre csak az alábbi a kívántnál gyengébb összefüggést tudjuk felírni (felhasználva, hogy a jobb-oldalon szereplő számok mind nem-negatívak:

$$u_j \ge \sum_{k\colon E_k\in\mathcal{E}} u_k P(m,k,j) \tag{15a}$$

Hasonlóan a $\sum\limits_{k\colon E_k\in\mathcal{E}} P(n,i,k) = 1$ azonossából kapjuk $n\to\infty$ határátmenettel, hogy

$$s = \sum_{k\colon E_k\in\mathcal{E}} P(n,i,k) \le 1. \tag{16}$$

Ezt az egyenlőtlenséget felhasználva, és összegezve j-re a (15a) egyenlőtlenségeket azt kapjuk, hogy

$$s = \sum_{j\colon E_j\in\mathcal{E}} u_j \ge \sum_{j\colon E_j\in\mathcal{E}} \sum_{k\colon E_k\in\mathcal{E}} u_k P(m,k,j) = \left(\sum_{j\colon E_j\in\mathcal{E}} P(m,k,j)\right)\left(\sum_{k\colon E_k\in\mathcal{E}} u_k\right) = s.$$

Mivel az utolsó egyenlőtlenségsor bal és jobb oldala egyenlő (és véges), a felhasznált (15a) egyenlőtlenségekben mindenütt azonosságnak kell állni, tehát a (15) formula érvényes.

Ha a (15) formulát $m = 1$ választással tekintjük, megkapjuk a (14) képletet. Tekintsük újból a (15) formulát tetszőleges j számmal, és tartsunk végtelenhez az m

paraméterrel. Ekkor felhasználva a (16) formulát (pontosabban csak azt, hogy az ott tekintett összeg véges) azt kapjuk, hogy

$$u_j = \sum_{k:E_k\in\mathcal{E}} u_k u_j = u_j \sum_{k:E_k\in\mathcal{E}} u_k.$$

Mivel $u_j = \frac{1}{\mu_j} > 0$, innen következik a (13) reláció.

Rátérek a b) rész bízonyítására. Megmutatom n szerinti teljes indukcióval a Chapman–Kolmogorov egyenlet segítségével, hogy a b) esetben teljesül a (14) képlet következő általánosítása is.

$$u_j = \sum_{i:E_i\in\mathcal{E}} u_i P(n,i,j) \quad \text{minden } j=1,2,\dots \text{ és } n=1,2,\dots \text{ számra.} \tag{14'}$$

Valóban, ha a (14) formula igaz az n számra, akkor

$$\begin{aligned}\sum_{i:E_i\in\mathcal{E}} u_i P(n+1,i,j) &= \sum_{i:E_i\in\mathcal{E}} u_i \sum_{k:E_k\in\mathcal{E}} P(n,i,k)P(k,j)\\ &= \sum_{k:E_k\in\mathcal{E}} P(k,j) \sum_{i:E_i\in\mathcal{E}} u_i P(n,i,k) = \sum_{k:E_k\in\mathcal{E}} P(k,j)u_k = u_j,\end{aligned}$$

tehát a (14′) azonosság igaz $n+1$-re is.

Alkalmazzunk $n\to\infty$ határátmenetet a (14′) képletben. Azt állítom, hogy a Markov lánc ergodikus, minden i és j indexre létezik a $\lim\limits_{n\to\infty} P(n,i,j) = \frac{1}{\mu_j} > 0$ (2) képletben definiált határérték, és

$$u_j = \sum_{i:E_i\in\mathcal{E}} u_i \frac{1}{\mu_j} \quad \text{minden } j=1,2,\dots \text{ számra.} \tag{17}$$

Valóban, mivel irreducibilis Markov láncot tekintünk, a Markov lánc minden állapota egyszerre tranziens, null-rekurrens vagy pozitív rekurrens. Viszont az első két esetben $\lim\limits_{n\to\infty} P(n,i,j) = 0$, ezért a (14′) formulában elvégzett határátmenet azt adná, hogy $u_j = 0$ minden j indexre. Ez viszont ellentmond a (13) relációnak. Ezért a Markov lánc állapotai pozitív rekurensek. Mivel feltettük, hogy a Markov lánc aperiódikus, ezért ergodikus, az átmenet valószínűségeknek az előző bekezdésben felírt határértékeik vannak, és teljesül a (17) formula. A (17) és (13) formulák alapján viszont

$$u_j = \sum_{i:E_i\in\mathcal{E}} u_i \frac{1}{\mu_j} = \frac{1}{\mu_j} \sum_{i:E_i\in\mathcal{E}} u_i = \frac{1}{\mu_j} \quad \text{minden } j=1,2,\dots \text{ számra.}$$

Ez azt jelenti, hogy egy ergodikus Markov láncnak $u_j = \frac{1}{\mu_j}$, $E_j\in\mathcal{E}_j$, az egyetlen stacionárius eloszlása.

Feladat: Lássuk be az előző feladat eredményének a segítségével, hogy egy $X_0, X_1, \dots,$ irreducibilis, ergodikus Markov lánc eloszlása tetszőleges kezdeti eloszlás esetén konvergál a Markov lánc stacionárius eloszlásához, ha $n \to \infty$, azaz $\lim\limits_{n\to\infty} P(X_n = E_j) = u_j$ minden u_j állapotra.

A korábbi eredményekben feltettem, hogy a tekintett Markov lánc aperiódikus. Ezt elsősorban kényelmi szempontok miatt tettem. Periódikus Markov láncok vizsgálata hasonlóan történhet, és az eredmények is hasonlóak, csak a jelölés válik kissé bonyolultabbá. Ezenkívül periódikus Markov láncok vizsgálata visszavezethető az aperiódikus Markov láncok esetére. Itt csak röviden áttekintem azt, hogy milyen eredmények érvényesek, és milyen észrevétel segítségével kaphatjuk meg azokat.

Legyen $X_0, X_1, \dots$ irreducibilis Markov lánc egy $\mathcal{E} = \{E_1, E_2, \dots\}$ állapottéren $P(n, j, k)$ átmenetvalószínűséggel. Tegyük fel, hogy a Markov lánc állapotai pozitív rekurrensek, és periódusok valamilyen $d \ge 1$. (Tudjuk, hogy egy irreducibilis Markov lánc minden állapota egyszerre, tranziens, null vagy pozitív rekurrens, és minden állapotnak ugyanaz a periódusa.) Tekintsük az $X_0, X_1, \dots$ Markov lánc mellett az $\bar{X}_n = X_{dn}$, $n = 0, 1, 2, \dots$ Markov láncot is, amelynek állapottere megegyezik az eredeti Markov lánc $\mathcal{E}$ állapotterével, és értéke az n időpontban egyenlő az eredeti Markov lánc értékével az nd időpontban. Ez utóbbi Markov lánc aperiódikus, $\bar{P}(j, k) = P(d, j, k)$ egy lépéses átmenetvalószínűségekkel, minden állapota pozitív rekurrens, de állapottere nem irreducibilis. Viszont meg tudjuk adni e Markov lánc állapotterének felbontását d darab $\mathcal{C}_0, \dots, \mathcal{C}_{d-1}$ irreducibilis osztály uniójaként.

Rögzítsük mondjuk az $E_1 \in \mathcal{E}$ állapotot, és tekintsük minden $E_k \in \mathcal{E}$ állapotra az $\mathcal{N}_k = \{n \colon P(n, 1, k) > 0\}$ halmazt, azaz azon időpontok halmazát, amely időpontok alatt az E_1 állapotból pozitív valószínűséggel jutunk az E_k állapotba. Nem nehéz belátni, hogy az $X_0, X_1, \dots$ Markov lánc d periódusa miatt létezik olyan $0 \le p \le d-1$ szám, hogy $n = p \bmod d$ minden $n \in \mathcal{N}_k$ számra. Nevezzük ezt a p számot az E_k állapot $r(k)$ rangjának. Definiáljuk a $\mathcal{C}_p \subset \mathcal{E}$, $1 \le p \le d$, halmazokat úgy, hogy $E_k \in \mathcal{C}_p$ akkor és csak akkor, ha $r(k) = p$. Némi munkával be lehet látni, hogy a $\bar{X}_0, \bar{X}_1, \dots$ Markov lánc $\mathcal{E}$ állapotterének felbontása irreducibilis zárt osztályokra a $\mathcal{C}_p$, $0 \le p \le d-1$, halmazokból áll. Továbbá, ha $E_k \in \mathcal{C}_p$, $E_l \in \mathcal{C}_q$ valamely p és q számokkal, akkor $P(n, k, l) > 0$ csak akkor lehetséges, ha $n = q - p \bmod d$. Ezenkívül, ha $E_k \in \mathcal{C}_p$, és $P(n, k, l) > 0$ valamely n időpontra, akkor $E_l \in \mathcal{C}_q$ azzal a q számmal, amelyre $n = q - p \bmod d$.

A fenti összefüggések segítségével, és használva a már bizonyított eredményeket az $\bar{X}_0, \bar{X}_1, \dots$ Markov lánc megszorításán az irreducibilis $\mathcal{C}_p$, $1 \le p \le d-1$ irreducibilis zárt osztályokra kapjuk, hogy létezik a

$$du_j = \lim_{n\to\infty} P(dn + q - p, i, j) = \frac{d}{\mu_j} \quad \text{ha } E_i \in \mathcal{C}_p \quad \text{ és } E_j \in E_q$$

határérték, ahol a μ_j szám a (2) képletben van definiálva, és $P(dn + r, i, j) = 0$, ha $r \ne q - p \bmod p$. Továbbá az is igaz, hogy az $u_j = \frac{1}{\mu_j}$ számok alkotják a Markov lánc (egyetlen) stacionárius eloszlását.

Rátérek véges állapotterű, azaz véges sok különböző értéket felvevő Markov láncok rövid tárgyalására. Véges állapotterű Markov láncok rendelkeznek néhány extra tulajdonsággal, amelyeket a következő tételben fogalmazok meg. Ezenkívül a sztochaszikus mátrixok néhány lineáris algebrai tulajdonsága segít véges állapotterű Markov láncok vizsgálatában.

Tétel véges állapotterű Markov láncok tulajdonságairól. *Egy véges állapotterű Markov láncnak mindig van pozitív rekurrens állapota, viszont nincs rekurrens null állapota. Egy tranziens állapotból elindított Markov lánc 1 valószínűséggel bekerül a Markov lánc rekurrens állapotaiból álló halmazba.*

A tétel bizonyítása. Mivel egy véges lánc Markov n lépéses átmenetvalószínűségét is egy sztochasztikus mátrix adja meg, amelynek mérete nem függ az n számtól, és egy sztochasztikus mátrix sorösszege 1, ezért, rögzítve a sztochasztikus mátrix i-edik sorát valamely i számmal, létezik olyan j index, amelyre $\limsup n \to \infty P(n,i,j) > 0$. Viszont egy ilyen j indexhez tartozó E_j állapot pozitív rekurrens, mert sem tranziens, sem null rekurrens nem lehet. Mivel a Markov lánc rekurrens állapotaiból álló halmaz felbontható diszjunkt irreducibilis zárt osztályok uniójára, és egy irreducibilis zárt osztály elemei egyidejűleg pozitív vagy null rekurrens állapotok, ezért tekintve a Markov lánc megszorítását egy irreducibilis zárt osztályra az előző érvelés alapján kapjuk, hogy a Markov láncnak nincs rekurrens null állapota. Végül, ha egy tranziens E_i állapotból elindított Markov lánc pozitív valószínűséggel elkerülné a Markov lánc rekurrens állapotaiból álló halmazt, akkor (újból a Markov lánc véges állapottere miatt létezne a Markov láncnak olyan E_j tranziens állapota, amelyre $\limsup\limits_{n\to\infty} P(n,i,j) > 0$, és ez ellentmondás.

Egy Markov lánc viselkedéséről nagyon hasznos információt ad az a Markov lánc átmenetvalószínűségeit megadó sztochasztikus mátrix spektrumának, azaz sajátértékeinek és sajátvektorainak az ismerete. Vegyük észre, hogy egy sztochasztikus mátrixnak, ha jobbról szorozzuk meg oszlopvektorokkal, akkor a csupa 1 koordinátából álló vektor sajátvektor 1 sajátértékkel. Továbbá 1-nél nagyobb sajátértékkel rendelkező (oszlop) sajátvektor nem lehetséges. Azt is tudjuk a lineáris algebrából, hogy egy mátrixot akár balról szorozzuk meg egy sorvektorral, akár jobbról egy oszlopvektorral, ugyanazok lesznek a sajátértékei. (A sajátvektorai lényegesen különbözhetnek.)

Feladat. Lássuk be, hogy egy sztochasztikus mátrixnak nem lehet 1-nél nagyobb sajátértékkel rendelkező (oszlop) sajátvektora.
Segítség: Mutassuk meg, hogy az eredeti vektor legnagyobb abszolut értékű koordinátájának az abszolut értéke nagyobb vagy egyenlő, mint a képvektor bármely koordinátájának az abszolut értéke.

Tudjuk tehát, hogy egy sztochasztikus mátrixnak létezik egy 1 sajátértékkel rendelkező (sor) sajátvektora. Enyhe plusz feltevések mellett azt is lehet tudni, hogy ennek összes koordinátája pozitív, és szép esetekben az is igaz, hogy a sztochasztikus mátrix összes többi sajátértéke szigorúan kisebb, mint 1. Ilyen esetben a sztochasztikus mátrix n-ik hatványának viselkedéséről nagyon értékes információt nyerhetünk. Ugyanis felírva a mátrix által meghatározott lineáris transzformációt olyan koordinátarendszerben, ahol annak a lehető legegyszerűbb az alakja (ez a Jordan féle normálalak használatát jelenti) be lehet látni, hogy a következő aszimptotikus reláció érvényes. Tekintsünk egy tetszőleges nem negatív értékű koordinátákból álló sorvektort, amelyben a koordináták összege 1. Ha alkalmazzuk erre a vektorra az átmenetvalószínűségek által meghatározott sztochasztikus mátrix n-ik hatványát, akkor olyan vektort kapunk, amely e mátrix 1 sajátértékű sajátvektorától az (n változó függvényében) exponenciálisan kis különbséggel tér el. Ez azt jelenti, hogy a mátrix sajátvektora adja meg a Markov lánc stacionárius eloszlását, és ha a Markov láncot tetszőleges kezdeti eloszlással indítjuk el, akkor a Markov lánc eloszlása az n idő függvényében exponenciális sebességgel konvergál a stacionárius eloszláshoz.

Az előbb vázolt gondolatmenet azt mutatja, hogy hasznos olyan eredményeket bizonyítani, amelyek megmondják, hogy egy sztochasztikus mátrixnak mikor van egyetlen 1 sajátértékű sajátvektora. A Markov láncok elméletében Perron egy tétele, illetve annak Frobenius által bizonyított általánosítása hasznos. Ezek az eredmények olyan mátrixokkal foglalkoznak, amelyeknek összes együtthatója nem negatív. Perron tételét fogom kimondani, és röviden jelzem, hogy milyen általánosítását adta ennek az eredménynek Frobenius.

Perron tétele. *Legyen egy A $n\times n$ méretű négyzetes mátrix minden eleme* szigorúan *pozitív szám. Az A mátrix* $\det(A-\lambda I)$ *karakterisztikus polinomjának, (ahol I az $n\times n$-es diagonális egységmátrix) legnagyobb abszolut értékű gyöke szigorúan pozitív, egyszeres gyök, és a karakterisztikus polinom összes többi gyökének az abszolut értéke szigorúan kisebb, mint ez a sajátérték. Az A mátrixnak a karakterisztikus polinom legnagyobb sajátértékéhez tartozó egyszeres sajátvektorának mindegyik koordinátája szigorúan pozitív szám. Ha az A mátrix sztochasztikus mátrix, azaz minden sorösszege 1-gyel egyenlő, akkor karakterisztikus polinomjának a legnagyobb gyöke 1.*

Perron tétele alkalmazható olyan véges állapotterű Markov lánc vizsgálatában, amelynek összes egy-lépéses átmenetvalószínűségei szigorúan pozitívak. Ebben az esetben a Markov lánc (egyetlen) stacionárius eloszlását a Markov lánc sztochasztikus mátrixának 1 sajátértékű alkalmasan normalizált sajátvektora adja meg, és a Markov lánc n időpontbeli eloszlása tetszőleges kezdeti eloszlás esetén exponenciális sebességgel tart a stacionárius eloszláshoz, ha az n lépésszám tart végtelenhez.

A Perron tételben tekintett mátrixok irreducibilis és 1 periódusú Markov láncok vizsgálatában használhatóak. Ugyancsak algebrai módon jellemezhetőek az irreducibilis Markov láncok átmenetvalószínűségei által meghatározott sztochasztikus mátrixok. Be lehet vezetni a nem negatív elemű úgynevezett irreducibilis mátrixok fogalmát az általános esetben, amely az irreducibilis Markov láncok által meghatározott sztochasztikus mátrixok természetes általánosítása. A Perron tétel Frobenius által bizonyított általánosítása irreducibilis, nem negatív elemű mátrixok legnagyobb abszolut értékű sajátértékeinek és a hozzájuk tartozó sajátvektorok jellemzését adja meg. Ennek megfogalmazása, amelyet ebben az ismertetésben elhagyok, bonyolultabb, mint az eredeti Perron tételé. Ez a bonyolultság azzal függ össze, hogy Frobenius eredménye nemcsak az aperiódikus, hanem az irreducibils, periódikus Markov láncok viselkedésének a leírását is lehetővé teszi.