Diszkrét és folytonos paraméterű Markov láncok

Csiszár Villő

Tartalomjegyzék

1. Bevezetés

Tekintsünk egy megszámlálható sok csúcspontú, irányított gráfot úgy, hogy minden élre egy nemnegatív szám van írva, és minden csúcs kimenő éleire írt számok összege 1 (hurokél is megengedett). Ezen a gráfon bolyongunk az élekre írt valószínűségek szerint. Ha egységnyi időközönként lépünk akkor diszkrét paraméterű Markov láncot kapunk. Ha pedig egy adott csúcsból exponenciális eloszlású idő elteltével lépünk tovább (ahol az eloszlás paramétere a csúcsra jellemző), akkor folytonos paraméterű Markov láncunk lesz. Mindkét esetben igaz, hogy a lánc jövőbeli fejlődése csak a pillanatnyi állapottól függ, a múlttól nem.

Ilyen jellegű folyamattal számos helyen találkozhatunk, például tömegkiszolgálási rendszerekben, populációk fejlődésének vizsgálatánál, diffúziós modelleknél. A folyamat általánosítása, ha a jövőbeli fejlődés csak a mostani és az előző néhány állapottól függ. Ez az általánosítás még szélesebb körű alkalmazásokat tesz lehetővé, pl. az írott nyelvek is tanulmányozhatók így.

A folyamattal kapcsolatban a következő kérdések merülhetnek fel: Honnan hová lehet eljutni? Mekkora eséllyel érünk vissza a kiindulási helyünkre? Mekkora az esélye, hogy végtelen sokszor visszatérünk? Átlagosan mennyi idő alatt érünk vissza a kiindulási helyre, vagy általánosabban egy másik csúcsba? Vannak-e elnyelő csúcsok? Ha igen, mekkora eséllyel nyelődünk el bennük? Van-e stacionárius kezdeti eloszlás? Hány? Tart-e a bolyongás egy stacionárius eloszláshoz? Milyen gyorsan? Érvényes-e valamilyen NSzT? Érvényes-e valamilyen CHT?

Ebben a jegyzetben először a diszkrét, majd a folytonos paraméterű Markov láncokkal foglalkozva próbálunk meg válaszolni a fenti kérdésekre. A tárgyhoz kapcsolódó további ajánlott irodalom:

1. W. Feller: Bevezetés a valószínűségszámításba és alkalmazásaiba (I. kötet), 15., 16., 17. fejezet.

2. S. Karlin, H. M. Taylor: Sztochasztikus folyamatok. 2., 3., 4. fejezet.

3. K. L. Chung: Markov Processes with Stationary Transition Probabilities.

4. Barczy Mátyás, Pap Gyula: Sztochasztikus folyamatok példatár és elméleti kiegészítések II. rész (diszkrét idejű Markov-láncok).

5. Pap Gyula: Sztochasztikus folyamatok.

I. rész

Diszkrét paraméterű Markov láncok

2. Visszatérőség

Ebben a fejezetben arra keressük a választ, hogy a lánc milyen gyakran tér vissza a kezdeti állapotba, illetve milyen sűrűn látogatja meg a különböző állapotokat. Megkérdezhetjük, hogy átlagosan mennyi ideig tart, míg egy adott állapotból a lánc először eljut egy másik adott állapotba (ha egyáltalán eljut). Rokon kérdés, hogy egy távoli időpontban mekkora eséllyel lesz a lánc az egyes állapotokban.

2.1. Markov-tulajdonság

Definiáljuk először pontosan a Markov láncot!

2.1. Definíció. *Legyen adott az $\{X_n\}_{n\in\mathbb{N}}$ folyamat, ahol $X_n : \Omega \to I$ valószínűségi változók az $(\Omega, \mathcal{F}, P)$ valószínűségi mezőn, I pedig megszámlálható halmaz. A folyamatot diszkrét paraméterű homogén Markov láncnak nevezzük, ha*

– *A folyamat Markov-tulajdonságú, azaz $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ jelöléssel minden $B \subseteq I$ és minden $m \geq n$ esetén*

$$P(X_m \in B \mid \mathcal{F}_n) = P(X_m \in B \mid X_n).$$

(Ezt a tulajdonságot általános mérhető állapottérre is definiálhatjuk.)

– *A Markov folyamat stacionárius (vagy homogén) átmenetvalószínűségű, azaz minden $i, j \in I$ esetén minden olyan n-re, melyre $P(X_n = i) > 0$,*

$$P(X_{n+1} = j \mid X_n = i) = p_{ij},$$

n-től függetlenül.

I az állapottér, elemeit a továbbiakban $i, j, k, \ldots$ jelöli. Esetünkben az $\mathcal{F}_n$ σ-algebra atomos, ezért a Markov-tulajdonság azzal ekvivalens, hogy

$$P(X_m \in B \mid X_{n_1} = i_1, \ldots, X_{n_k} = i_k) = P(X_m \in B \mid X_{n_k} = i_k), \qquad \text{ha } n_1 < \cdots < n_k \leq m.$$

Feladat: A Markov-tulajdonság ekvivalens megfogalmazásai. Mutassuk meg, hogy a Markov-tulajdonság ekvivalens a következőkkel:

1. $P(X_{n+1} \in B \mid \mathcal{F}_n) = P(X_{n+1} \in B \mid X_n)$.
2. Minden $A \in \mathcal{F}_{n-1}$ és $B \in \mathcal{F}^{n+1} = \sigma\{X_{n+1}, X_{n+2}, \ldots\}$ esetén
$$P(A \cap B \mid X_n) = P(A \mid X_n) P(B \mid X_n).$$
3. Minden $\mathcal{F}^{n+1}$-mérhető Y-ra $E(Y \mid \mathcal{F}_n) = E(Y \mid X_n)$, ha a baloldal értelmes.

A $P = (p_{ij})$ mátrixot átmenetmátrixnak nevezzük (nem keverendő össze a valószínűséggel), elemei az átmenetvalószínűségek. Az X_0 eloszlását kezdeti eloszlásnak hívjuk, és $p = (p_i)$-vel jelöljük. Ez a két objektum már meghatározza a folyamatot, hiszen a véges dimenziós eloszlásokat a Markov-tulajdonságot felhasználva kapjuk:

$$\begin{aligned} & P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \\ & \quad = P(X_0 = i_0) P(X_1 = i_1 | X_0 = i_0) \cdots P(X_n = i_n | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = \\ & \quad = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}. \end{aligned}$$

2.2. Definíció. *A P mátrix*

1. *sztochasztikus, ha $p_{ij} \geq 0$ és minden sor összege 1,*
2. *duplán sztochasztikus, ha sztochasztikus, és minden oszlop összege 1,*
3. *szubsztochasztikus, ha $p_{ij} \geq 0$ és minden sor összege legfeljebb 1.*

2.3. Tétel. *Tetszőleges I-n adott p eloszláshoz és $|I| \times |I|$ méretű P sztochasztikus mátrixhoz, létezik I állapotterű Markov lánc, melynek kezdeti eloszlása p, átmenetmátrixa P.*

Bizonyítás. (vázlat) A bizonyítás a Kolmogorov alaptételen múlik. Ez a következőt mondja ki:

Legyen $\mathcal{X}$ teljes szeparábilis metrikus tér, $\mathcal{B}$ a Borel halmazok σ-algebrája, Θ pedig tetszőleges halmaz. Jelölje $(\mathcal{X}^n, \mathcal{B}^{(n)})$ a tér n-edik hatványát. Tegyük fel, hogy minden n-re és minden $\theta_1, \ldots, \theta_n \in \Theta$-ra adott $\mathcal{B}^{(n)}$-en a $P_{\theta_1,\ldots,\theta_n}$ valószínűségi mérték, melyek eleget tesznek az alábbi konzisztenciafeltételeknek:
(i) $P_{\theta_1,\ldots,\theta_n,\theta_{n+1},\ldots,\theta_{n+m}}(A^{(n)} \times \mathcal{X}^m) = P_{\theta_1,\ldots,\theta_n}(A^{(n)})$ minden $A^{(n)} \in \mathcal{B}^{(n)}$-re,
(ii) Minden $\pi \in S_n$ permutációra $P_{\theta_1,\ldots,\theta_n}(A^{(n)}) = P_{\theta_{\pi(1)},\ldots,\theta_{\pi(n)}}(\pi(A^{(n)}))$.
Ekkor létezik valószínűségi mező és azon X_θ valószínűségi változók, melyek véges dimenziós eloszlásai az adottak.

Legyen $\mathcal{X} = I$, $\Theta = \mathbb{N}$, és

$$P_{0,1,\ldots,n}(i_0, i_1, \ldots, i_n) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n},$$

belátható, hogy ezek eleget tesznek a konzisztenciafeltételeknek. Ekkor a Kolmogorov alaptétel által garantált X_n folyamat valóban Markov lánc a kívánt kezdeti eloszlással és átmenetvalószínűségekkel. ∎

2.4. Állítás. *Legyenek $p_{ij}^{(n)} = P(X_{n+m} = j | X_m = i)$ az n-edrendű átmenetvalószínűségek (feltesszük, hogy $P(X_m = i) > 0$), ezekre teljesül a Chapman-Kolmogorov egyenlőség:*

$$p_{ij}^{(n+m)} = \sum_k p_{ik}^{(n)} p_{kj}^{(m)}.$$

Ez azt jelenti, hogy a a $p_{ij}^{(n)}$ mennyiségek éppen a P^n mátrix megfelelő elemei.

Bizonyítás.

$$P(X_{n+m} = j \mid X_0 = i) = \sum_k P(X_{n+m} = j, X_n = k \mid X_0 = i) =$$
$$\sum_k P(X_{n+m} = j \mid X_n = k, X_0 = i) P(X_n = k \mid X_0 = i) = \sum_k p_{kj}^{(m)} p_{ik}^{(n)}.$$

∎

2.2. Állapotok osztályozása

2.5. Definíció. *1. Azt mondjuk, hogy az i állapotból elérhető j, $(i \to j)$, ha van olyan $n \geq 0$, hogy $p_{ij}^{(n)} > 0$. Ez reflexív ($p_{ii}^{(0)} = 1$) és tranzitív (Chapman-*

Kolmogorov) reláció.

2. *Azt mondjuk, hogy i és j közlekednek, ha $i \to j$ és $j \to i$. Ez ekvivalenciareláció, tehát osztályokra bontja az állapotteret (csak az átmenetmátrixtól függ). A Markov lánc irreducibilis, ha egyetlen osztályból áll.*

3. *Az i állapot lényeges, ha $i \to j$ esetén $j \to i$ is teljesül.*

Az állapotokra értelmezett valamely tulajdonság *osztálytulajdonság*, ha egy osztálynak vagy minden eleme ilyen tulajdonságú, vagy egy sem. Triviálisan látszik, hogy a lényegesség osztálytulajdonság, azaz beszélhetünk lényeges és lényegtelen osztályokról. (Biz.: Tegyük fel, hogy i lényeges, j lényegtelen, és $i \to j$. Létezik k, hogy $j \to k$, de $k \not\to j$. Másrészt $i \to j \to k$, tehát i lényegessége miatt $k \to i \to j$, ami ellentmondás.) A lényeges osztályokból nem lehet kijutni (mert akkor vissza is tudnánk jönni, azaz osztályon belül maradnánk), a lényegtelenekből viszont igen: ha elhagytuk őket, akkor többé már nem térhetünk vissza. A lényegtelen osztályok között parciális rendezés van: $C >> D$, ha $i \in C, j \in D$ esetén $i \to j$ (tranzitív, reflexív, antiszimmetrikus).

2.6. Definíció. *Az $\{n > 0 : p_{ii}^{(n)} > 0\}$ halmaz legnagyobb közös osztója az i periódusa, jelölése $d(i)$. Ha a halmaz üres, akkor a periódust nem értelmezzük. Ha $d(i) = 1$, akkor az állapot aperiodikus.*

2.7. Állítás. *Egy osztály minden állapotának ugyanannyi a periódusa.*

Bizonyítás. Legyen $i, j \in C$ azonos osztálybeliek. Ekkor létezik n, m, hogy $p_{ij}^{(n)} > 0$, $p_{ji}^{(m)} > 0$. Ha valamely s-re $p_{jj}^{(s)} > 0$, akkor $p_{ii}^{(n+s+m)} > 0$, $p_{ii}^{(n+2s+m)} > 0$. Emiatt $d(i)|n+s+m$ és $d(i)|n+2s+m$, amiből $d(i)|s$ következik. $d(i)$ tehát közös osztója az ilyen s számoknak, azaz $d(i)|d(j)$. Mivel i és j szerepe felcserélhető, az állítást beláttuk. ■

2.8. Tétel. *(Részosztályok) Legyen C osztály d periódussal, és $i \in C$ tetszőleges. Ekkor C felbomlik d darab $C_0(i), C_1(i), \ldots, C_{d-1}(i)$ részosztályra úgy, hogy ha $j \in C_r(i)$ és $p_{ij}^{(n)} > 0$, akkor szükségképpen $n \equiv r \mod d$. Továbbá létezik $N(j)$ küszöbindex, hogy $n \geq N(j)$ esetén $p_{ij}^{(nd+r)} > 0$.*

Bizonyítás. Legyen $j \in C$. Létezik k, hogy $p_{ji}^{(k)} > 0$. Ha n-re és m-re $p_{ij}^{(n)} > 0$ és $p_{ij}^{(m)} > 0$, akkor $d|k + n$ és $d|k + m$, azaz $n \equiv m \mod d$.

A második állításra rátérve, a legnagyobb közös osztó előáll $d = \sum_{k=1}^{K} c_k n_k$ alakban, ahol c_k egész, és $p_{ii}^{(n_k)} > 0$. Legyen $n_0 = \sum n_k$, és $N = n_0^2 \max |c_k|$. Osszuk el az $n \geq N$ számot maradékosan n_0-val: $n = ln_0 + q$. Ekkor

$$nd = \sum_{k=1}^{K} (ld + qc_k) n_k,$$

és ebben a lineáris kombinációban az együtthatók már nemnegatívak. Ezért $p_{ii}^{(nd)} > 0$, és $p_{ij}^{((n+m_0)d+r)} > 0$, ha $p_{ij}^{(m_0 d+r)} > 0$. Tehát $N(j) = N + m_0$ jó lesz. ■

Vegyük észre, hogy a részosztályok függetlenek az i állapottól, csak az indexelésük függ tőle: ha $j \in C_r(i)$ és $k \in C_s(i)$, akkor $k \in C_{s-r}(j)$. A fenti tétel segítségével belátható, hogy legtöbbször elég irreducibilis és aperiodikus Markov láncokat vizsgálni. A nem lényeges állapotokat elhagyva ugyanis, ha egyszer belekerülünk valamelyik osztályba, akkor végleg ott is maradunk. Ha tehát C lényeges osztály d periódussal, és $P(X_0 \in C_r) = 1$, akkor az $\{X_{nd}\}_{n=0,\ldots}$ egy irreducibilis, aperiodikus ML C_r állapottérrel és $q_{ij} = p_{ij}^{(d)}$ átmenetvalószínűségekkel.

2.9. Példa. Legyen az átmenetmátrix a következő:

$$\begin{pmatrix} 0 & 1/4 & 1/4 & 0 & 0 & 0 & 1/2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/3 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Adjuk meg a ML osztályait, keressük meg közülük a lényegeseket! Mennyi az egyes osztályok periódusa?

2.3. Visszatérőség

Vezessük be a következő jelöléseket:

$$f_{ij}^{(0)} = 0, \quad f_{ij}^{(n)} = P(X_n = j,\, X_k \neq j : \ k = 1,2,\ldots n-1 \mid X_0 = i) \quad n \geq 1,$$

Az $f_{ij}^{(n)}$ mennyiség tehát annak valószínűsége, hogy az i állapotból indulva a lánc először az n. lépésben ér el a j állapotba. Legyen még

$$f_{ij}^* = \sum_{n=1}^{\infty} f_{ij}^{(n)}, \text{ és } g_{ij} = P(X_n = j \text{ végtelen sok } n\text{-re } \mid X_0 = i).$$

2.10. Definíció. *Az i állapot* visszatérő vagy rekurrens, *ha $f_{ii}^* = 1$, egyébként pedig* átmeneti vagy tranziens. *Ha i visszatérő, akkor az átlagos visszatérési idő $m_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$. Az i állapot* pozitív rekurrens, *ha $m_i < \infty$, ha pedig $m_i = \infty$, akkor* nulla rekurrens.

2.11. Állítás. *1. $i \to j$ akkor és csak akkor, ha $f_{ij}^* > 0$.*

2. $g_{ij} = f_{ij}^* g_{jj}$.

3. Ha i rekurrens állapot, akkor $g_{ii} = 1$, ha pedig i tranziens, akkor $g_{ii} = 0$.

4. A nem lényeges állapotok tranziensek (fordítva azonban nem igaz).

5. Ha $g_{ii} = 1$ és $f_{ij}^* > 0$, akkor $g_{ij} = 1$.

Bizonyítás. A fenti állítások az utolsó kivételével egyszerűen láthatók.

1. $f_{ij}^{(n)} \le p_{ij}^{(n)}$ és $p_{ij}^{(n)} \le \sum_{m=1}^{n} f_{ij}^{(m)}$, ezért $\sup_{n\ge 1} p_{ij}^{(n)} \le f_{ij}^* \le \sum_{n=1}^{\infty} p_{ij}^{(n)}$.

2. A „végtelen sok n-re” kifejezésre a v.s. rövidítést használva,

$$\begin{aligned} g_{ij} &= P(X_n = j \text{ v.s. } \mid X_0 = i) = \\ &\sum_{k=1}^{\infty} P(X_n = j \text{ v.s.}, X_k = j, X_l \ne j:\ l = 1, \dots, k-1 \mid X_0 = i) = \\ &\sum_{k=1}^{\infty} P(X_n = j \text{ v.s. } \mid X_k = j, X_l \ne j:\ l = 1, \dots, k-1, X_0 = i) f_{ij}^{(k)} = \\ &\sum_{k=1}^{\infty} g_{jj} f_{ij}^{(k)} = g_{jj} f_{ij}^*. \end{aligned}$$

3. Legyen $g_{ii}(m) = P(X_n = i$ legalább m különböző $n > 0-$ra $\mid X_0 = i)$, ha $m \ge 1$. Ekkor $g_{ii}(1) = f_{ii}^*$, és $g_{ii} = \lim_{m\to\infty} g_{ii}(m)$. Viszont $m \ge 2$-re

$$g_{ii}(m) = \sum_{k=1}^{\infty} f_{ii}^{(k)} g_{ii}(m-1) = f_{ii}^* g_{ii}(m-1).$$

Tehát $g_{ii}(m) = (f_{ii}^*)^m$.

4. Legyen j olyan, hogy $i \to j$, de $j \not\to i$. Létezik olyan állapotsorozat, melyre $\hat{p} = p_{ii_1}p_{i_1i_2}\cdots p_{i_nj} > 0$, és a belső állapotok egyike sem i. Ekkor $f_{ii}^* \le 1 - \hat{p}$.

■

Az utolsó állítás bizonyításához tegyünk egy kis kitérőt a megállási idők világába!

Legyenek adva az X_n valószínűségi változók, és most is $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$. A $\tau : \Omega \to \mathbb{N}$ valószínűségi változó *megállási idő*, ha $\{\tau = n\} \in \mathcal{F}_n$ minden n-re. Ekkor értelmes az $X_\tau : \omega \mapsto X_{\tau(\omega)}(\omega)$ valószínűségi változóról beszélni. Definiálhatjuk még azt a σ-algebrát, amely a megállás időpontjáig megfigyelhető eseményekből áll:

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau = n\} \in \mathcal{F}_n \, \forall n\}.$$

Erről könnyen látható, hogy valóban σ-algebra.

Azt mondjuk, hogy az X_n folyamatra teljesül az *erős Markov-tulajdonság*, ha minden τ véges megállási időre, $k \ge 0$-ra és B mérhető halmazra

$$P(X_{\tau+k} \in B \mid \mathcal{F}_\tau) = P(X_{\tau+k} \in B \mid X_\tau). \tag{1}$$

Vegyük észre, hogy ha $P(\tau = n) = 1$, azaz τ determinisztikus megállási idő, akkor $X_\tau = X_n$ és $\mathcal{F}_\tau = \mathcal{F}_n$, azaz az erős Markov tulajdonságból következik a közönséges. Fordítva általában nem igaz ez az állítás, de a mi esetünkben igen.

2.12. Tétel. *Az $\{X_n\}_{n\ge0}$ Markov láncra teljesül az erős Markov tulajdonság.*

Bizonyítás. Először belátjuk, hogy (1) jobb oldala $\mathcal{F}_\tau$-mérhető, azaz $\sigma(X_\tau) \subseteq \mathcal{F}_\tau$.

$$\{X_\tau \in B\} \cap \{\tau = n\} = \{X_n \in B\} \cap \{\tau = n\} \in \mathcal{F}_n.$$

Másrészt, $\mathcal{F}_\tau$ is atomos, tehát elég az egyenlőséget az atomokon bizonyítani. Egy atom általános alakja:

$$C = \{\tau = m, X_0 = i_0, \dots, X_m = i_m\}.$$

Ezek nyilván $\mathcal{F}_\tau$-beliek, diszjunktak, kiadják a teljes Ω-t, és ha $A \in \mathcal{F}_\tau$, akkor $A \cap C =$ $= (A \cap \{\tau = m\}) \cap C$. Mivel a metszet első tagja $\mathcal{F}_m$-beli, ami atomos, van olyan Z halmaz, melyre

$$A \cap \{\tau = m\} = \cup_{(j_0,\dots,j_m)\in Z}\{X_0 = j_0, \dots, X_m = j_m\}.$$

Ezt C-vel elmetszve, a metszet vagy üres, vagy maga C. Azt kell tehát belátni, hogy

$$P((X_{\tau+k} \in B) \cap C) = \int_C I(X_{\tau+k} \in B) dP =$$
$$\int_C P(X_{\tau+k} \in B | X_\tau) dP = P(X_{\tau+k} \in B | X_\tau = i_m) P(C).$$

A pozitív valószínűségű C atomon a Markov tulajdonság szerint

$$P(X_{\tau+k} \in B \mid C) = \sum_{j \in B} p_{i_m j}^{(k)}.$$

A fentiek szerint azt kell belátni, hogy $P(X_{\tau+k} \in B \mid X_\tau = i_m)$ is ugyanennyi. Ehhez egyrészt azt kell látni, hogy

$$P(X_{\tau+k} \in B \mid X_\tau = i_m, \tau = n) = \sum_{j \in B} p_{i_m j}^{(k)}$$

a közönséges Markov tulajdonság szerint, másrészt legyen $A_n = \{X_\tau = i_m, \tau = n\}$ és ezen diszjunkt események unióját jelölje $A = \cup A_n = \{X_\tau = i_m\}$, ekkor

$$P(X_{\tau+k} \in B \mid A) = \frac{\sum P(X_{\tau+k} \in B \mid A_n) P(A_n)}{\sum P(A_n)} = \sum_{j \in B} p_{i_m j}^{(k)}.$$

Vegyük észre, hogy beláttuk, hogy a τ megállási időtől a ML a múlttól függetlenül újraindul, azaz $P(X_{\tau+k} = j | X_\tau = i) = p_{ij}^{(k)}$. ■

A még bizonyítandó állításra visszatérve, definiáljuk azt a τ_n időpontot, amikor n-edszer térünk vissza az i állapotba. Ekkor τ_n megállási idő, amely 1 valószínűséggel véges, és $X_{\tau_n} = i$. Legyen

$$A_n = \{X_{\tau_n+1}, X_{\tau_n+2}, \ldots, X_{\tau_{n+1}-1} \text{ valamelyike } j\} \in \mathcal{F}_{\tau_{n+1}}.$$

Ekkor A_n független az $\mathcal{F}_{\tau_n}$ σ-algebrától, melybe az $A_1, \ldots, A_{n-1}$ események tartoznak, mivel az erős Markov tulajdonság szerint

$$P(A_n \mid \mathcal{F}_{\tau_n}) = P(A_n \mid X_{\tau_n}) = P(A_n),$$

hiszen $\sigma(X_{\tau_n})$ a triviális σ-algebra. A lánc újraindulási tulajdonságából következik, hogy $P(A_n) = p > 0$, n-től függetlenül (p annak az esélye, hogy i-ből indulva előbb érünk j-be, mint vissza i-be). A Borel-Cantelli lemma szerint 1 valószínűséggel az A_n események közül végtelen sok következik be, azaz $g_{ij} = 1$. Érdemes megjegyezni, hogy ekkor az is igaz, hogy $f_{ij}^* = 1$ és $g_{jj} = 1$, azaz a visszatérőség osztálytulajdonság. ■

2.13. Tétel. *$g_{ij} = 0$ akkor és csak akkor, ha $\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$.*

A bizonyítás előtt egy nagyon hasznos kis lemmát látunk be, mely a Toeplitz szummációs tétel speciális esete.

2.14. Lemma. *(Nörlund) Legyenek $a_n \geq 0$ és b_n valós sorozatok, és $\lim_{n\to\infty} b_n = b$. Ha $\lim_{n\to\infty} a_n / \sum_{k=0}^{n} a_k = 0$, akkor*

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n} a_k b_{n-k}}{\sum_{k=0}^{n} a_k} = b.$$

Bizonyítás. A bizonyítást arra az esetre végezzük el, amikor b véges, a végtelen eset is hasonlóan intézhető el. Legyen B olyan, hogy $|b_n - b| < B$ minden n-re, és adott ϵ-hoz N olyan küszöbindex, hogy $n \geq N$ esetén $|b_n - b| < \epsilon$.

$$|\sum_{k=0}^{n} a_k (b_{n-k} - b)| \leq (\sum_{k=0}^{n-N} a_k)\epsilon + (\sum_{k=n-N+1}^{n} a_k)B.$$

Ebből

$$\limsup_{n\to\infty} \left| \frac{\sum_{k=0}^{n} a_k b_{n-k}}{\sum_{k=0}^{n} a_k} - b \right| \leq \epsilon + B \lim_{n\to\infty} \frac{\sum_{k=n-N+1}^{n} a_k}{\sum_{k=0}^{n} a_k} = \epsilon.$$

■

Megjegyzés: Ha az a_n sorozat korlátos, akkor biztosan teljesíti a lemma feltételét.

Bizonyítás. (2.13 Tételé.)

$$\sum_{r=1}^{n} p_{ij}^{(r)} = \sum_{r=1}^{n} \sum_{k=0}^{r-1} f_{ij}^{(r-k)} p_{jj}^{(k)} = \sum_{k=0}^{n-1} p_{jj}^{(k)} \sum_{s=1}^{n-k} f_{ij}^{(s)} = \sum_{k=0}^{n} a_k b_{n-k},$$

ahol

$$a_k = p_{jj}^{(k)}, \quad b_k = \sum_{s=1}^{k} f_{ij}^{(s)}, \quad b_0 = 0.$$

Ekkor az előző lemmát alkalmazva, $b = f_{ij}^*$, és

$$\lim_{n\to\infty} \frac{\sum_{m=1}^{n} p_{ij}^{(m)}}{\sum_{m=0}^{n} p_{jj}^{(m)}} = f_{ij}^*. \tag{2}$$

Ha ezt az $i = j$ esetre alkalmazzuk, akkor megkapjuk, hogy i akkor és csak akkor visszatérő, ha $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$. Most tegyük fel, hogy $i \neq j$ és $g_{ij} = 0$. Ekkor vagy $f_{ij}^* = 0$, vagy $g_{jj} = 0$. Az első esetben $p_{ij}^{(n)} = 0$ minden n-re, azaz a sor összege nulla. A második esetben j tranziens állapot, ezért $\sum p_{jj}^{(n)}$ véges, és így $\sum p_{ij}^{(n)}$ is. Ha most $g_{ij} > 0$, akkor $g_{jj} = 1$, azaz j rekurrens, és $f_{ij}^* > 0$, amiből $\sum p_{ij}^{(n)} = \infty$ adódik. ■

Egy táblázatban foglalhatjuk össze, hogy tetszőleges két állapot esetén mi mondható el az $f_{ij}^*, g_{ij}, \sum p_{ij}^{(n)}$ mennyiségekről:

	f_{ij}^*	g_{ij}	$\sum p_{ij}^{(n)}$
i, j ugyanabban a rekurrens osztályban	1	1	∞
i, j ugyanabban a tranziens osztályban	> 0	0	$< \infty$
i, j különböző osztályban, $i \to j$, j tranziens	> 0	0	$< \infty$
i, j különböző osztályban, $i \to j$, j rekurrens	$c > 0$	$c > 0$	∞
$i \not\to j$	0	0	0

Mindezek segítségével kiszámolható például az, hogy az egydimenziós bolyongás akkor és csak akkor visszatérő, ha szimmetrikus.

2.15. Példa. Bolyongás a számegyenesen. Legyen a jobbra lépés valószínűsége p, a balra lépésé $q = 1 - p$ $(p, q > 0)$. A lánc irreducibilis, periódusa 2. Vizsgáljuk a visszatérőséget! Elég a 0 állapottal foglalkozni.

$$p_{00}^{(2n)} = \binom{2n}{n} p^n (1-p)^n.$$

Felírható, hogy

$$\sum_{n=0}^{\infty} \binom{2n}{n} x^n = (1-4x)^{-1/2}, \text{ ha } 0 < 4x < 1.$$

Ha tehát $p \neq 1/2$, akkor az átmenetvalószínűségek sorösszege $|1-2p|^{-1} < \infty$, azaz a lánc tranziens. Szimmetrikus esetben viszont a lánc rekurrens ($p_{00}^{(2n)} \sim \frac{1}{\sqrt{\pi n}}$). Tranziens esetben érdekesek az f_{ij}^* valószínűségek. Tegyük fel először, hogy $i > j$. Nyilvánvaló, hogy $f_{ij}^* = (f_{10}^*)^{i-j}$. (2) szerint

$$f_{10}^* = \frac{\sum_{n=1}^{\infty} p_{10}^{(n)}}{\sum_{n=0}^{\infty} p_{00}^{(n)}}.$$

Mármost

$$p_{10}^{(2n+1)} = \binom{2n+1}{n} p^n (1-p)^{n+1} = \frac{1}{2p}\binom{2n+2}{n+1} p^{n+1}(1-p)^{n+1},$$

ezért

$$f_{10}^* = \frac{\frac{1}{2p}\left(\frac{1}{|1-2p|} - 1\right)}{\frac{1}{|1-2p|}} = \frac{1}{2p}(1 - |1-2p|) = \begin{cases} 1, & \text{ha} \quad p < 1/2 \\ (1-p)/p, & \text{ha} \quad p > 1/2 \end{cases}.$$

Hasonlóan járhatunk el, ha $i < j$, ehhez csak az

$$f_{01}^* = \begin{cases} p/(1-p), & \text{ha} \quad p < 1/2 \\ 1, & \text{ha} \quad p > 1/2 \end{cases}$$

mennyiség kell. Végül pedig $f_{ii}^* = f_{00}^*$, és

$$f_{00}^* = \frac{\sum_{n=1}^{\infty} p_{00}^{(2n)}}{\sum_{n=0}^{\infty} p_{00}^{(2n)}} = 1 - |1-2p| = \begin{cases} 2p, & \text{ha} \quad p < 1/2 \\ 2(1-p), & \text{ha} \quad p > 1/2 \end{cases}.$$

2.4. Az átmenetvalószínűségek konvergenciája

Az eddigiekből az következik, hogy ha j átmeneti állapot, akkor $p_{ij}^{(n)} \to 0$. Vajon rekurrens állapot esetén mit mondhatunk erről a határértékről?

2.16. Tétel. *Ha i rekurrens állapot d periódussal, akkor* $\lim_{n\to\infty} p_{ii}^{(nd)} = \frac{d}{m_i}$.

Bizonyítás. Az egyszerűség kedvéért vezessük be a következő jelöléseket: $p_n = p_{ii}^{(n)}$, $f_n = f_{ii}^{(n)}$. Először belátjuk, hogy a

$$\{n \geq 1 : f_n > 0\}$$

halmaz legnagyobb közös osztója d. Legyen ugyanis ez a legnagyobb közös osztó d_f. Mivel

$$\{n \geq 1 : f_n > 0\} \subset \{n \geq 1 : p_n > 0\},$$

$d | d_f$. Másrészt, ha $p_n > 0$, akkor n előáll $n = n_1 + \ldots + n_k$ alakban, ahol $f_{n_i} > 0$. Ezért $d_f | n$, azaz d_f közös osztó, és így $d_f | d$.

Legyen ezután $r_n = \sum_{k=n+1}^{\infty} f_k$, azaz annak a valószínűsége, hogy az első n lépés alatt nem térünk vissza i-be. Igazak a következő összefüggések:

$$\sum_{k=0}^{\infty} r_k = \sum_{k=1}^{\infty} k f_k = m_i, \quad \sum_{k=0}^{n} r_k p_{n-k} = 1,$$

a második összefüggés úgy adódik, hogy a lánc lehetséges meneteit az n. lépésig felbontjuk aszerint, hogy hányadik lépésben járt utoljára az i állapotban (éppen az $n-k$. lépésben).

Ha $\limsup_{n\to\infty} p_{nd} = \lambda$, akkor válasszunk egy olyan részsorozatot, melyre

$$\lim_{k\to\infty} p_{n_k d} = \lambda.$$

Tetszőleges olyan s-re, melyre $f_s > 0$,

$$\lambda = \lim_{k\to\infty} p_{n_k d} = \lim_{k\to\infty} (f_s p_{n_k d - s} + \sum_{\nu=1,\nu\neq s}^{n_k d} f_\nu p_{n_k d-\nu}),$$

mivel $p_n = \sum_{j=1}^{n} f_j p_{n-j}$. Tetszőleges $\epsilon > 0$-hoz legyen N olyan nagy, hogy $\sum_{n=N}^{\infty} f_n < \epsilon$. Ekkor

$$\sum_{\nu=1,\nu\neq s}^{n_k d} f_\nu p_{n_k d-\nu} = \sum_{\nu=N,\nu\neq s}^{n_k d} f_\nu p_{n_k d-\nu} + \sum_{\nu=1,\nu\neq s}^{N-1} f_\nu p_{n_k d-\nu} \leq$$

$$\epsilon + \left(\sum_{\nu=1,\nu\neq s}^{N-1} f_\nu\right) \sup_{\nu<N} p_{n_k d-\nu}.$$

Felhasználva, hogy $\lim(a_n + b_n) \leq \liminf a_n + \limsup b_n$, folytathatjuk a fenti sort:

$$\lambda \leq f_s \liminf_{k\to\infty} p_{n_k d-s} + \left(\sum_{\nu=1,\nu\neq s}^{\infty} f_\nu\right) \lambda = f_s \liminf_{k\to\infty} p_{n_k d-s} + (1-f_s)\lambda.$$

Ezt átrendezve azonnal következik, hogy $\liminf_{k\to\infty} p_{n_k d-s} \geq \lambda$, azaz $\lim_{k\to\infty} p_{n_k d-s} = \lambda$. Ugyanez igaz akkor is, ha $s = \sum c_j s_j$ alakú, ahol $c_j > 0$ egész szám, és $f_{s_j} > 0$. Mivel minden elég nagy t esetén td felírható ilyen alakban (ezt a részosztályokról szóló tételnél bizonyítottuk), elmondhatjuk hogy $t \geq t_0$ esetén $\lim_{k\to\infty} p_{(n_k-t)d} = \lambda$. Mármost

$$1 = \sum_{h=0}^{(n_k-t_0)d} r_h p_{(n_k-t_0)d-h} = \sum_{\nu=0}^{n_k-t_0} r_{\nu d} p_{(n_k-t_0-\nu)d},$$

hiszen p_n csak akkor lehet pozitív, ha $d \mid n$. Egy olyan jellegű összeg jelent meg, mint a Nörlund lemmában, azaz $\sum_{\nu=0}^{m_k} a_\nu b_{m_k-\nu}$, csakhogy most a b_m sorozatról nem tudjuk, hogy konvergens, csak a $b_{m_k-\nu}$ sorozatokról, minden rögzített ν-re.

Tegyük fel először, hogy $\sum_{\nu=0}^{\infty} r_{\nu d} = \infty$, ekkor minden N-re

$$1 \geq \lim_{k\to\infty} \sum_{\nu=0}^{N} r_{\nu d} p_{(n_k-t_0-\nu)d} = \lambda \sum_{\nu=0}^{N} r_{\nu d},$$

tehát $\lambda = 0$. Ha viszont $\sum_{\nu=0}^{\infty} r_{\nu d} < \infty$, akkor a Nörlund lemmához hasonló bizonyítás működik: válasszuk N-et olyan nagyra, hogy $\sum_{\nu=N+1}^{\infty} r_{\nu d} < \epsilon$ legyen. Elég nagy k-ra

$$\left| \sum_{\nu=0}^{n_k-t_0} r_{\nu d}(p_{(n_k-t_0-\nu)d} - \lambda) \right| \leq \sum_{\nu=0}^{N} r_{\nu d} \mid p_{(n_k-t_0-\nu)d} - \lambda \mid +\epsilon \leq 2\epsilon.$$

Tehát

$$1 - \lambda \sum_{\nu=0}^{n_k-t_0} r_{\nu d} \to 0,$$

amiből $\lambda = 1/\sum r_{\nu d}$ következik (és ez a végtelen esetben is érvényes).

Írjuk át a tört nevezőjét:

$$\sum_{\nu=0}^{\infty} r_{\nu d} = \sum_{\nu=0}^{\infty} \frac{1}{d} \sum_{j=\nu d}^{\nu d+d-1} r_j = \frac{1}{d} \sum_{j=0}^{\infty} r_j = \frac{m_i}{d}.$$

Megkaptuk tehát, hogy a sorozat limsupja a kívánt érték. Ha most $\liminf_{n\to\infty} p_{nd} = \eta$, akkor az előző gondolatmenet lemásolásával azt kapjuk, hogy $\eta = \lambda$, azaz $\lim_{n\to\infty} p_{nd}$ létezik. ∎

Azt kaptuk tehát, hogy az i állapot akkor és csak akkor pozitív rekurrens, ha a $\lim_{n\to\infty} p_{ii}^{(nd)}$ határérték pozitív. Az is könnyen látszik, hogy a pozitivitás osztálytulajdonság: ha i és j ugyanabban a rekurrens osztályban vannak, akkor van olyan n és m, hogy $p_{ij}^{(n)} > 0$ és $p_{ji}^{(m)} > 0$. A $p_{ii}^{(m+kd+n)} \geq p_{ij}^{(n)} p_{jj}^{(kd)} p_{ji}^{(m)}$ kifejezésben k-val végtelenhez tartva azt kapjuk, hogy ha j pozitív állapot, akkor szükségképpen i is az.

Most már könnyen bebizonyítható a következő általános tétel.

2.17. Tétel. *Legyen i, j két tetszőleges állapot, és jelölje j periódusát d. Ekkor $r = 1,2,\ldots,d$ esetén*

$$\lim_{n\to\infty} p_{ij}^{(nd+r)} = f_{ij}^*(r)\frac{d}{m_j},$$

ahol $f^*_{ij}(r) \geq 0$ *és* $\sum_{r=1}^d f^*_{ij}(r) = f^*_{ij}$. *(Tranziens állapotra legyen* $m_j = \infty$.*)*

Bizonyítás. Legyen $f^*_{ij}(r) = \sum_{n=0}^\infty f_{ij}^{(nd+r)}$. Ekkor

$$p_{ij}^{(nd+r)} = \sum_{k=0}^n f_{ij}^{(kd+r)} p_{jj}^{(nd-kd)}.$$

A Nörlund lemmából azonnal következik az állítás, $a_k = f_{ij}^{(kd+r)}$, $b_k = p_{jj}^{(kd)}$ szereposztással. ■

2.18. Következmény. *Minden* i, j *állapotpárra* $\lim n^{-1} \sum_{k=1}^n p_{ij}^{(k)} = f^*_{ij}/m_j$. *Itt a bal oldal azt fejezi ki, hogy* i*-ből indulva, a lánc várhatóan a lépések hányad részét tölti a* j *állapotban.*

2.19. Példa. Az egydimenziós szimmetrikus bolyongás nulla rekurrens.

2.20. Példa. Vegyünk egy bolyongást a nemnegatív számokon, a nullában egy visszaverő fallal. Legyen a jobbra lépés esélye $p < 1/2$, a balra lépésé $q = 1 - p$. A lánc irreducibilis, periódusa $d = 2$. Korábbi számolásunk alapján a lánc visszatérő. A magasabb rendű átmenetvalószínűségeket nem könnyű kiszámítani, viszont az m_{ij} átlagos elérési időkre (átlagosan ennyi lépés alatt érünk i-ből j-be) fel tudunk írni egyenleteket.

$$m_{00} = 1 + m_{10},\ m_{10} = q + p(1 + m_{20}),\ m_{20} = m_{21} + m_{10},\ m_{21} = m_{10}.$$

Ezeknek az egyenleteknek a végtelen is megoldása, tegyük azonban fel, hogy a mennyiségek végesek. Ekkor az

$$m_{00} = 1 + \frac{1}{1-2p},\ \frac{1}{m_{00}} = \frac{1-2p}{2-2p},\ \frac{d}{m_{00}} = 1 - \frac{p}{q}$$

megoldás adódik. Ha tehát a lánc 1/4 eséllyel lép jobbra, és 3/4 eséllyel balra, akkor a nulla állapotból átlagosan 3 lépés alatt ér vissza a nullába, egy távoli páros időpontban ránézve a láncra, kb. 2/3 eséllyel lesz éppen a nulla állapotban, és a láncot sokáig futtatva, nagyjából a lépések 1/3-át tölti a nulla állapotban (és ez utóbbi nem függ a kezdeti eloszlástól).

2.5. Stacionárius eloszlás

2.21. Definíció. *Legyen* P *egy átmenetvalószínűségmátrix. A* $(p_i)_{i \in I}$ *eloszlás stacionárius, ha* $p_i = \sum_{k \in I} p_k p_{ki}$ *minden* $i \in I$*-re, azaz a* p_i *kezdeti eloszlású,* P *átmenetva-*

lószínűségű X_n Markov láncra $P(X_n = i) = p_i$ minden $i \in I, n \geq 0$-ra. Ez utóbbi esetben azt mondjuk, hogy a Markov lánc stacionárius (ez megfelel a szokásos definíciónak, azaz hogy a véges dimenziós eloszlások eltolás-invariánsak).

2.22. Tétel. *Legyen C lényeges osztály, és jelölje $\pi_i = 1/m_i$. Ekkor az*

$$u_i = \sum_{k\in C} u_k p_{ki}, \ i \in C$$

egyenletrendszer $\sum_{i\in C} |u_i| < \infty$ megoldásai: $u_i = c\pi_i$.

Bizonyítás. Először belátjuk, hogy ezek megoldások. Ha C átmeneti vagy nulla rekurrens, akkor $\pi_i = 0$. Legyen C pozitív rekurrens, jelölje periódusát d.

$$p_{ii}^{(nd)} = \sum_{k\in C} p_{ik}^{(nd-1)} p_{ki} = \sum_{k\in C_{-1}(i)} p_{ik}^{(nd-1)} p_{ki}.$$

Tartson n végtelenhez:

$$d\pi_i = \lim p_{ii}^{(nd)} \geq \sum_{k\in C_{-1}(i)} \lim p_{ik}^{(nd-1)} p_{ki} = \sum_{k\in C_{-1}(i)} d\pi_k p_{ki},$$

azaz $\pi_i \geq \sum_{k\in C} \pi_k p_{ki}$. Másrészt, mivel

$$1 = \lim_{n\to\infty} \sum_{j\in C_r(i)} p_{ij}^{(nd+r)} \geq \sum_{j\in C_r(i)} \lim_{n\to\infty} p_{ij}^{(nd+r)} = \sum_{j\in C_r(i)} d\pi_j,$$

ezért $\sum_{i\in C} \pi_i \leq 1$.

$$\sum_{i\in C} \pi_i \geq \sum_{i\in C} \sum_{k\in C} \pi_k p_{ki} = \sum_{k\in C} \pi_k \sum_{i\in C} p_{ki} = \sum_{k\in C} \pi_k,$$

ami a $\sum \pi_i$ végessége miatt csak úgy lehet, ha $\pi_i = \sum_{k\in C} \pi_k p_{ki}$ minden i-re.

Ezután meg kell mutatni, hogy nincs más megoldás. Legyen u_i egy abszolút konvergens megoldás. Ekkor

$$u_i = \sum_{k\in C_{-1}(i)} u_k p_{ki} = \sum_{k\in C_{-1}(i)} \sum_{l\in C_{-1}(k)} u_l p_{lk} p_{ki} = \sum_{l\in C_{-2}(i)} u_l \sum_{k\in C_{-1}(i)} p_{lk} p_{ki} = \sum_{l\in C_{-2}(i)} u_l p_{li}^{(2)},$$

(a szummák az abszolút konvergencia miatt felcserélhetők), ezt iterálva kapjuk, hogy minden n, r-re

$$u_i = \sum_{k \in C_{-r}(i)} u_k p_{ki}^{(nd+r)}.$$

Mivel az összeg tagjainak van konvergens majoránsa, határértéket véve kapjuk, hogy

$$u_i = \left(\sum_{k \in C_{-r}(i)} u_k \right) d\pi_i, \forall r.$$

Ha C átmeneti vagy nulla rekurrens, akkor $u_i = 0$, ha pedig pozitív rekurrens, akkor szükségképpen $\sum_{k \in C_{-r}(i)} u_k = K$ konstans, és $u_i = K d\pi_i$. ∎

Megjegyzés: az $u_i = \pi_i$ megoldásra tehát visszahelyettesítéssel:

$$\pi_i = \left(\sum_{k \in C_{-r}(i)} \pi_k \right) d\pi_i, \text{ azaz } \sum_{k \in C_{-r}(i)} \pi_k = \frac{1}{d},$$

amiből $\sum_{i \in C} \pi_i = 1$ is következik. A C pozitív osztályon tehát π_i stacionárius kezdeti eloszlás.

2.23. Tétel. *Legyen adott a P sztochasztikus mátrix az I állapottéren. Jelölje a pozitív osztályokat $D_\alpha : \alpha \in A$, és legyen $D = \cup_\alpha D_\alpha$ a pozitív állapotok halmaza. Ekkor a p_i eloszlás akkor és csak akkor stacionárius, ha*

$$p_i = \begin{cases} 0 & ha \quad i \notin D, \\ \lambda_\alpha \pi_i & ha \quad i \in D_\alpha, \end{cases}$$

ahol $\lambda_\alpha \geq 0$, $\sum_\alpha \lambda_\alpha = 1$.

Bizonyítás. $\Rightarrow$: Ha p_i stacionárius eloszlás, akkor minden n-re $p_i = \sum_{j \in I} p_j p_{ji}^{(n)}$. Ha $i \notin D$, akkor határátmenettel

$$p_i = \sum_{j \in I} p_j \lim p_{ji}^{(n)} = 0,$$

ha pedig $i \in D_\alpha$, akkor az előző miatt

$$p_i = \sum_{j \in D} p_j p_{ji}^{(n)} = \sum_{j \in D_\alpha} p_j p_{ji}^{(n)},$$

hiszen másik pozitív osztályból nem lehet i-be jutni. Ezért minden N-re

$$p_i = \sum_{j \in D_\alpha} p_j \frac{1}{N} \sum_{n=1}^{N} p_{ji}^{(n)},$$

amiből határátmenettel $p_i = (\sum_{j \in D_\alpha} p_j)\pi_i$, azaz $\lambda_\alpha = \sum_{j \in D_\alpha} p_j$.
$\Leftarrow$: Ha $i \notin D$, akkor $P(X_n = i) = 0$ minden n-re, hiszen a pozitív osztályokból nem lép ki a lánc. Ha viszont $i \in D_\alpha$, akkor

$$P(X_n = i) = \sum_{k \in D_\alpha} p_k p_{ki}^{(n)} = \lambda_\alpha \sum_{k \in D_\alpha} \pi_k p_{ki}^{(n)} = \lambda_\alpha \pi_i = p_i,$$

hiszen az előző tétel bizonyításánál láttuk, hogy π_i stacionárius eloszlás D_α-n, azaz ha $i \in D_\alpha$, akkor

$$\pi_i = \sum_{k \in D_\alpha} \pi_k p_{ki}^{(n)}.$$

■

Összefoglalva azt kaptuk, hogy egy stacionárius Markov lánc csak pozitív osztályokból áll. Egy osztályban egyértelműen létezik stacionárius eloszlás, a Markov lánc eloszlása ezen stacionárius eloszlások keveréke. Az osztályokat még részosztályokra is szét lehet bontani ($nd+r$ alakú időpontokban ránézve), így irreducibilis, pozitív rekurrens, aperiodikus láncokat kapunk, egyértelmű stacionárius eloszlással, és tetszőleges kezdeti eloszlás esetén X_n eloszlása a stacionárius eloszláshoz tart. (Tetszőleges Markov láncot tetszőleges eloszlásból elindítva vizsgálhatjuk, hogy X_n eloszlása hová tart.)

2.24. Példa. A most belátottak segítségével bizonyítható, hogy a korábbi példában szereplő egydimenziós bolyongás a 0-ban visszaverő fallal pozitív rekurrens. Megoldhatók ugyanis a stacionárius eloszlás egyenletei, és kapjuk, hogy

$$\pi_0 = \frac{1-2p}{2-2p}, \quad \pi_i = \pi_0 \frac{p^{i-1}}{q^i},\ i \geq 1.$$

2.25. Példa. Egy érmét, melyen a fej valószínűsége p, dobálva, jelölje X_n, hogy az első n dobásból alkotott sorozat végén hány fej van. X_n irreducibilis, aperiodikus Markov láncot alkot, $P(X_0 = 0) = 1$, és $p_{k,k+1} = p$, $p_{k,0} = 1 - p = q$. Keressünk stacionárius eloszlást!

(i) Oldjuk meg a $\pi = \pi P$ egyenletrendszert:

$$\pi_0 = q \sum_{k=0}^{\infty} \pi_k = q,\ \ \pi_i = p\pi_{i-1} = p^i q,$$

azaz a stacionárius kezdeti eloszlás $Geo(q) - 1$.
(ii) Tekintsük a $p_{ik}^{(n)}$ átmenetvalószínűségek határértékét. Ha $n > k$, akkor $p_{ik}^{(n)} = p^k q$. Ennek segítségével könnyen kiszámítható, hogy átlagosan hány dobás kell ahhoz, hogy egy k hosszú fejsorozat megjelenjen. Legyen ez a mennyiség μ_k. Erre

$$m_k = m_{k0} + m_{0k} = 1/q + \mu_k,$$

és mivel $m_k = 1/\pi_k$, kapjuk, hogy $\mu_k = \frac{1-p^k}{qp^k}$.

3. Pozitív rekurrens Markov láncok

3.1. Definíció. *Nevezzük az X_n Markov láncot ergodikusnak, ha irreducibilis és pozitív rekurrens. (Figyelem: az irodalomban nem feltétlenül ezt nevezik ergodikusnak!!)*

Láttuk, hogy az ilyen Markov láncok hosszú távú viselkedését különösen egyszerű leírni (ha még aperiodikusak is, akkor ez még inkább igaz).

3.1. Nagy számok törvénye

3.2. Tétel. *Legyen X_n ergodikus Markov lánc π_i stacionárius eloszlással, és $f : I \to \mathbb{R}$ tetszőleges függvény. Ha $\mathcal{I}(f) = \sum_{i \in I} f(i)\pi_i$ sor abszolút konvergens, akkor*

$$\frac{1}{n} \sum_{k=0}^{n} f(X_k) \to \mathcal{I}(f)\ 1 \text{ valószínűséggel.}$$

Megjegyzés: A $Z_n = f(X_n)$ sorozat nem feltétlenül Markov lánc.

Bizonyítás. Legyen $i \in I$ tetszőleges rögzített állapot, definiáljuk azokat a megállási időket, melyek az i-be tett látogatások időpontjait adják meg:

$$0 = \tau_0 < \tau_1 < \dots,$$

τ_n az n. elérési időpont. Ezek 1 valószínűséggel véges megállási idők. Minden n-re definiáljuk azt az $l(n)$ valószínűségi változót, amely megmondja, hogy az n. lépésig hányszor járt a lánc i-ben:

$$\tau_{l(n)} \leq n < \tau_{l(n)+1}.$$

A Z_k-k összegére a következő felbontási formulát használjuk:

$$\sum_{k=0}^{n} Z_k = \sum_{k=0}^{\tau_1 - 1} Z_k + \sum_{h=1}^{l(n)-1} \sum_{s=\tau_h}^{\tau_{h+1}-1} Z_s + \sum_{k=\tau_{l(n)}}^{n} Z_k = Y' + \sum_{h=1}^{l(n)-1} Y_h + Y''(n).$$

Pozitív és negatív részre való felbontással feltehető, hogy $f \geq 0$. Ekkor

$$\frac{1}{l(n)} \sum_{h=1}^{l(n)-1} Y_h \leq \frac{1}{l(n)} \sum_{k=0}^{n} Z_k \leq \frac{1}{l(n)} \left(Y' + \sum_{h=1}^{l(n)} Y_h \right).$$

Korábban már meggondoltuk, hogy az erős Markov-tulajdonság miatt az Y_h változók függetlenek, és azonos eloszlásúak. Mivel 1 valószínűséggel végtelen sokszor jár a lánc i-ben, $l(n)$ 1 valószínűséggel végtelenhez tart. Ezért a nagy számok erős törvénye szerint

$$\frac{1}{l(n)} \sum_{h=1}^{l(n)} Y_h \to E(Y_h) \text{ 1 valószínűséggel.}$$

Másrészt $Y'/l(n) \to 0$ (1 valószínűséggel), tehát $\frac{1}{l(n)} \sum_{k=0}^{n} Z_k$ is tart $E(Y_h)$-hoz 1 valószínűséggel. Számítsuk ki ezt a várható értéket! Az egyszerűség kedvéért tegyük fel, hogy a lánc i-ből indul. Ekkor

$$E(Y_h) = E\left(\sum_{k=0}^{\infty} Z_k \chi\{k < \tau_1\} \right) = \sum_{k=0}^{\infty} \sum_{j \in I} f(j) P(X_k = j, k < \tau_1) = \sum_{j \in I} f(j) u_j,$$

ahol $u_j = \sum_{k=0}^{\infty} P(X_k = j, k < \tau_1)$, azt fejezi ki, hogy két i-ben tett látogatás között a lánc átlagosan hányszor jár j-ben (a kezdőpontot beleszámítva, a végpontot nem). Ezért $\sum_{j \in I} u_j = m_i < \infty$. Megmutatjuk, hogy u_j kielégíti a stacionárius eloszlás egyenletrendszerét.

$$\sum_j u_j p_{jl} = \sum_j \sum_{k=0}^{\infty} P(X_k = j, k < \tau_1) P(X_{k+1} = l \mid X_k = j) =$$
$$\sum_{k=0}^{\infty} E\left[P(X_{k+1} = l \mid X_k)\chi\{k < \tau_1\}\right] = \sum_{k=0}^{\infty} E\left[P(X_{k+1} = l \mid \mathcal{F}_k)\chi\{k < \tau_1\}\right],$$

ahol az utolsó lépésben a Markov tulajdonságot használtuk. Felhasználva, hogy $\chi\{k < < \tau_1\}$ $\mathcal{F}_k$-mérhető, folytathatjuk:

$$\sum_j u_j p_{jl} = \sum_{k=0}^{\infty} E\left[E(\chi\{X_{k+1} = l, k < \tau_1\} \mid \mathcal{F}_k)\right] =$$
$$\sum_{k=0}^{\infty} P(X_{k+1} = l, k < \tau_1) = \sum_{k=1}^{\infty} P(X_k = l, k \leq \tau_1).$$

Ez pedig tényleg u_l, hiszen a kezdőpontot kihagytuk, a végpontot viszont bevettük, de mivel mindkettőben az i állapotban van a lánc, az összeg nem változott. Ezért $u_j = c\pi_j$. Ha most $i = j$, akkor $1 = u_i = c\pi_i$, azaz $u_j = \pi_j/\pi_i$. Tehát

$$E(Y_h) = \sum_{j \in I} f(j) u_j = \frac{\mathcal{I}(f)}{\pi_i}.$$

Ha most $f(j) = 1$ minden j-re, akkor azt kapjuk, hogy

$$\frac{n+1}{l(n)} = \frac{1}{l(n)} \sum_{k=0}^{n} Z_k \to \frac{1}{\pi_i},$$

azaz

$$\frac{1}{n+1} \sum_{k=0}^{n} Z_k = \frac{l(n)}{n+1} \frac{1}{l(n)} \sum_{k=0}^{n} Z_k \to \mathcal{I}(f) \text{ 1 valószínűséggel.}$$

■

3.3. Példa. Legyen $\underline{i} = (i_1, \ldots, i_k)$ egy k hosszú állapotsorozat, ahol $p_{i_1 i_2} \cdot p_{i_2 i_3} \cdots p_{i_{k-1} i_k} > > 0$. Hová tart az $\underline{i}$ sorozat relatív gyakorisága n lépésből?

Válasz: $\pi_{i_1} \cdot p_{i_1 i_2} \cdot p_{i_2 i_3} \cdots p_{i_{k-1} i_k}$-hez. (Készítünk egy új ML-ot, melynek állapotai a k hosszú sorozatok, ez is ergodikus lesz, tehát alkalmazható rá a tétel).

3.2. Centrális határeloszlás-tétel

Ugyanabban a felállásban, mint az előbb, vizsgáljuk, hogy a részletösszegeket hogyan kell normálni, hogy normális határeloszlást kapjunk. Két lemmára lesz szükségünk.

3.4. Lemma. *$P(n - \tau_{l(n)} \geq t) \leq c_t$, ahol $\lim_{t\to\infty} c_t = 0$.*

Bizonyítás. Legyen $n \geq t$. Jelölje az i-be való visszatérés lépésszámát τ.

$$P(n - \tau_{l(n)} \geq t) = \sum_{s=t}^{n} P(n - \tau_{l(n)} = s) =$$
$$\sum_{s=t}^{n-1} P(X_{n-s} = i)P(\tau > s) + P(\tau_1 > n) \leq \sum_{s=t}^{\infty} P(\tau > s) + P(\tau_1 > t) = c_t,$$

és ez valóban 0-hoz tart, mivel mind τ várható értéke véges, τ_1 pedig 1 valószínűséggel véges. ■

3.5. Lemma. *A felbontási formula jelölésével, $\frac{Y''(n)}{\sqrt{n}} \to 0$ sztochasztikusan.*

Bizonyítás. Legyen ϵ, δ adott. Minden t-re

$$P(|Y''(n)| \geq \sqrt{n}\epsilon) \leq P(n - \tau_{l(n)} > t) + P\left(\max_{0<s\leq t} |\sum_{j=\tau_{l(n)}}^{\tau_{l(n)}+s} Z_j| \geq \sqrt{n}\epsilon\right).$$

Ha t elég nagy, akkor az első tag $< \delta/2$ minden n-re, ezek után ha n elég nagy, akkor a második tag is $< \delta/2$, mivel a zárójelben egy 1 valószínűséggel véges, n-től független valószínűségi változó áll. ■

Bizonyítás nélkül idézzük fel a Kolmogorov egyenlőtlenséget!

3.6. Lemma. *Legyenek V_i független, nulla várható értékű, véges szórású valószínűségi változók. Ekkor minden $c > 0$-ra*

$$P(\max_{1\leq k\leq n} |\sum_{i=1}^{k} V_i| \geq c) \leq \frac{\sum_{i=1}^{n} D^2(V_i)}{c^2}.$$

3.7. Tétel. *Teljesüljenek az előző tétel feltételei, és legyen* $V_h = Y_h - \mathcal{I}(f)(\tau_{h+1} - \tau_h)$. *Ha* $0 < E(V_h^2) < \infty$, *akkor*

$$\frac{\sum_{k=0}^{n} f(X_k) - n\mathcal{I}(f)}{\sqrt{n\pi_i E(V_h^2)}} \to N(0{,}1).$$

Bizonyítás. A V_h valószínűségi változók függetlenek, azonos eloszásúak, 0 várható értékűek. A felbontási formula szerint

$$\sum_{k=0}^{n} Z_k - n\mathcal{I}(f) = \sum_{h=1}^{l(n)-1} V_h + Y' + Y''(n) - \mathcal{I}(f)(n - \tau_{l(n)} + \tau_1).$$

Azt már tudjuk, hogy $Y'/\sqrt{n}$ sztochasztikusan 0-hoz tart. A 3.4 lemma szerint ugyanez igaz az $(n - \tau_{l(n)})/\sqrt{n}$ tagra, a 3.5 lemma miatt pedig $Y''(n)/\sqrt{n}$ is sztochasztikusan 0-hoz tart. A Cramér-Szlujkij lemma szerint ezért elég csak a V_h-k összegével foglalkozni, azaz megmutatni, hogy

$$\frac{\sum_{h=1}^{l(n)-1} V_h}{\sqrt{n\pi_i E(V_h^2)}} \to N(0{,}1).$$

Szeretnénk a CHT-re hivatkozni, azonban itt az összegzés egy véletlen indexig történik, ettől kellene megszabadulni.

Legyen $n^* = [n\pi_i]$. Azt tudjuk, hogy

$$\frac{\sum_{h=1}^{n^*} V_h}{\sqrt{n^* E(V_h^2)}} \to N(0{,}1),$$

itt kellene az összegzésben n^*-ot $l(n) - 1$-re cserélni. Tudjuk, hogy ezek nagy valószínűséggel közel vannak egymáshoz. Legyen ϵ, δ adott, $n' = n\pi_i(1 - \delta)$, $n'' = n\pi_i(1 + \delta)$, és

$$A_m = \{n' < l(n) - 1 < n'',\ \forall n \geq m\}.$$

Ezek bővülő események, és

$$\left\{\frac{l(n)}{n} \to \pi_i\right\} \subset \cup_m A_m.$$

Mivel a bal oldali esemény valószínűsége 1, ezért van olyan m, melyre $P(A_n) > 1 - \epsilon$ minden $n \geq m$-re. Ha pedig $l(n) - 1$ és n^* már közel vannak egymáshoz, akkor használhatjuk a Kolmogorov-egyenlőtlenséget.

$$P\left(\left|\sum_{h=1}^{l(n)-1} V_h - \sum_{h=1}^{n^*} V_h\right| \geq c\sqrt{n^* E(V_h^2)}\right) \leq \epsilon + \frac{2\delta n \pi_i E(V_h^2)}{c^2 n^* E(V_h^2)} < 2\epsilon,$$

ha δ elég kicsi, és n elég nagy. Tehát $l(n)-1$-et n^*-ra cserélve, a különbség sztochasztikusan 0-hoz tart, így a Cramér-Szluckij lemmára való ismételt hivatkozással készen vagyunk. ∎

Vajon hogyan lehet az $E(V_h^2)$ mennyiséget kiszámítani? Már a NSzT-e bizonyításánál észrevehettük, hogy Y_h várható értéke kiszámolásakor olyan valószínűségek bukkantak fel, hogy n lépés alatt i-ből j-be megyünk, de közben nem járunk i-ben.

3.3. Tabu állapotok

Legyen X_n ML, és $H \subseteq I$ tetszőleges.

3.8. Definíció. *Átmenetvalószínűségek tabu állapotokkal:*

$${}_Hp_{ij}^{(n)} = P(X_n = j, X_m \notin H \ 0 < m < n | X_0 = i).$$

Jelölje ${}_Hf_{ij}^{(n)} = {}_{j,H}p_{ij}^{(n)}$. *Legyen* ${}_Hp_{ij}^{(0)} = I(i = j, i \notin H)$.

Legyen még ${}_Hp_{ij}^* = \sum_{n=1}^{\infty} {}_Hp_{ij}^{(n)}$, mely azt adja meg, hogy i-ből indulva, várhatóan hányszor jár a lánc j-ben, míg H-ba beér (a beérést is beszámítva).

3.9. Lemma. *Alapformulák: legyen* $n \geq 1$, $k \notin H$. *Ekkor*

$$\begin{array}{rclcll}
{}_Hp_{ij}^{(n)} & = & {}_{k,H}p_{ij}^{(n)} & + & \sum_{s=1}^{n-1} {}_{k,H}p_{ik}^{(s)} \cdot {}_Hp_{kj}^{(n-s)} & (1A) \\
{}_Hp_{ij}^{(n)} & = & {}_{k,H}p_{ij}^{(n)} & + & \sum_{s=1}^{n-1} {}_Hp_{ik}^{(s)} \cdot {}_{k,H}p_{kj}^{(n-s)} & (2A) \\
{}_Hp_{ij}^{*} & = & {}_{k,H}p_{ij}^{*} & + & {}_{k,H}p_{ik}^{*} \cdot {}_Hp_{kj}^{*} & (1B) \\
{}_Hp_{ij}^{*} & = & {}_{k,H}p_{ij}^{*} & + & {}_Hp_{ik}^{*} \cdot {}_{k,H}p_{kj}^{*} & (2B)
\end{array}$$

Bizonyítás. Az 1-es formulák a k első, a 2-esek a k utolsó elérése szerinti felbontásból adódnak, a B formulák pedig az A-k összegzésével keletkeznek. ∎

3.10. Definíció. *Legyen* $m_{ij} = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$ *az átlagos elérési idő, általában pedig* ${}_Hm_{ij} = \sum_{n=1}^{\infty} n\, {}_Hf_{ij}^{(n)}$.

Megjegyzések: ${}_H f^*_{ij}$: annak valószínűsége, hogy i-ből indulva, a lánc előbb ér j-be, mint H-ba. ${}_H m_{ij} \le m_{ij}$. Ha C pozitív osztály, akkor $m_{ij} < \infty$ minden $i, j \in C$.

3.11. Lemma. *Legyen $i, j, k \in C$, C pozitív osztály, és $j \ne k$. Ekkor*

$$ {}_k f^*_{ij} = \frac{m_{jk} + m_{kj}}{m_{ik} + m_{kj} - m_{ij}}, $$

továbbá

$$ {}_k p^*_{ij} = \frac{m_{ik} + m_{kj} - m_{ij}}{m_{jj}}. $$

Bizonyítás. Először is, ${}_k f^*_{ij} + {}_j f^*_{ik} = 1$. Másrészt az $(1A)$ formula szerint

$$ f^{(n)}_{ij} = {}_j p^{(n)}_{ij} = {}_{k,j} p^{(n)}_{ij} + \sum_{s=1}^{n-1} {}_{k,j} p^{(s)}_{ik} \cdot {}_j p^{(n-s)}_{kj} = {}_k f^{(n)}_{ij} + \sum_{s=1}^{n-1} {}_j f^{(s)}_{ik} \cdot f^{(n-s)}_{kj}. $$

Emiatt

$$ \begin{aligned} m_{ij} &= {}_k m_{ij} + \sum_{n=1}^{\infty} n \sum_{s=1}^{n-1} {}_j f^{(s)}_{ik} \cdot f^{(n-s)}_{kj} = \\ &= {}_k m_{ij} + \sum_{s=1}^{\infty} {}_j f^{(s)}_{ik} \sum_{n=s+1}^{\infty} ((n-s) f^{(n-s)}_{kj} + s f^{(n-s)}_{kj}) = {}_k m_{ij} + m_{kj} \cdot {}_j f^*_{ik} + {}_j m_{ik}. \end{aligned} $$

Megcserélve j és k szerepét, kapjuk hogy

$$ m_{ik} + m_{kj} - m_{ij} = m_{kj} + m_{jk} \cdot {}_k f^*_{ij} - m_{kj} \cdot {}_j f^*_{ik} = m_{jk} \cdot {}_k f^*_{ij} + m_{kj}(1 - {}_j f^*_{ik}) = {}_k f^*_{ij}(m_{jk} + m_{kj}), $$

ami az első bizonyítandó formula. Ebből speciális esetként kapjuk, hogy

$$ {}_j f^*_{jk} = \frac{m_{jj}}{m_{jk} + m_{kj}}. $$

Másrészt, az $(1B)$ formula szerint

$$ {}_k p^*_{ij} = {}_{j,k} p^*_{ij} + {}_{j,k} p^*_{ij} \cdot {}_k p^*_{jj} = {}_k f^*_{ij}(1 + {}_k p^*_{jj}). $$

A $(2B)$ formulából pedig

$$ 1 = f^*_{jk} = {}_k p^*_{jk} = {}_{j,k} p^*_{jk} + {}_k p^*_{jj} \cdot {}_{j,k} p^*_{jk} = {}_j f^*_{jk}(1 + {}_k p^*_{jj}). $$

A fenti kettőt egymással elosztva,

$$ {}_kp^*_{ij} = \frac{{}_kf^*_{ij}}{{}_jf^*_{jk}} = \frac{(m_{ik}+m_{kj}-m_{ij})/(m_{jk}+m_{kj})}{m_{jj}/(m_{jk}+m_{kj})} = \frac{m_{ik}+m_{kj}-m_{ij}}{m_{jj}}. $$

■

Megjegyzés: $k = i$ helyettesítéssel ismét megkapjuk, (amit már eddig is tudtunk), hogy ${}_ip^*_{ij} = m_{ii}/m_{jj}$.

3.4. A CHT-ben szereplő szórásnégyzet kiszámítása

Visszatérve az eredeti feladathoz, a

$$ V_h = Y_h - \mathcal{I}(f)(\tau_{h+1}-\tau_h) = \sum_{n=\tau_h}^{\tau_{h+1}-1} (f(X_n) - \mathcal{I}(f)) $$

mennyiség négyzetének várható értékét keressük (emlékezzünk arra, hogy a τ_h megállási idők a rögzített i állapotba tett látogatások időpontjai). Ehhez először a $\sum_{n=\tau_h}^{\tau_{h+1}-1} g(X_n)$ mennyiség négyzetének várható értékét számoljuk ki, majd ezt a $g = f - \mathcal{I}(f)$ függvényre alkalmazzuk. Az egyszerűség kedvéért tegyük most is fel, hogy $X_0 = i$. Ekkor

$$ \begin{aligned} & E\left[\left(\sum_{n=0}^{\tau_1-1} g(X_n)\right)^2\right] = \\ & E\left[\left(\sum_{n=0}^{\infty} g(X_n)\chi\{n<\tau_1\}\right)^2\right] = E\left[\left(\sum_{n=1}^{\infty} g(X_n)\chi\{n\le\tau_1\}\right)^2\right] = \\ & E\left[\sum_{n=1}^{\infty} g^2(X_n)\chi\{n\le\tau_1\} + 2\sum_{n<m} g(X_n)g(X_m)\chi\{m\le\tau_1\}\right] = \\ & \sum_{j\in I} g^2(j)\sum_{n=1}^{\infty} {}_ip^{(n)}_{ij} + 2\sum_{j\in I, j\ne i}\sum_{l\in I} g(j)g(l)\sum_{n<m} {}_ip^{(n)}_{ij}\,{}_ip^{(m-n)}_{jl} = \\ & \sum_{j\in I} g^2(j)\frac{\pi_j}{\pi_i} + 2\sum_{j\in I, j\ne i}\sum_{l\in I} g(j)g(l)\frac{\pi_j}{\pi_i}\pi_l(m_{ji}+m_{il}-m_{jl}), \end{aligned} \tag{3} $$

a 3.11. Lemmát használva. Tovább számolva,

$$\pi_i E\left[\left(\sum_{n=0}^{\tau_1-1} g(X_n)\right)^2\right] =$$

$$\mathcal{I}(g^2) + 2\sum_{j\in I}\sum_{l\in I} g(j)g(l)\pi_j\pi_l(m_{ji}+m_{il}-m_{jl}) - 2g(i)\sum_{l\in I} g(l)\pi_l =$$

$$\mathcal{I}(g^2) - 2g(i)\mathcal{I}(g) + 2\mathcal{I}(g)\sum_{j\in I} g(j)\pi_j m_{ji} + 2\mathcal{I}(g)\sum_{l\in I} g(l)\pi_l m_{il} -$$

$$-2\sum_{j\in I}\sum_{l\in I} g(j)g(l)\pi_j\pi_l m_{jl}.$$

Ha most $g = f - \mathcal{I}(f)$, akkor

$$\pi_i E(V_h^2) = \mathcal{I}\left\{(f-\mathcal{I}(f))^2\right\} - 2\sum_{j\in I}\sum_{l\in I}\{f(j)-\mathcal{I}(f)\}\{f(l)-\mathcal{I}(f)\}\pi_j\pi_l m_{jl}.$$

A most kiszámolt képlet segítségével az i-be való visszatérési idő második momentumát is megkaphatjuk. Ehhez válasszuk a $g = 1$ függvényt, a (3) egyenletből és a 3.11. Lemmából:

$$E(\tau^2) = \frac{1}{\pi_i} + 2\sum_{j\in I, j\neq i}\sum_{l\in I}\frac{\pi_j}{\pi_i}\cdot {}_ip^*_{jl} =$$

$$m_{ii} + 2m_{ii}\sum_{j\in I, j\neq i}\pi_j\sum_{l\in I} {}_ip^*_{jl} = m_{ii} + 2m_{ii}\sum_{j\in I, j\neq i}\pi_j m_{ji} = m_{ii}\left[2\sum_{j\in I}\frac{m_{ji}}{m_{jj}} - 1\right]$$

3.12. Példa. Legyen egy Markov lánc állapottere $\mathbb{N}$, és a folyamat olyan, hogy 0-ból átugrunk valamelyik állapotba, geometriai eloszlású ideig ott maradunk, majd visszaugrunk 0-ba. Tehát:

$$p_{0j} = \alpha_j, \quad p_{j0} = \beta_j, \quad p_{jj} = 1-\beta_j,\ j \geq 1,$$

ahol $\alpha_j > 0$, $\sum_{j=1}^{\infty}\alpha_j = 1$, $0 < \beta_j < 1$. Ez irreducibilis, aperiodikus. A lánc akkor pozitív rekurrens, ha van π_j stacionárius eloszlása. Az egyenletek:

$$\pi_j = (1-\beta_j)\pi_j + \alpha_j\pi_0,\ j \geq 1, \quad \pi_0 = \sum_{j=1}^{\infty}\beta_j\pi_j,$$

amit a $\pi_j = \frac{\alpha_j}{\beta_j}\pi_0$ számok elégítenek ki. Ez akkor lehet eloszlás, ha $\sum_{j=1}^{\infty} \frac{\alpha_j}{\beta_j} < \infty$, és ekkor

$$\pi_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \frac{\alpha_k}{\beta_k}}, \ \pi_j = \frac{\alpha_j}{\beta_j}\pi_0.$$

Számoljuk ki az átlagos elérési időket! Egyrészt $m_{jj} = 1/\pi_j$, $(j \in \mathbb{N})$, másrészt nyilván $m_{j0} = 1/\beta_j$, $(j \geq 1)$ a geometriai eloszlás miatt. Felhasználva, hogy $m_{00} = {}_0f_{0j}^*(m_{0j} + {} + m_{j0})$, kapjuk, hogy

$$m_{0j} = \frac{m_{00}}{{}_0f_{0j}^*} - m_{j0} = \frac{1 + \sum_{k=1}^{\infty} \frac{\alpha_k}{\beta_k}}{\alpha_j} - \frac{1}{\beta_j} = \frac{1}{\alpha_j}\left(1 + \sum_{k \neq j} \frac{\alpha_k}{\beta_k}\right).$$

Ha most $j \neq k$, $j, k \neq 0$, akkor

$$m_{jk} = m_{j0} + m_{0k} = \frac{1}{\beta_j} + \frac{1}{\alpha_k}\left(1 + \sum_{l \neq k} \frac{\alpha_l}{\beta_l}\right).$$

Most már megkaphatjuk a 0-ba való visszatérési idő második momentumát:

$$E(\tau^2) = \sum_{j=1}^{\infty} \alpha_j \left(1 + \frac{1}{\beta_j} + \frac{2}{\beta_j^2}\right).$$

Mj.: ezt, és m_{0j}-t is ki lehet egyszerűbben is számolni.

4. Reguláris mérték

A (nemnegatív) reguláris mértékek a stacionárius eloszlások általánosításai. Eleget tesznek a stacionárius eloszlásra vonatkozó egyenletrendszernek, de nem követeljük meg, hogy végesek legyenek. Legyen C tetszőleges osztály, és jelölje P_C az átmenet-mátrix megszorítását a C-beli állapotokra. Legyen $u = (u_i)_{i \in C}$ sorvektor, $v = (v_i)_{i \in C}$ oszlopvektor. Azt mondjuk, hogy u reguláris (szuperreguláris, szubreguláris) mérték C-n, ha $u = uP_C$ ($u \geq uP_C$, $u \leq uP_C$). Hasonlóan, v reguláris (szuperreguláris, szubreguláris) függvény C-n, ha $v = P_C v$ ($v \geq P_C v$, $v \leq P_C v$).

4.1. Reguláris függvény

4.1. Lemma. *Legyen C tetszőleges osztály. Ha létezik $v = (v_i)_{i\in C} \geq 0$, és $j \in C$, melyre*

$$v_i \geq \sum_{k\in C, k\neq j} p_{ik} v_k + p_{ij},\ i \in C, \tag{4}$$

akkor $v_i \geq f^*_{ij}$ minden $i \in C$-re, és a $v_i = f^*_{ij}$ egyenlőséggel teljesíti (4)*-et.*

Vegyük észre, hogy ha a $v = (v_i)$ nemnegatív szuperreguláris függvény, és $v_j > 0$, akkor a $w_i = v_i/v_j$ számok (a j állapottal) kielégítik a (4) rendszert.

Bizonyítás. (4) szerint $v_i \geq p_{ij} = f^{(1)}_{ij}$. Indukcióval megmutatjuk, hogy $v_i \geq \sum_{m=1}^n f^{(m)}_{ij}$ minden n-re. Legyen most $m \geq 1$, és $i \in C$. Az

$$f^{(m+1)}_{ij} = \sum_{k\in C, k\neq j} p_{ik} f^{(m)}_{kj}$$

egyenlőséget felhasználva, ha az állítást n-re már tudjuk, akkor

$$v_i \geq f^{(1)}_{ij} + \sum_{k\in C,k\neq j} p_{ik} v_k \geq f^{(1)}_{ij} + \sum_{k\in C,k\neq j} p_{ik} \sum_{m=1}^{n} f^{(m)}_{kj} =$$
$$f^{(1)}_{ij} + \sum_{m=1}^{n} \sum_{k\in C,k\neq j} p_{ik} f^{(m)}_{kj} = f^{(1)}_{ij} + \sum_{m=1}^{n} f^{(m+1)}_{ij} = \sum_{m=1}^{n+1} f^{(m)}_{ij}.$$

Ezért $v_i \geq f^*_{ij}$.

Másrészt,

$$p_{ij} + \sum_{k\in C,k\neq j} p_{ik} f^*_{kj} = p_{ij} + \sum_{m=1}^{\infty} \sum_{k\in C,k\neq j} p_{ik} f^{(m)}_{kj} = f^{(1)}_{ij} + \sum_{m=1}^{\infty} f^{(m+1)}_{ij} = f^*_{ij},\ i \in C.$$

■

4.2. Tétel. *Ha C rekurrens osztály, akkor minden nemnegatív szuperreguláris függvény konstans. Ha C tranziens, és $|C| \geq 2$, akkor van rajta* 0 *és* 1 *közötti nem-konstans szuperreguláris függvény.*

Bizonyítás. Legyen $v = (v_i)_{i\in C}$ nemnegatív szuperreguláris függvény. Ekkor

$$v_i \geq \sum_{k\in C} p_{ik} v_k \geq \sum_{k\in C} p_{ik} \sum_{j\in C} p_{kj} v_j = \sum_{j\in C} p^{(2)}_{ij} v_j \geq \ldots \geq \sum_{k\in C} p^{(n)}_{ik} v_k,\ \forall n.$$

Ebből látszik, hogy ha valamelyik v_k pozitív, akkor az összes v_i is az. Legyen $v > 0$ szuperreguláris. Ekkor a megjegyzés szerint tetszőleges $j \in C$-re $w_i = v_i/v_j$ a lemma szerinti függvény, tehát $v_i/v_j \geq f^*_{ij}$. Ha C rekurrens, akkor $f^*_{ij} = 1$, ezért v_i konstans. Ha most C tranziens, akkor legyen $j \in C$ tetszőleges, $v_j = 1$, és $v_i = f^*_{ij}$, ha $i \neq j$. A lemma szerint ekkor

$$v_i = f^*_{ij} = \sum_{k\in C, k\neq j} p_{ik} f^*_{kj} + p_{ij} = \sum_{k\in C} p_{ik} v_k, \text{ ha } i \neq j,$$

és

$$v_j = 1 > f^*_{jj} = \sum_{k\in C, k\neq j} p_{jk} f^*_{kj} + p_{jj} = \sum_{k\in C} p_{jk} v_k, \text{ ha } i = j.$$

Ez a függvény nem konstans, mivel ha minden $i \neq j$-re $f^*_{ij} = 1$ lenne, akkor $f^*_{jj} = 1$ lenne. ■

4.2. Reguláris mérték

Térjünk rá a reguláris mértékek vizsgálatára! Vezessük be az $e_{hi} = {}_h p^*_{hi}$ jelölést, ez tehát azt fejezi ki, hogy két h-ban tett látogatás között a lánc átlagosan hányszor jár i-ben (a végpontot beleszámítva, a kezdőpontot nem). A reguláris mértékekről szóló tétel előtt bizonyítsunk be egy hasznos lemmát.

4.3. Lemma. *Legyen C rekurrens osztály, $i, j, k, l, h \in C$ tetszőleges állapotok. Ekkor*

$$\lim_{N\to\infty} \frac{\sum_{n=0}^N p^{(n)}_{ij}}{\sum_{n=0}^N p^{(n)}_{kl}} = \frac{e_{hj}}{e_{hl}}.$$

Megjegyzések:

– Feladat: mutassuk meg, hogy tetszőleges osztályban $0 < e_{hi} < \infty$. Általánosan, ha $j \to H$ (j-ből el lehet jutni H-ba), akkor ${}_H p^*_{ij} < \infty$.

– Pozitív rekurrens osztályra az állítás már korábban szerepelt. Ekkor ugyanis $e_{hi} = m_h/m_i$ és $(1/N)\sum_0^N p^{(n)}_{ij} \to 1/m_j$.

– Mivel h tetszőleges lehet, és $e_{ll} = 1$, kapjuk, hogy

$$\frac{e_{hj}}{e_{hl}} = \frac{e_{lj}}{e_{ll}} = e_{lj},$$

azaz $e_{hj} = e_{hl} e_{lj}$.

Bizonyítás. Azt már korábban láttuk, hogy tetszőleges i, j állapotpárra

$$\lim_{N\to\infty} \frac{\sum_{n=1}^{N} p_{ij}^{(n)}}{\sum_{n=0}^{N} p_{jj}^{(n)}} = f_{ij}^*. \tag{5}$$

Ezért, mivel az osztály rekurrens,

$$\lim_{N\to\infty} \frac{\sum_{n=0}^{N} p_{ij}^{(n)}}{\sum_{n=0}^{N} p_{kl}^{(n)}} = \lim_{N\to\infty} \frac{\sum_{n=0}^{N} p_{jj}^{(n)}}{\sum_{n=0}^{N} p_{ll}^{(n)}},$$

erről kell belátni, hogy e_{lj}-vel egyenlő (ezt a $\sum_{n=0}^{N} p_{hh}^{(n)}$ taggal bővítve kapjuk az általános eredményt).

A kulcslépés egy olyan hányados határértékének meghatározása, ahol a számlálóban és a nevezőben az *első* helyen álló állapotok egyeznek meg. A (2A) alapformulából

$$p_{lj}^{(n)} = {}_l p_{lj}^{(n)} + \sum_{s=1}^{n-1} p_{ll}^{(s)} {}_l p_{lj}^{(n-s)},$$

amiből összegzéssel

$$\sum_{n=1}^{N} p_{lj}^{(n)} = \sum_{n=1}^{N} {}_l p_{lj}^{(n)} + \sum_{s=1}^{N-1} p_{ll}^{(s)} \sum_{t=1}^{N-s} {}_l p_{lj}^{(t)}.$$

Alkalmazzuk a Nörlund lemmát $a_0 = 0$, $a_n = p_{ll}^{(n)}$, $b_n = \sum_{t=1}^{n} {}_l p_{lj}^{(t)}$ szereposztással!

$$\lim_{N\to\infty} \frac{\sum_{n=1}^{N} p_{lj}^{(n)}}{\sum_{n=1}^{N} p_{ll}^{(n)}} = \frac{{}_l p_{lj}^*}{p_{ll}^*} + {}_l p_{lj}^*.$$

Mivel az osztály rekurrens,

$$\lim_{N\to\infty} \frac{\sum_{n=1}^{N} p_{lj}^{(n)}}{\sum_{n=1}^{N} p_{ll}^{(n)}} = \lim_{N\to\infty} \frac{\sum_{n=1}^{N} p_{lj}^{(n)}}{\sum_{n=0}^{N} p_{ll}^{(n)}} = {}_l p_{lj}^* = e_{lj}.$$

Ebből (5) által már következik az állítás. ■

4.4. Tétel. *Legyen C rekurrens osztály, ekkor a nemnegatív reguláris mértékek általános alakja: $u_i = ce_{hi}$, ahol $h \in C$ tetszőleges. Másrészt, az $u_i = e_{hi}$ mérték tetszőleges osztályon szuperreguláris, és akkor és csak akkor reguláris, ha az osztály rekurrens.*

Megjegyzések:

- Rekurrens osztályban $\sum_{i\in C} e_{hi} = m_{hh}$, attól függően véges vagy végtelen, hogy az osztály pozitív vagy nulla rekurrens. Pozitív rekurrens osztályban tehát nem kapunk új megoldást (itt $e_{hi} = \pi_i/\pi_h$).
- A multiplikatív tulajdonságból következik, hogy a különböző h választásokkal kapott mértékek csak konstans szorzóban térnek el egymástól.

Bizonyítás. Először belátjuk, hogy $u_i = ce_{hi}$ reguláris mérték, azaz megoldása az egyenletrendszernek. Felhasználva, hogy tetszőleges C osztályban

$$ {}_hp_{hi}^{(n+1)} = \sum_{k\in C, k\neq h} {}_hp_{hk}^{(n)} p_{ki}, $$

kapjuk, hogy

$$ \sum_{k\in C} e_{hk}p_{ki} = \sum_{k\in C}\sum_{n=1}^{\infty} {}_hp_{hk}^{(n)}p_{ki} = \sum_{n=1}^{\infty}\sum_{k\in C} {}_hp_{hk}^{(n)}p_{ki} = $$
$$ \sum_{n=1}^{\infty}\left({}_hp_{hi}^{(n+1)} + {}_hp_{hh}^{(n)}p_{hi}\right) = e_{hi} - {}_hp_{hi}^{(1)} + {}_hp_{hh}^{*}p_{hi} = e_{hi}, $$

mivel ${}_hp_{hi}^{(1)} = p_{hi}$, és ${}_hp_{hh}^{*} = f_{hh}^{*} = 1$. Rögtön látszik, hogy ha az osztály tranziens, akkor $f_{hh}^{*} < 1$ miatt $\sum_{k\in C} e_{hk}p_{ki} < e_{hi}$, azaz a mérték szuperreguláris, de nem reguláris.

Másrészt, legyen most $u_i \geq 0$ megoldása az egyenletrendszernek. Az egyenletrendszer iterációjával $u_i = \sum_{k\in C} u_k p_{ki}^{(n)}$ minden n-re, azaz a mérték vagy konstans nulla, vagy szigorúan pozitív. Ez utóbbi esetben legyen

$$ q_{ij}^{(n)} = \frac{u_j}{u_i} p_{ji}^{(n)}, \quad i, j \in C. $$

Megmutatjuk, hogy ezek egy C-n adott sztochasztikus mátrix n-dik hatványának elemei. Ehhez azt kell látni, hogy nemnegatívak (triviális), a sorok összege 1 (a regularitás miatt), és teljesül a Chapman-Kolmogorov egyenlőség:

$$ \sum_{k\in C} q_{ik}q_{kj}^{(n)} = \sum_{k\in C} \frac{u_k}{u_i}p_{ki}\frac{u_j}{u_k}p_{jk}^{(n)} = \frac{u_j}{u_i}p_{ji}^{(n+1)} = q_{ij}^{(n+1)}. $$

A q_{ij} átmenetvalószínűségekkel definiált Markov-lánc irreducibilis C-n, és rekurrens (mert $q_{ii}^{(n)} = p_{ii}^{(n)}$). Ezért a lemma alapján

$$1 = \lim_{N\to\infty} \frac{\sum_{n=0}^{N} q_{ij}^{(n)}}{\sum_{n=0}^{N} q_{jj}^{(n)}} = \frac{u_j}{u_i} \lim_{N\to\infty} \frac{\sum_{n=0}^{N} p_{ji}^{(n)}}{\sum_{n=0}^{N} p_{jj}^{(n)}} = \frac{u_j}{u_i} e_{ji},$$

azaz $u_i = u_h e_{hi}$. ■

Rekurrens osztályban tehát konstans szorzó erejéig egyértelműen létezik nemnegatív reguláris mérték, pozitív rekurrens esetben összege véges, nulla rekurrens esetben összege végtelen. Tranziens (lényeges) osztályban véges összegű reguláris mérték nincs, végtelen összegű vagy van, vagy nincs.

4.5. Példa. A P mátrix duplán sztochasztikus, ha sztochasztikus, és az oszlopok összege is 1. (Tegyük fel, hogy a hozzá tartozó Markov lánc irreducibilis.) Ekkor az $u_i = c$ mérték reguláris.

Az egy dimenziós egyszerű bolyongás átmenetmátrixa ilyen. Ha $p = 1/2$, akkor a lánc rekurrens, tehát ez az egyetlen reguláris mérték. Ha viszont $p \neq 1/2$, akkor az $u_i = (p/q)^i$ ettől különböző reguláris mérték.

4.6. Példa. Kártyakeverés: tegyük fel, hogy egy pakli kártyában N lap van. Egy keverést a lapok egy átrendezése, azaz egy S_N-beli permutáció határoz meg. Ha egy paklit elkezdünk keverni, az egyes keverések utáni állapotok (kártya-sorrendek) egy Markov láncot alkotnak. Tegyük fel, hogy a keverő a π-szerinti átrendezést (keverést) r_π valószínűséggel alkalmazza. Ekkor a Markov lánc átmenetvalószínűségei: $p_{\sigma\rho} = r_{\rho\circ\sigma^{-1}}$. Ha tehát a lánc irreducibilis, akkor stacionárius eloszlása egyenletes, aperiodikus esetben sok keverés után a pakli jól meg lesz keverve.

4.3. Megfordított láncok

4.7. Definíció. *Legyen P tetszőleges átmenetmátrix, és $u_k > 0$ ($k \in I$) reguláris mérték P-re a teljes állapottéren. Ekkor a $q_{ij} = p_{ji}\frac{u_j}{u_i}$ elemű Q mátrix is átmenetmátrix ugyanezen az állapottéren. Q a P u szerinti megfordítása.*

4.8. Állítás. *A magasabbrendű átmenetvalószínűségekre igaz, hogy $q_{ij}^{(n)} = p_{ji}^{(n)}\frac{u_j}{u_i}$. A két mátrix által meghatározott Markov láncok osztályai megegyeznek, továbbá a két láncban egyszerre nulla rekurrensek, pozitív rekurrensek, vagy tranziensek. Az u_k mérték a Q által meghatározott Markov láncra is reguláris.*

Az állítás bizonyítása az előző szakaszban szerepelt (kivéve az utolsó állítást, ami triviális). Vegyük még észre, hogy ha most Q-t megfordítjuk u szerint, akkor visszakapjuk P-t.

Tegyük fel, hogy $u_k > 0$ stacionárius eloszlás a P átmenetmátrixra. (Ez akkor lehetséges, ha a Markov láncnak minden osztálya pozitív rekurrens.) Tegyük még fel, hogy $P(X_0 = k) = u_k$. Ekkor

$$P(X_m = j | X_{m+n} = i) = \frac{P(X_m = j, X_{m+n} = i)}{P(X_{m+n} = i)} = \frac{u_j p_{ji}^{(n)}}{u_i} = q_{ij}^{(n)},$$

azaz minden N-re az $Y_n = X_{N-n}$ $(n = 0, \ldots, N)$ sorozat (véges) stacionárius Markov láncot alkot, u_k kezdeti eloszlással, q_{ij} átmenetvalószínűségekkel. (Mivel a Markov tulajdonság azzal ekvivalens, hogy a jelenre nézve a múlt és a jövő feltételesen független, az Y_n sorozat is Markov tulajdonságú.)

Azt mondjuk, hogy az X_n Markov lánc megfordítható (u-ra nézve), ha megfordítása önmaga.

Tegyük most fel ismét, hogy $u_i > 0$ stacionárius eloszlás, és a Markov lánc megfordítható. Sorsoljuk ki az X_0 értékét az u_i eloszlás szerint. Ezután egymástól függetlenül indítsunk el egy-egy Markov láncot előre és hátra P szerint. Ekkor egy kétirányban végtelen stacionárius Markov láncot kapunk. Ehhez azt kell látni, hogy a Markov tulajdonság a teljes folyamatra érvényben marad (feladat: lássuk be!), valamint, hogy az átmenetvalószínűségeket a $p_{ij}^{(n)}$ értékek adják. Ez utóbbi külön-külön a két félegyenesen igaz, és

$$P(X_n = j | X_{-m} = i) = \frac{P(X_{-m} = i, X_n = j)}{P(X_{-m} = i)} = \frac{\sum_k u_k p_{ki}^{(m)} p_{kj}^{(n)}}{u_i} = \sum_k p_{ik}^{(m)} p_{kj}^{(n)} = p_{ij}^{(n+m)}.$$

4.9. Példa. Egyszerű bolyongás egy dimenzióban (jobbra p, balra q valószínűséggel lép). Tegyük fel, hogy $p \neq 1/2$. Láttuk, hogy két reguláris mérték van. Ha az $u_k = 1$ mérték szerint fordítjuk meg a láncot, akkor $q_{i,i+1} = p_{i+1,i} = q$, és $q_{i,i-1} = p_{i-1,i} = p$, azaz ezzel a mértékkel a lánc nem megfordítható. Az $u_k = (p/q)^k$ mértékkel megfordítva a láncot,

$$q_{i,i+1} = \frac{u_{i+1}}{u_i} p_{i+1,i} = \frac{p}{q} q = p,$$

és hasonlóan $q_{i+1,i} = q$, azaz a lánc megfordítható.

4.10. Példa. Kártyakeverés. Az $u_k = 1$ mértékkel megfordítva a láncot,

$$q_{\sigma\rho} = p_{\rho\sigma} = r_{\sigma\circ\rho^{-1}} = r_{(\rho\circ\sigma^{-1})^{-1}},$$

tehát a lánc akkor megfordítható, ha $r_\pi = r_{\pi^{-1}}$ minden π-re.

4.11. Példa. Fejek száma egy dobássorozat végén (p a fej valószínűsége). Láttuk, hogy a Markov lánc irreducibilis, pozitív rekurrens. Hogy néz ki a megfordítása? $u_k = p^k q$, tehát

$$q_{k+1,k} = \frac{u_k}{u_{k+1}} p_{k,k+1} = 1, \quad q_{0,k} = \frac{u_k}{u_0} p_{k,0} = p^k q.$$

A lánc nyilván nem megfordítható.

4.12. Állítás. *Legyen X_n irreducibilis Markov lánc, melynek létezik egy $u > 0$ reguláris mértéke. Ekkor X_n pozitív reguláris függvényei kölcsönösen egyértelmű megfeleltetésben állnak az u szerint megfordított Y_n Markov lánc pozitív reguláris függvényeivel.*

Bizonyítás. Ha $v > 0$ reguláris függvény X_n-re, akkor $s_k = u_k v_k$ pozitív reguláris mérték Y_n-re:

$$\sum_k s_k q_{kj} = \sum_k u_k v_k \frac{u_j}{u_k} p_{jk} = u_j \sum_k p_{jk} v_k = u_j v_j = s_j.$$

Ugyanígy, ha $s > 0$ reguláris mérték Y_n-re, akkor $v_k = s_k / u_k$ pozitív reguláris függvény X_n-re. ∎

Ha az irreducibilis Markov lánc rekurrens, akkor már korábban láttuk, hogy pozitív reguláris függvényei csak a konstans függvények, pozitív reguláris mértékei pedig az $u_i = c e_{hi}$ alakú mértékek. A fenti állítás tehát tranziens esetben érdekes.

4.13. Példa. egyszerű bolyongás egy dimenzióban (jobbra p, balra q valószínűséggel lép). Keressük meg a pozitív reguláris függvényeket! Láttuk, hogy ha az $u_k = (p/q)^k$ mértékkel fordítjuk meg a láncot, akkor önmagát kapjuk. Tehát a lánc pozitív reguláris függvényei

$$v_k = \frac{s_k}{u_k} = \left(\frac{q}{p}\right)^k s_k$$

alakúak, ahol s_k pozitív reguláris mérték. Ebből a függvények (konstans szorzó erejéig): $v_k = (q/p)^k$, illetve $v_k = 1$.

5. Elnyelődési valószínűségek

Legyen X_n Markov lánc, jelölje $C_1, C_2, \dots$ a rekurrens osztályokat, T pedig a tranziens állapotok unióját. Tudjuk, hogy rekurrens osztályból nem lehet kilépni, tranziensből pedig vagy lehet, vagy nem, aszerint, hogy az osztály lényeges-e. Egy tranziens állapotból indulva tehát vagy örökké tranziens állapotban marad a lánc, vagy elnyelődik valamelyik rekurrens osztályba.

Jelölje $\sigma_i^{(n)}$ annak valószínűségét, hogy az i (tranziens) állapotból indulva, n lépés múlva a lánc tranziens állapotban van. Ekkor $\sigma_i^{(n)}$ monoton csökkenő, és $\lim_{n\to\infty} \sigma_i^{(n)} = \sigma_i$ annak esélye, hogy i-ből indulva a lánc örökre tranziens állapotban marad. Minden $n \geq 0$-ra igaz, hogy

$$\sigma_i^{(n+1)} = \sum_{j\in T} p_{ij}^{(n+1)} = \sum_{j\in T}\sum_{k\in T} p_{ik} p_{kj}^{(n)} = \sum_{k\in T} p_{ik}\sigma_k^{(n)}.$$

Határértéket véve kapjuk (van szummábilis majoráns), hogy

$$\sigma_i = \sum_{k\in T} p_{ik}\sigma_k \quad i \in T. \tag{6}$$

Azaz σ reguláris függvény a tranziens állapotok halmazán.

5.1. Állítás. *σ a legnagyobb nulla és egy közötti megoldása a* (6) *egyenletrendszernek.*

Bizonyítás. Ha $0 \leq x_i \leq 1$ megoldás, akkor indukcióval megmutatható, hogy $x_i \leq \sigma_i^{(n)}$. $n = 0$-ra a feltevés szerint igaz. Ha n-re már tudjuk, akkor

$$\sigma_i^{(n+1)} = \sum_{k\in T} p_{ik}\sigma_k^{(n)} \geq \sum_{k\in T} p_{ik} x_k = x_i.$$

Ebből pedig határátmenettel $x_i \leq \sigma_i$. ■

5.2. Állítás. *Az X_n Markov láncban legyen C valahány rekurrens osztály uniója, és jelölje θ_i annak valószínűségét, hogy az i (tranziens) állapotból indulva, előbb-utóbb C-be nyelődik el a lánc. Ekkor θ_i a*

$$y_i = \sum_{j\in T} p_{ij} y_j + \sum_{j\in C} p_{ij}, \quad i \in T \tag{7}$$

egyenletrendszer minimális nemnegatív megoldása.

Bizonyítás. Jelölje $\theta_i^{(n)}$ annak valószínűségét, hogy az n. lépésig beér a lánc C-be (i-ből indulva). Ez a sorozat monoton növő, határértéke θ_i. Ugyanakkor

$$\theta_i^{(n+1)} = \sum_{j\in C} p_{ij} + \sum_{j\in T} p_{ij}\theta_j^{(n)},$$

amiből határátmenettel kapjuk, hogy θ_i megoldása (7)-nek. Ha y_i is megoldás, akkor

$$y_i \geq \sum_{j\in C} p_{ij} = \theta_i^{(1)},$$

és az indukciós lépés:

$$\theta_i^{(n+1)} = \sum_{j\in C} p_{ij} + \sum_{j\in T} p_{ij}\theta_j^{(n)} \leq \sum_{j\in C} p_{ij} + \sum_{j\in T} p_{ij}y_j = y_i.$$

■

5.3. Példa. Legyen egy Markov lánc állapottere $I = \{0,\dots,a\}$. Tegyük fel, hogy a lánc martingál, azaz $E(X_{n+1}|X_n) = X_n$, ami azt jelenti, hogy $\sum_k kp_{ik} = i$ minden i-re. Mivel

$$0 = \sum_{k=0}^{a} kp_{0k} = \sum_{k=1}^{n} kp_{0k},$$

$p_{0k} = 0$, ha $k \geq 1$, azaz $p_{00} = 1$, a nulla elnyelő állapot. Hasonlóan az a állapot is lenyelő. Tegyük fel, hogy az $\{1,\dots,a-1\}$ állapotok egy osztályt alkotnak, ez ekkor szükségképpen tranziens (lényegtelen) osztály. Vizsgáljuk meg, mekkorák az elnyelődési valószínűségek! A martingáltulajdonság miatt $\sum_k kp_{ik}^{(n)} = i$, de $p_{ik}^{(n)} \to 0$, ha $k \in$ $\in \{1,\dots,a-1\}$. Ezért $\lim p_{ia}^{(n)} = i/a$, azaz ekkora valószínűséggel nyelődik a lánc a-ba (és a komplementer valószínűséggel 0-ba).

Végül lássunk még egy tételt, melynek segítségével eldönthető, hogy egy osztály rekurrens-e!

5.4. Állítás. *Legyen X_n irreducibilis Markov lánc. Az i állapot akkor és csak akkor rekurrens, ha az*

$$x_j = \sum_{k\neq i} p_{jk}x_k \quad j \neq i \tag{8}$$

egyenletrendszer egyetlen nulla és egy közötti megoldása $x_j = 0$.

Bizonyítás. Készítsünk egy új Markov láncot, melyben $p_{ii} = 1$, és a p_{jk} $(j \neq i)$ valószínűségek változatlanok. Ebben a láncban i rekurrens, az összes többi állapot tranziens. A korábbi állítás szerint (8) legnagyobb nulla és egy közötti megoldása σ_j, ami annak a valószínűsége, hogy j-ből indulva soha nem ér a lánc i-be, azaz $\sigma_j = 1 - f^*_{ji}$. Ha ez minden j-re 0, akkor $f^*_{ji} = 1$ minden $j \neq i$-re, amiből $f^*_{ii} = 1$. Fordítva, ha a lánc rekurrens, akkor $f^*_{ji} = 1$ minden $j \neq i$-re, azaz $\sigma_j = 0$. ■

6. Véges állapotterű Markov láncok

Ebben a szakaszban véges állapotterű Markov láncokkal foglalkozunk. Ilyenkor az átmenetmátrix egy véges, négyzetes sztochasztikus mátrix. Ezeket algebrai eszközökkel vizsgálva, újra levezethetjük a korábban kapott eredményeket, illetve megkaphatunk speciálisan erre az esetre érvényes állításokat.

6.1. Állítás. *Legyen X_n véges állapotterű Markov lánc. Ekkor (i) Minden rekurrens osztály pozitív, (ii) minden tranziens osztály lényegtelen, és a lánc 1 valószínűséggel előbb-utóbb elhagyja.*

Bizonyítás. (i) Legyen C rekurrens osztály. Ekkor minden $i \in C$-re és $n \geq 1$-re

$$1 = \sum_{j \in C} p_{ij}^{(n)}.$$

Ha viszont az osztály nulla lenne, akkor $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ lenne minden $j \in C$-re, ami C végessége miatt ellentmond a fentieknek.

(ii) Legyen C tranziens osztály. Minden $i \in C$-re és $n \geq 1$-re

$$1 = \sum_{j \in C} p_{ij}^{(n)} + \sum_{j \notin C} p_{ij}^{(n)},$$

amiből

$$1 = \sum_{j \in C} \lim_{n\to\infty} p_{ij}^{(n)} + \lim_{n\to\infty} \sum_{j \notin C} p_{ij}^{(n)} = \lim_{n\to\infty} P(X_n \notin C | X_0 = i).$$

■

A fentiekből adódik, hogy ha a Markov láncirreducibilis, akkor pozitív rekurrens, valamint, hogy minden Markov láncnak van legalább egy pozitív osztálya.

A továbbiakban A nemnegatív elemű, négyzetes mátrix. A mátrix legnagyobb sajátértékét, és a hozzá tartozó sajátvektorokat fogjuk elemezni. Összefoglalóan nevezhetjük az itt szereplő eredményeket Perron-Frobenius tételkörnek. Mátrixokra, illetve vektorokra a következő jelöléseket alkalmazzuk:

– $\geq$: minden koordináta nagyobb vagy egyenlő

– $>$: minden koordináta nagyobb vagy egyenlő, és legalább egy koordináta nagyobb

– $\gg$: minden koordináta nagyobb

6.2. Definíció. *Legyen* $A \geq 0$, *ekkor*

$$\lambda_0 = \lambda_0[A] = \sup\{\lambda : \ \exists x > 0 : \ Ax \geq \lambda x\}. \tag{9}$$

A továbbiakban $\mathbf{1}$ jelöli azt a vektort, melynek minden koordinátája 1.

6.3. Állítás. $\min_i(A\mathbf{1})_i \leq \lambda_0 \leq \max_i(A\mathbf{1})_i$.

Bizonyítás. Legyen egyrészt $x > 0$, λ (9)-beli pár, és $x_k = \max_i x_i > 0$. Ekkor

$$\lambda x_k \leq (Ax)_k = \sum_j a_{kj} x_j \leq x_k (A\mathbf{1})_k \leq x_k \max_i (A\mathbf{1})_i,$$

azaz $\lambda \leq \max_i(A\mathbf{1})_i$. Másrészt, a $\lambda = \min_i(A\mathbf{1})_i$, $x = \mathbf{1}$ pár (9)-beli, hiszen minden j-re

$$(Ax)_j = (A\mathbf{1})_j \geq \min_i (A\mathbf{1})_i = \lambda = \lambda x_j.$$

∎

6.4. Lemma. $\lambda_0 = 0$ *akkor és csak akkor, ha létezik* m, *melyre* $A^m = 0$. *(A nilpotens)*

Bizonyítás. Egyrészt, ha az x, λ pár (9)-beli, akkor $A^m x \geq \lambda^m x$, amiből az „akkor" irány következik. Másrészt, tegyük fel, hogy $\lambda_0 = 0$. Jelölje C_n az $A^n\mathbf{1}$ pozitív koordinátáinak halmazát. C_n nyilván csökkenő halmazsorozat, megmutatjuk, hogy szigorúan csökkenő (legalábbis amíg nem üres). Ha ugyanis $C_{n+1} = C_n \neq \emptyset$ lenne, akkor $A^n\mathbf{1}$ és $A^{n+1}\mathbf{1}$ ugyanazon koordinátái pozitívak, tehát létezik $\lambda > 0$, melyre

$$A(A^n\mathbf{1}) \geq \lambda(A^n\mathbf{1}),$$

és itt $A^n\mathbf{1} > 0$, ami ellentmondás. ∎

6.1. Irreducibilis, aperiodikus mátrixok

6.5. Tétel. *Legyen $A \gg 0$. Ekkor*

1. *λ_0 egyszeres sajátérték, melyhez tartozik $x^0 \gg 0$ sajátvektor.*

2. *Minden más sajátérték abszolút értéke kisebb λ_0-nál.*

Bizonyítás. 1. Minden nemnegatív elemű mátrix esetén létezik $x^0 > 0$, melyre $Ax^0 \geq \lambda_0 x^0$, ugyanis válasszunk olyan λ_k, x^k (9)-beli párokat, melyekre $\lambda_k \to \lambda_0$, $\|x_k\| = 1$, ekkor kiválasztható olyan k_n részsorozat, melyre $x_{k_n} \to x^0$, azaz $Ax^0 \geq \lambda_0 x^0$, és $x^0 > 0$. Tegyük most fel, hogy $Ax^0 > \lambda_0 x^0$, ekkor az $A \gg 0$ mártixszal rászorozva kapnánk, hogy $A(Ax^0) \gg \lambda_0(Ax^0)$, és mivel $Ax^0 > 0$, így ellentmondásra jutunk, mivel λ_0 növelhető lenne. Azt kaptuk, hogy $Ax^0 = \lambda_0 x^0$, és mivel $Ax^0 \gg 0$, így $x^0 \gg 0$. Rátérve a sajátérték multiplicitására, tegyük fel, hogy van egy másik sajátvektor, y, melyre tehát $y \neq cx^0$ semmilyen komplex c-re. Ekkor y valós és képzetes része is sajátvektor, és legalább az egyikükre igaz, hogy x^0-nak nem valós konstansszorosa. Tegyük fel tehát, hogy y valós. Ekkor létezik c valós szám, melyre $y - cx^0 > 0$, de $y - cx^0 \not\gg 0$. A-val rászorozva,

$$0 \ll A(y - cx^0) = \lambda_0 y - c\lambda_0 x^0 = \lambda_0(y - cx^0),$$

ami ellentmondás.

2. Az minden nemnegatív elemű mátrixra igaz, hogy a sajátértékek abszolút értéke legfeljebb λ_0, ugyanis $Az = \lambda z$-ből $A|z| \geq |\lambda| \cdot |z|$, azaz $|\lambda| \leq \lambda_0$. Tegyük most fel, hogy $|\lambda| = \lambda_0$. Legyen $\delta > 0$ olyan, hogy $A - \delta I \gg 0$ maradjon, erre könnyen látszik, hogy $\lambda_0[A - \delta I] = \lambda_0[A] - \delta$. Viszont $\lambda - \delta$ sajátértéke $A - \delta I$-nek ($(A - \delta I)z = (\lambda - \delta)z$), tehát az előzőek szerint $|\lambda - \delta| \leq \lambda_0 - \delta$. A háromszög-egyenlőtlenségből viszont $|\lambda| \leq |\lambda - \delta| + \delta$. Ezért $\lambda = \lambda_0$.

∎

Mielőtt a fenti tételt általánosítanánk, lássunk be egy lemmát!

6.6. Lemma. *Egy nemnegatív elemű A mátrixnak legfeljebb egy pozitív sajátértékéhez tartozhat pozitív elemű sajátvektor.*

Bizonyítás. Legyen $Ax = \lambda x$, $Ay = \mu y$, $x, y \gg 0$, és $\mu > \lambda > 0$. Ekkor van olyan $c > 0$, melyre $x - cy > 0$, de $x - cy \not\gg 0$, erre

$$0 \leq A(x - cy) = \lambda x - c\mu y = \lambda(x - cy) - c(\mu - \lambda)y \ngeq 0,$$

ami ellentmondás. ■

6.7. Tétel. *Az előző tétel akkor is igaz, ha van olyan m, melyre $A^m \gg 0$.*

Bizonyítás.

1. Az előzőek szerint létezik $x^0 > 0$, melyre $Ax^0 \geq \lambda_0 x^0$. Tegyük most fel, hogy $Ax^0 > \lambda_0 x^0$, ekkor az $A^m \gg 0$ mártixszal rászorozva kapnánk, hogy $A(A^m x^0) \gg \lambda_0(A^m x^0)$, és mivel $A^m x^0 > 0$, így ellentmondásra jutunk, mivel λ_0 növelhető lenne. Ezen kívül, mivel $A^m x^0 = \lambda_0^m x^0$ és $A^m x^0 \gg 0$, így $x^0 \gg 0$. A multiplicitásra rátérve, először belátjuk, hogy $\lambda_0[A^m] = \lambda_0[A]^m$. Legyen $\lambda_0[A^m] = \mu$. Frobenius első tétele szerint létezik $z^0 \gg 0$, melyre $A^m z^0 = \mu z^0$, ugyanakkor $A^m x^0 = \lambda_0^m x^0$, és $x^0 \gg 0$. A lemma szerint tehát $\mu = \lambda_0^m$. Ezután, ha $Az = \lambda_0 z$, akkor $A^m z = \lambda_0^m z$, de mivel Frobenius első tétele szerint A^m-nek a λ_0^m egyszeres sajátértéke, $z = cx^0$.

2. Legyen λ sajátérték, azaz $Az = \lambda z$. Tudjuk már, hogy $|\lambda| \leq \lambda_0$, tegyük most fel, hogy $|\lambda| = \lambda_0$. Mivel λ^m sajátértéke A^m-nek (z sajátvektorral), és $|\lambda^m| = \lambda_0^m$, Frobenius első tételéből $\lambda^m = \lambda_0^m$, és az egyszeresség miatt $z = cx^0$. Tehát $\lambda = \lambda_0$. ■

Ezt a tételt olyan Markov láncokra alkalmazhatjuk, melyek átmenetmátrixának valamelyik hatványa pozitív elemű. Ezek éppen az irreducibilis, aperiodikus Markov láncok. Markov láncok esetén $\lambda_0[P] = 1$, ezekre tehát kijött, hogy egyértelműen létezik stacionárius eloszlás, mely pozitív, valamint minden szuperreguláris függvény reguláris, és konstans.

6.2. Irreducibilis, periodikus mátrixok

6.8. Tétel. *Legyen $A \geq 0$ irreducibilis mátrix (azaz minden i, j-re létezik m, hogy $(A^m)_{ij} > 0$). Ekkor*

1. *λ_0 egyszeres sajátérték, melyhez tartozik $x^0 \gg 0$ sajátvektor.*

2. *Minden más sajátérték abszolút értéke kisebb vagy egyenlő, mint λ_0.*

Eszerint a tétel előtt elmondottak minden irreducibilis láncra is teljesülnek.

Bizonyítás. Csak az első állítást kell bizonyítani. Azt, hogy λ^0 sajátérték valamilyen pozitív sajátvektorral, az eddigiekhez hasonlóan bizonyítjuk, azaz keresünk egy olyan pozitív elemű mátrixot, mely A-val felcserélhető. Ebben az esetben ez a mátrix $(A+ +I)^m$, ahol m olyan, melyre $(A+I)^m \gg 0$ (ilyen van az irreducibilitás miatt). Ezért az $Ax^0 > \lambda_0 x^0$ feltételezésből rászorzással ellentmondásra jutunk. Ha már tudjuk, hogy λ_0 sajátérték x^0 sajátvektorral, akkor $(A+I)^m x^0 = (1+\lambda_0)^m x^0$ miatt $x^0 \gg 0$. Az egyszeresség belátásához megint azt kell látni, hogy $\mu = \lambda_0[(A+I)^m] = (1+\lambda_0[A])^m$. Ezt is ugyanúgy láthatjuk be, mint korábban: $(A+I)^m$-nek mind μ, mind $(1+\lambda_0)^m$ olyan pozitív sajátértéke, melyhez tartozik pozitív sajátvektor, ezért egyenlőek. Ha tehát $Az = \lambda_0 z$, akkor $(A+I)^m z = (1+\lambda_0)^m z$, de mivel Frobenius első tétele szerint $(1+\lambda_0)^m$ egyszeres sajátértéke $(A+I)^m$-nek, $z = cx^0$. ∎

6.9. Tétel. *Legyen $A \geq 0$ irreducibilis, d periódusú mátrix. Ekkor A azon sajátértékei, melyekre $|\lambda| = \lambda_0$:*

$$\lambda_k = \lambda_0 \exp\{2\pi i k/d\},\ k = 0, \dots, d-1.$$

Bizonyítás. Jelölje A legnagyobb sajátértékét λ_0, a hozzá tartozó pozitív sajátvektort x^0. A periódus definíciója irreducibilis mátrixra: legyen

$$d_i = \mathrm{lnko}\{n > 0 : (A^n)_{ii} > 0\}.$$

Ha $A \geq 0$ irreducibilis mátrix, akkor nincs benne csupa 0 sor, ezért minden sort 1-re normálva egy Markov lánc P átmenetmátrixát kapjuk, és A-ban és P-ben pontosan ugyanott vannak pozitív elemek. Emiatt a Markov láncokra kapott eredmények érvényben maradnak, azaz $d_i = d$, és a sorok/oszlopok halmaza rész-osztályokra bomlik. A sorokat és oszlopokat egyszerre alkalmasan átrendezve (ez a sajátértékeken nem változtat), egy $d \times d$-s blokkmátrixot kapunk, melyben a nem-nulla blokkok: $B_{12}, B_{23}, \dots, B_{d-1,d}, B_{d,1}$. Az A mátrix d-dik hatványa pedig blokkdiagonális, melynek blokkjai:

$$C_i = B_{i,i+1} B_{i+1,i+2} \cdots B_{i+d-1,i+d},\ i = 1, \dots, d,$$

ahol az indexelés modulo d értendő.

Korábbi eredményeink szerint a C_i blokkok alkalmas hatványa már $\gg 0$, ezekre tehát alkalmazható Frobenius második tétele. Jelölje $\lambda_0[C_i] = \mu_i$, és legyen $y^i \gg 0$ a hozzá tartozó sajátvektor. Ezekre

$$\mu_i(B_{i-1,i}y^i) = B_{i-1,i}(\mu_i y^i) = B_{i-1,i}C_i y^i = C_{i-1}(B_{i-1,i}y^i).$$

Itt $B_{i-1,i}y^i \gg 0$, amiből $\mu_{i-1} \geq \mu_i$ minden i-re, azaz $\mu_i = \mu$ minden i-re. Legyen $y \gg 0$ az y^i-kből összefűzött vektor, erre $A^d y = \mu y$.

Belátjuk még, hogy $\lambda_0[A^d] = \mu$. Ha ugyanis a λ, x párra $A^d x \geq \lambda x$, és $x > 0$, akkor az x vektort feldarabolva az x^i részekre, van olyan i, melyre $x^i > 0$, és $C_i x^i \geq \lambda x^i$. Tehát $\lambda \leq \mu_i = \mu$.

Mivel $A^d x^0 = \lambda_0^d x^0$, a lemmából $\mu = \lambda_0^d$. Tegyük most fel, hogy $Az = \lambda z$, és $|\lambda| = = \lambda_0$. Ekkor $A^d z = \lambda^d z$, és λ^d mindazon C_i blokkoknak is sajátértéke, melyekre z i-dik darabja nem azonosan nulla, legalább egy ilyen C_i biztosan van. Frobenius második tétele szerint, mivel $|\lambda^d| = \lambda_0^d = \lambda_0[C_i]$, így $\lambda^d = \lambda_0^d$. Azaz λ valóban csak az állításbeli alakú lehet.

Be kell még látni, hogy λ_i valóban sajátérték. Tekintsük az x^0 vektor feldarabolását: $x^0 = (x^1, \dots x^d)$. $Ax^0 = \lambda_0 x^0$ miatt $B_{j,j+1}x^{j+1} = \lambda_0 x^j$. Ha most tekintjük a $z = = (z^1, \dots, z^d)$ vektort, melyre $z^j = \exp\{2\pi ijk/d\}x^j$, erre

$$B_{j,j+1}z^{j+1} = \lambda_0 \exp\{2\pi i(j+1)k/d\}x^j = \lambda_0 \exp\{2\pi ik/d\}z^j = \lambda_k z^j.$$

■

6.3. Konvergencia és sebessége

A következő tétel A^n viselkedéséről szól, ha $n \to \infty$.

6.10. Tétel. *Legyen A irreducibilis, aperiodikus mátrix, x^0 illetve f^{0T} a λ_0 sajátértékhez tartozó pozitív elemű jobb- illetve baloldali sajátvektorok úgy normálva, hogy $f^{0T}x^0 = = 1$, és legyen $R = x^0 f^{0T}$. Ekkor $\lim_{n\to\infty} A^n/\lambda_0^n = R$.*

Bizonyítás. Könnyen felírható, hogy $R^2 = R$, $Rx^0 = x^0$, $f^{0T}R = f^{0T}$, $AR = RA = = \lambda_0 R$. Legyen $B = A - \lambda_0 R$, erre az előzőek szerint

$$B^n = (A-\lambda_0 R)^n = A^n + \sum_{k=0}^{n-1} \binom{n}{k} A^k(-\lambda_0 R)^{n-k} = A^n + \sum_{k=0}^{n-1} \binom{n}{k} (-1)^{n-k}\lambda_0^n R = A^n - \lambda_0^n R.$$

Ezért $A^n/\lambda_0^n - R = B^n/\lambda_0^n$, erről kell belátni, hogy nullához tart. Jelölje egy M mátrix sperktrálrádiuszát $r(M) = \max\{|\lambda| : \lambda \text{ az } M \text{ sajátértéke}\}$. Megmutatjuk, hogy

$r(B) < \lambda_0$. Tegyük fel ugyanis, hogy $Bz = \lambda z$ valamilyen $\lambda \neq 0$-ra. Mivel $RB = RA - \lambda_0 R^2 = 0$, ezért $0 = RBz = \lambda Rz$, azaz $Rz = 0$. Emiatt $\lambda z = Bz = Az$, azaz λ sajátértéke A-nak. Korábbi tételeink szerint tehát $|\lambda| \leq \lambda_0$. Továbbá $|\lambda| = \lambda_0$ csak úgy lehetne, hogy $\lambda = \lambda_0$ és $z = cx^0$. Ekkor viszont $Rz = cRx^0 = cx^0 = z \neq 0$ lenne, ami ellentmondás. Jelölje $\rho = r(B/\lambda_0) < 1$ (vegyük észre, hogy $\rho \leq \lambda_1/\lambda_0$, ahol λ_1 jelöli A sajátértékeinek abszolút értékei közül a második legnagyobbat). Használjuk fel azt a tételt a spektrálrádiuszról, hogy $r(M) = \lim_{n\to\infty} \|M^n\|^{1/n}$ (ahol $\|\cdot\|$ tetszőleges norma, mivel véges dimenzióban minden norma ekvivalens), ebből kapjuk, hogy minden $\epsilon > 0$-ra, elég nagy n-re $\|A^n/\lambda_0^n - R\| < (\rho + \epsilon)^n$. ■

Ha ezt a tételt irreducibilis, aperiodikus Markov láncra alkalmazzuk, akkor $x^0 = \mathbf{1}$, f^0 a stacionárius eloszlás, R pedig az a mátrix, amelynek minden sora f^{0T}. Ha az $\|M\| = \max |m_{ij}|$ normát választjuk, akkor azt kapjuk, hogy minden $\epsilon > 0$-ra, elég nagy n-re, minden i, j-re

$$|p_{ij}^{(n)} - \pi_j| \leq (\lambda_1 + \epsilon)^n.$$

6.4. Perron-Frobenius tétel

6.11. Tétel. (Perron-Frobenius tétel.) *Legyen $A > 0$ mátrix. Ekkor*

1. *λ_0 sajátérték, melyhez tartozik $x^0 > 0$ sajátvektor.*
2. *Minden más sajátérték abszolút értéke kisebb vagy egyenlő, mint λ_0.*
3. *Ha $|\lambda| = \lambda_0$ sajátérték, akkor $\eta = \lambda/\lambda_0$ egységgyök, és $\lambda_0\eta^m$ is sajátérték minden m-re.*
4. *Ha $x^0 \gg 0$, akkor $\frac{1}{n}\sum_{k=1}^{n}(A/\lambda_0)^k$ konvergens.*

Bizonyítás. Jelölje E a csupa egyesből álló mátrixot, és legyen $A_\delta = A + \delta E \gg 0$. Erre

$$\lambda_\delta = \lambda_0[A_\delta] = \sup\{\lambda : \exists x > 0 : A_\delta x \geq \lambda x\},$$

melyből látható, hogy $\lambda_0 = \lim_{\delta\to 0} \lambda_\delta$. Mivel létezik $x_\delta^0 \gg 0$, $\|x_\delta^0\| = 1$, melyre $A_\delta x_\delta^0 = \lambda_\delta x_\delta^0$, konvergens részsorozatot kiválasztva kapjuk az első állítást.

A harmadik állításhoz feltehetjük, hogy $\lambda_0 = 1$ (az A/λ_0 mátrixra áttérve). Az első állítás szerint van f^0, melyre $f^{0T}A = f^{0T}$.

Tegyük fel először, hogy $f^0 \gg 0$. Legyen $|\lambda| = 1$, $\lambda \neq 1$, és $Ax = \lambda x$. Ekkor egyrészt $A|x| \geq |x|$. Ha $A|x| > |x|$ lenne, akkor $f^{0T}|x| < f^{0T}A|x| = f^{0T}|x|$, ami ellentmondás, azaz $A|x| = |x|$. Ebből minden i-re

$$\sum_j a_{ij}|x_j| = |x_i| = |\frac{1}{\lambda}\sum_j a_{ij}x_j|,$$

azaz a háromszög-egyenlőtlenség egyenlőséggel teljesül. Van tehát minden i-re egy $|\mu_i| = 1$ szám, hogy

$$a_{ij}x_j/\lambda = \mu_i|a_{ij}x_j|, \ \forall j \quad (1).$$

Adott i-re j-ben összegezve kapjuk, hogy

$$x_i = \mu_i|x_i| \quad (2).$$

Tekintsük az $x \cdot \mu^r$ vektort $r \geq 0$, ahol a szorzást koordinátánként értjük. Megmutatjuk, hogy ez sajátérték λ^{r+1} sajátértékkel. Ha $r-1$-re már tudjuk, akkor a fenti (1), (2) egyenleteket használva

$$\sum_j a_{ij}x_j\mu_j^r = \sum_j \lambda\mu_i a_{ij}|x_j|\mu_j^r = \sum_j \lambda\mu_i a_{ij}x_j\mu_j^{r-1} = \lambda\mu_i\lambda^r x_i\mu_i^{r-1} = \lambda^{r+1}x_i\mu_i^r.$$

Térjünk rá arra az esetre, amikor $f^0 \not\gg 0$. Tegyük fel, hogy az első r koordináta pozitív, a többi nulla, azaz $f^{0T} = (f^{1T}, 0)$. Ekkor $f^{0T}A = f^{0T}$ miatt

$$A = \begin{pmatrix} A_1 & 0 \\ B & A_2 \end{pmatrix}$$

alakú, és A sajátértékeinek halmaza az A_1 és az A_2 sajátérték-halmazainak uniója. Tegyük fel először, hogy $\lambda_0[A_1] = 1$. Ekkor, mivel $f^{1T}A_1 = f^{1T}$, és $f^1 \gg 0$, az előzőek szerint készen vagyunk. Ha viszont $\lambda_0[A_2] = 1$, akkor A_2-re van $f^2 > 0$, hogy $f^{2T}A_2 = f^{2T}$, amely megint vagy $\gg 0$ (ebben az esetben készen vagyunk), vagy $\not\gg 0$, ebben az esetben pedig A_2 tovább bontható.

A negyedik állításhoz a rövidség kedvéért vezessük be a $T = A/\lambda_0$ jelölést. Belátjuk, hogy ha $x^0 \gg 0$, akkor T^n egyenletesen korlátos, ugyanis $T^n x^0 = x^0$ szerint minden i, j-re

$$\max x_k^0 \geq x_i^0 = \sum_l (T^n)_{il}x_l^0 \geq (T^n)_{ij}x_j^0 \geq (T^n)_{ij}\min x_k^0,$$

azaz
$$(T^n)_{ij} \le \frac{\max x_k^0}{\min x_k^0}.$$
Legyen $T_n = \frac{1}{n}\sum_{k=1}^n T^k$. Azt kell belátni, hogy a T_n operátornak létezik T_0 limesze, azaz minden x-re létezik a $T_0x = \lim_{n\to\infty} T_nx$ határérték.

Először belátjuk, hogy $\mathrm{Im}(I-T) = \{x : \lim T_nx = 0\}$. Ugyanis $T_n(I-T) = \frac{1}{n}(T - T^{n+1})$, amiből, ha $w = (I-T)v \in \mathrm{Im}(I-T)$, akkor $T_nw = \frac{1}{n}(T - T^{n+1})v \to 0$, T^n egyenletes korlátossága miatt. Fordítva, tegyük fel, hogy $T_nx \to 0$. Vegyük észre, hogy
$$I - T_n = \frac{1}{n}\sum_{k=1}^n (I - T^k) = \frac{1}{n}(I-T)\sum_{k=1}^n(\sum_{j=0}^{k-1} T^j) = \frac{1}{n}(I-T)\sum_{j=0}^{n-1}(n-j)T^j.$$
Emiatt $x - T_nx \in \mathrm{Im}(I-T)$, amiből $\lim(x - T_nx) = x \in \mathrm{Im}(I-T)$.

Minden x-re $\{T_nx : \ n \ge 1\}$ korlátos halmaz, azaz kiválasztható belőle konvergens részsorozat: $T_{n'}x \to x_0$. Be kell látni, hogy $T_nx \to x_0$. Mivel $x - T_{n'}x \in \mathrm{Im}(I-T)$, $x - x_0 \in \mathrm{Im}(I-T)$, ezért $T_n(x - x_0) \to 0$. Kell még, hogy T_nx_0 konvergens, azt látjuk be, hogy x_0 fixpontja T-nek, és így T_n-nek is. Ugyanis
$$Tx_0 = \lim TT_{n'}x = \lim(TT_{n'}x - T_{n'}x) + \lim T_{n'}x = \lim(T_{n'}(T-I)x) + x_0 = x_0.$$
Emiatt T_nx is konvergál, és $\lim T_nx = x_0$. ∎

Vegyük még észre, hogy T és T_0 felcserélhető, és $TT_0 = T_0T = T_0$. Ezért $T_nT_0 = T_0$ és $T_0^2 = T_0$. Könnyen belátható, hogy $\mathrm{Im}(T_0) = \mathrm{Ker}(I-T)$ és $\mathrm{Ker}(T_0) = \mathrm{Im}(I-T) = \mathrm{Im}(I-T_0)$.

A tétel első pontját egy Markov lánc átmenetmátrixának transzponáltjára alkalmazva kapjuk, hogy véges állapottér esetén mindig van stacionárius eloszlás (mivel van legalább egy pozitív osztály). A negyedik pont tetszőleges Markov lánc átmenetmátrixára alkalmazható, hiszen $x^0 = \mathbf{1} \gg 0$. Megkaptuk, hogy $\frac{1}{n}\sum_{k=1}^n p_{ij}^{(k)}$ minden i, j-re konvergens, azaz tetszőleges p kezdeti eloszlásra, ha $q_p^{(n)}$ jelöli X_n eloszlását, akkor $\frac{1}{n}\sum_{k=1}^n q_p^{(n)}$ konvergens. Az is kijött, hogy ilyen határértékként éppen a stacionárius eloszlások állnak elő.

6.5. A konvergenciasebesség becslése megállási időkkel

Legyen az X_n Markov lánc irreducibilis, aperiodikus, pozitív rekurrens, stacionárius eloszlását jelölje π. Tegyük fel, hogy a τ „ügyes” megállási idő olyan, hogy ha tudjuk, hogy a lánc a k-adik lépésben áll meg, akkor ebben az időpontban az eloszlása épp a stacionárius eloszlás, azaz minden k-ra és minden i állapotra

$$P(X_k = i | \tau = k) = \pi_i.$$

Legyen továbbá $\| \cdot \|_{TV}$ a teljes variáció.

6.12. Tétel. *Ha τ „ügyes” megállási idő, akkor minden p^0 kezdeti eloszlásra a $p_i^k = P(X_k = i)$ eloszlás és a stacionárius eloszlás variációs távolságára*

$$\|p^k - \pi\|_{TV} \leq P(\tau > k).$$

Bizonyítás. Legyen $A \subset I$, a variációs távolság definíciója alapján azt kell belátni, hogy $|p^k(A) - \pi(A)| \leq P(\tau > k)$. Mármost

$$p^k(A) = P(X_k \in A) = \sum_{j \leq k} P(X_k \in A, \tau = j) + P(X_k \in A, \tau > k).$$

Könnyen látszik, hogy $P(X_k \in A | \tau = j) = \pi(A)$, mivel a $\{\tau = j\}$ feltétel mellett X_j eloszlása épp a stacionárius eloszlás, ezért $k \geq j$ miatt X_k eloszlása is a stacionárius eloszlás. Ezért

$$p^k(A) = \sum_{j \leq k} \pi(A) P(\tau = j) + P(X_k \in A, \tau > k) = \pi(A) + P(\tau > k)[P(X_k \in A | \tau > k) - \pi(A)].$$

Ebből azonnal következik a bizonyítandó állítás. ■

Nézzünk egy alkalmazást a fenti módszerre. Egy m lapból álló kártyapaklit úgy keverünk, hogy a legfelső lapot levesszük, és belekeverjük a többi közé (azaz m helyre kerülhet a levett lap). Mivel az így meghatározott irreducibilis, aperiodikus Markov lánc átmenetmátrixa duplán sztochasztikus, a stacionárius eloszlás egyenletes lesz. Megmutatjuk, hogy a stacionárius eloszlás eléréséhez $m \ln m$ keverés elég.

6.13. Tétel. *Tekintsük a mondott példát, jelölje π az egyenletes eloszlást. Legyen $c > 0$ tetszőleges, és $k = m(\ln m + c)$. Ekkor $\|p^k - \pi\|_{TV} \leq e^{-c}$.*

Bizonyítás. Legyen τ az az időpont, amikor az eredetileg legalul lévő kártyát először keverjük bele a pakliba (mivel a pakli tetejére került). Ez megállási idő, sőt, „ügyes” megállási idő. Ez azért igaz, mert az eredetileg legalul lévő lap alá kerülő lapok minden sorrendje egyformán valószínű. Jelölje τ_n azt az időpontot, amikor az eredetileg alul lévő lap alá bekerül az n-edik lap. Ekkor

$$\tau = \tau_1 + (\tau_2 - \tau_1) + \ldots + (\tau - \tau_{m-1}),$$

ahol a tagok függetlenek, és $\tau_{i+1} - \tau_i \sim \mathrm{Geo}(\frac{i+1}{m})$. Vegyük észre, hogy τ eloszlása ugyanaz, mint a kupongyűjtési problémában (m-féle kupont kell összegyűjteni), csak ott egyre több és több kísérletre van szükség az újabb féle kuponok összegyűjtéséhez, ahogy az idő telik, míg a mi feladatunkban egyre gyorsabban fognak gyűlni a lapok az eredetileg legalul lévő lap alatt, ahogy az idő telik. Koncentráljunk tehát a kupongyűjtési feladatra, és legyen A_i az az esemény, hogy az i-edik típusú kupont nem szereztük be k kísérlet alatt.

$$P(\tau > k) = P(\cup_{i=1}^m A_i) \leq \sum_{i=1}^m P(A_i) = m(1-1/m)^k \leq m e^{-k/m} = e^{-c}.$$

Az előző tétel szerint kész vagyunk. ■

7. MCMC módszerek

Markov lánc Monte Carlo (MCMC) módszer alatt olyan eljárást értünk, amikor egy adott eloszlású valószínűségi változó előállításához, vagy az eloszlás vizsgálatához Markov láncot hívunk segítségül. Ebben a szakaszban csak néhány egyszerű példát tekintünk át.

7.1. A Hastings-Metropolis algoritmus

Tegyük fel, hogy adottak a $b_1, \ldots, b_m > 0$ számok, és legyen $B = \sum_{i=1}^m b_i$. Szeretnénk olyan valószínűségi változót generálni, melyre $P(X = i) = b_i/B$. Itt m általában nagy, B nehezen számolható, tehát a generáláshoz csak a b_i számokat szeretnénk felhasználni.

Legyen $I = \{1, \ldots, m\}$, szeretnénk ezen az állapottéren egy irreducibilis, aperiodikus Markov láncot definiálni, melynek stacionárius eloszlása $\pi_i = b_i/B$, és átmenetmátrixában csak a b_i számok szerepelnek. Ekkor tetszőleges kezdeti eloszlásból futtatva a láncot, X_n eloszlása exponenciális gyorsasággal tart a π_i eloszláshoz. A nagy számok törvénye miatt egyszersmint $E(h(X))$ is becsülhető az $\frac{1}{n}\sum_{k=N}^{N+n} h(X_k)$ átlaggal.

Vegyünk először egy tetszőleges P átmenetmátrixot, mellyel a Markov lánc irreducibilis, és állítsuk elő belőle a következő Q átmenetmátrixot:

$$q_{ij} = \alpha_{ij} p_{ij}, \text{ ha } j \neq i, \quad q_{ii} = p_{ii} + \sum_{j \neq i} (1 - \alpha_{ij}) p_{ij}.$$

Itt $0 \leq \alpha_{ij} \leq 1$ jelöli, hogy ha P szerint i-ből j-be lépne a lánc, azt mekkora valószínűséggel fogadjuk el. Ahhoz, hogy Q irreducibilis legyen, elég pl., hogy $p_{ij} > 0$ esetén $\alpha_{ij} > 0$ teljesüljön. Az aperiodicitáshoz elég pl., hogy legyen i, melyre $p_{ii} > 0$, vagy legyen $i \neq j$, melyre $p_{ij} > 0, \alpha_{ij} < 1$.

Vizsgáljuk most a stacionárius eloszlást! Az is elérhető, hogy Q megfordítható legyen π stacionárius eloszlással, azaz $\frac{\pi_j}{\pi_i} q_{ji} = q_{ij}$ minden i, j-re (ha ez igaz, akkor π_i már stacionárius eloszlás). Kell tehát, hogy

$$\pi_i p_{ij} \alpha_{ij} = \pi_j \alpha_{ji} p_{ji} \quad \forall i \neq j.$$

7.1. Állítás. *A fenti egyenlőség teljesül, ha*

$$\alpha_{ij} = \min\left(\frac{\pi_j p_{ji}}{\pi_i p_{ij}}, 1\right) = \min\left(\frac{b_j p_{ji}}{b_i p_{ij}}, 1\right), \text{ ha } p_{ij} > 0.$$

Bizonyítás. Ha $p_{ij} = p_{ji} = 0$, akkor az egyenlőség teljesül. Ha $p_{ij} > 0, p_{ji} = 0$, akkor $\alpha_{ij} = 0$, és az egyenlőség megint teljesül. Ha $p_{ij}, p_{ji} > 0$, akkor legyen pl. $\pi_j p_{ji} \leq \pi_i p_{ij}$, ekkor $\alpha_{ij} = \frac{\pi_j p_{ji}}{\pi_i p_{ij}}$, és $\alpha_{ji} = 1$, tehát

$$\pi_i p_{ij} \alpha_{ij} = \pi_j p_{ji} = \pi_j \alpha_{ji} p_{ji}.$$

■

Ezekkel az α-kkal vizsgáljuk meg, mi elég az irreducibilitáshoz! A fentiek szerint elég, hogy $p_{ij} > 0$ esetén $p_{ji} > 0$ is teljesüljön. Az aperiodikussághoz elég, hogy $p_{ii} > > 0$ valamilyen i-re, vagy hogy legyen $i \neq j$, melyre $p_{ij} > 0$ (és ezért $p_{ji} > 0$), és $p_{ji}/p_{ij} \neq \pi_i/\pi_j$. (Ezek persze nem szükséges feltételek.)

A fenti konstrukció egy speciális esete az, ha a P mátrix szimmetrikus, azaz $p_{ij} = p_{ji}$ minden i, j-re. Ekkor $\alpha_{ij} = \min(b_j/b_i, 1)$. Ebben az esetben egyébként az $\alpha_{ij} = \frac{b_j}{b_i + b_j}$ is jó választás, hiszen

$$\pi_i p_{ij} \alpha_{ij} = \frac{b_i}{B} p_{ij} \frac{b_j}{b_i + b_j} = \frac{1}{B} \frac{b_i b_j}{b_i + b_j} p_{ij} = \pi_j p_{ji} \alpha_{ji}.$$

7.2. Példa. Egy (nagy) összefüggő G gráfból szeretnénk véletlen csúcsot kiválasztani, azaz $b_s = 1$ minden s csúcsra. Jelölje egy s csúcs szomszédainak halmazát $N(s)$, ekkor legyen a P mátrix a véletlen bolyongást leíró átmenetmátrix, azaz

$$p_{st} = \frac{1}{|N(s)|}, \text{ ha } t \in N(s), \text{ és 0 egyébként.}$$

Ez irreducibilis Markov láncot határoz meg, és aperiodikus is, ha nem minden csúcsnak ugyanannyi szomszédja van. Ha viszont minden csúcsnak ugyanannyi szomszédja van, akkor a P mátrix duplán sztochasztikus, azaz már P stacionárius eloszlása egyenletes. Tehát $\alpha_{st} = \min(|N(s)|/|N(t)|, 1)$, ha $t \in N(s)$.

Ez alkalmazható pl. arra a konkrét feladatra, amikor egy olyan n elemű véletlen x permutációt szeretnénk előállítani, melyre

$$x \in \mathcal{P}_a = \{x : \sum_{j=1}^{n} jx_j > a\}$$

valamilyen nagy a számra. Ehhez szomszédsági relációt definiálunk az ilyen permutációkon: két permutáció akkor legyen szomszédos, ha transzpozícióval kaphatók egymásból. A gráf összefüggő, mert minden permutációból el lehet jutni az $(1, \ldots, n)$ permutációba az inverzióban álló elemek kicserélésével (mely a $\sum_{j=1}^{n} jx_j$ összeget növeli).

Persze $N(s)$ meghatározása nem könnyű. Könnyen látszik azonban, hogy a Hastings-Metropolis algoritmus akkor is működik, ha bizonyos b_i-k nullával egyenlők. Legyen pl. $b_1, \ldots, b_k > 0$, és $b_{k+1} = \cdots = b_m = 0$, és indítsuk a láncot az $\{1, \ldots, k\}$ halmazból. Ekkor végig itt is marad a lánc, hiszen $i \le k$, $j > k$, $p_{ij} > 0$-ra $\alpha_{ij} = 0$. Tehát ha mi a teljes $n!$ csúcsú gráfnak csak a $\mathcal{P}_a$ részhalmazából akarunk véletlenszerű csúcsot választani, akkor vehetjük a

$$p_{st} = \frac{1}{\binom{n}{2}}, \ s, t \in S_n$$

átmenetmátrixot, és az

$$\alpha_{st} = 1, \text{ ha } t \in \mathcal{P}_a, \text{ és 0 egyébként}$$

módosítást.

7.2. Gibbs mintavételező

Az előző algoritmust arra a feladatra alkalmazzuk, ha egy n dimenziós eloszlásból kell valószínűségi változót generálni, és az egydimenziós feltételes eloszlásokból könnyen tudunk generálni. Itt két n dimenziós vektor akkor lesz szomszédos, ha csak egy koordinátában különböznek.

Legyen tehát $f(x) = f(x_1, \ldots, x_n) = cb(x)$ n dimenziós eloszlás, c-t esetleg nem ismerjük. Működtessük a következő átmenetvalószínűségű Markov láncot:

$$p_{xy} = \frac{1}{n} P(X_i = y_i | X_j = x_j,\ j \neq i)$$

ha x és y szomszédosak, és épp az i. koordinátában különböznek (egyébként nulla). Megmutatjuk, hogy erre már $f(x)$ stacionárius eloszlás. Ha x és y szomszédos, és az i. koordinátában különböznek, akkor

$$f(x)p_{xy} = f(x)\frac{1}{n}\frac{f(y)}{P(X_j = x_j,\ j \neq i)} = f(y)\frac{1}{n}\frac{f(x)}{P(X_j = y_j,\ j \neq i)} = f(y)p_{yx}.$$

Természetesen az irreducibilitást, aperiodikusságot vizsgálni kell.

7.3. Példa. Nézzünk egy folytonos állapotterű példát! Legyen az (X, Y, Z) vektor sűrűségfüggvénye

$$f(x, y, z) = C \exp\{-(x + y + z + xy + xz + yz\}, \quad x, y, z > 0.$$

Szeretnénk meghatározni az $E(XYZ)$ várható értéket. Ehhez a Gibbs mintavételezővel generálunk egy (X_i, Y_i, Z_i) mintát, majd kiszámoljuk az

$$\frac{1}{N} \sum_{i=n+1}^{n+N} X_i Y_i Z_i$$

átlagot. A mintavételezéshez a mintának mindig csak az egyik, véletlenszerűen választott koordinátáját változtatjuk meg, a feltételes eloszlás szerint. Például X feltételes eloszlása az (Y, Z) párra nézve $\mathrm{Exp}(1 + Y + Z)$.

II. rész

Folytonos paraméter

Az állapottér most is diszkrét, azaz $X_t : \Omega \to I$ valószínűségi változók, és $t \geq 0$ valós szám. Legyen

$$p_{ij}(t) = P(X_{s+t} = j | X_s = i), \quad t > 0,$$

ahol az átmenetvalószínűség megint nem függ s-től. A $p_{ij}(t)$ értékekből alkotott mátrixot jelölje $P(t)$. Ekkor minden pozitív t-re $P(t)$ sztochasztikus mátrix, és a Markov tulajdonság miatt teljesül a Chapman-Kolmogorov összefüggés, azaz $P(s+t) = P(s)P(t)$. Ha adottak a $P(t)$ mátrixokat valamilyen $(0, \varepsilon)$ intervallumon, akkor a Chapman-Kolmogorov egyenletek miatt $P(t)$ tetszőleges $t > 0$ értékre kiszámítható. Azonban most nincs „legrövidebb lépésköz," ezért bonyolultabb a helyzet, mint a diszkrét paraméter esetén, ahol elegendő volt az egylépéses átmenetmátrixot megadni. Vizsgáljuk az analízis eszközeivel, milyenek lehetnek a $p_{ij}(t)$ átmenetvalószínűség-függvények!

8. Infinitezimális generátor

1. Feltevés. *Minden i, j párra a $t \mapsto p_{ij}(t)$ függvény mérhető.*

8.1. Tétel. *(i) Minden i-re és $h > 0$-ra $\Delta(t, h) = \sum_{j \in I} |p_{ij}(t + h) - p_{ij}(t)|$ t-ben monoton csökkenő. (ii) $h \to 0$ esetén $\Delta(t, h) \to 0$ a $t \in [\delta, \infty)$ félegyenesen egyenletesen, minden $\delta > 0$-ra.*

Bizonyítás. (i) Ha $0 < s < t$, akkor

$$\Delta(t, h) = \sum_j \left| \sum_k p_{ik}(s + h) p_{kj}(t - s) - \sum_k p_{ik}(s) p_{kj}(t - s) \right| \leq$$
$$\sum_k |p_{ik}(s + h) - p_{ik}(s)| \sum_j p_{kj}(t - s) = \Delta(s, h).$$

(ii) Másrészt, ha $0 < h \leq \delta \leq t$, akkor (i) miatt

$$\Delta(t, h) \leq \sum_j \frac{1}{\delta} \int_0^\delta |p_{ij}(u + h) - p_{ij}(u)| du,$$

amiből

$$\lim_{h\to 0} \Delta(t,h) \le \lim_{h\to 0} \sum_j \frac{1}{\delta} \int_0^\delta |p_{ij}(u+h) - p_{ij}(u)| du = \sum_j \frac{1}{\delta} \lim_{h\to 0} \int_0^\delta |p_{ij}(u+h) - p_{ij}(u)| du$$

a dominált konvegencia szerint, ugyanis van összegezhető majoráns:

$$\frac{1}{\delta} \int_0^\delta |p_{ij}(u+h) - p_{ij}(u)| du \le \frac{1}{\delta} \int_0^\delta p_{ij}(u+h) + p_{ij}(u) du \le \frac{2}{\delta} \int_0^{2\delta} p_{ij}(u) du.$$

Ugyanakkor $\lim_{h\to 0} \int_0^\delta |p_{ij}(u+h) - p_{ij}(u)| du = 0$, mivel az eltolás L^1-ben folytonos. ■

8.2. Következmény. *Minden $i,j \in I$ és $\delta > 0$ esetén $p_{ij}(t)$ egyenletesen folytonos a $[\delta, \infty)$ félegyenesen.*

8.3. Tétel. *Minden i,j-re létezik a $\lim_{t\to 0} p_{ij}(t) = u_{ij}$ határérték.*

Bizonyítás. Tegyük fel, hogy $t_n, t'_n \to 0$, és $p_{ij}(t_n) \to u_{ij}$, $p_{ij}(t'_n) \to u'_{ij}$. A Fatou lemma szerint ekkor $\sum_j u_{ij} \le 1$. Másrészt dominált konvergenciát használva

$$p_{ij}(t) = \lim_{n\to\infty} p_{ij}(t+t_n) = \lim_{n\to\infty} \sum_k p_{ik}(t) p_{kj}(t_n) = \sum_k p_{ik}(t) u_{kj}. \tag{10}$$

Ebből

$$1 = \sum_j p_{ij}(t) = \sum_k p_{ik}(t) \sum_j u_{kj}.$$

Ezért, ha valamely k-ra $\sum_j u_{kj} < 1$, akkor minden i-re és minden $t > 0$-ra $p_{ik}(t) = 0$ kell legyen, azaz ekkor $u'_{ik} = 0$ minden i-re. Ha (10)-ben a Chapman-Kolmogorov egyenletet fordítva írjuk fel, és a t'_n sorozatra, akkor nem hivatkozhatunk a dominált konvergenciára, viszont a Fatou lemmából kapjuk, hogy

$$p_{ij}(t) \ge \sum_k u'_{ik} p_{kj}(t).$$

Írjunk most t helyébe t_n-et, és tartson n végtelenhez:

$$u_{ij} \ge \lim_{n\to\infty} \sum_k u'_{ik} p_{kj}(t_n) = \sum_k u'_{ik} u_{kj},$$

megint dominált konvergenciát használva. Hasonlóan, ha (10)-ben t helyébe t'_n-t írunk, akkor határátmenettel $u'_{ij} \ge \sum_k u'_{ik} u_{kj}$. Ezt j-ben összegezve, a korábbi észrevételt felhasználva

$$\sum_j u'_{ij} \geq \sum_k u'_{ik} \sum_j u_{kj} = \sum_k u'_{ik},$$

azaz minden j-re $u'_{ij} = \sum_k u'_{ik} u_{kj}$ kell teljesüljön. Ebből kapjuk, hogy $u_{ij} \geq u'_{ij}$, és szimmetria miatt készen vagyunk. ■

8.4. Definíció. *A $P(t)$ család standard, ha $\lim_{t\to 0} p_{ij}(t) = \delta_{ij}$. Legyen ekkor $p_{ij}(0) = = \delta_{ij}$, azaz $P(0)$ az egységmátrix.*

2. Feltevés. *Mindig feltesszük, hogy a Markov láncunk átmenetvalószínűsége standard.*

8.5. Állítás. *Ha $P(t)$ standard, akkor $|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h)$.*

Bizonyítás.

$$p_{ij}(t+h) - p_{ij}(t) = \sum_k p_i k(h) p_{kj}(t) - p_{ij}(t) = (p_{ii}(h) - 1) p_{ij}(t) + \sum_{k\neq i} p_{ik}(h) p_{kj}(t).$$

A második tagot felülről becsülhetjük a

$$\sum_{k\neq i} p_{ik}(h) = 1 - p_{ii}(h)$$

összeggel. Tehát

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h),$$

mivel az abszolút értékben álló kifejezés két 0 és $1 - p_{ii}(h)$ közé eső tag különbsége. ■

8.6. Állítás. *Legyen P standard. A $p_{ii}(t)$ mennyiségnek létezik a $-p'_{ii}(0)$ jobboldali deriváltja, mely nemnegatív, de esetleg $+\infty$.*

Bizonyítás. Először belátjuk, hogy minden t-re $p_{ii}(t) > 0$. Mivel $\lim_{t\to 0} p_{ii}(t) = 1$, elég nagy n-re a Chapman-Kolmogorov egyenlőségből

$$p_{ii}(t) \geq p_{ii}(t/n)^n > 0.$$

Legyen most $\phi(t) = -\log p_{ii}(t)$, az előzőek szerint ez jól definiált nemnegatív, folytonos függvény a $[0,\infty)$ félegyenesen, és $\phi(0) = 0$. A $p_{ii}(s+t) \geq p_{ii}(t) p_{ii}(s)$ Ch-K egyenlőtlenségből $\phi(s+t) \leq \phi(s) + \phi(t)$. Legyen

$$q_i = \sup_{t>0} \frac{\phi(t)}{t} \leq +\infty.$$

Belátjuk, hogy $q_i = \lim_{t\to 0} \phi(t)/t$. Legyen először $q_i < \infty$, és t_0 olyan, hogy $\phi(t_0)/t_0 > > q_i - \varepsilon$. Legyen most t tetszőleges (kicsi) szám, írjuk fel t_0-t $t_0 = nt + \delta$ alakban, ahol $0 \le \delta < t$.

$$q_i - \varepsilon < \frac{\phi(t_0)}{t_0} \le \frac{n\phi(t) + \phi(\delta)}{t_0} = \frac{nt}{t_0} \cdot \frac{\phi(t)}{t} + \frac{\phi(\delta)}{t_0}.$$

Tartson t nullához:

$$q_i - \varepsilon \le \liminf_{t\to 0} \left(\frac{nt}{t_0} \cdot \frac{\phi(t)}{t} + \frac{\phi(\delta)}{t_0} \right) = \liminf_{t\to 0} \frac{\phi(t)}{t},$$

mivel $nt/t_0 \to 1$ és $\phi(\delta)/t_0 \to 0$. Azaz megkaptuk, hogy $\lim_{t\to 0} \phi(t)/t = q_i$. Ha $q_i = = \infty$, akkor hasonlóan látható be az állítás. Végül felhasználva, hogy $h \to 0$ esetén $\log(1+h) = h(1+o(1))$,

$$q_i = \lim_{t\to 0} \frac{-\log(1-(1-p_{ii}(t)))}{t} = \lim_{t\to 0} \frac{(1-p_{ii}(t))(1+o(1))}{t} = -p_{ii}'(0).$$

■

8.7. Állítás. *Ha $i \neq j$, akkor a $p_{ij}'(0)$ jobboldali nemnegatív derivált létezik, és véges.*

Bizonyítás. Legyen h tetszőleges, és vizsgáljuk a folytonos paraméterű Markov lánc h-vázát: $Y_n = X_{nh}$. Ez diszkét paraméterű Markov lánc lesz, átmenetmátrixa $P(h)$. Tegyük fel, hogy ebben a láncban n lépés alatt i-ből j-be érünk. Ennél szűkebb az az esemény, hogy úgy érünk n lépés alatt i-ből j-be, hogy j-be először i-ből lépünk (jelölje ezen lépés indexét $k+1$). Felírhatjuk, hogy

$$p_{ij}(nh) \ge \sum_{k=0}^{n-1} {}_j p_{ii}^{(k)}(h) p_{ij}(h) p_{jj}((n-k-1)h).$$

Továbbá

$$p_{ii}(kh) = {}_j p_{ii}^{(k)}(h) + \sum_{m=1}^{k-1} f_{ij}^{(m)}(h) p_{ji}((k-m)h).$$

Mivel $\sum_{m=1}^{k-1} f_{ij}^{(m)}(h) \le 1$, így

$${}_j p_{ii}^{(k)}(h) \ge p_{ii}(kh) - \max_{1\le m\le k-1} p_{ji}((k-m)h).$$

Ha most $\varepsilon > 0$ fix, és t_0 elég kicsi, akkor

$$\max_{0\le t\le t_0} p_{ji}(t) < \varepsilon, \ \min_{0\le t\le t_0} p_{ii}(t) > 1-\varepsilon, \ \min_{0\le t\le t_0} p_{jj}(t) > 1-\varepsilon.$$

Ezért, ha $nh < t_0$, akkor az összegben szereplő k értékekre

$${}_j p_{ii}^{(k)}(h) > 1-2\varepsilon, \quad p_{jj}((n-k-1)h) > 1-\varepsilon,$$

amiből

$$p_{ij}(nh) > (1-2\varepsilon)(1-\varepsilon)\sum_{k=0}^{n-1} p_{ij}(h) \ge (1-3\varepsilon)np_{ij}(h).$$

Átrendezve,

$$\frac{p_{ij}(nh)}{nh} > (1-3\varepsilon)\frac{p_{ij}(h)}{h}.$$

Jelölje $q_{ij} = \limsup_{t\to 0} p_{ij}(t)/t$, továbbá válasszunk egy fix $0 < t < t_0$ számot, és végezzük el az előző képletben a $h \to 0$, $nh \to t$ határátmenetet:

$$\frac{p_{ij}(t)}{t} \ge (1-3\varepsilon)q_{ij}.$$

Ebből azonnal következik, hogy $q_{ij} < \infty$ és

$$\liminf_{t\to 0} \frac{p_{ij}(t)}{t} \ge q_{ij},$$

vagyis létezik a mondott határérték. ■

A továbbiakban használjuk a $q_{ij} = p'_{ij}(0)$ jelölést, az ezekből alkotott mátrix legyen Q. A Q mátrix elnevezése: a Markov lánc infinitezimális generátora.

8.8. Állítás. *Minden i-re* $\sum_j q_{ij} \le 0$.

Bizonyítás.

$$\frac{1-p_{ii}(t)}{t} = \sum_{j:j\ne i} \frac{p_{ij}(t)}{t}.$$

Tartson t nullához, és használjuk a Fatou lemmát:

$$-q_{ii} = -p'_{ii}(0) \ge \sum_{j:j\ne i} p'_{ij}(0) = \sum_{j:j\ne i} q_{ij}.$$

■

9. Kolmogorov-féle differenciálegyenletek

A következő kérdés, hogy a $p_{ij}(t)$ függvények máshol is deriválhatók-e?

9.1. Tétel. *Tegyük fel, hogy* $-q_{ii} < \infty$ *minden* i*-re. Ekkor* $p_{ij}(t)$ *folytonosan differenciálható a* $[0, \infty)$ *félegyenesen.*

Ezt a tételt nem bizonyítjuk, a bizonyítása meglehetősen hosszadalmas és technikás. Vázlatosan arról van szó, hogy először

$$p_{ii}(t) \geq p_{ii}(t/n)^n = \left(p_{ii}(h)^{1/h}\right)^t,$$

ahol $h = t/n$. Mivel

$$p_{ii}(h) = 1 - h\left(\frac{1 - p_{ii}(h)}{h}\right) = 1 - h(-q_{ii} + o(h)),$$

így $p_{ii}(h)^{1/h} \to e^{q_{ii}}$. Ebből kapjuk, hogy $p_{ii}(t) \geq e^{tq_{ii}} \geq 1 + tq_{ii}$. A 8.5 Állítás miatt kijön, hogy

$$|p_{ij}(t+h) - p_{ij}(t)| \leq -hq_{ii},$$

vagyis $p_{ij}(t)$ Lipschitz-folytonos. Vegyük rögtön észre, hogy $q_{ii} = 0$ esetén $p_{ii}(t) = 1$ minden t-re, vagyis i elnyelő állapot; ez az eset triviális.

Könnyen megkapjuk, hogy

$$p_{ij}(t+h) - p_{ij}(t) \geq p_{ii}(h)p_{ij}(t) - p_{ij}(t) = (p_{ii}(h) - 1)p_{ij}(t) \geq hq_{ii}p_{ij}(t),$$

ahol a Chapman–Kolmogorov azonosságot, majd a pár sorral feljebb kiszámolt becslést használtuk. Legyen $D^+ f(t) = \liminf_{h \to 0}(f(t+h) - f(t))/h$ az f függvény alsó Dini deriváltja. Ekkor

$$D^+(e^{-tq_{ii}}p_{ij}(t)) = e^{-tq_{ii}}(D^+ p_{ij}(t) - q_{ii}p_{ij}(t)) \geq 0.$$

Dini tétele szerint $e^{-tq_{ii}}p_{ij}(t)$ monoton növő, és ezért majdnem mindenütt deriválható. Innentől további technikai jellegű meggondolásokkal adódik, hogy p_{ij} mindenütt deriválható.

Ebben a szakaszben feltesszük, hogy az infinitezimális generátor *konzervatív*, azaz $-q_{ii} < \infty$ minden i-re, továbbá minden i-re $\sum_k q_{ik} = 0$. Láttuk, hogy létezik a $P'(t)$

derivált.

$$p_{ij}(t+h) = \sum_{k\in I} p_{ik}(h)p_{kj}(t),$$

átrendezve

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = \frac{p_{ii}(h)-1}{h}p_{ij}(t) + \sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t) \to \sum_k q_{ik}p_{kj}(t),$$

ha $h \to 0$, az alábbi lemma szerint.

9.2. Lemma. *Legyenek $c_k(h) \geq 0$ és $0 \leq a_k \leq 1$ adottak, minden k-ra $c_k(h) \to c_k$, ha $h \to 0$, sőt $\sum_k c_k(h) \to \sum_k c_k < \infty$. Ekkor*

$$\sum_k c_k(h)a_k \to \sum_k c_k a_k \quad (h \to 0).$$

Bizonyítás. A Fatou lemma szerint

$$\liminf_{h\to 0} \sum_k c_k(h)a_k \geq \sum_k c_k a_k.$$

Másrészt, feltehetjük, hogy $I = \mathbb{N}$, ekkor

$$\limsup_{h\to 0} \sum_k c_k(h)a_k \leq \sum_{k=0}^{N} c_k a_k + \limsup_{h\to 0} \sum_{k=N+1}^{\infty} c_k(h)a_k \leq$$
$$\sum_{k=0}^{\infty} c_k a_k + \limsup_{h\to 0} \sum_{k=N+1}^{\infty} c_k(h) = \sum_{k=0}^{\infty} c_k a_k + \sum_{k=N+1}^{\infty} c_k,$$

és a második tag tetszőlegesen kicsi, ha N elég nagy. ∎

Mátrix alakba írva kaptuk, hogy

$$P'(t) = QP(t) \text{ Kolmogorov hátrafelé (backward) egyenletei.}$$

Mi történik, ha a $(0, t+h)$ intervallumot a $(0,t)$ és $(t, t+h)$ intervalumokra osztjuk fel?

$$p_{ij}(t+h) = \sum_{k\in I} p_{ik}(t)p_{kj}(h),$$

átrendezve

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = p_{ij}(t)\frac{p_{jj}(h)-1}{h} + \sum_{k\neq j} p_{ik}(t)\frac{p_{kj}(h)}{h}.$$

Itt már nem feltétlenül tudjuk az összegzést felcserélni a limesszel, csak ha további feltételeket teszünk, például azt, hogy rögzített j esetén a $\frac{p_{kj}(h)}{h} \to q_{kj}$ konvergencia k-ban egyenletes. Ha most $h \to 0$, akkor azt kaptuk, hogy

$$P'(t) = P(t)Q \text{ Kolmogorov előrefelé (forward) egyenletei.}$$

A fő kérdés az, hogy adott konzervatív Q mátrix esetén a Kolmogorov-féle differenciálegyenleteknek mikor létezik megoldásuk, és a megoldás mikor egyértelmű (a $P(0) = I$ kezdeti feltétel mellett). Mielőtt ezt tovább vizsgálnánk, nézzünk konkrét példákat!

10. Születési-halálozási folyamatok

A továbbiakban olyan folytonos paraméterű Markov láncokat vizsgálunk, melyek állapottere $\mathbb{N}$, továbbá az állapot mindig csak eggyel nő vagy csökken.

10.1. A Poisson folyamat

Legyenek $Z_i \sim \mathrm{Exp}(\lambda)$ független változók, $i = 1,2,\ldots$, és $X_t = \max\{k \geq 0 : \sum_{i=1}^{k} Z_i \leq t\}$. Képzelhetjük azt, hogy időnként valami történik, és két egymás utáni történés között exponenciális idő telik el. Ekkor X_t azt fejezi ki, hogy a $[0,t]$ intervallumban hány történés volt. Könnyű megmutatni, hogy X_t folytonos paraméterű Markov lánc az $\mathbb{N}$ állapottéren, mégpedig

$$p_{ij}(t) = e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!}, \text{ ha } j \geq i.$$

Ez azt jelenti, hogy egy t hosszú intervallumba eső történések száma Poisson(λt) eloszlású. Az így kapott $P(t)$ átmenetmátrixokról könnyen ellenőrizhetjük, hogy $P(t+s) = P(t)P(s)$, és a család standard, $\lim_{t\to 0} P(t) = I$. Hogy fog kinézni az infinitezimális generátor?

$$q_i = -p'_{ii}(0) = \lim_{t\to 0}\frac{1-p_{ii}(t)}{t} = \lim_{t\to 0}\frac{1-e^{-\lambda t}}{t} = \lambda.$$

Nyilván $i > j$ esetén $q_{ij} = 0$, ha pedig $i < j$,

$$q_{ij} = \lim_{t\to 0} \frac{p_{ij}(t)}{t} = \lim_{t\to 0} \frac{e^{-\lambda t}(\lambda t)^{j-i}}{t(j-i)!} = \begin{cases} \lambda & j = i+1 \\ 0 & j > i+1 \end{cases}$$

10.2. Születési folyamatok

A Poisson folyamat enyhe általánosítása, ha a Q infinitezimális generátorban csak a főátló és a fölötte lévő elemek nem nullák, azaz

$$q_i = q_{i,i+1} = \lambda_i, \quad i = 0,1,\ldots.$$

λ_i az i állapothoz tartozó születési intenzitás, ezekről feltesszük, hogy mind pozitívak. Feltesszük továbbá, hogy $p_{ij}(t) = 0$, ha $j < i$. Tegyük fel, hogy $X_0 = 0$, és legyen $r_n(t) = P(X_t = n) = p_{0n}(t)$, valamint $r(t) = (r_0(t), r_1(t), \ldots)$. Ismerjük $r(0)$-t, kérdés, hogy meg tudjuk-e határozni $r(t)$-t? Tudjuk, hogy $r_n(t)$ deriválható, sőt, az általános tételből rögtön fel is írhatnánk a deriváltat. Talán érdemes azonban közvetlenül meghatározni: nézzük a jobb oldali deriváltat, legyen először $n \geq 1$:

$$r_n(t+h) = \sum_{i=0}^{n} r_i(t)p_{in}(h) = r_{n-1}(t)(\lambda_{n-1}h + o(h)) + r_n(t)(1 - \lambda_n(h) + o(h)) + \\ + o(h) = r_{n-1}(t)\lambda_{n-1}h + r_n(t)(1 - \lambda_n h) + o(h).$$

Ebből

$$\lim_{h\to 0} \frac{r_n(t+h) - r_n(t)}{h} = \lambda_{n-1}r_{n-1}(t) - \lambda_n r_n(t).$$

Ha $n = 0$, akkor teljesen hasonlóan kapjuk, hogy $\lim_{h\to 0}(r_0(t+h) - r_0(t))/h = -\lambda_0 r_0(t)$. A baloldali deriváltak pedig megegyeznek a jobboldaliakkal, mivel tudjuk, hogy $r_n(t)$ folytonos, ezért $\lim_{h\to 0} r_n(t-h) = r_n(t)$.

A következő rendszert kell tehát megoldanunk:

$$r_0'(t) = -\lambda_0 r_0(t),\ r_n'(t) = -\lambda_n r_n(t) + \lambda_{n-1}r_{n-1}(t),\ r(0) = (1,0,0,\ldots).$$

Próbáljunk rekurzívan haladni! Az első differenciálegyenletnek az egyértelmű megoldása

$$r_0(t) = e^{-\lambda_0 t},$$

(ebből rögtön látszik, hogy a 0 állapotban $\mathrm{Exp}(\lambda_0)$ ideig tartózkodik a lánc). Legyen $v_n(t) = e^{\lambda_n t} r_n(t)$, erre $v_n'(t) = e^{\lambda_n t}\lambda_{n-1}r_{n-1}(t)$. Ennek megoldása

$$v_n(t) = e^{\lambda_n t} r_n(t) = \lambda_{n-1}\int_0^t e^{\lambda_n x} r_{n-1}(x)dx,$$

azaz

$$r_n(t) = \lambda_{n-1}e^{-\lambda_n t}\int_0^t e^{\lambda_n x} r_{n-1}(x)dx. \tag{11}$$

Kérdés, hogy vajon sztochasztikus mátrixot kapunk-e a megoldásból? Nem feltétlenül!

10.1. Tétel. *Minden i-re és $t > 0$-ra, a $\sum_j p_{ij}(t) = 1$ egyenlőség akkor és csak akkor teljesül, ha $\sum_{n=0}^\infty 1/\lambda_n = \infty$.*

Bizonyítás. A bizonyítást az $i = 0$ esetre végezzük el, ez nyilván elég. Legyen $S_k(t) = \sum_{i=0}^k r_i(t)$. Erre

$$S_k'(t) = \sum_{i=0}^k r_i'(t) = -\lambda_k r_k(t),$$

mivel a teleszkópos összegből a többi tag kiesik. Ebből integrálással kapjuk, hogy

$$1 - S_k(t) = \lambda_k \int_0^t r_k(s)ds, \text{ mivel } S_k(0) = 1.$$

Azonnal látszik, hogy $1 - S_k(t) \geq 0$, azaz a $\mu(t) = \lim_{k\to\infty}(1 - S_k(t))$ (monotonitás miatt létező) határérték nemnegatív. A

$$\mu(t) \leq 1 - S_k(t) \leq 1$$

egyenlőtlenségből

$$\lambda_k^{-1}\mu(t) \leq \int_0^t r_k(s)ds \leq \lambda_k^{-1}.$$

Összegezzük ezeket 0-tól n-ig:

$$\mu(t)\sum_{k=0}^n \lambda_k^{-1} \leq \int_0^t S_n(s)ds \leq \sum_{k=0}^n \lambda_k^{-1}.$$

Ha $\sum_{k=0}^\infty \lambda_k^{-1} = \infty$, akkor a baloldali egyenlőtlenségre tekintettel $\mu(t) = 0$ minden t-re. Fordítva, ha $\sum_{k=0}^\infty \lambda_k^{-1} < \infty$, akkor a jobboldali egyenlőtlenség miatt nem lehet $\mu(t) = 0$ minden t-re (lásd a következő lemmát). ■

10.2. Lemma. *A $\sum_j p_{ij}(t) = 1$ egyenlőség vagy minden i és $t > 0$ esetén teljesül, vagy egyre sem.*

Bizonyítás. Tegyük fel, hogy $\sum_{j=i}^{\infty} p_{ij}(t) = 1$, és legyen $0 < s < t$.

$$p_{ij}(t) = \sum_{k=i}^{j} p_{ik}(s) p_{kj}(t-s).$$

Ebből összeadással

$$1 = \sum_{j=i}^{\infty} p_{ij}(t) = \sum_{k=i}^{\infty} p_{ik}(s) \sum_{j=k}^{\infty} p_{kj}(t-s).$$

Mivel azt már láttuk, hogy minden sorösszeg legfeljebb 1, továbbá $p_{ik}(s) > 0$ minden $i \leq k$ és $s > 0$ esetén, a fenti egyenlőség csak akkor teljesülhet, ha

$$\sum_{j=k}^{\infty} p_{kj}(t-s) = 1 \text{ és } \sum_{k=i}^{\infty} p_{ik}(s) = 1.$$

Tehát $0 < s < t$ esetén a $P(s)$ mátrix sztochasztikus. Mivel sztochasztikus mátrixoknak a szorzata is az, ezért az állítás kiterjed minden t-re. ■

Ha a $P(t)$ mátrix nem sztochasztikus, az szemléletesen azt jelenti, hogy a lánc véges idő alatt kimegy a végtelenbe, azaz felrobban. Ezt a jelenséget később részletesebben is tárgyaljuk.

Nézzük példaként a Yule-folyamatot. Ez egy populáció méretét írja le, melyben minden egyed λ intenzitással szaporodik. Ez azt jelenti, hogy $\lambda_n = n\lambda$, ha $n \geq 1$. Itt nincs értelme nullából indítani a folyamatot, helyette 1 egyedből induljunk ki! A (11) képletet használhatjuk a $p_{1n}(t)$ valószínűségek kifejezésére, melyek az előző tétel szerint sztochasztikus mátrixot adnak majd. Először is, $p_{11}(t) = e^{-\lambda t}$. Ha elkezdjük a rekurziót kiszámolni, láthatóvá válik, hogy a megoldás

$$p_{1n}(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{n-1}.$$

Azt kaptuk, hogy t idő elteltével a populáció mérete $\mathrm{Geo}(e^{-\lambda t})$ eloszlású. Ha k egyedből indulunk, akkor ők egymástól függetlenül szaporodnak, vagyis a $P(t)$ átmenetmátrix k-adik sorában a $\mathrm{Negbin}(k, e^{-\lambda t})$ eloszlás jelenik meg.

Feladat: hogyan tudnánk számolás nélkül kihozni ezt az eredményt, felhasználva az exponenciális eloszlású minta rendezett mintájáról szóló ismereteket?

10.3. Születési-halálozási folyamatok

A Q olyan konzervatív mátrix, melyben $q_{i,i+1} = \lambda_i > 0$ a születési intenzitások, és $q_{i,i-1} = \mu_i > 0$ a halálozási intenzitások, $q_{ii} = -(\lambda_i + \mu_i)$, és minden más elem nulla.

10.3. Tétel. *Legyen* $\rho_0 = 1$, $\rho_n = \prod_{i=1}^{n} \frac{\lambda_{i-1}}{\mu_i}$. *Ha*

$$\sum_{n=0}^{\infty} \rho_n \sum_{k=0}^{n} \frac{1}{\lambda_k \rho_k} = \infty,$$

akkor Q *meghatározza a Markov láncot, azaz csak egy olyan* $P(t)$ *Markov-átmenet család van, melynek generátora* Q.

10.4. Példa. lineáris növekedés bevándorlással. Legyen $\lambda_n = n\lambda + a$, $\mu_n = n\mu$, ahol $\lambda, \mu, a > 0$. Legyen $M(t) = (M_0(t), M_1(t), \ldots)$, ahol

$$M_i(t) = E(X(t)|X(0) = i) = \sum_{j=0}^{\infty} j p_{ij}(t)$$

jelöli, hogy ha kezdetben i tagú volt a populáció, akkor t idő múlva várhatóan hány tagú lesz. Legyen $\mathbf{1}$ a csupa egyesből álló vektor, és $\mathbf{f} = (0,1,2,\ldots)$. Ezzel a jelöléssel $M(t) = P(t)\mathbf{f}$. Ha teljesülnek a Kolmogorov-féle előre egyenletek, akkor $P'(t) = P(t)Q$, azaz

$$M'(t) = (P(t)\mathbf{f})' = P'(t)\mathbf{f} = P(t)Q\mathbf{f}.$$

Számítsuk ki $Q\mathbf{f}$ i-edik koordinátáját:

$$(Q\mathbf{f})_i = \sum_j q_{ij}\mathbf{f}_j = \mu_i(i-1) - (\lambda_i + \mu_i)i + \lambda_i(i+1) = \lambda_i - \mu_i = a + (\lambda - \mu)i.$$

Mátrixos alakban, $Q\mathbf{f} = a\mathbf{1} + (\lambda - \mu)\mathbf{f}$. Kaptuk tehát, hogy

$$M'(t) = P(t)Q\mathbf{f} = P(t)(a\mathbf{1} + (\lambda - \mu)\mathbf{f}) = a\mathbf{1} + (\lambda - \mu)M(t).$$

A differenciálegyenlet kezdeti feltétele $M(0) = \mathbf{f}$, a megoldás pedig

$$M(t) = \begin{cases} at\mathbf{1} + \mathbf{f}, & \text{ha } \lambda = \mu, \\ \frac{a}{\lambda-\mu}(e^{(\lambda-\mu)t} - 1)\mathbf{1} + e^{(\lambda-\mu)t}\mathbf{f}, & \text{ha } \lambda \neq \mu. \end{cases}$$

11. Visszatérőség

11.1. Az állapotok osztályozása

Folytonos idejű Markov láncoknál is mondhatjuk, hogy a j állapot elérhető i-ből, mégpedig akkor, ha van $t \geq 0$, melyre $p_{ij}(t) > 0$. Két állapot akkor érintkezik, ha kölcsönösen elérhetők egymásból. Ez nyilván ekvivalenciareláció, mely osztályokra bontja az állapotteret.

11.1. Lemma. *Ha $P(t)$ standard, akkor $p_{ii}(t) > 0$ minden t-re, és $p_{ij}(t_0) > 0$ esetén $p_{ij}(t) > 0$ minden $t \geq t_0$.*

Bizonyítás. Az első állítás már szerepelt, a második pedig triviális a

$$p_{ij}(t) \geq p_{ij}(t_0)p_{jj}(t - t_0) > 0$$

egyenlőtlenség miatt. ■

Megjegyezzük, hogy ennél több is igaz: ha $p_{ij}(t_0) > 0$ valamely t_0-ra, akkor $p_{ij}(t) > 0$ minden pozitív t-re.

11.2. Következmény. *Ha a folytonos idejű Markov láncban i-ből elérhető j, akkor a lánc h-diszkretizáltjában is, minden $h > 0$ esetén. Fordítva, ha valamely $h > 0$ számra a lánc h-diszkretizáltjában i-ből elérhető j, akkor a folytonos idejű Markov láncban is.*

Tehát az állapotok osztályozása és lényegessége megegyezik az összes h-diszkretizáltban és a folytonos idejű láncban. Továbbá a h-diszkretizáltakban minden állapot aperiodikus ($p_{ii}(h) > 0$).

11.2. A Markov lánc trajektóriái

Eddig csak a $P(t)$ átmenetmátrixok családjával foglalkoztunk, vizsgáljuk most meg a Markov lánc trajektóriáit. A Kolmogorov alaptétel segítségével bizonyítható, hogy adott $P(t)$ családhoz és kezdeti eloszláshoz létezik Markov lánc (ezt nem részletezzük). A továbbiakban tegyük fel, hogy $I \subset \overline{\mathbb{R}}$, az állapotok kiterjesztett valós számok.

11.3. Tétel. *Tegyük fel, hogy a $P(t)$ család standard. Ekkor van olyan X_t Markov lánc, melynek átmenetmátrixa $P(t)$, és X_t sztochasztikusan folytonos, jól-szeparálható, és mérhető.*

A tételt nem bizonyítjuk, viszont tisztázzuk a benne szereplő fogalmakat! Mérhetőség alatt azt értjük, hogy az $(\omega, t) \mapsto X_t(\omega)$ hozzárendelés mérhető (a szorzatmérték szerint). A sztochasztikus folytonosság jelentése, hogy $t \to t_0$ esetén X_t sztochasztikusan tart X_{t_0}-hoz.

11.4. Definíció. *Az X_t folyamat jól-szeparálható, ha minden $R \subset [0, \infty)$ megszámlálható sűrű halmazhoz van olyan nullmértékű $N \subset \Omega$ esemény, hogy a következő teljesül: Minden $A \subset \overline{\mathbb{R}}$ zárt részhalmazra és minden G nyílt intervallumra*

$$\{\omega : X_t(\omega) \in A\ \forall t \in G \cap R\} \setminus \{\omega : X_t(\omega) \in A\ \forall t \in G\} \subset N.$$

Egy jól-szeparálható folyamat majdnem minden realizációja teljesül, hogy ha R megszámlálható sűrű halmaz, akkor $X_t(\omega)$ torlódási pontja az

$$\{X_r(\omega) : r \in R \cap (t - 1/n, t + 1/n), r \neq t\}$$

halmaznak, minden n-re.

Jelölje

$$S_i(\omega) = \{t : X_t(\omega) = i\}$$

azokat az időpontokat, amikor a lánc az i állapotban tartózkodik. Rögzített i állapotra legyen $\zeta(t, \omega) = I(X_t(\omega) = i)$, és jelölje a számegyenesen a Lebesgue-mértéket μ. Ha feltesszük, hogy $X_0 = i$, akkor

$$E(\mu(S_i)) = \int_\Omega \mu(S_i(\omega))dP = \int_\Omega \int_0^\infty \zeta(t, \omega)d\mu dP = \int_0^\infty \int_\Omega \zeta(t, \omega)dPd\mu = \int_0^\infty p_{ii}(t)dt.$$

11.5. Tétel. *Ha az X_t Markov lánc jól-szeparálható, akkor*

$$P(X_s = i\ \forall 0 \leq s \leq t | X_0 = i) = e^{-q_i t}.$$

Bizonyítás. A feltétel miatt

$$P(X_s = i\ \forall 0 \leq s \leq t | X_0 = i) = \lim_{n\to\infty} p_{ii}(t/2^n)^{2^n} = e^{q_{ii}t}.$$

■

Ez a tétel a $q_i = \infty$ esetre is érvényes. Ha most feltesszük, hogy $q_i < \infty$, akkor megkaptuk, hogy az i-ben való tartózkodás ideje $\mathrm{Exp}(q_i)$ eloszlású. Az észrevétel lehetőséget teremt arra, hogy adott Q konzervatív mátrixhoz megkonstruáljuk a *minimális Markov láncot*: a lánc realizációját a $Z_0, Z_1, \dots$ iid. egységnyi paraméterű exponenciális változók segítségével építjük fel. A kiinduló állapotot a kezdeti eloszlás szerint választjuk (jelölje ezt i), majd ott tartózkodunk Z_0/q_i ideig. Ekkor átugrunk valamelyik másik állapotba, méghozzá a j állapotot q_{ij}/q_i valószínűséggel választjuk. Ha az n-edik ugrás után a k állapotba jutottunk, ott Z_n/q_k ideig tartózkodunk. Ugyanezt a konstrukciót elmondhatjuk úgy is, hogy ha egy ugrás az i állapotba vitt, akkor az összes többi állapotban azonnal ketyegni kezd egy-egy óra, a j állapot órája $\mathrm{Exp}(q_{ij})$ eloszlású idő után csörög. Amikor az első óra megszólal, átugrunk a hozzá tartozó állapotba. A $q_i < \infty$ feltétel biztosítja, hogy az esetleg végtelen sok csörgés között biztosan lesz első.

11.3. Visszatérőség

11.6. Tétel. *Konzervatív Q mátrix esetén (i) $\exists h > 0 : \sum_n p_{ii}(nh) = \infty \Rightarrow \int_0^\infty p_{ii}(t)dt = = \infty \Rightarrow \forall h > 0 : \sum_n p_{ii}(nh) = \infty$.*
(ii) A $P_i(\mu(S_i) = \infty)$ valószínűség 0 vagy 1 aszerint, hogy $\int_0^\infty p_{ii}(t)dt$ véges vagy végtelen.

Bizonyítás. (i) Tetszőleges $h > 0$-ra legyen $\delta(h) = \min_{0\leq s\leq h} p_{ii}(s)$. Tudjuk, hogy $\delta(h) > 0$.

$$\min_{nh\leq t\leq(n+1)h} p_{ii}(t) = \min_{0\leq s\leq h} p_{ii}(nh+s) \geq p_{ii}(nh)\delta(h).$$

Ugyanakkor, ha $nh \leq t \leq (n+1)h$, akkor

$$p_{ii}((n+1)h) \geq p_{ii}(t)p_{ii}((n+1)h - t) \geq p_{ii}(t)\delta(h),$$

amiből

$$p_{ii}((n+1)h) \geq \max_{nh\leq t\leq(n+1)h} p_{ii}(t)\delta(h).$$

Összerakva kapjuk, hogy

$$\delta(h)h\sum_{n=0}^{N-1}p_{ii}(nh)\leq\int_0^{Nh}p_{ii}(t)dt\leq\frac{1}{\delta(h)}h\sum_{n=1}^{N}p_{ii}(nh).$$

(ii) Ha $\int_0^\infty p_{ii}(t)dt$ véges, akkor $E(\mu(S_i))$ véges, tehát $\mu(S_i)$ 1 valószínűséggel véges. Ha viszont az integrál végtelen, akkor az (i) pont szerint i rekurrens állapot az X_n diszkrét paraméterű Markov láncban, ezért 1 valószínűséggel $X_n = i$ végtelen sok n egészre. Legyen τ_n az i-be való n-edik visszatérés ideje az X_n láncban, és

$$A_n = \{X_t = i,\ \tau_n \leq t \leq \tau_n + h\},$$

ahol $h < 1$. Mivel az A_n események függetlenek (erős Markov tulajdonság) és valószínűségük nullánál nagyobb konstans, 1 valószínűséggel végtelen sok teljesül közülük. ■

11.7. Következmény. *Az i állapot vagy az összes h-diszkretizáltban visszatérő, vagy egyben sem.*

A folytonos idejű Markov láncban azt mondjuk, hogy az i állapot visszatérő, ha $\int_0^\infty p_{ii}(t)dt = \infty$.

11.4. Stacionárius eloszlás

11.8. Tétel. *Legyen $P(t)$ standard. Minden i, j-re létezik a $\lim_{t\to\infty} p_{ij}(t) = \pi_{ij}$ határérték.*

Bizonyítás. Nézzük a lánc h-diszkretizáltját, ebben minden állapot aperiodikus, ezért létezik a

$$\lim_{n\to\infty} p_{ij}(nh) = \frac{f^*_{ij}(h)}{m_j(h)} = \pi_{ij}(h)$$

határérték. Belátjuk, hogy $p_{ij}(t)$ Cauchy sorozat, azaz minden $\varepsilon > 0$ esetén van t_0, hogy $t, t' > t_0$ esetén

$$|p_{ij}(t) - p_{ij}(t')| < \varepsilon.$$

A 8.5 Állítás szerint $p_{ij}(t)$ egyenletesen folytonos, azaz létezik h, hogy $|s-u| < h$ esetén

$$|p_{ij}(s) - p_{ij}(u)| < \varepsilon/3.$$

Rögzítsük ezt a h-t, ekkor $t = nh + s$, $t' = n'h + s'$, és

$$|p_{ij}(t) - p_{ij}(t')| \le |p_{ij}(t) - p_{ij}(nh)| + |p_{ij}(nh) - p_{ij}(n'h)| + |p_{ij}(n'h) - p_{ij}(t')|.$$

Ha $n, n' > n_0(h)$, akkor a középső tag kisebb, mint $\varepsilon/3$, így a baloldal kisebb, mint ε. Tehát a $t_0 = n_0(h)h$ választás megfelelő. ∎

Megkaptuk, hogy $\pi_{ij}(h) = \pi_{ij}$. Tegyük most fel, hogy a folytonos idejű Markov lánc irreducibilis, és rekurrens. Ekkor ez minden diszkretizáltra is teljesül, azaz $f^*_{ij}(h) = 1$ minden i, j, h esetén. Ebből kapjuk, hogy az $m_j(h)$ átlagos visszatérési idő nem függ h-tól! Tehát vagy minden diszkretizált pozitív, vagy mindegyik nulla rekurrens. Pozitív rekurrens esetben tehát

$$\lim_{t\to\infty} p_{ij}(t) = \pi_j,$$

ahol π_j az összes diszkretizált közös stacionárius eloszlása. Hogyan tudjuk vajon a stacionárius eloszlást a generátor mátrixból kiszámítani?

A $\pi_i = \sum_j \pi_j p_{ji}(t)$ egyenletet átrendezve kapjuk, hogy

$$0 = \pi_i(p_{ii}(t) - 1) + \sum_{j\neq i} \pi_j p_{ji}(t),$$

majd

$$0 = \pi_i \frac{p_{ii}(t) - 1}{t} + \sum_{j\neq i} \pi_j \frac{p_{ji}(t)}{t}.$$

Tegyük fel, hogy minden i-re a $\lim_{t\to 0} p_{ji}(t)/t = q_{ji}$ konvergencia j-ben egyenletes (ezt a feltételt már a Kolmogorov előremenő egyenleteknél is láttuk), ekkor érvényes a

$$0 = \sum_j \pi_j q_{ji}$$

összefüggés, mátrix alakban $\pi Q = 0$. Ezt az összefüggést a Kolmogorov előre egyenletekből is megkaphatjuk, ha a $P'(t) = P(t)Q$ egyenletben $t \to \infty$ határértéket veszünk.

Megjegyezzük, hogy ha találunk olyan π_i mennyiségeket, melyekre $\pi_i q_{ij} = \pi_j q_{ji}$, akkor $\pi Q = 0$. Ezt például a születési-halálozási folyamatokra felírva, $\pi_i \lambda_i = \pi_{i+1}\mu_{i+1}$ adódik, melynek megoldása $\pi_i = \rho_i \pi_0$. (Emlékeztető: $\rho_0 = 1$, $\rho_n = \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i}$.)

Legyen most még $\tilde\rho_0 = 1$, $\tilde\rho_n = \prod_{i=1}^n \frac{\mu_i}{\lambda_i}$

11.9. Tétel. (Karlin-McGregor tétel) *Tekintsünk egy születési-halálozási folyamatot a természetes számokon, és legyen $\lambda_i, \mu_i > 0$ minden szóbajövő i-re.*

(i) A Kolmogorov-egyenleteknek akkor és csak akkor létezik egyértelmű megoldása, ha

$$S = \sum_{i=0}^{\infty} \rho_i = \infty \text{ vagy } \tilde{S} = \sum_{i=0}^{\infty} \tilde{\rho}_i = \infty.$$

Ebben az esetben a lánc
(ii) tranziens, ha $S = \infty$ és $\tilde{S} < \infty$,
(iii) nulla rekurrens, ha $S = \infty$ és $\tilde{S} = \infty$,
(iv) pozitív rekurrens, ha $S < \infty$ és $\tilde{S} = \infty$.

11.10. Példa. (M/M/1 sor) Egy rendszerbe λ-Poisson folyamat szerint érkeznek igények, melyeket egy kiszolgáló egység szolgál ki (érkezési sorrendben). A kiszolgálási idő eloszlása $\mathrm{Exp}(\mu)$. Ha X_t jelöli azt, hogy a t időpontban hány igény tartózkodik a rendszerben, születési-halálozási folyamatot kapunk, melyre $\lambda_i = \lambda$, $\mu_i = \mu$. Az eddigiek alapján $\lambda > \mu$ esetén a lánc tranziens, $\lambda = \mu$ esetben nulla rekurrens, $\lambda < \mu$ esetén pozitív rekurrens. Utóbbi esetben a stacionárius eloszlás

$$\pi_i = \left(\frac{\lambda}{\mu}\right)^i \left(1 - \frac{\lambda}{\mu}\right), \quad i = 0,1,\ldots.$$

Stacionaritás esetén a rendszerben tartózkodó igények várható száma $\frac{\lambda}{\mu-\lambda}$.

11.11. Példa. (M/M/∞ sor) Az előző példán annyit módosítsunk, hogy végtelen sok kiszolgáló egység van, tehát minden beérkező igényt azonnal elkezdhetünk kiszolgálni. Most a $\lambda_i = \lambda$, $\mu_i = i\mu$ paraméterek szerepelnek a generátor mátrixban. Ez a lánc pozitív rekurrens lesz, stacionárius eloszlása $\mathrm{Poisson}(\lambda/\mu)$.

12. A Kolmogorov-egyenletek megoldhatóságáról

Legyen Q konzervatív. Ebben a szakaszban megkonstruáljuk a Kolmogorov-féle egyenletek $\overline{P}(t)$ *minimális megoldását*. Ez olyan standard szubsztochasztikus átmenetmátrix-család lesz, mely kielégíti mindkét egyenletrendszert, továbbá ha a $P(t)$ szubsztochasztikus család generátora szintén Q, akkor $p_{ij}(t) \geq \overline{p}_{ij}(t)$ minden i, j, t-re. Ebből az is következik, hogy ha a minimális megoldás sztochasztikus, akkor a kétféle egyenletrendszernek nincs más (szubsztochasztikus) megoldása.

Nézzük példaként a születési folyamatot! Láttuk, hogy az előrefelé egyenlet megoldása egyértelmű. Ha $\sum \frac{1}{\lambda_i} = \infty$, akkor ez a megoldás sztochasztikus, és ez az egyetlen megoldása mindkét egyenletrendszernek. Ha viszont $\sum \frac{1}{\lambda_i} < \infty$, akkor az előrefelé egyenletek megoldása szubsztochasztikus, de a hátrafelé egyenleteknek vannak sztochasztikus megoldásaik is (melyek nem elégítik ki az előrefelé egyenleteket).

Szemléletesen $\overline{p}_{ij}(t)$ annak valószínűsége lesz, hogy t idő alatt véges sok ugrással jut a lánc i-ből j-be. Vagyis

$$\overline{p}_{ij}(t) = \sum_{n=0}^{\infty} p_{ij}^{\langle n\rangle}(t),$$

ahol $p_{ij}^{\langle n\rangle}(t)$ annak valószínűsége, hogy t idő alatt pontosan n ugrással jut a lánc i-ből j-be. A $p_{ij}^{\langle n\rangle}(t)$ mennyiségek rekurzívan kaphatók (a Q mátrixból):

$$p_{ij}^{\langle 0\rangle}(t) = \delta_{ij} e^{-q_i t},$$

hiszen nulla ugrással csak úgy juthatunk az i állapotból j-be, ha $i = j$, és végig i-ben voltunk, melynek esélye az exponenciális tartózkodási időből számolható. A rekurziós képlet pedig

$$p_{ij}^{\langle n+1\rangle}(t) = \sum_{k\neq i} \int_0^t e^{-q_i(t-s)} q_{ik} p_{kj}^{\langle n\rangle}(s) ds, \tag{12}$$

ahol k jelöli azt az állapotot, ahová először ugrott a lánc, és $t - s$ jelöli ennek az első ugrásnak az időpontját. A fenti képlet n szerinti összegzésével a következő integrálegyenletet kapjuk $\overline{p}$-ra:

$$\overline{p}_{ij}(t) = \delta_{ij} e^{-q_i t} + \sum_{k\neq i} \int_0^t e^{-q_i(t-s)} q_{ik} \overline{p}_{kj}(s) ds. \tag{13}$$

12.1. Tétel. *$\overline{P}$ szubsztochasztikus mátrix minden t-re.*

Bizonyítás. A definícióból azonnal világos, hogy $p_{ij}^{\langle n\rangle}(t)$, és így $\overline{p}_{ij}(t)$ nemnegatív. Másrészt n szerinti indukcióval megmutatjuk, hogy

$$\sum_j \sigma_{ij}^{\langle n\rangle}(t) \leq 1,$$

ahol $\sigma_{ij}^{\langle n\rangle}(t) = \sum_{m=0}^{n} p_{ij}^{\langle m\rangle}(t)$. Ebből az $n \to \infty$ határátmenettel kapjuk, hogy $\sum_j \overline{p}_{ij}(t) \leq$ ≤ 1. Az indukciót elkezdhetjük, mivel

$$\sum_j \sigma_{ij}^{\langle 0\rangle}(t) = e^{-q_i t} \leq 1.$$

Az indukciós lépéshez először fel kell írnunk a $\sigma_{ij}^{\langle n\rangle}(t)$ mennyiségekre vonatkozó rekurziót:

$$\sigma_{ij}^{\langle n+1\rangle}(t) = \delta_{ij}e^{-q_i t} + \sum_{k\neq i}\int_0^t e^{-q_i(t-s)}q_{ik}\sigma_{kj}^{\langle n\rangle}(s)ds.$$

Ezt felhasználva,

$$\sum_j \sigma_{ij}^{\langle n+1\rangle}(t) = e^{-q_i t} + \sum_{k\neq i}\int_0^t e^{-q_i(t-s)}q_{ik}\sum_j \sigma_{kj}^{\langle n\rangle}(s)ds.$$

A $\sum_j \sigma_{kj}^{\langle n\rangle}(s) \leq 1$ indukciós feltevést beírva, és az integrálást elvégezve kapjuk, hogy

$$\sum_j \sigma_{ij}^{\langle n+1\rangle}(t) \leq e^{-q_i t} + \sum_{k\neq i} q_{ik}\frac{1}{q_i}(1-e^{-q_i t}) = e^{-q_i t} + \frac{1}{q_i}(1-e^{-q_i t})q_i = 1.$$

■

12.2. Tétel. *A $\overline{P}(t)$ család megoldása a Kolmogorov-féle hátrafelé egyenleteknek.*

Bizonyítás. A $\overline{P}(t)$ családra vonatkozó (13) integrálegyenletet deriváljuk le! Emlékeztető: ha

$$F(u,v) = \int_0^v e^{-q_i(u-s)}\overline{p}_{kj}(s)ds,$$

akkor

$$F'(t,t) = \frac{\partial}{\partial u}F\Big|_{u=t,v=t} + \frac{\partial}{\partial v}F\Big|_{u=t,v=t}.$$

$$\overline{p}_{ij}'(t) = -q_i\delta_{ij}e^{-q_i t} + \sum_{k\neq i} q_{ik}\left[\overline{p}_{kj}(t) + \int_0^t -q_i e^{-q_i(t-s)}\overline{p}_{kj}(s)ds\right] =$$

$$= q_{ii}\overline{p}_{ij}(t) + \sum_{k\neq i} q_{ik}\overline{p}_{kj}(t) = \sum_k q_{ik}\overline{p}_{kj}(t).$$

(Felhasználtuk, hogy $-q_i = q_{ii}$.) Az összeget az egyenletes konvergencia miatt lehetett tagonként deriválni. ■

12.3. Tétel. *A $\overline{P}(t)$ család megoldása a Kolmogorov-féle előrefelé egyenleteknek.*

Bizonyítás. Ezt a tételt csak vázlatosan bizonyítjuk. Először is, fel kellene írnunk újra annak valószínűségét, hogy t idő alatt pontosan n ugrással jutunk i-ből j-be, de most aszerint bontjuk fel ezt az eseményt, hogy az utolsó ugrásnál melyik állapotból jutottunk j-be. Legyen tehát

$$p_{ij}^{(0)}(t) = \delta_{ij} e^{-q_j t},$$

és

$$p_{ij}^{(n+1)}(t) = \sum_{k \neq j} \int_0^t p_{ik}^{(n)}(s) q_{kj} e^{-q_j(t-s)} ds.$$

Megmutatható (n szerinti indukcióval), hogy $p_{ij}^{(n)}(t) = p_{ij}^{\langle n \rangle}(t)$, azaz tényleg ugyanazt a valószínűséget definiáltuk kétféleképp. Így $\overline{P}(t)$-re kapunk egy második integrálegyenletet:

$$\overline{p}_{ij}(t) = \delta_{ij} e^{-q_i t} + \sum_{k \neq j} \int_0^t \overline{p}_{ik}(s) q_{kj} e^{-q_j(t-s)} ds. \tag{14}$$

Ezt lederiválva pont a Kolmogorov-féle előre-egyenleteket kapjuk (illetve itt kicsit trükközni kell az összeg tagonkénti deriválásával). ■

12.4. Tétel. *A $\overline{P}(t)$ szubsztochasztikus család standard, és teljesíti a Chapman-Kolmogorov egyenleteket.*

Bizonyítás. A család standardságát könnyű belátni:

$$\liminf_{t \to 0} \overline{p}_{ii}(t) \geq \lim_{t \to 0} p_{ii}^{\langle 0 \rangle}(t) = 1.$$

Azaz $\lim_{t \to 0} \overline{p}_{ii}(t) = 1$, amiből már következik, hogy $\lim_{t \to 0} \overline{p}_{ij}(t) = 0$, ha $i \neq j$.

A Chapman-Kolmogorov egyenletekre rátérve, azt fogjuk belátni, hogy

$$p_{ij}^{\langle n \rangle}(s+t) = \sum_{m=0}^{n} \sum_k p_{ik}^{\langle m \rangle}(s) p_{kj}^{\langle n-m \rangle}(t). \tag{15}$$

Vagyis aszerint bontjuk fel a baloldali valószínűséghez tartozó eseményt, hogy az s időpontban melyik állapotban van a lánc, és addig hányat ugrott. Ha a (15) egyenlőséget n szerint összegezzük, épp a $\overline{p}_{ij}(t+s) = \sum_k \overline{p}_{ik}(s) \overline{p}_{kj}(t)$ Chapman-Kolmogorov egyenletet kapjuk.

Könnyen ellenőrizhető, hogy (15) teljesül $n = 0$-ra. Az indukciós lépés pedig:

$$\sum_{m=0}^{n+1}\sum_k p_{ik}^{\langle m\rangle}(s)p_{kj}^{\langle n+1-m\rangle}(t) =$$

$$\sum_k p_{ik}^{\langle 0\rangle}(s)p_{kj}^{\langle n+1\rangle}(t) + \sum_{m=1}^{n+1}\sum_k\sum_{\ell\neq i}\int_0^s e^{-q_iu}q_{i\ell}p_{\ell k}^{\langle m-1\rangle}(s-u)p_{kj}^{\langle n+1-m\rangle}(t)du =$$

$$= e^{-q_is}p_{ij}^{\langle n+1\rangle}(t) + \sum_{\ell\neq i}\int_0^s e^{-q_iu}q_{i\ell}p_{\ell j}^{\langle n\rangle}(s-u+t)du$$

az indukciós feltétel szerint. Az első tag másképp (az integrálban az u helyébe $u+s$-et írva):

$$e^{-q_is}p_{ij}^{\langle n+1\rangle}(t) = e^{-q_is}\sum_{\ell\neq i}\int_0^t e^{-q_iu}q_{i\ell}p_{\ell j}^{\langle n\rangle}(t-u)du = \sum_{\ell\neq i}\int_s^{s+t} e^{-q_iu}q_{i\ell}p_{\ell j}^{\langle n\rangle}(t+s-u)du.$$

Ezért

$$\sum_{m=0}^{n+1}\sum_k p_{ik}^{\langle m\rangle}(s)p_{kj}^{\langle n+1-m\rangle}(t) = \sum_{\ell\neq i}\int_0^{s+t} e^{-q_iu}q_{i\ell}p_{\ell j}^{\langle n\rangle}(s-u+t)du = p_{ij}^{\langle n+1\rangle}(s+t).$$

■

12.5. Tétel. *Legyen $P(t)$ olyan szubsztochasztikus átmenetmátrix család, melynek generátora Q. Ekkor $p_{ij}(t) \geq \overline{p}_{ij}(t)$ minden i, j, t-re.*

Bizonyítás. A korábbi tételek alapján $p_{ij}(t)$ folytonosan differenciálható, és

$$p'_{ij}(t) \geq -q_ip_{ij}(t) + \sum_{k\neq i} q_{ik}p_{kj}(t). \tag{16}$$

Ezt ugyan csak sztochasztikus családra láttuk be, de az eredmények igazak szubsztochasztikus esetben is. Ha ugyanis a $P(t)$ szubsztochasztikus mátrix, akkor vezessünk be egy új, elnyelő állapotot, jelölje ezt ∞ (azaz $p_{\infty\infty}(t) = 1$). Legyen még

$$p_{i\infty}(t) = 1 - \sum_{j=0}^{\infty} p_{ij}(t).$$

Az így kiterjesztett $P(t)$ család már sztochasztikus, és teljesíti a Chapman-Kolmogorov egyenleteket.

A (16) becslés alkalmazásával

$$(e^{q_i t} p_{ij}(t))' = q_i e^{q_i t} p_{ij}(t) + e^{q_i t} p'_{ij}(t) \geq e^{q_i t} \sum_{k \neq i} q_{ik} p_{kj}(t),$$

amiből integrálással kapjuk, hogy

$$e^{q_i t} p_{ij}(t) - \delta_{ij} \geq \sum_{k \neq i} \int_0^t e^{q_i s} q_{ik} p_{kj}(s) ds.$$

A $p_{ij}(t)$ mennyiségre átrendezéssel a

$$p_{ij}(t) \geq e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t e^{-q_i (t-s)} q_{ik} p_{kj}(s) ds \tag{17}$$

becslést kapjuk. Mivel a jobboldalon csupa nemnegatív tag áll, rögtön adódik, hogy $p_{ij}(t) \geq e^{-q_i t} \delta_{ij} = p_{ij}^{\langle 0 \rangle}(t)$. Indukcióval bizonyítjuk, hogy $p_{ij}(t) \geq \sum_{m=0}^{n} p_{ij}^{\langle m \rangle}(t)$ minden n-re igaz, tehát $p_{ij}(t) \geq \overline{p}_{ij}(t)$. A (17) egyenlőtlenség jobboldalára alkalmazva az indukciós feltevést:

$$p_{ij}(t) \geq e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t e^{-q_i (t-s)} q_{ik} \sum_{m=0}^{n} p_{kj}^{\langle m \rangle}(s) ds = p_{ij}^{\langle 0 \rangle}(t) + \sum_{m=0}^{n} p_{ij}^{\langle m+1 \rangle}(t) = \sum_{m=0}^{n+1} p_{ij}^{\langle m \rangle}(t),$$

ahol a második lépésben előrevettük az m szerinti összegzést, és a (12) összefüggést használtuk. ■