Diszkrét és folytonos paraméterű Markov láncok

Csiszár Villő

Tartalomjegyzék

1. Bevezetés **1**

I. Diszkrét paraméterű Markov láncok **2**

2. Markov-tulajdonság **2**

3. Állapotok osztályozása **4**

4. Visszatérőség **6**

5. Az átmenetvalószínűségek konvergenciája **12**

6. Stacionárius eloszlás **15**

7. Pozitív rekurrens Markov láncok **19**

7.1. Nagy számok törvénye 20

7.2. Centrális határeloszlás-tétel 22

7.3. Tabu állapotok 24

7.4. A CHT-ben szereplő szórásnégyzet kiszámítása 26

8. Reguláris mérték **28**

8.1. Megfordított láncok 31

9. Véges állapotterű Markov láncok **32**

9.1. Perron-Frobenius tételkör 32

9.2. A konvergenciasebesség becslése megállási időkkel 37

10. MCMC módszerek **38**

10.1. A Hastings-Metropolis algoritmus 38

10.2. Gibbs mintavételező 40

II. Folytonos paraméter **42**

11. Infinitezimális generátor **42**

12. Kolmogorov-féle differenciálegyenletek **47**

13. Születési-halálozási folyamatok **49**
13.1. A Poisson folyamat 49
13.2. Születési folyamatok 50
13.3. Születési-halálozási folyamatok 53

14. Visszatérőség **54**
14.1. Az állapotok osztályozása 54
14.2. A Markov lánc trajektóriái 54
14.3. Visszatérőség 56
14.4. Stacionárius eloszlás 57

15. A Kolmogorov-egyenletek megoldhatóságáról **59**

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. 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

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

A Markov-tulajdonság ekvivalens megfogalmazásai: Megmutatható, 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:

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

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.

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

■

3. Állapotok osztályozása

3.1. 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).

3.2. 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.*

3.3. 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. ■

3.4. 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: $n = \ell n_0 + q$. Ekkor

$$nd = \ell d n_0 + qd = \ell d \sum_{k=1}^{K} n_k + \sum_{k=1}^{K} c_k n_k q = \sum_{k=1}^{K} (\ell d + q c_k) n_k,$$

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

3.5. 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?

4. Visszatérőség

Vezessük be a következő jelöléseket:

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

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

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

4.1. 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.

4.2. Állítás. *Igazak a következő összefüggések:*

1. *$i \neq j$ esetén $i \to j$ akkor és csak akkor, ha $f_{ij}^{*} > 0$.*
2. *A nem lényeges állapotok tranziensek (fordítva azonban nem igaz).*

3. $g_{ij} = f^*_{ij} g_{jj}$.

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

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

Bizonyítás. A fenti állítások az utolsó kivételével egyszerűen láthatók.

1. $f_{ij}^{(n)} \leq p_{ij}^{(n)}$ és $p_{ij}^{(n)} \leq \sum_{m=1}^{n} f_{ij}^{(m)}$, ezért $\sup_{n\geq 1} p_{ij}^{(n)} \leq f^*_{ij} \leq \sum_{n=1}^{\infty} p_{ij}^{(n)}$.

2. Legyen j olyan, hogy $i \to j$, de $j \not\to i$. Létezik olyan állapotsorozat, melyre $\hat{p} = p_{ii_1} p_{i_1 i_2} \cdots p_{i_n j} > 0$, és a belső állapotok egyike sem i. Ekkor $f^*_{ii} \leq 1 - \hat{p}$.

3. A „végtelen sok n-re" kifejezésre a v.s. rövidítést használva,

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

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

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

Tehát $g_{ii}(m) = (f^*_{ii})^m$.

■

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

Legyenek adva az X_n valószínűségi változók, és most is $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$. A $\tau : \Omega \to \mathbb{N}$ valószínűségi változó (véges) *megállási idő*, ha $\{\tau \leq n\} \in \mathcal{F}_n$ 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:

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

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, $k \geq 0$-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.

4.3. Tétel. *Az $\{X_n\}_{n\geq 0}$ 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:

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

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

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

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

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

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

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

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. ∎

4.4. 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.

4.5. 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. (4.4 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.

4.6. 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 (A Stirling formula szerint $n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$, és ebből $p_{00}^{(2n)} \sim \frac{1}{\sqrt{\pi n}}$). 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

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

5. Az átmenetvalószínűségek konvergenciája

Az eddigiekből az következik, hogy ha j átmeneti állapot, akkor $p_{ij}^{(n)} \to 0$. Ha tehát a Markov lánc már jó ideje tart, akkor elhanyagolható valószínűséggel tartózkodik a tranziens j állapotban, függetlenül attól, hogy melyik i állapotból indult. Vajon rekurrens állapot esetén mit mondhatunk erről a határértékről?

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

Bizonyítás. Vegyük először észre, hogy elég a $d = 1$ esetet igazolni. Ha ugyanis a periódus $d > 1$, akkor az $\{X_0, X_d, X_{2d}, \ldots\}$ láncra áttérve, ott már i aperiodikus állapot lesz, és az átlagos visszatérési idő nyilván az eredetinek d-edrésze.

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

$$d_f \doteq \text{lnko}\{n \geq 1 : f_n > 0\} = 1.$$

Ha ugyanis $p_n > 0$, akkor n előáll $n = n_1 + \ldots + n_k$ alakban, ahol $f_{n_i} > 0$, tehát $d_f | n$. Korábban azonban megmutattuk, hogy minden elég nagy n-re $p_n > 0$, így $d_f = 1$.

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 tudnánk, hogy p_n konvergens, akkor a Nörlund lemma alapján a határértékére azonnal adódna $1/m_i$. Azonban nem tudjuk, hogy a sorozat konvergens, ezért többet kell dolgoznunk a bizonyítással.

Ha $\limsup_{n\to\infty} p_n = \lambda$, akkor válasszunk egy olyan részsorozatot, melyre $\lim_{k\to\infty} p_{n_k} = = \lambda$. Tudjuk, hogy $p_n = \sum_{j=1}^{n} f_j p_{n-j}$. Vegyünk egy olyan s-et, melyre $f_s > 0$, és válasszuk le ezt a tagot a szummából:

$$\lambda = \lim_{k\to\infty} p_{n_k} = \lim_{k\to\infty} (f_s p_{n_k-s} + \sum_{\nu=1,\nu\neq s}^{n_k} f_\nu p_{n_k-\nu}). \tag{3}$$

Ezt szeretnénk felülről becsülni, felhasználva, hogy $\lim(a_n+b_n) \leq \liminf a_n + \limsup b_n$. Tehát a második tag limszupját kell felülről becsülnünk. Tetszőleges $\epsilon > 0$-hoz legyen N olyan nagy, hogy $\sum_{n=N}^{\infty} f_n < \epsilon$. Ekkor

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

Mivel $\lim_{k\to\infty} \sup_{\nu<N} p_{n_k-\nu} \leq \lambda$, ezért (3)-ból:

$$\lambda \leq f_s \liminf_{k\to\infty} p_{n_k-s} + (1-f_s)\lambda.$$

Ezt átrendezve azonnal következik, hogy $\liminf_{k\to\infty} p_{n_k-s} \geq \lambda$, azaz
$\lim_{k\to\infty} p_{n_k-s} = \lambda$. 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 felírható ilyen alakban (ezt a részosztályokról szóló tételnél bizonyítottuk), elmondhatjuk hogy $t \geq t_0$ esetén $\lim_{k\to\infty} p_{n_k-t} = \lambda$. Mármost

$$1 = \sum_{h=0}^{n_k-t_0} r_h p_{n_k-t_0-h}.$$

Egy olyan jellegű összeg jelent meg, mint a Nörlund lemmában, azaz $\sum_{h=0}^{m_k} a_h b_{m_k-h}$, csakhogy most a b_m sorozatról nem tudjuk, hogy konvergens, csak a b_{m_k-h} sorozatokról, minden rögzített h-ra.

Tegyük fel először, hogy $\sum_{h=0}^{\infty} r_h = \infty$, ekkor minden N-re

$$1 \geq \lim_{k\to\infty} \sum_{h=0}^{N} r_h p_{n_k-t_0-h} = \lambda \sum_{h=0}^{N} r_h,$$

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

$$\mid \sum_{h=0}^{n_k-t_0} r_h(p_{n_k-t_0-h} - \lambda) \mid \leq \sum_{h=0}^{N} r_h \mid p_{n_k-t_0-h} - \lambda \mid + \epsilon \leq 2\epsilon.$$

Tehát

$$1 - \lambda \sum_{h=0}^{n_k-t_0} r_h \to 0,$$

amiből $\lambda = 1/\sum r_h$ következik (és ez a végtelen esetben is érvényes).

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

5.2. 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. ■

5.3. 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.*

Nézzük meg az általános tételünk néhány speciális esetét!

- Ha j tranziens vagy nulla rekurrens, akkor $p_{ij}^{(n)} \to 0$.
- Ha i és j ugyanabban az aperiodikus, pozitív rekurrens osztályban vannak, akkor $p_{ij}^{(n)} \to \frac{1}{m_j}$.
- Ha i és j ugyanabban a d periódusú, pozitív rekurrens osztályban vannak, méghozzá $j \in C_r(i)$, akkor $p_{ij}^{(nd+r)} \to \frac{d}{m_j}$.

5.4. Példa. Az egydimenziós szimmetrikus bolyongás nulla rekurrens.

5.5. Példa. Tegyük fel, hogy egy társasjátékban minden körben annyi mezőt ugrik előre a bábunk, ahányat egy dobókockával dobtunk. Jelölje p_n annak esélyét, hogy rálépünk az n-edik mezőre. Mutassuk meg, hogy $p_n \to 2/7$.

6. Stacionárius eloszlás

6.1. Definíció. *Legyen P egy átmenetmátrix. 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 átmenetvalószínűségű 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).*

6.2. 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: $u_i = K\pi_i$.

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 $\sum_{k \in C_{-r}(i)} u_k$ konstans, és $u_i = K\pi_i$. ∎

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.

6.3. 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*

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

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 $P(X_1 = i) = 0$, hiszen a pozitív osztályokból nem lép ki a lánc. Ha viszont $i \in D_\alpha$, akkor

$$P(X_1 = i) = \sum_{k \in D_\alpha} p_k p_{ki} = \lambda_\alpha \sum_{k \in D_\alpha} \pi_k p_{ki} = \lambda_\alpha \pi_i = p_i,$$

hiszen π_i stacionárius eloszlás D_α-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.)

6.4. Példa. Vegyünk egy bolyongást a nemnegatív számokon, ahol a nullában egy visszaverő fal van. 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. Belátjuk, hogy pozitív rekurrens. Megoldhatók ugyanis a stacionárius eloszlás egyenletei (vagy: használjuk a példa utáni észrevételt $A = \{0,1,\ldots,i\}$ választással), és kapjuk, hogy

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

A stacionárius eloszlás egyenletrendszerének megoldását néha egyszerűsíti a következő észrevétel:

6.5. Tétel. *Ha π stacionárius eloszlás, akkor minden $A \subset I$ esetén*

$$\sum_{i\in A, j\notin A} \pi_i p_{ij} = \sum_{i\in A, j\notin A} \pi_j p_{ji}.$$

Bizonyítás. Legyen

$$\mathrm{flux}(A,B) = \sum_{i\in A, j\in B} \pi_i p_{ij} = P_\pi(X_0 \in A, X_1 \in B).$$

Először is

$$\mathrm{flux}(\{i\}, I) = \pi_i = \sum_{j\in I} \pi_j p_{ji} = \mathrm{flux}(I, \{i\}).$$

Mivel a fluxus nyilvánvalóan additív a diszjunkt unióra mindkét argumentumára, ebből következik, hogy minden $A \subset I$ esetén $\mathrm{flux}(A, I) = \mathrm{flux}(I, A)$. Ebből

$$\mathrm{flux}(A,A) + \mathrm{flux}(A,A^c) = \mathrm{flux}(A,I) = \mathrm{flux}(I,A) = \mathrm{flux}(A,A) + \mathrm{flux}(A^c,A),$$

azaz $\text{flux}(A, A^c) = \text{flux}(A^c, A)$, amit bizonyítani kellett. ■

6.6. Példa. (Ehrenfest diffúziós modell) Képzeljünk el egy tartályt, benne N darab molekulát. A tartályt képzeletben két egyenlő részre osztjuk, és azt vizsgáljuk, hogy hány molekula van a két félben. A diffúziót úgy modellezzük, hogy minden lépésben egy véletlenszerűen választott molekula átmegy a másik felébe a tartálynak. Jelölje X_n, hogy n lépés után hány molekula van a tartály első felében. Ez nyilván Markov lánc a $\{0, 1, \ldots, N\}$ állapottéren, és

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

Legyen $A = \{0, 1, \ldots, i-1\}$, ekkor

$$\text{flux}(A, A^c) = \pi_{i-1} \frac{N-i+1}{N} = \pi_i \frac{i}{N} = \text{flux}(A^c, A),$$

ami a $\pi_i = \frac{N-i+1}{i} \pi_{i-1}$ rekurziót adja. Ennek megoldása éppen az N rendű, 1/2 paraméterű binomiális eloszlás.

6.7. 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 $\mu_k = \frac{1-p^k}{qp^k}$.

7. Pozitív rekurrens Markov láncok

Az irreducibilis és pozitív rekurrens Markov láncok hosszú távú viselkedését különösen egyszerű leírni (ha még aperiodikusak is, akkor ez még inkább igaz).

7.1. Nagy számok törvénye

7.1. Tétel. *Legyen* $\{X_n\}_{n\geq 0}$ *irreducibilis és pozitív rekurrens Markov lánc* π *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*

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

Megjegyzés: A $Z_n = f(X_n)$ sorozat nem feltétlenül Markov lánc. Tekintsük például az $I = \{0,1,2\}$ állapotteret, ezen a bolyongást ($p_{01} = p_{21} = 1, p_{10} = p_{12} = 1/2$). Legyen a kezdeti eloszlás $P(X_0 = 0) = P(X_0 = 1) = 1/2$. Definiáljuk az f függvényt úgy, hogy $f(0) = 0, f(1) = f(2) = 1$, és legyen $Z_n = f(X_n)$. Ekkor könnyen adódik, hogy

$$1/2 = P(Z_2 = 0|Z_1 = 1, Z_0 = 0) \neq P(Z_2 = 0|Z_1 = 1, Z_0 = 1) = 0,$$

azaz Z_n nem Markov lánc.

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

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

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

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

A Z_k-k összegére a következő felbontási formulát használjuk:

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

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

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

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, $\ell(n)$ 1 valószínűséggel végtelenhez tart. Ezért a nagy számok erős törvénye szerint

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

Másrészt $Y'/\ell(n) \to 0$ (1 valószínűséggel), tehát $\frac{1}{\ell(n)}\sum_{k=0}^{n} Z_k$ 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:

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

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

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

azaz

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

■

7.2. Példa. Ha f olyan függvény, mely egy állapotra 1, a többire nulla (pl. $f(i) = 1$ és $f(j) = 0$, ha $j \neq i$), akkor azt kaptuk, hogy az i állapot relatív gyakorisága az első n lépésből tart π_i-hez.

Ezt általánosíthatjuk állapotsorozatokra is: legyen $\underline{i} = (i_1, \ldots, i_k)$ egy k hosszú állapotsorozat. Ekkor az $\underline{i}$ sorozat relatív gyakorisága a $\pi_{i_1} \cdot p_{i_1 i_2} \cdot p_{i_2 i_3} \cdots p_{i_{k-1} i_k}$ szorzathoz tart. (Ehhez készítenünk kell egy új ML-ot, melynek állapotai a k hosszú sorozatok, ez is ergodikus lesz, tehát alkalmazható rá a tétel).

7.2. Centrális határeloszlás-tétel

Ugyanabban a felállásban, mint az előbb, megmutatható, hogy a részletösszegeket alkalmasan normálva, normális határeloszlást kapunk.

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

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

Vegyük észre, hogy $V_h = \sum_{s=\tau_h}^{\tau_{h+1}-1} g(X_s)$, ahol $g(j) = f(j) - \mathcal{I}(f)$, és $E(V_h) = 0$.

A bizonyításhoz két lemmára lesz szükségünk.

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

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

$$P(n-\tau_{\ell(n)} \geq t) = \sum_{s=t}^{n} P(n-\tau_{\ell(n)} = s) =$$

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

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

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

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

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

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!

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

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

7.7. 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

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

Azt már tudjuk, hogy $Y'/\sqrt{n}$ sztochasztikusan 0-hoz tart. A 7.4 lemma szerint ugyanez igaz az $(n -$ $- \tau_{\ell(n)})/\sqrt{n}$ tagra, a 7.5 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

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

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 $\ell(n)-1$-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

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

Ezek bővülő események, és

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

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

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

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

7.3. Tabu állapotok

Vajon hogyan lehet az V_h szórásnégyzetét 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. Ezt általánosítjuk most, tehát úgy számolunk ki átmenetvalószínűségeket, hogy előírjuk, hogy bizonyos állapotokban nem járhat a lánc.

7.8. Definíció. *(Átmenetvalószínűségek tabu állapotokkal.) Legyen $H \subseteq I$ tetszőleges. Az i-ből j-be menő n lépéses átmenetvalószínűség a H tabuhalmazzal:*

$$_Hp_{ij}^{(n)} = P(X_n = j, X_m \notin H\ 0 < m < n | X_0 = i)\ \text{ha}\ n \geq 1.$$

Jelölje $_Hf_{ij}^{(n)} = {}_{j,H}p_{ij}^{(n)}$. 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). Legyen $m_{ij} = \sum_{n=1}^\infty nf_{ij}^{(n)}$ az átlagos elérési idő, általában pedig $_Hm_{ij} = \sum_{n=1}^\infty n_Hf_{ij}^{(n)}$.

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

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

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

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. ■

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

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

továbbá

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

Bizonyítás. Az $(1A)$ formula szerint (k első elérése szerinti felbontás)

$$ f_{ij}^{(n)} = {}_jp_{ij}^{(n)} = {}_{k,j}p_{ij}^{(n)} + \sum_{s=1}^{n-1} {}_{k,j}p_{ik}^{(s)} \cdot {}_jp_{kj}^{(n-s)} = {}_kf_{ij}^{(n)} + \sum_{s=1}^{n-1} {}_jf_{ik}^{(s)} \cdot f_{kj}^{(n-s)}. $$

Beszorozva n-nel, majd összegezve egytől végtelenig

$$ \begin{aligned} m_{ij} = {}_km_{ij} + \sum_{n=1}^{\infty} n \sum_{s=1}^{n-1} {}_jf_{ik}^{(s)} \cdot f_{kj}^{(n-s)} = \\ = {}_km_{ij} + \sum_{s=1}^{\infty} {}_jf_{ik}^{(s)} \sum_{n=s+1}^{\infty} ((n-s)f_{kj}^{(n-s)} + sf_{kj}^{(n-s)}) = {}_km_{ij} + m_{kj} \cdot {}_jf_{ik}^* + {}_jm_{ik}, \end{aligned} $$

ahol az utolsó lépésben felhasználtuk, hogy $f_{kj}^* = 1$. Megcserélve j és k szerepét, majd kivonva a kettőt egymásból kapjuk hogy

$$ m_{ik} + m_{kj} - m_{ij} = m_{kj} + m_{jk} \cdot {}_kf_{ij}^* - m_{kj} \cdot {}_jf_{ik}^* = m_{jk} \cdot {}_kf_{ij}^* + m_{kj}(1 - {}_jf_{ik}^*) = {}_kf_{ij}^*(m_{jk} + m_{kj}), $$

mivel ${}_kf_{ij}^* + {}_jf_{ik}^* = 1$. Ezzel megkaptuk az első bizonyítandó formulát. Ebből speciális esetként kapjuk, hogy

$$ {}_jf_{jk}^* = \frac{m_{jj}}{m_{jk} + m_{kj}}. $$

Másrészt megadhatjuk, hogy i-ből indulva, mi lesz a j-ben tett látogatások számának eloszlása, mielőtt a lánc beér k-ba. Ha ugyanis a lánc előbb ér k-ba, mint j-be (ennek valószínűsége ${}_jf^*_{ik}$), akkor ez a szám 0, ha viszont előbb ér j-be, mint k-ba (ennek valószínűsége ${}_kf^*_{ij}$), akkor a j-ben tett látogatások száma geometriai eloszlású ${}_jf^*_{jk}$ paraméterrel. Ezért a várható érték:

$$ {}_kp^*_{ij} = {}_jf^*_{ik} \cdot 0 + {}_kf^*_{ij} \cdot \frac{1}{{}_jf^*_{jk}} = \frac{m_{ik} + m_{kj} - m_{ij}}{m_{jj}}. $$

Máshogy, formálisan bizonyítva, használjuk az összegre vonatkozó formulákat. Az $(1B)$ formula szerint

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

A $(2B)$ formulából pedig

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

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 ${}_ip^*_{ij} = m_{ii}/m_{jj}$. Konkrét (véges állapotterű) Markov láncra az m_{ij} mennyiségek egy lineáris egyenletrendszer megoldásával kaphatók meg.

7.4. 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 $g = f - \mathcal{I}(f)$ függványre alkalmazzuk. Az egyszerűség kedvéért tegyük most is fel, hogy $X_0 = i$. Ekkor

$$E\left[\left(\sum_{n=0}^{\tau_1-1} g(X_n)\right)^2\right] = E\left[\left(\sum_{n=0}^{\infty} g(X_n)\chi\{n<\tau_1\}\right)^2\right] = E\left[\left(\sum_{n=1}^{\infty} g(X_n)\chi\{n\le\tau_1\}\right)^2\right] =$$
$$E\left[\sum_{n=1}^{\infty} g^2(X_n)\chi\{n\le\tau_1\} + 2\sum_{n<m} g(X_n)g(X_m)\chi\{m\le\tau_1\}\right]$$

Vegyünk tagonként várható értéket, és írjuk tovább:

$$\sum_{j\in I} g^2(j)\sum_{n=1}^{\infty} {}_ip_{ij}^{(n)} + 2\sum_{j\in I, j\neq i}\sum_{l\in I} g(j)g(l)\sum_{n<m} {}_ip_{ij}^{(n)}\,{}_ip_{jl}^{(m-n)} =$$
$$\sum_{j\in I} g^2(j)\,{}_ip_{ij}^{*} + 2\sum_{j\in I, j\neq i}\sum_{l\in I} g(j)g(l)\,{}_ip_{ij}^{*}\,{}_ip_{jl}^{*} =$$
$$\sum_{j\in I} g^2(j)\frac{\pi_j}{\pi_i} + 2\sum_{j\in I, j\neq i}\sum_{l\in I} g(j)g(l)\frac{\pi_j}{\pi_i}\pi_l(m_{ji}+m_{il}-m_{jl}), \quad (4)$$

a 7.10. 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 $\mathcal{I}(g) = 0$, és

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

A most kiszámolt képlet segítségével az i-be való visszatérési idő második momentumát is megkaphatjuk. Ehhez válasszuk a $g = 1$ függvényt, a (4) egyenlet középső sorából:

$$E(\tau^2) = m_{ii} + 2 \sum_{j\in I, j\neq i} \sum_{l\in I} \frac{m_{ii}}{m_{jj}} \cdot {}_i p^*_{jl} =$$

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

(Felhasználtuk, hogy ${}_i p^*_{ij} = m_{ii}/m_{jj}$, $\sum_j {}_i p^*_{ij} = m_{ii}$, és $\sum_l {}_i p^*_{jl} = m_{ji}$.)

8. Reguláris mérték

A (nemnegatív) reguláris mértékek a stacionárius eloszlások általánosításai. Eleget tesznek a stacionárius eloszlásra vonatkozó egyenletrendszernek, de nem követeljük meg, hogy végesek legyenek. Az egyszerűség kedvéért tegyük fel mostantól, hogy a Markov lánc irreducibilis.

8.1. Definíció. *Az $u = (u_i)_{i\in I}$ nemnegatív elemű sorozat reguláris mérték, ha*

$$u_i = \sum_{j\in I} u_j p_{ji} \quad \forall i \in I.$$

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.

8.2. Lemma. *Legyenek i, j, k, l egy irreducibilis, rekurrens Markov lánc tetszőleges állapotai. Ekkor*

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

Megjegyzések:

- Megmutatható, 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 $(1/N)\sum_0^N p^{(n)}_{ij} \to 1/m_j$.

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

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

Mivel most a Markov lánc rekurrens, $f_{ij}^* = 1$, továbbá $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$, tehát mindegy, hogy az összegzéseket $n = 1$-től vagy $n = 0$-tól kezdjük. Ezért

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

erről kell belátni, hogy e_{lj}-vel egyenlő.

A kulcslépés tehát 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!

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

mivel a rekurrencia miatt $p_{ll}^* = \infty$. ■

8.3. Tétel. *Irreducibilis, rekurrens Markov láncban a nemnegatív reguláris mértékek általános alakja: $u_i = c e_{ji}$, ahol j tetszőleges rögzített állapot.*

Bizonyítás. Azt, hogy $u_i = c e_{ji}$ reguláris mérték, azaz megoldása az egyenletrendszernek, már láttuk a nagy számok törvényének bizonyításánál..

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

$$q_{ij}^{(n)} = \frac{u_j}{u_i} p_{ji}^{(n)}.$$

Megmutatjuk, hogy ezek egy 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 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 $u_i = u_j e_{ji}$. ∎

Megjegyzések:

– Az előző tétel bizonyításánál láttuk, hogy

$$\lim_{N\to\infty} \frac{\sum_{n=0}^N p_{ii}^{(n)}}{\sum_{n=0}^N p_{jj}^{(n)}} = e_{ji},$$

ezt a $p_{kk}^{(n)}$ valószínűségek összegével bővítve kapjuk, hogy $e_{ji} = e_{ki} e_{jk} = e_{jk} e_{ki}$. Ebből a multiplikatív tulajdonságból következik, hogy a különböző j választásokkal kapott mértékek csak konstans szorzóban térnek el egymástól.

– Rekurrens Markov láncban $\sum_i e_{ji} = m_{jj}$, attól függően véges vagy végtelen, hogy a lánc pozitív vagy nulla rekurrens. Pozitív rekurrens esetben tehát nem kapunk új megoldást (itt $e_{ji} = \pi_i/\pi_j$).

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

8.4. 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.

8.1. Megfordított láncok

Legyen P tetszőleges átmenetmátrix, és $u_i > 0$ $(i \in I)$ 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-t a P u szerinti megfordításának nevezzük. 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_i mérték a Q által meghatározott Markov láncra is reguláris. Vegyük még észre, hogy ha most Q-t megfordítjuk u szerint, akkor visszakapjuk P-t.

A megfordításnak pozitív rekurrens esetben a következő jelentése van. Legyen P irreducibilis, pozitív rekurrens Markov lánc átmenetmátrixa, ekkor $u = \pi$ a stacionárius eloszlás, legyen ez a kezdeti eloszlás: $P(X_0 = i) = u_i$. 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.) Sorsoljuk most ki az X_0 értékét az u_i eloszlás szerint. Ezután egymástól függetlenül indítsunk el egy-egy Markov láncot előre P szerint és hátra Q 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

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

Azt mondjuk, hogy az X_n Markov lánc megfordítható (u-ra nézve), ha megfordítása önmaga. Pozitív rekurrens Markov lánc akkor és csak akkor megfordítható, ha stacionáius eloszlására

$$\pi p_{ij} = \pi_j p_{ji} \quad \forall i, j.$$

Korábban láttuk, hogy például az Ehrenfest diffúziós modell ilyen.

9. Véges állapotterű Markov láncok

Ebben a szakaszban véges állapotterű Markov láncokkal foglalkozunk. Ilyenkor az átmenetmátrix egy véges, négyzetes sztochasztikus mátrix.

9.1. Állítás. *Legyen $\{X_n\}_{n\geq 0}$ véges állapotterű Markov lánc. Ekkor (i) minden rekurrens osztály pozitív, és (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

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

■

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

9.1. Perron-Frobenius tételkör

A négyzetes sztochasztikus mátrixok vizsgálhatók algebrai eszközökkel. Kezdjük általánosabban: a továbbiakban A nemnegatív elemű, négyzetes mátrix. A mátrix legnagyobb sajátértékét, és a hozzá tartozó sajátvektorokat fogjuk elemezni. 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

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

$$\lambda_0 = \lambda_0[A] = \sup\{\lambda : \ \exists x > 0 : \ Ax \geq \lambda x\}. \tag{5}$$

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

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

Bizonyítás. Legyen egyrészt $x > 0$, λ (5)-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 (5)-beli, hiszen minden j-re

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

■

9.4. 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 λ_0-nál.*
3. *Ha A aperiodikus is (azaz van olyan m, melyre $A^m \gg 0$), akkor Minden más sajátérték abszolút értéke kisebb λ_0-nál.*

Bizonyítás. A bizonzítást csak arra az esetre végezzük el, ha $A \gg 0$.

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 (5)-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 $|\lambda| \leq \lambda_0$.

3. 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 $\lambda = \lambda_0$.

■

Egy Markov lánc éppen akkor irreducibilis, ha a P átmenetmátrix irreducibilis. Erre $\lambda_0[P] = 1$. Azt is tudjuk, hogy melyik szigorúan pozitív sajátvektor tartozik hozzá: éppen az **1**. Érdekesebbek nekünk az 1 sajátértékhez tartozó baloldali sajátvektorok, melyek épp a stacionárius eloszlások. Mivel egy mátrixnak és transzponáltjának megegyeznek a sajátértékei, a szintán irreducibilis P^T mátrixra $\lambda_0[P^T] = 1$. Irreducibilis Markov láncokra tehát az előző tételt úgy fordíthatjuk le (P^T-ra alkalmazva az állítást), hogy egyértelműen létezik stacionárius eloszlásuk, mely szigorúan pozitív.

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

9.5. 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

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

9.6. 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 $x^0 \gg 0$, 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 második állítást már korábban bizonyítottuk. A harmadik 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

$$(T^n)_{ij} \leq \frac{\max x_k^0}{\min x_k^0}.$$

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 $T_0 x = \lim_{n\to\infty} T_n x$ határérték.

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

Minden x-re $\{T_n x : \ n \geq 1\}$ korlátos halmaz, azaz kiválasztható belőle konvergens részsorozat: $T_{n'} x \to x_0$. Be kell látni, hogy $T_n x \to x_0$. 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 $T_n x_0$ konvergens, azt látjuk be, hogy x_0 fixpontja T-nek, és így T_n-nek is. Ugyanis

$$T x_0 = \lim T T_{n'} x = \lim(T T_{n'} x - T_{n'} x) + \lim T_{n'} x = \lim(T_{n'}(T - I)x) + x_0 = x_0.$$

Emiatt $T_n x$ is konvergál, és $\lim T_n x = x_0$. $\blacksquare$

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 harmadik 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.

9.2. 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.

9.7. 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.

9.8. 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.

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

Az előző tétel szerint kész vagyunk. ■

10. 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.

10.1. 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

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

10.1. 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 elfogadási valószínűségekkel 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.)

10.2. 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. Tehát $\alpha_{st} = \min(|N(s)|/|N(t)|, 1)$, ha $t \in N(s)$.

10.2. 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 $f(x) = f(x_1, \dots, x_n) = cb(x)$ 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.

10.3. 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!

11. Infinitezimális generátor

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

11.1. 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. ■

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

11.3. 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

$$p_{ij}(t)=\lim_{n\to\infty}p_{ij}(t+t_n)=\lim_{n\to\infty}\sum_k p_{ik}(t)p_{kj}(t_n)=\sum_k p_{ik}(t)u_{kj}. \tag{6}$$

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 (6)-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 (6)-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. ■

11.4. 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.*

11.5. 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. ■

11.6. 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 $0 \leq \delta < t$.

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

Tartson t nullához:

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

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 $\log(1+h) = h(1+o(1))$,

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

■

11.7. 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

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

Továbbá

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

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

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

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

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

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.

11.8. 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:

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

■

12. Kolmogorov-féle differenciálegyenletek

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

12.1. 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 11.5 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.

12.2. 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!

13. Születési-halálozási folyamatok

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

13.1. 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 $i < j$,

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

13.2. 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

$$r_n(t) = \lambda_{n-1} e^{-\lambda_n t} \int_0^t e^{\lambda_n x} r_{n-1}(x)dx. \tag{7}$$

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

13.1. 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). ■

13.2. 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 (7) 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?

13.3. 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.

13.3. 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.

13.4. 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

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

14. Visszatérőség

14.1. 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.

14.1. 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.

14.2. 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 ($p_{ii}(h) > 0$).

14.2. 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.

14.3. 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.

14.4. 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

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

14.5. 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 $Z_0, Z_1, \ldots$ 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.

14.3. Visszatérőség

14.6. 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

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

(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

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

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

14.7. 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 $\int_0^\infty p_{ii}(t)dt = \infty$.

14.4. Stacionárius eloszlás

14.8. 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 11.5 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

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

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

14.9. 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 $\tilde{S} = \infty$.

14.10. 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 $\text{Exp}(\mu)$. 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 $\frac{\lambda}{\mu-\lambda}$.

14.11. 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 $\text{Poisson}(\lambda/\mu)$.

15. A Kolmogorov-egyenletek megoldhatóságáról

Legyen Q konzervatív. Ebben a szakaszban megkonstruáljuk a Kolmogorov-féle egyenletek $\overline{P}(t)$ *minimális megoldását.* Ez olyan standard szubsztochasztikus átmenet-mátrix-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

$$p_{ij}^{\langle n+1 \rangle}(t) = \sum_{k \neq i} \int_0^t e^{-q_i (t-s)} q_{ik} p_{kj}^{\langle n \rangle}(s) ds, \tag{8}$$

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:

$$\overline{p}_{ij}(t) = \delta_{ij} e^{-q_i t} + \sum_{k \neq i} \int_0^t e^{-q_i (t-s)} q_{ik} \overline{p}_{kj}(s) ds. \tag{9}$$

15.1. 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:

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

Ezt felhasználva,

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

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

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

■

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

Bizonyítás. A $\overline{P}(t)$ családra vonatkozó (9) integrálegyenletet deriváljuk le! Emlékeztető: ha

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

akkor

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

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

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

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

15.3. 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:

$$\overline{p}_{ij}(t) = \delta_{ij} e^{-q_i t} + \sum_{k \neq j} \int_0^t \overline{p}_{ik}(s) q_{kj} e^{-q_j(t-s)} ds. \tag{10}$$

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

15.4. 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

$$p_{ij}^{\langle n \rangle}(s+t) = \sum_{m=0}^{n} \sum_k p_{ik}^{\langle m \rangle}(s) p_{kj}^{\langle n-m \rangle}(t). \tag{11}$$

Vagyis aszerint bontjuk fel a baloldali valószínűséghez tartozó eseményt, hogy az s időpontban melyik állapotban van a lánc, és addig hányat ugrott. Ha a (11) 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 (11) teljesül $n = 0$-ra. Az indukciós lépés pedig:

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

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

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

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

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

Ezért

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

■

15.5. 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

$$p'_{ij}(t) \geq -q_i p_{ij}(t) + \sum_{k \neq i} q_{ik} p_{kj}(t). \tag{12}$$

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 (12) 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

$$p_{ij}(t) \geq e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t e^{-q_i (t-s)} q_{ik} p_{kj}(s) ds \tag{13}$$

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 (13) 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 (8) összefüggést használtuk. ∎