

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Felsmann Dániel

ÁLTALÁNOS ÁLLAPOTTERŰ MARKOV LÁNCOK STABILITÁSA

MSc diplomamunka

Témavezető:

Dr. Gerencsér Balázs

egyetemi adjunktus

Valószínűségelméleti és Statisztika Tanszék



2022 december

Tartalomjegyzék

1. Bevezetés	2
2. Alapfogalmak	4
2.1. Markov kernelek	5
2.2. A kanonikus Markov lánc	6
2.3. Markov láncok stabilitásának alapfogalmai	8
2.3.1. φ -irreducibilitás	8
2.3.2. Tranziens és rekurrens Markov láncok	10
2.3.3. Periódus	11
2.3.4. Invariáns mértékek	11
3. Általános állapotterű Markov láncok ergodikus tulajdonságai	13
3.1. A geometriai ergodikus tétel	14
3.2. Ergodikus Markov láncok, mint ergodikus dinamikai rendszerek	20
4. A Centrális határeloszlás-tétel	26
4.1. Keverő folyamatok	26
4.2. Keverő Markov láncok	28
4.3. Centrális határeloszlás-tétel Markov láncokra	31
5. Markov láncok véletlen környezetben	35
5.1. A hosszú távú kontraktivitási feltétel vizsgálata	37
5.1.1. Keverés és a hosszú távú kontraktivitási feltétel	37
5.2. Példák Markov láncokra véletlen környezetben	39
5.2.1. Bilineáris rendszer véletlen környezetben	39
5.2.2. SETAR modell véletlen környezetben	41
Kiemelt jelölések	44

1. fejezet

Bevezetés

A Markov láncok elmélete régóta ismert és széleskörűen alkalmazott területe a matematikának. Sok más sztochasztikus folyamathoz hasonlóan az elmélet egyik legfontosabb kérdése az, hogy mit tudunk mondani Markov láncok hosszú távú viselkedéséről. Milyen feltételek mellett konvergálnak az eloszlások az idő előrehaladtával valamilyen adott eloszláshoz? Ha konvergálnak, akkor hogyan becsülhető meg a konvergencia sebessége? Igazak-e a nagy számok törvénye és a centrális határeloszlás-tétel? Az, hogy a halmaz, amelyben a Markov lánc az értékeit felveszi megszámlálható-e vagy sem, meghatározó abból a szempontból, hogy ezeket a kérdéseket milyen eszközökkel válaszolhatjuk meg. Ha a Markov lánc állapottere megszámlálható, akkor a hosszú távú viselkedés kombinatorikus és lineáris algebrai eszközökkel is vizsgálható, míg nem megszámlálható állapottér esetén erre nincs lehetőség. A dolgozatban célunk, hogy betekintést adjunk az általános állapotterű Markov láncok hosszú távú viselkedésének elméletébe. A teljesség igénye nélkül áttekintjük a legfontosabb fogalmakat és eredményeket, amelyek közül többet részletesen be is bizonyítunk.

Először Revuz [10] felépítését követve bevezetjük azokat az alapfogalmakat, amelyek a hosszú távú viselkedés vizsgálatának legfontosabb eszközei közé tartoznak. Ezután Meyn és Tweedie [9] alapján bemutatjuk a megszámlálható állapottér esetéből ismert olyan fogalmak általánosításait, mint az irreducibilitás, a periódus, illetve a tranziens és rekurrens tulajdonságok. A harmadik fejezetben precízen értelmezzük a Markov lánc eloszlásainak konvergenciáját. Hairer és Mattingly [5] gondolatmenetét követve bebizonyítjuk a geometriai ergodikus tételt, amely az eloszlások konvergenciájának sebességére is használható becslést szolgáltat. A fejezet zárásaként Douc et al. [2] segítségével összekapcsoljuk Markov láncok és dinamikai rendszerek ergodikusság-fogalmait és elégséges feltételeket adunk a nagy számok törvényének teljesülésére. A negyedik fejezetben a centrális határeloszlás-tételt tárgyaljuk. Először ismertetjük sztochasztikus folyamatok keverésének fogalmát, majd megmutatjuk, hogy ez a fogalom hogyan használható Markov láncokkal kapcsolatban. A fejezet végén a keverő folyamatokkal kapcsolatos centrális határeloszlás-tételeket alkalmazzuk Markov láncokra. Az ötödik fejezetben azt vizsgáljuk meg, hogy mit tudunk mondani a hosszú távú

viselkedésről, ha a Markov lánc időben nem homogén, hanem az átmenet-valószínűségeket egy másik sztochasztikus folyamat határozza meg. Ezzel kapcsolatban bemutatjuk Lovas és Rásonyi [8] eredményeit, majd a szerzők egyik feltételét egy saját példán keresztül is megvizsgáljuk. Végül pedig mutatunk két olyan modellt, amelyben Lovas és Rásonyi eredményei alkalmazhatók.

A dolgozatban használt nem sztenderd jelöléseket igyekeztünk minden esetben az első előfordulásukkor bevezetni. Ezeket a jelöléseket a dolgozat végén külön fejezetben foglaltuk össze.

2. fejezet

Alapfogalmak

A dolgozatban általános állapotterű Markov láncok ergodikus tulajdonságaival foglalkozunk. Ahhoz, hogy ezt megtehessük, áttekintjük azokat az alapvető fogalmakat, amelyekre az elmélet épül. Először a Markov tulajdonság segítségével definiáljuk az általános állapotterű Markov láncokat. Ezután megmutatjuk, hogy a homogén Markov láncok hogyan jellemezhetők sztochasztikus kernelekkel, és bevezetjük a kanonikus Markov lánc fogalmát. A fejezet második felében a megszámlálható állapotterű Markov láncokkal kapcsolatos olyan fogalmak általánosítását mutatjuk be, mint az irreducibilitás, a tranzienst és rekurrens tulajdonságok és a periódus, amelyek kulcsfontosságúak a Markov láncok hosszú távú viselkedésének szempontjából.

A fejezetben végig legyen adott egy $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mező és egy $(\mathcal{X}, \mathcal{B})$ mérhető tér. Az $X = (X_n)_{n \in \mathbb{N}}$ $\mathcal{X}$ -értékű sztochasztikus folyamaton az $X_n : \Omega \rightarrow \mathcal{X}$ mérhető függvényeket értjük. Legyen $\mathcal{F}_n := \sigma(X_k, k \leq n)$.

2.1. Definíció. Azt mondjuk, hogy az $(X_n)_{n \in \mathbb{N}}$ sztochasztikus folyamatra teljesül a Markov tulajdonság, ha minden $m < n$ és minden $A \in \mathcal{B}$ estén

$$\mathbb{P}[X_n \in A \mid \mathcal{F}_m] = \mathbb{P}[X_n \in A \mid X_m].$$

Az $\mathcal{X}$ -értékű X sztochasztikus folyamat Markov lánc, ha teljesül rá a Markov tulajdonság.

Az $\mathcal{X}$ halmazzá a folyamat állapotterének nevezzük. Az állapotter és a $\mathcal{B}$ σ -algebra tulajdonságai meghatározzák, hogy milyen eszközökkel vizsgálhatjuk a Markov láncokat, ezért az állapottereket két típusba soroljuk.

2.2. Definíció. Az $\mathcal{X}$ állapotter

- (i) megszámlálható vagy diszkrét, ha $\mathcal{X}$ megszámlálható, $\mathcal{B}$ pedig az $\mathcal{X}$ összes részhalmazából álló σ -algebra,
- (ii) általános, ha $\mathcal{B}$ megszámlálhatóan generált.

2.1. Markov kernelek

Megszámlálható állapotter esetén a Markov láncot leíró legfontosabb fogalom az átmenet-valószínűségek mátrixa. Általános állapotterű Markov láncok jellemzésekor az átmenetmátrix szerepét a sztochasztikus kernelek veszik át.

2.3. Definíció. Egy $P : \mathcal{X} \times \mathcal{B} \rightarrow [0, 1]$ függvényt sztochasztikus kernelnek vagy Markov kernelnek nevezünk, ha teljesül rá a következő két tulajdonság:

- (i) $P(x, \cdot)$ valószínűségi mérték $\mathcal{B}$ -n minden $x \in \mathcal{X}$ esetén,
- (ii) $P(\cdot, A)$ mérhető függvény minden $A \in \mathcal{B}$ esetén.

Hasonlóan az átmenetmátrixok esetéhez $P(x, A)$ szemléletesen annak a valószínűségét jelenti, hogy a lánc a következő lépésben az A halmazban tartózkodik, ha jelenleg x -ben van. Mielőtt rátérnénk arra, hogyan írhatók le sztochasztikus kernelekkel az általános állapotterű Markov láncok, előbb a Markov kernelek számunkra legfontosabb tulajdonságait foglaljuk össze.

A P Markov kernel az $\mathcal{X}_+ := \{f : \mathcal{X} \rightarrow [0, \infty] : f \text{ mérhető}\}$ függvénytérre ható operátorként is értelmezhető. Ha $f \in \mathcal{X}_+$, akkor a $Pf \in \mathcal{X}_+$ függvényt úgy definiáljuk, hogy

$$[Pf](x) := \int_{\mathcal{X}} f(y)P(x, dy). \quad (2.1)$$

Természetesen a Pf hozzárendelés minden olyan f esetén is értelmezhető, amikor az $\int_{\mathcal{X}} f(y)P(x, dy)$ integrál értelmes minden x -re.

A P Markov kernelre az $\mathcal{M}(\mathcal{X})$ mértékek terén értelmezett operátorként is tekinthetünk. Legyen $\mu \in \mathcal{M}(\mathcal{X})$. A μP mértéket úgy definiáljuk, hogy

$$[\mu P](A) := \int_{\mathcal{X}} P(x, A)\mu(dx). \quad (2.2)$$

Az operátorként értelmezett Markov kernelek egy hasznos tulajdonságát mutatja a P operátor „önadjungáltsága” kifejező összefüggés, amely szerint

$$(\mu P)f = \mu(Pf). \quad (2.3)$$

Ez az összefüggés természetesen nem jelenti azt, hogy P önadjungált lenne, hiszen μ és f különböző terekből származnak.

Markov kerneleket össze is szorozhatunk egymással.

2.4. Definíció. A P és Q Markov kernelek szorzatán a

$$PQ(x, A) := \int_{\mathcal{X}} P(x, dy)Q(y, A)$$

módon definiált Markov kernelt értjük.

A fenti integrálban a szokásostól eltérően az integráló mérték a bal oldalon, míg az integrálandó függvény a jobb oldalon szerepel. Ezt a jelölést a továbbiakban sokszor fogjuk használni, mivel nagyon leegyszerűsíti a többszörös integrálok felírását.

A 2.4 definíció alapján iteratívan értelmezhetjük egy Markov kernel n -edik hatványát is:

$$\begin{aligned} P^n(x, A) &= \int_{\mathcal{X}} P(x, dy)P^{n-1}(y, A) = \\ &= \int_{\mathcal{X}} P(x, dy_1) \int_{\mathcal{X}} P(y_1, dy_2) \cdots \int_{\mathcal{X}} P(y_{n-2}, dy_{n-1})P(y_{n-1}, A). \end{aligned}$$

Ebből a felírásból látszik, hogy a P Markov kernel hatványira igaz a Chapman-Kolmogorov egyenlőség, amelyet tömören úgy fejezhetünk ki, hogy tetszőleges $k \leq n$ esetén

$$P^n = P^k P^{n-k}.$$

A P kernel nulladik hatványát úgy értelmezzük, hogy a fenti azonosság igaz maradjon. Ezzel összhangban legyen $P^0(x, A) = \delta_x(A)$, ahol δ_x az x pontra koncentrált Dirac mérték.

2.2. A kanonikus Markov lánc

Megszámlálható állapotterű homogén Markov láncokkal kapcsolatban ismert, hogy az $X = (X_n)_{n \in \mathbb{N}}$ Markov lánc eloszlását meghatározza X_0 eloszlása és a P átmenetmátrix. Revuz [10] könyvének 1.2 fejezete alapján bemutatjuk a kanonikus Markov lánc konstrukcióját. A konstrukció egyik fontos következménye, hogy hasonló állítás igaz általános állapotterű Markov láncokra is. A konstrukciónak köszönhetően X_t eloszlása kifejezhető lesz X_0 eloszlásával és egy P Markov kernellel.

X_0 eloszlását a kitüntetett szerepe miatt a továbbiakban kezdeti eloszlásnak hívjuk, és μ -vel jelöljük. Formálisan a

$$\mu(A) := \mathbb{P}[X_0 \in A]$$

mértéket az X Markov lánc kezdeti eloszlásának nevezzük.

Egy diszkrét állapotterű Markov láncot akkor nevezünk homogénnek, ha tetszőleges $m < n$ esetén $\mathbb{P}[X_n = x_j \mid X_m = x_i]$ értéke csak $n - m$ értékétől függ. Ehhez hasonlóan értelmezzük az általános állapotterű homogén Markov láncokat is.

2.5. Definíció. Az $(X_n)_{n \in \mathbb{N}}$ sztochasztikus folyamatot homogén Markov láncnak hívjuk P átmenet-valószínűségekkel, ha létezik egy olyan P Markov kernel, hogy minden $m < n$ és minden $f \in \mathcal{X}_+$ nemnegatív, illetve $f \in b\mathcal{X}$ korlátos mérhető függvényre

$$\mathbb{E}[f(X_n) \mid \mathcal{F}_m] = [P^{n-m}f](X_m). \quad (2.4)$$

2.6. Megjegyzés. Ha (2.4)-ben az f függvény egy $A \in \mathcal{B}$ esemény indikátorra, akkor azt kapjuk, hogy

$$\mathbb{P}[X_n \in A \mid \mathcal{F}_m] = P^{n-m}(X_m, A), \quad (2.5)$$

vagyis a fenti definíció valóban a diszkrét állapotterű Markov láncok homogén tulajdonságának általánosítása. Az is megmutatható, hogy a (2.5) összefüggés ekvivalens (2.4)-el.

A továbbiakban Markov láncon mindig homogén Markov láncot értünk.

A kanonikus Markov lánc konstrukciójának segítségével megmutatjuk, hogy tetszőleges adott P Markov-kernelhez, és μ kezdeti eloszláshoz készíthető olyan homogén Markov lánc, amelynek az átmenet-valószínűségei P -vel egyeznek meg. Legyen $\mathcal{X}^{\mathbb{N}}$ az $\mathcal{X}$ elemeiből készített sorozatok halmaza, és $\mathcal{B}^{\otimes \mathbb{N}}$ a szorzat σ -algebra. Az $A \in \mathcal{B}^{\otimes \mathbb{N}}$ halmazt mérhető téglának nevezzük, ha

$$A = A_0 \times A_1 \times \cdots \times A_n \times \mathcal{X} \times \cdots$$

alakú, ahol minden $0 \leq k \leq n$ esetén $A_k \in \mathcal{B}$. A Kolmogorov kiterjesztési tétel miatt minden $x \in \mathcal{X}$ kezdeti feltételhez egyértelműen létezik egy olyan $\mathbb{P}_x$ valószínűségi mérték az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}})$ szorzattéren, amelyre tetszőleges $A = A_0 \times \cdots \times A_n \times \mathcal{X} \times \cdots$ mérhető téglán esetén

$$\mathbb{P}_x[A] = \mathbb{I}_{A_0}(x) \int_{A_1} P(x, dx_1) \int_{A_2} P(x_1, dx_2) \cdots \int_{A_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, A_n).$$

Továbbá az is igaz, hogy tetszőleges $A \in \mathcal{B}^{\otimes \mathbb{N}}$ eseményre az $x \mapsto \mathbb{P}_x[A]$ függvény mérhető. A $\mathbb{P}_x[A] : \mathcal{X} \rightarrow [0, 1]$ függvény mérhetősége miatt tetszőleges $\mu \in \mathcal{M}_1(\mathcal{X})$ valószínűségi mérték esetén a

$$\mathbb{P}_\mu[A] := \int_{\mathcal{X}} \mathbb{P}_x[A] \mu(dx)$$

hozzárendelés valószínűségi mértéket definiál az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}})$ téren. Figyeljük meg, hogy ha μ az $x \in \mathcal{X}$ pontra koncentrált Dirac-mértékkel egyezik meg, akkor $\mathbb{P}_\mu = \mathbb{P}_x$.

Definiáljuk az $X_t : \mathcal{X}^{\mathbb{N}} \rightarrow \mathcal{X}$ függvényeket a t -edik koordinátára történő vetítésként $(x_0, x_1, \dots, x_t, \dots) \mapsto x_t$ módon. Az $(X_t)_{t \in \mathbb{N}}$ sztochasztikus folyamatot kanonikus folyamatnak nevezzük. Erre a kanonikus folyamatra igaz a következő tétel, amely Revuz [10] 1.2.4 állításának segítségével egyszerűen igazolható, ezért mi itt nem bizonyítjuk.

2.7. Tétel. Az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}}, \mathbb{P}_\mu)$ valószínűségi mértéktéren értelmezett $(X_t)_{t \in \mathbb{N}}$ kanonikus folyamat homogén Markov lánc P átmenet-valószínűségekkel és μ kezdeti eloszlással. Ezt a folyamatot a P Markov kernelhez tartozó kanonikus Markov láncnak nevezzük.

A homogén Markov tulajdonság (2.4) alakja és $\mathbb{P}_\mu$ definíciója miatt a kanonikus Markov láncra igaz, hogy

$$\begin{aligned} & P_\mu[X_0 \in A_0, X_1 \in A_1, \dots, X_n \in A_n] = \\ &= \int_{A_0} \mu(dx_0) \int_{A_1} P(x_0, dx_1) \int_{A_2} P(x_1, dx_2) \cdots \int_{A_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, A_n). \end{aligned} \quad (2.6)$$

X_n eloszlása leírható a P^n Markov kernellel abban az egyszerűbb esetben, amikor a kezdeti eloszlás δ_x , mivel a fenti összefüggés miatt ilyenkor $\mathbb{P}_x[X_n \in A] = P^n(x, A)$. A továbbiakban $\mathbb{E}_x$ és $\mathbb{E}_\mu$ mindig a $\mathbb{P}_x$, illetve $\mathbb{P}_\mu$ szerinti várható értékeket jeleni.

A kanonikus Markov lánc segítségével megfogalmazzuk a (2.4) homogén Markov tulajdonság egy ekvivalens alakját, amely a későbbiekben különösen hasznos lesz számunkra. Legyen $\theta : \mathcal{X}^\mathbb{N} \rightarrow \mathcal{X}^\mathbb{N}$ az eltolás transzformáció, amelyre

$$\theta((x_0, x_1, \dots)) = (x_1, x_2, \dots). \quad (2.7)$$

A θ transzformáció θ^k -val jelölt k -adik hatványán az önmagával vett k -szoros kompozícióját értjük.

2.8. Állítás. Az $(X_n)_{n \in \mathbb{N}}$ kanonikus folyamat pontosan akkor homogén Markov lánc P átmenet-valószínűségekkel és μ kezdeti eloszlással, ha minden $n \in \mathbb{N}_+$ és minden $f \in b\mathcal{X}^\mathbb{N}$ függvényre

$$\mathbb{E}_{X_n}[f] = \mathbb{E}_\mu[f \circ \theta^n \mid \mathcal{F}_n]. \quad (2.8)$$

A továbbiakban kizárólag a P Markov kernelhez tartozó kanonikus Markov láncsal foglalkozunk, ami Revuz [10] 1.2.14 állítása szerint nem jelent valódi megszorítást.

2.3. Markov láncok stabilitásának alapfogalmai

A diszkrét állapotterű Markov láncok számos tulajdonságának létezik megfelelője általános állapottér esetén is. Az irreducibilitás, a periódus, illetve a tranzien és rekurrens állapotok fogalmainak mindegyike általánosítható. Ezek a fogalmak amellet, hogy összekapcsolják a diszkrét- és az általános állapotterű Markov láncok elméletét, az általános állapottérű láncok hosszú távú viselkedésének tanulmányozásához is nagy segítséget nyújtanak.

2.3.1. φ -irreducibilitás

Markov láncok egyik legfontosabb tulajdonsága az irreducibilitás, ami diszkrét állapottér esetén azt fejezi ki, hogy bármelyik állapotból elérhető bármelyik másik állapot véges idő alatt. Általános állapottér esetén az egyes állapotok helyett csak az $\mathcal{X}$ állapottér mérhető részhalmazai állnak rendelkezésünkre. Ennek megfelelően irreducibilitáson azt értjük, hogy

az állapottér mérhető részhalmazai közül a „nagy” halmazok bármely állapotból elérhetők. Az elérhetőség precíz értelmezéséhez minden $A \in \mathcal{B}$ mérhető halmazhoz bevezetjük a

$$\tau_A := \inf\{n \geq 1 : X_n \in A\}$$

megállási időt, ami az első olyan (nem nulla) időpontot jelenti, amelyre $X_n \in A$. A „nagy” halmazokat egy úgynevezett irreducibilitási mérték segítségével értelmezve definiáljuk az irreducibilitás fogalmát.

2.9. Definíció. Az X Markov lánc φ -irreducibilis, ha létezik egy olyan nem azonosan nulla φ mérték $\mathcal{B}$ -n, amelyre igaz, hogy ha valamilyen $A \in \mathcal{B}$ halmazra $\varphi(A) > 0$, akkor minden $x \in \mathcal{X}$ esetén

$$\mathbb{P}_x[\tau_A < \infty] > 0. \quad (2.9)$$

Ilyenkor a φ mértéket irreducibilitási mértéknek hívjuk.

Ez a definíció tehát azt fejezi ki, hogy a φ pozitív mértékű „nagy” halmazok bármelyik $x \in \mathcal{X}$ állapotból pozitív valószínűséggel elérhetők véges idő alatt. Könnyen belátható, hogy (2.9) ekvivalens azzal, ha azt követeljük meg, hogy $\varphi(A) > 0$ esetén

$$\sum_{k=1}^{\infty} P^k(x, A) > 0$$

teljesüljön.

Az irreducibilitási mérték nem egyértelmű, viszont minden φ -irreducibilis Markov láncnak létezik olyan ψ irreducibilitási mértéke, amelyre igaz, hogy a lánc tetszőleges másik irreducibilitási mértéke abszolút folytonos ψ -re. Ennek köszönhetően a ψ mértéket maximális irreducibilitási mértéknek nevezzük, mert tetszőleges φ -irreducibilitási mérték esetén ha $\varphi(A) > 0$, akkor $\psi(A) > 0$ is igaz. Erről szól Meyn és Tweedie [9] 4.2.2 tétele.

2.10. Tétel. Legyen az X Markov lánc φ -irreducibilis, és legyen φ' tetszőleges irreducibilitási mérték. Ekkor létezik egy olyan $\psi \in \mathcal{M}(\mathcal{X})$ mérték, amelyre a Markov lánc ψ -irreducibilis és $\varphi' \ll \psi$. Az is igaz, hogy a ψ mérték ekvivalens a

$$\psi'(A) = \int_{\mathcal{X}} \varphi'(dx) \sum_{k=0}^{\infty} 2^{-(k+1)} P^k(x, A) \quad (2.10)$$

mértékkel.

A továbbiakban ψ -vel mind a maximális irreducibilitási mértéket jelöljük. A tétel miatt a maximális irreducibilitási mértékek ekvivalensek, ezért a ψ pozitív mértékű halmazok rendszere egyértelműen meghatározott. Ezeket a halmazokat $\mathcal{B}^+$ -al jelöljük.

2.3.2. Tranziens és rekurrens Markov láncok

Diszkrét állapottér esetén a Markov lánc állapotait tranziensnek vagy rekurrensnek nevezzük aszerint, hogy egy valószínűséggel visszatérünk-e az adott állapotba ugyan ebből az állapotból indulva. Átlános állapottéren az irreducibilitásnál látotthoz hasonlóan az egyes állapotok helyett mérhető halmazok segítségével fogalmazzuk meg ezeket a tulajdonságokat. Jelölje η_A azoknak az időpontoknak a számát, amikor a lánc A -ban tartózkodik:

$$\eta_A := \sum_{k=1}^{\infty} \mathbb{I}_A(X_k).$$

A tranziens és rekurrens tulajdonságokat az η_A valószínűségi változó segítségével fogalmazzuk meg.

2.11. Definíció (tranziens és rekurrens halmazok). Az $A \in \mathcal{B}$ halmaz egyenletesen tranziens, ha

$$\sup_{x \in A} \mathbb{E}_x [\eta_A] < \infty.$$

Az A halmaz rekurrens, ha minden $x \in A$ esetén

$$\mathbb{E}_x [\eta_A] = \infty,$$

és Harris rekurrens, ha minden $x \in A$ -ra

$$\mathbb{P}_x [\eta_A = \infty] = 1.$$

A definíció szerint tehát akkor nevezzük egy A halmazt egyenletesen tranziensnek, ha A akármelyik pontjából indulva várhatóan csak véges sokszor térünk vissza A -ba. Az A halmaz rekurrens tulajdonsága azt jelenti, hogy A -ból indulva várhatóan végtelen sokszor visszatérünk, a Harris rekurrens tulajdonság pedig úgy értelmezhető, hogy egy valószínűséggel végtelen sokszor térünk vissza.

ψ -irreducibilis láncokra Meyn és Tweedie [9] 8.0.1 tétele szerint a diszkrét állapottérhez hasonlóan igaz a tranziens-rekurrens dichotómia. A tétel szerint, ha a Markov lánc ψ -irreducibilis, akkor vagy minden $A \in \mathcal{B}^+$ halmaz rekurrens vagy $\mathcal{X}$ lefedhető megszámlálhatóan sok egyenletesen tranziens halmazzal. Ez alapján egy Markov láncot (Harris) rekurrensnek hívunk, ha ψ -irreducibilis és minden $A \in \mathcal{B}^+$ (Harris) rekurrens. Harris rekurrens láncokkal kapcsolatban fontos megjegyeznünk, hogy [9] 9.1.4 tétele szerint ha a lánc Harris rekurrens, akkor az a látszólag erősebb tulajdonság is igaz, hogy ha $A \in \mathcal{B}^+$, akkor

$$\mathbb{P}_x [\eta_A = \infty] = 1 \tag{2.11}$$

minden $x \in \mathcal{X}$ -re teljesül, nem csak A elemeire.

Bár nem minden rekurrens Markov lánc Harris rekurrens, [9] 9.0.1 tétele szerint a rekurrens láncok „majdnem” Harris rekurrenssek. A tétel szerint ψ -irreducibilis rekurrens Markov láncok esetén léteznek olyan $H, N \in \mathcal{B}$ halmazok, amelyekre $\mathcal{X} = H \cup N$, $\psi(N) = 0$ és minden $A \in \mathcal{B}^+$ Harris rekurrens, ha $A \subset H$. Ezen felül az is igaz, hogy H elnyelő halmaz, amit úgy értünk, hogy $P(x, H) = 1$, ha $x \in H$.

2.3.3. Periódus

Szintén fontos tulajdonság, hogy a Markov lánc periodikus-e. Ciklikus halmazoknak nevezzük az olyan $(D_1, \dots, D_n)$ páronként diszjunkt $\mathcal{B}$ -beli halmazokat, amelyekre teljesül, hogy

$$P(x, D_{k+1}) = 1$$

minden $x \in D_k$ esetén, ahol $k + 1$ -et modulo n értjük. Ilyenkor tehát ha a lánc egy adott időpontban a D_k halmazban tartózkodik, akkor a következő időpontban egy valószínűséggel D_{k+1} -ben lesz.

Meyn és Tweedie [9] 5.4.4. tétele szerint ψ -irreducibilis láncok esetén ψ -m.m. egyértelműen létezik egy $(D_1, \dots, D_d)$ ciklikus felbontása az $\mathcal{X}$ állapotternek, amely maximális abban az értelemben, hogy ha $(D'_1, \dots, D'_n)$ az előzőektől különböző ciklikus halmazok, akkor n osztója d -nek. Az egyértelműséget úgy értjük, hogy bármely más $(D'_1, \dots, D'_d)$ d elemű ciklikus felbontás esetén a halmazok sorrendjének esetleges átrendezésével minden $1 \leq k \leq d$ esetén ψ -m.m. $D'_k = D_k$. Emellett az is igaz, hogy az $\bigcup_{k=1}^d D_k$ halmaz komplementere ψ nullmértékű. A láncot aperiodikusnak nevezzük, ha ebben a maximális ciklikus felbontásban csak egy halmaz szerepel, azaz $d = 1$.

2.3.4. Invariáns mértékek

Az egyik legerősebb stabilitási tulajdonság, amit egy sztochasztikus folyamattól elvárhatunk az, hogy a folyamat stacionárius legyen. Ilyenkor X_n eloszlása nem függ n értékétől, amit úgy értelmezhetünk, hogy a folyamat ugyan véletlenszerű, de ez a véletlenszerűség nem függ n -től. Ahhoz, hogy egy Markov lánc stacionárius legyen természetesen szükséges, hogy tetszőleges n -re X_n és X_0 eloszlása megegyezzen. Amikor X_0 eloszlása valamilyen adott π kezdeti eloszlás, akkor X_n eloszlása πP^n alakban fejezhető ki. Vagyis ahhoz, hogy X stacionárius legyen szükséges, hogy $\pi = \pi P$ teljesüljön. Ez motiválja az invariáns mértékek fogalmának bevezetését.

2.12. Definíció (invariáns mérték). A $\pi \in \mathcal{M}(\mathcal{X})$ nem azonosan nulla mértéket a P kernel invariáns mértékének hívjuk, ha

$$\pi = \pi P.$$

A továbbiakban mindig az invariáns mértékeket jelöljük π -vel. A fenti definícióból természetesen következik, hogy ha π invariáns, akkor $\pi = \pi P^n$ minden $n \geq 1$ esetén is teljesül. A (2.6) összefüggésből látható, hogy ha a kezdeti eloszlás egy invariáns valószínűségi mérték, akkor a Markov lánc stacionárius lesz, vagyis az X Markov lánc pontosan akkor stacionárius, ha a kezdeti eloszlás invariáns valószínűségi mérték. Emiatt kulcsfontosságú kérdés, hogy létezik-e a P átmenet-valószínűségeknek invariáns valószínűségi mértéke, és ha igen, akkor egyértelmű-e? Ezzel a kérdéssel kapcsolatban a Meyn és Tweedie [9] könyvéből származó 10.0.1 tételt emeljük ki.

2.13. Tétel. Ha az X Markov lánc rekurrens, akkor létezik invariáns mértéke, és ez az invariáns mérték konstans szorzó erejéig egyértelmű.

A stacionárius tulajdonságnál gyengébb stabilitási fogalom, ha csak azt várjuk el, hogy a $\mathbb{P}_\mu[X_n \in \cdot]$ eloszlások konvergáljanak. Az eloszlások konvergenciájához nincs szükség arra, hogy a kezdeti eloszlás invariáns legyen, ezért ez a tulajdonság sokkal szélesebb körben alkalmazható. Az eloszlások konvergenciáját többféleképpen is lehet értelmezni. Markov láncokkal kapcsolatban az egyik leggyakrabban használt konvergenciafogalom a totális variációs metrika szerinti konvergencia.

2.14. Definíció. A μ és λ mértékek totális variációs távolságán a $\mu - \lambda$ előjeles mérték totális variációját értjük:

$$\|\mu - \lambda\|_{TV} := |\mu - \lambda|(\mathcal{X}) = \sup_{|f| \leq 1, f \text{ mérhető}} \mu(f) - \lambda(f). \quad (2.12)$$

A definícióban a $\mu(f)$ jelölést $\int_{\mathcal{X}} f d\mu$ rövid jelölésére használjuk. Az invariáns valószínűségi mértékek szerepe nem merül ki abban, hogy segítségükkel stacionárius Markov láncokat adhatunk meg. Az is igaz, hogy az eloszlások totális variáció szerinti határértéke csak invariáns valószínűségi mérték lehet, amit úgy értelmezhetünk, hogy az invariáns valószínűségi mérték a hosszú távú viselkedést is meghatározza.

2.15. Állítás. Ha valamilyen μ kezdeti eloszlás és π valószínűségi mérték esetén

$$\lim_{n \rightarrow \infty} \|\mathbb{P}_\mu[X_n \in \cdot] - \pi\|_{TV} = 0,$$

akkor π invariáns.

Bizonyítás. Az állítást a (2.6) összefüggés segítségével látjuk be. A totális variációs konvergencia miatt tetszőleges $A \in \mathcal{B}$ halmazra

$$\begin{aligned} \pi(A) &= \lim_{n \rightarrow \infty} \int_{\mathcal{X}} \mu(dx) P^n(x, A) = \\ &= \lim_{n \rightarrow \infty} \int_{\mathcal{X}} \mu(dx) \int_{\mathcal{X}} P^{n-1}(x, dy) P(y, A) = \int_{\mathcal{X}} \pi(dx) P(x, A). \end{aligned}$$

Az utolsó lépésben azt használtuk fel, hogy a (2.3) összefüggés miatt $(\mu(P^{n-1}P(\cdot, A))) = (\mu P^{n-1})P(\cdot, A)$, és azt, hogy a $\mathbb{P}_\mu[X_n \in \cdot]$ valószínűségi mértékek totális variációban vett konvergenciájából következik a $P(\cdot, A)$ korlátos mérhető függvény integráljainak konvergenciája. $\square$

3. fejezet

Általános állapotterű Markov láncok ergodikus tulajdonságai

Az eddig bevezetett fogalmak segítségével megkezdjük az általános állapotterű Markov láncok hosszú távú viselkedésének vizsgálatát. Ezzel kapcsolatban a legfontosabb kérdés az, hogy az X Markov lánc X_n marginálisainak eloszlásai konvergálnak-e a totális variációs metrikában. Ha igen, akkor a Markov láncot ergodikusnak nevezzük. Az ergodikus Markov láncok fontos jellemzője a konvergencia sebessége. Ha például a konvergencia megfelelően gyors, akkor a Markov láncra centrális határeloszlás-tételek is teljesülhetnek. A fejezetben először definiáljuk, hogy pontosan mit értünk egy Markov lánc ergodikusságán. Ezután Hairer és Mattingly [5] gondolatmenetét követve bebizonyítjuk a gemoetria ergodikus tételt. Végül pedig Douc et al. [2] segítségével bemutatjuk, hogy Harris rekurrens Markov láncok ergodikussága azt is jelenti, hogy a Markov láncnak megfeleltethető dinamikai rendszer is ergodikus. Ennek következménye, hogy Birkhoff ergodikus tétele miatt ilyenkor igaz a nagy számok törvénye.

Először tehát definiáljuk mit is értünk ergodikusságon.

3.1. Definíció. A P Markov kernelt ergodikusnak nevezzük, ha létezik egy olyan $\pi \in \mathcal{M}_1(\mathcal{X})$ valószínűségi mérték, hogy minden $x \in \mathcal{X}$ esetén

$$\|P^n(x, \cdot) - \pi\|_{TV} \rightarrow 0 \quad \text{ha } n \rightarrow \infty.$$

Az ergodikusság fenti definíciója szoros kapcsolatban áll az előző fejezet stabilitási fogalmaival. Az invariáns mértékeknél a 2.15 állításban láttuk, hogy a 3.1 definíció π mértéke csak invariáns lehet. Az ergodikusság szükséges feltétele, hogy a Markov lánc irreducibilis legyen, mivel a fenti tulajdonságú π mérték irreducibilitási mérték is egyben. Valóban, ha $\pi(A) > 0$, akkor a totális variációs konvergencia miatt minden x esetén minden elég nagy n -re $P^n(x, A) > 0$, és így $\mathbb{P}_x[\tau_A < \infty] > 0$. A maximális irreducibilitási mérték (2.10)

karakterizációjának köszönhetően π maximális irreducibilitási mérték is, mert

$$\int_{\mathcal{X}} \pi(dx) \sum_{k=0}^{\infty} 2^{-(k+1)} P^k(x, A) = \sum_{k=0}^{\infty} 2^{-(k+1)} [\pi P^k](A) = \pi(A).$$

Emiatt a $\mathcal{B}^+$ halmazok rendszere a π pozitív mértékű halmazokból áll. Ennek köszönhetően az is igaz, hogy ergodikus Markov lánc rekurrens, ugyanis ha $\pi(A) > 0$, akkor minden x -re

$$\mathbb{E}_x [\eta_A] = \sum_{k=1}^{\infty} P^k(x, A) = \infty.$$

Az is nyilván igaz, hogy ha P ergodikus, akkor aperiodikus, mert ilyenkor a maximális ciklikus felbontás legalább egy halmazára $\pi(D_i) > 0$. Ezeket az észrevételeket egy állításban foglaljuk össze.

3.2. Állítás. Ha a P Markov kernel ergodikus a 3.1 definíció értelmében, akkor π az egyértelmű invariáns valószínűségi mérték, a P kernel π -irreducibilis, rekurrens és aperiodikus.

Meyn és Tweedie [9] 13.0.1 tétele szerint ezek a szükséges feltételek majdnem elégségesek is.

3.3. Tétel. Legyen a P Markov kernel ψ -irreducibilis, aperiodikus és Harris rekurrens. Ha P -nek van invariáns valószínűségi mértéke, akkor P ergodikus.

Ergodikus Markov láncokkal kapcsolatban sok esetben hasznos, ha a totális variációs konvergencia sebességéről is rendelkezünk információkkal. Erről szól a következő definíció.

3.4. Definíció. A P Markov kernelt geometriailag ergodikusnak nevezzük, ha létezik egy olyan $r \in (0, 1)$ konstans, és egy $M : \mathcal{X} \rightarrow [0, \infty)$ függvény, amelyekkel

$$\|P^n(x, \cdot) - \pi\|_{TV} \leq M(x)r^n.$$

Ha ez az M függvény korlátos is, akkor azt mondjuk, hogy P egyenletesen ergodikus. Ezt úgy is megfogalmazhatjuk, hogy léteznek olyan $C \geq 0$ és $\rho \in (0, 1)$ konstansok, amelyekkel

$$\sup_{x \in \mathcal{X}} \|P^n(x, \cdot) - \pi\|_{TV} \leq C\rho^n.$$

3.1. A geometriai ergodikus tétel

Régóta ismertek olyan tételek, amelyek a totális variációs konvergenciára adnak elégséges feltételeket. Egy ilyen tételre Hairer és Mattingly [5] adtak a korábbiaknál sokkal elemibb bizonyítást, itt ezt mutatjuk be. A korábbi ismert bizonyítások alapgondolata az volt, hogy egy konstrukció segítségével az általános állapotter esetét a megszámlálható állapotter esetéhez hasonlóan kezelhetővé tegyék. Mi itt ezt a klasszikus bizonyítást a terjedelme és összetettsége miatt nem részletezzük. Szeretnénk viszont megemlíteni, hogy ez a korábbi bizonyítás motivált sok olyan fogalmat és módszert, amelyek azóta is kulcsfontosságúak,

többek között a drift és minorizálási feltételek használatát. A következő drift feltétel azt fogalmazza meg, hogy létezik egy olyan V teszt függvény, amelyre a P kernel „kontraktívan” hat.

3.5. Feltétel (drift). Létezik egy $V : \mathcal{X} \rightarrow [0, \infty)$ függvény és $K \geq 0$, $\gamma \in (0, 1)$ konstansok, amelyekkel minden x -re

$$[PV](x) \leq \gamma V(x) + K$$

A minorizálási feltétel szerint informálisan fogalmazva ennek a V függvénynek egy szinthal-mazába pozitív valószínűséggel vissza tud térni a folyamat. Ezt precízen a következő feltétel fogalmazza meg:

3.6. Feltétel (minorizálás). Létezik egy olyan $\alpha \in (0, 1)$ konstans és egy $\nu \in \mathcal{M}_1(\mathcal{X})$ valószínűségi mérték, amelyekre minden $A \in \mathcal{B}$ esetén teljesül, hogy

$$\inf_{x \in \mathcal{C}} P(x, A) \geq \alpha \nu(A),$$

ahol $\mathcal{C} = V^{-1}([0, R])$ és $R > 2K/(1 - \gamma)$. K és γ a 3.5 drift feltétel konstansai.

A következő tételt fogjuk belátni:

3.7. Tétel (geometriai ergodikus tétel). Ha az X Markov láncra teljesülnek a 3.5 drift és 3.6 minorizálás feltételek, akkor X -nek létezik egyértelmű invariáns valószínűségi mértéke, és X geometriailag ergodikus a 3.4 definíció értelmében. Ha a 3.5 feltételből származó V függvény korlátos is, akkor X egyenletesen ergodikus.

Hairer és Mattingly bizonyításának alapötlete az, hogy a totális variációs metrika helyett egy ahhoz hasonlóan definiált, súlyozott totális variációs metrikaként értelmezhető metrika családdal dolgozzunk. A dolgozatnak ebben a szakaszában V mindig a 3.5 feltétel V függvényét jelenti. Először minden $\beta > 0$ konstanshoz bevezetünk egy súlyozott szuprémum normát:

$$\|f\|_\beta := \sup_{x \in \mathcal{X}} \frac{|f(x)|}{1 + \beta V(x)}. \quad (3.1)$$

Ez a hozzárendlés normát definiál azoknak az $f : \mathcal{X} \rightarrow \mathbb{R}$ mérhető függvényeknek a vektor-terén, amelyre $\|f\|_\beta$ véges. A $\|\cdot\|_\beta$ norma segítségével egy metrikát adunk meg az $\mathcal{M}_1(\mathcal{X})$ téren. Legyen

$$\rho_\beta(\mu, \lambda) := \sup_{\|f\|_\beta \leq 1} \int_{\mathcal{X}} f(x)(\mu - \lambda)(dx). \quad (3.2)$$

Azt, hogy ρ_β -t súlyozott totális variációként értelmezzük, az az előzővel ekvivalens definíció indokolja, amely szerint $\rho_\beta(\mu, \lambda)$ úgy kapható meg, hogy az $1 - \beta V$ függvényt a $|\mu - \lambda|$ totális variációs mérték szerint integráljuk:

$$\rho_\beta(\mu, \lambda) := \int_{\mathcal{X}} 1 + \beta V(x) |\mu - \lambda|(dx). \quad (3.3)$$

Fontos megjegyeznünk, hogy erre a metrikára a $\|\cdot\|_\beta$ norma definíciója miatt igaz, hogy tetszőleges V és $\beta > 0$ esetén $\|\mu - \lambda\|_{TV} \leq \rho_\beta(\mu, \lambda)$.

Mielőtt ténylegesen hozzákezdénénk a geometriai ergodikus tétel bizonyításához, további elkőkészületeket teszünk. Először egy félnormát definiálunk az $f : \mathcal{X} \rightarrow \mathbb{R}$ mérhető függvényeken:

$$\|f\|_\beta := \sup_{x \neq y} \frac{|f(x) - f(y)|}{2 + \beta V(x) + \beta V(y)}.$$

Megjegyezzük, hogy hasonlóan $\|\cdot\|_\beta$ esetéhez $\|\cdot\|_\beta$ csak azoknak az f függvényeknek a vektorterén definiál valóban félnormát, amelyekre $\|f\|_\beta < \infty$. Az így bevezetett félnorma segítségével $\mathcal{M}_1(\mathcal{X})$ -en is megadunk egy súlyozott totális variációs távolságot, amit d_β -val jelölünk:

$$d_\beta(\mu, \lambda) := \sup_{\|f\|_\beta \leq 1} \int_{\mathcal{X}} f(x)(\mu - \lambda)(dx). \quad (3.4)$$

Megmutatjuk, hogy valószínűségi mértékek esetén a d_β és ρ_β metrikák megegyeznek.

3.8. Lemma. Ha $\mu, \lambda \in \mathcal{M}_1(\mathcal{X})$ valószínűségi mértékek, akkor $\rho_\beta(\mu, \lambda) = d_\beta(\mu, \lambda)$.

Bizonyítás. A definíciókból látszik, hogy $\|f\|_\beta \leq \|f\|_\beta$, ami miatt d_β esetén (3.4)-ben bővebb halmazon vesszük a szupréumot, mindt ρ_β esetén (3.2)-ben. Ezért $\rho_\beta \leq d_\beta$.

A másik irányú egyenlőtlenség igazolásához azt mutatjuk meg, hogy tetszőleges olyan f függvényhez, amelyre $\|f\|_\beta \leq 1$, megadható olyan c konstans, amely esetén $\|f + c\|_\beta \leq 1$. Mivel μ és λ valószínűségi mértékek, ezért

$$\int_{\mathcal{X}} f(x)(\mu - \lambda)(dx) = \int_{\mathcal{X}} (f(x) + c)(\mu - \lambda)(dx),$$

amiből következik, hogy $d_\beta \leq \rho_\beta$.

Az adott f függvényhez a

$$c := \inf_{x \in \mathcal{X}} 1 + \beta V(x) - f(x) \quad (3.5)$$

konstans megfelelő lesz. Először azt mutatjuk meg, hogy a fenti infimum véges. $\|f\|_\beta \leq 1$ miatt tetszőleges x -re és y -ra

$$f(x) \leq |f(y)| + |f(x) - f(y)| \leq |f(y)| + 2 + \beta V(x) + \beta V(y).$$

Ezt átrendezve kapjuk, hogy

$$1 + \beta V(x) - f(x) \geq -1 - \beta V(y) - |f(y)|.$$

y tetszőleges volt, ezért kaptunk egy alsó korlátot c -re, vagyis valóban konstanst definiáltunk.

A lemma igazolásához már csak azt kell belátnunk, hogy a (3.5)-ben megadott c konstanssal valóban $\|f + c\|_\beta \leq 1$, ha $\|f\|_\beta \leq 1$. c definíciója miatt

$$f(x) + c \leq f(x) + 1 + \beta V(x) - f(x) = 1 + \beta V(x).$$

A másik irányú egyenlőtlenséghez figyeljük meg, hogy $\|f\|_\beta \leq 1$ miatt

$$f(x) - f(y) \geq -2 - \beta V(x) - \beta V(y)$$

mindig teljesül. Ezt a megfigyelést felhasználva így folytatjuk a számolást:

$$\begin{aligned} f(x) + c &= \inf_y f(x) + 1 + \beta V(y) - f(y) \\ &\geq \inf_y 1 + \beta V(y) - 2 - \beta V(x) - \beta V(y) \geq -(1 + \beta V(x)), \end{aligned}$$

vagyis valóban $|f(x) + c| \leq 1 + \beta V(x)$, amivel a lemmát beláttuk. $\square$

A geometriai ergodikus tétel bizonyításának legfontosabb lépése során azt látjuk be, hogy a P kernel a $\|\cdot\|_\beta$ félnormában kontraktívan hat minden olyan függvényre, amelyre $\|f\|_\beta < \infty$. A lemma fontosságát az adja, hogy egyszerű következményként megkapjuk belőle azt, hogy P kontrakció a ρ_β metrikával ellátott $\mathcal{M}_1(\mathcal{X})$ téren. Ebből pedig egyrészt az invariáns valószínűségi mérték létezését, másrészt a totális variációs konvergenciára vonatkozó eredményeket is megkapjuk. A következő lemmát szeretnénk tehát belátni.

3.9. Lemma. Ha 3.5 drift és 3.6 minorizálás feltételek teljesülnek, akkor léteznek olyan $r \in (0, 1)$ és $\beta > 0$ konstansok, amelyek esetén minden mérhető f függvényre

$$\|Pf\|_\beta \leq r\|f\|_\beta. \quad (3.6)$$

Bizonyítás. Ha $\|f\|_\beta = \infty$, akkor (3.6) automatikusan teljesül, ezért feltehető, hogy $\|f\|_\beta$ véges. Figyeljük meg, hogy f helyett $f/\|f\|_\beta$ -ra áttérve elég azt megmutatni, hogy minden $\|f\|_\beta = 1$ függvényhez létezik olyan (f -től nem függő) $r \in (0, 1)$, amellyel

$$\|Pf\|_\beta \leq r.$$

Legyen $\|f\|_\beta = 1$ tetszőleges rögzített függvény. Azt szeretnénk belátni, hogy minden x és y esetén

$$|Pf(x) - Pf(y)| \leq r(2 + \beta V(x) + \beta V(y)). \quad (3.7)$$

Az első esethez tegyük fel, hogy

$$V(x) + V(y) \geq R.$$

Az előző, 3.8 lemma bizonyításában láttuk, hogy f -hez található olyan c konstans, amellyel $\|f + c\|_\beta = \|f\|_\beta$. Mivel $\|f + c\|_\beta = \|f\|_\beta$ és $\|P(f + c)\|_\beta = \|Pf\|_\beta$, ezért feltehető, hogy $\|f\|_\beta = 1$. Ezt használva azt kapjuk, hogy

$$|Pf(x) - Pf(y)| \leq P|f|(x) + P|f|(y) \leq P(1 + \beta V(x)) + P(1 + \beta V(y)) = 2 + \beta PV(x) + \beta PV(y).$$

A 3.5 drift feltétel miatt, és mert feltettük, hogy $V(x) + V(y) \geq R$ úgy folytatjuk a becsléseinket, hogy

$$\begin{aligned} |Pf(x) - Pf(y)| &\leq 2 + 2\beta K + \beta\gamma(V(x) + V(y)) \\ &\leq 2 + \left(\frac{2\beta K}{R} + \beta\gamma\right)(V(x) + V(y)) \\ &= 2 + \beta\gamma_0 V(x) + \beta\gamma_0 V(y), \end{aligned}$$

ahol az utolsó lépésben bevezettük a $\gamma_0 := \gamma + 2K/R$ konstans. A 3.6 minorizálási feltétel miatt az is igaz, hogy $\gamma_0 \in (0, 1)$. Ahhoz, hogy az első esetben igazoljuk a (3.7) egyenlőtlenséget, megmutatjuk, hogy van egy olyan $\gamma_1 \in (\gamma_0, 1)$ konstans, amellyel

$$2 + \beta\gamma_0 V(x) + \beta\gamma_0 V(y) \leq 2\gamma_1 + \gamma_1\beta V(x) + \gamma_1\beta V(y).$$

Az egyenlőtlenséget átrendezve kapjuk, hogy

$$\gamma_1 \geq \frac{2 + \beta R \gamma_0}{2 + \beta R}.$$

Mivel a lemma állítása annál erősebb, minél kisebb γ_1 , ezért legyen

$$\gamma_1 := (2 + \beta R \gamma_0) / (2 + \beta R).$$

A második esethez tegyük fel, hogy $V(x) + V(y) \leq R$. Ilyenkor $x, y \in V^{-1}([0, R])$, ezért használhatjuk a 3.6 minorizálási feltételt. Vezessük be a $\tilde{P}$ Markov kernelt:

$$\tilde{P}(x, \cdot) = \frac{1}{1 - \alpha} P(x, \cdot) - \frac{\alpha}{1 - \alpha} \nu(\cdot).$$

Ezzel a jelöléssel

$$Pf(x) = (1 - \alpha)\tilde{P}f(x) + \alpha\nu(f) \quad \text{és} \quad Pf(y) = (1 - \alpha)\tilde{P}f(y) + \alpha\nu(f).$$

Felhasználva hogy V nemnegativitása miatt $\tilde{P}V(x) \leq \frac{1}{1 - \alpha} PV(x)$, majd a 3.5 drift feltételt, azt kapjuk, hogy

$$\begin{aligned} |Pf(x) - Pf(y)| &= (1 - \alpha)|\tilde{P}f(x) - \tilde{P}f(y)| \\ &\leq (1 - \alpha)2 + (1 - \alpha)\beta(\tilde{P}V(x) + \tilde{P}V(y)) \\ &\leq (1 - \alpha)2 + \beta(PV(x) + PV(y)) \\ &\leq 2(1 - \alpha + \beta K) + \beta\gamma V(x) + \beta\gamma V(y). \end{aligned}$$

Ha β értékét α_0/K -nak választjuk, ahol $\alpha_0 \in (0, \alpha)$ tetszőleges konstans, akkor $\gamma_2 := \max\{1 - (\alpha - \alpha_0), \gamma\}$ választással a második esetben

$$|Pf(x) - Pf(y)| \leq \gamma_2(2 + \beta V(x) + \beta V(y)).$$

Ezzel a lemmát bebizonyítottuk az $r = \max\{\gamma_1, \gamma_2\}$ konstanssal. □

Ennek a lemmának egyszerű következménye, hogy a P operátor kontrakció $\mathcal{M}_1(\mathcal{X})$ -en a ρ_β metrika szerint.

3.10. Tétel. Minden $\mu, \lambda \in \mathcal{M}_1(\mathcal{X})$ valószínűségi mértékre

$$\rho_\beta(\mu P, \lambda P) \leq r \rho_\beta(\mu, \lambda),$$

ahol β és r a 3.9 lemma konstansai.

Bizonyítás. A 3.8 lemma szerint valószínűségi mértékeken $\rho_\beta = d_\beta$. Ezt kombinálva a (2.3) összefüggéssel azt kapjuk, hogy

$$\rho_\beta(\mu P, \lambda P) = d_\beta(\mu P, \lambda P) = \sup_{\|f\|_\beta \leq 1} (\mu P)(f) - (\lambda P)(f) = \sup_{\|f\|_\beta \leq 1} \mu(Pf) - \lambda(Pf).$$

Mivel a 3.9 lemma miatt P a $\|\cdot\|_\beta$ félnormában kontraktív, ezért igaz a következő egyenlőtlenség, mert a jobb oldalán bővebb halmazon veszünk szuprémumot:

$$\sup_{\|f\|_\beta \leq 1} \mu(Pf) - \lambda(Pf) \leq \sup_{\{f : \|Pf\|_\beta \leq r\}} \mu(Pf) - \lambda(Pf).$$

Egyszerű átalakításokkal adódik, hogy

$$\begin{aligned} \sup_{\{f : \|Pf\|_\beta \leq r\}} \mu(Pf) - \lambda(Pf) &= r \sup_{\{f : \|Pf\|_\beta \leq r\}} \mu\left(\frac{Pf}{r}\right) - \lambda\left(\frac{Pf}{r}\right) \\ &= r \sup_{\left\{f : \left\|\frac{Pf}{r}\right\|_\beta \leq 1\right\}} \mu\left(\frac{Pf}{r}\right) - \lambda\left(\frac{Pf}{r}\right). \end{aligned}$$

A bizonyítás utolsó lépéséhez figyeljük meg, hogy

$$r \sup_{\left\{f : \left\|\frac{Pf}{r}\right\|_\beta \leq 1\right\}} \mu\left(\frac{Pf}{r}\right) - \lambda\left(\frac{Pf}{r}\right) \leq r \sup_{\|f\|_\beta \leq 1} \mu(f) - \lambda(f) = r d_\beta(\mu, \lambda),$$

mert az egyenlőtlenség jobb oldalán bővebb halmazon veszünk szuprémumot. $\square$

A geometriai ergodikus tétel bizonyításának befejezéséhez az invariáns valószínűségi mérték létezését igazoljuk.

3.11. Lemma. Ha a 3.5 kontraktivitási és 3.6 minorizálási feltételek teljesülnek, akkor létezik egy π invariáns valószínűségi mérték, amelyre $\pi(V) < \infty$.

Bizonyítás. Legyen $\mu_n = \delta_x P^n$ minden $n \geq 0$ egészhez. A 3.7 tételből tudjuk, hogy vannak olyan $r \in (0, 1)$ és $\beta > 0$ konstansok, amelyekre

$$\rho_\beta(\mu_{n+1}, \mu_n) \leq r^n \rho_\beta(\mu_1, \delta_x),$$

vagyis (μ_n) Cauchy-sorozat. A totális variációs metrika teljes a véges mértékeken, ezért ρ_β is teljes azokon a valószínűségi mértékeken, amelyek esetén V integrálható. Emiatt a (μ_n) sorozat konvergál egy π valószínűségi mértékhez a ρ_β metrikában, és ρ_β definíciója miatt a sorozat a totális variációs metrikában is π -hez konvergál. Az, hogy π invariáns, abból következik, hogy P kontrakció a totális variációs metrikában, ezért

$$\pi P = \left(\lim_{n \rightarrow \infty} \mu_n \right) P = \lim_{n \rightarrow \infty} \mu_n P = \lim_{n \rightarrow \infty} \mu_{n+1} = \pi.$$

□

Az eddig igazolt tételek segítségével sikerült belátnunk a szakasz elején kimondott 3.7 tételt is, amely szerint a 3.5 és 3.6 feltételek teljesülése esetén az X Markov lánc geometriailag ergodikus, korlátos V esetén pedig egyenletesen ergodikus.

3.12. Tétel. Ha 3.5 és 3.6 feltételek teljesülnek, akkor létezik a π invariáns valószínűségi mérték, és

$$\|P^n(x, \cdot) - \pi\|_{TV} \leq r^n(2 + \beta V(x) + \beta \pi(V)),$$

ahol r és β a 3.9 lemma konstansai.

Bizonyítás. A tétel egyszerű következménye annak, hogy tetszőleges $\mu, \lambda \in \mathcal{M}_1(\mathcal{X})$ valószínűségi mértékekre $\|\mu - \lambda\|_{TV} \leq \rho_\beta(\mu, \lambda)$, mivel

$$\|P^n(x, \cdot) - \pi\|_{TV} = \|\delta_x P^n - \pi\|_{TV} \leq \rho_\beta(\delta_x P^n, \pi) \leq r^n \rho_\beta(\delta_x, \pi).$$

A $\rho_\beta(\delta_x, \pi)$ mennyiséget a (3.3) összefüggéssel számolhatjuk ki:

$$\rho_\beta(\delta_x, \pi) = 2 + \beta V(x) + \beta \pi(V).$$

□

3.2. Ergodikus Markov láncok, mint ergodikus dinamikai rendszerek

Az ergodikusság fogalmát eddig csak Markov láncok egy, a hosszú távú viselkedést jellemző tulajdonságaként értettük. Ergodikusságról viszont nem csak Markov láncok, hanem dinamikai rendszerek esetén is beszélhetünk. A célunk, hogy Markov láncok ergodikusságának fogalmát összekapcsoljuk a dinamikai rendszerek ergodikusságával. Ennek köszönhetően Markov láncokra is alkalmazhatjuk a dinamikai rendszerekkel kapcsolatosan ismert ergodikus tételeket, amelyek segítségével megmutatjuk, hogy ha egy Markov lánc ergodikus és Harris rekurrens, akkor tetszőleges kezdeti eloszlással teljesül rá a nagy számok törvénye.

Először dinamikai rendszerekkel kapcsolatban vezetünk be néhány alapfogalmat. Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mező, és $\tau : \Omega \rightarrow \Omega$ egy olyan mérhető függvény, amelyre minden $A \in \mathcal{A}$ esetén igaz, hogy $\mathbb{P}[\tau^{-1}(A)] = \mathbb{P}[A]$. Ilyenkor azt mondjuk, hogy a τ transzformáció

mértéktartó, az $(\Omega, \mathcal{A}, \mathbb{P}, \tau)$ négyest pedig mértéktartó dinamikai rendszernek nevezzük. Az $A \in \mathcal{A}$ eseményt invariáns eseménynek hívjuk, ha $\tau^{-1}(A) = A$. Az invariáns események σ -algebrát alkotnak, amelyet $\mathcal{I}$ -vel jelölünk:

$$\mathcal{I} := \{A \in \mathcal{A} : \tau^{-1}(A) = A\}.$$

3.13. Definíció (ergodikus dinamikai rendszer). Az $(\Omega, \mathcal{A}, \mathbb{P}, \tau)$ mértéktartó dinamikai rendszert ergodikusnak nevezzük, ha az invariáns események $\mathcal{I}$ σ -algebrája $\mathbb{P}$ -triviális, vagyis minden $A \in \mathcal{I}$ eseményre $\mathbb{P}[A] \in \{0, 1\}$.

Ergodikus mértéktartó dinamikai rendszerekkel kapcsolatban az egyik legfontosabb állítás Birkhoff tétele (Klenke [7] 20.14 tétel), amely a nagy számok törvényének dinamikai rendszerekre vonatkozó megfelelője. A τ leképezés k -adik hatványán az önmagával vett k -szoros kompozícióját értjük, τ^0 pedig az identikus leképezés.

3.14. Tétel (Birkhoff). Legyen $f \in \mathcal{L}^1(\mathbb{P})$. Ekkor

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau^k \xrightarrow{n \rightarrow \infty} \mathbb{E}[f | \mathcal{I}] \quad \mathbb{P}\text{-m.m.},$$

és ha a mértéktartó dinamikai rendszer ergodikus, akkor

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau^k \xrightarrow{n \rightarrow \infty} \mathbb{E}[f] \quad \mathbb{P}\text{-m.m.}$$

Sztocasztikus folyamatokra is tekinthetünk dinamikai rendszerekként. Legyen $(X_n)_{n \in \mathbb{N}}$ a kanonikus Markov lánc az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}}, \mathbb{P}_\mu)$ szorzattéren, és $\theta : \mathcal{X}^{\mathbb{N}} \rightarrow \mathcal{X}^{\mathbb{N}}$ a (2.7) összefüggéssel definiált eltolás transzformáció. Könnyen látható, hogy az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}}, \mathbb{P}_\mu, \theta)$ dinamikai rendszer pontosan akkor mértéktartó, ha a $\mathbb{P}_\mu$ valószínűségi mértékkel a kanonikus Markov lánc stacionárius, vagyis a μ kezdeti eloszlás invariáns valószínűségi mérték. Douc et al. [2] felépítését követve a következő tételt fogjuk belátni.

3.15. Tétel. Legyen X Harris rekurrens, ergodikus Markov lánc π invariáns valószínűségi mértékkel. Ekkor tetszőleges μ kezdeti eloszlás esetén minden olyan $f : \mathcal{X}^{\mathbb{N}} \rightarrow \mathbb{R}$ függvényre, amelyre $\mathbb{E}_\pi[|f|] < \infty$ teljesül, hogy

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ \theta^k \xrightarrow{n \rightarrow \infty} \mathbb{E}_\pi[f] \quad \mathbb{P}_\mu\text{-m.m.} \quad (3.8)$$

Speciálisan, ha f csak a nulladik koordinátától függ, akkor ezt úgy is érhetjük, hogy f -re teljesül a nagy számok törvénye, azaz

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \xrightarrow{n \rightarrow \infty} \pi(f) \quad \mathbb{P}_\pi\text{-m.m.}$$

A tételt két részben bizonyítjuk be. Először csak azzal az esettel foglalkozunk, amikor a kezdeti eloszlás a π invariáns valószínűségi mérték. Ebben az esetben azt mutatjuk meg, hogy az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}}, \mathbb{P}_\pi, \theta)$ mértéktartó dinamikai rendszer ergodikus, amiből a Birkhoff tétel miatt következik (3.8). Ezután azt fogjuk belátni, hogy ha van olyan kezdeti eloszlás, amellyel (3.8) teljesül, akkor tetszőleges kezdeti eloszlással is teljesül. Mindkét részben hasznosak lesznek számunkra a harmonikus függvények.

3.16. Definíció. A $h : \mathcal{X} \rightarrow \bar{\mathbb{R}}$ függvényt harmonikusnak nevezzük, ha minden x -re

$$Ph(x) = h(x).$$

A fenti definícióba természetesen azt is beleértjük, hogy a bal oldal mindig értelmes. Harmonikus függvényekre fontos példát jelentenek az invariáns események indikátoraiból képzett várható értékek.

3.17. Lemma. Ha az $A \in \mathcal{B}$ esemény invariáns a θ transzformációra, akkor a $h(x) := \mathbb{E}_x[\mathbb{I}_A]$ függvény harmonikus.

Bizonyítás. Az A esemény invariáns, ezért $\mathbb{I}_A = \mathbb{I}_A \circ \theta$. Ennek köszönhetően, felhasználva a Markov tulajdonság (2.4) és (2.8) alakjait

$$Ph(x) = \mathbb{E}_x[h(X_1)] = \mathbb{E}_x[\mathbb{E}_{X_1}[\mathbb{I}_A]] = \mathbb{E}_x[\mathbb{E}_\pi[\mathbb{I}_A \circ \theta \mid \mathcal{F}_1]] = \mathbb{E}_x[\mathbb{I}_A] = h(x).$$

□

Először tehát azt fogjuk belátni, hogy a 3.1 definíció értelmében ergodikus Markov lánc esetén az $(\mathcal{X}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}}, \mathbb{P}_\pi, \theta)$ mértéktartó dinamikai rendszer ergodikus. Ehhez először két lemmát bizonyítunk, amelyek [2] 1.4.5 és 5.2.2 állításai. Emlékeztetőül idézzük fel, hogy egy $A \in \mathcal{B}$ halmazt akkor hívtunk elnyelő halmaznak, ha minden $x \in A$ esetén $P(x, A) = 1$.

3.18. Lemma. Ha a P Markov kernelnek létezik egyértelmű invariáns valószínűségi mértéke, akkor minden $A \in \mathcal{B}$ elnyelő halmazra $\pi(A) \in \{0, 1\}$.

Bizonyítás. Legyen π az egyértelmű invariáns valószínűségi mérték és $A \in \mathcal{B}$ tetszőleges adott elnyelő halmaz. Megmutatjuk, hogy $\pi_A := \pi(A \cap \cdot)$ is invariáns mérték. Legyen $B \in \mathcal{B}$ tetszőleges. Felhasználva, hogy $\pi_A \leq \pi$ és hogy A elnyelő halmaz, ezért $[\pi_A P](A^c) = 0$, a következő egyenlőtlenséget kapjuk:

$$[\pi_A P](B) = [\pi_A P](A \cap B) + [\pi_A P](A^c \cap B) \leq [\pi P](A \cap B) + [\pi_A P](A^c) = \pi_A(B).$$

Ezt az egyenlőtlenséget B^c -re alkalmazva megkapjuk a másik irányú egyenlőtlenséget is. Ehhez azt használjuk fel, hogy $\pi_A(\mathcal{X}) = [\pi_A P](\mathcal{X}) < \infty$, ezért

$$[\pi_A P](\mathcal{X}) - [\pi_A P](B) = [\pi_A P](B^c) \leq \pi_A(B^c) = \pi_A(\mathcal{X}) - \pi_A(B).$$

Ha az A elnyelő halmazra $\pi(A) > 0$, akkor $\pi_A/\pi(A)$ is invariáns valószínűségi mérték. π egyértelműsége miatt $\pi_A(\cdot) = \pi(A)\pi(\cdot)$, ami csak úgy lehetséges, ha $\pi(A) = 1$. □

A 3.15 tétel bizonyításához szükséges másik lemma invariáns események indikátorainak egy számunkra hasznos felírását adja meg.

3.19. Lemma. Ha $A \in \mathcal{I}$ invariáns esemény, akkor $\mathbb{I}_A = \mathbb{E}_{X_0}[\mathbb{I}_A] \mathbb{P}_\pi$ -m.m.

Bizonyítás. Először figyeljük meg, hogy ha A invariáns esemény, akkor $\mathbb{I}_A = \mathbb{I}_A \circ \theta$. Ennek, és a Markov tulajdonságnak köszönhetően az $(\mathbb{E}_{X_k}[\mathbb{I}_A], \mathcal{F}_k)_{k \in \mathbb{N}}$ folyamat martingál, mert

$$\mathbb{E}_\pi[\mathbb{E}_{X_{k+1}}[\mathbb{I}_A] \mid \mathcal{F}_k] = \mathbb{E}_\pi[\mathbb{E}_\pi[\mathbb{I}_A \mid \mathcal{F}_{k+1}] \mid \mathcal{F}_k] = \mathbb{E}_\pi[\mathbb{E}_{X_k}[\mathbb{I}_A] \mid \mathcal{F}_{k+1}] = \mathbb{E}_{X_k}[\mathbb{I}_A],$$

ahol az első átalakításban a Markov tulajdonság (2.8) alakját használtuk. Az $(\mathbb{E}_{X_k}[\mathbb{I}_A], \mathcal{F}_k)$ martingál korlátos, ezért konvergens, így

$$\lim_{k \rightarrow \infty} \mathbb{E}_{X_k}[\mathbb{I}_A] = \lim_{k \rightarrow \infty} \mathbb{E}_\pi[\mathbb{I}_A \mid \mathcal{F}_k] = \mathbb{I}_A \quad \mathbb{P}_\pi\text{-m.m. és } L^1\text{-ben.} \quad (3.9)$$

Felhasználva, hogy θ mértéktartó, és $\mathbb{I}_A = \mathbb{I}_A \circ \theta$ azt kapjuk, hogy

$$\mathbb{E}_\pi[|\mathbb{I}_A - \mathbb{E}_{X_0}[\mathbb{I}_A]|] = \mathbb{E}_\pi[|\mathbb{I}_A - \mathbb{E}_{X_0}[\mathbb{I}_A]| \circ \theta^k] = \mathbb{E}_\pi[|\mathbb{I}_A - \mathbb{E}_{X_k}[\mathbb{I}_A]|].$$

Ebből k -val végtelenhez tartva következik a lemma állítása, mert (3.9) miatt

$$\lim_{k \rightarrow \infty} \mathbb{E}_\pi[|\mathbb{I}_A - \mathbb{E}_{X_k}[\mathbb{I}_A]|] = 0.$$

□

A 3.18 és 3.19 lemmák segítségével igazoljuk Douc et al. [2] 5.2.6 tételét, amely elégséges feltételt ad a Markov lánchoz asszociált dinamikai rendszer ergodikusságára.

3.20. Tétel. Ha a P Markov kernelnek létezik egyértelmű invariáns valószínűségi mértéke, akkor az $(\mathcal{X}^\mathbb{N}, \mathcal{B}^{\otimes \mathbb{N}}, \mathbb{P}_\pi, \theta)$ dinamikai rendszer ergodikus.

Bizonyítás. Legyenek $A \in \mathcal{I}$ egy tetszőleges invariáns esemény, $h(x) := \mathbb{E}_x[\mathbb{I}_A]$ és legyen $B = \{x : h(x) = 1\}$. A 3.17 lemmában láttuk, hogy ez a h függvény harmonikus, ezért minden $x \in B$ esetén

$$1 = h(x) = Ph(x) = \mathbb{E}_x[h(X_1)].$$

Emiatt $\mathbb{P}_x[h(X_1) = 1] = 1$, ha $x \in B$. B definíciója miatt minden B -beli x -re

$$1 = \mathbb{P}_x[h(X_1) = 1] = \mathbb{P}_x[X_1 \in B] = P(x, B),$$

vagyis B elnyelő halmaz. A 3.18 lemmában láttuk, hogy ilyenkor $\pi(B) \in \{0, 1\}$.

A 3.19 lemma miatt $\mathbb{P}_\pi[h(X_0) = \mathbb{I}_A] = 1$, ezért $\mathbb{P}_\pi[h(X_0) \in \{0, 1\}] = 1$. Ezt összevetve az eddigiekkel

$$\mathbb{P}_\pi[A] = \mathbb{E}_\pi[\mathbb{I}_A] = \mathbb{E}_\pi[h(X_0)] = \int_{\mathcal{X}} h(x) \pi(dx) = \int_{\mathcal{X}} \mathbb{I}_B \pi(dx) = \pi(B) \in \{0, 1\}.$$

□

A nagy számok törvényét megfogalmazó 3.15 tételünk bizonyításának utolsó lépéséhez már csak azt kell megmutatnunk, hogy Harris rekurrens Markov láncok esetén valóban elég azzal az esettel foglalkoznunk, amikor a kezdeti eloszlás π . Ehhez Meyn és Tweedie [9] 17.1.5 tételét használjuk fel.

3.21. Lemma. Ha X ψ -irreducibilis, Harris rekurrens Markov lánc, és h korlátos harmonikus függvény, akkor h konstans.

Bizonyítás. A h függvény korlátossága miatt elég azt belátni, hogy minden c konstans esetén vagy $h \geq c$ vagy $h \leq c$. Jelölje ψ az X Markov lánc maximális irreducibilitási mértékét. Az $\{x : h(x) \geq c\}$ és $\{x : h(x) \leq c\}$ halmazok legalább egyike pozitív ψ mértékű.

Tegyük fel, hogy $\psi(\{x : h(x) \geq c\}) > 0$ (a másik eset teljesen hasonló módon kezelhető). A Markov lánc (2.11)-ben megfogalmazott Harris-rekurrens tulajdonsága miatt minden x -re $\limsup_{k \rightarrow \infty} h(X_k) \geq c$ $\mathbb{P}_x$ -m.m., mert minden $x \in \mathcal{X}$ esetén

$$\mathbb{P}_x \left[\limsup_{k \rightarrow \infty} h(X_k) \geq c \right] = \mathbb{P}_x \left[\sum_{k=1}^{\infty} \mathbb{I}_{\{x : h(x) \geq c\}}(X_k) = \infty \right] = 1.$$

Megmutatjuk, hogy $h(x) = \mathbb{E}_x \left[\limsup_{k \rightarrow \infty} h(X_k) \right]$ és ezért minden x -re $h(x) \geq c$.

Ehhez azt használjuk fel, hogy a $(h(X_k), \mathcal{F}_k)_{k \in \mathbb{N}}$ folyamat martingál. Valóban, h harmonikusságát és a Markov tulajdonság (2.4) alakját felhasználva kapjuk, hogy tetszőleges μ kezdeti eloszlás esetén

$$h(X_k) = Ph(X_k) = \mathbb{E}_\mu[h(X_{k+1}) \mid \mathcal{F}_k].$$

Mivel h korlátos, a $h(X_k)$ martingál konvergens, ezért

$$\begin{aligned} \mathbb{E}_x \left[\limsup_{k \rightarrow \infty} h(X_k) \right] &= \mathbb{E}_x \left[\lim_{k \rightarrow \infty} h(X_k) \right] = \lim_{k \rightarrow \infty} \mathbb{E}_x[h(X_k)] \\ &= \lim_{k \rightarrow \infty} \mathbb{E}_x[\mathbb{E}_x[h(X_k) \mid \mathcal{F}_0]] = \mathbb{E}_x[h(X_0)] = h(x), \end{aligned}$$

ahol az utolsó lépésben a martingál tulajdonságot használtuk. $\square$

A 3.21 lemmának köszönhetően a nagy számok törvényének kezdeti eloszlástól való függetlenségét egy megfelelő harmonikus függvény megadásával mutatjuk meg.

3.22. Tétel. Legyen X ψ -irreducibilis, Harris rekurrens Markov lánc. Ha a nagy számok törvényének (3.8) alakja teljesül valamilyen kezdeti eloszlással, akkor minden kezdeti eloszlással teljesül.

Bizonyítás. Legyen $f \in \mathcal{L}^1(\mathbb{P}_\pi)$ tetszőleges függvény, és legyen

$$A := \left\{ \omega \in \mathcal{X}^\mathbb{N} : \lim_{k \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \theta^k = \mathbb{E}_\pi[f] \right\}$$

azoknak a pontoknak a halmaza, ahol teljesül a nagy számok törvénye. Az A esemény invariáns, ezért a 3.17 lemma miatt a $h(x) := \mathbb{E}_x[\mathbb{I}_A]$ függvény harmonikus. A feltételeink miatt valamilyen μ kezdeti eloszlás esetén $\mu(h) = 1$. A 3.21 lemmának köszönhetően ez a h függvény konstans 1, vagyis $\mathbb{P}_x[A] = 1$ minden x -re, és így minden $\lambda \in \mathcal{M}_1(\mathcal{X})$ -re is $\mathbb{P}_\lambda[A] = 1$. $\square$

Mivel ergodikus Markov láncnak létezik egyértelmű invariáns valószínűségi mértéke, ezért a 3.20 tétel miatt az ergodikusság elégséges feltétel arra, hogy a megfelelő dinamikai rendszer ergodikus legyen. Ekkor alkalmazható a Birkhoff ergodikus tétel, és teljesül a (3.8) nagy számok törvénye, abban az esetben, ha a kezdeti eloszlás π . A 3.22 tételben megmutattuk, hogy ez a kezdeti eloszlásra vonatkozó feltétel nem szükséges, és a 3.15 tétel valóban tetszőleges kezdeti eloszlás mellett is igaz.

4. fejezet

A Centrális határeloszlás-tétel

A centrális határeloszlás-tétel (a továbbiakban CHT) a valószínűségi számítás fontos és széleskörűen alkalmazott eredménye. A CHT hagyományosan csak független és azonos eloszlású valószínűségi változók sorozatára vonatkozik. A célunk a fejezetben, hogy elégséges feltételeket mutassunk arra, hogy egy Markov lánc valós értékű függvényére mikor teljesül a CHT? Azt mutatjuk meg, hogy Markov láncok esetén sem a függetlenség, sem az azonos eloszlás nem szükséges feltételek. A függetlenségre vonatkozó feltétel helyett csak egyfajta „gyenge összefüggőségre” lesz szükségünk, amelyet formálisan sztochasztikus folyamatok keverő tulajdonságának segítségével fogalmazzunk meg. Először ismertetjük sztochasztikus folyamatok keverésének különböző változatait, majd bemutatjuk, hogy ezek a fogalmak hogyan alkalmazhatók Markov láncokra. Stacionárius, keverő folyamatokra több különböző centrális határeloszlás-tétel is ismert, emiatt ha a Markov lánc stacionárius és keverő tulajdonságú, akkor a CHT is teljesül az alkalmasság momentumfeltételeket kielégítő függvényeire. Azonban az sem szükséges, hogy a Markov lánc stacionárius legyen, a nagy számok törvényéhez hasonlóan ebben az esetben is igaz, hogy ha a lánc Harris rekurrens, akkor ha van olyan kezdeti eloszlás, amellyel a CHT teljesül, akkor mindegyikkel teljesül. Az itt ismertetett tételek többsége megtalálható Jones [6] összegző munkájában, amely nagy segítségünkre volt a fejezet írása során.

4.1. Keverő folyamatok

A keverés fogalma σ -algebrák közötti összefüggés erősségét számszerűsíti. A formális definícióhoz legyenek adottak az $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mező, és az $\mathcal{F}, \mathcal{G} \subset \mathcal{A}$ rész σ -algebrák. Az $\mathcal{F}$ és $\mathcal{G}$ σ -algebrák közötti összefüggőség mérésére használható α , ϕ és ρ keverési együtthatókat a következőképpen definiáljuk:

$$\alpha(\mathcal{F}, \mathcal{G}) = \sup\{|\mathbb{P}[A \cap B] - \mathbb{P}[A]\mathbb{P}[B]| : A \in \mathcal{F}, B \in \mathcal{G}\}, \quad (4.1)$$

$$\phi(\mathcal{F}, \mathcal{G}) = \sup\{|\mathbb{P}[B | A] - \mathbb{P}[B]| : A \in \mathcal{F}, \mathbb{P}[A] \neq 0, B \in \mathcal{G}\}, \quad (4.2)$$

$$\rho(\mathcal{F}, \mathcal{G}) = \sup\{\text{Corr}(f, g) : f \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}), g \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})\}. \quad (4.3)$$

A ρ keverési együttható definíciójában $\text{Corr}(f, g)$ az f és g függvények közötti korrelációt jelenti, amit úgy értünk, hogy

$$\text{Corr}(f, g) = \frac{\text{Cov}(f, g)}{\sqrt{\text{Var}(f)}\sqrt{\text{Var}(g)}} = \frac{\mathbb{E}[fg] - \mathbb{E}[f]\mathbb{E}[g]}{\sqrt{\mathbb{E}[f^2] - \mathbb{E}^2[f]}\sqrt{\mathbb{E}[g^2] - \mathbb{E}^2[g]}}.$$

Mindhárom keverési együttható nemnegatív, és pontosan akkor nulla, ha az $\mathcal{F}$ és $\mathcal{G}$ σ -algebrák függetlenek. Ez alapján a keverési együtthatókat úgy értelmezhetjük, hogy minél kisebb az értékük, annál kevésbé összefüggők az $\mathcal{F}$ és $\mathcal{G}$ σ -algebrák. A fentiekén kívül többféle keverési együttható is definiálható, amelyek hasonló módon mérik az összefüggés erősségét. Mi csak az α , ϕ és ρ együtthatókkal foglalkozunk, mivel a centrális határeloszlástétel szempontjából ezek bizonyulnak a leginkább hasznosnak.

Ellentétben az α és ρ együtthatókkal $\phi(\mathcal{F}, \mathcal{G})$ és $\phi(\mathcal{G}, \mathcal{F})$ nem mindig egyeznek meg. A fenti keverési együtthatók kapcsolatáról szól Doukhan [3] könyvének 1.1.1 állítása, amely szerint

$$2\alpha(\mathcal{F}, \mathcal{G}) \leq \phi(\mathcal{F}, \mathcal{G}) \quad (4.4)$$

és

$$4\alpha(\mathcal{F}, \mathcal{G}) \leq \rho(\mathcal{F}, \mathcal{G}) \leq 2\phi^{1/2}(\mathcal{F}, \mathcal{G})\phi^{1/2}(\mathcal{G}, \mathcal{F}). \quad (4.5)$$

A keverési együtthatók segítségével definiálhatjuk sztochasztikus folyamatok keverő tulajdonságát is. Legyen adott egy $(X_n)_{n \in \mathbb{N}}$ sztochasztikus folyamat. Vezessük be az $\mathcal{F}_n := \sigma(X_k : k \leq n)$ „múlt” és az $\mathcal{F}^n := \sigma(X_k : k \geq n)$ „jövő” σ -algebrákat.

4.1. Definíció. Az $(X_n)_{n \in \mathbb{N}}$ sztochasztikus folyamat

(i) α -keverő vagy erősen keverő, ha

$$\alpha(n) := \sup_{k \geq 0} \alpha(\mathcal{F}_k, \mathcal{F}^{k+n}) \rightarrow 0 \quad \text{ha } n \rightarrow \infty,$$

(ii) ρ -keverő vagy aszimptotikusan korrelálatlan, ha

$$\rho(n) := \sup_{k \geq 0} \rho(\mathcal{F}_k, \mathcal{F}^{k+n}) \rightarrow 0 \quad \text{ha } n \rightarrow \infty,$$

(iii) ϕ -keverő vagy egyenletesen keverő, ha

$$\phi(n) := \sup_{k \geq 0} \phi(\mathcal{F}_k, \mathcal{F}^{k+n}) \rightarrow 0 \quad \text{ha } n \rightarrow \infty.$$

Ha a folyamat a fenti definíció szerint keverő, azt úgy is mondjuk, hogy teljesíti a megfelelő keverési feltételt. Sztochasztikus folyamatok keverési tulajdonsága mindhárom esetben az $\mathcal{F}_k$ „múlt” és az $\mathcal{F}^{k+n}$ „jövő” σ -algebrák aszimptotikus függetlenségeként értelmezhető. Ez a tulajdonság teszi lehetővé olyan aszimptotikus eredmények igazolását összefüggő folyamatokra, amelyek hagyományosan csak független folyamatokra vonatkoznak.

4.2. Megjegyzés. A (4.4) és (4.5) egyenlőtlenségek közvetlen következménye, hogy $4\alpha(n) \leq \rho(n)$ és $\rho(n) \leq 2\phi^{1/2}(n)$, tehát ha egy folyamat egyenletesen keverő, akkor aszimptotikusan korrelálatlan, és mind az egyenletes keverésből, mind az aszimptotikus korrelálatlanságból következik, hogy a folyamat erősen keverő. Fontos megjegyeznünk, hogy ha az $(X_n)_{n \in \mathbb{N}}$ folyamat teljesíti valamelyik fenti keverési felételt, akkor tetszőleges f mérhető függvénnyel az $(f(X_n))_{n \in \mathbb{N}}$ folyamat is teljesíti ugyanezt a feltételt, hiszen a megfelelő keverési együttható definíciójában szűkebb σ -algebrákon vesszük a szuprémumot.

4.2. Keverő Markov láncok

Ebben a szakaszban azt a kérdést vizsgáljuk meg, hogy milyen feltételek mellett lesz egy Markov lánc erősen, illetve egyenletesen keverő. A ρ -keverés esetét itt nem tárgyaljuk, mivel az általunk használtaktól eltérő módszereket igényel. Azonban a teljesség kedvéért szeretnénk megemlíteni, hogy ebben az esetben is léteznek elégséges feltételek arra, hogy egy Markov lánc ρ keverő legyen, illetve ρ -keverő folyamatokra is ismert centrális határeloszlás-tétel. Ezekről bővebben [6]-ban olvashatunk. Az α és a ϕ keverés esetében úgy járunk el, hogy a keverési együtthatókat a P^n átmenet-valószínűségek és a π invariáns valószínűségi mérték totális variációs távolságával becsüljük felül. Ezzel a módszerrel nem csak azt állapítjuk meg, hogy a megfelelő keverési együtthatók nullához tartanak, hanem a konvergencia sebességére is kapunk eredményeket, amelyeket a keverő folyamatokra vonatkozó centrális határeloszlás-tételek Markov láncokra történő alkalmazásához fogunk használni.

A bizonyításunk lényegében az Ibragimov és Linnik [12] könyvének 367-368-ik oldalán megtalálható gondolatmenetet követi. A szerzők az idézett könyben Markov láncok egyenletesen keverő tulajdonságával foglalkoznak. Mi az ott megtalálható érvelés kis módosításával az α -keverés esetét is a ϕ -keverés esetéhez hasonlóan kezeljük. Először egy lemmát igazolunk, ami megteremti a kapcsolatot az $\alpha(n)$ és $\phi(n)$ keverési együtthatók és a $\|P^n(x, \cdot) - \pi\|_{TV}$ totális variációs távolság között.

4.3. Lemma. Legyen π a Markov lánc invariáns valószínűségi mértéke. Ha $A = \{X_0 \in A_0, \dots, X_k \in A_k\}$ és $B = \{X_{k+n} \in B_{k+n}, \dots, X_m \in B_m\}$ alakú események, akkor tetszőleges μ kezdeti eloszlás esetén

$$\begin{aligned} & |\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A]\mathbb{P}_\mu[B]| \leq \\ & \leq \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) \|P^n(x_k, \cdot) - \pi\|_{TV} + \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) \|P^{k+n}(x_0, \cdot) - \pi\|_{TV}. \end{aligned}$$

Bizonyítás. A bizonyításunk során a (2.6) azonosság és a totális variációs távolság (2.12) alakjának felhasználásával számolunk.

Legyen μ tetszőleges kezdeti eloszlás. A (2.6) azonosság miatt ha A és B mérhető téglák, akkor

$$|\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A]\mathbb{P}_\mu[B]| =$$

$$\left| \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) \int_{B_{k+n}} P^n(x_k, dx_{k+n}) \cdots \int_{B_{m-1}} P(x_{m-2}, dx_{m-1}) P(x_{m-1}, B_m) \right. \\ \left. - \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) \int_{B_{k+n}} P^{k+n}(x_0, dx_{k+n}) \cdots \int_{B_{m-1}} P(x_{m-2}, dx_{m-1}) P(x_{m-1}, B_m) \right|.$$

Az $f(y) = \mathbb{I}_{B_{k+n}}(y) \int_{B_{k+n+1}} P(y, dx_{k+n+1}) \cdots \int_{B_{m-1}} P(x_{m-2}, dx_{m-1}) P(x_{m-1}, B_m)$ függvény bevezetésével egyszerűsödik a jelölés:

$$\begin{aligned} & |\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A] \mathbb{P}_\mu[B]| = \\ & = \left| \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) [P^n f](x_k) - \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) [P^{k+n} f](x_0) \right| \\ & = \left| \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) ([P^n f](x_k) - \pi(f) + \pi(f)) - \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) [P^{k+n} f](x_0) \right|. \end{aligned}$$

A lemma bizonyításának befejezéséhez a háromszög-egyenlőtlenséget és a totális variációs norma definícióját használjuk fel:

$$\begin{aligned} & |\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A] \mathbb{P}_\mu[B]| \leq \\ & \leq \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) |[P^n f](x_k) - \pi(f)| + \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) |[P^{k+n} f](x_0) - \pi(f)| \\ & \leq \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) \|P^n(x_k, \cdot) - \pi\|_{TV} + \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) \|P^{k+n}(x_0, \cdot) - \pi\|_{TV}. \end{aligned}$$

□

A lemmának köszönhetően mérhető téglák esetén használható becslést kapunk az $\alpha(n)$ és $\phi(n)$ keverési együttthatók definíciójában szereplő $|\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A] \mathbb{P}_\mu[B]|$ mennyiségre. A következő megjegyzés szerint elegendő kizárólag mérhető téglák esetén becsülni ezt a mennyiséget.

4.4. Megjegyzés. A keverési együttthatók meghatározásához valójában nem minden esetben szükséges (4.1)-ben és (4.2)-ben a szuprémumot a teljes $\mathcal{F}$ és $\mathcal{G}$ σ -algebrák eseményein vennünk. Ehelyett elég, ha egy-egy $\mathcal{F}$ -et illetve $\mathcal{G}$ -t generáló félgyűrű eseményeire szorítkozunk. Ezt az a mértékelméletről ismert approximációs tétel (Klenke [7] 1.65 tétele) teszi lehetővé, amely szerint ha $\mathcal{A}$ félgyűrű, μ pedig egy σ -véges mérték $\sigma(\mathcal{A})$ -n, akkor tetszőleges $\varepsilon > 0$ és $A \in \sigma(\mathcal{A})$ esetén vannak olyan $A_1, \dots, A_n \in \mathcal{A}$ páronként diszjunkt félgyűrűbeli halmazok, amelyekkel $\mu(A \Delta \cup_{i=1}^n A_i) < \varepsilon$. Ebből az approximációs tételből egyszerű számolással adódik, hogy az $\alpha(n)$ és $\phi(n)$ keverési együttthatók esetén elég a megfelelő mérhető téglákkal foglalkozni. A számolás teljesen elemi, ezért itt nem részletezzük.

A 4.3 lemma segítségével először azt mutatjuk meg, hogy minden, a 3.1 definíció értelmében ergodikus Markov lánc α -keverő. Az itt bebizonyított tételünk első része Athreya és Pantula [1] tételének egy gyengébb változata, mivel mi csak az egyszerűség kedvéért csak az aperiodikus esettel foglalkozunk. A tétel második felében $\alpha(n)$ konvergenciájának sebességét is megbecsüljük P^n totális variációs konvergenciájának segítségével.

4.5. Tétel. Ha X ergodikus Markov lánc π invariáns valószínűségi mértékkel, akkor X α -keverő tetszőleges μ kezdeti eloszlás esetén. Ha

$$\|P^n(x, \cdot) - \pi\|_{TV} \leq M(x)\gamma(n) \quad (4.6)$$

teljesül valamilyen $M(x)$ és $\gamma(n)$ nemnegatív függvényekkel, akkor

$$\alpha(n) \leq \sup_{k \geq 0} \mathbb{E}_\mu[M(X_k)]\gamma(n) + \mathbb{E}_\mu[M(X_0)]\gamma(k+n).$$

Bizonyítás. A 4.4 megjegyzés miatt $\alpha(n)$ becslésekor feltehető, hogy $A \in \mathcal{F}_0^k$ és $B \in \mathcal{F}_{k+n}^\infty$ mérhető téglák. A 4.3 lemma becslését használva kapjuk, hogy

$$\begin{aligned} & |\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A]\mathbb{P}_\mu[B]| \leq \\ & \leq \int_{\mathcal{X}} \mu(dx) \int_{\mathcal{X}} P^k(x, dy) \|P^n(y, \cdot) - \pi\|_{TV} + \int_{\mathcal{X}} \mu(dx) \|P^{k+n}(x, \cdot) - \pi\|_{TV}. \end{aligned}$$

Figyeljük meg, hogy $\int_{\mathcal{X}} \mu(dx) \int_{\mathcal{X}} P^k(x, A)$ éppen X_k eloszlása $\mathbb{P}_\mu$ szerint, ezért

$$|\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A]\mathbb{P}_\mu[B]| \leq \mathbb{E}_\mu[\|P^n(X_k, \cdot) - \pi\|_{TV} + \|P^{k+n}(X_0, \cdot) - \pi\|_{TV}].$$

Itt az A és B eseményeknek már nincs szerepe, ezért

$$\alpha(n) \leq \sup_{k \geq 0} \mathbb{E}_\mu[\|P^n(X_k, \cdot) - \pi\|_{TV} + \|P^{k+n}(X_0, \cdot) - \pi\|_{TV}]. \quad (4.7)$$

A $\|P^n(x, \cdot) - \pi\|_{TV}$ totális variációs távolság 0-hoz tart minden x esetén, ezért a dominált konvergencia miatt valóban $\alpha(n) \rightarrow 0$, ha $n \rightarrow \infty$, vagyis a Markov lánc α -keverő. A konvergencia sebességére vonatkozó eredmény a (4.7) becslés következménye. $\square$

4.6. Megjegyzés. Az $\alpha(n)$ -re kapott becslés fontos speciális esete az, amikor a kezdeti eloszlás a stacionárius eloszlás. Ilyenkor $\alpha(n) \leq 2\mathbb{E}_\pi[M] \max\{\gamma(n), \gamma(k+n)\}$. Ha a Markov lánc geometriailag ergodikus, akkor (4.6)-ban $\gamma(n) = r^n$ valamilyen $r \in (0, 1)$ -el, és így $\alpha(n) \leq 2\mathbb{E}_\pi[M]r^n$. Ha $\mathbb{E}_\pi[M] < \infty$, akkor pedig $\alpha(n) = O(r^n)$.

A következő tétel Markov láncok egyenletes keverési tulajdonságára ad elégséges feltételt.

4.7. Tétel. Ha az X Markov lánc egyenletesen ergodikus, akkor X egyenletesen keverő. Ilyenkor a $\phi(n)$ keverési együtthatók exponenciális sebességgel tartanak nullához, azaz léteznek olyan K és $\rho \in (0, 1)$ konstansok, amelyekkel $\phi(n) \leq K\rho^n$ minden n -re.

Bizonyítás. Emlékeztetõül idézzük fel, hogy X egyenletes ergodikusságán azt értjük, hogy léteznek olyan C és $\rho \in (0, 1)$ konstansok, hogy minden $n \geq 1$ esetén

$$\sup_{x \in \mathcal{X}} \|P^n(x, \cdot) - \pi(\cdot)\|_{TV} \leq C\rho^n.$$

A bizonyításunk során az α -keverés esetéhez hasonlóan szintén a 4.3 lemmát használjuk, amit a 4.4 megjegyzés tesz lehetővé. A lemma becslése szerint

$$\begin{aligned} & |\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A]\mathbb{P}_\mu[B]| \leq \\ & \leq \int_{A_0} \mu(dx_0) \cdots \int_{A_k} P(x_{k-1}, dx_k) \|P^n(x_k, \cdot) - \pi\|_{TV} + \mathbb{P}_\mu[A] \int_{\mathcal{X}} \mu(dx_0) \|P^{k+n}(x_0, \cdot) - \pi\|_{TV}. \end{aligned}$$

Az egyenletes ergodikusság miatt pedig

$$|\mathbb{P}_\mu[A \cap B] - \mathbb{P}_\mu[A]\mathbb{P}_\mu[B]| \leq \mathbb{P}_\mu[A]Cr^n + \mathbb{P}_\mu[A]Cr^{k+n} \leq \mathbb{P}_\mu[A]2Cr^n.$$

A és B tetszőleges mérhető téglák voltak, ezért valóban $\phi(n) \leq 2Cr^n$, amivel a tételt beláttuk. $\square$

4.3. Centrális határeloszlás-tétel Markov láncokra

A Markov láncokra vonatkozó centrális határeloszlás-tételek precíz kimondásához először is definiáljuk mit értünk azon, hogy a Markov lánc egy függvényére teljesül a centrális határeloszlás-tétel. Ezután megmutatjuk, hogy Harris rekurrens Markov láncok esetén a kezdeti eloszlásnak nincs szerepe a CHT szempontjából. Ez azt jelenti, hogy ha van olyan kezdeti eloszlás, amely esetén teljesül a CHT, akkor tetszőleges kezdeti eloszlással is teljesül. Végül kimondjuk a keverő folyamatokra vonatkozó centrális határeloszlás-tételeket, majd a kezdeti eloszlástól való függetlenség, és az előző szakasz eredményei segítségével Markov láncokra alkalmazzuk ezeket a tételeket.

Először tehát definiáljuk, hogyan értjük azt, hogy a Markov lánc egy függvényére teljesül a CHT.

4.8. Definíció. Legyen X olyan Markov lánc, amelynek létezik π invariáns valószínűségi mértéke. Azt mondjuk, hogy az $f : \mathcal{X} \rightarrow \mathbb{R}$ függvényre teljesül a centrális határeloszlás-tétel a μ kezdeti eloszlással, ha létezik egy $0 < \sigma^2 < \infty$ konstans, amellyel

$$\lim_{n \rightarrow \infty} \mathbb{P}_\mu \left[\frac{S_n(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{u^2}{2}} du = \Phi(t), \quad (4.8)$$

ahol $S_n(\bar{f}) = \sum_{k=0}^{n-1} (f(X_k) - \pi(f))$, Φ pedig a standard normális eloszlás eloszlásfüggvényét jelöli.

Fontos megjegyezni, hogy a centrális határeloszlás-tétel teljesülésekor feltesszük, hogy a Markov láncnak van invariáns valószínűségi mértéke.

A Markov láncokra vonatkozó centrális határeloszlás-tételt a Markov lánc keverő tulajdonságainak segítségével vezetjük le. A keverő folyamatokra vonatkozó centrális határeloszlás-tételek feltételezik, hogy a folyamat stacionárius. Ez a feltétel nem fontos, ha a Markov

lánc Harris rekurrens. Ilyenkor ugyanis, ha a centrális határeloszlás-tétel teljesül valamilyen kezdeti eloszlással, akkor tetszőleges kezdeti eloszlás mellett is teljesül. Erről szól Meyn és Tweedie [9] 17.1.6 állítása, amelyet a nagy számok törvényénél látott 3.22 tételhez hasonló módon, harmonikus függvények segítségével fogunk belátni.

4.9. Tétel. Ha az X ψ -irreducibilis, Harris rekurrens Markov láncra teljesül a centrális határeloszlás-tétel valamilyen kezdeti eloszlással, akkor tetszőleges kezdeti eloszlással teljesül.

Bizonyítás. Tegyük fel, hogy valamilyen μ kezdeti eloszlással teljesül a centrális határeloszlás-tétel. Azt fogjuk megmutatni, ekkor (4.8) $\mathbb{P}_x$ -el is teljesül minden x esetén, amiből már következik a tétel állítása. Vezessük be a

$$g_n^t(x) = \mathbb{P}_x \left[\frac{S_n(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right]$$

és a $g_\infty^t = \lim_{n \rightarrow \infty} g_n^t$ függvényeket. Ezekkel a jelölésekkel a centrális határeloszlás-tétel (4.8) alakja így írható át:

$$\lim_{n \rightarrow \infty} \mathbb{P}_\mu \left[\frac{S_n(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right] = \lim_{n \rightarrow \infty} \int_{\mathcal{X}} g_n^t(x) \mu(dx) = \int_{\mathcal{X}} g_\infty^t(x) \mu(dx) = \Phi(t).$$

A 3.21 lemma szerint minden korlátos harmonikus függvény konstans, ezért elég azt megmutatnunk, hogy a g_∞^t függvény harmonikus és így $g_\infty^t \equiv \Phi(t)$. Egyszerű számolással kapjuk, hogy

$$Pg_\infty^t = \int_{\mathcal{X}} g_\infty^t(y) P(x, dy) = \mathbb{E}_x[g_\infty^t(X_1)] = \lim_{n \rightarrow \infty} \mathbb{E}_x[g_n^t(X_1)] = \lim_{n \rightarrow \infty} \mathbb{E}_x \left[\mathbb{P}_{X_1} \left[\frac{S_n(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right] \right].$$

A Markov tulajdonság (2.8) alakját felhasználva folytatjuk a számolást:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_x \left[\mathbb{P}_{X_1} \left[\frac{S_n(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right] \right] &= \lim_{n \rightarrow \infty} \mathbb{E}_x \left[\mathbb{P}_x \left[\frac{\sum_{i=1}^n (f(X_{i+1}) - \pi(f))}{\sqrt{n\sigma^2}} \leq t \mid \mathcal{F}_1 \right] \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{P}_x \left[\frac{\sum_{i=1}^n (f(X_{i+1}) - \pi(f))}{\sqrt{n\sigma^2}} \leq t \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{P}_x \left[\frac{(\pi(f) - f(X_1)) + S_{n+1}(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right] \\ &= \lim_{n \rightarrow \infty} g_{n+1}^t(x) = g_\infty^t(x). \end{aligned}$$

Ezzel beláttuk, hogy g_∞^t valóban harmonikus, és így $\lim_{n \rightarrow \infty} \mathbb{P}_x \left[\frac{S_n(\bar{f})}{\sqrt{n\sigma^2}} \leq t \right] = \Phi(t)$ minden t -re és minden x -re. $\square$

Keverő stacionárius folyamatokra a keverési együtthatók konvergenciájára és a folyamat momentumaira tett megfelelő feltételek esetén teljesül centrális határeloszlás-tétel. Ez biztosítja, hogy alkalmas feltételek mellett a CHT Markov láncokra is teljesüljön. Az α -keverő stacionárius folyamatokra vonatkozó elégséges feltételeket egy tételben foglaltuk össze, amelyet itt nem bizonyítottunk.

4.10. Tétel. Legyen $(X_n)_{n \in \mathbb{N}}$ olyan centrált stacionárius folyamat, amelyre a következő feltételek legalább egyike teljesül:

- (i) $\text{ess sup } |X_0| < \infty$ és $\sum_{n=0}^{\infty} \alpha(n) < \infty$,
- (ii) $\mathbb{E}[X_0^{2+\delta}] < \infty$ valamilyen $\delta > 0$ esetén és $\sum_{n=0}^{\infty} \alpha(n)^{\delta/(2+\delta)} < \infty$,
- (iii) $\mathbb{E}[X_0^2 \log^+ |X_0|] < \infty$ és vannak olyan K és $\rho \in (0, 1)$ konstansok, amelyekkel $\alpha(n) \leq K\rho^n$.

Ekkor $\sigma^2 = \mathbb{E}[X_0^2] + 2 \sum_{k=1}^{\infty} \mathbb{E}[X_0 X_k] < \infty$. Ha $\sigma^2 > 0$, akkor az $(X_n)_{n \in \mathbb{N}}$ folyamatra teljesül a centrális határeloszlás-tétel, azaz

$$\frac{\sum_{k=0}^{n-1} X_k}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0, 1) \quad \text{ha } n \rightarrow \infty.$$

A tétel (i)-es és (ii)-es pontja Ibragimov és Linnik [12] könyvének 18.5.3 és 18.5.4 tételei, a tétel (iii)-as pontja pedig Doukhan, Massart és Rio [4] cikkében található meg.

Ezeket az α -keverő folyamatokra vonatkozó elégséges feltételeket Markov láncokra is alkalmazhatjuk. A 4.5 tételünkben láttuk, hogy minden ergodikus Markov lánc α -keverő, és az $\alpha(n)$ keverési együtthatókat is sikerült megbecsülnünk a Markov lánc totális variációban vett konvergenciájának sebességével. A következő tételben elégséges feltételeket adunk a CHT teljesülésére geometriailag ergodikus Markov láncok esetén. A 4.10 tétel természetesen akkor is alkalmazható, ha a totális variációs konvergencia sebessége ennél lassabb, mi azonban itt csak a geometriai ergodikusság esetét említjük meg. A tétel kimondásában M a 4.5 tételből származó függvény.

4.11. Tétel. Legyen X olyan geometriailag ergodikus, Harris rekurrens Markov lánc, amelyre $\mathbb{E}_{\pi}[M] < \infty$. Ekkor teljesül a CHT minden olyan $f : \mathcal{X} \rightarrow \mathbb{R}$ függvényre, amelyre $\sigma_f^2 = \mathbb{E}_{\pi}[f(X_0)^2] + 2 \sum_{k=1}^{\infty} \mathbb{E}_{\pi}[f(X_0)f(X_k)] > 0$ és a következő feltételek legalább egyike igaz:

- (i) $\text{ess sup}_{\pi} |f(X_0)| < \infty$,
- (ii) $\mathbb{E}_{\pi}[f(X_0)^{2+\delta}] < \infty$ valamilyen $\delta > 0$ esetén,
- (iii) $\mathbb{E}_{\pi}[f(X_0)^2 \log^+ |f(X_0)|] < \infty$.

Bizonyítás. A Markov lánc Harris rekurrens, ezért a 4.9 tétel miatt elég azzal az esettel foglalkozni, amikor a kezdeti eloszlás π . Az $\alpha(n)$ együtthatókat a (4.7) összefüggéssel becsülve a geometriai ergodikusság és $\mathbb{E}_\pi[M]$ végeessége miatt a tétel az α -keverő folyamatokra vonatkozó 4.10 tétel következménye. $\square$

Az $\mathbb{E}_\pi[M]$ végeességére vonatkozó feltétellel kapcsolatban a geometriai ergodikus tétel bizonyításában láttuk, hogy ha a geometriailag ergodikus tulajdonságot a 3.5 drift és 3.6 minorizálási feltételek segítségével ellenőrizzük, akkor $\mathbb{E}_\pi[M] < \infty$ automatikusan teljesül.

A következő tétel az egyenletesen keverő stacionárius folyamatokra vonatkozó CHT, amely Ibragimov és Linnik [12] könyvének 18.5.2 tétele.

4.12. Tétel. Legyen $(X_n)_{n \in \mathbb{N}}$ olyan centrált stacionárius folyamat, amelyre $\mathbb{E}[X_0^2] < \infty$. Ha

$$\sum_{n=0}^{\infty} \sqrt{\phi(n)} < \infty,$$

akkor $\sigma^2 = \mathbb{E}[X_0^2] + 2 \sum_{k=1}^{\infty} \mathbb{E}[X_0 X_k] < \infty$. Ha $\sigma^2 > 0$, akkor az $(X_n)_{n \in \mathbb{N}}$ folyamatra teljesül a centrális határeloszlás-tétel, azaz

$$\frac{\sum_{k=0}^{n-1} X_k}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0, 1) \quad \text{ha } n \rightarrow \infty.$$

Ennek a tételnek közvetlen következménye az egyenletesen ergodikus Markov láncokra vonatkozó CHT.

4.13. Következmény. Ha X egyenletesen ergodikus, Harris rekurrens Markov lánc, akkor tetszőleges kezdeti eloszlás esetén teljesül a CHT minden olyan $f : \mathcal{X} \rightarrow \mathbb{R}$ függvényre, amelyre $\mathbb{E}_\pi[f(X_0)^2] < \infty$ és $\sigma_f^2 = \mathbb{E}_\pi[f(X_0)^2] + 2 \sum_{k=1}^{\infty} \mathbb{E}_\pi[f(X_0)f(X_k)] > 0$.

Bizonyítás. A 4.9 tétel miatt elég azt megmutatnunk, hogy teljesül a CHT, ha a kezdeti eloszlás π . A 4.7 tétel szerint a Markov lánc egyenletesen keverő, és a $\phi(n)$ együtthatókra igaz, hogy $\phi(n) \leq 2Cr^n$, ahol C és $r \in (0, 1)$ az egyenletes ergodikusság konstansai. Ez alapján a tétel a 4.12 ϕ -keverő folyamatokra vonatkozó CHT következménye. $\square$

5. fejezet

Markov láncok véletlen környezetben

A dolgozatban eddig csak homogén Markov láncokkal foglalkoztunk, amikor a $P^k(x, A) = \mathbb{P}[X_{n+k} \in A \mid X_n = x]$ átmenet-valószínűségek csak k értékétől függtek, n értékétől nem. Ebben az esetben a Markov láncot a P Markov kernel és a kezdeti eloszlás segítségével tudtuk jellemezni. A homogén Markov láncok esetét általánosíthatjuk, ha az egyetlen P kernel helyett Markov kernelek egy $\{Q_y : y \in \mathcal{Y}\}$ családjával foglalkozunk, amelyeket valamilyen $\mathcal{Y}$ halmaz elemeivel indexelünk.

A precíz definícióhoz legyen adott egy $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező, egy $(\mathcal{Y}, \mathcal{A})$ mérhető tér, és egy $Y = (Y_k)_{k \in \mathbb{Z}}$ stacionárius folyamat, amely az értékeit $\mathcal{Y}^{\mathbb{Z}}$ -ben veszi fel, azaz $Y_k : \Omega \rightarrow \mathcal{Y}$ mérhető függvény minden k esetén. A továbbiakban ezt az Y folyamatot környezetnek fogjuk nevezni. Legyen ezen kívül adott az $(\mathcal{X}, \mathcal{B})$ mérhető tér, és egy olyan $Q : \mathcal{Y} \times \mathcal{X} \times \mathcal{B} \rightarrow [0, 1]$ függvény, amelyre

- (i) $Q(\cdot, \cdot, B)$ $\mathcal{A} \otimes \mathcal{B}$ -mérhető minden $B \in \mathcal{B}$ esetén,
- (ii) $Q(y, x, \cdot)$ valószínűségi mérték $(\mathcal{X}, \mathcal{B})$ -n minden y és x esetén.

Világos, hogy rögzített $y \in \mathcal{Y}$ mellett $Q(y, \cdot, \cdot)$ Markov-kernel az $(\mathcal{X}, \mathcal{B})$ téren, így a Q függénnyel Markov kernelek egy $\mathcal{Y}$ -al indexelt családját adtuk meg. Legyen $X = (X_k)_{k \in \mathbb{N}}$ olyan sztochasztikus folyamat, amelyben minden k -ra $X_k : \Omega \rightarrow \mathcal{X}$ mérhető. Azt mondjuk, hogy az X folyamat Markov lánc az Y véletlen környezetben Q átmenet-valószínűségekkel, ha minden $k \in \mathbb{N}_+$ esetén

$$\mathbb{P}[X_{k+1} \in B \mid X_0, \dots, X_k, Y] = Q(Y_k, X_k, B).$$

Ezt a definíciót úgy értelmezhetjük, hogy az egy lépéses átmenet-valószínűségeket a $Q(Y_k, \cdot, \cdot)$ kernel adja meg, vagyis a környezetet jelentő Y folyamat választja ki, hogy az adott lépésben mik legyenek az átmenet-valószínűségek.

A homogén Markov láncok esetéhez hasonlóan itt is azt a kérdést vizsgáljuk meg, hogy mit tudunk mondani a hosszú távú viselkedésről.

Lovas és Rásonyi [8] igazolták, hogy a geometriai ergodikus tételnél látott 3.5 drift és 3.6 minorizálás feltételekhez hasonló feltételek mellett a $\mu_n := \mathbb{P}[X_n \in \cdot]$ eloszlások konvergálnak a totális variációs metrikában. A 3.5 feltétel megfelelője véletlen környezet esetén az

5.1. Feltétel (drift feltétel). Léteznek olyan $V : \mathcal{X} \rightarrow [0, \infty)$ és $\gamma, K : \mathcal{Y} \rightarrow (0, \infty)$ mérhető függvények, amelyekre minden x és y esetén

$$[Q(y)V](x) \leq \gamma(y)V(x) + K(y).$$

Fontos különbség a 3.5 drift feltételhez képest, hogy nem követeljük meg, hogy $\gamma(y) < 1$ legyen minden y -ra. Ehelyett azt várjuk el, hogy γ „hosszú távon” legyen 1-nél kisebb.

5.2. Feltétel (hosszú távú kontrakтивitás). Az 5.1 feltétel γ és K függvényeire teljesül, hogy

$$\tilde{\gamma} := \limsup_{n \rightarrow \infty} \mathbb{E}^{1/n} \left(K(Y_0) \prod_{i=1}^n \gamma(Y_i) \right) < 1.$$

A 3.6 minorizálási feltételt a következő feltétel általánosítja.

5.3. Feltétel (minorizálás). Valamilyen $\varepsilon \in (0, \tilde{\gamma}^{-1/2} - 1)$ számhoz létezik egy olyan $\alpha \in [0, 1)$ konstans és egy $\kappa : \mathcal{Y} \times \mathcal{B} \rightarrow [0, 1]$ Markov-kernel, amelyekre minden y és minden $B \in \mathcal{B}$ esetén

$$\inf_{x \in V^{-1}([0, R(y)])} Q(y, x, B) \geq (1 - \alpha)\kappa(y, B). \quad (5.1)$$

ahol $R(y) = K(y)/(\varepsilon\gamma(y))$ és minden y -ra $V^{-1}([0, R(y)]) \neq \emptyset$.

Lovas és Rásonyi [8]-ban ezek mellett a feltételek mellett a következő tételt bizonyították be.

5.4. Tétel. Ha az 5.1 drift, az 5.2 hosszú távú kontrakтивitási és az 5.3 minorizálási feltételek teljesülnek, akkor létezik egy μ_* valószínűségi mérték $\mathcal{X}$ -en, amelyre

$$\lim_{n \rightarrow \infty} \|\mu_n - \mu_*\|_{TV} = 0.$$

A konvergencia sebességére az is igaz, hogy $\|\mu_n - \mu_*\|_{TV} = O(e^{-\zeta n^{1/3}})$ valamilyen $\zeta > 0$ konstanssal.

Ezen felül az is igaz, hogy ergodikus környezet esetén a nagy számok törvénye is teljesül a Markov láncra.

5.5. Tétel. Ha Y ergodikus, és az 5.1, 5.2 és 5.3 feltételek teljesülnek, akkor minden $f : \mathcal{X} \rightarrow \mathbb{R}$ korlátos mérhető függvényre

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \rightarrow \int_{\mathcal{X}} f d\mu_* \quad L^p\text{-ben}$$

minden $1 \leq p < \infty$ esetén.

5.1. A hosszú távú kontraktivitási feltétel vizsgálata

A 5.4 tétel feltételei közül a legnehezebben az 5.2 hosszú távú kontraktivitási feltétel ellenőrizhető. Ahhoz, hogy 5.2 teljesüljön elég azt megkövetelnünk, hogy

$$\text{ess sup } \gamma < 1$$

legyen, ez azonban egy meglehetősen szigorú feltétel γ -ra. Ha az $(Y_k)_{k \in \mathbb{Z}}$ stacionárius folyamat marginálisai függetlenek, akkor a hosszú távú kontraktivitási feltétel nyilvánvalóan teljesül minden olyan γ függvényre, amelyre $\mathbb{E}[\gamma(Y_1)] < 1$. Kérdés, hogy mennyire enyhíthető a függetlenség feltétele, ahhoz, hogy a kontraktivitási feltétel továbbra is minden 1-nél kisebb várható értékű γ függvényre teljesüljön. Függetlenség helyett a gyenge összefüggőséget keverési együtthatókkal mérve az derül ki, hogy Y keverő tulajdonsága önmagában nem elég a hosszú távú kontraktivitáshoz, amit egy példán keresztül mutatunk meg.

5.1.1. Keverés és a hosszú távú kontraktivitási feltétel

A következő példában azt mutatjuk meg, hogy α -keverő környezet esetén megadható olyan γ függvény, amelynek a várható értéke $1/2$ -hez tetszőlegesen közeli, azonban γ nem teljesíti az 5.2 feltételt. Emiatt nem igaz az, hogy keverő Y esetén a γ -ra tett $\mathbb{E}[\gamma(Y_1)] < 1$ feltétel elégséges lenne a hosszú távú kontraktivitáshoz.

Legyen $(Z_n)_{n \in \mathbb{N}}$ független, azonos eloszlású pénzfeldobás-sorozat, $\mathbb{P}[Z_0 = 0] = \mathbb{P}[Z_0 = 1] = 1/2$ eloszlással, és $Y_n = Z_{n-1} + Z_n$. Ekkor Y_n és Y_m függetlenek, ha $|n - m| > 1$, ezért az $(Y_n)_{n \in \mathbb{N}}$ stacionárius folyamat nyilván α -keverő, hiszen $\alpha(n) = 0$, ha $n \geq 2$. Válasszunk egy olyan γ függvényt, ami háromféle értéket vehet fel. Jelölje ezeket a pozitív számokat x, y és z , és legyen $\gamma(0) = x$, $\gamma(1) = y$ és $\gamma(2) = z$. Tegyük fel, hogy

$$\mathbb{E}[\gamma(Y_1)] = \frac{x}{4} + \frac{y}{2} + \frac{z}{4} < 1. \quad (5.2)$$

Belátjuk, hogy γ ilyen megválasztása esetén előfordulhat, hogy az 5.2 hosszú távú kontraktivitási feltétel nem teljesül. Ehhez először vegyük az $a_n := \mathbb{E}[\prod_{k=1}^n \gamma(Y_k)]$ sorozatot. Ahhoz, hogy az 5.2 feltétel ne teljesüljön, a Hölder-egyenlőtlenség miatt elég, ha $a_n \rightarrow \infty$, ha $n \rightarrow \infty$.

A $\prod_{k=1}^n \gamma(Y_k)$ valószínűségi változóra tekinthetünk úgy, mint ami az $\Omega^{(n)} := \{0, 1\}^{n+1}$ halmazon van értelmezve. Bontsuk fel $\Omega^{(n)}$ -t két részre. Legyen $\Omega_0^{(n)} := \{\omega \in \Omega^{(n)} : \omega_{n+1} = 0\}$ és $\Omega_1^{(n)} := \{\omega \in \Omega^{(n)} : \omega_{n+1} = 1\}$ a 0-ra illetve az 1-re végződő $n + 1$ hosszú 0-1 sorozatok halmaza. Az eseménytér ilyen felbontásával az a_n -et definiáló integrál is két részre bontható. Legyen

$$b_n := \int_{\Omega_0^{(n)}} \prod_{k=1}^n \gamma(Y_k) d\mathbb{P}_n \text{ és } c_n := \int_{\Omega_1^{(n)}} \prod_{k=1}^n \gamma(Y_k) d\mathbb{P}_n,$$

ahol a $\mathbb{P}_n$ mérték a Z folyamat által meghatározott $n + 1$ és $1/2$ paraméterekkel leírható binomiális eloszlás $\Omega^{(n)}$ -en. Ekkor $a_n = b_n + c_n$, és igaz a következő rekurzió:

$$b_{n+1} = \frac{xb_n + yc_n}{2} \text{ és } c_{n+1} = \frac{yb_n + zc_n}{2}. \quad (5.3)$$

A rekurziót az magyarázza, hogy az $\omega \in \Omega_0^{(n+1)}$ elemek kétféleképpen állhatnak elő: az első esetben $\omega_{n+1} = 0$, tehát az utolsó előtti koordináta is 0, a másodikban $\omega_{n+1} = 1$. Ha például ω olyan, ahol az utolsó előtti és az utolsó koordináta is 0, akkor az utolsó koordinátát elhagyva kapott $\omega' \in \Omega_0^{(n)}$, és $Y_{n+1} = 0$, tehát az ilyen elemi eseményeken az integrál éppen $xb_n/2$. A többi esetre ugyanezt végiggondolva pedig megkapjuk a 5.3 rekurziót.

b_1 és c_1 könnyen kiszámolható, értékük $b_1 = x/4 + y/4$ és $c_1 = y/4 + z/4$. Bevezetve az

$$A = \begin{pmatrix} x & y \\ y & z \end{pmatrix}$$

mátrixot pedig egyszerűsödik a jelölés:

$$\begin{pmatrix} b_1 \\ c_1 \end{pmatrix} = \frac{1}{4}A$$

Sőt, a rekurzió is egyszerűbb alakban írható fel:

$$\begin{pmatrix} b_{n+1} \\ c_{n+1} \end{pmatrix} = \frac{1}{2}A \begin{pmatrix} b_n \\ c_n \end{pmatrix},$$

amiből azt kapjuk, hogy

$$\begin{pmatrix} b_n \\ c_n \end{pmatrix} = \frac{1}{2^{n+1}}A^n.$$

Ha feltesszük, hogy x, y, z nemnegatívak, és A -nak van két különböző valós sajátértéke, akkor A spektrálfelbontásából látjuk, hogy az a_n sorozat pontosan tart végtelenhez, ha az A mátrix nagyobb λ_1 sajátértéke nagyobb, mint 2. A karakterisztikus polinom gyökei

$$\lambda_{1,2} = \frac{x + z \pm \sqrt{(x - z)^2 + y^2}}{2},$$

így $\lambda_1 > 2$ pontosan akkor teljesül, ha

$$2x + 2z - xz + y^2 - 4 > 0. \quad (5.4)$$

Ezt a feltételt több megengedett γ függvény is teljesíti. γ várható értéke lineáris függvénye (x, y, z) -nek, így a minimális várható értéket adó ponthármas nemnegatív x, y, z esetén a (5.4) egyenlettel megadott halmaz határán van. Egy ilyen globális minimumhely az $x = 0$, $y = 0$, $z = 2$. Ehhez figyeljük meg, hogy (5.2) miatt minden minimumhelyen $y = 0$, mivel ha nem így lenne, akkor y értékét csökkentve, majd x vagy z értékét ugyanennyivel növelve

a (5.4) feltétel továbbra is teljesül, de γ várható értéke csökkent. Az $y = 0$ esetben (5.2)-t nyilván úgy minimalizálhatjuk a (5.4) feltétel teljesítése mellett, ha vagy $x = 0$ vagy $z = 0$.

Ezen a minimumhelyen $\mathbb{E}[\gamma(Y_1)] = 1/2$. y -t tetszőlegesen kicsi pozitív számnak választva kapjuk, hogy olyan γ függvény is van, amelyre $\mathbb{E}[\gamma(Y_1)]$ $1/2$ -hez tetszőlegesen közel van, de γ nem teljesíti a hosszú távú kontraktivitási feltételt.

5.2. Példák Markov láncokra véletlen környezetben

Lovas és Rásonyi [8] cikkében három példát is láthatunk az 5.4 és 5.5 tételek alkalmazhatóságra. Mi további két példát mutatunk olyan Markov láncokra véletlen környezetben, amelyekre teljesülnek ezek a tételek. Ehhez azt fogjuk ellenőrizni, hogy a példákban szereplő Markov láncok teljesítik-e az 5.1, 5.2 és 5.3 drift, hosszú távú kontraktivitás és minorizálás feltételeket. Mindkét példánkban a bizonyítások a [8] cikkben megtalálható lineáris rendszer példájához hasonlóan történnek. Környezet alatt továbbra is egy $(Y_k)_{k \in \mathbb{Z}}$, $\mathcal{Y}$ -értékű stacionárius folyamatot értünk.

5.2.1. Bilineáris rendszer véletlen környezetben

Legyenek adottak az $a, b, c : \mathcal{Y} \rightarrow \mathbb{R}$ mérhető függvények, és egy $\varepsilon = (\varepsilon_n)_{n \in \mathbb{N}}$ független, azonos eloszlású sztochasztikus folyamat. Feltesszük azt is, hogy ε független a környezettől is. Az a, b és c függvények és ε a környezettel együtt meghatároznak egy $X = (X_n)_{n \in \mathbb{N}}$ Markov láncot véletlen környezetben, amit bilineáris rendszernek nevezünk véletlen környezetben és a következőképpen definiálunk:

$$X_{n+1} = a(Y_n)X_n + b(Y_n)X_n\varepsilon_{n+1} + c(Y_n)\varepsilon_{n+1}. \quad (5.5)$$

Ez a modell a bilineáris rendszer általánosítása, ami csak abban különbözik (5.5)-től, hogy a, b és c konstansok, így a környezetnek nincs szerepe.

A célunk, hogy olyan egyszerűen ellenőrizhető feltételeket adjunk meg az a, b és c függvényekre, valamint ε -ra, amelyekből levezethetők az 5.1, 5.2 és 5.3 feltételek. Ehhez szükség van a Q átmenet-valószínűségek pontos felírására. A (5.5) bilineáris rendszer Q átmenet-valószínűségei a definíciók alapján így írhatók fel:

$$Q(y, x, A) = \mathbb{P}[X_{n+1} \in A \mid Y_n = y, X_n = x] = \mathbb{P}[a(y)x + b(y)x\varepsilon_1 + c(y)\varepsilon_1 \in A].$$

Mivel $Q(y, x, \cdot)$ nem más, mint ε_1 egy függvényének eloszlása, ezért a $Q(y)V$ függvény is egyszerűen kiszámolható:

$$Q(y)V(x) = \int_{\mathbb{R}} V(z)Q(y, x, dz) = \mathbb{E}[V(a(y)x + b(y)x\varepsilon_1 + c(y)\varepsilon_1)].$$

Ez alapján úgy gondoljuk, hogy $V = |x|$ jó választás lehet a 5.1 drift feltétel igazolásához.

5.6. Lemma. Ha $\mathbb{E}[|\varepsilon_1|] < \infty$ és minden y -ra $|a(y)| + |b(y)| > 0$, akkor teljesül a 5.1 drift feltétel a $V(x) = |x|$, $\gamma(y) = |a(y)| + |b(y)|\mathbb{E}[|\varepsilon_1|]$ és $K(y) = \max\{1, |c(y)|\mathbb{E}[|\varepsilon_1|]\}$ függvényekkel.

Bizonyítás. A háromszög egyenlőtlenség felhasználásával kapjuk, hogy

$$\begin{aligned} Q(y)V(x) &= \mathbb{E}[|a(y)x + b(y)x\varepsilon_1 + c(y)\varepsilon_1|] \leq \mathbb{E}[|a(y)x| + |b(y)x\varepsilon_1| + |c(y)\varepsilon_1|] \\ &= (|a(y)| + |b(y)|\mathbb{E}[|\varepsilon_1|])|x| + |c(y)|\mathbb{E}[|\varepsilon_1|] = \gamma(y)V(x) + K(y). \end{aligned}$$

□

Sajnos a 5.2 hosszú távú kontraktivitási feltétel teljesüléséhez nem tudtunk kényelmesen ellenőrizhető kellően általános feltételeket találni az a és b függvényekre, így azt kell feltennünk, hogy 5.2 teljesül a fenti γ függvényre. Ezt a konkrét modellben adott a és b függvények, ε folyamat és az $(Y_n)_{n \in \mathbb{Z}}$ környezet esetén külön kell ellenőrizni.

A minorizálási feltétel ellenőrzéséhez további feltételeket teszük az ε folyamatra és az a, b, c függvényekre. Ezeket egy lemmában foglalkozunk össze.

5.7. Lemma. Tegyük fel, hogy ε_1 abszolút folytonos a λ Lebesgue-mértékre, és az f sűrűségfüggvényének van pozitív alsó korlátja minden korlátos halmazon. A 5.3 minorizálási feltétel teljesül, ha a, b és c korlátos, $\inf_{y \in \mathcal{Y}} |a(y)| + |b(y)| > 0$ és

$$\inf_{x \in [-K, K], y \in \mathcal{Y}} |b(y)x + c(y)| > 0,$$

ahol $K = \sup_{y \in \mathcal{Y}} R(y)$, $R(y)$ pedig a 5.3 feltételben bevezetett függvény.

Bizonyítás. Először azt látjuk be, hogy $K < \infty$, azaz (5.1)-ben x egy korlátos halmazba esik. A c függvény korlátossága miatt $K(y)$ is korlátos. $\inf_{y \in \mathcal{Y}} |a(y)| + |b(y)| > 0$ miatt pedig $1/\gamma(y)$ is korlátos, ezért $R(y) = K(y)/\delta\gamma(y)$ szintén korlátos tetszőleges $\delta > 0$ esetén. Emiatt $V^{-1}([0, R(y)]) = [-R(y), R(y)]$ lefedhető egy megfelelően nagy, y -től nem függő $[-K, K]$ intervallummal.

Vezessük be a $g_{xy}(t) = a(y)x + b(y)xt + c(y)t$ függvényt. Legyen $B \in \mathcal{B}(\mathbb{R})$ tetszőleges, és $A = B \cap [0, 1]$. Egyszerű számolással kapjuk, hogy

$$\begin{aligned} Q(y, x, B) &= \mathbb{P}[a(y)x + b(y)x\varepsilon_1 + c(y)\varepsilon_1 \in B] \geq \mathbb{P}[a(y)x + b(y)x\varepsilon_1 + c(y)\varepsilon_1 \in A] = \\ &= \mathbb{P}[g_{xy}(\varepsilon_1) \in A] = \mathbb{E}[\mathbb{I}_A(g_{xy}(\varepsilon_1))] = \int_{\mathbb{R}} \mathbb{I}_A(g_{xy}(t))f(t)\lambda(dt) = \int_{g_{xy}^{-1}(A)} f(t)\lambda(dt). \end{aligned}$$

Az $A + u = \{u + v : v \in A\}$ és a $uA = \{uv : v \in A\}$ jelölések segítségével $g_{xy}^{-1}(A)$ így írható fel:

$$g_{xy}^{-1}(A) = \frac{A - a(y)x}{b(y)x + c(y)}.$$

Az a, b, c függvényekre tett feltételeink és x korlátossága miatt van olyan korlátos H halmaz, amelyre $g_{xy}^{-1}(A) \subset H$ és H nem függ x -től, y -től és A -tól, ezért

$$\int_{g_{xy}^{-1}(A)} f(t) \lambda(dt) \geq \lambda(g_{xy}^{-1}(A)) \inf_{t \in H} f(t) \geq \lambda(A) \inf_{t \in H} f(t) \inf_{x \in [-K, K], y} \frac{1}{|b(y)x + c(y)|}.$$

Az utolsó lépésben a Lebesgue mérték eltolás-invariáns tulajdonságát használtuk fel, illetve azt, hogy $\lambda(uA) = |u|\lambda(A)$. Az ε_1 sűrűségfüggvényére vonatkozó feltétel miatt pedig

$$\alpha := \inf_{t \in H} f(t) \inf_{x \in [-K, K], y} \frac{1}{|b(y)x + c(y)|} \in (0, 1),$$

vagyis a $[0, 1]$ intervallumra megszorított Lebesgue-mérték α -szorosa megfelelő lesz minorizáló mértéknek. $\square$

A véletlen környezetbe helyezett bilineáris rendszer stabilitásával kapcsolatos megállapításainkat egy tételben foglaljuk össze.

5.8. Tétel. A (5.5) bilineáris rendszerre teljesül a 5.4 tétel, és ergodikus környezet esetén a 5.5 tétel, ha

- (i) $\mathbb{E}[|\varepsilon_1|] < \infty$, ε_1 abszolút folytonos a Lebesgue-mértékre, és a sűrűségfüggvényének van pozitív alsó korlátja minden korlátos halmazon.
- (ii) $\limsup_{n \rightarrow \infty} \mathbb{E}^{1/n} [\Pi_{i=1}^n |a(Y_i)| + |b(Y_i)| \mathbb{E}[|\varepsilon_1|]] < 1$
- (iii) a, b és c korlátosak
- (iv) $\inf_{y \in \mathcal{Y}} |a(y)| + |b(y)| > 0$
- (v) $\inf_{x \in [-K, K], y \in \mathcal{Y}} |b(y)x + c(y)| > 0$, ahol $K = \sup_{y \in \mathcal{Y}} R(y)$

5.2.2. SETAR modell véletlen környezetben

Az idősorok elemzésének témaköréből ismert SETAR (self-exciting threshold autoregressive) modell az autoregresszív modell általánosítása. Az egyszerűség kedvéért a modellt mi itt csak a legegyszerűbb alakjában definiáljuk, ahol X_n valós értékű és csak X_{n-1} -en keresztül függ a korábbi állapotoktól. Tong [11] cikkében további változatok és számos példa is megtalálható.

Legyenek $-\infty = t_0 < t_1 < \dots < t_m = \infty$ adottak. Az $(X_n)_{n \in \mathbb{N}}$ folyamat SETAR modell, ha

$$X_{n+1} = a_i X_n + b_i \varepsilon_n, \text{ ha } X_n \in (t_{i-1}, t_i], \quad (5.6)$$

ahol $a_1, a_2, \dots, a_m$ és $b_1, b_2, \dots, b_m$ konstansok, $(\varepsilon_n)_{n \in \mathbb{N}}$ pedig egy független, azonos eloszlású folyamat.

Az (5.6) modellt általánosíthatjuk, ha az a_i és b_i konstansok helyett $\mathcal{Y} \rightarrow \mathbb{R}$ függvényeket írunk. Ezzel a módosítással (a környezet alatt továbbra is egy $(Y_k)_{k \in \mathbb{Z}}$ stacionárius folyamatot értve) az

$$X_{n+1} = a_i(Y_n)X_n + b_i(Y_n)\varepsilon_n, \text{ ha } X_n \in (t_{i-1}, t_i] \quad (5.7)$$

folyamat SETAR modell véletlen környezetben. Az (5.7) összefüggéssel definiált modell Markov lánc véletlen környezetben, így az a_i és b_i függvényekre valamint az $(\varepsilon_n)_{n \in \mathbb{Z}_+}$ folyamatra tett feltételek segítségével alkalmazhatjuk az X_n eloszlásának konvergenciáját biztosító 5.4 és 5.5 tetteleket. A feltételeinket egy tételben foglaljuk össze.

5.9. Tétel. Legyen $(X_n)_{n \in \mathbb{N}}$ SETAR modell véletlen környezetben. $(X_n)_{n \in \mathbb{N}}$ -re teljesül a 5.4 tétel, és ergodikus környezet esetén a 5.5 tétel, ha

- (i) $\mathbb{E}[|\varepsilon_1|] < \infty$, ε_1 abszolút folytonos a Lebesgue-mértékre, és a sűrűségfüggvényének van pozitív alsó korlátja minden korlátos halmazon.
- (ii) $\limsup_{n \rightarrow \infty} \mathbb{E}^{1/n} \left[\prod_{j=1}^n \max_{1 \leq i \leq m} |a_i(Y_j)| \right] < 1$
- (iii) mindegyik a_i függvény korlátos
- (iv) $\inf_{y \in \mathcal{Y}} \max_{1 \leq i \leq m} |a_i(y)| > 0$
- (v) mindegyik b_i függvényre $0 < k \leq |b_i| \leq K < \infty$

Bizonyítás. A tétel bizonyítása a bilineáris rendszerre vonatkozó 5.8 tételnél látotthoz hasonlóan történik. Először a Q kernelt írjuk fel. Ebben segít, ha a modellt definiáló (5.7) egyenletet a $(t_{i-1}, t_i]$ intervallumok indikátorainak segítségével írjuk fel:

$$X_{n+1} = \sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(X_n) \left(a_i(Y_n)X_n + b_i(Y_n)\varepsilon_n \right).$$

Ebből

$$Q(y, x, A) = \mathbb{P}[X_{n+1} \in A \mid X_n = x, Y_n = y] = \mathbb{P} \left[\sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(x) \left(a_i(y)x + b_i(y)\varepsilon_n \right) \in A \right].$$

Miután a Q kernelt már ismerjük, $[Q(y)V](x)$ -et számoljuk ki, először tetszőleges V függvényre:

$$[Q(y)V](x) = \mathbb{E} \left[V \left(\sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(x) \left(a_i(y)x + b_i(y)\varepsilon_n \right) \right) \right].$$

Ahogy a 5.8 tétel esetén $V(x) = |x|$ itt is kényelmes választásnak tűnik, így ezzel a V függvénnyel folytatjuk a számolást.

Először a drift feltételt igazoljuk:

$$\begin{aligned}
[Q(y)V](x) &= \mathbb{E} \left[\left| \sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(x) (a_i(y)x + b_i(y)\varepsilon_n) \right| \right] \\
&\leq |x| \sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(x) |a_i(y)| + \mathbb{E}[|\varepsilon_1|] \sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(x) |b_i(y)| \\
&\leq |x| \max_{1 \leq i \leq m} |a_i(y)| + \mathbb{E}[|\varepsilon_1|] \sum_{i=1}^m |b_i(y)|,
\end{aligned}$$

amivel a 5.1 drift feltételt beláttuk a $\gamma(y) = \max_{1 \leq i \leq m} |a_i(y)|$ és $K(y) = \max\{1, \sum_{i=1}^m |b_i(y)|\}$ függvényekkel.

A minorizáláshoz vegyük észre, hogy a b_i függvények korlátossága miatt $K(y)$ is korlátos. A (iii)-as feltétel második része miatt $1/\gamma(y)$ is korlátos, így $\sup_{y \in \mathcal{Y}} R(y) < \infty$, és ezért

$V^{-1}([0, R(y)])$ korlátos.

Vezessük be a

$$g_{xy}(t) = \sum_{i=1}^m \mathbb{I}_{(t_{i-1}, t_i]}(x) (a_i(y)x + b_i(y)t)$$

függvényt.

A 5.8 tétel bizonyításához hasonlóan itt is elég megmutatni, hogy minden $A \subset \mathcal{B}(\mathbb{R}) \cap [0, 1]$ halmazra $g_{xy}^{-1}(A)$ lefedhető egy x -től, y -től és A -tól nem függő korlátos halmazzal. Ezt az a_i és $1/b_i$ függvények korlátossága biztosítja, mert $g_{xy}(t) \in A$ elégséges feltétele, ha

$$t \in \frac{A - a_i(y)x}{b_i(y)} \quad \forall 1 \leq i \leq m.$$

□

Kiemelt jelölések

$\mathbb{N}$	a nemnegatív egészek halmaza
$\mathbb{N}_+$	a pozitív egészek halmaza
$\bar{\mathbb{R}}$	a bővített valós számok halmaza
$(\mathcal{X}, \mathcal{B})$	tetszőleges mérhető tér
$\mathcal{X}_+$	az $f : \mathcal{X} \rightarrow [0, \infty]$ nemnegatív mérhető függvények halmaza
$b\mathcal{X}$	az $f : \mathcal{X} \rightarrow \mathbb{R}$ korlátos függvények halmaza
$\mathcal{M}(\mathcal{X})$	az $(\mathcal{X}, \mathcal{B})$ téren értelmezett σ -véges mértékek halmaza
$\mathcal{M}_1(\mathcal{X})$	az $(\mathcal{X}, \mathcal{B})$ téren értelmezett valószínűségi mértékek halmaza
$ \mu $	a μ mérték totális variációja
$\mu(f)$	az f függvény μ mérték szerinti integrálja
$\ \mu - \lambda\ _{TV}$	a μ és λ mértékek totális variációs távolsága
$\mathcal{L}^p(\mathbb{P}), \mathcal{L}^p(\Omega, \mathcal{A}, \mathbb{P})$	az $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mezőn értelmezett olyan $f : \Omega \rightarrow \bar{\mathbb{R}}$ mérhető függvények halmaza, amelyekre $\int_{\Omega} f ^p d\mathbb{P} < \infty$
$\text{ess sup } f$	az f függvény lényeges felső korlátja, ami azoknak a c számoknak az infimuma, amelyekre $\mathbb{P}[f \leq c] = 1$

Irodalomjegyzék

- [1] Krishna B Athreya and Sastry G Pantula. Mixing properties of Harris chains and autoregressive processes. *Journal of applied probability*, 23(4):880–892, 1986.
- [2] Randal Douc, Eric Moulines, Pierre Priouret, and Philippe Soulier. *Markov chains*. Springer, 2018.
- [3] Paul Doukhan. *Mixing*. Springer, 1994.
- [4] Paul Doukhan, Pascal Massart, and Emmanuel Rio. The functional central limit theorem for strongly mixing processes. In *Annales de l’IHP Probabilités et statistiques*, volume 30, pages 63–82, 1994.
- [5] Martin Hairer and Jonathan C Mattingly. Yet another look at Harris’ ergodic theorem for Markov chains. pages 109–117, 2011.
- [6] Galin L Jones. On the Markov chain central limit theorem. *Probability surveys*, 1:299–320, 2004.
- [7] Achim Klenke. *Probability theory: a comprehensive course*. Springer, third edition, 2020.
- [8] Miklós Lovas, Attila és Rásonyi. Markov chains in random environment with applications in queuing theory and machine learning. *Stochastic Processes and their Applications*, 137:294–326, 2021.
- [9] Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, second edition, 2009.
- [10] Daniel Revuz. *Markov chains*. North-Holland Mathematical Library, 1984.
- [11] Howell Tong. Threshold models in time series analysis—30 years on. *Statistics and its Interface*, 4(2):107–118, 2011.
- [12] I. A. Ibragimov és Yu. A. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff, 1971.