# REJTETT MARKOV MODELLEK II.

Gerencsér László   Michaletzky György   Molnár-Sáska Gábor
Vágó Zsuzsanna

(Oktatási segédlet)

Budapest, Eötvös Loránd Tudományegyetem
Valószínűségelméleti és Statisztika Tanszék

2002. április

# 1. Véletlen mátrixok szorzata I.

**1.1. Tétel.** *Fürstenberg-Kesten: Legyen $X_1, \ldots, X_n$ véletlen mátrixok olyan stacionárius ergodikus sorozata, melyre $E \log^+ \|X_1\| < \infty$. Ekkor*

$$\frac{1}{n} \log \|X_n \ldots X_1\| \longrightarrow \lambda$$

*1-valószínűséggel*

**1.1. Megjegyzés.** *$\lambda$ akár $\infty$ is lehet*

A tételt több lemmán keresztül fogjuk bizonyítani.

**1.1. Definíció.** *Az $a_n$ sorozat szubadditív, ha $a_{n+m} \leq a_n + a_m$.*

**1.1. Lemma.** *Ha az $a_n$ sorozat szubadditív, akkor*

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\to\infty} \frac{a_n}{n}$$

*létezik (akár $-\infty$ is lehet)*

*Bizonyítás.* Legyen először

$$\gamma = \inf_{n\to\infty} \frac{a_n}{n} > -\infty.$$

Az infimum definíciója miatt ekkor rögzített $\epsilon > 0$-ra $\exists k$, hogy

$$\frac{a_k}{k} < \gamma + \epsilon.$$

A szubadditívitás miatt ekkor $a_{nk} \leq n a_k$, azaz

$$\frac{a_{nk}}{nk} \leq \frac{a_k}{k} < \gamma + \epsilon.$$

Másfelől tetszőleges elegendő nagy $m$-re ($m = nk + r$) $a_m \leq a_{nk} + a_r$ alakban írható, így

$$\frac{a_m}{m} \leq \frac{a_{nk}}{nk}\frac{nk}{m} + \frac{a_r}{m} < \gamma + 2\epsilon.$$

Ezzel a lemma állítását bebizonyítottuk.

**1.2. Lemma.** *Ha $X_k$ egy stacionárius mátrixfolyamat, akkor*

$$\lim_{n\to\infty} \frac{1}{n} E \log \|X_n \dots X_1\|$$

*határérték létezik. (A határértéket jelöljük $\lambda$-val.)*

*Bizonyítás.* Kihasználva az $X_n$ sorozat stacionaritását, az

$$a_n = E \log \|X_n \dots X_1\|$$

sorozat szubadditív, így a 1.1 lemma alapján kapjuk az állítást.

Most folytassuk a Fürstenberg-Kesten tétel bizonyítását a következő állítással.

**1.1. Állítás.** *A fenti feltételek mellett*

$$\limsup \frac{1}{n} \log \|X_n \dots X_1\| \leq \lambda$$

*Bizonyítás.* Legyen ${}^nY^1 = X_n \dots X_1$, valamint $N$ olyan nagy, hogy rögzített $\epsilon > 0$-ra $\frac{1}{N} E \log \|{}^nY^1\| < \lambda + \epsilon$ teljesüljön.

Legyen $U^k = \log \|{}^{k+N}Y^{k+1}\|$. $U^k$ egy stacionárius és ergodikus folyamat, és nyilván $U^k = U^0(T^k\omega)$.

Mivel $a_n = E \log \|X_n \dots X_1\|$ szubadditív folyamat és $a_1 = E \log \|X_1\| \leq E \log^+ \|X_1\| < \infty$, így $\lambda$ nem lehet $+\infty$.

Most tegyük fel, hogy $\lambda > -\infty$. Így $-\infty < EU^k < +\infty$.

Az ergod-tétel szerint

$$\lim_r \frac{1}{r} \sum_{k=0}^{r-1} \frac{1}{N} U^k = E \frac{1}{N} U^1 < \lambda + \epsilon,$$

ahol az egyenlőség 1-valószínűségű.

Most legyen $n = rN$. Bontsuk a $n$ tényezős szorzatot $r + 1$ részre az alábbi módon

$$\|X_n \dots X_1\| = \|(X_n \dots X_{n-N+d+1}) \dots (X_{d+N} \dots X_{d+1})(X_d \dots X_1)\|,$$

ahol $0 \leq d \leq N-1$.

Mindkét oldalnak véve a logaritmusát, kapjuk, hogy

$$\log \|X_n \dots X_1\| \leq \sum_{i=0}^{N-d-1} \log \|X_{n-i}\| + \sum_{i=0}^{r-2} U^{d+iN} + \sum_{i=1}^{d} \log \|X_i\| \leq$$

$$\leq \sum_{i=0}^{N-1} \log \|X_{n-i}\| + \sum_{i=0}^{r-2} U^{d+iN} + \sum_{i=1}^{N} \log \|X_i\|,$$

azaz

$$\frac{1}{n} \log \|X_n \dots X_1\| \leq \frac{1}{rN} \sum_{i=0}^{N-1} \log \|X_{n-i}\| + \frac{1}{rN} \sum_{i=0}^{r-2} U^{d+iN} + \frac{1}{r} \frac{\sum_{i=1}^{N} \log \|X_i\|}{N}$$

Tekintsük ezt az egyenlőtlenséget minden $d$-re, és adjuk össze a $N$ darab egyenlőtlenséget. Így kapjuk, hogy

$$\frac{N}{n} \log \|X_n \dots X_1\| \leq \frac{1}{r} \sum_{i=0}^{N-1} \log \|X_{n-i}\| + \frac{1}{n} \sum_{i=0}^{n-N-1} U^k + \frac{1}{r} \sum_{i=1}^{N} \log \|X_i\| \qquad (1)$$

Mivel $N$ rögzített, így az (1) összeg harmadik tagja $n \longrightarrow \infty$ (és így $r \longrightarrow \infty$) esetén nyilván 0-hoz tart. Most tekintsük az (1) összeg első tagját. Mivel egy $N$ tagú összegről van szó, ahol $N$ egy rögzített konstans, így elegendő azt megmutatni, hogy

$$\frac{\log \|X_{n-i}\|}{r} \longrightarrow 0,$$

ahol $0 \leq i \leq N-1$, amihez elég azt látni, hogy

$$\frac{\log \|X_n\|}{n} \longrightarrow 0.$$

Ehhez belátjuk, hogy ha $E \log^+ \|X_1\| < \infty$, akkor $n \geq n_0(\epsilon)$ esetén $\|X_n\| \leq e^{n\epsilon}$ 1-valószínűséggel.

Mivel

$$P(\|X_n\| > Ce^{\epsilon n}) = P(\log^+ \|X_n\| > c + \epsilon n) = P(\log^+ \|X_1\| > c + \epsilon n) = 1 - F(\epsilon n)$$

és így az $E \log^+ \|X_1\| < \infty$ feltételből kapjuk, hogy

$$\sum_{i=0}^{\infty} (1 - F(\epsilon n)) < \infty,$$

azaz a Borel-Cantelli lemma alapján az $\|X_n\| > Ce^{\epsilon n}$ csak véges sokszor teljesül, amivel a szükséges állítást bebizonyítottuk. (Felhasználtuk, hogy $\Psi \geq 0$ és $E\Psi < \infty$ esetén $E\Psi = \int_0^\infty (1 - F(x))dx$.)

Végül nézzük a (1) egyenlőtlenség jobb oldalának középső tagját.

$$\frac{1}{n} \sum_{i=0}^{n-N-1} U^k = \frac{n-N}{n} \quad \frac{1}{n-N} \sum_{i=0}^{n-N-1} U^k$$

A jobboldali szorzat első tényezője 1-hez tart, míg a második tényező a klasszikus ergod-tétel alapján $EU^1$-hez. Így $N$ választása alapján kapjuk, hogy

$$\frac{1}{n} \log \|X_n \dots X_1\| \leq \lambda + \epsilon.$$

A fenti egyenlőtlenség $\forall\epsilon$-ra fennáll, így kapjuk, hogy

$$\limsup \frac{1}{n} \log \|X_n \dots X_1\| \leq \lambda,$$

azaz az állítást bebizonyítottuk.

Folytassuk a Fürstenberg-Kesten tétel bizonyítását. Definiáljuk a $Z_n$ folyamatot a következő módon:

$$Z_n = \frac{X_n \dots X_1}{\|X_n \dots X_1\|}, \quad \text{ha} \quad \|X_n \dots X_1\| \neq 0, \quad \text{és}$$

$$Z_n = 0 \quad \text{egyébként.}$$

Így a $Z_n$ folyamatra a

$$Z_{n+1} = \frac{X_{n+1}Z_n}{\|X_{n+1}Z_n\|} \tag{2}$$

rekurziót kapjuk.

($n = 1$-re nem követeljük meg, hogy $Z_1$ normált legyen.)

Jelölje $\mathcal{X} = \{x_1, \dots\}$ az $X_n$ illetve $\mathcal{Z} = \{z_1, \dots\}$ a $Z_n$ folyamatok terét, továbbá legyen $P$ a folyamat generálta mérték $\mathcal{X}$-n.

Legyen $\Omega' = \mathcal{X} \times \mathcal{Z}$, és legyen $\Omega'_0 \subset \Omega'$ azon $(x, z)$ sorozatok halmaza, melyre (2) teljesül.

Mivel $X_n$ meghatározza $Z_n$-t, így az $(X_n, Z_n)$ eloszlást meghatározza a $P$ mérték.

Most definiáljuk a $T$ eltolásoperátort (shiftoperátor). Ha $x = (x_1, x_2, \dots)$ és $z = (z_1, z_2, \dots)$, akkor legyen

$$(T(x, z))_n = (x_{n+1}, z_{n+1}).$$

Végül definiáljuk mértékek egy családját a következő módon: Legyen $\pi : \mathcal{X} \times \mathcal{Z} \longrightarrow \mathcal{X}$ a természetes projekció,

$$\mu_1 = P \circ \pi,$$

és így tetszőleges $C \subset \Omega'$ részhalmazra legyen

$$\mu_k(C) = \mu_1(T^{-k+1}(C)).$$

**1.2. Állítás.** *A $\mu_k$ mértéksorozat feszes, azaz $\forall \epsilon > 0$-ra $\exists C \subset \Omega'$ kompakt halmaz, hogy $\mu_k(C) > 1 - \epsilon \quad \forall k$-ra.*

*Bizonyítás.* Definíció szerint $\mu_1$ képe $\pi$ alatt $P$, azaz $\mu_1 \circ \pi^{-1} = P$. Innen könnyen látható, hogy $P \circ T^{-k+1} = \mu_k \circ \pi^{-1}$, azaz általában $\mu_k$ vetülete $\pi$ alatt $P$.

Most legyen $C_X \subset \mathcal{X}$ olyan kompakt halmaz, melyre $P(C_X) > 1 - \epsilon$, illetve $C = C_X \times \mathcal{Z}$. Mivel kompakt halmazok szorzata is kompakt, és $\mathcal{Z}$ kompakt, így $C$ is kompakt halmaz lesz. Másfelől

$$\mu_k(C) = \mu_k(C_X \times \mathcal{Z}) =$$
$$= \mu_k(\pi^{-1}(C_X)) =$$
$$= P(T^{-k+1}(C_X)) = P(C_X) \geq 1 - \epsilon,$$

hiszen $P \circ T^{-k+1}$ stacionárius. Ezzel az állítást bebizonyítottuk.

Az állítás közvetlen következménye, hogy $\exists \mu_{u_i}$ részsorozata a mértékeknek, hogy $\mu_{u_i} \longrightarrow \mu$ gyengén, azaz minden korlátos és folytonos $f$ függvényre $\int_{\Omega'} f d\mu_{u_i} \longrightarrow \int_{\Omega'} f d\mu$.

Most legyen

$$\nu_n = \frac{1}{n} \sum_{k=1}^{n} \mu_k$$

az átlagmérték.

**1.3. Állítás.** *A $\nu_n$ mértékcsalád feszes.*

*Bizonyítás.* A $\mu_n$ mértékcsalád feszességére vonatkozó bizonyításban definiált $C$ halmaz megfelelő lesz olyan értelemben, hogy $\nu_n(C) > 1 - \epsilon$ teljesül.

A bizonyítás alapján $\nu_{u_i} \longrightarrow \mu$ gyengén.

**1.4. Állítás.** $\lambda = \int_{\Omega'} \log \|x_2 z_1\| d\mu$

*Bizonyítás.* Azt használjuk fel, hogy ha $\eta_n \longrightarrow \eta$ gyengén, $f$ folytonos függvény (nem feltétlenül korlátos), valamint $\eta_n$ egyenletesen integrálható, azaz $\forall \epsilon > 0$-hoz $\exists K$, hogy $\int_{\|f\| > K} \|f\| d\eta_n < \epsilon \quad \forall n$-re, akkor $\int f d\eta_n \longrightarrow \int f d\eta$ teljesül.

Ehhez először azt kell látni, hogy $\log^+ \|X_2 Z_1\|$ egyenletesen integrálható $\mu_n$ szerint. Mivel $\log^+ \|X_2 Z_1\| \leq \log^+ \|X_2\|$ és $\log^+ \|X_2\|$ eloszlása $\mu_k$ szerint $k$-tól független, így a rögzített $\epsilon$-hoz vegyük azt a $K$-t, melyre

$$\int_{\log^+ \|X_2\| > K} \log^+ \|X_2\| d\mu_k < \epsilon$$

teljesül (ez független $\mu_k$-tól), és ezzel

$$\int_{\Omega'} \log^+ \|X_2 Z_1\| d\nu_{u_i} \longrightarrow \int_{\Omega'} \log^+ \|X_2 Z_1\| d\mu$$

Másfelől a Fatou-lemma alapján tudjuk, hogy

$$\liminf \int_{\Omega'} \log^- \|X_2 Z_1\| d\nu_n \leq \int_{\Omega'} \log^- \|X_2 Z_1\| d\mu$$

Tehát kapjuk, hogy

$$\int_{\Omega'} \log \|X_2 Z_1\| d\mu \geq \liminf \int_{\Omega'} \log \|X_2 Z_1\| d\nu_{u_i}$$

Alakítsuk át az egyenlőtlenség jobb oldalát: $\nu_n$ definíciója alapján

$$\int_{\Omega'} \log \|X_2 Z_1\| d\nu_n = \frac{1}{n} \sum_{k=1}^{n} \int_{\Omega'} \log \|X_2 Z_1\| d\mu_k = \frac{1}{n} \int_{\Omega'} \sum_{k=1}^{n} \log \|X_{k+1} Z_k\| d\mu_1,$$

továbbá mivel

$$\log \|X_{k+1} Z_k\| = \log \|\frac{X_{k+1} \dots X_1}{\|X_k \dots X_1\|}\| = \log \|X_{k+1} \dots X_1\| - \log \|X_k \dots X_1\|,$$

így

$$\sum_{k=1}^{n} \log \|X_{k+1} Z_k\| = \log \|X_{n+1} \dots X_1\| - \log \|X_1\|,$$

azaz

$$\int_{\Omega'} \log \|X_2 Z_1\| d\nu_n = \frac{1}{n} \int_{\Omega'} \log \|X_{n+1} \dots X_1\| - \log \|X_1\| d\mu_1.$$

Az utolsó egyenlőség jobb oldala azonban már csak $X_n$-től függ, így $d\mu_1$ helyére írható $dP$ is. Tehát kaptuk, hogy

$$\int_{\Omega'} \log \|X_2 Z_1\| d\mu \geq \liminf \frac{1}{u_i} \int_{\mathcal{X}} \log \|X_{u_i+1} \dots X_1\| dP =$$

$$= \liminf \frac{1}{u_i} E_P \log \|X_{u_i+1} \dots X_1\| = \lambda,$$

azaz a bizonyítandó egyenlőségből egyenlőre még csak azt tudjuk, hogy

$$\int_{\Omega'} \log \|X_2 Z_1\| d\mu \geq \lambda \tag{3}$$

Most bizonyítsuk a fordított irányú egyenlőtlenséget.

Legyen $\xi_k = \log \|X_{k+1} Z_k\|$ $\Omega'$-n. Ekkor $\xi_k$ egy stacionárius folyamat $\mu$ szerint, hiszen $(T\xi)_k = \xi_{k+1}$, és a $\mu$ mérték $T$-invariáns.

Az ergod-tétel alapján $\liminf \frac{1}{n} \sum_{k=1}^{n} \xi_k = \psi$ létezik $\mu$-majdnem mindenütt, feltéve, hogy $E_\mu \xi_1$ véges. Ez utóbbi viszont teljesül, hiszen egyrészt $\log \|X_2 Z_1\| \leq \log \|X_2\|$ és $\int_{\mathcal{X} \times \mathcal{Z}} \log \|X_2\| d\mu = \int_{\mathcal{X}} \log^+ \|X_2\| dP < \infty$ a feltétel miatt, másfelől azt már láttuk, hogy $\int_{\mathcal{X} \times \mathcal{Z}} \log \|X_2 Z_1\| d\mu \geq \lambda$. Tehát az ergod-tétel azt mondja, hogy

$$\int_{\mathcal{X} \times \mathcal{Z}} \psi d\mu = E\xi_1 = \int_{\mathcal{X} \times \mathcal{Z}} \log \|X_2 Z_1\| d\mu \tag{4}$$

Az egyenlőség bal oldalán elég a $\mu$ tartóján integrálni (a legszűkebb zárt halmazon, amelynek a $\mu$ mértéke 1), azaz a $\mathrm{supp}\mu_k \subset \Omega_0'$ halmazon. Felhasználjuk azt az állítást, hogy ha $\eta_k \longrightarrow \eta$ gyengén és $F$ zárt halmaz, akkor $\eta(F) \geq \limsup \eta_k(F)$. Mivel $\Omega_0'$ zárt halmaz $\mathcal{X} \times \mathcal{Z}$-n és $\mu_k(\Omega_0') = 1$, így $\mu(\Omega_0') \geq 1$, azaz $\mu(\Omega_0') = 1$.

Az ergod-tétel miatt

$$\psi = \lim \frac{1}{n} \sum_{k=1}^{n} \log \|X_{k+1} Z_k\| =$$

$$= \lim \frac{1}{n} \log \| \frac{X_{n+1} \dots X_1}{\|X_1\|} \leq \liminf \frac{1}{n} \log \|X_{n+1} \dots X_2\|$$

$\mu$-majdnem mindenütt. De a jobboldal ismét csak az $X$ függvénye, így

$$\int_{\mathcal{X} \times \mathcal{Z}} \psi d\mu \leq \int_{\mathcal{X} \times \mathcal{Z}} \liminf \frac{1}{n} \log \|X_{n+1} \dots X_2\| d\mu \leq$$

$$\leq \int_{\mathcal{X}} \liminf \frac{1}{n} \log \|X_{n+1} \dots X_2\| dP \leq \int_{\mathcal{X}} \lambda dP = \lambda,$$

ahol az utolsó egyenlőtlenség a 1.1 állítás miatt teljesül.

Tehát kaptuk, hogy

$$\int_{\mathcal{X} \times \mathcal{Z}} \psi d\mu \leq \lambda.$$

Összevetve (3)-mal

$$\int_{\mathcal{X} \times \mathcal{Z}} \psi d\mu = \lambda,$$

ami (4) miatt éppen a bizonyítandó állítás.

# 2. Szubadditív folyamatok

Tekintsük most a $g_n = \log \|X_n \dots X_1\|$ sztochasztikus folyamatot, ahol $X_n$ egy stacionárius mátrixfolyamat, $X_n = T^{n-1} X_1$ és $T$ egy mátrixtranszformáció az $(\Omega, \mathcal{F}, P)$ alaptéren. Ekkor nyilván

$$\log \|X_{n+m} \dots X_1\| \leq \log \|X_{n+m} \dots X_{n+1}\| + \log \|X_n \dots X_1\|,$$

azaz

$$g_{n+m}(\omega) \leq g_n(T^m \omega) + g_m(\omega)$$

**2.1. Definíció.** *Azt mondjuk, hogy egy $g_n$ sztochasztikus folyamat szubadditív, ha $\exists T$ mértéktartó transzformáció, hogy $\forall n, m \geq 1$-re*

$$g_{n+m}(\omega) \leq g_n(\omega) + g_m(T^n \omega).$$

**2.1. Tétel.** *(Szubadditív ergodtétel-Kingman, Hammersley) Legyen $g_n$ szubadditív sztochasztikus folyamat, és $E|g_1| < \infty$. Ekkor*

$$\lim_{n \to \infty} \frac{g_n(\omega)}{n} = g(\omega) \geq -\infty$$

*létezik 1-valószínűséggel, és itt $g$ $T$-invariáns függvény.*

*Bizonyítás.* Legyen
$$g'_n(\omega) = g_n(\omega) - \sum_{i=0}^{n-1} g_1(T^i\omega)$$
Ekkor $g_n$ szubadditivitása miatt
$$g_n(\omega) \le g_1(\omega) + g_1(T\omega) + \cdots + g_1(T^{n-1}\omega),$$
és így $g'_n(\omega) \le 0$. Másfelől $g'_n(\omega)$ nyilván szubadditív. A Birkhoff féle ergod-tétel alapján tudjuk, hogy
$$\frac{1}{n}\sum_{i=0}^{n-1} g_1(T^i\omega) \longrightarrow h(\omega)$$
1-valószínűséggel, ahol $h$ $T$-invariáns, méghozzá előáll
$$h = E(g_1|\mathcal{F}_0)$$
alakban, ahol $\mathcal{F}_0$ jelöli a $T$-invariáns halmazok $\sigma$-algebráját.

Tehát elegendő a tétel állítását olyan $g_n$ szubadditív folyamatra igazolni, mely nem pozitív.

Legyen
$$g(\omega) = \liminf_n \frac{g_n(\omega)}{n}$$
Először belátjuk, hogy az így definiált $g$ függvény $T$-invariáns. $g_n$ szubadditivitása miatt
$$g_{n+1}(\omega) \le g_1(\omega) + g_n(T\omega).$$
Mindkét oldalnak véve a limesz inferiorját,
$$g(\omega) \le g(T\omega),$$
azaz
$$\{\omega : g(\omega) > \alpha\} \subset \{\omega : g(T\omega) > \alpha\} = T^{-1}\{y : g(y) > \alpha\}.$$
Tehát a
$$\{\omega : g(T\omega) > \alpha\} \backslash \{\omega : g(\omega) > \alpha\}$$
egy 0-mértékű halmaz, és így
$$\{\omega : g(T\omega) \ne g(\omega)\}$$
szintén 0-mértékű, ami $g$ $T$-invarianciáját mutatja.

Most lássuk be, hogy a limesz inferiorként definiált $g$ igazából limeszként is előáll.

Legyen $N$ egy rögzített természetes, illetve $\epsilon > 0$ egy rögzített valós szám. Nevezzük a $B_N \subset \Omega$ halmazt "rossznak", melyre

$$B_N = \{\omega : \frac{g_l(\omega)}{l} > g(\omega) + \epsilon \quad \forall\, 1 \leq l \leq N\text{-re}\},$$

illetve legyen a $B_N^C = A_N$ "jó" halmaz. Tetszőleges $\omega \in \Omega$-ra felosztjuk az $[1, n)$ intervallumot "jó" és "rossz" intervallumokra a következőképpen: Legyen $k$ a legkisebb még le nem fedett időpont. Ha $T^k\omega \in A_N$, akkor $\exists 1 \leq l \leq N$, hogy

$$\frac{g_l(T^k\omega)}{l} \leq g(T^k\omega) + \epsilon,$$

és ekkor legyen a $[k, k+l)$ egy "jó" intevallum, feltéve, hogy $k+l \leq n$. Ha $k+l > n$ vagy $T^k\omega \in B_N$, akkor a $[k, k+1)$ intervallum legyen "rossz". Jelöljük a "jó" intervallumokat $[\tau_i, \tau_i + l_i)$-vel, ahol $1 \leq l_i \leq N$, illetve a "rossz" intervallumokat $[\sigma_i, \sigma_i + 1)$-gyel. Azokat az egységnyi "rossz" intervallumokat, melyek azért rosszak, mert a $k + l > n$ teljesült, különböztessük meg a többi "rossz" intervallumtól. Ezek legfeljebb az utolsó $w \leq N$ egységnyi szakaszon lehetnek. $g_n$ szubadditivitása miatt a fenti jelölések mellett

$$g_n(\omega) \leq \sum_i g_{l_i}(T^{\tau_i}\omega) + \sum_j g_1(T^{\sigma_j}\omega) + \sum_{k=1}^{w} g_1(T^{n-k}\omega).$$

Mivel $g_n \leq 0$, így

$$g_n(\omega) \leq \sum_i g_{l_i}(T^{\tau_i}\omega).$$

Mivel

$$g_{l_i}(T^{\tau_i}\omega) \leq l_i(g(T^{\tau_i}\omega) + \epsilon),$$

és $g$ $T$-invariáns, így kapjuk, hogy

$$g_n(\omega) \leq (g(\omega) + \epsilon) \sum_i l_i \leq g(\omega) \sum_i l_i + n\epsilon.$$

A bizonyítás további részében azt látjuk be, hogy a "jó" intervallumok kitöltik $[1, n)$-t, azaz $\sum_i l_i \sim n$. A fentiek alapján

$$\sum_i l_i \geq n - \sum_{k=1}^{n} \mathbf{1}_{B_N}(T^k\omega) - N.$$

Birkhoff ergod-tétele szerint így

$$\liminf \frac{1}{n} \sum_i l_i \geq 1 - E(\mathbf{1}_{B_N}|\mathcal{F}_0)$$

majdnem mindenütt, ahol $\mathcal{F}_0$ jelöli a $T$-invariáns halmazok $\sigma$-algebráját. Továbbra is figyelembe véve, hogy $g_n \leq 0$, kapjuk, hogy

$$\limsup \frac{g_n(\omega)}{n} \leq g(\omega)(1 - E(\mathbf{1}_{B_N}|\mathcal{F}_0)) + \epsilon. \tag{5}$$

Ha $N \longrightarrow \infty$, akkor $\mathbf{1}_{B_N} \longrightarrow 0$ 1-valószínűséggel, hiszen $\mathbf{1}_{B_N}(\omega) = 0$ ha $N \geq N_0(\omega)$, és így $E(\mathbf{1}_{B_N}|\mathcal{F}_0) \longrightarrow 0$ 1-valószínűséggel. (5) egyenlőtlenséget tetszőleges $\epsilon > 0$ esetén tekintve kapjuk, hogy

$$\limsup \frac{g_n(\omega)}{n} \leq g(\omega) \quad \text{majdnem minden} \quad \omega\text{-ra},$$

amit $g$ definíciójával összevetve kapjuk a bizonyítandó állítást.

A 1.2 lemma és a 2.1 tétel kapcsolatát mutatja a következő állítás.

**2.1. Állítás.** *Legyen $g_n$ szubadditív sztochasztikus folyamat, és $E|g_1| < \infty$. Ekkor*

$$Eg = \inf \frac{Eg_n}{n} \quad (= \lim \frac{Eg_n}{n}).$$

*Bizonyítás.* Legyen

$$\gamma = \inf E\frac{g_n}{n},$$

és tegyük fel, hogy $\gamma > -\infty$. Tekintsük a

$$\overline{g}_n(\omega) = g_n(\omega) - \sum_{k=0}^{n-1} g_1(T^k\omega)$$

folyamatot. Mivel $g_n$ szubadditív folyamat, így

$$g_n(\omega) \leq \sum_{k=0}^{n-1} g_1(T^k\omega),$$

azaz $\overline{g}_n(\omega) \leq 0$, továbbá $\overline{g}_n(\omega)$ maga is szubadditív, hiszen egy additív és egy szubadditív folyamat összege. Mivel

$$\lim \frac{\overline{g}_n}{n} = \lim \frac{g_n}{n} - \lim \frac{1}{n} \sum_{k=0}^{n-1} g_1(T^k\omega),$$

így a hagyományos ergod-tétel alapján elegendő az állítást $\overline{g}_n$-ra, azaz nem pozitív $g_n$-re igazolni. 2.1 tétel alapján $\lim \frac{g_n}{n} = g$, így a Fatou-lemma alapján

$$\liminf E\frac{g_n}{n} \leq Eg.$$

Másfelől $g_n$ szubadditivitása miatt

$$g_{nk} \leq \sum_{i=0}^{n-1} g_k(T^{ik}\omega).$$

Mindkét oldalt szorozva $\frac{1}{nk}$-val és véve $n$-ben a határértéket kapjuk, hogy

$$g = \lim \frac{g_{nk}}{nk} \leq \lim \frac{1}{n}\sum_{i=0}^{n-1} \frac{g_k(T^{ik}\omega)}{k}.$$

Vegyük mindkét oldalnak a várhatóértékét, használjuk ki $g_n$ nem pozitivitását, valamint a hagyományos ergod tétel alapján rögzített $\epsilon$ esetén ha $k$ elég nagy, akkor minden tagra

$$E\frac{g_k(T^{ik}\omega)}{k} \leq \gamma + \epsilon,$$

és így

$$Eg \leq \gamma + \epsilon,$$

azaz az állítást bebizonyítottuk.

A 2.1 tétel és a 2.1 állítás valamivel gyengébb feltételek mellett is kimondható.

**2.2. Tétel.** *Legyen $g_n$ szubadditív sztochasztikus folyamat, $Eg_1^+ < \infty$. Ekkor*

$$g = \lim_{n\to\infty} \frac{g_n}{n}$$

*létezik 1-valószínűséggel, és*

$$Eg = \lim_{n\to\infty} \frac{Eg_n}{n} = \inf_n \frac{Eg_n}{n}.$$

*Bizonyítás.* Először belátjuk, hogy ha $g_n$ és $h_n$ szubadditív folyamatok, akkor $\max(g_n, h_n)$ is szubadditív folyamat lesz. A szubadditivitás definíciója alapján

$$g_{n+m}(\omega) \leq g_n(\omega) + g_m(T^n\omega), \quad \text{illetve}$$

$$h_{n+m}(\omega) \leq h_n(\omega) + h_m(T^n\omega).$$

Így

$$\max(g_{n+m}(\omega), h_{n+m}(\omega)) \leq \max(g_n(\omega), h_n(\omega)) + \max(g_m(T^n\omega), h_m(T^n\omega)),$$

ami igazolja az állítást.

Most legyen $K$ egy tetszőlegesen rögzített szám, és $h_n = -nK$. Ekkor nyilván

$$h_{n+m} = h_n + h_m.$$

Tekintsük a

$$g_n^K = \max(g_n, -nK)$$

folyamatot. Az előzőek alapján ez is szubadditív folyamat lesz, $(g_1^K)^+ = g_1^+$ miatt $E(g_1^K)^+ < \infty$, valamint $g_1^K$ definíciója miatt $Eg_1^K > -\infty$ is teljesül.

Tehát $g_1^K$ integrálható, azaz a szubadditív ergod-tétel alapján (2.1 tétel) minden rögzített $K$-ra 1-valószínűséggel létezik a

$$\lim \frac{g_n^K}{n} = \lim \max(\frac{g_n}{n}, -K) = \xi^K$$

határérték. Ekkor azonban a $\lim \frac{g_n}{n}$ határérték is létezik 1-valószínűséggel, ugyanis indirekten tegyük fel, hogy $\limsup \frac{g_n}{n} > \liminf \frac{g_n}{n}$, és így $-\limsup g_n n < K$ választással azt kapnánk, hogy a $\xi^K$ határérték nem létezik.

Tehát legyen

$$\lim \frac{g_n}{n} = \xi, \quad \text{és} \quad \xi^K = \max(\xi, -K)$$

Azt kell belátni, hogy $E\xi = \gamma$, ahol $\gamma = \inf \frac{Eg_n}{n}$.

Ha $Eg_1 > -\infty$, akkor ez nyilvánvaló a 2.1 tétel alapján. Tehát tegyük fel, hogy $Eg_1 = -\infty$, és ekkor persze $\gamma = -\infty$. Nyilván

$$E\xi \leq E\xi^K = \inf \frac{Eg_n^K}{n} \leq Eg_1^K.$$

Ha $K \longrightarrow \infty$, akkor $g_1^K \longrightarrow g_1$ monoton csökkenően, így $E(g_1^K)^+ < \infty$ és a Lebesgue-tétel miatt $Eg_1^K \longrightarrow Eg_1$. Tehát kaptuk, hogy

$$E\xi \leq Eg_1,$$

azaz $E\xi = -\infty$ teljesül. Ezzel a tétel állítását bebizonyítottuk.

Végül alkalmazzuk a szubadditív folyamatokra vonatkozó tételeinket a Fürstenberg-Kesten tételre (1.1 tétel).

*Bizonyítás.* A

$$g_n(\omega) = \log \|X_1 \dots X_n\|$$

sorozat szubadditív, és a tétel feltételei szerint $Eg_1^+ < \infty$. A 2.2 tétel szerint így

$$g = \lim_{n\to\infty} \frac{g_n}{n}$$

létezik 1-valószínűséggel, és

$$Eg = \lim_{n\to\infty} \frac{Eg_n}{n} = \inf_n \frac{Eg_n}{n}.$$

Mivel $T$ ergodikus és $g$ $T$-invariáns, így $g$ nyilván konstans, azaz $g = \lambda$.

# 3. Külső szorzatok

Tekintsük az $\mathbb{R}^3$ vektorteret és benne egy $\mathbf{x}, \mathbf{y}, \mathbf{z}$ ortonormált bázist.

**3.1. Definíció.** *Legyen $Ad\mathbf{x} + Bd\mathbf{y} + Cd\mathbf{z}$ egy vonalelem, illetve $(Ad\mathbf{x} + Bd\mathbf{y} + Cd\mathbf{z}) \wedge (Ed\mathbf{x} + Fd\mathbf{y} + Gd\mathbf{z})$ egy felületelem, ahol $A, B, C, E, F, G \in \mathbb{R}$, és $\wedge$ a külső szorzat jele.*

Általában, ha van két vektorunk $\mathbf{u}$ és $\mathbf{v}$, akkor az $\mathbf{u} \wedge \mathbf{v}$ külső szorzatra a következő elvárásaink vannak:

$$(\mathbf{u}_1 + \mathbf{u}_2) \wedge \mathbf{v} = \mathbf{u}_1 \wedge \mathbf{v} + \mathbf{u}_2 \wedge \mathbf{v}$$

$$(\lambda \mathbf{u}) \wedge \mathbf{v} = \lambda(\mathbf{u} + \mathbf{v})$$

$$\mathbf{u} \wedge \mathbf{u} = 0$$

**3.1. Következmény.**

$$\mathbf{u} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{u}$$

*Bizonyítás.* A fenti definíció alapján

$$(\mathbf{u} + \mathbf{v}) \wedge (\mathbf{u} + \mathbf{v}) = 0,$$

amit kifejtve kapjuk, hogy

$$(\mathbf{u} \wedge \mathbf{u}) + (\mathbf{v} \wedge \mathbf{v}) + (\mathbf{u} \wedge \mathbf{v}) + (\mathbf{v} \wedge \mathbf{u}) = 0.$$

Az összeg első két tagja definíció szerint 0, így kapjuk a bizonyítandó állítást.

**3.2. Következmény.** *Legyen* $\mathbf{u} = Ad\mathbf{x}+Bd\mathbf{y}+Cd\mathbf{z}$, *illetve* $\mathbf{v} = Ed\mathbf{x}+Fd\mathbf{y}+Gd\mathbf{z}$. *Ekkor*

$$\mathbf{u} \wedge \mathbf{v} = (AG - CF)d\mathbf{y} \wedge d\mathbf{z} + (CE - AG)d\mathbf{z} \wedge d\mathbf{x} + (AF - BE)d\mathbf{x} \wedge d\mathbf{y}.$$

**3.1. Megjegyzés.** *Ez analóg a jól ismert geometriai vektoriális szorzattal.*

*Példa.* Legyen $\mathbf{u}_1 = A_{11}d\mathbf{x} + A_{12}d\mathbf{y} + A_{13}d\mathbf{z}$, $\mathbf{u}_2 = A_{21}d\mathbf{x} + A_{22}d\mathbf{y} + A_{23}d\mathbf{z}$, és $\mathbf{u}_3 = A_{31}d\mathbf{x} + A_{32}d\mathbf{y} + A_{33}d\mathbf{z}$. Ekkor

$$\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \mathbf{u}_3 = (\det A)d\mathbf{x} \wedge d\mathbf{y} \wedge d\mathbf{z}.$$

Most tekintsük az $n$-dimenziós Euklideszi vektorteret, és benne a $\sigma_1, \ldots, \sigma_n$ ortonormált bázist. Legyen $\alpha = \sum_{i=1}^{n} a_i\sigma_i$, és $\beta = \sum_{i=1}^{n} b_i\sigma_i$.

Definiáljuk a külső szorzatot a következő módon:

$$\alpha \wedge \beta = \sum_{i<j=1} (a_ib_j - a_jb_i)\sigma_i \wedge \sigma_j. \tag{6}$$

Nyilvánvaló, hogy ekkor a külső szorzat bilineáris, azaz mindkét változójában lineáris.

**3.2. Megjegyzés.** *A tényleges menete a külső szorzat definiálásának a következő: Vegyük először az $\mathbb{R}^n \otimes \mathbb{R}^n$-nel jelölt teret, melynek elemei*

$$\alpha \otimes \beta = \sum_{i,j} (a_i b_j) \sigma_i \otimes \sigma_j.$$

*Ezt tulajdonképpen úgy kapjuk, hogy vesszük a $H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \otimes \mathbb{R}^n$ bilineáris leképezést (azaz megköveteljük, hogy mindkét változójában lineáris legyen) olyan módon, hogy a lehetséges bilineáris leképezések között ő legyen a legszűkebb abban az értelemben, hogy bármely $\mathbb{R}^n \times \mathbb{R}^n$-ből képező bilineáris leképezés keresztülvihető legyen $\mathbb{R}^n \otimes \mathbb{R}^n$-n, azaz ha $F$ egy tetszőleges bilineáris függvény, akkor létezik olyan $f$ lineáris függvény, hogy*

$$F = f \circ H.$$

*Az így kapott $\mathbb{R}^n \otimes \mathbb{R}^n$ térben vegyük azt az alteret, melyet az $\alpha \otimes \alpha$ alakú elemek generálnak. E szerint az altér szerint faktorizálva kapjuk meg a fent leírt külső szorzat teret, melynek elemei (6) alakúak.*

A külső szorzat teret jelöljük $\wedge^2 L$-lel.

**3.1. Állítás.** *A $\wedge^2 L$ tér egy bázisát adják a $\sigma_i \wedge \sigma_j$ alakú elemek.*

**3.3. Megjegyzés.** *Tehát $\wedge^2 L$ elemei a fentiek alapján nem csak $\sum\limits_r \lambda_r (\alpha_r \wedge \beta_r)$ alakban írhatóak, hanem*

$$\sum_{i<j} a_{ij} \sigma_i \wedge \sigma_j$$

*alakban is.*

**3.4. Megjegyzés.** *A 3.1 állítás miatt a $\wedge^2 L$ tér dimenziója $\binom{n}{2}$.*

Végül az eddigiekhez teljesen hasonló módon tudjuk definiálni az $\wedge^p L$ teret, ahol $p \geq 2$: Legyen $\alpha_i = \sum\limits_{r=1}^{n} a_{ir} \sigma_r$, és így

$$\alpha_1 \wedge \dots \wedge \alpha_p = \sum_{r_1, \dots r_p} a_{1 r_1} \dots a_{p r_p} \sigma_{r_1} \wedge \dots \wedge \sigma_{r_p},$$

ahol teljesül a linearitás, és $\alpha_i = \alpha_j \quad (i \neq j)$ esetén $\alpha_1 \wedge \dots \wedge \alpha_p = 0$.

**3.2. Állítás.**

$$\alpha_1 \wedge \dots \wedge \alpha_i \wedge \dots \wedge \alpha_j \wedge \dots \wedge \alpha_p = -\alpha_1 \wedge \dots \wedge \alpha_j \wedge \dots \wedge \alpha_i \wedge \dots \wedge \alpha_p,$$

*azaz teljesül az antiszimmetria.*

*Bizonyítás.* A fentiek alapján

$$\alpha_1 \wedge \dots \wedge (\alpha_i + \alpha_j) \wedge \dots \wedge (\alpha_j + \alpha_i) \wedge \dots \wedge \alpha_p = 0,$$

amit kifejtve kapjuk a bizonyítandó állítást.

Az $\wedge^p L$ tér bázisát alkotják a $\sigma_{h_1} \wedge \dots \wedge \sigma_{h_p}$ alakú elemek, ahol $1 \leq h_1 < \dots < h_p \leq n$.

Legyen $H = \{h_1, \dots h_p : 1 \leq h_1 < \dots < h_p \leq n\}$, és így $\wedge^p L$ tér elemei $\sum\limits_H a_H \sigma_H$ alakúak lesznek. $\wedge^p L$ dimenziója nyilván $\binom{n}{p}$ lesz.

Végül legyen $\pi$ az $\{1, \dots p\}$ halmaz egy permutációja, ahol $i$ képe $\pi(i)$. Ekkor

$$\alpha_{\pi(1)} \wedge \dots \alpha_{\pi(p)} = (\mathrm{sgn}\pi)\alpha_1 \wedge \dots \wedge \alpha_p.$$

$$\wedge^p L$$

Legyen $f$ egy $\wedge^p L \longrightarrow \mathbb{R}$ lineáris függvény. Definiáljuk a $g$ függvényt a $\bigotimes\limits_{i=1}^{p} L$ Descartes-szorzaton a következőképpen:

$$g(\alpha_1, \dots \alpha_p) = f(\alpha_1 \wedge \dots \wedge \alpha_p).$$

**3.3. Állítás.** *$g$ multilineáris és alternáló függvény a $\bigotimes\limits_{i=1}^{p} L$ szorzattéren.*

*Bizonyítás.*

$$g(\alpha_1' + \alpha_1'', \alpha_2, \dots, \alpha_p) = f((\alpha_1' + \alpha_1'') \wedge \alpha_2 \wedge \dots \wedge \alpha_p) =$$

$$= f(\alpha_1' \wedge \alpha_2 \wedge \cdots \wedge \alpha_p) + f(\alpha_1'' \wedge \alpha_2 \wedge \cdots \wedge \alpha_p)$$

bizonyítja a multilinearitást (a skalár is nyilván kivihető).

$g(\alpha_2, \alpha_1, \ldots \alpha_p) = f(\alpha_2 \wedge \alpha_1 \wedge \cdots \wedge \alpha_p) = -f(\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_p) = -g(\alpha_1, \alpha_2, \ldots, \alpha_p)$

bizonyítja az alternálást.

**3.4. Állítás.** *Fordítva: ha $g$ egy multilineáris és alternáló funkcionál a $\bigotimes_{i=1}^{p} L$ szorzattéren, akkor ez meghatároz egy $f$ lineáris funkcionált $\wedge^p L$-n.*

*Bizonyítás.* Definiáljuk a $f$ funkcionált a következő módon: Tetszőleges $\lambda \in \wedge^p L$ esetén, ahol $\lambda = \sum \lambda_r(\alpha_1^r \wedge \ldots \alpha_p^r)$, legyen

$$f(\lambda) = \sum_r \lambda_r g(\alpha_1^r, \ldots, \alpha_p^r).$$

$f$ ekkor nyilván lineáris, így azt kell még látni, hogy egyértelmű is lesz. A külső szorzat alternáló tulajdonsága és a $g$-re szabott feltételek miatt ez is teljesül, hiszen a $g$ függvény $(\sigma_{h_1}, \ldots, \sigma_{h_p})$ helyeken felvett értékei meghatározzák tetszőleges helyen felvett értéket.

Most legyenek $M$ és $N$ lineáris terek, és legyenek $\sigma_1, \ldots, \sigma_m$ illetve $\tau_1, \ldots, \tau_n$ a megfelelő bázisok. Továbbá legyen $A$ egy $M \longrightarrow N$ lineáris leképezés. Definiáljuk a $\wedge^p M \longrightarrow \wedge^p N$ lineáris leképezést a következő módon: Először vegyük az

$$(\alpha_1, \ldots, \alpha_p) \longrightarrow A\alpha_1 \wedge \cdots \wedge A\alpha_p$$

$\bigotimes_{i=1}^{p} M \longrightarrow \wedge^p N$ leképezést, ahol $\alpha_i \in M$. Ez nyilván minden változójában lineáris lesz, hiszen $A$ is és a külsőszorzat is lineáris. Másfelől alternáló is lesz a külsőszorzat alternálósága miatt. Tehát ez a leképezés meghatároz egy $\wedge^p M \longrightarrow \wedge^p N$ leképezést, melyet jelöljünk $\wedge^p A$-val.

*Példa.* : $\mathbb{R}^3$-ban természetes konstrukció a felületelem transzformálása:

$$\mathbf{u} \wedge \mathbf{v} \longrightarrow A\mathbf{u} \wedge A\mathbf{v},$$

ahol $\mathbf{u}, \mathbf{v}, A\mathbf{u}, A\mathbf{v} \in \mathbb{R}^3$.

Egy egyszerű állítás a következő:

**3.5. Állítás.**

$$\wedge^p(AB) = (\wedge^p A)(\wedge^p B)$$

*Bizonyítás.* Tetszőleges $(\alpha_1, \dots, \alpha_p)$-re az

$$(\alpha_1, \dots, \alpha_p) \longrightarrow AB\alpha_1 \wedge \cdots \wedge AB\alpha_p$$

leképezés előáll úgy is, hogy először alkalmazzuk csak $B$-t, majd $A$-t.

Legyen $(\alpha, \beta)$ belső szorzat $L$-en, ahol feltesszük, hogy $L$ egy $\mathbb{R}$ feletti vektortér, azaz $(\alpha, \beta)$ egy bilineáris és pozitív definit függvény. Terjesszük ezt ki $\wedge^p L$-re. Legyen $\lambda = \alpha_1 \wedge \cdots \wedge \alpha_p \in \wedge^p L$ és $\mu = \beta_1 \wedge \cdots \wedge \beta_p \in \wedge^p L$. Definiáljuk a $(\lambda, \mu)$ belső szorzatot a következő módon:

$$(\lambda, \mu) = \det((\alpha_i, \beta_j)_{i,j=1}^p).$$

**3.6. Állítás.** *A fent definiált $\mathbb{R}$-be képező függvény valóban belső szorzat lesz, azaz egy bilineáris, pozitív definit függvényt definiáltunk.*

Most keressünk ortonormált bázist $\wedge^p L$-ben. Legyen $\sigma_1, \dots, \sigma_n$ ortonormált bázis $L$-ben, azaz $(\sigma_i, \sigma_j) = \delta_{ij}$.

**3.7. Állítás.** *$\sigma_H$, ahol $H \subset \{1, \dots, p\}$ ortonormált bázis $\wedge^p L$-ben.*

*Bizonyítás.* Legyen $H \neq K$ két részhalmaza az indexeknek. Ekkor

$$(\sigma_H, \sigma_K) = \det(\sigma_{h_i}, \sigma_{k_j}).$$

Ha $h_1 \notin K$, akkor van egy 0 sora a mátrixnak, így a determináns nyilván 0 lesz. Ugyanakkor $H = K$ esetén $(\sigma_H, \sigma_H) = \det(\sigma_{h_i}, \sigma_{h_j}) = 1$.

**3.8. Állítás.** *Ha $A : L \longrightarrow L$ lineáris leképezés, akkor*

$$(\wedge^p A)^T = \wedge^p(A^T).$$

*Bizonyítás.* Legyen $\lambda = \alpha_1 \wedge \dots \wedge \alpha_p, \quad \mu = \beta_1 \wedge \dots \wedge \beta_p$. A transzponált definíciója alapján

$$((\wedge^p A)\lambda, \mu) = (\lambda, (\wedge^p A)^T \mu). \tag{7}$$

Másfelől

$$(A\alpha_1 \wedge \dots \wedge A\alpha_p, \beta_1 \wedge \dots \wedge \beta_p) = \det((A\alpha_i, \beta_j)_{i,j=1}^p) =$$
$$= \det((\alpha_i, A^T\beta_j)_{i,j=1}^p) = (\lambda, \wedge^p(A^T)\mu),$$

melyet összehasonlítva a (7) egyenlőséggel, kapjuk a bizonyítandó állítást.

**3.3. Következmény.** *Ha $U : L \longrightarrow L$ ortonormált leképezés (azaz $UU^T = I$), akkor $\wedge^p U$ is ortonormált leképezés $\wedge^p L$-ben.*

*Bizonyítás.*

$$(\wedge^p U)(\wedge^p U)^T = (\wedge^p U)(\wedge^p U^T) = \wedge^p(UU^T) = \wedge^p(I),$$

ami éppen az identitás mátrix az $\wedge^p L$ térben.

**3.9. Állítás.** *Legyen $D = diag(d_i)$ diagonális mátrix ($i = 1, \dots, n$). Ekkor $\wedge^p D$ is diagonális mátrix lesz, mégpedig*

$$\wedge^p D = diag(d_{i_1} \dots d_{i_p}),$$

*ahol $(i_1, \dots i_p) = H \subset \{1, \dots, n\}$, és $H$ minden lehetséges részhalmazon végigfut.*

*Bizonyítás.* Vizsgáljuk meg $\wedge^p D$ hatását $\wedge^p L$ bázisán.

$$\wedge^p D(\sigma_{i_1} \wedge \dots \wedge \sigma_{i_p}) = D\sigma_{i_1} \wedge \dots \wedge D\sigma_{i_p} = d_{i_1}\sigma_{i_1} \wedge \dots \wedge d_{i_p}\sigma_{i_p},$$

ami igazolja az állításunkat.

**3.1. Tétel.** *Legyen $A$ egy $n \times n$-es valós mátrix. Ekkor $\exists U, V$ ortonormált és $D$ diagonális mátrix ($D = diag(d_i)$, $\quad 0 \le d_1 \le \cdots \le d_n$), hogy $A = UDV^T$.*

**3.5. Megjegyzés.** *Ha $A$ szimmetrikus, akkor előáll $A = U\overline{D}U^T$ alakban, ahol $\overline{D}$ diagonális mátrix. Ezt hívják főtengely-transzformációnak.*

*Bizonyítás.* Tegyük fel, hogy létezik a tétel szerinti felbontása az $A$ mátrixnak. Tekintsük az $A$ mátrixot, mint $\widetilde{V} \longrightarrow \widetilde{U}$ leképezést. Tekintsük a $\widetilde{V}$ vektortérnek azt a bázisát, amit a $V$ mátrix oszlopvektorai határoznak meg (jelöljük $\mathbf{v}_i$-vel), valamint az $\widetilde{U}$ vektortérnek azt a bázisát, amit az $U$ mátrix oszlopvektorai határoznak meg (jelöljük $\mathbf{u}_j$-vel). $U, V$ ortonormáltsága miatt ezek valóban bázisok lesznek, és ráadásul ortonormált bázisok lesznek. Ebben a bázisban az $A$ leképezés mátrixa diagonális lesz, hiszen

$$U^T AV = U^T UDV^T V = D$$

Ismert, hogy

$$\|A\| = \max_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|},$$

és így

$$\|A\| = d_n$$

Ennek megfelelően a tételben szereplő felbontást a következőképpen kaphatjuk: Legyen $\mathbf{v}_n$ egy olyan vektor, amelyre teljesül, hogy

$$A\mathbf{v}_n = \|A\|\|\mathbf{v}_n\|, \quad \text{és} \quad \|\mathbf{v}_n\| = 1,$$

valamint legyen

$$A\mathbf{v}_n = \mathbf{u}_n.$$

Belátjuk, hogy ez a választás megfelelő lesz. Ehhez az kell, hogy tetszőleges $\mathbf{x}$ vektorra, mely merőleges $\mathbf{v}_n$-re ($\mathbf{x} \perp \mathbf{v}_n$), teljesül az, hogy $A\mathbf{x} \perp \mathbf{u}_n$.

Tegyük fel, hogy $\mathbf{x} \perp \mathbf{v}_n$ és $(A\mathbf{x}, \mathbf{u}_n) > 0$. Tekintsük a $\mathbf{v}_n + \epsilon\mathbf{x}$ vektort. Vegyük a

$$\frac{\|A(\mathbf{v}_n + \epsilon\mathbf{x})\|}{\|\mathbf{v}_n + \epsilon\mathbf{x}\|}$$

hányadost. Belátjuk, hogy ez nem $\epsilon = 0$ esetén maximális, és így ellentmondást kapunk $\mathbf{v}_n$ definíciója miatt. Ehhez persze az kell, hogy felhasználva az $A$ linearitását is

$$\frac{\partial}{\partial \epsilon} \frac{\|\mathbf{u}_n + \epsilon A\mathbf{x}\|}{\|\mathbf{v}_n + \epsilon \mathbf{x}\|}|_{\epsilon=0} > 0.$$

Mivel

$$\frac{\partial}{\partial \epsilon} \frac{\|\mathbf{u}_n + \epsilon A\mathbf{x}\|}{\|\mathbf{v}_n + \epsilon \mathbf{x}\|} =$$

$$= \frac{(2(A\mathbf{x}, \mathbf{u}_n) + 2\epsilon(A\mathbf{x}, A\mathbf{x}))((\mathbf{v}_n, \mathbf{v}_n) + \epsilon^2(\mathbf{x}, \mathbf{x})) - 2\epsilon(\mathbf{x}, \mathbf{x})((\mathbf{u}_n, \mathbf{u}_n)}{((\mathbf{v}_n, \mathbf{v}_n) + \epsilon^2(\mathbf{x}, \mathbf{x}))^2} +$$

$$+ \frac{2\epsilon(A\mathbf{x}, \mathbf{u}_n) + \epsilon^2(A\mathbf{x}, A\mathbf{x}))}{((\mathbf{v}_n, \mathbf{v}_n) + \epsilon^2(\mathbf{x}, \mathbf{x}))^2},$$

így $\epsilon \longrightarrow 0$ esetén ez a hányados pozitív.

**3.10. Állítás.** *Legyen $A = UDV^T$ az $A$ SVD felbontása. Ekkor*

$$\wedge^p A = (\wedge^p U)(\wedge^p D)(\wedge^p V)^T$$

*a $\wedge^p A$ SVD felbontása.*

*Bizonyítás.* Az, hogy $\wedge^p A = (\wedge^p U)(\wedge^p D)(\wedge^p V)^T$ egyenlőség fennáll, az nyilvánvaló a 3.5 állítás alapján. Kérdés, hogy ez valóban SVD felbontás-e. Azonban itt a 3.3 következmény miatt $\wedge^p U$ és $\wedge^p V$ ortonormált, illetve 3.9 állítás miatt $\wedge^p D$ diagonális, ami bizonyítja az állítást.

Legyen $D = \mathrm{diag}(d_i)$, ahol $d_1 \leq \cdots \leq d_n$. Ekkor rendezzük lexikografikusan a $H \subset \{1, \ldots, n\}$ részhalmazokat, és legyenek $d_H = d_{i_1} \ldots d_{i_p}$ a megfelelő diagonális elemek. Ekkor $H_1 < H_2 < \cdots < H_{\binom{n}{p}}$ esetén $d_{H_1} \leq d_{H_2} \leq \cdots \leq d_{H_{\binom{n}{p}}}$.

**3.11. Állítás.** *Legyen $\mathbf{u}, \mathbf{v} \in L$. Ekkor*

$$\|\mathbf{u} \wedge \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\||\sin \gamma|,$$

*ahol $\gamma$ jelöli az $\mathbf{u}, \mathbf{v}$ vektorok által bezárt szöget.*

*Bizonyítás.* Legyen $\alpha = \beta = \mathbf{u} \wedge \mathbf{v}$. Ekkor

$$(\alpha, \beta) = \begin{vmatrix} (\mathbf{u}, \mathbf{u}) & (\mathbf{u}, \mathbf{v}) \\ (\mathbf{v}, \mathbf{u}) & (\mathbf{v}, \mathbf{v}) \end{vmatrix} = \begin{vmatrix} \|\mathbf{u}\|^2 & \|\mathbf{u}\|\|\mathbf{v}\| \cos\gamma \\ \|\mathbf{u}\|\|\mathbf{v}\| \cos\gamma & \|\mathbf{v}\|^2 \end{vmatrix} =$$

$$= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \cos^2\gamma = \|\mathbf{u}\|\|\mathbf{v}\||\sin\gamma|.$$

# 4. Véletlen mátrixok szorzata II.

**4.1. Definíció.** *Legyenek $A_1, A_2, \dots$ tetszőleges $r \times r$-es mátrixok. Egy $\mathbf{x}$ vektor Ljapunov-exponense legyen*

$$\lambda(\mathbf{x}) = \limsup_{n\to\infty} \frac{1}{n} \log \|A_n \dots A_1 \mathbf{x}\|.$$

**4.1. Állítás.** *Tetszőleges valós $c$-re $L_c = \{\mathbf{x} : \lambda(\mathbf{x}) < c\}$ egy altér lesz.*

*Bizonyítás.* Az $A_i$ mátrixok linearitása, valamint $n \longrightarrow \infty$ miatt ha $\mathbf{x} \in L_c$, akkor nyilván $k\mathbf{x} \in L_c$ is teljesül minden valós $k$-ra.

Másfelől, ha $\mathbf{x}, \mathbf{y} \in L_c$, akkor mivel

$$\frac{1}{n} \log \|A_n \dots A_1(\mathbf{x} + \mathbf{y})\| \le \frac{1}{n} \log(\|A_n \dots A_1\mathbf{x}\| + \|A_n \dots A_1\mathbf{y}\|) \le$$
$$\le \frac{1}{n} \log(2 \max(\|A_n \dots A_1\mathbf{x}\|, \|A_n \dots A_1\mathbf{y}\|)) \le$$
$$\le \frac{\log 2}{n} + \frac{1}{n} \log \max(\|A_n \dots A_1\mathbf{x}\|, \|A_n \dots A_1\mathbf{y}\|),$$

így $\mathbf{x} + \mathbf{y} \in L_c$ is nyilván teljesül.

**4.1. Következmény.** *Az előző állítás alapján ha $c < c'$, akkor $L_c \subset L_{c'}$. Mivel a teljes tér dimenziója $r$ (legyen a teljes tér $L$), így létezik véges sok altér, melyekre*

$$E_1 \subset \dots \subset E_p = L \textit{(a tartalmazás szigorú tartalmazást jelent), és}$$

$$\textit{ha} \quad \mathbf{x} \in E_{k+1} \backslash E_k, \quad \textit{akkor} \quad \lambda(\mathbf{x}) = \lambda_k,$$

*azaz a Ljapunov-exponens konstans két szomszédos $E_i$ között.*

Most legyen $(A_n)$ $r \times r$-es mátrixokból álló stacionárius folyamat, amely ergodikus ($A_n(\omega) = A_1(T^{n-1}\omega)$, ahol $T$ mértéktartó transzformáció és $TX \subset X$ esetén $P(X) = 0$ vagy 1.) Tegyük fel, hogy $E \log^+ \|A_1\| < \infty$, és vezessük be az $u(n,x) = A_n(x) \dots A_1(x)$ jelölést.

**4.1. Tétel.** *Legyen $u(n,x) = M'(n,x)D(n,x)M(n,x)$ SVD-felbontás, ahol $D(n,x) = diag(\overline{d}_i(n,x))_{i=1}^r$. Ekkor*

$$\lim_{n\to\infty} \frac{1}{n} \log \overline{d}_i(n,x)$$

*létezik 1-valószínűséggel.*

**4.1. Megjegyzés.** *$\overline{d}_r(n,x) = \|u(n,x)\|$, és a Fürstenberg-Kesten tétel alapján azt tudjuk, hogy $\lim\limits_{n\to\infty} \frac{1}{n} \log \|u(n,x)\|$ létezik 1-valószínűséggel.*

*Bizonyítás.* Bizonyítsuk be az állítást $\overline{d}_{r-1}(n,x)$-re. Tekintsük az $(A_n \wedge A_n)$ stacionárius folyamatot.

$$(A_n \wedge A_n) \dots (A_1 \wedge A_1) = u(n,x) \wedge u(n,x),$$

és így a szorzat SVD-felbontása

$$(M'(n,x) \wedge M'(n,x))(D(n,x) \wedge D(n,x))(M(n,x) \wedge M(n,x)).$$

Azonban

$$D(n,x) \wedge D(n,x) = \mathrm{diag}(\overline{d}_{i_1}(n,x)\overline{d}_{i_2}(n,x))_{i_1 \le i_2},$$

azaz a legnagyobb szinguláris érték $\overline{d}_r(n,x)\overline{d}_{r-1}(n,x)$ Mivel az $(A_n \wedge A_n)$ stacionárius, ergodikus folyamat és $E \log^+ \|A_1 \wedge A_1\| < \infty$, így alkalmazható erre a folyamatra is a Fürstenberg-Kesten tétel, azaz

$$\lim_{n\to\infty} \frac{1}{n} \log \|(u(n,x) \wedge u(n,x))\| = \lim_{n\to\infty} \frac{1}{n} \log(\overline{d}_r(n,x)\overline{d}_{r-1}(n,x))$$

létezik 1-valószínűséggel. Azonban

$$\lim_{n\to\infty} \frac{1}{n} \log(\overline{d}_r(n,x)\overline{d}_{r-1}(n,x)) = \lim_{n\to\infty} \frac{1}{n} \log(\overline{d}_r(n,x)) + \lim_{n\to\infty} \frac{1}{n} \log(\overline{d}_{r-1}(n,x))$$

és az összeg első tagja a fentiek alapján (lásd a megjegyzést) konvergens 1-valószínűséggel.

Tehát $\overline{d}_{r-1}(n,x)$-re az állítást bebizonyítottuk. Ha $\overline{d}_{i+1}$-ig tudjuk a konvergenciát 1-valószínűséggel, akkor $i$-re a fentihez hasonló gondolatmenettel igazolható (véve az $A_n \wedge \cdots \wedge A_n \quad r-i+1$-szeres külsőszorzatot).

Legyen $\lim\limits_{n\to\infty} \frac{1}{n} \log \overline{d}_i(n,x) = d_i$.

Most kimondjuk a fejezet fő tételét, melyet több lépésben fogunk bizonyítani.

**4.2. Tétel.** *(Oseledec) A fenti jelölések mellett létezik az indexeknek egy* $I = \{1 = i_1 < \cdots < i_p < i_{p+1} = r+1\}$ *felbontása és* $f_1 \leq f_2 \leq \cdots \leq f_r$ *valós számok* ($f_{i_q} < f_{i_{q+1}}$, *illetve* $f_i = f_{i_q} \quad i_q \leq i < i_{q+1}$ *esetén), hogy* $E(I,q,x)$ *növekvő alterek, és amelyekre igaz, hogy majdnem minden* $x$*-re* $\forall \mathbf{v} \in E(I,q,x) - E(I,q-1,x)$*-re*

$$\lim_{n\to\infty} \frac{1}{n} \log \|u(n,x)\mathbf{v}\| = f_{i_q}$$

Tekintsük az $u(n,x)$ véletlen mátrixot, mint egy $\mathbb{R}^r \longrightarrow \mathbb{R}^r$ leképezést. Tekintsük ezt a leképezést az $M^{-1}(n,x)\mathbf{e}_i$ bázisban, ahol $\mathbf{e}_i$ az $i$. kanonikus bázisvektor ($\mathbf{e}_i = (0,\ldots 0,1,0,\ldots 0)^T$). Legyen

$$\mathbf{v}_i(n,x) = M^{-1}(n,x)\mathbf{e}_i.$$

Másfelől legyen a képtér bázisa

$$\mathbf{z}_j(n,x) = M'(n,x)\mathbf{e}_j.$$

Így

$$\begin{gathered}(\mathbf{z}_j(n,x), u(n,x)\mathbf{v}_i(n,x)) = \mathbf{z}_j(n,x)^T u(n,x)\mathbf{v}_i(n,x) = \\ = \mathbf{e}_j^T M'(n,x)^T M'(n,x) D(n,x) M(n,x) M^{-1}(n,x)\mathbf{e}_i = \mathbf{e}_j^T \overline{d}_i(n,x)\mathbf{e}_i = \\ = \overline{d}_i(n,x)\delta_{ij}\end{gathered}$$

Most definiáljuk a véges sok tagú, véletlen mátrixok szorzataként előállt véletlen leképezést, és az általa generált felbontását a térnek.

**4.2. Definíció.**

$$E(I,q,n) = \{\sum_{i<i_{q+1}} k_i M^{-1}(n,x)\mathbf{e}_i\},$$

*ahol* $k_i \in \mathbb{R}$

Legyen $\mathbf{v} \in E(I,q,n) \quad (\|\mathbf{v}\| = 1)$.

**4.1. Lemma.**

$$\|u(n,x)\mathbf{v}\| \leq \exp\{n(d_{i_q} + \epsilon)\}$$

*$\forall \epsilon > 0$-ra ha $n \geq n_0(\epsilon, x)$, 1-valószínűséggel.*

*Bizonyítás.*

$$\|M'(n,x)D(n,x)M(n,x)\mathbf{v}\| = \|M'(n,x)D(n,x)M(n,x)\sum_{i<i_{q+1}} k_i M^{-1}(n,x)\mathbf{e}_i\| =$$

$$= \|M'(n,x)D(n,x)\sum_{i<i_{q+1}} k_i\mathbf{e}_i\| = \|M'(n,x)\|\|\sum_{i<i_{q+1}} k_i\overline{d}_i(n,x)\mathbf{e}_i\| =$$

$$= \|\sum_{i<i_{q+1}} k_i\overline{d}_i(n,x)\mathbf{e}_i\|$$

4.1 határeloszlás tétel miatt $\overline{d}_i(n,x) \leq \exp\{n(d_i + \epsilon)\} \quad n \geq n_0(\epsilon,x)$-re 1-valószínűséggel. Így felhasználva, hogy $d_i$-k növekedő sorrendben vannak definiálva,

$$\|\sum_{i<i_{q+1}} k_i\overline{d}_i(n,x)\mathbf{e}_i\| \leq \sum_{i<i_{q+1}} |k_i| \exp\{n(d_{i_q} + \epsilon)\}.$$

Felhasználva a Cauchy-Bunyakovszkij egyenlőtlenséget, valamint azt, hogy $\|v\| = 1$

$$\sum_{i<i_{q+1}} |k_i| \leq (\sum_{i<i_{q+1}} k_i^2)^{1/2}(\sum_{i<i_{q+1}} 1^2)^{1/2} \leq r^{1/2}$$

Összevetve az előző egyenlőtlenséggel kapjuk a bizonyítandó állítást.

Most nézzük meg mi a helyzet "kilógó" $\mathbf{v}$-re, azaz ha $\mathbf{v} \notin E(I,q,n)$. Legyen $\mathbf{v} = \sum_{i=1}^{r} b_i M^{-1}(n,x)\mathbf{e}_i$.

**4.2. Lemma.**

$$\|u(n,x)\mathbf{v}\| \geq |b_i| \exp\{n(d_i - \epsilon)\}$$

*$\forall \epsilon > 0$-ra $\forall i$-re $n \geq n_0(\epsilon, x)$ esetén, 1-valószínűséggel .*

*Bizonyítás.* Az előző lemma bizonyításához hasonlóan

$$\|u(n,x)\mathbf{v}\| = \|M'(n,x)D(n,x)\sum_{j=1}^{r} b_j\mathbf{e}_j\| = \|\sum_{j=1}^{r}\overline{d}_j(n,x)b_j\mathbf{e}_j\|,$$

melyet alulról lehet becsülni $\|\overline{d}_i(n,x)b_i\mathbf{e}_i\|$-vel $\forall i$-re. Szintén a 4.1 határeloszlás-tétel miatt

$$\overline{d}_i(n,x) \geq \exp\{n(d_i-\epsilon)\}$$

$\forall\epsilon > 0$-ra $n > n_0(\epsilon, x)$ esetén, amiből adódik a lemma állítása.

**4.2. Megjegyzés.** *Az $E\log^+\|A_1\| < \infty$ feltételből nyilvánvalóan következik, hogy $\|A_n\| \leq \exp(n\epsilon)\quad \forall\epsilon > 0$-ra $n > n_0(\epsilon, x)$ esetén.*

**4.3. Lemma.** *Legyen $\mathbf{v}_n \in E(I,q,n)$, $(\|\mathbf{v}_n\| = 1)$. $\mathbf{v}_n$ vetülete az $E(I,q,n+1)$ altérre legyen $\mathbf{z}_{n+1}$. Ekkor*

$$\|\mathbf{v}_n - \mathbf{z}_{n+1}\| \leq \exp\{-n(d_{i_{q+1}} - d_{i_q} - \epsilon)\}$$

*$\forall\epsilon > 0$-ra, $n > n_0(\epsilon, x)$ esetén, 1-valószínűséggel.*

*Bizonyítás.* $\mathbf{v}_n$-t írjuk fel az $\{M^{-1}(n+1,x)\mathbf{e}_i\}$ bázisban:

$$\mathbf{v}_n = \sum_{j<i_{q+1}} b_jM^{-1}(n+1,x)\mathbf{e}_j + \sum_{j\geq i_{q+1}} b_jM^{-1}(n+1,x)\mathbf{e}_j =$$

$$= \mathbf{z}_{n+1} + \sum_{j\geq i_{q+1}} b_jM^{-1}(n+1,x)\mathbf{e}_j.$$

4.1 lemma alapján

$$\|u(n,x)\mathbf{v}_n\| \leq \exp\{n(d_{i_q}+\epsilon)\}$$

$\forall\epsilon > 0$-ra $n > n_0(\epsilon, x)$ esetén 1-valószínűséggel. Így a 4.2 megjegyzés alapján

$$\|u(n+1,x)\mathbf{v}_n\| = \|A_{n+1}u(n,x)\mathbf{v}_n\| \leq \|A_{n+1}\|\|u(n,x)\mathbf{v}_n\| \leq \exp\{n(d_{i_q}+2\epsilon)\} \tag{8}$$

$\forall\epsilon > 0$-ra $n > n_0(\epsilon, x)$ esetén 1-valószínűséggel.

Másfelől 4.2 lemma alapján $j \geq i_{q+1}$ esetén

$$\|u(n+1,x)\mathbf{v}_n\| \geq |b_j|\exp\{n(d_j-\epsilon)\} \geq |b_j|\exp\{n(d_{i_{q+1}}-\epsilon)\} \tag{9}$$

$\forall\epsilon > 0$-ra $n > n_0(\epsilon, x)$ esetén 1-valószínűséggel.

Összevetve a (8) és a (9) egyenlőtlenségeket, kapjuk, hogy $j \geq i_{q+1}$-re

$$|b_j| \leq \exp\{n(d_{i_q} - d_{i_{q+1}} + 3\epsilon)\}, \text{ azaz}$$

$$\| \sum_{j \geq i_{q+1}} b_j M^{-1}(n+1, x)\mathbf{e}_j \| \leq \sum_{j \geq i_{q+1}} |b_j| \leq r \exp\{n(d_{i_q} - d_{i_{q+1}} + 3\epsilon)\},$$

így a lemma állítását bebizonyítottuk.

Az $\mathbb{R}^r$ alterei között bevezetünk egy távolságfogalmat. Legyen $A$ és $B$ két altér $\mathbb{R}^r$-ben.

**4.3. Definíció.**

$$gap(A, B) = \|P_A - P_B\|,$$

*ahol $P_A$, illetve $P_B$ jelöli az $A$, illetve a $B$ altérre történő projekciót.*

Bizonyítás nélkül mondunk ki a gap metrikára vonatkozóan néhány állítást.

**4.2. Állítás.** *A gap metrika jól definiált, azaz valóban teljesülnek a metrika tulajdonságok.*

**4.3. Megjegyzés.**

$$gap(A, B) \leq 1$$

**4.3. Állítás.** *A gap metrika teljes.*

**4.3. Tétel.** *Tegyük fel, hogy $gap(A, B) < 1$. Ekkor*

$$gap(A, B) = \sup_{\mathbf{u} \in A, \|\mathbf{u}\|=1} dist(\mathbf{u}, B)$$

Legyen $A = E(I, q, n)$ és $B = E(I, q, n+1)$. Ekkor a 4.3 lemmában éppen azt láttuk be, hogy ez a két altér a gap metrikában exponenciálisan közel van egymáshoz.

**4.2. Következmény.**

$$gap(E(I,q,n),E(I,q,n+1)\leq \exp\{-n(d_{i_q}-d_{i_{q+1}}-\epsilon)\}$$

Mivel tetszőleges $n<m$ esetén

$$\mathrm{gap}(E(I,q,n),E(I,q,m))\leq \exp(-nc),$$

így az $E(I,q,n)$ sorozat Cauchy sorozat lesz $\mathbb{R}^r$-ben, azaz a gap metrika teljessége miatt létezik az

$$E(I,q)=\lim_{n\to\infty}E(I,q,n)$$

határérték majdnem minden $x$-re.

**4.4. Lemma.**

$$\mathbf{v}\in E(I,q)\backslash E(I,q-1)$$

*esetén*

$$\liminf_{n\to\infty}\frac{1}{n}\log\|u(n,x)\mathbf{v}\|\geq d_{i_q}$$

*Bizonyítás.* Írjuk fel a $\mathbf{v}$ vektort az $\{M^{-1}(n,x)\mathbf{e}_i\}$ bázisban.

$$\mathbf{v}=\mathbf{z}_n+\sum_{i\geq i_q}c_iM^{-1}(n,x)\mathbf{e}_i.$$

Mivel $\mathbf{v}\notin E(I,q-1)$, így $\mathrm{dist}(\mathbf{v},E(I,q-1)>0$, azaz elég nagy $n$-re valamilyen $\delta>0$-ra

$$\mathrm{dist}(\mathbf{v},E(I,q-1,n))>\delta$$

teljesül. Ez azt jelenti, hogy minden elég nagy $n$-re van olyan $c_i$ együttható ($i\geq i_q$), hogy $|c_i|\geq\delta'$, ahol $\delta'>0$ egy rögzített szám, mely független $n$-től. 4.2 lemma alapján így

$$\|u(n,x)\mathbf{v}\|\geq|c_i|\exp\{n(d_i-\epsilon)\}\geq\delta'\exp\{n(d_{i_q}-\epsilon)\}.$$

Mindkét oldal logaritmusának $n$-edrészének liminf-jét véve kapjuk a bizonyítandó lemma állítását.

Most tekintsük a $q=p$ esetet, azaz a teljes tér esetét. Ekkor ha $\mathbf{v}\notin E(I,p-1)$, akkor az előző lemma miatt

$$\liminf_{n\to\infty}\frac{1}{n}\log\|u(n,x)\mathbf{v}\|\geq d_{i_p}.$$

Másfelől tetszőleges $\mathbf{v}$-re fennáll, hogy

$$\|u(n,x)\mathbf{v}\| \leq \|u(n,x)\|\|\mathbf{v}\|,$$

ahol mindkét oldalnak véve a logaritmusának $n$-edrészét, kapjuk, hogy

$$\limsup_{n\to\infty} \frac{1}{n} \log \|u(n,x)\mathbf{v}\| \leq d_{i_p} \quad \forall v - \text{re}.$$

Ez azt jelenti, hogy

**4.3. Következmény.** $\mathbf{v} \notin E(I,p-1)$ *esetén*

$$\lim_{n\to\infty} \log \|u(n,x)\mathbf{v}\| = d_{i_p}.$$

A többi altérre itt nem bizonyítunk hasonló állítást. (lásd Raganathan)

Most tekintsük az $x \longrightarrow E(I,p-1)$ leképezést.

**4.4. Állítás.** *Az* $x \longrightarrow E(I,q)$ *leképezés Borel-mérhető.*

*Bizonyítás.* Mivel $E(I,q) = \lim_n E(I,q,n)$, így elegendő $E(I,q,n)$ mérhetősége. Ehhez viszont

$$E(I,q,n) = \{\sum_{i<i_{q+1}} \lambda_i M^{-1}(n,x)\mathbf{e}_i\}$$

miatt elég $M(n,x)$ mérhetősége. Tehát elegendő azt belátni, hogy tetszőleges $A$-ra

$$A = UDV^T$$

SVD felbontás esetén $U, D, V$ választható $A$ mérhető függvényeként. Ez viszont igaz, hiszen $A = UDV^T$ esetén $AA^T = UD^2U^T$, így ha $AA^T$ sajátértékei egyszeresek, akkor $U$ és $D$ egyértelmű, folytonos függvényei $AA^T$-nak, ha viszont van többszörös sajátértéke, akkor ez $A$-ban csak 0-mértékű lehet.

Az Oseledec-tétel egy fontos állításához jutottunk.

**4.5. Állítás.** $\lambda$*-majdnem minden* $\mathbf{v} \in \mathbb{R}^r$ *esetén*

$$\lim_{n\to\infty} \frac{1}{n} \log \|u(n,x)\mathbf{v}\| = d_{i_p},$$

*ahol* $\lambda$ *jelöli a Lebesgue-mértéket.*

*Bizonyítás.* Tekintsük az

$$E^* = \{(x, \mathbf{v}) : x \in \Omega, \mathbf{v} \in E(I, p-1, x)\}$$

halmazt. Megmutatható, hogy $E^*$ mérhető halmaz $\Omega \times \mathbb{R}^r$-ben. Ekkor

$$(\lambda \times P)E^* = \int_\Omega \lambda(E(I, p-1, x))dP = 0,$$

hiszen hipersík Lebesgue-mértéke 0. Így Fubini-tétel alapján

$$\int_{\mathbb{R}^r} P(x : (x, \mathbf{v}) \in E^*)d\lambda = 0,$$

azaz $P(x : (x, \mathbf{v}) \in E^*) = 0 \quad \lambda$-majdnem minden $\mathbf{v}$-re, vagyis $P(x : (x, \mathbf{v}) \notin E^*) = 1$, tehát

$$P(x : \mathbf{v} \notin E(I, p-1)) = 1 \quad \lambda - \text{majdnem minden} \quad \mathbf{v} - \text{re}.$$

Ezzel az állítást beláttuk.

A következőkben szintén egy, az Osedelec-tételhez tartozó állítást tárgyalunk.

**4.4. Tétel.**

$$\lim_{n\to\infty} (u^T(n, x)u(n, x))^{\frac{1}{2n}}$$

*létezik 1-valószínűséggel.*

Az állítást több részállítás és lemmán keresztül fogjuk bizonyítani.

Legyen $u(n, x) = M'(n, x)D(n, x)M(n, x)$ az SVD-felbontás a korábbiakhoz hasonlóan.

**4.4. Megjegyzés.** *$M(n, x)$ nem egyértelmű.*

*Bizonyítás.* Legyen egy $A$ mátrix SVD-felbontása $A = UDV^T$, ahol $D = \text{diag}(\lambda_i)$ Ekkor $A$ felírható

$$A = U\text{diag}(\lambda_i)IV^T = UW\text{diag}(\lambda_i)W^TV^T$$

alakban, ahol $WW^T = I$. Így ha $A$-nak vannak többszörös sajátértékei, akkor az ő SVD-felbontása nyilván nem egyértelmű.

Legyen

$$F(I,q,n) = \{ \sum_{i_q \le i < i_{q+1}} \lambda_i M^{-1}(n,x)\mathbf{e}_i \}.$$

**4.6. Állítás.** *Az $F(I,q,n)$ altér konvergens.*

*Bizonyítás.*

$$P_{F(I,q,n)} = P_{E(I,q+1,n)} - P_{E(I,q,n)},$$

és a jobb oldalon 4.3 lemma miatt van konvergencia.

**4.7. Állítás.** *Ha $E_n \longrightarrow E$ a gap-metrikában, akkor $\exists (\mathbf{v}_n^i)_{i=1}^k$ ortonormált bázisok $E_n$-ben, hogy $\mathbf{v}_n^i \longrightarrow \mathbf{v}^i$ és $\mathbf{v}^i$ ortonormált bázis $E$-ben. ($k = \dim E_n = \dim E$)*

*Bizonyítás.* Válasszunk egy $\mathbf{v}^i$ bázist $E$-ben, és legyen $\mathbf{v}_n^i = P_{E_n}\mathbf{v}^i$. Ekkor könnyen látható, hogy $\mathbf{v}_n^i \longrightarrow \mathbf{v}^i$, és $\mathbf{v}_n^i$-k lineárisan független egymásra merőleges rendszert alkotnak.

Az $F(I,q,n)$ altér konvergens, és a 4.7 állítás bizonyításának megfelelően konstruált ortonormált bázisa legyen $V_q(n,x)$, azaz $V_q(n,x) \longrightarrow V_q(x)$ ortonormált bázis $F(I,q)$-ban. Legyen

$$(M^{-1}(n,x))_q = (M^{-1}(n,x))_{i_q \le i < i_{q+1}}.$$

**4.8. Állítás.** *Létezik $V_q(n,x)$ ortonormált bázis, hogy $(M^{-1}(n,x))_q V_q(n,x)$ konvergens.*

Most tekintsük egyben az egész teret, és legyen

$$\mathrm{diag}(V_q(n,x)) = V(n,x)$$

Nyilván $V(n,x)$ is ortogonális lesz és a

$$V'(n,x)^T = M^T(n,x)V(n,x)$$

ortogonális mátrixsorozat konvergens lesz.

*Bizonyítás.* (4.4 tétel)

$$u(n,x) = M'(n,x)D(n,x)M(n,x)$$

az u(n,x) SVD felbontása, és így

$$u(n,x)^T u(n,x) = V'(n,x)V^T(n,x)D^2(n,x)V(n,x)V'(n,x) \tag{10}$$

a fenti jelölések mellett. Vegyük az $\sqrt[2n]{\cdot}$ gyököt. Felhasználva, hogy ha $A$ SVD-felbontása $A = U^T DU$, akkor $\sqrt[2n]{A} = U^T \sqrt[2n]{D}U$, kapjuk, hogy

$$\sqrt[2n]{u^T(n,x)u(n,x)} = V'(n,x)V^T(n,x)\sqrt[2n]{D^2}V(n,x)V'(n,x)$$

A 4.1 tételt felhasználva kapjuk, hogy

$$G_n = \sqrt[2n]{D^2(n,x)} = \operatorname{diag}(\exp(d_i + \epsilon_i(n,x))) \longrightarrow \operatorname{diag}(\exp d_i) = G.$$

(10)-ben szerepel még $V(n,x)$, melyről nem tudjuk, hogy konvergens-e, csak azt, hogy ortonormált.

Azonban egyfelől

$$V^T(n,x)G_nV(n,x) - V^T(n,x)GV(n,x) \longrightarrow 0$$

1-valószínűséggel, másfelől

$$V^T(n,x)GV(n,x) = \operatorname{diag}(V_q^T(n,x))G\operatorname{diag}(V_q(n,x)).$$

A q. blokkot tekintve

$$V_q^T(n,x)e^{d_{iq}}IV_q(n,x) = e^{d_{iq}}I,$$

azaz

$$(V^T(n,x)D^2(n,x)V(n,x))^{\frac{1}{2n}} \longrightarrow \operatorname{diag}(\exp(d_i))$$

Tehát kaptuk, hogy

$$\lim_{n\to\infty}(u^T(n,x)u(n,x))^{\frac{1}{2n}} = V'^T(x)\operatorname{diag}(\exp(d_i))V^T(x),$$

ahol $V'(x) = \lim_{n\to\infty} V'(n,x)$ ortonormált. Ezzel a tételt bebizonyítottuk.

# 5. Baum-egyenlet

Tekintsük az $(X_n, Y_n)$ rejtett Markov-folyamatot, ahol $X_n$ egy Markov folyamat a $\mathcal{X}$ véges állapottéren, $G_{ij} = P(X_{n+1} = j | X_n = i)$ az átmenetvalószínűség-mátrix (feltesszük, hogy $G$ primitív, azaz $G^k > 0$ valamilyen $k$ egész számra), és az $Y_n$ a kiolvasás, ahol

$$P(Y_n = j | X_n = i) = b^i(j) > 0 \quad \forall i, j\text{-re.}$$

Legyen

$$p_n(i) = P(X_n = i | Y_n, \dots, Y_1), \text{ és } p_n = (p_n(i))_{i=1}^r.$$

Ekkor a Baum-egyenlet

$$p_{n+1} = \pi(G^T B(Y_n) p_n),$$

ahol $\pi(x)_i = \frac{x_i}{\sum x_i}$, ahol $x \geq 0, x \neq 0$ és $B(y) = \mathrm{diag}(b^i(y))$.

Végül legyen $p_0 = q$ a kezdeti eloszlás, és így kapunk egy $p_n^q$ folyamatot. Bizonyítás nélkül mondjuk ki a következő tételt.

**5.1. Tétel.** *(Legland-Mevel)*

$$\|p_n^q - p_n^{q'}\|_{TV} \leq C\alpha^n,$$

*ahol* $0 < \alpha < 1$*, és* $TV$ *a totális variációt jelenti.*

Egy másik állítás a két különböző kezdetiértékkel induló folyamat eltérésére a

**5.2. Tétel.** *(Atar-Zeitouni)*

$$\lim_{n \to \infty} \frac{1}{n} \log \|p_n^q - p_n^{q'}\| = \gamma \quad \text{létezik 1-valószínűséggel Lebesgue-szerint majdnem minden} \quad q, q' \quad \text{párra},$$

*ahol* $\gamma$ *konstans.*

**5.1. Megjegyzés.** *Az 5.1 tétel szerint* $\lim\limits_{n \to \infty} \frac{1}{n} \log \|p_n^q - p_n^{q'}\| \leq \gamma \quad \forall q, q'$*-re.*

*Bizonyítás.* A Baum-egyenlet alapján

$$p_{n+1} = \pi(G^T B(Y_n)p_n) = \cdots = \pi(G^T B(Y_n)\ldots G^T B(Y_0)p_0).$$

Legyen $A_n = G^T B(Y_n)$ és $\overline{p}_{n+1} = A_n A_{n-1}\ldots A_0 p_0$. A $p_n^q$ folyamatra közvetlenül nem, de a $\overline{p}_n^q$ folyamatra tudjuk az Oseledec-tételt alkalmazni. Azonban mivel léteznek olyan $c_1, c_2 > 0$ konstansok, hogy

$$c_2 \sin(a,b) < \|a-b\|_{TV} < c_1 \sin(a,b)$$

(azaz a két norma ekvivalens), továbbá felhasználva, hogy

$$\sin(p_n^q, p_n^{q'}) = \sin(\overline{p}_n^q, \overline{p}_n^{q'}),$$

ez elegendő.

Az Oseledec-tétel szerint

$$\lim_{n\to\infty} \frac{1}{n} \log \|\overline{p}_n^q\| = \lambda_1 \quad \text{ha} \quad q \notin S_1(\omega),$$

$$\lim_{n\to\infty} \frac{1}{n} \log \|\overline{p}_n^{q'}\| = \lambda_1 \quad \text{ha} \quad q' \notin S_1(\omega), \text{illetve}$$

$$\lim_{n\to\infty} \frac{1}{n} \log \|\overline{p}_n^q \wedge \overline{p}_n^{q'}\| = \lambda_1 + \lambda_2 \quad \text{ha} \quad (q,q') \notin S_2(\omega),$$

ahol $\lambda_1$, illetve $\lambda_2$ az $A$ mátrix szinguláris értékfelbontásából adódó két legnagyobb sajátérték, míg $S_1(\omega), S_2(\omega)$ a 4.3 következmény szerinti "kilógó" valódi alterek.

Tehát

$$\lim_{n\to\infty} \frac{1}{n} \log \sin(\overline{p}_n^q, \overline{p}_n^{q'}) = (\lambda_1 + \lambda_2) - \lambda_1 - \lambda_1 = \lambda_2 - \lambda_1 \leq 0,$$

ha $q, q' \notin S_1(\omega)$, $(q,q') \notin S_2(\omega)$. Végül azt látjuk be, hogy egy $q$ kezdeti eloszlás vektor esetén $q \notin S_1(\omega)$.

**5.1. Állítás.** *Ha $q \geq 0$, akkor $q \in S_1^C(\omega)$.*

*Bizonyítás.* Először legyen $q > 0$. Indirekt feltevésként tegyük fel, hogy $q \in S_1(\omega)$. Tekintsük a

$$C = \lim_n (A_1^T \ldots A_n^T A_n \ldots A_1)^{\frac{1}{2n}}$$

leképezést. Ennek invariáns altere lesz $S_1(\omega)$. Legyen $v \perp q$ az a vektor, melyre $Cv = \lambda_1 v$, ahol $\lambda_1$ maximális és $v \neq 0$. Mivel $C \geq 0$ ($C_{ij} \geq 0$), így a Perron-Frobenius tétel szerint $v$ megválasztható úgy is, hogy $v \geq 0$ legyen. Azonban ez a $v$ egyrészt benne van $S_1(\omega)^{\perp}$-ben, másrészt a $(v, q)$ skalárszorzat pozitív, ami ellentmondás.

Most legyen $q \geq 0$. Mivel $G^k > 0$ valamilyen $k$ egész számra, így $A_{k-1} \ldots A_0 = G^T B(Y_{k-1}) \ldots G^T B(Y_0) > 0$, amit alkalmazva egy $q \geq 0$ vektorra, kapjuk, hogy

$$A_{k-1} \ldots A_0 q > 0,$$

és így a bizonyítás első felét erre alkalmazva kapjuk az állítást.

**5.1. Következmény.** *Ha* $(q, q') \in S_2^C(\omega)$, *akkor*

$$\lim_{n\to\infty} \frac{1}{n} \log \|p_n^q - p_n^{q'}\|_{TV} = \lambda_2 - \lambda_1.$$

Legyen $(X_n, Y_n)$ Rejtett Markov folyamat, ahol $X_n \in \mathcal{X}$ véges állapottér, $Y_n \in \mathcal{Y}$ véges vagy folytonos $\mathbb{R}^k$-ban. Tegyük fel, hogy a folyamatot egy $\Theta \in \mathbb{R}^m$ paraméter hajtja, az $f(y|X_i = j, \Theta)$ feltételes sűrűségfüggvények léteznek, valamint legyen az $X_n$ Markov folyamat átmenetvalószínűség mátrixa $G(\Theta)$.

Legyen $y_1, \ldots, y_n$ egy megfigyelés sorozat. A loglikelihood függvény a $\Theta$ paraméter mellett:

$$V_n(\Theta) = \log p(y_1, \ldots, y_n|\Theta) = \sum_{i=2}^{n} \log p(y_i|y_{i-1}, \ldots y_1, \Theta) + \log p(y_1|\Theta).$$

A maximum-likehood becslés konzisztenciájához az első kérdés az, hogy egyáltalán mikor létezik a

$$\lim_{n\to\infty} \frac{1}{n} V_n(\Theta)$$

határérték.

Heurisztikusan szemlélve a folyamatot, a $\log p(y_i|y_{i-1}, \dots y_1, \Theta)$ folyamat aszimptotikusan stacionárius, azaz approximálhatjuk a folyamatunkat a $\log p(y_i|y_{i-1}, \dots, \Theta)$ folyamattal, ahol ez utóbbi folyamatnál végtelen hosszú a múlt.

Legyen $\Theta = \Theta^*$ a valódi paraméter. Ekkor a Shannon-McMillan tétel alapján

$$\lim_{n\to\infty} \frac{1}{n} \log p(y_1, \dots y_n|\Theta^*) = H$$

létezik, ahol $\log p(y_1, \dots, y_n|\Theta^*)$ az optimális kódhossz a $\Theta^*$ mellett. És így az Ergod-tétel alapján

$$\frac{1}{n} \sum_{i=1}^{n} p(y_i|y_{i-1}, \dots, \Theta^*)$$

létezik.

Tekintsük az $x_{n+1} = ax_n + e_n$ AR(1) (autoregresszív) folyamatot, ahol $|a| < 1$, $(e_n)$ független azonos eloszlású sorozat és $E\|e_n\|^q < \infty \quad \forall q$-ra. Ismert, hogy ennek a folyamatnak van stacionárius megoldása.

**5.2. Állítás.** *Ha* $E\|X_0\|^q < \infty \quad \forall q$*-ra, akkor*

$$\sup_{m\geq 0} E\|X_m\|^q < \infty.$$

*Bizonyítás.*

$$x_{n+1} = a^{n+1}x_0 + \sum_{i=0}^{n-1} a^i e_{n-i},$$

és így

$$\|X_{n+1}\|^q \leq |a|^{(n+1)q}\|X_0\|^q + \sum_{i=0}^{n-1} |a|^{iq}\|e_{n-i}\|^q,$$

amiből adódik az állítás.

Ezek után vezessük be az $M$-korlátosság, majd az $L$-keverőség definícióját.

**5.1. Definíció.** *Egy $X_n \geq 0$ sztochasztikus folyamat $M$-korlátos, ha $\forall q \geq 1$-re*

$$M_q(x) = \sup E^{\frac{1}{q}} \|X_n\|^q < \infty.$$

Legyenek $(\mathcal{F}_n)$ monoton növő, illetve $(\mathcal{F}_n^+)$ monoton fogyó $\sigma$-algebra sorozatok olyanok, hogy $\mathcal{F}_n$ és $\mathcal{F}_n^+$ függetlenek $\forall n$-re. Egy egyszerű példa ilyen $\sigma$-algebrákra a fenti eset, amikor $(e_n)$ egy független azonos eloszlású valószínűségváltozó sorozat és $\mathcal{F}_n = \sigma\{e_i : i \leq n\}$, valamint $\mathcal{F}_n^+ = \sigma\{e_i : i \geq n+1\}$.

**5.2. Definíció.** *Az $X_n$ folyamat $L$-keverő, ha $M$-korlátos és*

$$\gamma_q(\tau) = \sup_{n \geq \tau} E^{1/q} \|X_n - E(X_n|\mathcal{F}_{n-\tau}^+)\|^q$$

*mellett*

$$\Gamma(q) = \sum_{\tau=0}^{\infty} \gamma_q(\tau) < \infty.$$

**5.3. Állítás.** *Ha $X_n$ $L$-keverő folyamat és $g(x)$ egy olyan függvény, amely legfeljebb polinomiálisan növekszik, akkor a $g(X_n)$ folyamat is $L$-keverő.*

**5.4. Állítás.** *$L$-keverő folyamatok összege, szorzata is $L$-keverő.*

Most nézzük meg, hogy az $L$-keverő folyamatok osztálya miért hasznos a számunkra.

**5.3. Tétel.** *Legyen $X_n$ egy olyan $L$-keverő folyamat, melyre $EX_n = 0$. Ekkor*

$$E^{\frac{1}{q}} \|\sum_{n=1}^{N} f_n X_n\|^q \leq C_q (\sum_{n=1}^{N} f_n^2)^{\frac{1}{2}} M_q(X) \Gamma_q(X) \quad \forall q \geq 2\text{-re.}$$

Alkalmazva a fenti tételt az $f_n \equiv 1$ esetre, kapjuk, hogy

**5.2. Következmény.**

$$E^{\frac{1}{q}} \|\frac{\sum_{n=1}^{N} X_n}{N}\|^q \leq C N^{-\frac{1}{2}}.$$

Tehát ha $X_n$ egy $L$-keverő folyamat és $EX_n = 0$ akkor

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = 0 \quad \text{majdnem mindenütt.}$$

Először legyen $Z$ egy véges értékkészletű valószínűségi változó. Legyen $P(Z = i) = p_i$

**5.3. Definíció.** *A $Z$ valószínűségi változó entrópiája legyen*

$$H(Z) = \sum_i -p_i \log p_i$$

**5.5. Állítás.** *Ha $Z$ és $Y$ két véges értékkészletű valószínűségi változó, akkor*

$$H(Z, V) \leq H(Z) + H(V)$$

Most legyen $Y_n$ egy szigorúan stacionárius, véges értékkészletű folyamat. Ekkor az előzőekhez hasonlóan definiálható a $H_n(Y_1, \ldots, Y_n)$ entrópia, és az 5.5 állítás következményeként

$$H_{n+m}(Y_1, \ldots Y_n, Y_{n+1}, \ldots, Y_{n+m}) \leq H_n(Y_1, \ldots, Y_n) + H_m(Y_{n+1}, \ldots, Y_{n+m}),$$

ami a stacionaritás miatt

$$= H_n(Y_1, \ldots, Y_n) + H_m(Y_1, \ldots, Y_m)$$

Tehát azt kaptuk, hogy a

$$g_n = H_n(Y_1, \ldots, Y_n)$$

egy szubadditív folyamat.

Ekkor a 2.2 tétel alapján

**5.6. Állítás.**

$$\lim_{n\to\infty} \frac{1}{n} H_n(Y_1, \ldots Y_n) = H$$

*létezik. (Ezt hívják a folyamat entrópia sebességének.)*

**5.2. Megjegyzés.**

$$H_n(Y_1,\ldots,Y_n) = \sum_{y_1,\ldots,y_n} -p(y_1,\ldots,y_n)\log p(y_1,\ldots,y_n) = E(-\log p(Y_1,\ldots,Y_n))$$

Most térjünk vissza a rejtett Markov modellünkhöz. Tehát legyen $(X_n, Y_n)$ rejtett Markov-folyamat, $X_n \in \mathcal{X}$ véges, a $G$ átmenetvalószínűség-mátrix primitív, $Y_n$ folytonos $\mathbb{R}$-értékű, és jelöljük $f(y|j)$-vel az $Y_n$ változó sűrűségfüggvényét az $X_n = j$ feltétel mellett. A továbbiakban tegyük fel, hogy

$$E|\log f(y|j)| < \infty \tag{11}$$

fennáll minden $j$-re. Rejtett Markov folyamat definíciója szerint az $Y_n$ változók feltételesen függetlenek, így kapjuk, hogy

**5.7. Állítás.** *$(Y_1,\ldots,Y_n)$ együttes sűrűségfüggvénye létezik.*

Jelöljük az együttes sűrűségfüggvényt $p(Y_n,\ldots,Y_1)$-gyel. Az 5.6 állításhoz hasonló állítást mondunk ki a folytonos esetre.

**5.8. Állítás.** *Létezik a*

$$\lim_{n\to\infty} \frac{1}{n} E(-\log p(Y_1,\ldots,Y_n))$$

*határérték.*

*Bizonyítás.* Teljes valószínűség tétel alapján

$$p(Y_1,\ldots,Y_n) = \prod_{i=2}^{n} p(Y_i|Y_{i-1},\ldots,Y_1)p(Y_1),$$

azaz a logaritmusának várhatóértékének $n$-edrészét véve kapjuk, hogy

$$\frac{1}{n}E\log p_n(Y_1,\ldots,Y_n) = \sum_{i=2}^{n}\frac{1}{n}E\log p(Y_i|Y_{i-1},\ldots,Y_1) + \frac{1}{n}E\log p(Y_1).$$

Ismert, hogy $a_n \longrightarrow a$ esetén $\frac{1}{n}\sum\limits_{i=1}^{n} a_i \longrightarrow a$ is fennáll, így elegendő azt belátni, hogy

$$E(-\log p(Y_n|Y_{n-1},\ldots,Y_1)) - \text{nek},$$

vagy ami ezzel ekvivalens

$$E(-\log p(Y_0|Y_{-1},\ldots,Y_{-n})) - \text{nek}$$

létezik határértéke $n \longrightarrow \infty$ esetén.

A továbbiakban vizsgáljuk ez utóbbi várhatóértéket konvergencia szempontjából. Rögzítsük le az $Y_0 = y_0$ pontot. Ekkor kihasználva, hogy $\mathcal{X}$ véges

$$p(y_0|Y_{-1},\ldots,Y_{-n}) = \sum_{j=1}^{m} P(X_0 = j|Y_{-1},\ldots,Y_{-n}) f(y_0|j).$$

Felhasználva, hogy

$$P(X_0 = j|Y_{-1},\ldots,Y_{-n}) = E(\chi_{X_0=j}|Y_{-1},\ldots,Y_{-n}),$$

így a Martingál-konvergencia alapján ennek van határértéke (ha $\mathcal{G}_i = \sigma(Y_{-1},\ldots,Y_{-i})$ és $\mathcal{G} = \sigma(\bigcup \mathcal{G}_i)$, valamint $\xi$ egy integrálható valószínűségi változó, akkor

$$E(\xi|\mathcal{G}_i) \longrightarrow E(\xi|\mathcal{G})$$

1-valószínűséggel).

Most tekintsük a

$$\log \sum_{j=1}^{m} P(X_0 = j|Y_{-1},\ldots,Y_{-n}) f(y_0|j) \tag{12}$$

függvényt. (12)-t felülről lehet becsülni egy integrálható függvénnyel,

$$\log \sum_{j=1}^{m} P(X_0 = j|Y_{-1},\ldots,Y_{-n}) f(y_0|j) \le \log \sum_{j=1}^{m} f(y_0|j).$$

Még az kell, hogy $E \log \sum\limits_{j=1}^{m} f(y_0|j) < \infty$. Felhasználva, hogy nemnegatív számokra teljesül, hogy

$$\log(a_1 + \cdots + a_m) \le m \log \max_{1\le j\le m} a_j,$$

így

$$E \log \sum_{j=1}^{m} f(y_0|j) \leq mE \log \max_j f(y_0|j) < \infty,$$

ami éppen a feltétel.

Másfelől (12)-t alulról lehet becsülni egy integrálható függvénnyel.

$$\log \sum_{j=1}^{m} P(X_0 = j|Y_{-1}, \ldots, Y_{-n}) f(y_0|j) \geq \log \frac{1}{m} \min_{1 \leq j \leq m} f(y_0|j),$$

és ez utóbbi (11) feltétel miatt szintén integrálható.

Tehát kaptuk, hogy (12) pontonként 1-valószínűséggel konvergens, valamint van alsó és felső integrálható majoránsa, azaz a Lebesgue-tétel szerint

$$E(-\log p(Y_0|Y_{-1}, \ldots, Y_{-n}))$$

is konvergens, így az állítást bebizonyítottuk.

Bizonyítás nélkül mondunk ki véges értékkészletű $Y_n$ esetén egy általános tételt várhatóérték nélkül.

**5.4. Tétel.** *(Shannon-McMillan-Breiman) Legyen $Y_n$ véges értékkészletű, szigorúan stacionárius, ergodikus folyamat. Ekkor*

$$\lim_{n \to \infty} -\frac{1}{n} \log P(Y_1, \ldots, Y_n) = H$$

*1-valószínűséggel.*

Most ismét térjünk vissza a rejtett Markov modellünkre.

**5.5. Tétel.** *Ha $E|\log f(y|j)| < \infty$ teljesül minden $j$-re, akkor*

$$\lim_{n \to \infty} \frac{1}{n}(-\log p(Y_1, \ldots, Y_n)) = H$$

*1-valószínűséggel.*

*Bizonyítás.* Az előző bizonyításhoz hasonlóan

$$\log p(Y_1, \ldots, Y_n) = \sum_{i=2}^{n} \log p(Y_i | Y_{i-1}, \ldots Y_1) + \log p(Y_1)$$

Tekintsük az

$$U_i = -\log p(Y_i | Y_{i-1}, \ldots)$$

közelítést. Ekkor $U_i$ szigorúan stacionárius folyamat, így rá érvényes az ergod-tétel, azaz

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} U_i = EU_1 = H.$$

Még vizsgálni kell az approximációs hibát. Az $i$-dik tagban ez

$$\log p(Y_i | Y_{i-1}, \ldots, Y_1) - \log p(Y_i | Y_{i-1}, \ldots) = \eta_i.$$

$\eta_i$ nyilván egy valószínűségi változó $\eta_i(T^i \omega)$ alakban, ahol $T$ egy mértéktartó transzformáció az $\Omega$ alaptérben, amivel $Y_n = Y_1(T^{n-1}\omega)$. Másfelől $\eta_i$-re tekinthetünk a jelenben is a következő módon:

$$\eta_i = \log p(Y_0 | Y_{-1}, \ldots, Y_{-i+1}) - \log p(Y_0 | Y_{-1}, \ldots)$$

Tételünk bizonyításához elegendő azt megmutatni, hogy

$$\frac{1}{n} \sum_{i=1}^{n} \eta_i(T^i \omega) \longrightarrow 0$$

1-valószínűséggel.

Ehhez használjuk a Breiman-féle ergodtételt.

**5.6. Tétel.** *(Breiman) Adott a $T$ mértéktartó ergodikus transzformáció $\Omega$-n, valamint az $f_i$ valószínűségi változók, melyekre $f_i \longrightarrow f$ 1-valószínűséggel. Továbbá létezik egy $g$ integrálható majoráns, azaz $|f_i| \leq g$, valamint $Eg < \infty$. Ekkor*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f_i(T^i \omega) = Ef$$

*1-valószínűséggel.*

**5.3. Megjegyzés.** *$f_i = f$ esetén ez éppen a klasszikus ergod-tétel.*

*Bizonyítás.* Legyen

$$h_i = f_i - f$$

Ekkor a tétel feltételei alapján egyrészt $h_i \longrightarrow 0$, másrészt $|h_i| \le |f_i| + |f| \le 2g$ miatt, létezik a $h_i$ sorozatnak is integrálható majoránsa. A klasszikus ergod-tétel alapján tudjuk, hogy

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(T^i\omega) = Ef$$

1-valószínűséggel, így elegendő azt belátni, hogy

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} h_i(T^i\omega) = 0$$

1-valószínűséggel. Azt fogjuk bebizonyítani, hogy a

$$\frac{1}{n}\sum_{i=1}^{n} h_i(T^i\omega)$$

sorozat liminf-je és limsup-ja is 0-hoz tart.

Tekintsük a $h_i^* = \sup_{k\ge i} h_k$ sorozatot. Ekkor a fentiek alapján egyrészt $|h_i^*| \le 2g$, másrészt $h_i^* \longrightarrow 0$ 1-valószínűséggel, azaz a Lebesgue-tétel alapján $Eh_i^* \longrightarrow 0$.

Most rögzítsünk le egy $j$ indexet. Ekkor $h_j^*$ definíciója alapján ha $i \ge j$, akkor $h_j^* \ge h_i$. Térjünk rá a fenti limsup-ra.

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} h_i(T^i\omega) \le \limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{j} h_i(T^i\omega) + \limsup_{n\to\infty} \frac{1}{n}\sum_{i=j+1}^{n} h_i(T^i\omega) \le$$

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{j} h_i(T^i\omega) + \limsup_{n\to\infty} \frac{1}{n}\sum_{j+1=1}^{n} h_j^*(T^i\omega)$$

Az utóbbi összeg első tagja 0-hoz tart, hiszen véges sok korlátos tag összege van $n$-nel osztva, míg a második tagnak a klasszikus ergod-tétel miatt van limesze, mégpedig $Eh_j^*$. Tehát

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} h_i(T^i\omega) \le Eh_j^*$$

Felhasználva, hogy $Eh_i^*$ tart 0-hoz, ha $i \longrightarrow \infty$, $\quad Eh_j^*$ tetszőlegesen kicsinek választható, azaz

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} h_i(T^i\omega) = 0$$

liminf-re a bizonyítás teljesen hasonlóan megy.

*Bizonyítás.* (folytatás) $f_i = \eta_i$ mellett alkalmazzuk a Breiman-tételt. $\eta_i$ definíciója alapján nyilvánvaló, hogy $\eta_i \longrightarrow 0$ 1-valószínűséggel. A közös majoránst, azaz a Breiman-tételben szereplő $g$ megtalálását az olvasóra bízzuk.