

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Keliger Dániel
Matematika BSc

Sztochasztikus folyamatok
átlagtér közelítése

Szakdolgozat

Témavezető: Kunszenti-Kovács Dávid
Alkalmazott Analízis és Számításmatematikai Tanszék



Budapest, 2019

1. Bevezetes

Az alkalmazások során gyakran találkozunk nagy méretű rendszerekkel, melyek sok komponensből állnak, melyek egymással véletlenszerűen interakcióba lépnek. Ilyen rendszerre példa a kémiában két anyag közti reakció, a hővezetés jelensége a fizikában, járványterjedés az epidemológiában, szerver-kliens interakciók az informatikában, két faj együttélésének alakulása a populációdinamikában, vagy az emberek véleményeinek alakulása a sociológiában.

Ami a fentiekben közös, hogy a komponensek állapotait valószínűségi változóként kezeljük, s véletlen interakciókat tételezünk fel köztük. A klasszikus valószínűségszámítási feltételezésekkel szemben itt viszont nem kezelhetjük függetlenként a komponenseket. Ennek hátránya, hogy így a valószínűségek egzakt kiszámítása már közepes méretknél is reménytelen, nagy méretű rendszereknél pedig a szimuláció bizonyul költségesnek.

A hátrányból előnyt kovácsolva mégis azt tapasztaljuk, hogy amennyiben a komponensek számával végtelenbe tartunk, akkor megfelelő skálázás mellett egy determinisztikus rendszerhez jutunk úgy nevezett átlagtér közelítéssel, mely legtöbb esetben egy differenciálegyenlet vagy egy rekuzió alakját ölti. Ez meglehetősen előnyös, mivel a szimulációval szemben jóval kevesebb számítással is adhatóak meg adekvát numerikus közelítések az ilyen determinisztikus folyamatokra.

További előnye ezen szemléletnek, hogy a szóban forgó dinamikai rendszereket leíró differenciálegyenletet vagy rekuziót könnyedén, s intuitív módon kaphatjuk meg a feltételes várhatóérték segítségével. Éppen ezért a gyakorlatban nagy népszerűségnek örvendenek ezek az eszközök.

Azonban az átlagtér közelítés népszerűsége és könnyen interpretálhatósága paradox módon egyben a hátránya is, mivel sok esetben fel sem tűnhet, hogy amit használunk az nem a várható érték, hanem csak annak közelítése, illetve felmerül a veszélye annak, hogy olyan helyzetekben is használjuk az átlagtér közelítést, amikor annak segítségével rossz következtetéseket vonnánk le a közelítendő véletlen rendszerről.

Éppen ezért releváns kérdés az átlagtér közelítés alkalmazhatósági körének vizsgálata.

A dolgozatban először ismertetjük magát a közelítést példákon keresztül, hangsúlyozva olyan eseteket, mikor a mean field közelítés rossz következtetésre jut, majd modellek viszonylag általános családjára bizonyítunk konvergencia tételeket. Külön fejezetben részletezzük a determinisztikus és a sztochasztikus folyamat aszimptotikus viselkedése közötti kapcsolatot. Végezetül vázoljuk, milyen egyéb sztochasztikus folyamatokra lehetne még kiterjeszteni a mean field közelítés konvergenciáját.

Tartalomjegyzek

| | |
|--|----|
| 1. Bevezetes | 2 |
| 2. Motivacio | 4 |
| 3. Diszkret eset | 12 |
| 4. Folytonos eset | 20 |
| 5. A diszkret es a folytonos eset kapcsolata | 28 |
| 6. Aszimptotikus viselkedes | 31 |
| 7. Kitekintes | 35 |

2. Motiváció

1. Definíció. Sztochasztikus folyamat alatt valószínűségi változók egy $(X_t)_{t \in \Gamma}$ rendezett halmazát értjük, ahol minden $t \in \Gamma$ -ra $X_t : \Omega \rightarrow H$ közös eseményhalmazból ugyanabba a halmazba kepez.

Ebben az írásban csupán a $H = \mathbb{R}^d$ esettel fogunk foglalkozni, vagy annak bizonyos részhalmazáival (például \mathbb{Z}^d). Legtöbb esetben Γ indexhalmaz Z -vel vagy $[0; T]$ -vel egyenlő, ahol $0 < T < \infty$. Utóbbi esetben helyett t változót, előbbinél pedig n -et használunk. X_t helyett kényelmesebb az $X(t)$ jelölés használata. Egyrészt így fenntartjuk az alsó index helyét más jelölésekre, másrészt ez jobban kifejezi azt az interpretációt, miszerint a folyamat időtől és a véletlentől - pontosabban az elemi eseménytől - függ.

Megjegyezzük, hogy sztochasztikus folyamat alatt gyakran csak a folytonos indexhalmazt értik. Diszkrét esetben inkább idősorokról beszélünk. Ebben az írásban mind a kettőre egyaránt a sztochasztikus folyamat (vagy röviden folyamat) nevet használjuk.

1. Példa. Determinisztikus folyamatok

Amennyiben $\Omega = \{f, g\}$, a folyamat egy determinisztikus sorozat. Hasonlóan a $[0; T]$ intervallumon értelmezett függvényeket is tekinthetjük speciális sztochasztikus folyamatnak.

2. Példa. Független komponensű folyamatok

Legyenek $X_k(t)$ sztochasztikus folyamatok, melyekre minden $k \in \{1, \dots, N\}$ -ra és minden $t \in [0; T]$ -re létezik $m(t) := E(X_k(t))$ és $\sigma(t) := D(X_k(t)) < \infty$. Ezen kívül tegyük fel, hogy $X_k(t)$ -k függetlenek egymástól! $X(t) := [X_1(t); \dots; X_N(t)]^T$ ekkor maga is sztochasztikus folyamat, ahogy $\bar{X}(t) := \frac{1}{N} \sum_{k=1}^N X_k(t)$ is az. $X_k(t)$ -ket ekkor az $X(t)$ folyamat komponenseinek nevezzük.

Vegyük észre, hogy a nagy számok erős törvénye szerint $\bar{X}(t) \rightarrow m(t)$ m.m. minden rögzített t -re, ha $N \rightarrow \infty$.

Ez az egyszerű észrevétel azért jelentős, mert $\bar{X}(t)$ eloszlásának időbeli alakulása rendkívül bonyolult is lehet. Explicit kiszámítása gyakran reménytelen feladat. Azonban mégis azt tapasztaljuk, hogy amennyiben N elég nagy, az átlag viselkedése leegyszerűsödik egy determinisztikus folyamattá. Ez a feltétel pedig igen gyakran tud teljesülni az alkalmazások során. Gondoljunk csak a bele, hogy például csupán egy mólyi gáz $6 \cdot 10^{23}$ molekulát tartalmaz, de betörések vagy járványok vizsgálatánál is gyakran milliós nagyságrendű egyedszámról beszélhetünk.

A függetlenség ilyen erős változata viszont igencsak leszűkíti az alkalmazhatóság körét. Felmerül tehát az igény a nagy számok törvényének általánosítására sztochasztikus folyamatokra. Előtte viszont érdemes megvizsgálni két olyan alkalmazást, ahol a függetlenség teljesül.

3. Példa. Radioaktív bomlás

Ismert tény, hogy egy radioaktív atom bomlásának ideje exponenciális eloszlást követ, s külső behatás nem befolyásolja ezt a folyamatot. Ezek alapján az atomokat egymástól függetlennek tekinthetjük.

Legyen $X_k(t)$ annak az indikátora, hogy a k : atom nem bomlott még el. Könnyen látható, hogy

$$X_k(t) = e^{-\lambda t}$$

Deriválás után láthatóvá válik, hogy a várható érték kielégíti az alábbi differenciálegyenletet:

$$\frac{d}{dt} \bar{X}(t) = -\lambda \bar{X}(t)$$

Ez analóg a fizikában használatos differenciálegyenlettel, mely az alábbi alakot ölti:

$$\frac{dN(t)}{dt} = -\lambda N(t);$$

ahol $N(t)$ a t időpontban meglévő részecskék számával, vagyis $N(t) = N\overline{X}(t)$.

Ez a megfogalmazás azonban több szempontból is magyarázatra szorul. Egyrészt $N(t)$ valójában egy lépcsős függvény, így a derivált nem értelmezhető, másrészt $N(t)$ nem egy determinisztikus mennyiség, s N növelésével a valódi realizáció nagy valószínűséggel eltér az egyenlet megoldásától. Szerencsére aszimptotikusan mégis egyenlő a két mennyiség, hiszen a de Moivre-Laplace-tétel alapján $O(\frac{1}{\sqrt{N}})$ -nél nagyobb hibát csak exponenciálisan kicsi valószínűséggel vesz fel a rendszer.

4. Pelda. Di uzio

Legyenek $X_i^N(n) : i \geq 1; \dots; N$ egész értékű valószínűségi változók, melyek egymástól független, szimmetrikus bolyongást végző részecskék helyzetét jelentik az n : lépésnél. $X_i^N(0)$ -kat azonosan 0-nak tételezzük fel. Legyen $X_{i;k}^N(n)$ annak az indikátora, hogy $X_i(n) = k$, valamint $x_k^N(n) := \frac{1}{N} \sum_{i=1}^N X_{i;k}^N(n)$. Utóbbi változó az n időpontban k helyen tartózkodó részecskék arányával egyenlő.

Tekintve, hogy i szerint minden $X_{i;k}^N(n)$ független, alkalmazható a nagy számok törvénye, így minden k -ra $x_k^N(n) \rightarrow \kappa(n)$ a fenti értelemben. Már csak $\kappa(n)$ -t kell meghatároznunk.

$$\begin{aligned} \kappa(n) &= E(X_{i;k}^N(n)) = P(X_{i;k}^N(n) = 1) = P(X_i^N(n) = k) = \\ &= P(X_i^N(n) = k | X_i^N(n-1) = l) P(X_i^N(n-1) = l) = \\ &= P(X_i^N(n) = k | X_i^N(n-1) = k-1) P(X_i^N(n-1) = k-1) + \\ &= P(X_i^N(n) = k | X_i^N(n-1) = k+1) P(X_i^N(n-1) = k+1) = \\ &= \frac{\kappa(n-1) + \kappa(n-1)}{2} \end{aligned}$$

A fentiek alapján $\kappa(n)$ -t az alábbi rekurzió határozza meg:

$$\kappa(n+1) = \frac{\kappa(n) + \kappa(n)}{2};$$

A kezdeti feltételt majd később határozzuk meg.

Skálázzuk át az időt az alábbi képpen:

$$\begin{aligned} t &= nh^2 \\ x &= \frac{D}{2kh} \end{aligned}$$

A motiváció a skála mögött az, hogy a szimmetrikus bolyongás n lépés után tipikusan $O(\sqrt{n})$ mértékben mozdul el az eredeti helyétől. Az arányosság miatt általánosan lehetett volna $x = Dkh$ -t is választani, ahol D -t diffúziós állandónak nevezik, de az egyszerűbb számolások miatt mi a $D = \frac{D}{2}$ -t választottuk. (Triviális koordináta-transzformációval egyébként is erre az alakra lehet hozni bármilyen D esetén.) A megfelelő skálán értelmezett $v^h(t; x) := \kappa(n)$ függvényre ez után az alábbi rekurziót kapjuk:

$$v^h(t+h^2; x) = \frac{v^h(t; x + \frac{D}{2h}) + v^h(t; x - \frac{D}{2h})}{2}$$

$$\frac{v^h(t+h^2; x) - v^h(t; x)}{h^2} = \frac{v^h(t; x + \frac{\rho}{2}h) - 2v^h(t; x) + v^h(t; x - \frac{\rho}{2}h)}{(\frac{\rho}{2}h)^2}$$

Vegyük észre, hogy v^h nem más, mint az véges elem módszerrel közelítése a hővezetés egyenletének, mely az alábbi alakot veszi fel:

$$@_t u = @_x^2 u$$

Nem is véletlenül, hiszen a bolyongásnál képzelhetünk "energiacsomagokat", melyek a részecskék között mozognak, ezzel egy egyszerű, viszont sok szempontból kielégítő közelítést kapjuk a diffúzió jelenségének.

A közelítés pontosságához még a kezdeti értékeket kell beállítanunk úgy, hogy $v^h(0; x) = u(0; x)$ fennáljon, ahol $u(0; x)$ előre adott. Belátható, hogy ez valójában $v^h(0; x)$ -re, hanem $w^h(t; x) := v^h(t; x) - \frac{\rho}{2}h$ -ra fog fennállni, s rá ugyanaz a rekurzió érvényes, s így $w^h(t; x)$ fogja $u(t; x)$ -et közelíteni.

A példa természetesen kiterjeszthető tetszőleges dimenzióra. Szerencsére a rekurzió ott is könnyen kezelhető, s tagonként vehetjük a különböző térbeli deriváltakat, kihozva a jól ismert egyenletet:

$$@_t u = \Delta u$$

A továbbiakban olyan esetekkel fogunk találkozni, ahol a függetlenséget már nem biztosíthatjuk csak lépéenként.

5. Pelda. Galton-modell

Ezt a modellt tekinthetjük a klasszikus Malthusi exponenciális növekedés sztochasztikus esetének. Érdekessége, hogy a determinisztikus esettel ellentétben a véletlen fluktuációk következtében ki is halhat egy faj. Ennek valószínűségét a generátorfüggvény nem 0 fixpontjával lehet meghatározni.

Érdekességként megjegyezzük, hogy a modell magyarázatot ad arra, miért lehet viszonylag kevés vezetőknévvel lefedni egy ország lakóinak nagy részét. [1]

Számunkra viszont a könnyen meghatározható várhatóérték miatt érdekes a modell, mely a következőképp fogalmazható meg:

Vegyünk egy állatot, melyből kezdetben egy pár van. Létrehoznak véletlen számú utódot, valamilyen diszkrét eloszlás szerint. Az előző generáció nem tud új utódokat létrehozni, s az utódok is ugyanazzal az eloszlással szaporodnak.

Legyen $N(n)$ a párok száma az n generációban 0-tól számozva. Így $N(0) = 1$. Legyen $Z_{n,k}$ a k : pár által létrehozott utódok száma az n : időpontban (valamilyen sorrendben). Ekkor:

$$N(n+1) = \sum_{k=1}^{N(n)} Z_{n,k}$$

Ezúttal nem tudunk olyan könnyedén várhatóértéket számolni, mint korábban, mivel most egy véletlen tagszámú összegről van szó. Viszont a folyamat rekurzív jellege ad egy ötletet a probléma leküzdésére, mi szerint, ha tudnánk, hogy $N(n)$ -ban hány pár élt, mondjuk m , akkor onnan könnyedén kiszámolhatnánk az átlagot, elvégre:

$$E \sum_{k=1}^{\infty} Z_{n,k} = NE(Z_{n,k}) =: \dots$$

Pontosan ezt teszi lehetővé a feltételes várható érték. Definiáljuk a következő függvényt: $m := E(N(n+1)|N(n) = N)$. A definíció értelmes, mivel a feltételes várható érték tényleg nem függ n konkrét értékétől, mivel $N(n)$ Markov-lánc. Az előző gondolatmenet alapján: $m = mq$. A teljes várható érték tételét felhasználva: $E(N(n)) = E(N(n+1)) = (n+1)$, s így

$$(n+1) = q(n) \quad (n) = q^n:$$

Vegyük észre, hogy a rekurzió ilyen alakra is hozható:

$$(n+1) = (n):$$

Ez több szempontból is előnyös felírás. Egyrészt egy zárt dinamikát ad a várhatóértékre, ami sokkal könnyebben kezelhető így, mint a teljes eloszlás. Másrészt a függvényt a folyamat rekurzív jellege miatt sok más alkalmazásnál is könnyedén elő tudjuk állítani. Harmadrészt pedig nagyon intuitív a képlet állítása: Ha N pár volt előzőleg, később várhatóan (N) pár lesz. Előzőleg várhatóan (n) pár volt, tehát várhatóan (n) pár lesz a következő generációban.

Az érvelés, bár rendkívül intuitív, valójában általában nem működik, mivel a teljes várható érték tétel alapján csak az alábbi állítás igaz:

$$E(N(n+1)) = E(N(n)) \neq (E(N(n))):$$

A fenti esetben valójában azért teljesült az egyenlőség, mert lineáris volt. Amennyiben a második deriváltja nem tűnik el, akkor például egy konvex függvénynél a Jensen-egyenlőtlenség az alábbi információval szolgál:

$$E(N(n+1)) = E(N(n)) > (E(N(n))):$$

Általános esetben, amikor (n) -nek konvex és konkáv szakaszai is vannak, nem tudunk semmilyen következtetés levonni $E(N(n+1))$ és $(E(N(n)))$ viszonyáról. Ez azt a sajnálatos tényt emeli ki, hogy tipikusan a függvény nem szolgáltat zárt rekurzív egyenletet a várható értékre.

Az alábbi heurisztikus gondolatmenet mégis motiválhat minket:

1. Heurisztika.

Legyen X tetszőleges, véges szórású valószínűségi változó. Amennyiben $D(X) = 0$, automatikusan teljesül, hogy $E(X) = (E(X))$ tetszőleges függvényre. Ez természetesen túlságosan erős feltételezés, viszont az alkalmazások során gyakran valamilyen kiátlagolt mennyiségek folyamatait vizsgáljuk, melyeknek ezért kis szórásuk van. Tehát elmondhatjuk, hogy X változó nagy eséllyel közel van az átlagához. Ez alapján egy sorfejtést végezve (feltéve, hogy elég sima):

$$(X) = (n) + \frac{1}{2} D^2(X) + \dots$$

$$E(X) = (E(X)) + \frac{1}{2} D^2(X) + \dots$$

Tehát, ha a változó szórása kicsi, akkor $E(\varphi(X)) \approx \varphi(E(X))$ is fennál, ami pont megfelel annak, hogy a várható értékre egy zárt alakot kapunk közelítőleg.

A fenti heurisztikát az alábbi két tétel segítségével fogalmazhatjuk meg matematikailag korrekt alapon:

1. Tétel. Legyen $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ Lipszitz-folytonos L Lipszitz-konstanssal! $X : \Omega \rightarrow \mathbb{R}^d$. Ekkor

$$E(\varphi(X)) \approx \varphi(E(X)) + L \sum_{i=1}^d \frac{\text{Cov}(X_i, \varphi(X))}{\sigma(X_i)};$$

amennyiben a várható értékek és a szórások léteznek.

1. Bizonyítás.

$$\begin{aligned} E(\varphi(X)) - \varphi(E(X)) &= \sum_{i=1}^d \frac{\partial \varphi_j}{\partial x_i}(E(X)) (X_i - E(X_i)) + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \varphi_j}{\partial x_i \partial x_j}(E(X)) (X_i - E(X_i))(X_j - E(X_j)) + \dots \\ E(\varphi(X)) - \varphi(E(X)) &= L \sum_{i=1}^d \frac{\partial \varphi_j}{\partial x_i}(E(X)) (X_i - E(X_i)) + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \varphi_j}{\partial x_i \partial x_j}(E(X)) (X_i - E(X_i))(X_j - E(X_j)) + \dots \\ E(\varphi(X)) - \varphi(E(X)) &= L \sum_{i=1}^d \frac{\partial \varphi_j}{\partial x_i}(E(X)) (X_i - E(X_i)) + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \varphi_j}{\partial x_i \partial x_j}(E(X)) (X_i - E(X_i))(X_j - E(X_j)) + \dots \end{aligned}$$

□

2. Tétel. Legyen φ egyváltozós, kétszer differenciálható valós függvény, melynek második deriváltja korlátos. $X : \Omega \rightarrow \mathbb{R}$ Ekkor

$$E(\varphi(X)) \approx \varphi(E(X)) + \frac{1}{2} \varphi''(E(X)) \text{Var}(X);$$

ha a várható értékek és a variancia léteznek, és ahol $K = \frac{1}{2} \max_{x \in \Omega} |\varphi''(x)|$

2. Bizonyítás.

Lagrange-féle középértéktételt felhasználva minden $X(\omega)$ számhoz létezik olyan $\xi(\omega) \in [E(X), X(\omega)]$, amire:

$$\begin{aligned} \varphi(X(\omega)) &= \varphi(E(X)) + \varphi'(\xi(\omega))(X(\omega) - E(X)) + \frac{1}{2} \varphi''(\xi(\omega))(X(\omega) - E(X))^2 \\ \varphi(X(\omega)) - \varphi(E(X)) &= \varphi'(\xi(\omega))(X(\omega) - E(X)) + \frac{1}{2} \varphi''(\xi(\omega))(X(\omega) - E(X))^2 \\ E(\varphi(X)) - \varphi(E(X)) &= E(\varphi'(\xi(X))(X - E(X)) + \frac{1}{2} \varphi''(\xi(X))(X - E(X))^2) \\ E(\varphi(X)) - \varphi(E(X)) &= \varphi'(E(X)) E(X - E(X)) + \frac{1}{2} \varphi''(E(X)) E((X - E(X))^2) \\ E(\varphi(X)) - \varphi(E(X)) &= \frac{1}{2} \varphi''(E(X)) \text{Var}(X) \end{aligned}$$

□

A fenti gondolatmenet alapján érdemes bevezetni a következő eljárást:

Legyen Lipsitz-folytonos függvény, melyre teljesül, hogy X értékészletén $(m) = E(X(n+1)|X(n) = m)$! Kezdetben (0) ismert a kezdeti feltételek alapján. A fenti közelítést felhasználva

$$(1) = E(X(0)) = (0)$$

$$(2) = E(X(1)) \quad (E(X(1)) = (1)) = ((0))$$

2. Definíció. Legyen $X(n)$ diszkrét idejű Markov-folyamat, melynek letezik minden n esetén a várható értéke, mely (n) -nel egyenlő. Legyen olyan függvény, mely $X(n)$ értékészletén meggyezik $E(X(n+1)|X(n))$ -nel. Az $m(n+1) = (m(n))$ rekurzio megoldását az $X(n)$ diszkrét idejű sztochasztikus folyamat átlagter közel tesenek nevezzük.

Hasonlóképp definiálhatjuk az átlagter közelítést folytonos idejű folyamatokra. Ekkor a véges differenciákból kiindulva definiálható annak feltételes várható értéke, megfelelő simasággal kiterjesztve.

$$M^h(x) := E(X(t+h) | X(t) = x)$$

A teljes várhatóérték tétele alapján ismét igaz, hogy

$$E(X(t+h) | X(t)) = E(M^h(X(t)))$$

Tegyük fel, hogy az alábbi határérték létezik:

$$M(x) := \lim_{h \rightarrow 0^+} \frac{M^h(x) - x}{h}$$

$M(x)$ -et a sztochasztikus folyamatok nyelvén intenzitásfüggvénynek szokás nevezni.

h -val leosztva és feltéve, hogy felcserélhetjük a várható értéket és a limeszt, megkapjuk, hogy

$$\frac{d}{dt} E(X(t)) = E(M(X(t)))$$

Az 1-es tétel alapján alapján kicsi szórás esetén $E(M(X(t))) \approx M(E(X(t)))$ Ez a következő definíciót motiválja:

3. Definíció. Legyen $X(t)$ folytonos idejű Markov-folyamat, melynek letezik minden $t \in [0; T]$ -re a várható értéke, mely (t) -vel egyenlő. Legyen $M^h(x)$ olyan függvény, melyre teljesül, hogy $X(t)$ értelmezési tartományán meggyezik $E(X(t+h) | X(t) = x)$ -szel. Feltesszük továbbá, hogy letezik az alábbi határérték: $\lim_{h \rightarrow 0^+} M^h(x) = h$. A $\frac{d}{dt} m(t) = M(m(t))$ differenciálegyenletet megoldását a folytonos idejű sztochasztikus folyamat átlagter közel tesenek nevezzük.

6. Példa. A kvantummechanika klasszikus közelítése

A kvantummechanikában egy részecske helyének (jele: x) és momentumának (jele: p) mért értéke valószínűségi változók. Az Ehrenfest-tétel alapján ezek az egyenletek mondhatóak el róluk:

$$m \frac{d}{dt} \langle x \rangle = \langle p \rangle$$

$$\frac{d}{dt} \langle p \rangle = - \langle \nabla F(x) \rangle$$

Ahol $\langle \cdot \rangle$ a várható értéket jelenti, F pedig a potenciális energia deriváltjának -1 -szerese.

Az első egyenletet deriválva még egyszer megkajuk, hogy:

$$m \frac{d^2}{dt^2} hxi = hF(x)i$$

Az átlagtér közelítés a következő egyenletet adja:

$$m \frac{d^2}{dt^2} hxi = F(hxi)$$

Ez pedig nem más, mint Newton második törvénye.

Megjegyezzük, hogy ez a fajta átlagtér közelítés nem teljesen ugyanaz, mint amit fentebb definiáltunk, mivel F nem egy feltételes várható érték. Tágabb értelemben mégis nevezhetjük átlagtér-közelítésnek, hisz a becslés ismét abból áll, hogy felcserélünk egy függvényt és a várható érték funkcionált, annak érdekében, hogy egy zárt formulát kapjunk a dinamikára.

Érdeemes azt is külön kiemelni, hogy a klasszikus mechanika nem mindig ad jó közelítést kvantummechanikai jelenségekre, így nem feltétlenül kell arra számítanunk, hogy az átlagtér közelítés mindig pontos megoldást ad.

Utóbbira mutat rá a következő:

7. Pelda.

Általános- és középiskolákban szokás játszani a következő játékot:

A gyerekek körbeállnak, s lesütik a fejüket, majd a tanár sípolására mindenki ránéz valakire a körből. Amennyiben szemkontaktus alakul ki, a két pár elkezd szaladni egymás felé. Aki később ér oda a másik helyére, kiesik.

Feltesszük, hogy mindenki mindenkit egyenlő valószínűséggel választ ki, s egyenlő valószínűséggel esik ki szemkontaktus esetén. (Utóbbi feltétel elhagyható. Valójában elég annyit feltennünk, hogy minden szemkontaktusnál pontosan egy gyerek esik ki, vagyis nincs döntetlen.)

Legyen $N(n)$ a gyerekek száma az n . körnél 0-tól számozva. Ekkor

$$N(n+1) = N(n) \prod_{i=1}^{N(n)} i(n)$$

ahol $i(n)$ annak az indikátor, a hogy az i . gyerek - valamilyen sorba rendezés szerint - az n . körben kiesik. Ez alapján $m = 2$ esetén:

$$E(N(n+1)|N(n) = m) = m \prod_{i=1}^{m-1} P(i(n) = 1) = m \frac{m}{2(m-1)};$$

mivel bárkire is nézett rá, az $\frac{1}{m-1}$ eséllyel nézett vissza rá, s utána még $\frac{1}{2}$ az esélye annak, hogy a futás során veszít.

$m = 1$ esetén pedig, mivel 1 gyerek esetén nem tud szemkontaktus kialakulni, így

$$E(N(n+1)|N(n) = 1) = 1$$

Tehát megválaszthatjuk $E(N(n+1)|N(n) = m)$ -t az alábbi képpen:

$$E(N(n+1)|N(n) = m) := \begin{cases} m \frac{m}{2(m-1)}; & \text{if } m \neq 1 \\ 1; & \text{if } m = 1 \end{cases}$$

Az átlagtér közelítés tehát, amennyiben feltesszük, hogy a rekurzió nem veszi fel az 1 értéket:

$$m(n+1) = m(n) \frac{m(n)}{2(m(n) - 1)}$$

$$m(0) = N(0)$$

Nyilván $N(n) = 1$ -nek kell teljesülnie, így ez a várható értékre is igaz marad. $m(n)$ -re viszont nem, sőt egészen nagy abszolút értékű negatív értékeket is fel tud venni. Az alábbi táblázat mutat erre egy példát:

| | | | | | | | | | | | |
|------|---|-----|------|------|------|------|------|------|--------|--------|--------|
| n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| m(n) | 6 | 5,4 | 4,79 | 4,15 | 3,50 | 2,80 | 2,02 | 1,03 | -19,21 | -19,68 | -20,17 |

Mindez óvatosságra int minket az átlagtér közelítés alkalmazása tekintetében, illetve azon feltételek felkutatására, melyek mellett a közelítés kielégítőnek bizonyul.

Különösen fontosnak tartjuk ennek a kihangsúlyozását, mivel az alkalmazások terén sok esetben fel sem merül, hogy $m(n)$ valójában nem ugyanaz, mint $E(X(n))$, felmerül, de nem elemzik mélyebben, illetve csupán heurisztikus érveléssel támasszák alá a validitását [2] [3] [4].

3. Diszkrét eset

Használjuk a 2. definícióban lévő jelöléseket! Vegyük észre, hogy a sorfejtéses gondolatmenet - még 0-hoz tartó szórás esetén sem - garantálja, hogy $m(n)$ (n) fennáljon. Ezt jól szemlélteti a hiba komponensekre bontása:

$$m(n+1) - (n+1) = (m(n) - E(X(n))) + (n) - E(X(n)):$$

A hiba második tényezője valóban kicsi a korábbi érvelés szerint, s azt mondja, ha eredetileg ismertük (n) -t, nem fogunk nagy hibát véteni egy lépésben. Az első tag viszont azt írja le, hogy mennyi hibát halmoztunk fel n lépés alatt. Ez még jobban látszik, ha feltesszük, hogy Lipsitz-folytonos. Ekkor ugyanis definíció szerint:

$$|(m(n) - (n)) - (m(n-1) - (n-1))| \leq L|m(n) - m(n-1)|$$

A korábbi érvelés nem mond semmit a felhalmozott hibáról, annak mértékéről további vizsgálódásra van szükségünk.

Definiáljuk a következő mennyiségeket ezúttal vektorváltozókra és tetszőleges normára, illetve ismét tegyük fel, hogy Lipsitz-folytonos!

$$H_n := \|m(n) - (n)\| \\ h_n := \|E(X(n)) - (n)\|$$

H_0 -t nem feltétlenül válasszuk meg 0-nak, viszont kicsi számként fogunk gondolni rá.

A hibák felhalmozására a következő tétel állít egy felsőbecslést:

3. Tétel. Legyen ν pozitív egész és legyen igaz minden n -re $\sum_{i=1}^n D^2(X_i(n)) < \nu^2$. Ezen ν mellett tegyük fel, hogy Lipsitz-folytonos. Ekkor létezik olyan K pozitív, ν -tól nem függő konstans, amire $H_n \leq K(\nu + 1)$; amennyiben $1 \leq n \leq \nu$:

3. Bizonyítás.

Az 1. tétel alapján van olyan $C > 0$ amire $h_n \leq C \|m(n) - (n)\|$.

$$H_{n+1} = \|m(n+1) - (n+1)\| \leq \|m(n) - (n)\| + \|m(n) - E(X(n))\| + \|E(X(n)) - (n)\| \leq H_n + h_n$$

vagyis

$$H_{n+1} \leq H_n + h_n$$

Az egyenlőtlenséget iterálva kijön, hogy:

$$H_n \leq L^n H_0 + h \sum_{k=0}^{n-1} L^k = L^n H_0 + h \frac{L^n - 1}{L - 1} = L^n H_0 + C \frac{L^n - 1}{L - 1} \leq K(\nu + 1);$$

ahol

$$K := \max\{L^n; C \frac{L^n - 1}{L - 1}\};$$

□

1. Megjegyzés. Kontrakciókra

Amennyiben kontrakció - vagyis $L < 1$ - elhagyva az felsőkorlátot n -re, még az alábbi egyenlőtlenség is fennál:

$$\limsup_{n \rightarrow \infty} H_n = \frac{h}{1-L}$$

Tehát, mivel kontrakció, m -nek van pontosan egy stabil egyensúlyi pontja, és az oda való konvergencia valamilyen értelemben öröklődik (n) -re is. (Elég nagy n esetén (n) közel lesz $m(n)$ limeszéhez.)

Ha nem kontrakció, csak a véges idejű, tranzien viselkedésekre tudunk jó becslést biztosítani (n) -re.

Vegyük észre, hogy e körülmények között nem csak azt mondhatjuk el, hogy $m(n)$ közel lesz (n) -hez, hanem magához $x(n)$ -hez is, mivel ismét komponenseire bonthatjuk a hibákat a háromszög-egyenlőtlenség segítségével:

$$\|jX(n) - m(n)\| \leq \|jX(n) - (n)\| + \|(n) - m(n)\|$$

A második hibát tárgyaltuk már. Az elsőre az alábbi tétel mond felsőbecslést:

4. Tétel. Legyen $\prod_{i=1}^d D^2(X_i(n)) \leq \frac{1}{2}$! Ekkor

$$P(\|jX(n) - (n)\| \geq \frac{C}{2}) < \frac{C^2}{2}$$

valamilyen C -tol és C -tol független $C > 0$ konstanssal.

4. Bizonyítás.

$$P(\|jX(n) - (n)\| \geq \frac{C}{2}) = P\left(\sum_{k=1}^d \|jX_k(n) - k(n)\| \geq \frac{C}{2}\right) \leq \sum_{k=1}^d P(\|jX_k(n) - k(n)\| \geq \frac{C}{2})$$

$$\sum_{k=1}^d \frac{D^2(X_k(n))}{\frac{1}{2}} < \frac{2}{2}$$

Mivel minden norma ekvivalens véges dimenziós vektortérben, így létezik C_1 , melyre $C_1 \|jX(n) - m(n)\| \leq \|jX(n) - (n)\|$ és így

$$P(\|jX(n) - m(n)\| \geq \frac{C}{C_1}) \leq P(\|jX(n) - (n)\| \geq \frac{C}{2})$$

$$P(\|jX(n) - m(n)\| \geq \frac{C}{C_1}) \leq \frac{C_1^2}{2} \frac{C^2}{2} =: \frac{C^2}{2}$$

□

Tehát megfelelően kicsi C esetén az átlagtól vett eltérés is kicsi marad.

Felmerül a kérdés, hogy mi garantálja nekünk, hogy C kicsi marad?

2. Heurisztika. Vegyünk egy folyamamatot, mely a 2. példához hasonlóan N komponensből áll. Tegyük fel, hogy a komponensek minden lépésben a multia feltételeken függetlenül lépnek egymástól. Így az N -nel normált vektorok minden egyes komponense $O\left(\frac{1}{N}\right)$ nagyságrendű varianciával rendelkezik. Ha d komponens van, ez csak egy konstans szorzót ad hozzá, így $= O\left(\frac{1}{N}\right)$.

Ez a heurisztika azonban nem feltétlen érvényes, mivel nem csak a hibák, de a varianciák is felhalmozódhatnak. Ennek az az oka, hogy az komponensek csak feltételeken függetlenek egymástól a jelen ismeretének feltételével. Amennyiben a jövőbeli eloszlásukra vagyunk kíváncsiak, nem feltétlen maradnak függetlenek, így a varianciákon kívül a kovarianciákat is figyelembe kell venni.

Abban az esetben, amikor sok komponensből álló rendszert vizsgálunk, a $X(n)$ olyan vektor, melynek koordinátái az adott állapotban lévő komponensek száma, az átlagtér közelítést az $x(n) = \frac{1}{N}X(n)$ folyamatra alkalmazzuk, a feltételes várhatóértéket leíró függvényt is ennek segítségével értelmezzük.

Érdekes ismét komponensekre bontanunk a hiba forrását, a varianciát, ezúttal a teljes variancia tételének segítségével.

$$D^2(x(n+1)) = E(D^2(x(n+1)|x(n))) + D^2(E(x(n+1)|x(n))) = E(D^2(x(n+1)|x(n))) + D^2(x(n))$$

Az $O\left(\frac{1}{N}\right)$ mennyiség itt valójában a feltételes varianciára vonatkozik, s onnan ered, hogy amennyiben ismerjük az előző lépésben a komponensek helyzetét, a következő lépésben valóban egymástól független változnak. Ez nem mondható el viszont a kettővel későbbi lépésre, ugyanis ekkor már számolunk kell a második taggal.

Viszont néhány plusz feltétel bevezetésével a sztochasztikus folyamatok egy széles osztályára érvényes marad a koncentráció. Ehhez be kell vezetnünk a sűrűségfüggő Markov-láncokat, melyeket példákön keresztül illusztrálunk először, majd megfogalmazzuk általánosan, végül a folyamat átlagtér közelítéséhez való konvergenciájára ismertetünk bizonyítást.

8. Példa. A parkapcsolatok minimalista modellje, diszkrét eset

Vegyünk a párkapcsolatoknak az alábbi toy modelljét:

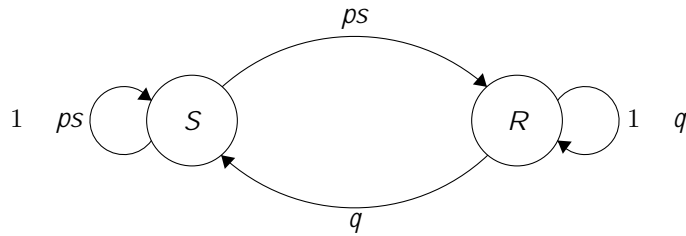
Legyen N egy populációban lévő fiúk száma, mely megegyezik a lányok számával. (Ez a feltétel relaxálható.) Legyen $R(n)$ az n . hónapban a párkapcsolatban lévők száma, $S(n)$ pedig az egyedülállókét! Heteroszexuális, monogám kapcsolatokat feltételezünk, így pontosan ugyanannyi fiúnak kell kapcsolatban lennie, mint lánynak. Nyilván teljesülnie kell annak, hogy $R(n) + S(n) = N$. Ez alapján elég $S(n)$ -t vizsgálunk.

$S(n)$ -t felírhatjuk indikátorok összegeként. Legyen $S_i(n)$ annak az indikátora, hogy az i . fiú (vagy lány) egyedülálló! Ekkor $S(n) = \sum_{i=1}^N S_i(n)$ Ezen kívül kényelmesebb áttérni abszolút mennyiségekről arányokra, melyet N -nel való osztással kaphatunk meg: $s(n) := \frac{1}{N}S(n)$.

A folyamat a következő szabályok szerint zajlik. Minden pár q valószínűséggel szakít egy hónap után, s aki egyedülálló, $ps(n)$ valószínűséggel kerül egy hónap alatt kapcsolatba, ahol $0 < p < 1$. Utóbbi feltevés azt foglalja magába, hogy ha sokan vannak párkapcsolatban, nehezebb találni olyat, akik egyedülálló, s így potenciális partner. Vagyis tetszőleges i -re:

$$\begin{aligned} P(S_i(n+1) = 1 | S_i(n) = 0) &= q \\ P(S_i(n+1) = 0 | S_i(n) = 0) &= 1 - q \\ P(S_i(n+1) = 0 | S_i(n) = 1) &= ps(n) \\ P(S_i(n+1) = 1 | S_i(n) = 1) &= 1 - ps(n): \end{aligned}$$

Mindezt szemléletesebb diagramon ábrázolni.



Fontos megjegyeznünk, hogy $S(n)$ maga Markov-lánc, viszont $S_i(n)$ nem az. Nem elég tudnunk pusztán $S_i(n-1)$ -et a valószínűségek meghatározásához, hanem $S(n)$ -re is szükség van. Épp ezért szokták az ilyen típusú folyamatot sűrűségfüggő Markov-láncnak hívni. A feltételes várhatóértéket kiszámolhatjuk a teljes valószínűség tétele segítségével:

$$(s) := E(s(n+1)|s(n) = s) = (1 - ps)s + q(1 - s)$$

Ez definiálja az átlagtér közelítés egyenletét, melyet érdemes az alábbi alakban felírni:

$$m(n+1) - m(n) = pm^2(n) + q(1 - m(n))$$

Hogy megtaláljuk a rekurzió fixpontját, az alábbi egyenletet kell megoldanunk:

$$pm^2 + q(1 - m) = 0$$

Ennek az egyetlen nemnegatív gyöke

$$m = \frac{p - \sqrt{p^2 + 4}}{2};$$

ahol

$$:= \frac{q}{p};$$

Könnyen ellenőrizhető, hogy

$$\theta(m) = 1 - \frac{p}{q^2 + 4pq}$$

Amennyiben p és q elég kicsik, de nem nullák, a fixpont stabil, ami egyezik is az intuíciónkkal.

9. Pelda. Diszkrét SIS modell

Adott N darab ember, s azoknak két csoportja: egészséges és fertőzött. Előbbiek számát $S(n)$, utóbbiakét $I(n)$ jelöli az n . időpontban.

Nyilván teljesülnie kell minden n -re, hogy $S(n) + I(n) = N$, vagy normált mennyiségekkel számolva: $s(n) + i(n) = 1$.

Elég tehát ismét csak az egyik változóval foglalkoznunk. Legyen ez I !

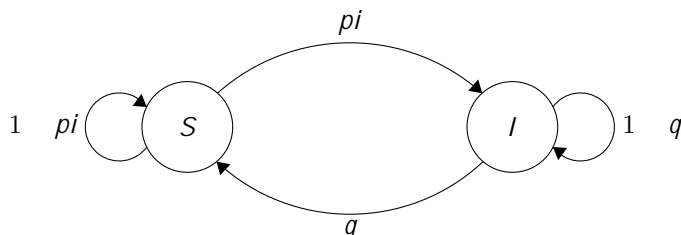
Legyen $p_i(n)$ annak az indikátora, hogy az i . ember az n . időpontban beteg! Az átmeneti valószínűségek a következők:

$$P(i_j(n+1) = 1 | i_j(n) = 0) = pi(n)$$

$$P(i(n+1) = 0 | i(n) = 1) = q$$

A fentiek azt a feltételezést tartalmazzák, hogy amennyiben több a fertőzött, az egészséges komponens könnyebben elkaphatja valakitől a betegséget, a gyógyulás viszont a sűrűségtől független.

Mindez ábrázolva:



Az átlagtér egyenlet ebben az esetben:

$$m(n+1) - m(n) = pm(n)(1 - m(n)) - qm(n);$$

ahol $m(n)$ $i(n)$ -et közelíti. A fixpontokra az egyenlet:

$$pm^2 + (p - q)m = 0$$

A megoldások pedig

$$m_1 = 0$$

$$m_2 = \frac{p - q}{p}.$$

Csupán az az érdekes eset, amikor $p > q$. Ez azt fejezi ki, hogy könnyebb megbetegedni, mint kigyógyulni belőle. Ellenkező esetben azt várnánk, hogy jóval gyorsabban fogyni kezd a betegek száma, mint amilyen gyorsan megbetegednek az emberek. Járványra aligha számíthatnánk.

Ez onnan is látszik, hogy

$$m(n+1) - m(n) < (p - q)m(n);$$

tehát legalább exponenciális gyorsasággal konvergál be $i(n)$ 0-hoz. $p = q$ esetén hasonlóan könnyen kezelhető.

A fenti feltétellel viszont behelyettesítve $i(n)$ -be

$$i(n+1) = 1 + q - p > 1$$

$$i(n+1) = 1 - (p - q)$$

Utóbbi elég kicsi, de nem nulla p és q esetén 0 és 1 között van, tehát a fixpont stabil.

3. Heurisztika. Tekintve, hogy $i(n)$ hamar bekonvergál a fixpontjához, a folyamat meg közel marad az átlagtér-közel téshez, hosszútávon $\frac{p-q}{p}$ betegarányra számíthatunk.

Ez meglehetősen hosszú intervallumon is teljesülhet elég nagy népesség esetén, aszimptotikusan mégis máshogy viselkedik a rendszer.

Ugyanis bármilyen állapotban is van a rendszer, mindig van egy nem 0 valószínűségű lehetőség arra, hogy mindenki kigyógyuljon a betegségből. Vagyis

$$:= \min_i P(i(n+1) = 0 | i(n) = i) > 0:$$

Ha már ebbe az állapotba eljutottunk, akkor onnantól kezdve örökké itt maradunk, mivel nincs beteg, aki meg tudna valakit fertőzni.

Tekintve, hogy

$$\prod_{n=0}^{\infty} P(i(n) > 0) = \prod_{n=0}^{\infty} (1 - p) = (1 - p)^n = \frac{1}{1 - p} < 1$$

Így a Borell-Canteli-lemma alapján

$$P \lim_{n \rightarrow \infty} i(n) = 0 = 1:$$

A fenti két példát általánosíthatjuk az alábbiaképpen:

4. Definíció. Adott N darab sztochasztikus folyamat, $(X_i^N(n))$ melyek értékeszlete az $\mathbb{R}^1; \dots; \mathbb{R}^K$ halmaz. Legyen a $f_i^N(n) = k$ esemény indikátora $X_{i;k}^N(n)$. Legyen $X^N(n) := [X_1^N(n); \dots; X_K^N(n)]^T$ az a vektor, melynek k . komponense $X_k^N(n) = \sum_{i=1}^N X_{i;k}^N(n)$. Az N -nel lenormalt értékeket kisbetűkkel jelöljük. Adott egy A átmeneti matrix, melynek komponensei:

$$A(x) = (p_{ij}(x))_{i,j=1}^K : p_{ij}(x) = P(X_j^N(n+1) = i | X_j^N(n) = j; X^N(n) = x);$$

ahol s egy tetszőleges index $\mathbb{R}^1; \dots; \mathbb{R}^K$ -ből.

Legyen $\Delta_N := \sum_{i=1}^K v_i = 1; N v_i \geq Ng$ és $\Delta := \sum_{i=1}^K v_i = 1g$!
 Megköveteljük, hogy letezzon egy folytonos Δ függvény, hogy

$$(x)_j^N = E(x(n+1) | x^N(n) = x) = A(x)x:$$

Az ilyen tulajdonságokkal rendelkező $x^N(n)$ -t diszkrét idejű sűrűségfüggő Markov-lancnak nevezzük $A(x)$ átmeneti matrixszal.

2. Megjegyzés.

Általánosan az A mátrix - és így a függvény - N -től való függését is megszokták engedni. Ez esetben azt szokás elvárni, hogy a megfelelő folytonos N függvény sorozatra fennálljon, hogy egyenletesen konvergáljon valamilyen függvényhez. Az alábbi konvergencia tételek mind általánosíthatóak erre az esetre is, csupán egy tetszőlegesen kicsi $\max_{x \geq 2} N(x) - (x)$ hibatag jelenik meg.

Megjegyezzük, hogy itt p_{ij} a j -ből i -be való átmenet valószínűségét jelenti. Néhány helyen fordítva írják az indexeket, mert intuitívabb az a jelölés, hogy i -ből j -be való átmenet, mint hogy j -ből i -be. Ez a sorrend azért előnyös mégis, mert megmaradhat az oszlopvektor jelölés és a mátrixszal való balról szorzás.

A folyamat átlagtér közelítése az alábbi rekurzió megoldása:

$$m(n+1) = (m(n)):$$

Az átlagtér közelítés konvergenciájának vizsgálatához előbb vezessük le az alábbi tételt:

5. **Tétel.** Legyen $x^N(n)$ sűrűségfüggő Markov-folyamat! Ekkor minden n és N -re fennal, hogy

$$D^2(x_k^N(n+1)|x^N(n)) = \frac{1}{N}$$

5. **Bizonyítás.**

$$D^2(x_k^N(n+1)|x^N(n)) = D^2 \left(\frac{1}{N} \sum_{i=1}^N X_{i;k}^N(n+1) | x^N(n) \right) = \frac{1}{N^2} \sum_{i=1}^N D^2(X_{i;k}^N(n+1)|x^N(n))$$

Az utolsó egyenlőségénél felhasználtuk, hogy a feltételes valószínűségi mezőn az indikátorok függetlenek egymástól.

$$D^2(X_{i;k}^N(n+1)|x^N(n)) = E(X_{i;k}^N(n+1)^2 | x^N(n)) - 1$$

Tekintve, hogy $X_{i;k}^N(n)$ indikátorváltozó

$$D^2(x_k^N(n+1)|x^N(n)) = \frac{1}{N^2} \sum_{i=1}^N 1 = \frac{N}{N^2} = \frac{1}{N}$$

□

A sűrűségfüggő Markov-láncok átlagtér közelítéséhez való konvergenciáját a lenti tétel biztosítja. Ennek bizonyítása megtalálható a [5] cikk harmadik fejezetében. Apró módosításokkal a tétel a következő:

6. **Tétel.** Legyen $x^N(n)$ sűrűségfüggő Markov-lánc Lipschitz-folytonos függvennyel. Tegyük fel, hogy $x^N(0) \neq m(0)$ sztochasztikusan, ha $N \rightarrow \infty$. Ekkor minden n -re $x^N(n) \rightarrow m(n)$ sztochasztikusan, ha $N \rightarrow \infty$.

3. **Megjegyzés.**

Mint majd a bizonyításból látni lehet, a konvergencia nem feltétlen egyenletes, így az aszimptotikus viselkedésről nem állít semmit a tétel.

6. **Bizonyítás.**

A bizonyításban minden konvergenciánál úgy értjük, hogy $N \rightarrow \infty$ esetén.

Teljes indukciót fogunk alkalmazni. Először is $n = 0$ esetben a feltétel szerint igaz az állítás. Tegyük fel most, hogy $x^N(n) \rightarrow m(n)$ sztochasztikusan. Tekintve, hogy $\frac{1}{k}(x)$ folytonos létezik minden $\epsilon_1 > 0$ -hoz olyan $\delta > 0$, hogy $||x - y|| < \delta \Rightarrow \frac{1}{k}(x) - \frac{1}{k}(y) < \epsilon_1$

Mivel $x^N(n) \rightarrow m(n)$ ezért minden $\epsilon_2 > 0$ -hoz létezik elég nagy N , melyre $P(||x^N(n) - m(n)|| < \epsilon_2) > 1 - \epsilon_2$.

$$\begin{aligned} E \frac{1}{k}(x^N(n)) - E \frac{1}{k}(m(n)) &= \int \frac{1}{k}(x^N(n)) - \frac{1}{k}(m(n)) dP \\ &= \int_{||x^N(n) - m(n)|| < \epsilon_2} \frac{1}{k}(x^N(n)) - \frac{1}{k}(m(n)) dP + \int_{||x^N(n) - m(n)|| \geq \epsilon_2} \frac{1}{k}(x^N(n)) - \frac{1}{k}(m(n)) dP \end{aligned}$$

$$\epsilon_1 + C_1 \epsilon_2$$

Ahol $C_1 = 2 \max_{x \geq 2} \sum_{j=2}^x \frac{1}{j}$.

Tekintve, hogy N elég nagyra választásával ϵ_1, ϵ_2 elég kicsire vehető, így

$$E\left(\sum_{k=1}^N x_k^N(n)\right) \approx E\left(\sum_{k=1}^N m(n)\right):$$

Teljesen analóg módon belátható, hogy

$$E\left(\sum_{k=1}^N x_k(n)\right) \approx E\left(\sum_{k=1}^N m(n)\right);$$

s így

$$\begin{aligned} D^2 \sum_{k=1}^N x_k^N(n) &= E \sum_{k=1}^N x_k^N(n) - E^2 \sum_{k=1}^N x_k^N(n) \approx \\ E \sum_{k=1}^N m(n) - E^2 \sum_{k=1}^N m(n) &= D^2 \left(\sum_{k=1}^N m(n)\right) = 0: \end{aligned}$$

Alkalmazva a teljes variancia tételt és az 5-ös tételt:

$$D^2(x_k^N(n+1)) = \frac{1}{N} + D^2\left(\sum_{k=1}^N x_k^N(n)\right) \approx 0$$

Elérhető tehát, hogy $\sum_{k=1}^K D^2(x_k^N(n)) < \epsilon^2$ legyen tetszőlegesen kicsi ϵ -ra. Hasonlóan az $n=0$ -ra való feltevésünk alapján $H_0 = \sum_{j=1}^N x_j^N(0) = \sum_{j=1}^N m(0)$ is tetszőlegesen kicsire vehető. Alkalmazva tehát a 3-as és 4-es tételt, kapjuk, hogy $x_k^N(n+1) \approx m(n+1)$.

□

4. Folytonos eset

Hogy megértsük a sűrűségfüggő Markov-lánccok folytonos változatát, előbb érdemes megismerkedni a folytonos idejű Markov-lánccal.

5. Definíció. Legyen $(X(t))_{t \geq 0}$ sztochasztikus folyamat \mathbb{R}^K értékeszettel. Folytonos idejű Markov-lánccnak nevezzük, ha teljesül rá, hogy:

$$P(X(t+s) = j | X(s) = i; \mathcal{F}(s)) = P(X(t+s) = j | X(s) = i) =: P_{k,l}(t)$$

A $P(t) := (P_{k,l}(t))$ mátrixot ezúttal is átmeneti mátrixnak fogjuk hívni. Vegyük észre, hogy $P(0) = I$.

Legyen $p_k(t) := P(X(t) = k)$ és $\rho(t) = [p_1(t); \dots; p_K(t)]^T$! A definícióból és a teljes valószínűség tételéből könnyen látható, hogy

$$\rho(t+s) = P(t)\rho(s)$$

A fenti egyenletnek triviális következménye, hogy

$$P(t+s) = P(t)P(s)$$

4. Heurisztika.

$$\dot{P}(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t+\Delta t) - P(t)}{\Delta t} = P(t) \lim_{\Delta t \rightarrow 0} \frac{P(\Delta t) - I}{\Delta t} =$$

$$P(t) \lim_{\Delta t \rightarrow 0} \frac{P(\Delta t) - P(0)}{\Delta t} = QP(t)$$

$$\frac{d}{dt}[P(t)\rho(0)] = \dot{P}(t)\rho(0) \Rightarrow \dot{\rho}(t) =: Q\rho(t)$$

A heurisztika érvényt nyer, ha be tudjuk bizonyítani, hogy $P(t)$ differenciálható. Sőt, a heurisztikából az is látszik, hogy elég $t=0$ pontban ellenőrizni a differenciálhatóságot.

Legyen $\lambda := \inf_{s > 0} \frac{P(s) - I}{s} =: Q$!

Az 5-ös definíció alapján

$$P(\lambda t + s | \lambda t) = P(s | \lambda t);$$

ami a jól ismert örökifjúsági tulajdonság. Ennek következménye, hogy λ exponenciális eloszlású. Legyen ennek a paramétere a_λ !

Egy fontos észrevétel, hogy

$$P(\lambda t) = 1 - e^{-a_\lambda t} = a_\lambda t + O(t^2)$$

Legyen λ az n . ugrás időpontja, és $Y(n) := (Y_n)!$

A Markov-tulajdonság miatt $Y(n)$ egy Markov-lánc. Legyen ennek az átmeneti mátrixa $R_{k,l}$! ($R_{k,k} = 0$.) Ezt felhasználva tekinthetünk a folytonos idejű Markov-lánccokra úgy, mint egy olyan folyamat, ami a_l paraméterű exponenciális eloszlású időt tartózkodik l állapotban majd $R_{k,l}$ valószínűséggel ugrik valamelyik másik k állapotba.

Ismeretes, hogy annak a valószínűsége, hogy legalább két ugrás történik egy Δt hosszú intervallumban $O(\Delta t^2)$.

A fentieket felhasználva $k \neq l$ re

$$\begin{aligned} \frac{P_{k;l}(\Delta t)}{\Delta t} &= \frac{P((\Delta t) = kj(0) = l)}{\Delta t} = \frac{P(l - \Delta t)R_{k;l}P(l + k > \Delta t) + O(\Delta t^2)}{\Delta t} \\ &= a_k R_{k;l} P(k + l > \Delta t) + O(\Delta t) \quad \lim_{t \rightarrow 0^+} \frac{P_{k;l}(\Delta t)}{\Delta t} = a_k R_{k;l} \end{aligned}$$

A második egyenlőség a következőkből jön: Legyen B az az esemény, hogy $[0; \Delta t]$ -n legalább 2 ugrás történik. Ekkor

$$\begin{aligned} P((\Delta t) = kj(0) = l) &= P((\Delta t) = k; B^j(0) = l) + P((\Delta t) = k; \bar{B}^j(0) = l) = \\ &= P(l - \Delta t)R_{k;l}P(k + l > \Delta t) + O(\Delta t^2) \end{aligned}$$

Mivel $P(\bar{B}^j(0) = l) = O(\Delta t^2)$ és ha legfeljebb 1 ugrás lehetett, akkor $(\Delta t) = k$ csak úgy fordulhat elő, hogy legfeljebb Δt idő alatt megtörténik az első ugrás, ami k -ba vezet, majd a következő ugrás már csak több mint Δt idő múlva történhet meg.

$$P_{l;l}(t) = 1 = \prod_{k \neq l} P_{k;l}(t) \quad \lim_{t \rightarrow 0^+} \frac{P_{l;l}(\Delta t) - 1}{\Delta t} = \prod_{k \neq l} R_{k;l} a_k$$

Összefoglalva Q valóban létezik és szerkezete olyan, hogy $q_{k;l} > 0$ ha $k \neq l$ és $q_{l;l} = \prod_{k \neq l} q_{k;l}$.

4. Megjegyzés. Mivel egyrészt

$$P((\Delta t) = lj(0) = l) = 1 - a_l \Delta t + O(\Delta t^2);$$

másrészt

$$P((\Delta t) = lj(0) = k) = 1 - \prod_{k \neq l} P((\Delta t) = kj(0) = l) = 1 - \prod_{k \neq l} q_{k;l} \Delta t + O(\Delta t^2);$$

ezért

$$\begin{aligned} a_l &= \prod_{k \neq l} q_{k;l} \\ R_{k;l} &= \frac{q_{k;l}}{a_k} = \prod_{s \neq l} \frac{q_{k;l}}{q_{s;l}} \end{aligned}$$

A megjegyzés azért fontos, mert lehetőséget ad arra, hogy Q komponenseiből meghatározzuk a_l -t és $R_{k;l}$ -t.

Segítségével definiálhatjuk az alábbiak is a folytonos idejű Markov-lánccokat:

6. Definíció. Adott (t) sztochasztikus folyamat $\{1, \dots, K\}$ értékeszlettel és egy $Q = (q_{k;l})$ matrix, melynek diagonális elemeire teljesül, hogy $q_{l;l} = \prod_{k \neq l} q_{k;l}$ és a nem diagonális elemei pozitívak.

Akkor nevezzük $P(t)$ -t folytonos idejű Markov-láncnak Q átmenet matrixszal, ha amennyiben l állapotban van, $\prod_{k \neq l} q_{k;l}$ parameterű exponenciális eloszlású ideig ebben az állapotban marad, majd egy l -től különböző k állapotba kerül $\prod_{s \neq l} \frac{q_{k;l}}{q_{s;l}}$ valószínűséggel.

A 6. definíció szellemében definiálhatjuk a sűrűségfüggő Markov-lánccal analóg módon a folytonos idejű sűrűségfüggő Markov-lánccot, melyet az egyszerűség kedvéért sűrűségfüggő Markov-folyamatnak fogunk hívni. Diszkrét idejű megfelelőjétől annyiban tér el, hogy itt az átmenetek nem egyenletesen történnek, hanem exponenciális eloszlású idő telik el két állapotváltás között, s az exponenciális eloszlás paramétere szintén függhet a sűrűségtől.

7. Definíció. Adott $\{N_i(t); t \geq 0; i = 1, \dots, N\}$ sztochasztikus folyamatok. Legyen $f_i^N(t) = \lambda_i$ a $k = 1, \dots, N$ indikátor $X_{i,k}^N(t)$, továbbá $X_k^N(t) = \sum_{i=1}^N X_{i,k}^N(t)$ és a belőlük alkotott vektor $X^N(t)$! Az N -nel normált változókat továbbra is kisbetűvel jelöljük. Adott ezen kívül egy $Q = Q(X^N(t))$ matrix a 6. definícióban felsorolt tulajdonságokkal. Ezen kívül meg azt is megköveteljük tőle, hogy folytonos legyen Δ -n.

$x^N(t)$ értéket az alábbi képp módosítjuk időben:

Minden λ_i -hez generalunk egy i exponenciális eloszlású valószínűségi változót, melynek paramétere λ_i amennyiben $X_{i,i}^N(t) = 1$.

Vegyük ezeknek a minimumát, mely legyen τ . Ez 1 valószínűséggel véges és egyértelmű. Ha megsem lenne egyértelmű, a minimum értékeket felvevő i -k közül egyenletes eloszlással kiválasztunk egyet.

Minden $s < \tau$ -ra $x^N(t+s) = x^N(t)$ marad és $x^N(t+\tau)$ -ot úgy módosítjuk, hogy a minimumhoz tartozó i -nel $P(X_i^N(t+\tau) = k) = \frac{q_{k,i}}{\sum_{s \in I} q_{k,i}}$, ahol $k \in I$.

Az ilyen tulajdonságú $x(t)$ folyamatot sűrűségfüggő Markov-folyamatnak nevezzük.

A diszkrét esethez hasonlóan itt is $x^N(t)$ viselkedéséről kívánunk állítani valamit állítani. Legfőképp $\bar{x}(t) := E(x^N(t))$ közelítésére szeretnénk egy zárt formulát kapni, valamint az a körüli fluktuációkról állítani valamit.

A folytonos idejű Markov-lánccok alapján a következőt várjuk:

5. Heurisztika.

$$\dot{\bar{x}}(t) = Q(\bar{x}(t)) \bar{x}(t) =: M(\bar{x}(t))$$

A 2. definícióval összevetve akkor lenne ez átlagtér közelítés, ha belátnánk az alábbi tételt:

7. Tétel. Legyen $x^N(t)$ sűrűségfüggő Markov-folyamat! Ekkor

$$\lim_{h \rightarrow 0^+} \frac{E(x^N(t+h) | x^N(t) = x) - x^N(t)}{h} = Q(x)x$$

A bizonyítás előtt elvégezzünk egy igen fontos átalakítást:

5. Megjegyzés. Poisson-reprezentáció:

Legyenek $Y_{i,k}(t)$ egymástól független, 1 intenzitású Poisson-folyamatok! Ekkor

$$x_k^N(t) = x_k^N(0) + \frac{1}{N} \sum_{i \neq k} \int_0^t Y_{i,k}(s) \sum_{l=0}^N q_{l,k}(x^N(s)) x_l^N(s) ds - \int_0^t Y_{k,i}(s) \sum_{l=0}^N q_{k,l}(x^N(s)) x_k^N(s) ds$$

A fenti a következők miatt igaz: három részből áll, hányan vannak k -ban az adott időpillanatban, mégpedig hányan voltak eredetileg, hányan jöttek be a k -ba, s hányan jöttek ki belőle.

Az első tagon nincs mit magyarázni. A másodiknál az integrál belseje egy lépcsős függvény. Tudjuk, hogy t -ből k -ba $q_{k,l}(x(s))$ rátával jönnek s -időpontban, és összesen $N x_l(s)$ -en vannak

l -ben. Tehát az l -ből érkezők száma épp egy Poisson-folyamatnak tekinthető, melyet minden ugrás során módosítunk az új sűrűségekkel, két ugrás között viszont konstans az intenzitás.

Az összeg azért kell, mert bármelyik más állapotból jöhetnek k -ba, az növekedést eredményez annál. N -nel pedig azért kell leosztani, mert a sűrűség $\frac{1}{N}$ -nel változik minden ugrásnál.

A harmadik tag analóg a másodikkal. Az egyetlen különbség, hogy ott a k állapotból távozókat tartjuk számon.

7. Bizonyítás.

Az általánosság csorbítása nélkül feltehető, hogy $t = 0$.

Jelölje $E(x^N(0) = x)$ -et $E_x(\cdot)$!

$$\begin{aligned} \frac{1}{N} \sum_{l \neq k} E_x \left(\sum_{i=1}^N Y_{k,l} \right) &= \frac{1}{N} \sum_{l \neq k} E_x \left(\sum_{i=1}^N \int_0^h q_{k,l}(x^N(s)) x_i^N(s) ds \right) \\ &= \sum_{l \neq k} E_x \left(\int_0^h q_{k,l}(x^N(s)) x_l^N(s) ds \right) \\ &= \sum_{l=1}^N E_x \left(\int_0^h q_{k,l}(x^N(s)) x_l^N(s) ds \right) \\ \lim_{h \rightarrow 0^+} \frac{E_x(x_k^N(h) - x_k)}{h} &= \sum_{l=1}^N \lim_{h \rightarrow 0^+} E_x \left(\frac{1}{h} \int_0^h q_{k,l}(x^N(s)) x_l^N(s) ds \right) = \sum_{l=1}^N q_{k,l}(x^N(0)) x_l^N(0) \end{aligned}$$

Az utolsó egyenlőség következő gondolatmenetből jön: Legyen B az az esemény, hogy h idő alatt egyik állapotban sem történik változás. Ekkor

$$\begin{aligned} E_x \left(\frac{1}{h} \int_0^h q_{k,l}(x^N(s)) x_l^N(s) ds \right) &= \\ E_x \left(\frac{1}{h} \int_0^h q_{k,l}(x^N(s)) x_l^N(s) ds \mid B \right) P(B) + E_x \left(\frac{1}{h} \int_0^h q_{k,l}(x^N(s)) x_l^N(s) ds \mid \bar{B} \right) P(\bar{B}) &= \\ \frac{1}{h} \int_0^h q_{k,l}(x^N(0)) x_l^N(0) ds (1 - O(h)) + O(h) &= q_{k,l}(x^N(0)) x_l^N(0) + O(h) \end{aligned}$$

□

A továbbiakban szeretnénk belátni, hogy az 5. heurisztika nem alaptalan. Ennek előkészítéséül belátjuk az alábbi tételt:

8. Tétel. Poisson-folyamatok nagy számok törvénye:

Legyen $N(t)$ egy 1 intenzitású Poisson-folyamat!

Ekkor minden $T > 0$ -ra

$$\sup_{0 \leq t \leq T} \frac{N(nt)}{n} \xrightarrow{m \rightarrow \infty} t \quad t \in [0, 1]$$

ahol $n \rightarrow \infty$.

8. Bizonyítás.

Legyen $M(t) := \frac{N(nt)}{n}$ és $E_m(\cdot) := E(\cdot | M(t) = m)$! (A Markov-tulajdonság miatt elég ezt vizsgálni, s nem kell a teljes filtráció.) Legyen továbbá $N^0(ns) := N(n(s+t)) - N(nt)$! Nyilván N^0 is 1 rátájú Poisson-folyamat, s független N -től.

$$\begin{aligned} E_m(M(t+s)) &= E_m \left(\frac{N(n(t+s))}{n} \right) = E_m \left(\frac{N^0(ns)}{n} + \frac{N(nt)}{n} \right) = \\ &= E_m \left(\frac{N^0(ns)}{n} \right) + E_m \left(\frac{N(nt)}{n} \right) = \\ &= E_m \left(\frac{N^0(ns)}{n} \right) + m \end{aligned}$$

Tehát $M(t)$ egy nemnegatív szupermartingál, így alkalmazható rá a Doob-egyenlőtlenség.

$$P \left(\sup_{0 \leq t \leq T} \frac{N(nt)}{n} \geq m \right) \leq \frac{E(M(T))}{m} = \frac{T}{n^2 m}$$

Legyen $n = k^2$! Mivel $\sum_{k=1}^{\infty} \frac{T}{k^2} < \infty$, így a Borel-Cantelli-lemma alapján

$$\sup_{0 \leq t \leq T} \frac{N(k^2 t)}{k^2} < \infty \quad m.m.;$$

ahol $k \rightarrow \infty$.

A maradék esetre tegyük fel, hogy $k^2 < n < (k+1)^2$!

Tekintve, hogy minden elemi esemény esetén

$$\begin{aligned} \frac{N(k^2 t)}{k^2} &\leq \frac{N(n t)}{n} \leq \frac{N((k+1)^2 t)}{(k+1)^2} \\ \frac{N(k^2 t)}{k^2} &\leq \frac{N(k^2 t)}{k^2} + \frac{N(n t)}{n} - \frac{N(k^2 t)}{k^2} \\ \frac{N(k^2 t)}{k^2} &\leq \frac{N(k^2 t)}{k^2} + \frac{N(n t)}{n} - \frac{N(k^2 t)}{k^2} \end{aligned}$$

Mivel mindkét oldal majdnem mindig 0-hoz tart, így a rendőrelv alapján

$$\sup_{0 \leq t \leq T} \frac{N(n t)}{n} < \infty \quad m.m.;$$

□

Ennek segítségével beláthatjuk a következő, alapvető tételt, mely Kurtztól származik [6]:

9. Tétel. Kurtz-tétel:

Legyen $X^N(t)$ szűrszűrt Markov-folyamat $Q(x)$ átmenetmatrixszal, es legyen $M(x) = Q(x)x$ Lipsitz-folytonos! Tegyük fel, hogy $X^N(0) = m(0)$ m.m. ! Ekkor

$$\lim_{N \rightarrow \infty} \sup_{0 \leq t \leq T} \|X^N(t) - m(t)\| = 0 \quad m.m.;$$

ahol $m(t)$ az alábbi differenciálegyenlet megoldása $m(0)$ kezdeti értékkel:

$$\dot{m}(t) = M(m(t)):$$

6. Megjegyzés. Az is belátható lenne, hogy létezik olyan $\epsilon > 0$ valós számmal korlátos, N -től független valós számú változó, melyre

$$\sup_{0 \leq t \leq T} \|jx^N(t) - m(t)\| \leq \frac{\epsilon}{N}$$

Ez egyben azt is jelenti, hogy csupán kis valós számmal fordulhat elő, hogy a hiba nagyobb, mint $O(\frac{1}{N})$, amennyiben hajlandóak vagyunk a konstanszt kellően nagyra választani [7].

Ennek bizonyítása azonban olyan apparátust használ, mely meghaladja jelen írás kereteit.

9. Bizonyítás.

Az 5. megjegyzést felhasználva

$$\begin{aligned} & \times \frac{1}{N} Y_{k;l} \int_0^t q_{k;l}(x^N(s)) x_l^N(s) ds - \int_0^t q_{k;l}(x^N(s)) x_l^N(s) ds + \\ & \times \frac{1}{N} Y_{k;l} \int_0^t q_{l;k}(x^N(s)) x_k^N(s) ds - \int_0^t q_{l;k}(x^N(s)) x_k^N(s) ds + \\ & \int_0^t M_k(x^N(s)) - M_k(m(s)) ds : \end{aligned}$$

Az első három tagot nevezzük el az egyszerűség kedvéért $H_k(t)$ -nek és legyen $H(t) := \max_k H_k(t)$!

$$\begin{aligned} \|jx(t) - m(t)\| & \leq H(t) + \int_0^t M(x(s)) - M(m(s)) ds \\ & \leq H(t) + L \int_0^t \|jx(s) - m(s)\| ds \end{aligned}$$

Így a Grönwall-egyenlőtlenséget alkalmazva:

$$\|jx(t) - m(t)\| \leq H(t) e^{Lt} \leq \sup_{0 \leq t \leq T} H(t) e^{LT} \leq \sup_{0 \leq t \leq T} H(t) e^{LT}$$

Elegendő tehát belátni, hogy $\sup_{0 \leq t \leq T} H(t) \rightarrow 0$ $m \rightarrow m$:

$x^N(0) \rightarrow m(0)$ m.m. miatt a kezdeti értékből eredő hiba 0-hoz tart.

Válasszunk ki tetszőleges $l \neq k$! Ekkor

$$\int_0^t q_{k;l}(x^N(s)) x_l^N(s) ds - \int_0^t q_{k;l}(x^N(s)) x_l^N(s) ds \leq \max_{k;l} \int_0^T q_{k;l}(x) ds$$

ahol $\max_{k;l} q_{k;l}(x) =: \max_{l \neq k} q_{l;k}(x)$. Legyen $L := \max_{l \neq k} q_{l;k}(x)$!

$$\sup_{0 \leq t \leq T} \frac{1}{N} Y_{k;l}(Nt) = \int_0^t q_{k;l}(x^N(s)) x_l^N(s) ds \quad \mathbb{E} \int_0^t q_{k;l}(x^N(s)) x_l^N(s) ds$$

$$\sup_{0 \leq t \leq T} \frac{1}{N} Y_{k;l}(Nt) \rightarrow \int_0^t m(s) ds$$

Utóbbi pedig a Poisson-folyamatok nagyszámok törvénye eszerint 0-hoz tart m.m. Tekintve, hogy ilyenek összegével becsülhető felülről $\sup_{0 \leq t \leq T} H(t)$ (ahol néha fel kell cserélni l és k szerepét a kilépő komponenseket vizsgáló tagoknál) plusz a kezdeti értékből eredő hibával, ezért az is 0-hoz tart m.m. Tehát

$$\sup_{0 \leq t \leq T} \frac{1}{N} Y_{k;l}(Nt) \rightarrow \int_0^t m(s) ds \quad m: m$$

Mivel a normák ekvivalensek véges dimenziós vektorterekben, minden más normában is igaz marad az állítás. □

7. Megjegyzes. *Bar azt gondolnánk, hogy korlatos intervallumon az idő szerinti szupremummal egy erősebb állítást bizonyítottunk be a diszkrét esetben látott pontonkenti konvergenciahoz képest, viszont mindketten esetekben felső korlatot szabtuk az időnek, s diszkrét esetben ez véges sok időpontot is jelent. Véges sok konvergáló elemmel pedig mindig választható közös.*

Kurtz-tételét felhasználva nézhetünk olyan modelleket, melyek analógiái a diszkrét idejű sűrűségfüggő Markov-láncoknak. A folytonos idejű modelleket sok esetben valóságosabbnak érezzük, mivel az a feltétel, hogy csak a periódus végén történjenek változások, erős megkötés tud lenni. Különösen akkor, ha nem elhanyagolhatóak azok az események, amikor egy intervallumon belül több mint egy változás történik.

10. Pelda. *A parkapcsolatok minimalista modellje, folytonos eset*

Feltesszük, hogy rátával lesz egy kapcsolatban lévő személyből egyedülálló, s amennyiben az egyedülállók aránya s , a kapcsolatba kerülés rátája legyen s !

Az átlagtér egyenlet az egyedülállókra ekkor

$$\dot{s}(t) = -s(t) + s^2(t)$$

Természetesen az egyenlet nagyon hasonló a diszkrét megfelelőjéhez. A pozitív fixpont például ezúttal is $s = \frac{2+4}{2}$, ahol $\frac{2+4}{2} = -$. Könnyen ellenőrizhető, hogy a fixpont ismét stabil lesz.

Egy pillanatra térjünk vissza a diszkrét modellre! Vezessük be az $u_N(t) := m(bNtc)$ függvényt. Legyen $\Delta t = \frac{1}{N}$! Ezen kívül még tegyük fel, hogy $p = \Delta t + O(\Delta t^2)$ és $q = \Delta t + O(\Delta t^2)$!

A fenti skálázás azt a problémát hivatott kiküszöbölni, hogy túl hosszú legyen egy intervallum. Ha besűrítjük a folyamatot úgy, hogy kis valószínűséggel történjen esemény egy intervallumban, ugyanakkor a lépések számát ezzel együtt megnöveljük, akkor arra számíthatunk, hogy a kettő kiegyensúlyozza egymást, s a folytonos modellt kapjuk vissza. Fordítva gondolkodhatunk úgy is, hogy folytonos folyamatot diszkrétizáltuk Δt hosszú intervallumokra, s ekkor az exponenciális eloszlásból épp a fenti átmenet valószínűségeket kapjuk.

Ez igazolódni látszik egyrészt abból, hogy

$$\frac{u_N(t + \Delta t) - u_N(t)}{\Delta t} = -u_N(t) + u_N^2(t) + O(\Delta t)$$

$$\lim_{N \rightarrow \infty} u_N(t) = s(t):$$

Viszont ez nem jelenti még feltétlen azt, hogy sztochasztikus folyamatunk is közel lesz $s(t)$ -hez, mivel egyrészt a konvergenciát csak korlátos lépésszámmal bizonyítottuk diszkrét esetben, itt meg $O(N)$ lépésre volna szükségünk; másrészt az átmenet valószínűségek is folyton változnak N függvényében, melyet nem engedtünk meg korábban.

Az állítás mégis igaz, ráadásul általánosan megfogalmazva is. Ezt a következő fejezetben fogjuk belátni.

A fejezet lezárásául megemlítjük még, hogy a folytonos idejű Markov-láncok kiterjeszhetőek pár, hármas, stb. interakciókra, s azok segítségével legitimizálhatóak azok a differenciálegyenletek, melyeket kémiai reakciók leírására alkalmaznak jól megkevert közeg esetén.

5. A diszkrét es a folytonos eset kapcsolata

A 10-es példa sejtését most megfogalmazzuk általánosan, s bizonyítást is adunk rá.

Adott egy $Q(x)$ átmenetmátrix, mely rendelkezik egy folytonos idejű Markov-lánchoz tartozó átmenetmátrix tulajdonságaival. Válasszuk meg $X^N(n)$ diszkrét idejű Markov-lánc $A(x)$ átmenetmátrixát úgy, hogy $k \neq l$ esetén legyen $p_{k,l}(x) = \frac{q_{k,l}(x)}{N} + o\left(\frac{1}{N}\right)$ és $k = l$ esetén legyen $p_{l,l}(x) = 1 - \sum_{k \neq l} p_{k,l}(x)$

Legyen $Y^N(t) := X^N(bNtc)$. valamint $y'(t)$ megoldása az alábbi differenciálegyenletnek :

$$y'(t) = F(y(t)) := Q(y(t))y'(t)$$

10. Tétel. Legyen F Lipsitz-folytonos, es tegyük fel, hogy $y^N(0) \rightarrow y(0)$ m.m.! Ekkor minden $T > 0$ -ra

$$\sup_{0 \leq t \leq T} \|y^N(t) - y(t)\| \rightarrow 0 \text{ m.m.};$$

ha $N \rightarrow \infty$.

10. Bizonyítás.

A bizonyítás egy átfogalmazása a [8]-ban látottaknak.

Legyen $\frac{1}{N}F^N(x) := E(x^N(n+1) - x^N(n) | x^N(n) = x)$!

$$\frac{1}{N}F^N(x) = A(x)x - x = (A - I)(x)x = \frac{1}{N}Q(x)x + o\left(\frac{1}{N}\right) = \frac{1}{N}F(x) + o\left(\frac{1}{N}\right)$$

Világos, hogy $F^N \rightarrow F$ Δ -n egyenletesen.

Legyen $\frac{1}{N}G(n+1) := x^N(n+1) - x^N(n)$ és $U(n+1) := G(n+1) - F(x(n))$!

$$x^N(n+1) - x^N(n) = \frac{1}{N} (U(n+1) + F(x(n)))$$

$$x^N(n) = x^N(0) + \frac{1}{N} \sum_{l=1}^n U(l) + \sum_{l=0}^{n-1} F(x^N(l))$$

$$y^N(t) - y'(t) = y^N(0) - y'(0) + \frac{1}{N} \sum_{l=1}^{bNtc} U(l) + \frac{1}{N} \sum_{l=0}^{bNtc-1} F(x^N(l)) - \int_0^t F(y'(s)) ds$$

$$y^N(t) - y'(t) = y^N(0) - y'(0) + \frac{1}{N} \sum_{l=1}^{bNtc} U(l) + \int_0^t F(y^N(s)) - F(y'(s)) ds$$

$$\|y^N(t) - y'(t)\|_2 \leq \|y^N(0) - y'(0)\|_2 + \frac{1}{N} \sum_{l=1}^{bNtc} \|U(l)\|_2 + L \int_0^t \|y^N(s) - y'(s)\|_2 ds$$

Alkalmazzuk a Grönwall-egyenlőtlenséget!

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T \dot{y}^N(t) \dot{y}^N(t) dt \mid \mathcal{F}_0 \right] = \mathbb{E} \left[\int_0^T \dot{y}^N(0) \dot{y}^N(0) dt + \frac{1}{N} \sum_{l=1}^{bNTc} U(l) e^{Ll} \right] \\
& \sup_{0 \leq t \leq T} \mathbb{E} \left[\int_0^T \dot{y}^N(t) \dot{y}^N(t) dt \mid \mathcal{F}_0 \right] = \mathbb{E} \left[\int_0^T \dot{y}^N(0) \dot{y}^N(0) dt + \frac{1}{N} \max_{n \leq bNTc} \sum_{l=1}^n U(l) e^{Ll} \right]
\end{aligned}$$

A kezdeti értékből eredő hiba a feltétel alapján 0-hoz tart. A maradék hibát ontuk két részre az alábbiaképp:

$$U(n+1) = G(n+1) - F(x^N(n)) = G(n+1) - F^N(x^N(n)) + F^N(x^N(n)) - F(x^N(n)) =: V(n+1) + W(n+1):$$

$$\begin{aligned}
& \frac{1}{N} \sum_{l=1}^n U(l) = \frac{1}{N} \sum_{l=1}^n V(l) + \frac{1}{N} \sum_{l=1}^n W(l) \\
& \frac{1}{N} \sum_{l=1}^n V(l) + T \sup_{x^2} \dot{y} F^N(x) - F(x) = \frac{1}{N} \sum_{l=1}^n V(l) + o(1)
\end{aligned}$$

Tehát már csak az első összeget kell közelíteni. Legyen $Z(n) := \sum_{l=1}^n V(l)$ és $F_n = (x^N(1); \dots; x^N(n))$!

$$E(Z(n+1) | \mathcal{F}_n) = Z(n) + E(V(n+1) | \mathcal{F}_n) = Z(n);$$

mivel

$$\begin{aligned}
E(G(n+1) | \mathcal{F}_n) &= E(G(n+1) | x^N(n)) = F^N(x^N(n)) \\
E(V(n+1) | \mathcal{F}_n) &= E(V(n+1) | x^N(n)) = 0:
\end{aligned}$$

Tehát $Z(n)$ martingál, s így $\int_0^T \dot{y}^N(t) \dot{y}^N(t) dt$ pozitív szupermartingál. Így a Doob-egyenlőtlenség alapján

$$\begin{aligned}
P\left(\frac{1}{N} \max_{n \leq bNTc} \int_0^T \dot{y}^N(t) \dot{y}^N(t) dt \leq 2N^2\right) &= P\left(\max_{n \leq bNTc} Z(n) \leq 2N^2\right) \\
&= \frac{1}{2N^2} \sum_{k=1}^{bNTc} E \left[\sum_{l=1}^k V_k(l) \right] = \frac{1}{2N^2} \sum_{k=1}^{bNTc} E \left[V_k^2(l) + \sum_{1 \leq s < l \leq bNTb} E(V_k(s)V_k(l)) \right]
\end{aligned}$$

Tekintve, hogy $E(V(l+1)) = 0$, így az első összeg valójában a varianciák összege. A második összeg pedig 0 a teljes kovariancia tétel alapján, tekintve, hogy $l > s$ esetén

$$E(V_k(s)V_k(l)) = E(E(V_k(s)V_k(l) | \mathcal{F}_{l-1})) = E(V_k(s)E(V_k(l) | \mathcal{F}_{l-1})) = E(V_k(s) \cdot 0) = 0;$$

A fentiek alapján így

$$P\left(\frac{1}{N} \max_{n \leq bNTc} \int_0^T \dot{y}^N(t) \dot{y}^N(t) dt \leq 2N^2\right) \geq \frac{1}{N^2} \sum_{l=1}^{bNTc} \sum_{k=1}^l D^2(V_k(l)):$$

A teljes variancia tétele szerint tehát

$$D^2(V_k(I+1)) = E D^2(V(I+1)jX^N(I)) + D^2 E(V_k(I+1)jX^N(I)) = E D^2(V_k(I+1)jX^N(I)) :$$

Más felől $D^2(V_k(I+1)jX^N(I)) = D^2(G_k(I+1)jX^N(I))$, ami felírható binomiálisak valószínűségi változók varianciájának összegeként.

$$\begin{aligned} \sum_{k=1}^{\infty} D^2(V_k(I)) &= \sum_{k=1}^{\infty} E D^2(G_k(I+1)jX^N(I)) \\ &= \sum_{k=1}^{\infty} \sum_{s \neq k} E N X_s^N(I) p_{k;s}(X^N(I)) + N X_k^N(I) p_{s;k}(X^N(I)) = \\ &= 2 \sum_{k=0}^{\infty} \sum_{s \neq k} E N X_s^N(I) p_{k;s}(X^N(I)) = 2 \sum_{k=1}^{\infty} \sum_{s \neq k} E q_{k;s}(X^N(I)) X_s^N(I) + o(1) \end{aligned}$$

Legyen $\tilde{q}_k := \max_{s \neq k} \sup_{X \geq 2} q_{k;s}(X)$!

$$2 \sum_{k=1}^{\infty} \sum_{s \neq k} q_{k;s}(X^N(I)) X_s^N(I) \leq 2 \sum_{k=1}^{\infty} \tilde{q}_k \sum_{s=1}^{\infty} X_s^N(I) = 2 \sum_{k=1}^{\infty} \tilde{q}_k := C$$

Tehát

$$P\left(\frac{1}{N} \max_{n \leq bNTc} \sum_{j=1}^m Z(n) \geq \frac{CNT + o(N)}{2N^2}\right) = O\left(\frac{1}{N}\right)$$

Legyen $B(N) := \max_{0 \leq n \leq bNTc} \sum_{j=1}^m Z(n)$!

A Borele-Cantelli-lemma alapján ismét elmondhatjuk, hogy $\frac{1}{k^2} B(k^2) \rightarrow 0$ m.m. Itt is teljesül, hogy $B(N) \leq B(N+1)$, tehát elvégezve ugyanazt, mint a Poisson-folyamatok nagy számok törvényénél, megkapjuk, hogy

$$\lim_{N \rightarrow \infty} \frac{1}{N} \max_{n \leq bNTc} \sum_{j=1}^m Z(n) = 0 \quad m:m:$$

Vagyis

$$\sup_{0 \leq t \leq T} \sum_{j=1}^m Y^N(t) \rightarrow 0 \quad m:m:$$

Természetesen a normák ekvivalenciája miatt a többi normában is ugyanúgy igaz marad az állítás.

□

6. Aszimptotikus viselkedés

A korábbi fejezetekben megalapoztuk az átlagtér közelítés használatát a tranziens szakaszon, vagyis olyan esetben, amikor az időt egy tetszőlegesen hosszú, de korlátos szakaszon vizsgáljuk. Ezen a szakaszon a rekurziókat vagy a differenciálegyenletünket vagy megoldását vagy meg tudjuk adni explicit alakban, vagy numerikusan tudjuk közelíteni jóval kevesebb számítási igénnyel, mintha a nagy méretű rendszert kellene szimulálnunk.

Azonban a dinamikus rendszerek vizsgálatánál nagyobb hangsúlyt szokott kapni az aszimptotikus viselkedés vizsgálata. Felmerül így a kérdés, milyen következtetést tudunk levonni az átlagtér dinamikájának aszimptotikus viselkedéséből a véletlen rendszer aszimptotikus viselkedésére? Egyenlettel megfogalmazva: igaz-e az alábbi?

$$\lim_{n \uparrow} \lim_{N \uparrow} x^N(n) = \lim_{N \uparrow} \lim_{n \uparrow} x^N(n)$$

Mint ahogy a diszkrét SIS esetében láttuk, ez általában nem tud teljesülni, mivel a várttal ellentétben nem a stabil fixponthoz konvergált a rendszer rögzített N esetén, hanem az instabilhoz 1 valószínűséggel.

Valamilyen szempontból extrémebb ellenpélda a következő:

11. Példa. Wright-Fisher modell

A modell részletes elemzéséhez lásd: [9].

Két féle állélt tételünk fel: a és A típusúakat. Tegyük fel, hogy minden generációban ugyanannyi, összesen N élőlény van, s az n . generációban jelölje az a -val rendelkezők arányát $x^N(n)$.

Feltesszük, hogy a következő generációban minden egyed $x^N(n)$ valószínűséggel örököl a állélt, és $1 - x^N(n)$ valószínűséggel A -t. Tehát $X^N(n+1)$ N -ed rendű $x(n)$ paraméterű binomiális valószínűségi változó. Ezek alapján:

$$\begin{aligned} \dot{x} &= x \\ m(n+1) &= m(n) \quad m(n) = m(0): \end{aligned}$$

A mean field egyenlet tehát rendkívül triviális. Minden pontja fixpont. A linearitás miatt ráadásul $m(n)$ megegyezik a várható értékkel.

Vegyük észre, hogy az SIS modellhez hasonlóan van két állapotunk, amikbe hogyha belekerül a rendszer, onnan nem tud többet kijönni. Ezek pedig az $x^N(n) = 0$ és az $x^N(n) = 1$ állapotok. Bármelyik másik állapotból nem nulla a valószínűsége annak, hogy a kettő valamelyikébe kerül, így 1 valószínűséggel igaz, hogy $y := \lim_{n \uparrow} x^N(n) \in \{0, 1\}$.

Legyen $P(y = 1) = q$! Mivel $x^N(n)$ korlátos és martingál, ezért

$$q = P(y = 1) = E(y) = E \left[\lim_{n \uparrow} x^N(n) \right] = \lim_{n \uparrow} E \left[x^N(n) \right] = x^N(0):$$

Összefoglalva azt kapjuk, hogy amennyiben rögzítjük $X^N(0)$ -t q -ra, akkor

$$q = \lim_{n \uparrow} \lim_{N \uparrow} x^N(n) \neq \lim_{N \uparrow} \lim_{n \uparrow} x^N(n) = y:$$

Ahol az utolsó N szerinti limesz eloszlásban való konvergenciaként értendő.

A fenti ellenpéldák ellenére a párcapcsolatok minimalista modelljében mégis az történik, amire heurisztikusan számítottunk, hogy a rekurzió stabil fixpontja körül marad a véletlen folyamat.

Ez azért is érdekes, mert ha csak nem 0-ból indítjuk az SIS és a párkapcsolati folyamatokat - ami az elsónél nem túl releváns -, akkor a rekurziók hasonló viselkedést mutatnak, a sztochasztikus folyamatok aszimptotikus viselkedése mégis radikálisan különböző.

A fenti állítás bizonyítását egy általánosabb tétel segítségével fogjuk megkapni. Némi módosítással a [10]-ben található gondolatmenetet írjuk most le.

Legyen M egy szeparábilis, teljes metrikus tér! Legyen $C_b(M)$ az M -ből \mathbb{R} -re képző korlátos, folytonos függvények halmaza! Később csupán az $M = \Delta$ esetre fogunk támaszkodni.

Legyen Y^N egy sztochasztikus folyamat M beli értékekkel, s a hozzá tartozó valószínűségi mérték P^N ! Az idő lehet itt folytonos és diszkrét is.

Legyen továbbá M^N az $Y^N(0)$ tartója, vagyis $P^N(Y^N(0) \in M^N) = 1$.

Feltesszük, hogy $E^N(h(Y^N(t) - jY^N(0) = y))$ folytonos y -ban, ahol $y \in M$, $E^N(jY^N(0) = y)$ a P^N szerinti feltételes várhatóérték, $h \in C_b(M)$ és $t \geq 0$.

Π^N M -en értelmezett valószínűségi mértéket Y^N -re invariánsnak mondunk, ha $\Pi^N(M^N) = 1$ és minden $t \geq 0$ -ra és $h \in C_b(M)$ -re teljesül, hogy

$$\int_M E^N(h(Y^N(t) - jY^N(0) = y)) d\Pi^N(y) = \int_M h(y) d\Pi^N(y):$$

Legyen $\phi(t, y)$ determinisztikus függvény, mely második koordinátájában M -en értelmezett, az elsőben pedig \mathbb{N} -en vagy $[0; \infty)$ -n.

Π M -en értelmezett valószínűségi mértéket ϕ -re invariánsnak mondunk, ha minden $t \geq 0$ -ra és $h \in C_b(M)$ -re teljesül, hogy

$$\int_M h(\phi(t, y)) d\Pi(y) = \int_M h(y) d\Pi(y):$$

A következő feltétel a mean fieldhez való konvergencia általánosítása:

Legyenek $h \in C_b(M)$ és $t \geq 0$ tetszőlegesek. Legyen y_0^N konvergens sorozat y_0 határértékkel, melyre minden N esetén $y_0^N \in M^N$. Ekkor teljesül, hogy

$$\lim_{N \rightarrow \infty} E^N(h(Y^N(t) - jY^N(0) = y_0^N)) = h(\phi(t, y_0)):$$

A fenti feltételekkel igaz az alábbi tétel:

11. Tétel. *Legyen Π^N sorozat Y^N -re invariáns! Legyen Π^{N_k} egy gyengén konvergens reszszorozata, melynek határértéke Π ! Ekkor Π ϕ -re invariáns.*

11. Bizonyítás.

A Skorohod reprezentációs tétel alapján létezik egy olyan közös $(\Omega; \mathcal{A}; P)$ valószínűségi mező, melyre tetszőleges mérhető $A \in M$ esetén

$$\begin{aligned} P(X^k \in A) &= \Pi^{N_k}(A) \\ P(X \in A) &= \Pi(A) \\ X^k &\xrightarrow{d} X \quad P \text{ m.m.} \end{aligned}$$

Legyenek $t \geq 0$ és $h \in C_b(M)$ tetszőlegesek és rögzítettek!

$$a^k(y) := E^N h(Y^N(t) | Y^N(0) = y)$$

A feltételeink alapján a^k folytonos és $y^k \in M^{N_k}$ esetén ha $y^k \rightarrow y$, akkor

$$\lim_{k \rightarrow \infty} a^k(y^k) = h(\cdot(t; y))$$

Mivel Π^{N_k} Y^{N_k} -re invariáns, ezért $P(X^k \in M^{N_k}) = 1$. Az is tudjuk, hogy $X^k \rightarrow X$ P m.m., így

$$\lim_{k \rightarrow \infty} a^k(X^k) = h(\cdot(t; X)) \quad P \text{ m.m.}$$

Mivel $\int a^k(X^k) d\Pi^{N_k} = \int h(\cdot(t; y)) d\Pi^{N_k}(y)$, így a Lebesgue-tétel szerint

$$\begin{aligned} \int_M h(y) d\Pi(y) &= \lim_{k \rightarrow \infty} \int_M h(y) d\Pi^{N_k}(y) = \lim_{k \rightarrow \infty} \int_M E^N h(Y^N(t) | Y^N(0) = y) d\Pi^{N_k}(y) = \\ &= \lim_{k \rightarrow \infty} \int_M a^k(y) d\Pi^{N_k}(y) = \lim_{k \rightarrow \infty} \int_M E a^k(X^k) = E \lim_{k \rightarrow \infty} a^k(X^k) = E(h(\cdot(t; X))) = \\ &= \int_M h(\cdot(t; y)) d\Pi(y). \end{aligned}$$

Mivel $t \geq 0$ és $h \in C_b(M)$ tetszőleges volt, így Π \cdot -re invariáns. □

A párhuzamosított minimalista modelljében csupán egyetlen fixpontja volt a rekurciónak, mely globálisan stabil volt, szemben az SIS modellel, ahol a 0 instabil fixpont is szerepel. Az előbbi úgy tudjuk általánosítani, hogy feltesszük, hogy létezik olyan $y \in M$, hogy minden $y \in M$ esetén

$$\lim_{t \rightarrow \infty} \cdot(t; y) = y$$

Ez esetben leegyszerűsödik a \cdot -re invariáns Π -k halmaza, mivel ekkor h korlátossága és folytonossága miatt teljesülnie kell, hogy

$$\int_M h(y) d\Pi(y) = \lim_{t \rightarrow \infty} \int_M h(\cdot(t; y)) d\Pi(y) = \int_M h \lim_{t \rightarrow \infty} \cdot(t; y) d\Pi(y) = \int_M h(y) d\Pi(y)$$

$$\Rightarrow \Pi = \delta_y;$$

ahol δ_y az y -ra koncentrált Dirac-mérték.

A fentiek értelmében ha Π^N Y^N -re invariáns mértékeknek van gyengén konvergens részsorozata, akkor a fenti feltételek mellett csak y -hoz tarthat az adott részsorozat.

Amennyiben ezen kívül azt is feltesszük még, hogy M kompakt - mely $M = \Delta$ esetén teljesül -, akkor ennél is többet tudunk mondani, nevezetesen, hogy Π^N konvergens, s így \cdot aszimptotikus viselkedése miatt $\Pi^N \rightarrow \delta_y$.

Ennek oka a következő: ha nem konvergens Π^N , akkor van olyan $h \in C_b(M)$ $\epsilon > 0$ és N_k részsorozat, hogy minden k -ra

$$\int_M h(y) d\Pi^{N_k}(y) = \int_M h(y) d_{y'}(y) \quad "$$

Mivel M kompakt, ezért Π^{N_k} feszes, így a Prohorov-tétel alapján létezik konvergens $\Pi^{N_{k_s}}$ részsorozat. Erről tudjuk, hogy $\Pi^{N_{k_s}} \rightarrow \Pi^{N_k}$. Ez viszont ellentmond a fenti egyenlőségnek.

Ezek alapján válasszuk X^N -t sűrűségfüggő Markov-láncnak folytonos vagy diszkrét időben. Maga az egész vektor Markov-folyamat, így ha még az ergodikusságot is feltesszük, akkor léteznek az X^N -hez tartozó Π^N stacionárius eloszlások, melyek leírják X^N aszimptotikus viselkedését. Vegyük az Y^N folyamatot, mely megegyezik X^N viselkedésével, csak a kezdeti értéket Π^N eloszlásból vesszük. Ekkor Π^N Y^N -re invariáns.

Legyen γ a folyamathoz tartozó átlejtés közelítés rekurziója vagy differenciálegyenlete. A feltételes várhatóértékek folytonos függését mind a 6., mind a 7. tételben feltettük, s ezek a tételek biztosítják a feltételes várhatóérték γ -hez való konvergenciáját $N \rightarrow \infty$ esetén.

Ha még azt is tudjuk, hogy a rekurciónak vagy differenciálegyenletnek egyetlen globálisan aszimptotikusan stabil egyensúlyi pontja vagy fixpontja van y értékkel, akkor a 11. tétel és az alatta lévő megjegyzések alapján $\Pi^N \rightarrow \delta_y$. Tekintve, hogy a határerloszlás egy pontra koncentrálódik, így a sztochasztikus konvergencia is igaz, vagyis

$$\lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} X^N(t) = y ;$$

ahol az N szerinti limesz sztochasztikusan értendő. Vagyis ebben az esetben felcserélhető a két limesz.

7. Kitekintés

A dolgozatban bemutatásra kerültek az átlagtér közelítés mögötti motivációk, intuíciók, valamint figyelmeztetések, hogy milyen esetben jutnak hamis következtetésre a heurisztikák. Sűrűségfüggő Markov-folyamatok esetén bizonyításra kerültek konvergenciátételek.

Ezek a folyamatok, bár elég általánosak ahhoz, hogy az alkalmazások széles rétegét elérjék, természetesen nem fednek le mindent. Ebben a fejezetben kitérünk pár általánosabb esetre, melyek elemezhetőek hasonló, vagy bonyolultabb eszközökkel.

A sűrűségfüggő Markov-folyamatokra jellemző volt, hogy az komponensek száma állandó volt a folyamat során, s ezzel az állandóval tartottunk a végtelenbe. Gyakran viszont nem ez a helyzet. Például a születés-halálozás folyamatoknál az egyedek létszáma folyamatosan változik.

12. Pelda. Logisztikus növekedés

A populáció dinamikában gyakori feltevés, hogy logisztikusan növekedik egy adott faj létszáma, hogyha a környezet nem változik, s nincs jelen másik faj, mellyel versenyezne vagy ragadozózsákmány viszonyban lenne. Az ezt leíró differenciálegyenlet:

$$\frac{dN(t)}{dt} = N(t)(K - N(t))$$

ahol $N(t)$ a populáció létszáma, K szaporodás vagy halálozás gyorsaságát szabályozó paraméter, K pedig a környezet eltartóképessége.

A radioaktív bomláshoz hasonlóan itt is kifogásolhatjuk, hogy $N(t)$ a valóságban mindig csak egész értékeket vehet fel. Ezt korábban úgy oldottuk meg, hogy áttértünk arányokra, s ha N -nel tartottunk a végtelenbe, akkor az $\frac{1}{N}$ -eket ugró véletlen folyamat is egyre jobban fog hasonlítani egy folytonos függvényhez.

Azonban itt most nem normálhatunk le az egyedek számával, mivel folyton változik, s nincs valami felső korlátja, ami betöltené a radioaktív bomlásnál N szerepét. (Elméletben tetszőlegesen nagyra nőhet a populáció.)

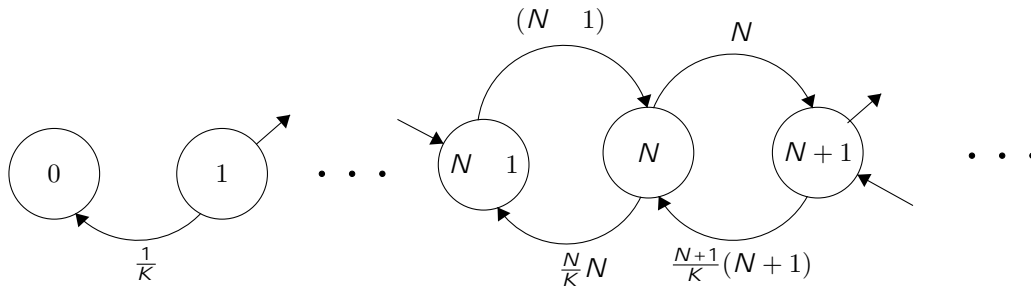
Hogy ezt megoldjuk, K -hoz fogjuk kötni a nagyságrendet. Érdemes tehát bevezetni az $n(t) := \frac{N(t)}{K}$ mennyiséget. Erre átírva a differenciálegyenletet azt kapjuk, hogy

$$\frac{dn(t)}{dt} = K n(t)(1 - n(t)):$$

$K \neq 1$ esetén várunk valamilyen aszimptotikus viselkedést, viszont a differenciálegyenlet jobb oldalán ott marad a K paraméter, s végtelenbe tartás esetén elszáll a jobb oldal. Hogy ezt kiküszöböljük, érdemes a n paramétert K nagyságrendjéhez igazítani, például $n = \frac{N}{K}$ megválasztásával. Ez a fajta paraméter skálázás az idő átskálázásával is elérhető, ha áttértünk $t = Kt$ változóra.

A sztochasztikus modell a következő:

Amennyiben N egyedből áll a populáció, akkor N -rátával kerüljön a populáció az $N + 1$ állapotba és $\frac{N}{K}N$ rátával $N - 1$ állapotba.



A motiváció a fenti modellre az, hogy feltesszük, hogy egy egyed konstans rátával tud szaporodni, viszont minnél többen vannak, annál nagyobb valószínűséggel halnak meg az erőforrások szűkössége miatt.

$$M(n) := \lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} E(n(t + \Delta t) | n(t) = n) = n(1 - \frac{1}{K})$$

Így az átlagtér közelítés valóban a várt differenciálegyenlet megoldása lesz.

$K \neq 1$ esetén a mean fieldhez való konvergenciát a Kurtz-tételben látottakhoz hasonlóan lehet belátni. Az egyetlen különbség az, hogy ügyelnünk kell arra, hogy $n(t)$ -t most nem tudjuk triviálisan felülbecsülni 1-el.

Helyette azt használhatjuk ki, hogy

$$P(n(t) \leq R) \leq \frac{1}{R} E(n(t)) = \frac{n(0)}{R} e^{-t}$$

Így R -et elég nagyra választva elérhetjük, hogy nagy valószínűséggel $n(t) \leq R$ fennáljon, s ezen feltétel mellett már kezelhetjük $n(t)$ -t korlátosként. Fontos, hogy a fenti korlát független K -tól.

Legyenek Y_+ és Y_- egymástól független 1 rátájú Poisson folyamatok! Ekkor

$$N(t) = N(0) + Y_+ - \int_0^t N(s) ds - \frac{1}{K} Y_- - \int_0^t \frac{N(s)}{K} N(s) ds$$

$$n(t) = n(0) + \frac{1}{K} Y_+ - \int_0^t n(s) ds - \frac{1}{K} Y_- - \int_0^t n^2(s) ds :$$

Legyen $\bar{n}(t)$ megoldása a logisztikus differenciálegyenletnek $\bar{n}(0) = n(0)$ kezdeti feltétellel. Ekkor

$$jn(t) - n(t)j = \frac{1}{K} Y_+ - \int_0^t n(s) ds - \int_0^t n(s) ds + \frac{1}{K} Y_- - \int_0^t n^2(s) ds - \int_0^t n^2(s) ds$$

$$+ \int_0^t jM(n(s)) - M(\bar{n}(s))j ds:$$

$O(\frac{1}{R})$ valószínűséggel biztosítani tudjuk, hogy $n(t) \leq R$ fennáljon a $[0; T]$ intervallumon. Feltesszük, hogy $R \neq 1$. Nyilván $\bar{n}(t)$ is korlátos ezen az intervallumon, s így a megfelelő tartományon Lipsitz-folytonos M . Legyen a Lipsitz-konstansa L ! Így tehát a fenti feltétellel mellett a Grönvall-lemma alapján

$$\sup_{0 \leq t \leq T} |n(t) - n(t)| \leq \sup_{0 \leq t \leq RT} \frac{1}{K} Y_+(Kt) + \sup_{0 \leq t \leq R^2 T} \frac{1}{K} Y_+(Kt) e^{-Lt}$$

A jobb oldal teszőlegesen kicsivé tehető teszőlegesen nagy valószínűséggel, amennyiben $K \gg 1$. Tehát

$$\sup_{0 \leq t \leq T} |n(t) - n(t)| \ll 0$$

sztochasztikusan, amennyiben $K \gg 1$.

Általában az olyan paramétert, mely ugyanazt a szerepet tölti be, mint az előző példánkban K , rend paraméternek szokás nevezni. Tipikusan, ha találunk egy ilyen paramétert, akkor a fentihez hasonló érvelést végezve elkerülhető az egyedszám korlátlanóságának problémája. A rend paraméter meghatározása viszont nem mindig nyilvánvaló feladat. (Például Lotka-Volterra modellnél.)

Kurtz 1978-as cikkében [7] foglalkozott ezzel az általános esettel. Ráadásul a cikkében általánosabb eredményeket is adott. A dolgozat folyamán mi csak $O(1)$ nagyságrendű tagokkal foglalkoztunk. Kurtz viszont belátta, hogyha $O(N^{-1})$ korrekciós tagokat is alkalmazunk, akkor egy sztochasztikus differenciálegyenlethez jutunk, melynek a megoldása $O(\frac{\log N}{N})$ -el tér el az eredeti folyamattól, szemben az $O(N^{-1})$ -es hibával, mely a determinisztikus becslésből ered.

További megkötést jelent a Markov-folyamatok feltételezése, mivel kizárja az olyan alkalmazások lehetőségét, ahol a folyamatnak van valamilyen memóriája, de a sűrűségfüggés jellege megmarad. Bizonyos esetekben itt is tudunk konvergenciát biztosítani, viszont ekkor késleltetett differenciálegyenlet megoldásai lesznek a határfolyamatok. Egy ilyen modellre példa: [11].

Egy másik irány, melyre általánosítható az átlagtér közelítés: a reakció-diffúziós folyamatok körére.

13. Pelda. SIS bolyongással

Ebben a modellben az emberek térbeli elhelyezkedését is figyelembe vesszük. Ismét beteg és egészséges csoportokra osztjuk az embereket, s feltesszük, hogy összlétszámuk N . A különbség az, hogy helykoordinátákat is bevezetünk. Az egyszerűség kedvéért a teret az egységnyezeten belül az $(\frac{n}{K}; \frac{m}{K})$ alakú rácspontokat vesszük csak figyelembe, ahol n, m egészek 0 és K között.

Jelölje $s(t; x)$ azok arányát t időpontban, akik x helyen vannak és egészségesek, $i(t; x)$ pedig ugyanígy a betegekkel. Teljesül, hogy

$$\sum_x s(t; x) + i(t; x) = 1$$

A következőt tételezzük fel:

$i(t; x)$ rátával lesz valaki beteg az x helyen, s $s(t; x)$ rátával gyógyul meg mindenhol. Feltesszük, hogy a megbetegedés vagy meggyógyulás során nem változtatja meg senki a helyét.

Ugyanakkor az emberek mozgást is végeznek. Feltesszük hogy az egészségesek λ s, a betegek λ_i rátával mennek át egy szomszédos mezőre egyenletesen választva az irányok közül.

Kurtz tétele alapján, ha N tart a végtelenbe, akkor a tranziens szakaszon az alábbi differenciálegyenlet megoldásához lesz közel a folyamat

$$\frac{ds(t; x)}{dt} = -\lambda s(t; x) + \lambda_i i(t; x) + \sum_{j \neq x} \lambda_{xy} s(t; y) - \lambda s(t; x)$$

$$\frac{di(t; x)}{dt} = i(t; x)s(t; x) - i(t; x) + \sum_{jy}^{\times} i(t; y) - \frac{1}{K} i(t; x)$$

A peremen más egyenleteket kellene felírunk, de ennek vizsgálatától most eltekintünk.

Tetszőleges $v(t; x)$ vetkorra legyen a diszkrét Laplace-operátor:

$$\Delta^K v(t; x) = \frac{1}{K} \sum_{jy}^{\times} (v(t; y) - v(t; x))$$

Ezen jelölés mellett az egyenlet az alábbira módosul:

$$\frac{ds(t; x)}{dt} = i(t; x)s(t; x) - i(t; x) + \frac{s}{K^2} \Delta^K s(t; x)$$

$$\frac{di(t; x)}{dt} = i(t; x)s(t; x) - i(t; x) + \frac{i}{K^2} \Delta^K i(t; x)$$

A következő lépésbenbe szeretnénk sűríteni a négyzetrácsot azzal, hogy K -val tartunk a végtelenbe. E mellett viszont le kell lassítanunk a bolyongásokat, hogy egy időegység alatt $O(1)$ távolságra érjen el a limeszben. Legyenek tehát $D_s := \frac{s}{K^2}$ és $D_i := \frac{i}{K^2}$ állandók.

Ekkor [12] alapján $K \rightarrow \infty$ esetén a differenciálegyenlet az alábbi parciális differenciálegyenlethez konvergál a 4. példában látott skálázás mellett:

$$\partial_t s = i s - i + D_s \Delta s$$

$$\partial_t i = i s - i + D_i \Delta i$$

Fontos megjegyezni, hogy itt először N -el tartottunk a végtelenbe, majd K -val. Nem szól arról a cikk, hogyan kell K -t N -hez mérve skálázni.

Megjegyezzük, hogy a $\partial_t u = f(u) + \Delta u$ alakú parciális differenciálegyenleteket reakció-diffúziós egyenleteknek szokás nevezni.

