

8. Phase-type eloszlások

Kommunikációs hálózatok teljesítményének elemzése

Horváth Illés

2024/10/30

1. Definíció, bevezetés
2. Nevezetes PH eloszlások
3. Relatív szórás
4. Általános közelítés
5. Zártsági tulajdonságok
6. Laplace-transzformált
7. Kitekintés: ME eloszlások

Bevezetés

Emlékeztető. Egy folytonos idejű Markov-lánccban egy adott állapotban eltöltött idő mindig EXP eloszlású.

Láttunk már rá példát, hogy ha egy állapotot lecserélünk egy 2 állapotból álló mini Markov láncra, akkor az abban a 2 állapotban együtt eltöltött idő eloszlása már nem EXP. Ezt akarjuk tovább vizsgálni.

Általában egy Markov-lánccban helyettesíthetünk egy állapotot egy több állapotból álló mini Markov-lánccal. Ilyenkor a mini Markov-lánccnak mindig van egy nyelő állapota.

Bevezetés

Például ha a mini Markov-lánc generátora

$$\begin{bmatrix} -1 & 1/3 & 2/3 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

ez annak felel meg, hogy amikor a nagy Markov-lánc belép a mini Markov-láncba, akkor az első és második állapot között ugrál egy ideig, majd amikor a harmadik (nyelő) állapotba lép, akkor valószínűleg tovább lép a nagy Markov-lánc egy másik állapotába.

Emiatt a harmadik állapotot nem is szokás feltüntetni:

$$A = \begin{bmatrix} -1 & 1/3 \\ 1 & -2 \end{bmatrix}$$

Bevezetés

Definíció. Az A négyzetes mátrix szubgenerátor mátrix, ha

- ▶ A főátlóbeli elemei negatívak, főátlón kívüli elemei nemnegatívak,
- ▶ A minden egyes sorösszege legfeljebb 0, és legalább egy sorösszege szigorúan kisebb, mint 0.

Ilyenkor A tekinthető úgy, mint egy nyelővel rendelkező Markov-lánc generátora. A nyelőbe lépés rátáját az egyes állapotokból a $-A1$ oszlopvektor adja meg, ahol 1 a csupa 1-esből álló vektor.

A szubgenerátor lényegében a szubsztokasztikus mátrix megfelelője folytonos idejű Markov-folyamatokra.

Phase-type eloszlások

Definíció. Legyen α n hosszú eloszlás vektor és A $n \times n$ -es szubgenerátor mátrix. Ekkor $\text{PH}(\alpha, A)$ az az eloszlás, ami egy α kezdeti vektorú, A generátorú, nyelővel rendelkező Markov-lánc elnyelődésének ideje.

A $\text{PH}(\alpha, A)$ eloszlás sűrűségfüggvénye

$$f(t) = \alpha e^{At}(-A\mathbf{1}),$$

eloszlásfüggvénye

$$F(t) = 1 - \alpha e^{At}\mathbf{1}.$$

Az A -nak megfelelő állapotokat (a mini Markov-lánc állapotait) fázisnak hívjuk. A PH-eloszlás rendje a fázisok száma, vagyis n .

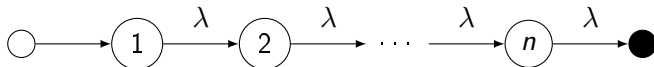
Exponenciális és Erlang eloszlás

Példa. Ha A egy 1×1 -es mátrix, akkor $\text{PH}(\alpha, A)$ az exponenciális eloszlást adja vissza.

Példa. Ha A $n \times n$ -es,

$$A = \begin{bmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -\lambda & \lambda & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & \dots & & 0 & -\lambda \end{bmatrix},$$

azaz a fázisok sorba vannak kötve, és $\alpha = (1 \ 0 \ \dots \ 0)$, akkor $\text{PH}(\alpha, A)$ éppen n darab független $\text{EXP}(\lambda)$ eloszlás összege, ami az $\text{Erlang}(\lambda, n)$ eloszlás.

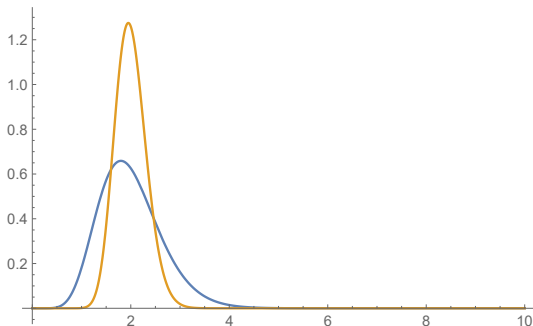


Az $\text{Erlang}(n, \lambda)$ eloszlás sűrűségfüggvénye

$$f(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}.$$

Erlang eloszlás

Az Erlang eloszlás koncentrált az n/λ pont környékére; minél nagyobb n , annál inkább koncentrált. Az alábbi ábrán a kék görbe az Erlang(10, 5) eloszlás sűrűségfüggvénye, a sárga pedig az Erlang(40, 20) eloszlás sűrűségfüggvénye:



Erlang eloszlás

Egy X valószínűségi változó relatív szórása (coefficient of variation)

$$cv(X) = \frac{D(X)}{E(X)}, \quad cv^2(X) = \frac{D^2(X)}{(E(X))^2} = \frac{\int_0^\infty t^2 f(t) dt}{\left(\int_0^\infty t f(t) dt\right)^2} - 1.$$

Például ha $X \sim \text{EXP}(\lambda)$, akkor $cv(X) = 1$.

Az adott rendű PH-eloszlások közül az Erlang-eloszlás a legkoncentráltabb:

Tétel (Aldous-Shepp)

$X \sim \text{Erlang}(n, \lambda)$ -ra

$$cv^2(X) = \frac{1}{n},$$

és bármilyen X n -rendű PH-eloszlásra

$$cv^2(X) \geq \frac{1}{n}.$$

Hiperexponenciális eloszlás

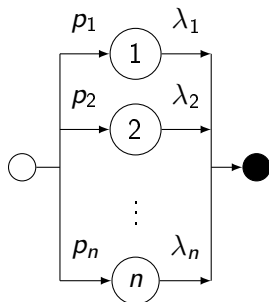
Példa. A hiperexponenciális eloszlásnál a fázisok párhuzamosan vannak kötve:

$$A = \begin{bmatrix} -\lambda_1 & 0 & 0 & \dots & 0 \\ 0 & -\lambda_2 & 0 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & \dots & 0 & & -\lambda_n \end{bmatrix}$$

és $\alpha = (p_1 \ p_2 \ \dots \ p_n)$.

Ilyenkor

$$f(t) = \sum_{i=1}^n p_i \lambda_i e^{-\lambda_i t}, \quad F(t) = 1 - \sum_{i=1}^n p_i e^{-\lambda_i t}.$$



Hiperexponenciális eloszlás

Példa. Időnként lefagyó szerverek kiszolgálási idejét lehet modellezni 2 állapotú hiperexponenciális eloszlással. Az egyik ág a normál kiszolgálás exponenciális idővel, a másik ág a lefagyás: a szerver olyan állapotba kerül, ahonnan nem (vagy csak nagyon sokára) tud kijönni. Ilyenkor a kiszolgálási ráta extrém kicsi.

Lemma

Ha X hiperexponenciális eloszlású, akkor

$$cv(X) > 1.$$

Nem biz., de a p_i, λ_i értékekből könnyen kiszámítható.

Általános eloszlás közelítése

Láttuk, hogy PH eloszlással lehet jól közelíteni a determinisztikus eloszlást (Erlang).

De vannak olyan PH eloszlások is, amik nagy szórásúak (hiperexponenciális).

Vajon lehet-e tetszőleges folytonos eloszlást közelíteni PH eloszlással?

Általános eloszlás közelítése

Láttuk, hogy PH eloszlással lehet jól közelíteni a determinisztikus eloszlást (Erlang).

De vannak olyan PH eloszlások is, amik nagy szórásúak (hiperexponenciális).

Vajon lehet-e tetszőleges folytonos eloszlást közelíteni PH eloszlással? Igen!

Tétel

Tetszőleges $g(t)$, $t \geq 0$ sűrűségfüggvényhez és $\varepsilon > 0$ számhoz létezik olyan (α, A) pár, hogy

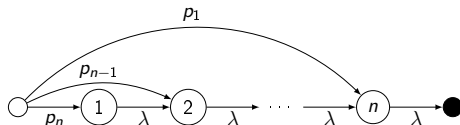
$$\int_0^{\infty} |f(t) - g(t)| < \varepsilon,$$

ahol

$$f(t) = \alpha e^{At}(-A1).$$

Általános eloszlás közelítése

Biz. (vázlat.) Koncentrált Erlang-sűrűségfüggvényekből szeretnénk kikeverni a g függvényt. Ehhez a következő struktúrát használjuk (ún. hipo-Erlang eloszlás):



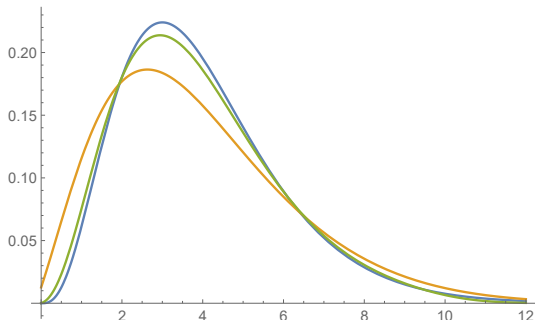
a következő paraméter-választásokkal:

$$\lambda = n/T,$$
$$p_i = \frac{g(Ti/n)}{\sum_{j=1}^n g(Tj/n)}, \quad i = 1, \dots, n;$$

a p_i valószínűségű állapotból az elnyelődés ideje $\text{Erlang}(i, \lambda)$, ami az $i/\lambda = Ti/n$ pontra koncentrálódik. Ha n és T elég nagy, akkor jó lesz a közelítés.

Erlang eloszlás

Az alábbi ábrán a kék sűrűségfüggvényt közelítjük; a sárga függvény az előbbi közelítés $n = 20$ és $T = 10$ választással, a zöld függvény pedig $n = 100$ és $T = 10$ választással.



Az ábrán látszik, hogy ahhoz, hogy jó közelítést kapjunk, n -et elég nagyra kell venni. Általában olyan PH eloszlást találni, ami alacsony rendű és jó közelítés, nehéz feladat. Kulcsszó: phase-type fitting.

Zártsági tulajdonságok

Tétel

Ha X és Y független PH eloszlású valószínűségi változók, akkor

- (a) cX is PH eloszlású, ha $c > 0$ konstans;*
- (b) $X + Y$ is PH eloszlású;*
- (c) $\min(X, Y)$ is PH eloszlású;*
- (d) $\max(X, Y)$ is PH eloszlású;*
- (e) X és Y keveréke is PH eloszlású:*

$$Z = \begin{cases} X & p \text{ valószínűséggel} \\ Y & 1 - p \text{ valószínűséggel} \end{cases}$$

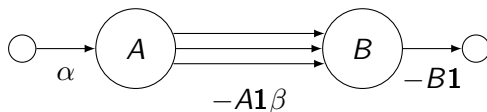
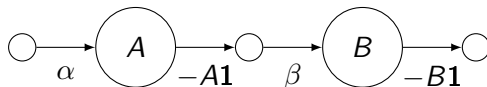
Zártsági tulajdonságok

- (a) cX PH eloszlású, ha $c > 0$ konstans.

Megoldás. Ha $X \sim \text{PH}(\alpha, A)$, akkor $cX \sim \text{PH}(\alpha, A/c)$. (Az időt c -edrészére lassítjuk.)

- (b) $X + Y$ is PH eloszlású.

Legyen $X \sim \text{PH}(\alpha, A)$, $Y \sim \text{PH}(\beta, B)$; sorba kapcsoljuk egymás után a két Markov-láncot.



Zártsági tulajdonságok

(b) Mátrix alakban összefoglalva $X + Y \sim \text{PH}(\gamma, C)$, ahol

$$C = \left[\begin{array}{c|c} A & -A\mathbf{1}\beta \\ \hline 0 & B \end{array} \right],$$
$$\gamma = (\alpha \ 0 \dots 0).$$

($-A\mathbf{1}\beta$ diádszorzat, azaz a $-A\mathbf{1}$ és a β vektorokat úgy szorozzuk össze, hogy 1 rangú mátrixot kapjunk.)

(c) $\min(X, Y)$ is PH eloszlású.

Megoldás. Ehhez próbáljuk meg párhuzamosan, függetlenül futtatni a két Markov-láncot, azzal a feltétellel, hogy ha bármelyik elnyelődik, akkor vége a dalnak.

Kitérő: Kronecker összeg és szorzat

Legyen $X(t)$ folytonos idejű Markov-lánc az $\{1, 2\}$ állapottéren,
 $Y(t)$ pedig az $\{a, b, c\}$ állapottéren; generátoraik

$$Q_X = \begin{bmatrix} -1/2 & 1/2 \\ 1 & -1 \end{bmatrix} \qquad Q_Y = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 3 & 0 & -3 \end{bmatrix}$$

$X(t)$ és $Y(t)$ párhuzamosan, egymástól függetlenül zajlanak.
Legyen $Z(t) = (X(t), Y(t))$. Gondoljuk meg, hogy $Z(t)$ is
folytonos idejű Markov-lánc. Mik az állapotai és hogy kaphatnánk
meg a generátorát?

Kitérő: Kronecker összeg és szorzat

A lehetséges állapotok $1a, 1b, 1c, 2a, 2b, 2c$. Az $1a$ állapotból a következő átmenetek lehetségesek:

- ▶ $1a \rightarrow 1b$. Ez akkor következik be, ha $Y(t)$ -ben történik egy $a \rightarrow b$ átmenet; ennek rátája 1.
- ▶ $1a \rightarrow 1c$ is 1 rátával.
- ▶ $1a \rightarrow 2a$ pedig akkor, ha $X(t)$ -ben történik egy $1 \rightarrow 2$ átmenet, ennek rátája $1/2$.

Összességében Z generátora, Q_Z a következő:

	$1a$	$1b$	$1c$	$2a$	$2b$	$2c$
$1a$	$-5/2$	1	1	$1/2$	0	0
$1b$	1	$-3/2$	0	0	$1/2$	0
$1c$	3	0	$-7/2$	0	0	$1/2$
$2a$	1	0	0	-3	1	1
$2b$	0	1	0	1	-2	0
$2c$	0	0	1	3	0	-4

Kitérő: Kronecker összeg és szorzat

Q_Z szerkezetét a következőképpen lehet megérteni: az $Y(t)$ -beli átmenetek pirosak, az $X(t)$ -beli átmenetek kék:

	1a	1b	1c	2a	2b	2c
1a	*	1	1	$\frac{1}{2}$	0	0
1b	1	*	0	0	$\frac{1}{2}$	0
1c	3	0	*	0	0	$\frac{1}{2}$
2a	1	0	0	*	1	1
2b	0	1	0	1	*	0
2c	0	0	1	3	0	*

A piros elemek tulajdonképpen Y generátora két példányban lemásolva. Ezt az ún. Kronecker-jelöléssel lehet felírni.

Kitérő: Kronecker összeg és szorzat

Egy A $n \times n$ -es és egy B $m \times m$ -es mátrix Kronecker-szorzata $mn \times mn$ -es mátrix a következő struktúrával:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{bmatrix}.$$

Ezzel a jelöléssel

$$I_2 \otimes Q_Y = \begin{bmatrix} * & 1 & 1 & 0 & 0 & 0 \\ 1 & * & 0 & 0 & 0 & 0 \\ 3 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 1 & 1 \\ 0 & 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 3 & 0 & * \end{bmatrix}.$$

Kitérő: Kronecker összeg és szorzat

Hasonlóan az $X(t)$ -beli átmenetekre

$$Q_X \otimes I_3 = \begin{bmatrix} * & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & * & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & * & 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 & * & 1 & 1 \\ 0 & 1 & 0 & 1 & * & 0 \\ 0 & 0 & 1 & 3 & 0 & * \end{bmatrix};$$

végül, bevezetve az

$$Q_Z \oplus Q_Y = I_2 \otimes Q_Y + Q_X \otimes I_3$$

Kronecker-összeget, azt kaptuk, hogy

$$Q_Z = Q_X \oplus Q_Y.$$

Zártsági tulajdonságok

Ha generátorok helyett szubgenerátorok vannak, akkor is pont ugyanez a helyzet:

$$C = A \oplus B;$$

ha A is és B is szubgenerátor, akkor mindkettőből a „hiányzó” ráta az elnyelődéshez járul hozzá $\rightarrow$ az elnyelődés ideje $\min(X, Y)$.

A kezdeti eloszlások függetlensége miatt a kezdeti vektor is Kronecker-szorzat alakú:

$$\gamma = \alpha \otimes \beta.$$

Zártsági tulajdonságok

(d) $\max(X, Y)$ is PH eloszlású.

Megoldás. Ezúttal ha az egyik Markov-lánc elnyelődik, attól még a másik tovább zajlik, és meg kell várni, amíg az is elnyelődik. Emiatt szükség van olyan állapotokra is, amikor már csak vagy az egyik, vagy a másik zajlik.

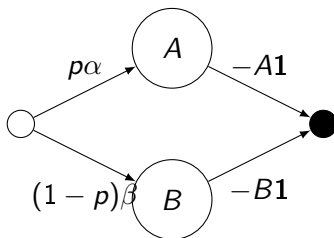
$$C = \left[\begin{array}{c|c|c} A \oplus B & -A\mathbf{1} \oplus I & I \oplus -B\mathbf{1} \\ \hline 0 & B & 0 \\ \hline 0 & 0 & A \end{array} \right],$$
$$\gamma = (\alpha \otimes \beta \ 0 \dots 0).$$

Zártsági tulajdonságok

(e) X és Y keveréke is PH eloszlású:

$$Z = \begin{cases} X & p \text{ valószínűséggel} \\ Y & 1 - p \text{ valószínűséggel} \end{cases}$$

Megoldás.



PH eloszlás sűrűségfüggvényének alakja

Az $\text{Erlang}(n, \lambda)$ eloszlás sűrűségfüggvénye

$$f(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}.$$

$f(t) = \alpha e^{At}(-A1)$ általános alakja

$$f(t) = \sum_i c_i p_i(t) \sin(\omega_i(t + \phi_i)) e^{-\lambda_i t},$$

ahol

- ▶ $p_i(t)$ polinom, és akkor jelenik meg, ha A -nak többszörös sajátértékei vannak;
- ▶ $\sin$ -os tag akkor van, ha a mini-Markov láncban van kör (ilyenkor A -nak van komplex sajátértéke).

Lásd még: mátrixok Jordan-felbontása, Jordan-blokk függvénye.

Momentumok, Laplace-transzformált

Az $X \sim \text{PH}(\alpha, A)$ eloszlás momentumai:

$$E(X^k) = k! \alpha (-A)^{-k} \mathbf{1}$$

A Laplace-transzformáltja

$$X^*(s) = \alpha (sI - A)^{-1} (-A\mathbf{1}),$$

ami mindig racionális törtfüggvény alakú (polinom per polinom alakú).

Vajon igaz-e visszafelé, vagyis ha $X^*(s)$ racionális törtfüggvény alakú, abból következik-e, hogy X phase-type eloszlású?

Majdnem, de nem egészen.

PH eloszlás sűrűségfüggvényének alakja

Tétel

Ha egy X nemnegatív folytonos valószínűségi változóra $X^(s)$ racionális törtfüggvény alakú, akkor X sűrűségfüggvénye*

$$f(t) = \alpha e^{At}(-A\mathbf{1})$$

alakú valamely A négyzet alakú mátrixra és α vektorra.

Nem biz.

A -bana főátlón kívül is lehetnek negatív elemek, illetve α -ban is lehetnek negatív elemek.

Kitekintés: ME eloszlások

És ez nagy baj amúgy? Végülis nem.

Definíció. Legyen α n hosszú vektor, melynek elemeinek összege 1, és A $n \times n$ -es mátrix. Ekkor, ha az

$$f(t) = \alpha e^{At}(-A\mathbf{1})$$

függvény nemnegatív minden $t \geq 0$ -ra, akkor ezt $\text{ME}(\alpha, A)$ eloszlásnak hívjuk (ME mint mátrix exponenciális).

Annyi a különbség a PH eloszlásokhoz képest, hogy itt megengedünk A és α elemei között negatív számokat is. Viszont emiatt nincs garancia arra, hogy $f(t)$ nemnegatív, emiatt külön ki kellett kötni. A mini Markov-láncos sztochasztikus interpretáció is elvész.

Viszont az ME eloszlásokra az összes analitikus formula pont ugyanúgy teljesül, mint PH eloszlásokra, és ME eloszlásokkal esetleg kisebb rendben is lehet jól közelíteni.