Tampereen yliopisto

Miranda Grönlund

MARKOVIN KETJUT

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Joulukuu 2020

Tiivistelmä

Miranda Grönlund: Markovin ketjut
Kandidaattitutkielma
Tampereen yliopisto
Matematiikan ja tilastotieteen tutkinto-ohjelma
Joulukuu 2020

Markovin ketjut (engl. Markov Chains) ovat 1900-luvun alkupuolella kehitettyjä stokastisia prosesseja, joissa uusi tila riippuu vain sitä edeltävästä tilasta. Ketjut ovat nimetty niiden kehittäjän, Andrey Markovin, mukaan, ja niitä sovelletaan nykyisin monilla eri tieteenaloilla esimerkiksi lääketieteessä ja kauppatieteissä. Suurin syy, miksi Markovin ketjuja käytetään eri tieteenaloilla, on niiden todennäköisyyslaskennallinen tausta. Markovin ketjujen avulla pystytään ennustamaan sekä jatkuvien että diskreettien stokastisten prosessien satunnaiskulkua, ja näin ollen tuottamaan arvioita ketjujen tulevista tiloista eri ajanhetkillä. Markovin ketjuista on olemassa melko vähän suomenkielisiä tieteellisiä tekstejä, ja siksi tämän tutkielman tarkoituksena onkin tuottaa lyhyt katsaus diskreettiaikaisten Markovin ketjujen teoriaan, ja selittää lukijalleen esimerkkien avulla, miten ketjuja voidaan hyödyntää käytännössä.

Työ on jaettu kahteen osaan. Ensimmäisessä osassa käsitellään Markovin ketjujen todennäköisyyslaskennallista taustaa ja esitellään kaikki määritelmät ja lauseet, joita tarvitaan Markovin ketjujen ymmärtämiseen. Tässä osassa tutkielmaa selitetään myös, mitä stokastiset prosessit ovat ja esitellään niihin liittyvät määritelmät, jotka ovat oleellisia Markovin ketjujen ymmärtämiseen. Tutkielman toinen osa käsittelee diskreettiaikaisia Markovin ketjuja. Siinä esitellään Markovin ketjujen määritelmiä, kuten Markovin ominaisuus sekä eräitä Markovin ketjujen ominaisuuksia, joita myöhemmin havainnollistetaan erilaisin esimerkein. Tutkielman toisessa osassa esitetään myös Markovin ketjujen siirtymämatriisiesitys sekä graafinen esitys, joiden tarkoitus on havainnollistaa Markovin ketjuja visuaalisessa muodossa.

Tutkielma päättyy Markovin ketjujen sovelluksiin, joissa kootaan yhteen kaikki tutkielmassa esille nousseet lauseet ja ominaisuudet, ja näytetään, kuinka ne käytännössä vaikuttavat ketjuihin. Tämä onkin yksi tutkielman tärkeimmistä osista, sillä esimerkkien avulla lukijan on helppo ymmärtää, kuinka Markovin ketjuja voidaan

käytännössä hyödyntää.

Avainsanat: Markovin ketjut, stokastiset prosessit, todennäköisyyslaskenta

Sisällys

1 Johdanto

Markovin ketjut ovat venäläisen matemaatikon, Andrey Markovin, kehittämä menetelmä, jolla mallinnetaan stokastisia prosesseja, joissa uusi tila riippuu vain sitä edeltävästä tilasta. Toisin sanoen, Markovin ketjuilla voidaan mallintaa sellaisia ajassa sattumanvaraisesti eteneviä ilmiöitä, joissa satunnaisilmiön tulos on riippuvainen vain ja ainoastaan sitä edeltävästä tuloksesta. Niitä hyödynnetään nykyään lukuisilla tieteenaloilla, esimerkiksi biologiassa, fysiikassa ja taloustieteessä.

Tämän tutkielma keskittyy käsittelemään diskreettiaikaisia Markovin ketjuja. Aloitamme tutkielman esittämällä Markovin ketjujen todennäköisyyslaskennallisen taustan ja selittämällä, mitä stokastiset prosessit ovat. Todennäköisyyslaskennasta esitämme Markovin ketjujen ymmärtämiseen ja soveltamiseen tarvittavia käsitteitä, kuten todennäköisyysavaruus, tapahtumien ja satunnaismuuttujien riippumattomuus sekä odotusarvot. Pykälässä 2.2 esitämme stokastisten prosessien ja satunnaiskulun määritelmät sekä eräitä esimerkkejä havainnollistamaan kyseisiä määritelmiä.

Luvussa 3 esitämme Markovin ominaisuuden määritelmän, määrittelemme Markovin ketjujen siirtymämatriisiesityksen ja esitämme Markovin ketjujen ominaisuuksia sekä niiden sovelluksia.

Lukijalta edellytämme joidenkin todennäköisyyslaskennan ja lineaarialgebran perusasioiden tuntemista. Edellytämme muun muassa, että lukija tuntee matriisilaskennan perusteet ja pystyy seuraamaan todennäköisyyslaskennan todistuksia. Ensisijaisena lähdeteoksena käytämme Nicolas Privaultin kirjaa Understanding Markov Chains [2].

2 Valmistelevia tarkasteluja

2.1 Todennäköisyyslaskennallinen tausta

Luvussa 2 esitämme lyhyesti muutamia pääaiheemme käsittelyssä tarvitsemiamme apuneuvoja. Tässä pykälässä esitämme Markovin ketjujen todennäköisyyslaskennallisen taustan (vrt. [1, s. 9–14, 18], [2, s. 7–30]).

Määritelmä 2.1. Joukkoa Ω, joka sisältää kaikki satunnaisilmiön mahdolliset tulokset, kutsutaan *perusjoukoksi*.

Määritelmä 2.2 (σ-algebra)**.** Olkoon Ω epätyhjä joukko. Osajoukkojen $A \subseteq \Omega$ joukkoperhettä $\mathcal{F}$ kutsutaan joukon Ω *σ-algebraksi*, jos seuraavat aksioomat ovat voimassa:

(1) $\emptyset, \Omega \in \mathcal{F}$,

(2) jos $A \in \mathcal{F}$, niin $A^C := \Omega \setminus A \in \mathcal{F}$,

(3) jos $A_1, A_2, \ldots \in \mathcal{F}$, niin $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Alkioita $A \in \mathcal{F}$ kutsutaan *tapahtumiksi*.

Esimerkki 2.1 (Nopanheitto)**.** Olkoon perusjoukko $\Omega = \{1, 2, 3, 4, 5, 6\}$, tällöin joukkoperhe

$$\mathcal{F} := \big\{\Omega, \emptyset, \{1, 2, 3\}, \{4, 5, 6\}\big\}$$

määrittelee perusjoukon Ω σ-algebran, joka vastaa satunnaisesti valittujen silmälukujen jakoa ”suuriin” ja ”pieniin” silmälukuihin.

Määritelmä 2.3. Kuvausta $\mathbb{P}\colon \mathcal{F} \to [0, 1]$ kutsutaan *todennäköisyysmitaksi*, jos

(1) $\mathbb{P}(\Omega) = 1$ ja

(2) $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$, kun $A_i \cap A_j \neq \emptyset$ ja $i \neq j$.

Kolmikkoa $(\Omega, \mathcal{F}, \mathbb{P})$ kutsutaan *todennäköisyysavaruudeksi*.

Määritelmä 2.4 (Ehdollinen todennäköisyys)**.** Olkoon $(\Omega, \mathcal{F}, \mathbb{P})$ todennäköisyysavaruus, jossa tapahtuma $A \in \mathcal{F}$ ja $\mathbb{P}(A) > 0$. Tällöin todennäköisyyttä

$$\mathbb{P}(B|A) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}, \qquad \text{kun } B \in \mathcal{F},$$

kutsutaan tapahtuman B *ehdolliseksi todennäköisyydeksi* ehdolla A.

Määritelmä 2.5. Tapahtuminen A ja B sanotaan olevan *riippumattomia*, jos

$$\mathbb{P}(A|B) = \mathbb{P}(A).$$

Määritelmä 2.6. Joukkoa

$$\{\mathbb{P}(X \in A)\colon \text{A } \textit{joukon } \mathbb{R} \text{ mitallinen osajoukko}\}.$$

kutsutaan satunnaismuuttujan $X\colon \Omega \to \mathbb{R}$ *todennäköisyysjakaumaksi*.

Määritelmä 2.7. Todennäköisyysmitalla $\mathbb{P}$ määritellyt satunnaismuuttujat X ja Y ovat *riippumattomia*, jos

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B),$$

pätee kaikilla $A, B \subseteq \mathbb{R}$.

Määritelmä 2.8 (Diskreetin satunnaismuuttujan odotusarvo)**.** Olkoon $X\colon \Omega \to \mathbb{N}$ diskreetti satunnaismuuttuja. Tällöin satunnaismuutujan X odotusarvo $\mathrm{E}(X)$ on

$$\mathrm{E}(X) = \sum_{k=0}^{\infty} k\mathbb{P}(X = k),$$

missä $k \in \mathbb{N}$.

Esimerkki 2.2 (Nopanheitto)**.** Kuvataan nopanheitossa saatua silmälukua satunnaismuuttujalla X. Perusjoukko $\Omega = \{1, 2, 3, 4, 5, 6\}$ ja jokaisen silmäluvun todennäköisyys $\mathbb{P}(X) = \frac{1}{6}$. Tällöin

$$\mathrm{E}(X) = \sum_{k=1}^{6} k\mathbb{P}(X = k) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3\frac{1}{2}.$$

Määritelmä 2.9 (Satunnaismuuttujan ehdollinen odotusarvo)**.** Satunnaismuuttujan $X\colon \Omega \to \mathbb{N}$ *ehdollinen odotusarvo* $\mathrm{E}(X|A)$ ehdolla A on

$$\mathrm{E}(X|A) = \sum_{k=0}^{\infty} k\mathbb{P}(X = k|A),$$

aina, kun sarja on suppeneva.

2.2 Stokastiset prosessit

Tässä pykälässä selitämme, mitä stokastiset prosessit ovat sekä esitämme niiden ymmärtämiseen tarvittavia käsitteitä (vrt. [2, s. 1, 61–64]).

Määritelmä 2.10. *Stokastinen prosessi* on matemaattinen työväline, jolla mallinnetaan ajassa sattumanvaraisesti eteneviä ilmiöitä.

Esimerkki 2.3 (Autojen käyttäytyminen risteysalueella)**.** Ajanhetkellä $t = 0$ autoja on kulkenut risteyksestä 0 kappaletta. Kun aikaa on kulunut x tuntia, autoja on kulkenut risteyksestä y kappaletta. Näistä autoista $\frac{1}{3}y$ kääntyi oikealle, $\frac{1}{4}y$ vasemmalle ja $\frac{5}{12}y$ jatkoi suoraan. Tätä autojen käyttäytymisen mallintamista voidaan pitää stokastisena prosessina.

Määritelmä 2.11. Rajoittamatonta *satunnaiskulkua* merkitään $(\mathrm{S}_n)_{n\geq 0}$ ja määritellään sen ensimmäisen tilan $\mathrm{S}_0 = 0$ avulla siten, että

$$\mathrm{S}_n = X_1 + X_2 + \cdots + X_n, \qquad n \geq 1,$$

missä $(X_k)_{k\geq 1}$ on ryhmä toisistaan riippumattomia $\{-1, 0, +1\}$-arvoisia satunnaismuuttujia.

Lisäksi oletamme, että ryhmä $(X_k)_{k\geq 1}$ koostuu toisistaan riippumattomista ja identtisesti jakautuneista satunnaismuuttujista, jotka noudattavat todennäköisyysjakaumaa

$$\begin{cases} \mathbb{P}(X_k = +1) = p, \\ \mathbb{P}(X_k = 0) = r, \\ \mathbb{P}(X_k = -1) = q, \end{cases}$$

$k \geq 1$, kun $p + q + r = 1$.

Määritelmä 2.12. Olkoon $r = 1 - p - q = 0$. Tällöin satunnaiskulkua $(\mathrm{S}_n)_{n\in\mathbb{N}}$ kutsutaan *Bernoullin satunnaiskuluksi*.

Tällöin satunnaismuuttujan X_n odotusarvo on

$$\mathrm{E}[X_n] = -1 \cdot q + 1 \cdot p = 2p - 1 = p - q$$

ja varianssi

$$\begin{aligned} \mathrm{Var}[X_n] &= \mathrm{E}[X_n^2] - (\mathrm{E}[X_n])^2 \\ &= 1 \cdot q + 1 \cdot p - (2p - 1)^2 \\ &= 4p(1 - p) = 4pq. \end{aligned}$$

3 Diskreettiaikaiset Markovin ketjut

Kuten luvussa 1 kerrotaan, tämä tutkielma keskittyy vain diskreettiaikaisiin Markovin ketjuihin. Tässä luvussa esitämme Markovin ketjujen tärkeimmät määritelmät, ominaisuudet ja niiden sovelluksia. Sivuutamme kuitenkin kaikki jatkuva-aikaisiin Markovin ketjuihin liittyvät taustatiedot, määritelmät ja ominaisuudet kokonaisuudessaan.

3.1 Diskreettiaikaisten Markovin ketjujen perusteita

Tässä pykälässä esitämme Markovin ketjujen ymmärtämiseen tarvittavia määritelmiä (vrt. [2, s. 77–166]).

Määritelmä 3.1 (Markovin ominaisuus)**.** Diskreettiaikaisella stokastisella prosessilla $(Z_n)_{n\in\mathbb{N}}$ tila-avaruudessa S sanotaan olevan *Markovin ominaisuus*, jos kaikilla $n \geq 1$ ja kaikilla $i_0, i_1, \ldots, i_n, j \in$ S pätee

$$\mathbb{P}(Z_{n+1} = j | Z_n = i_n, Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0) = \mathbb{P}(Z_{n+1} = j | Z_n = i_n).$$

Määritelmä 3.2 (Siirtymämatriisi)**.** Markovin ketjun $(Z_n)_{n\in\mathbb{N}}$ satunnaiskehitys määritetään datan

$$P_{i,j} := \mathbb{P}(Z_{n+1} = j | Z_n = i), \qquad i, j \in \mathrm{S},$$

avulla. Tämä vastaa siirtymätodennäköisyyksiä $\mathbb{P}(Z_{n+1} = j | Z_n = i)$, jotka ovat riippumattomia muuttujasta $n \in \mathbb{N}$.

Data $P_{i,j}$ voidaan muuttaa $|\,\mathrm{S}\,|^2 = |\,\mathrm{S} \times \mathrm{S}\,|$ kokoiseksi matriisiksi, missä $|S|$ on tila-avaruuden alkioiden lukumäärä, ja näin saadaan Markovin ketjun *siirtymämatriisi*:

$$[P_{i,j}]_{i,j\in\mathrm{S}} = [\mathbb{P}(Z_1 = j | Z_0 = i)]_{i,j\in\mathrm{S}'}$$

toisin sanoen

$$[P_{i,j}]_{i,j\in\mathrm{S}} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & ⋰ \\ \ldots & P_{-2,-2} & P_{-2,-1} & P_{-2,0} & P_{-2,1} & P_{-2,2} & \ldots \\ \ldots & P_{-1,-2} & P_{-1,-1} & P_{-1,0} & P_{-1,1} & P_{-1,2} & \ldots \\ \ldots & P_{0,-2} & P_{0,-1} & P_{0,0} & P_{0,1} & P_{0,2} & \ldots \\ \ldots & P_{1,-2} & P_{1,-1} & P_{1,0} & P_{1,1} & P_{1,2} & \ldots \\ \ldots & P_{2,-2} & P_{2,-1} & P_{2,0} & P_{2,1} & P_{2,2} & \ldots \\ ⋰ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Relaatiosta

$$\sum_{j \in S} \mathbb{P}(Z_1 = j | Z_0 = i) = 1, \qquad i \in \mathbb{N}$$

seuraa, että siirtymämatriisin rivit toteuttavat ehdon

$$\sum_{j \in S} P_{i,j} = 1,$$

jokaisella rivin indeksillä $i \in S$.

Tällaisen Markovin ketjun $(Z_n)_{n \in \mathbb{N}}$ sanotaan olevan *aikahomogeeninen* eli ajasta riippumaton.

Määritelmä 3.3. Tilaa $k \in S$ kutsutaan absorvoivaksi, jos $P_{k,k} = 1$.

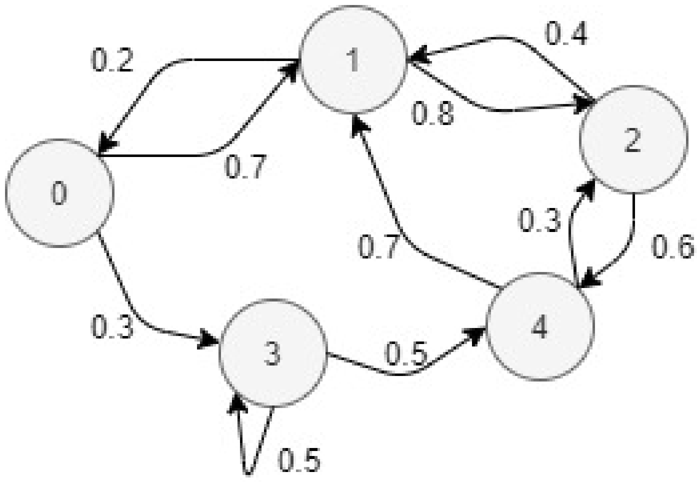
Esimerkki 3.1. Mallinnetaan henkilön luottoluokituksia satunnaisprosessina tilajoukossa $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, missä tila 1 = AAA, 2 = AA, 3 = A, 4 = BBB, 5 = BB, 6 = B, 7 = CCC, 8 = D ja 9 = Ei luokitusta (E.L.). Luottoluokituksen tila voidaan esittää Markovin ketjuna $(Z_0, Z_1, \dots)$, jonka siirtymämatriisi on

Luokitus	Luokitus vuoden lopussa									
vuoden alussa	AAA	AA	A	BBB	BB	B	CCC	D	E.L.	Yht.
AAA	91.32	4.64	0.42	0.07	0.02	0	0	0	3.4	100
AA	0.6	88.48	7	0.49	0.07	0.11	0.03	0	3.22	100
A	0.09	2.17	82.91	6.07	1.85	1.54	0.95	0.06	4.36	100
BBB	0.04	0.36	4.43	81.25	6.22	1.98	0.31	0.24	5.17	100
BB	0.01	0.07	0.63	7.02	77.93	5.67	1.34	0.96	6.37	100
B	0	0.08	0.32	0.45	6.7	72.71	4.06	5.0	10.68	100
CCC	0.02	0	0.33	0.82	1.72	9.05	54.23	20.26	13.57	100
D	0	0	0	0	1.2	4.7	7.43	56.87	29.8	100
E.L.	0	0	0	0	0	0.2	7.94	76.19	15.67	100

(Huom. siirtymätodennäköisyydet ovat ilmaistu prosentteina)

Esimerkki 3.2. Markovin ketjut voidaan esittää myös graafisessa muodossa. Alla on esitetty erään Markovin ketjun siirtymämatriisi P ja sen graafinen esitys.

$$P = \begin{bmatrix} 0 & 0.7 & 0 & 0.3 & 0 \\ 0.2 & 0 & 0.8 & 0 & 0 \\ 0.4 & 0 & 0 & 0 & 0.6 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0.7 & 0.3 & 0 & 0 \end{bmatrix}.$$



Kuva 3.1. Markovin ketju

Lause 3.1. *Olkoon Z_n Markovin ketju tila-avaruudessa* S*. Tällöin kaikilla ajanhetkillä $k \in \mathbb{N}$ pätee*

$$[\mathbb{P}(Z_{n+k} = j|Z_n = i)]_{i,j\in S} = [[P^k]_{i,j}]_{i,j\in S} = P^k,$$

kun $i, j \in$ S.

Todistus. Määritelmän 3.1 mukaan kaikilla ajanhetkillä $n \geq 1$ on voimassa $\mathbb{P}(Z_{n+1} = j|Z_n = i)$. Tällöin ehdollisen todennäköisyyden määritelmiä soveltamalla saadaan

$$\begin{aligned}
\mathbb{P}(Z_{n+2} = j|Z_n = i) &= \sum_{l\in S} \mathbb{P}(Z_{n+2} = j, Z_{n+1} = l|Z_n = i) \\
&= \sum_{l\in S} \frac{\mathbb{P}(Z_{n+2} = j, Z_{n+1} = l, Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l\in S} \frac{\mathbb{P}(Z_{n+2} = j, Z_{n+1} = l, Z_n = i)}{\mathbb{P}(Z_{n+1} = l, Z_n = i)} \frac{\mathbb{P}(Z_{n+1} = l, Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l\in S} \mathbb{P}(Z_{n+2} = j|Z_{n+1} = l, Z_n = i)\mathbb{P}(Z_{n+1} = l|Z_n = i) \\
&= \sum_{l\in S} \mathbb{P}(Z_{n+2} = j|Z_{n+1} = l)\mathbb{P}(Z_{n+1} = l|Z_n = i) \\
&= \sum_{l\in S} P_{i,l}P_{l,j} \\
&= [P^2]_{i,j'} \qquad i, j \in S\,.
\end{aligned}$$

Näin ollen kaikilla $k \in \mathbb{N}$ pätee

$$\begin{aligned}
\mathbb{P}(Z_{n+k+1} = j, Z_{n+k} = l | Z_n = i) &= \sum_{l \in \mathrm{S}} \mathbb{P}(Z_{n+k+1} = j, Z_{n+k} = l | Z_n = i) \\
&= \sum_{l \in \mathrm{S}} \frac{\mathbb{P}(Z_{n+k+1} = j, Z_{n+k} = l, Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l \in \mathrm{S}} \frac{\mathbb{P}(Z_{n+k+1} = j, Z_{n+k} = l, Z_n = i)}{\mathbb{P}(Z_{n+k} = l, Z_n = i)} \frac{\mathbb{P}(Z_{n+k} = l, Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l \in \mathrm{S}} \mathbb{P}(Z_{n+k+1} = j | Z_{n+k} = l, Z_n = i) \mathbb{P}(Z_{n+k} = l | Z_n = i) \\
&= \sum_{l \in \mathrm{S}} \mathbb{P}(Z_{n+k+1} = j | Z_{n+k} = l) \mathbb{P}(Z_{n+k} = l | Z_n = i) \\
&= \sum_{l \in \mathrm{S}} \mathbb{P}(Z_{n+k} = l | Z_n = i) P_{l,j}.
\end{aligned}$$

Siis matriisit

$$[\mathbb{P}(Z_{n+k} = j | Z_n = i)]_{i,j \in \mathrm{S}}, \quad k \geq 1,$$

toteuttavat ehdon

$$[P^{k+1}]_{i,j} = \sum_{l \in \mathrm{S}} [P^k]_{i,l} P_{l,j},$$

ja näin ollen yhtäsuuruus

$$[\mathbb{P}(Z_{n+k} = j | Z_n = i)]_{i,j \in \mathrm{S}} = [[P^k]_{i,j}]_{i,j \in \mathrm{S}} = P^k,$$

pätee kaikilla $k \in \mathbb{N}$. $\square$

Lause 3.2. *Kaksitilaisen Markovin ketjun, jonka siirtymämatriisi on muotoa*

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix},$$

kun $a \in [0,1]$ ja $b \in [0,1]$, n:s potenssi on muotoa

$$P^n = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}^n = \frac{1}{a+b} \begin{bmatrix} b + a(1-a-b)^n & a(1-(1-a-b)^n) \\ b(1-(1-a-b)^n) & a + b(1-a-b)^n \end{bmatrix}, \quad n \in \mathbb{N}, a = b \neq 0.$$

Todistus. Todistamme tuloksen käyttämällä matriisien diagonalisoituvuutta. Matriisilla P on kaksi lineaarisesti riippumatonta ominaisvektoria

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{ja} \quad \begin{bmatrix} -a \\ b \end{bmatrix},$$

joiden ominaisarvot ovat $\lambda_1 = 1$ ja $\lambda_2 = 1 - a - b$.

Täten matriisi P voidaan kirjoittaa diagonaalimuodossa

$$P = M \times D \times M^{-1}, \tag{3.1}$$

toisin sanoen

$$P = \begin{bmatrix} 1 & -a \\ 1 & b \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \times \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{-1}{a+b} & \frac{1}{a+b} \end{bmatrix}.$$

Nyt yhtälöstä (3.1) seuraa, että

$$\begin{aligned} P^n &= (M \times D \times M^{-1})^n = (M \times D \times M^{-1}) \dots (M \times D \times M^{-1}) \\ &= M \times D \times \cdots \times D \times M^{-1} = M \times D^n \times M^{-1}, \end{aligned}$$

missä

$$D^n = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2^n \end{bmatrix}, \quad n \in \mathbb{N}.$$

Täten

$$\begin{aligned} P^n &= \begin{bmatrix} 1 & -a \\ 1 & b \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2^n \end{bmatrix} \times \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{-1}{a+b} & \frac{1}{a+b} \end{bmatrix} \\ &= \begin{bmatrix} \frac{b+a\lambda_2^n}{a+b} & \frac{a-a\lambda_2^n}{a+b} \\ \frac{b-b\lambda_2^n}{a+b} & \frac{a+b\lambda_2^n}{a+b} \end{bmatrix} \\ &= \frac{1}{a+b} \begin{bmatrix} b + a\lambda_2^n & a(1-\lambda_2^n) \\ b(1-\lambda_2^n) & a + b\lambda_2^n \end{bmatrix} \\ &= \frac{1}{a+b} \begin{bmatrix} b + a(1-a-b)^n & a(1-(1-a-b)^n) \\ b(1-(1-a-b)^n) & a + b(1-a-b)^n \end{bmatrix}. \end{aligned}$$

□

3.2 Tilojen luokittelu

Tässä pykälässä esitämme muita Markovin ketjujen soveltamiseen tarvittavia määritelmiä (vrt. [2, s. 117–148]).

Määritelmä 3.4. Tilan $j \in \mathrm{S}$ *yhteyttä tilaan* $i \in \mathrm{S}$ merkitään kaaviolla

$$\text{(i)} \longmapsto \text{(j)},$$

jos on olemassa sellainen äärellinen luku $n \geq 0$, että

$$[P^n]_{i,j} = \mathbb{P}(Z_n = j | Z_0 = i) > 0.$$

Jos on voimassa sekä (i) $\longmapsto$ (j) että (j) $\longmapsto$ (i), sanotaan, että tilat (i) ja (j) ovat yhteydessä keskenään. Tätä merkitään (i) $\longleftrightarrow$ (j).

Määritelmä 3.5. Markovin ketjun sanotaan olevan *yhtenäinen*, jos

$$\text{(i)} \longmapsto \text{(j)}$$

kaikilla $i, j \in \mathrm{S}$. Muulloin se on *epäyhtenäinen*.

Määritelmä 3.6. Tila $i \in \mathrm{S}$ on *palautuva*, jos Markovin ketju palautuu tilasta i tilaan i äärellisellä määrällä askeleita todennäköisyydellä $\mathbb{P} = 1$. Toisin sanoen

$$p_{i,i} := \mathbb{P}(Z_n = i \text{ jollakin } n \geq 1 | Z_0 = i) = 1.$$

Määritelmä 3.7. Tila $i \in \mathrm{S}$ on *väistyvä*, jos se ei ole palautuva.

Määritelmä 3.8. Tilan $i \in \mathrm{S}$ palautumisaikaa merkitään $\mu_i(i)$ ja määritellään

$$\begin{aligned}\mu_i(i) :&= \mathrm{E}[T_i^r | Z_0 = 1] \\ &= \sum_{n=1}^{\infty} n\mathbb{P}([T_i^r = n | Z_0 = 1),\end{aligned}$$

jossa T_i^r merkitsee satunnaista ajanhetkeä, jolloin on päädytty ketjun tilaan i kulkemalla r askelta. Tilan sanotaan olevan positiivisesti palautuva, jos

$$\mu_i(i) = \mathrm{E}[T_i^r | Z_0 = 1] < \infty,$$

ja tyhjästi palautuva, jos

$$\mu_i(i) = \mathrm{E}[T_i^r | Z_0 = 1] = +\infty.$$

Määritelmä 3.9. Tilan $i \in \mathrm{S}$ *jakso* on joukon

$$\{n \geq 1 : [P^n]_{i,i} > 0\}$$

suurin yhteinen jakaja.

Tilan sanotaan olevan *jaksoton*, jos sen jakso on 1. Muulloin tilan sanotaan olevan *jaksollinen*.

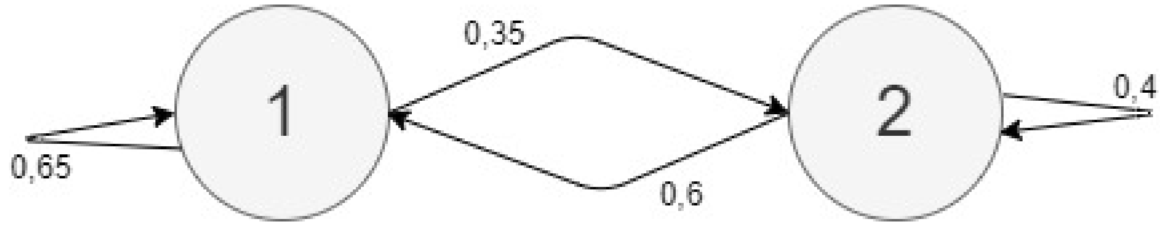
3.3 Sovelluksia

Tässä pykälässä esitämme esimerkkejä Markovin ketjujen ominaisuuksia soveltamisesta.

Esimerkki 3.3. Mallinnetaan syyspäivän $t = 0, 1, 2, \ldots$ säätilaa satunnaisprosessina tilajoukossa $\mathrm{S} = 1, 2$, jossa tila 1 = sataa ja 2 = ei sada. Oletetetaan, että jos päivänä n sataa, niin päivänä $n+1$ ei sada todennäköisyydellä $a = 0,35$, ja että kuivaa päivää seuraa sateinen päivä todennäköisyydellä $b = 0,6$. Tällöin säätä voidaan mallintaa Markovin ketjuna $(Z_0, Z_1, \ldots)$, jonka siirtymämatriisi P on

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} 0,65 & 0,35 \\ 0,6 & 0,4 \end{bmatrix}$$

ja graafinen esitys on kuvassa 3.2.



Kuva 3.2. Säätilan graafinen kuvaaja

Oletetaan, että maanantaina ($t = 0$) sataa. Tällöin säämallin mukaan tiistaina sataa todennäköisyydellä $1 - a$ ja ei sada todennäköisyydellä a, eli

$$\mathbb{P}(Z_1 = 1 | Z_0 = 1) = 1 - a \quad \text{ja} \quad \mathbb{P}(Z_1 = 2 | Z_0 = 1) = a.$$

Näin ollen todennäköisyys sille, että keskiviikkona sataa, voidaan laskea tiistain säätilojen avulla

$$\begin{aligned} \mathbb{P}(Z_2 = 1 | Z_0 = 1) &= (1-a)\mathbb{P}(Z_2 = 1 | Z_1 = 1) + a\mathbb{P}(Z_2 = 1 | Z_1 = 2) \\ &= (1-a)(1-a) + ab \\ &= (1-a)^2 + ab. \end{aligned}$$

Siis todennäköisyys sille, ettei keskiviikkona sada on $(1-a)^2 + ab = 0,6325$.

Lauseen 3.2 avulla voidaan myös sateen todennäköisyys mille tahansa päivälle n. Esimerkiksi se, sataako kahden viikon (14 vuorokauden) päästä aloitushetkestä voidaan laskea kaavalla

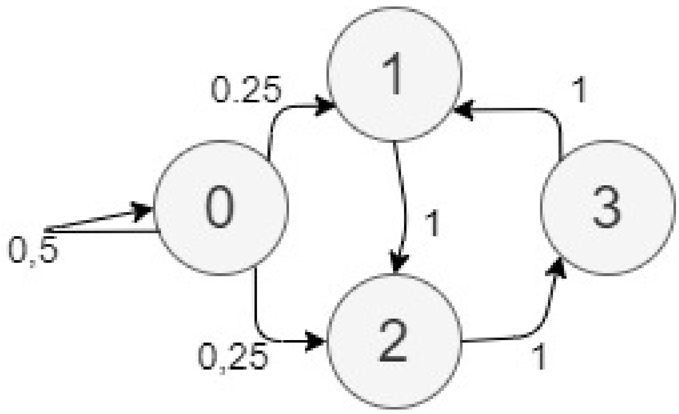
$$
\begin{aligned}
P^{15} &= \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}^{15} \\
&= \frac{1}{a+b} \begin{bmatrix} b + a(1-a-b)^{15} & a(1-(1-a-b)^{15}) \\ b(1-(1-a-b)^{15}) & a + b(1-a-b)^{15} \end{bmatrix} \\
&= \frac{1}{0,35+0,6} \begin{bmatrix} 0,6 + 0,35(0,65-0,6)^{15} & 0,35(1-(0,65-0,6)^{15}) \\ 0,6(1-(0,65-0,6)^{15}) & 0,35 + 0,6(0,65-0,6)^{15} \end{bmatrix} \\
&= \frac{1}{0,95} \begin{bmatrix} 0,6 & 0,35 \\ 0,6 & 0,35 \end{bmatrix} \\
&= \begin{bmatrix} 0,63157\ldots & 0,36842\ldots \\ 0,63157\ldots & 0,36842\ldots \end{bmatrix}.
\end{aligned}
$$

Esimerkki 3.4 (Tilojen luokittelu)**.** Olkoon $(Z_n)_{n\geq 0}$ Markovin ketju tila-avaruudessa $S = \{0, 1, 2, 3\}$ siten, että sen siirtymämatriisi on

$$
P = \begin{bmatrix} 0,5 & 0,25 & 0,25 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.
$$

Tällöin kejun Z_n graafinen esitys on

.



Kuva 3.3. Esimerkki 3.4 kuvaaja

Ketjulle pätee yhteydet

$$
(0) \longmapsto (1), (0) \longmapsto (2), (1) \longmapsto (2), (2) \longmapsto (3) \text{ ja } (3) \longmapsto (1),
$$

ja se on yhtenäinen, sillä sen tila-avaruus voidaan jakaa kahteen yhteydessä olevaan tilaan $S = \{0\} \cup \{1, 2, 3\}$. Ketjun tilan 0 jakso on 1 ja tilojen 1, 2 ja 3 jaksot ovat 3. Ketjussa ei ole absorvoivia tiloja. Tila 0 on väistyvä ja tilat 1, 2 ja 3 ovat palautuvia.

Lähteet

[1] Geiss, C. & Geiss, S. *An introduction to probability theory*. Department of Mathematics and Statistics University of Jyväskylä, 2009.

[2] Privault, N. *Understanding Markov Chains*. New York: Springer, 2013.