

kandi työotsikko

Topias Karjalainen

11. maaliskuuta 2020

Sisältö

1	Johdanto	2
2	Yleisiä tuloksia	3
2.1	Perusmääritelmiä	3
2.2	Markovin ketjut	4
2.2.1	Äärellinen tilajoukko	4
3	Metropolis–Hastings algoritmi	6

Luku 1

Johdanto

Tilastotieteissä *frekventistinen* koulukunta oli pitkään vallitseva koulukunta. Bayesiläinen päättely ei päässyt leviämään, sillä toisin kuin frekventistinen koulukunta, Bayesiläisyys ei tarjonnut suurinpaan osaan kysymyksiä analyyttisiä ratkaisuja. Vasta tietokoneiden aikakautena *Markovin ketju Monte Carlo -menetelmät* (MCMC-menetelmät) ovat antaneet mahdollisuuden ratkaista *posteriori*-jakaumat monimutkaisemmilta malleilta.

Tässä kandidutkielmassa aion selventää MCMC-menetelmien teoriaa, sekä esittää yleisimmät kaksi algoritmia: *Gibbs*- ja *Metropolis–Hastings* algoritmit.

Luku 2

Yleisiä tuloksia

2.1 Perusmääritelmiä

Määritellään ensiksi todennäköisyys.

Määritelmä 2.1. *σ -algebra.* Olkoot Ω mielivaltainen epätyhjä joukko. Sigma-algebra perusjoukolla Ω on sen osajoukkojen joukkoperhe \mathcal{F} , joka toteuttaa ehdot:

1. $\emptyset \in \mathcal{F}$
2. jos $A \in \mathcal{F}$, niin $A^c \in \mathcal{F}$
3. jos jos $A_k \in \mathcal{F}$, kaikilla $k \in K$, missä K on numeroituva joukko, niin $\bigcup_{k \in K} A_k \in \mathcal{F}$

Määritelmä 2.2. Kuvaus \mathbf{P} liittää kuhunkin tapahtumaan A *todennäköisyyden*, joka on luku suljetulla välillä $[0,1]$ ja sille pätee:

1. $\mathbf{P}(\Omega) = 1$
2. Jos A on tapahtuma, niin sen komplementtitapahtuman A^c todennäköisyys on $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$
3. Jos $(A_k)_{k \in \mathbb{N}}$ ovat erillisiä tapahtumia, niin

$$\mathbf{P}\left(\bigcup_{k \in \mathbb{N}} A_k\right) = \sum_{k \in \mathbb{N}} \mathbf{P}(A_k)$$

Määritelmä 2.3. Kolmikkoa $(\Omega, \mathcal{F}, \mathbf{P})$ kutsutaan *todennäköisyysavaruuudeksi*.

Määritelmä 2.4. *Satunnaismuuttuja* X on (lähes) mielivaltainen kuvaus $X : \Omega \rightarrow S$, jossa S on *tilajoukko*.

2.2 Markovin ketjut

2.2.1 Äärellinen tilajoukko

Määritelmä 2.5. Jono $(X_n : n = 1, 2, 3, \dots)$ satunnaismuuttujia on diskreettiaikainen stokastinen prosessi.

Merkintä 2.6. Merkitään stokastista prosessia merkinnällä $\{X_n\}_{n \in T}$

Määritelmä 2.7. Stokastinen prosessi $\{X_n\}$ on *Markovin ketju*, jos kaikilla alkuhetkillä m, n ja tiloilla $i, j \in S$ on voimassa

$$(2.8) \quad \begin{aligned} \mathbf{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ = \mathbf{P}(X_{n+1} = j | X_n = i) \end{aligned}$$

ja *siirtymätodennäköisyyksille* on voimassa

$$(2.9) \quad p_{ij} = \mathbf{P}(X_{n+1} = j | X_n = i) = \mathbf{P}(X_{m+1} = j | X_m = i)$$

Yhtälöä 2.8 kutsutaan *Markovin-ehdoksi* ja yhtälöä 2.9 taas kutsutaan *stationarisuusehdoksi*, mikä tarkoittaa, että siirtymätodennäköisyys tilojen i ja j välillä ei riipu ajasta m ja n , vaan pelkästään tiloista i ja j .

Määritelmä 2.10. Satunnaismuuttujan X_0 jakaumaa kutsutaan *alkujakaumaksi*.

Lause 2.11. *Ajanhetkellä $n \geq 1$ polun (i_0, \dots, i_n) todennäköisyys on*

$$(2.12) \quad \mathbf{P}(X_0 = i_0, \dots, X_n = i_n) = p_{i_0} p_{i_0, i_1} p_{i_1, i_2} \dots p_{i_{n-1}, i_n}$$

Todistus. Käyttäen ehdollisen todennäköisyyden kaavaa, saadaan 2:lle tapahtumalle

$$\mathbf{P}(A_0, A_1) = \mathbf{P}(A_0) \mathbf{P}(A_1 | A_0)$$

Jos tapahtumia on kolme, saadaan

$$\mathbf{P}(A_0, A_1, A_2) = \mathbf{P}(A_0) \mathbf{P}(A_1 | A_0) \mathbf{P}(A_2 | A_1, A_0)$$

neljä

$$\mathbf{P}(A_0, A_1, A_2, A_3) = \mathbf{P}(A_0) \mathbf{P}(A_1 | A_0) \mathbf{P}(A_2 | A_1, A_0) \mathbf{P}(A_3 | A_2, A_1, A_0)$$

ja n

$$(2.13) \quad \mathbf{P}(A_0, \dots, A_n) = \mathbf{P}(A_0) \mathbf{P}(A_1 | A_0) \dots \mathbf{P}(A_n | A_{n-1}, \dots, A_0)$$

Tämä on yleinen ehdollinen todennäköisyys. Merkataan $A_n := (X_i = i_n)$. Koska käsittelemme Markovin ketjua, niin yhtälö 2.8 pätee, jolloin yhtälöstä 2.13 saadaan

$$\mathbf{P}(X_0 = i_0, \dots, X_n = i_n) = \mathbf{P}(X_0 = i_0) \mathbf{P}(X_1 = i_1 | X_0 = i_0) \dots \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1})$$

jossa $\forall n = 0, 1, 2, \dots, n : \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1})$ on siirtymätodennäköisyys p_{i_{n-1}, i_n} jolloin tulos seuraa substituomalla termit. \square

Merkintä 2.14.

$$(2.15) \quad p_{ij}^{(m)} := \mathbf{P}(X_m = j | X_0 = i), \quad i, j \in S, m \in T$$

on siirtymätodennäköisyys tilasta i tilaan j , kun aikaa kuluu m yksikköä.

Määritelmä 2.16. *Siirtymämatriisi* on matriisi

$$(2.17) \quad \mathbf{P}^{(m)} := (p_{ij}^{(m)})_{i,j} = \begin{pmatrix} p_{00}^{(m)} & p_{01}^{(m)} & \dots & p_{0n}^{(m)} \\ p_{10}^{(m)} & p_{11}^{(m)} & \dots & p_{1n}^{(m)} \\ \vdots & & \ddots & \vdots \\ p_{n0}^{(m)} & p_{n1}^{(m)} & \dots & p_{nn}^{(m)} \end{pmatrix}$$

Lause 2.18. *Kaikilla ajanhetkillä on voimassa*

$$(2.19) \quad \mathbf{P}^{(m)} = \mathbf{P}^m$$

Todistus. Todistus on melko pitkä, joten ohitetaan se. \square

Määritelmä 2.20. Todennäköisyysjakauma $\pi = (\pi)_{i \in S}$ on Markovin ketjun $\{X_n\}$ tasapainojakauma, jos

$$(2.21) \quad \sum_{i \in S} \pi_i p_{ij} = \pi_j, \quad \forall j \in S$$

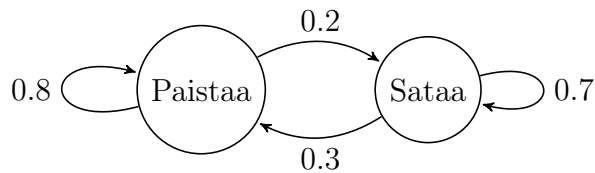
Yhtälö 2.21 voidaan kirjoittaa myös muotoon

$$(2.22) \quad \pi \mathbf{P} = \pi$$

Esimerkki 2.23. Pohditaan lyhyttä esimerkkiä, jossa tilajoukko on $S = \{ \text{"sataa"}, \text{"paistaa"} \}$. Määritellään siirtymätodennäköisyydet siirtymämatriisilla

$$\mathbf{P}^{(1)} = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}$$

Tämä voidaan visualisoida seuraavanlaisesti:



Luku 3

Metropolis–Hastings algoritmi

Kirjallisuutta