8. Yleinen äärellinen Markovin prosessi

Olemme jo päässeet niin pitkälle, että voimme määritellä yleisen Markovin prosessin, kun tilajoukko on äärellinen. Ajatus on tarkalleen sama kuin SK-prosessin tapauksessa. Korvaamme siirtymätodennäköisyydet p_{ij} *siirtymäintensiteeteillä* q_{ij}, kunhan $i \neq j$.

8.1. **Oletus.** Annettu parametrit $q_{ij} \geq 0$, kun $i, j \in S = \{1, 2, \ldots, d\}$ ja $i \neq j$. Näitä nimitämme *siirtymäintensiteeteiksi tilasta i tilaan j*.

8.2. **Oletus.** Ehdollisille todennäköisyyksille on arviot

$$\mathbf{P}\,(\,X(t+h) = j \,|\, X(t) = i\,) \approx q_{ij}h$$

Kuten SK-prosessin tapauksessa, voimme tulkita todennäköisyydet "kilpailuksi" $d-1$ riippumattoman eksponenttijakautuneen satunnaismuuttujan välillä siten, että mikä ehtii tapahtua ensin. Eli jos $i = 1$, niin $T_{12} \sim \text{Exp}\,(\,q_{12}\,), \ldots, T_{1d} \sim \text{Exp}\,(\,q_{1d}\,)$ kilpailevat siitä, mikä on pienin eli $T_1 := \min\{T_{12}, T_{13}. \ldots .T_{1d}\}$. Satunnainen aika T_1 on tilassa 1 viipymisaika, ja se on $\text{Exp}\,(\,q_1\,)$-jakautunut, kun

$$q_1 := \sum_{j \neq 1} q_{1j}.$$

Vastaavasti käymällä läpi muut tilat $2, 3, \ldots, d$ voimme määritellä viipymisajan $T_i := \min\{\,T_{ij} : \; j \neq i\,\}$ tilassa i, joka on vastaavasti $\text{Exp}\,(\,q_i\,)$-jakautunut ja

$$q_i := \sum_{j \neq i} q_{ij}.$$

Siispä tilasta i siirtymistodennäköisyys aikavälillä $(t, t+h]$ on

$$\mathbf{P}\,(\,X(t+h) \neq i \,|\, X(t) = i\,) \approx q_i h$$

ja vastaava viipymistodennäköisyys tilassa i on siten

$$\mathbf{P}\,(\,X(t+h) = i \,|\, X(t) = i\,) \approx 1 - q_i h$$

8.1. **DY Markovin prosessille.** Aivan kuten SK-prosessin tapauksessa voimme määrätä differentiaaliyhtälön prosessin jakaumalle. Tarvitsimme DY:n johtamiseen siirtymäintensiteettejä sekä tietoa viipymisestä. Koska nyt muuttujia on enemmän, on merkintöjä pyrittävä tiivistämään. Merkitsemme aluksi

$$p_{ij}(t) := \mathbf{P}\,(\,X(t) = j \,|\, X(0) = i\,) = \mathbf{P}_i\,(\,X(t) = j\,)\,,$$

mikä suoraan yleistää Markovin ketjun yleistetyn siirtymätodennäköisyyden $p_{ij}^{(n)}$. Markovin ketjun tapauksessa havaitsimme matriisimerkinnän hyödylliseksi, joten vastaavasti merkitsemme

$$P(t) := \begin{pmatrix} p_{11}(t) & \dots & p_{1d}(t) \\ \vdots & & \vdots \\ p_{d1}(t) & \dots & p_{dd}(t) \end{pmatrix}$$

Merkitsemme vielä $q_{ii} := -q_i$, sillä siirtymäintensiteettiä tilasta i itseensä emme ole määritelleet ja kuten muistamme, se on kyseenalainen käsite. Myös merkinnästä havaitsemme, että asetamme "intensiteetin" negatiiviseksi! Mutta tästä merkinnästä on kuitenkin hyötyä! Tämän merkinnän johdosta voimme koota myös siirtymäintensiteetit matriisimuotoon:

$$Q := \begin{pmatrix} q_{11} & q_{12} & \dots & q_{1d} \\ \vdots & & \vdots & \\ q_{d1} & q_{d2} & \dots & q_{dd} \end{pmatrix}$$

Tätä matriisia nimitämme Markovin prosessin *intensiteettimatriisiksi.* Aivan kuten aikasemminkin laskemme seuraavasti:

$$\begin{aligned} p_{ij}(t) = \mathbf{P}_i\,(\,X(t) = j\,) &= \sum_{k=1}^{d} \mathbf{P}_i\,(\,X(t-h) = k, X(t) = j\,) \\ &= \sum_{k=1}^{d} \mathbf{P}_i\,(\,X(t-h) = k\,)\,\mathbf{P}\,(\,X(t) = j \,|\, X(0) = i, X(t-h) = k\,) \\ &= \sum_{k \neq j} \mathbf{P}_i\,(\,X(t-h) = k\,)\,\mathbf{P}\,(\,X(t) = j \,|\, X(t-h) = k\,) \\ &\quad + \mathbf{P}_i\,(\,X(t-h) = k\,)\,\mathbf{P}\,(\,X(t) = j \,|\, X(t-h) = j\,) \\ &\approx \sum_{k \neq j} p_{ik}(t-h) q_{kj} h + p_{ij}(t-h)(1 - q_j h) \\ &= \sum_{k=1}^{d} p_{ik}(t-h) q_{kj} h + p_{ij}(t-h) \end{aligned}$$

Toiselta riviltä kolmannelle siirtymisessä käytimme Markov-ominaisuutta, jota emme ole siis aivan tarkasti määritelleet. Oli määritelmä kuitenkin mikä hyvänsä niin kyseisen yhtälön tulisi olla voimassa. Viimeisessä identiteetissä käytimme hyväksi sopimusta $q_{jj} = -q_j$. Siirtämällä termi $p_{ij}(t-h)$ vasemmalle puolelle, voimme päätellä että $p_{ij}(t)$ on jatkuva funktio, sillä

$$\lim_{h \to 0^+} p_{ij}(t) - p_{ij}(t-h) = 0$$

Jakamalla edelleen muuttujalla h ja antamalla $h \to 0^+$ saamme yhtälön

$$p'_{ij}(t) = \sum_{k=1}^{d} p_{ik}(t) q_{kj}$$

Olemme osoittaneet seuraavan tuloksen

8.3. **Lause.** *Yleisen äärellisen Markovin prosessin jakauma toteuttaa differentiaaliyhtälön*

$$p'_{ij}(t) = \sum_{k=1}^{d} p_{ik}(t) q_{kj} \tag{8.4}$$

jokaisella $i, j \in \{1, \dots, d\}$. Matriisimuodossa tämä yhtälö on

$$P'(t) = P(t)Q. \tag{8.5}$$

Voimme ratkaista tämän ensimmäisen kertaluvun differentiaaliyhtälön matriiseille varsin suoraviivaisesti.

8.6. **Lause.** *Siirtymämatriisi $P(t)$ voidaan esittää muodossa*

$$P(t) = e^{tQ} =: I + \sum_{n=1}^{\infty} \frac{Q^n t^n}{n!}.$$

Todistus. Havaitsemme välittömästi, että $P(0) = I$, joten väite on ainakin järkevä. Jos matriisin dimensio on 1×1, matriisiyhtälöä (8.5) vastaa tarkalleen yksi tavallinen DY. Jos merkitsemme $p(t) := p_{11}(t)$ ja $q := q_{11}$, niin tapauksessa $d = 1$ saamme

$$\begin{cases} p'(t) & = p(t)q, \\ p(0) & = 1 \end{cases} \implies \quad p(t) = e^{tq}.$$

Siispä väite on järkevä myös, kun $d = 1$.

Yleisessä tilanteessa voimme osoittaa väitteen derivoimalla. Merkitsemme $S(t)$:llä väitteen oikeaa puolta. Laskemalla saamme, että

$$S'(t) = D\Big(I + \sum_{n=1}^{\infty} \frac{Q^n t^n}{n!}\Big) = \sum_{n=1}^{\infty} D\Big(\frac{Q^n t^n}{n!}\Big)$$

Derivaatan viemien matriisipotenssisarjan sisälle vaatisi hieman perusteluja (HT. mieti perustelut tarkastelemalla osasummia ja derivoimalla komponenteittain), mutta niin vaatisi myös sarjan suppeneminenkin. Jos A on mikä tahansa $d \times d$-matriisi, niin matriisin

$$\frac{At^n}{n!}$$

derivaatta on

$$D\Big(\frac{At^n}{n!}\Big) = \frac{At^{n-1}}{(n-1)!}.$$

Siispä

$$S'(t) = D\Big(I + \sum_{n=1}^{\infty} \frac{Q^n t^n}{n!}\Big) = \sum_{n=1}^{\infty} D\Big(\frac{Q^n t^n}{n!}\Big) = \sum_{n=1}^{\infty} \frac{Q^n t^{n-1}}{(n-1)!}$$

Vaihtamalla summausmuuttuja n muuttujaksi $n' = n - 1$, saamme

$$S'(t) = Q + \sum_{n=1}^{\infty} \frac{Q^{n+1} t^n}{n!} = \Big(I + \sum_{n=1}^{\infty} \frac{Q^n t^n}{n!}\Big)Q = S(t)Q$$

Siispä myös $S(t)$ totetuttaa differentiaaliyhtälön (8.4). Lisäksi $S(0) = I = P(0)$. Ei ole erityisen hankalaa osoittaa, että tästä seuraa, että $S(t) = P(t)$ jokaisella t.

Eräs tapa on seuraavanlainen: Huomaamme, että jos $A(t)$ toteuttaa yhtälön (8.4) ja $A(0) = 0$, niin

$$A(t) = \int_0^t A'(s)\,\mathrm{d}s = \int_0^t A(s)Q\,\mathrm{d}s,$$

missä integrointi suoritetaan komponenteittain. Voimme nyt ottaa tästä itseisarvot ja arvioida

$$|A_{ij}(t)| \le \int_0^t |(A(s)Q)_{ij}|\,\mathrm{d}s.$$

Matriisituloa voimme arvioida maksimin (eli normin avulla), sillä

$$|(A(s)Q)_{ij}| \le \sum_{k=1}^{d} \max_{ij} |A_{ij}(s)||Q_{ij}| =: d\| A(s) \| \times \| Q \|$$

Saimme siis arvion

$$\| A(t) \| = \max_{i,j} |A_{ij}(t)| \le d\| Q \| \int_0^t \| A(s) \|\,\mathrm{d}s = C \int_0^t \| A(s) \|\,\mathrm{d}s.$$

Funktio $u(t) := \| A(t) \| \ge 0$ on tavallinen jatkuva yhden muuttujan reaaliarvoinen funktio. Iteroimalla arviota

$$u(t) \le C \int_0^t u(s)\,\mathrm{d}s$$

voimme päätellä, että

$$u(t) \le \frac{C^{n+1}}{n!} \int_0^t (t-s)^n u(s)\,\mathrm{d}s \to 0,$$

kun $n \to \infty$. Siispä $u(t) = 0$ jokaisella t ja myös $A(t) = 0$. Koska $S(t) - P(t) =: A(t)$ toteuttaa nämä ehdot, niin olemme näyttäneet, että $P(t) = S(t)$. $\square$

8.2. **Markovin prosessin TP-jakauma.** Määrittelimme SK-prosesseille tasapainojakauman. Voimme yleistää tämän nyt myös äärellisille Markovin prosesseille.

8.7. **Määritelmä.** Oletetaan, että $\mathbf{P}\left(X(0) = i \right) = \pi_i$. Jakauma π on *tasapainojakauma*, jos

$$\sum_i \pi_i p_{ij}(t) = \pi_j$$

jokaisella $j \in \{1, \ldots, d\}$ ja jokaisella t. Matriisimuodossa tämä voidaan esittää

$$\pi P(t) = \pi,$$

kun $\pi = (\pi_1 \pi_2 \ldots \pi_d)$ on vaakavektori.

SK-prosessien tapauksessa huomasimme, että tasapainojakauma toteutti differenssiyhtälön, jonka johdimme SK-prosessin differentiaaliyhtälöstä. Voimme nyt lopulta päättää yleisten äärellisten Markovin prosessien tarkastelun johtamalla (ajasta t riippumattoman) lineaarisen yhtälöryhmän Markovin prosessin tasapainojakaumalle.

Koska

$$\pi P(t) = \pi,$$

niin derivoimalla havaitsemme, että

$$\pi P'(t) = 0.$$

Differentiaaliyhtälön (8.5) mukaan siis

$$0 = \pi P'(t) = \pi\big(P(t)Q\big) = \big(\pi P(t)\big)Q = \pi Q.$$

Differenssiyhtälön johtaminen onnistui tarkalleen samalla tavalla kuin SK-prosessin yhteydessä. Oli vain huomattava, että matriisitulojen järjestystä ei saa muuttaa (sillä $AB = BA$ ei ole yleensä matriiseilla voimassa), mutta ryhmitellä niitä kyllä saa.

Haluttu yhtälö on siis

$$\pi Q = 0 \Longrightarrow \sum_{i=1}^{d} \pi_i q_{ij} = 0 \text{ jokaisella } j = 1, \ldots, d. \tag{8.8}$$