# Kapitel 4
# Diskrete Markov-Prozesse

# 4.1 Definition und grundlegende Eigenschaften

- Stetige Zeit aber weiterhin diskreter Zustandsraum.
- Beispiele:
  - Klinische Studien, Krankenverläufe
  - Markenwahlprozesse
  - Poisson-Prozess
  - Stochastische Modelle für die molekulare Evolution, d.h. die durch Mutation verursachte Veränderung der Nukleotide A, T, C, G an bestimmten Positionen der DNA. Vereinfachende Annahme: Die Evolution an verschiedenen Positionen verläuft unabhängig. Dann genügt es, Modelle separat für die Positionen zu entwickeln.

    Weitere Annahme: Homogenität und Stationarität. Verschiedene Modelle unterscheiden sich durch Annahmen an die Raten $\lambda_{ij}$ für Übergänge $i \to j$.

- Definition: Diskreter Markov-Prozess

Ein stochastischer Prozess $X = \{X(t), t \geq 0\}$ mit abzählbarem Zustandsraum heißt (diskreter) Markov-Prozess $:\Leftrightarrow$ $\forall n \geq 0$, $\forall t > s > s_n > \ldots > s_0 \geq 0$ gilt

$$P(X(t) = j | X(s) = i, X(s_n) = i_n, \ldots, X(s_0) = i_0) = P(X(t) = j | X(s) = i)$$

für beliebige $j, i, i_n, \ldots, i_0 \in S$.

$$p_{ij}(s,t) := P(X(t) = j | X(s) = i)$$

heißt Übergangswahrscheinlichkeit(-sfunktion) und

$$P(s,t) = (p_{ij}(s,t))$$

Übergangsmatrix. Ein Markov-Prozess heißt homogen (oder besitzt stationäre Übergangswahrscheinlichkeiten) $:\Leftrightarrow$

$$P(X(t+s) = j | X(s) = i) = P(X(t) = j | X(0) = i) = p_{ij}(0,t) =: p_{ij}(t),$$

d.h. nur die Zeitdifferenz ist maßgeblich.

- Bemerkungen:
  - Wie bei den Markov-Ketten werden wir uns im Folgenden meist auf homogene Markov-Prozesse beschränken.
  - $p_{ij}(t)$ entspricht $p_{ij}^{(t)}$ ($t$-schrittige Übergangswahrscheinlichkeit) bei Markov-Ketten.
  - Für die Übergangswahrscheinlichkeiten gilt

$$p_{ij}(t) \geq 0, \qquad \sum_{j \in S} p_{ij}(t) = 1$$

- Chapman-Kolmogorov-Gleichung (für homogene Markov-Prozesse)

$$p_{ij}(s+t) = \sum_{k \in S} p_{ik}(s) p_{kj}(t)$$

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

- Gemeinsame Verteilung: $\forall n \in \mathbb{N}$, $\forall 0 \leq t_0 < t_1 < \ldots < t_n \in \mathbb{R}_+$, $i_0, i_1, \ldots, i_n \in S$ gilt

$$\begin{aligned} P(X(t_n) &= i_n, \ldots, X(t_1) = i_1 | X(t_0) = i_0) \\ &= p_{i_0 i_1}(t_1 - t_0) \cdot \ldots \cdot p_{i_{n-1} i_n}(t_n - t_{n-1}) \end{aligned}$$

$$\begin{aligned} P(X(t_n) &= i_n, \ldots, X(t_1) = i_1, X(t_0) = i_0) \\ &= P(X(t_0) = i_0) \cdot p_{i_0 i_1}(t_1 - t_0) \cdot \ldots \cdot p_{i_{n-1} i_n}(t_n - t_{n-1}). \end{aligned}$$

$\Rightarrow$ Satz von Kolmogorov & Likelihood.

- Im Folgenden setzen wir stets voraus

$$p_{ij}(0) = \lim_{h \downarrow 0} p_{ij}(h) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

d.h. ein Übergang von $i$ nach $j$ ($i \neq j$) beansprucht eine positive Zeitspanne.

- Definition: (Übergangs-)Intensität, Rate

Für $i, j \in S,\ i \neq j$, heißt

$$\begin{aligned}\lambda_{ij} &= \lim_{h \downarrow 0} \frac{p_{ij}(h)}{h} = \lim_{h \downarrow 0} \frac{P(X(t+h) = j | X(t) = i)}{h} \\ &= \lim_{h \downarrow 0} \frac{P(X(t+h) = j | X(t) = i, \text{Vorgeschichte})}{h} = p'_{ij}(0)\end{aligned}$$

Übergangsintensität bzw. Rate für den Übergang von $i$ nach $j$.

Für $i = j$ definieren wir

$$\lambda_{ii} = \lim_{h \downarrow 0} \frac{p_{ii}(h) - p_{ii}(0)}{h} = \lim_{h \downarrow 0} \frac{p_{ii}(h) - 1}{h}.$$

Allgemein gilt also für $i, j \in S$

$$\lambda_{ij} = p'_{ij}(0) = \lim_{h \downarrow 0} \frac{p_{ij}(h) - p_{ij}(0)}{h}$$

mit $0 \leq \lambda_{ij} < \infty$ für $i \neq j$, aber $\lambda_{ii} \leq 0$.

- $\Lambda = (\lambda_{ij})$ heißt Intensitätsmatrix.

- In $o(h)$–Schreibweise gilt also für $i \neq j$

$$\begin{aligned} p_{ij}(h) &= P(X(t+h) = j | X(t) = i) = \lambda_{ij} h + o(h), \\ p_{ii}(h) &= 1 + \lambda_{ii} h + o(h). \end{aligned}$$

- Beispiel: Poisson-Prozess

$$\begin{aligned} p_{i,i+1}(h) &= \lambda h + o(h) & \lambda_{i,i+1} &= \lambda \\ p_{ij}(h) &= \begin{cases} 0, & 0 \leq j \leq i-1 \\ o(h), & j \geq i+2 \end{cases} & \lambda_{ij} &= \begin{cases} 0, & 0 \leq j \leq i-1 \\ 0, & j \geq i+2 \end{cases} \\ p_{ii}(h) &= 1 - \lambda h + o(h) & \lambda_{ii} &= -\lambda \end{aligned}$$

# 4.2 Geburts- und Todesprozesse

- Definition: Geburtsprozess

  Ein stochastischer Prozess $X = \{X(t), t \geq 0\}$ mit Zustandsraum $S \subseteq \mathbb{N}_0$ heißt Geburtsprozess $:\Leftrightarrow$ $X$ ist ein homogener, diskreter MP mit

$$\begin{aligned} p_{i,i+1}(h) &= \lambda_i h + o(h), \\ p_{i,i}(h) &= 1 - \lambda_i h + o(h), \qquad \lambda_i \geq 0 \qquad i \geq i_0 \\ p_{i,j}(h) &= \begin{cases} 0 & , \quad 0 \leq j \leq i-1 \\ o(h) & , \quad j \geq i+2 \end{cases} \end{aligned}$$

- Datenstruktur wie beim Poisson-Prozess, aber zustandsabhängige Intensitäten $\lambda_i$.

- Analog zum Poisson-Prozess gilt

$$\lambda_{i,i+1} = \lambda_i \ , \quad \lambda_{ii} = -\lambda_i, \quad \lambda_{ij} = 0 \ \text{für} \ j \notin \{i, i+1\}.$$

- Herleitung der Übergangswahrscheinlichkeiten

$$p_{0n}(t) = P(X(t) = n \mid X(0) = 0)$$

aus den Intensitäten.

- Eine Wahrscheinlichkeitsverteilung

$$\sum_{n=0}^{\infty} p_{0n}(t) = 1 \quad \forall\, t \geq 0$$

ergibt sich genau dann, wenn

$$\sum_{i=0}^{\infty} \frac{1}{\lambda_i} = \infty.$$

Gegenbeispiel: Rate $\lambda_i = 2^i$.

- Exponentialverteilte Verweildauern

  Die Verweildauern $T_i$, $i \in \mathbb{N}$, sind unabhängig und exponentialverteilt mit Parameter $\lambda_i$, d.h.

$$T_i \sim Ex(\lambda_i).$$

- Interpretation: $\sum_{i=0}^{\infty} \frac{1}{\lambda_i}$ ist die erwartete Zeit bevor $\{X(t) = \infty\}$ eintritt. Falls die Summe divergiert scheint plausibel, dass $P\{X(t) = \infty\} = 0$ gilt.

- Yule-Prozess: Population mit Anfangsbestand von $X(0) = i_0$ Individuen.

  Unabhängige Vermehrung, wobei jedes Individuum im Zeitintervall $h$ mit Wahrscheinlichkeit $\lambda h + o(h)$ ein weiteres Mitglied der Population erzeugt und die Wahrscheinlichkeit für die Erzeugung mehrerer Individuen $o(h)$ sei.

  Der binomische Lehrsatz liefert

$$\begin{aligned} p_{i,i+1}(h) &= \binom{i}{1}(\lambda h + o(h))(1 - \lambda h + o(h))^{i-1} = i\lambda h + o(h) \\ p_{i,j}(h) &= \binom{i}{j-i}(\lambda h + o(h))^{j-i}(1 - \lambda h + o(h))^{i-(j-i)} = o(h) \text{ für } j - i \geq 2. \end{aligned}$$

$\Rightarrow$ Geburtsprozess mit Rate $\lambda_i = i\lambda$.

$X(t) - i_0$ besitzt eine negative Binomialverteilung mit Parametern $i_0$ und $\exp(-\lambda t)$.

$$\begin{aligned}
E(X(t) - i_0) &= \frac{i_0(1 - \exp(-\lambda t))}{\exp(-\lambda t)}, \\
Var(X(t) - i_0) &= \frac{i_0(1 - \exp(-\lambda t))}{\exp(-2\lambda t)}
\end{aligned}$$

$\Rightarrow$ exponentielles Wachstum:

$$\begin{aligned}
E(X(t)) &= i_0 \exp(\lambda t), \\
Var(X(t)) &= i_0 \exp(\lambda t)(\exp(\lambda t) - 1).
\end{aligned}$$

- Definition: Geburts- und Todesprozess

  Ein stochastischer Prozess $X = \{X(t), t \geq 0\}$ mit Zustandsraum $S \subseteq \mathbb{N}_0$ heißt Geburts- und Todesprozess $:\Leftrightarrow$ $X$ ist ein homogener, diskreter MP mit

$$\begin{aligned} p_{i,i+1}(h) &= \lambda_i h + o(h) && i \geq 0 \\ p_{i,i-1}(h) &= \mu_i h + o(h) && i \geq 1 \\ p_{ii}(h) &= 1 - (\lambda_i + \mu_i)h + o(h) && i \geq 0 \\ p_{ij}(h) &= o(h) && \mid i - j \mid \geq 2, \quad i, j \geq 0 \end{aligned}$$

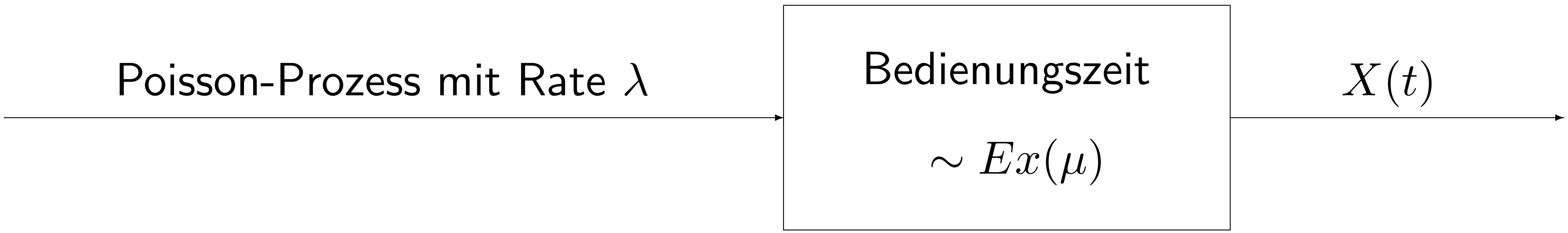
  mit

$$\begin{aligned} \mu_0 &= 0, \\ \lambda_0 &\geq 0 \quad (> 0 \text{ reflektierend}, = 0 \text{ absorbierend}), \\ \mu_i, \lambda_i &\geq 0 \quad \text{für } i \geq 1. \end{aligned}$$

Für die Intensitäten gilt also

$$
\begin{aligned}
\lambda_{i,i+1} &= \lambda_i\,,\\
\lambda_{i,i-1} &= \mu_i\,,\\
\lambda_{ij} &= 0 \text{ für } \mid i-j \mid \geq 2\,,\\
\lambda_{ii} &= -(\lambda_i+\mu_i)
\end{aligned}
$$

- Die explizite Berechnung der Übergangsmatrix $P(t)$ aus $\Lambda$ ist i.A. nicht möglich.
- Das Wartesystem M/M/1/$\infty$




Die Anzahl der im System befindlichen Kunden $X(t)$ ist ein spezieller Geburts-Todes-Prozess mit $\lambda_i = \lambda$ und $\mu_i = \mu$.

Für $\lambda < \mu$ existiert die Grenzverteilung $p_j$:

$$p_j = \lim_{t\to\infty} p_{ij}(t) = (1-r)r^j, \quad r = \frac{\lambda}{\mu}$$

$\Rightarrow \lim_{t\to\infty} X(t) = X(\infty)$ geometrisch verteilt mit

$$E(X(\infty)) = \frac{r}{1-r}, \quad Var(X(\infty)) = \frac{r}{(1-r)^2}$$

- Varianten:
  - begrenzte Kapazität $N$ des Warteraums (M/M/1/N)

$$\lambda_i = \begin{cases} \lambda & 0 \le i \le N \\ 0 & i > N \end{cases}$$

  - von der Länge der angetroffenen Warteschlange verursachte Entmutigung, z.B.

$$\lambda_i = \frac{\lambda}{(i+1)}, \; i = 0, 1, \ldots$$

- Das Wartesystem M/M/s/$\infty$: $s$ Schalter jeweils mit Rate $\mu$

  $X(t)$ Geburts-Todes-Prozess mit

$$\lambda_i = \lambda \quad \text{und} \quad \mu = \begin{cases} \mu i, & i = 1, \ldots, s \\ \mu s, & i = s+1, \ldots \end{cases}$$

  Grenzverteilung existiert für $\lambda < s\mu$.

- Modelle zur Bevölkerungsentwicklung

  Verallgemeinerung des Yule-Prozess mit Geburtsrate $\lambda_i = i\lambda$ und Sterberate $\mu_i = i\mu$

  Für $X(0) = 1$ kann man zeigen:

$$E(X(t)) = \exp((\lambda - \mu)t) \quad (\text{für } \lambda \neq \mu)$$

$$\text{Wahrscheinlichkeit des Aussterbens } = \begin{cases} 1, & \lambda \leq \mu \\ \frac{\mu}{\lambda}, & \lambda > \mu \end{cases}$$

- Struktur der Pfade von Geburts- und Todesprozessen

  Wahrscheinlichkeit für Verlassen des Zustands $i$ im Intervall $(t, t+h]$:

$$(\lambda_i + \mu_i)h + o(h).$$

  $\Rightarrow Exp(\lambda_i + \mu_i)$-verteilte Verweildauern (analog zum Poisson-Prozess).

  Wahrscheinlichkeit für Zustandsänderung $i \to i+1$: $\lambda_i h + o(h)$

  Wahrscheinlichkeit für Zustandsänderung $i \to i-1$: $\mu_i h + o(h)$

  Unter der Bedingung, dass eine Zustandsänderung stattfindet, erhält man also die bedingten Wahrscheinlichkeiten

$$\frac{\lambda_i}{(\lambda_i + \mu_i)} \qquad \text{Übergang } i \to i+1$$

$$\frac{\mu_i}{(\lambda_i + \mu_i)} \qquad \text{Übergang } i \to i-1$$

- Vorgehen zur Simulation:

  Der aktuelle Zustand sei $i$. Realisiere eine $Exp(\lambda_i + \mu_i)$-verteilte Zufallsvariable, um die Verweildauer in $i$ zu erhalten. Entscheide dann anhand eines Bernoulli-Experiments mit den obigen Wahrscheinlichkeiten, welcher der zwei möglichen Zustandswechsel eintritt.

# 4.3 Skizze der allgemeinen Theorie

- Voraussetzungen: $p_{ij}(0) = \lim_{h\downarrow 0} p_{ij}(h) = \delta_{ij}$, Pfade (mit W'keit 1) rechtsseitig stetig.

  Bezeichnungen:

  - $S_n$ Zeitpunkt der $n$-ten Zustandsänderung, $S_0 = 0$,
  - $Y_n = X(S_n)$ zum Zeitpunkt $S_n$ angenommener Zustand,
  - $T_{n+1} = S_{n+1} - S_n$ Verweildauer in $Y_n$,
  - $S_{n+1} = S_n + T_{n+1}$.
  - Falls absorbierende Zustände $Y_n$ möglich sind, d.h. $S_{n+1} = \infty$,

$$Y_{n+1} = \begin{cases} Y_n, & S_{n+1} = \infty \\ X(S_{n+1}), & S_{n+1} < \infty. \end{cases}$$

  - $V(t)$ Vorwärtsrekurrenzzeit: Dauer von $t$ an bis zur nächsten Zustandsänderung

- Exponentialverteilte Verweildauern:

  Die Verweildauern $T_i$, $i \in \mathbb{N}$ sind exponentialverteilt mit Parameter $\lambda_i$. Für jeden Zeitpunkt $t$ gilt

$$P(V(t) \leq s | X(t) = i) = 1 - e^{-\lambda_i s}, \quad s \geq 0,$$

  die Vorwärtsrekurrenzzeit ist also ebenfalls exponentialverteilt mit Parameter $\lambda_i$.

- Definition: Zustand $i$ ist

$$\begin{aligned} \text{absorbierend} \quad &:\Leftrightarrow \quad \lambda_i = 0 \\ \text{stabil} \quad &:\Leftrightarrow \quad 0 < \lambda_i < \infty \\ \text{instabil} \quad &:\Leftrightarrow \quad \lambda_i = \infty \end{aligned}$$

- Falls $i$ absorbierend ist, gilt für alle $s \geq 0$

$$P(V(t) > s \mid X(t) = i) = P(X(t+s) = i \mid X(t) = i) = 1.$$

Der Zustand $i$ wird mit W'keit 1 nicht mehr verlassen.

Falls $i$ stabil ist, gilt

$$P(0 < V(t) < \infty \mid X(t) = i) = 1.$$

$X$ verweilt in $i$ eine positive, aber endliche Zeit (mit W'keit 1).

Falls $i$ instabil ist, gilt

$$P(V(t) = 0 \mid X(t) = i) = 1.$$

Der MP verlässt einen instabilen Zustand sofort wieder. Dies führt zu nicht rechtsseitig stetigen Pfaden. Daher sind instabile Zustände im weiteren ausgeschlossen.

- Struktursatz: Sei $X$ ein Markov-Prozess mit rechtsseitig stetigen Pfaden. Dann gilt $\forall n \in \mathbb{N}_0\ \forall j, i, \ldots, i_0 \in S\ \forall t, t_1, \ldots, t_n \in \mathbb{R}\ \forall 0 < s_1 \ldots < s_n$

  1. $P(Y_{n+1} = j, T_{n+1} > t | Y_n = i, \ldots, Y_0 = i_0, S_n = s_n, \ldots, S_0 = 0) = q_{ij} e^{-\lambda_i t}$

     mit $q_{ij} \geq 0,\ \sum\limits_{j \in S} q_{ij} = 1, q_{ii} = \begin{cases} 0 & , i \text{ stabil} \\ 1 & , i \text{ absorbierend} \end{cases}$

  2. $\{Y_n, n \in \mathbb{N}\}$ ist eine Markov-Kette mit Übergangsmatrix $Q = (q_{ij})$

  3. $P(T_{n+1} > t | Y_n = i, Y_{n+1} = j) = P(T_{n+1} > t | Y_n = i) = e^{-\lambda_i t}$

  4. $P(T_1 > t_1, \ldots, T_n > t_n | Y_n = i, \ldots, Y_0 = i_0) = e^{-\lambda_{i_0} t_1} \ldots e^{-\lambda_{i_{n-1}} t_n}$

- Bemerkungen:

(a) $\{Y_n,\ n \in \mathbb{N}\}$ heißt eingebettete Markov-Kette. Erreichbarkeits- / Rekurrenzeigenschaften der eingebetteten Markov-Kette übertragen sich auf diskreten Markov-Prozess.

(b) Aussage 3 besagt, dass $\lambda_i$ nur vom eben besuchten Zustand, nicht aber vom nächsten abhängt.

(c) Aussage 4 besagt, dass die Verweildauern exponentialverteilt und bedingt unabhängig sind, wenn die Folge der besuchten Zusände gegeben ist.

(d) Der Struktursatz liefert auch die gemeinsame Verteilung von $\{T_n, Y_n\}$ und damit die Likelihood zur ML-Schätzung der Strukturparameter $\lambda_i$ und $q_{ij}$ (bzw. der Intensitäten $\lambda_{ij}$):

$$\begin{aligned}
&P(T_1 > t_1, \ldots, T_n > t_n, Y_0 = i_0, Y_1 = i_1, \ldots, Y_n = i_n) \\
&= \quad P(T_1 > t_1, \ldots, T_n > t_n \mid Y_0 = i_0, Y_1 = i_1, \ldots, Y_n = i_n) \\
&\qquad \cdot P(Y_0 = i_0, Y_1 = i_1, \ldots, Y_n = i_n) \\
&= \quad e^{-\lambda_0 t_1} \cdot \ldots \cdot e^{-\lambda_{i_{n-1}} t_n} \cdot P(Y_0 = i_0) \cdot q_{i_0 i_1} \cdot \ldots \cdot q_{i_{n-1} i_n}
\end{aligned}$$

- Konstruktion bzw. Simulation der Pfade eines Markov-Prozesses:
  - Startwert $X(\omega, 0) = Y_0(\omega) = i_0$
  - Im Zustand $Y_n(\omega) = i_n,\ n \in \mathbb{N}_0$, bestimmt man gemäß der Verteilung $q_{i_n j},\ j \in S$, den nächsten Zustand $Y_{n+1}(\omega) = i_{n+1}$.
  - Die Verweildauer $T_{n+1}(\omega) = t_{n+1}$ in $i_n$ bis zum Sprung nach $i_{n+1}$ wird gemäß einer $\lambda_{i_n}$-Exponentialverteilung realisiert.

- Definition: Regulärer Markov-Prozess

  Ein Markov-Prozess heißt regulär, wenn er die folgenden Eigenschaften besitzt:

  1. $p_{ij}(0) = \lim\limits_{h \downarrow 0} p_{ij}(h) = \delta_{ij} = \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases}$
  2. Pfade (mit Wahrscheinlichkeit 1) rechtsseitig stetig.
  3. $\sup_n S_n = +\infty$.

- Regularitätskriterien

  Ein Markov-Prozess, der die Eigenschaften 1. und 2. erfüllt ist regulär, falls eine der folgenden Eigenschaften gilt:

  1. $S$ ist endlich.
  2. $\lambda_i \leq c$ für alle $i \in S$.
  3. Alle Zustände der eingebetteten Markov-Kette $Y$ sind rekurrent.
  4. Die Markov-Kette $Y$ verbleibt mit Wahrscheinlichkeit $0$ in einer transienten Klasse.

- Intensitäten und Strukturparameter

  $X$ sei ein regulärer MP. Dann sind die Übergangswahrscheinlichkeiten $p_{ij}(t)$ stetig differenzierbar und es gilt

$$p'_{ij}(0) = \lim_{h \downarrow 0} \frac{p_{ij}(h) - p_{ij}(0)}{h} = \lambda_{ij} = \begin{cases} -\lambda_i, & i = j \\ \lambda_i q_{ij}, & i \neq j, \end{cases}$$

bzw. in $o(h)$-Schreibweise

$$
\begin{aligned}
P(X(t+h)=j|X(t)=i) &= h\lambda_i q_{ij} + o(h), \quad i \neq j \\
P(X(t+h)=i|X(t)=i) &= 1-\lambda_i h + o(h).
\end{aligned}
$$

In Matrixschreibweise ($S$ endlich):

$$P(h) = I + \Lambda h + o(h).$$

bzw.

$$\lim_{h\to\infty} \frac{P(h)-I}{h} = P'(0) = \Lambda.$$

Die Zeilensummen der Intensitätsmatrix $\Lambda$ sind gleich $0$, d.h. für alle $i \in S$ gilt

$$\lambda_{ii} = -\lambda_i = -\sum_{j\in S, j\neq i} \lambda_{ij}.$$

- Bestimmung der Übergangswahrscheinlichkeiten $P(t)$ aus $\Lambda$

  Sei $X$ ein regulärer Markov-Prozess. Dann ist die Übergangsmatrix $P(t) = (p_{ij}(t))$ durch die Übergangsmatrix $Q$ der eingebetteten Markov-Kette $Y$ und die Parameter $\lambda_i$ der Verweildauern bzw. die Intensitätsmatrix $\Lambda$ eindeutig bestimmt. Die Übergangswahrscheinlichkeiten $p_{ij}(t)$ sind Lösung von

  1. $p'_{ij}(t) = \sum_{k \in S} \lambda_{ik} p_{kj}(t)$ bzw. $P'(t) = \Lambda P(t),\ P(0) = I$

     (Rückwärtsgleichungen von Kolmogorov)

  2. $p'_{ij}(t) = \sum_{k \in S} p_{ik}(t) \lambda_{kj}$ bzw. $P'(t) = P(t)\Lambda,\ P(0) = I$

     (Vorwärtsgleichungen von Kolmogorov)

  3. Für endliches $S$ ist die Lösung gegeben durch

$$P(t) = e^{t\Lambda} = I + t\Lambda + \frac{(t\Lambda)^2}{2!} + \frac{(t\Lambda)^3}{3!} + \dots$$

Die (numerische) Berechnung erfolgt mit Hilfe der Spektralzerlegung von $\Lambda$

$$\Lambda = U \text{ diag}(d_1, \ldots, d_m) U^{-1},$$

mit den Eigenwerten $d_1, \ldots, d_m$ und der Matrix $U$ der Eigenvektoren. Dann gilt

$$P(t) = U \text{ diag}(e^{d_1 t}, \ldots, e^{d_m t}) U^{-1}.$$

- Beispiel: Markov-Prozess mit 2 Zuständen

$$\Lambda = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}$$

Eigenwertzerlegung von $\Lambda$:

$$D = \begin{pmatrix} 0 & 0 \\ 0 & -3 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \quad U^{-1} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}.$$

$$
\begin{aligned}
P(t) &= U\ \mathsf{diag}(e^{d_1 t}, e^{d_2 t})U^{-1} \\
&= \ldots \\
&= \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} + e^{-3t} \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{pmatrix}
\end{aligned}
$$

Bemerkung: Es existiert der Grenzwert

$$
\lim_{t\to\infty} P(t) = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}.
$$

- Beispiel: Allgemeiner Markov-Prozess mit 2 Zuständen

$$
\Lambda = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.
$$

Die Kolmogorov-Gleichung $P'(t) = P(t)\Lambda$ ergibt

$$\begin{pmatrix} p'_{00}(t) & p'_{01}(t) \\ p'_{10}(t) & p'_{11}(t) \end{pmatrix} = \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix} \cdot \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

$$\begin{aligned} &\Rightarrow \quad p'_{00}(t) = -\lambda p_{00}(t) + \mu p_{01}(t), \qquad p_{01}(t) = 1 - p_{00}(t), \\ &\Rightarrow \quad p'_{00}(t) = +\mu - (\lambda + \mu) p_{00}(t). \end{aligned}$$

Lösung:

$$\begin{aligned} p_{00}(t) &= \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} \\ p_{01}(t) &= 1 - p_{00}(t) \\ &= \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} \end{aligned}$$

Aus Symmetriegründen erhält man

$$\begin{aligned} p_{11}(t) &= \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} \\ p_{10}(t) &= \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} \end{aligned}$$

Grenzwert:

$$\lim_{t\to\infty} P(t) = \begin{pmatrix} \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \\ \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \end{pmatrix}.$$

- Die Zustände eines Markov-Prozesses werden anhand der eingebetteten Markov-Kette klassifiziert (Rekurrenz, Erreichbarkeit, Periodizität, etc.).
- Für transientes $j$ folgt aus $\lim_{t\to\infty} q_{ij}^{(t)} = 0$ für die $t$-Schritt ÜW $q_{ij}^{(t)}$ auch $\lim_{t\to\infty} p_{ij}(t) = 0$. Für Grenzwertsätze daher Beschränkung auf irreduzible, rekurrente MP.

- Grenzwertsatz:

  Sei $X$ ein regulärer, positiv-rekurrenter und irreduzibler Markov-Prozess. Dann gilt:

  1. $\lim\limits_{t \to \infty} P(X(t) = j) = \pi_j$ existiert $\forall\ j \in S$, und ist von der Anfangsverteilung unabhängig.

  2. Die Grenzverteilung $\pi$ läßt sich berechnen über

     (a) $\pi_j = \dfrac{\mu_j}{\sum_i \mu_i}$ mit $\mu\Lambda = 0$ bzw.

     (b) $\pi_j = \dfrac{\nu_j/\lambda_j}{\sum_i \nu_i/\lambda_i}$ mit $\nu = \nu Q$.

  3. $\lim\limits_{t \to \infty} P(t) = \begin{pmatrix} \pi \\ \vdots \\ \pi \\ \vdots \\ \pi \end{pmatrix}$

  4. $\pi = (..., \pi_j, ...)$ ist strikt positiv.

5. $\pi$ ist die einzige stationäre Verteilung von $X$, d.h. es gilt

$$\pi = \pi \cdot P(t) \quad \forall\, t \geq 0\,.$$

- Beispiel: Markov-Prozess mit zwei Zuständen

$$\Lambda = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$$

Möglichkeit 1: Über die Intensitätsmatrix $\mu\Lambda = 0$

$$\begin{aligned} -\mu_1\lambda_1 + \mu_2\lambda_2 = 0 \quad &\Leftrightarrow \quad -\mu_1\lambda_1 + (1-\mu_1)\lambda_2 = 0 \\ \mu_1\lambda_1 - \mu_2\lambda_2 = 0 \quad &\Leftrightarrow \quad \lambda_2 - \mu_1(\lambda_1+\lambda_2) = 0 \end{aligned}$$

$$\Rightarrow \mu_1 = \frac{\lambda_2}{\lambda_1+\lambda_2}, \quad \mu_2 = \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

Möglichkeit 2: Über die eingebettete Markov-Kette $\nu = \nu Q$

$$Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

stationäre Verteilung:

$$\nu = \left(\frac{1}{2}, \frac{1}{2}\right).$$

Grenzverteilung ist proportional zu $\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}\right)$:

$$(\pi_1, \pi_2) = \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right).$$

- Peter Prinzip: Drei berufliche Kategorien in einem Betrieb: $T$ Trainee, $J$ Junior, $S$ Senior

$$\Lambda = \begin{array}{c} \\ T \\ J \\ S \end{array} \begin{array}{c} \begin{array}{ccc} T & J & S \end{array} \\ \begin{pmatrix} -\lambda_T & \lambda_T & 0 \\ \lambda_{JT} & -\lambda_J & \lambda_{JS} \\ \lambda_S & 0 & -\lambda_S \end{pmatrix} \end{array}.$$

Grenzverteilung: $\pi = (\pi_T, \pi_J, \pi_S)$

$$\begin{aligned} \lambda_T \pi_T &= \lambda_{JT}\pi_J + \lambda_S \pi_S \\ \lambda_J \pi_J &= \lambda_T \pi_T \\ \lambda_S \pi_S &= \lambda_{JS} \pi_J \\ 1 &= \pi_T + \pi_J + \pi_S. \end{aligned}$$

Lösung ist

$$\pi_T = \frac{\lambda_S \lambda_J}{N}, \quad \pi_J = \frac{\lambda_S \lambda_T}{N}, \quad \pi_S = \frac{\lambda_T \lambda_{JS}}{N}.$$

mit $N = \lambda_S\lambda_J + \lambda_S\lambda_T + \lambda_T\lambda_{JS}$.

# 4.4 Inferenz

- Zwei mögliche Situationen in Anwendungen:
  1. Vollständige Kenntnis über einen oder mehrere Pfade.
  2. Beobachtungen nur zu Zeitpunkten $t_1 < t_2 < \ldots < t_n$, d.h. Pfade nicht vollständig beobachtet.
- In beiden Situationen: Likelihood–basierte Inferenz für die Strukturparameter $\Lambda = \{\lambda_{ij}\}$ bzw. $Q = \{q_{ij}\}$, $\lambda = \{\lambda_i\}$.
- Situation 1:
  Die Likelihood wird gemäß Struktursatz bestimmt:

$$P(Y_n = i_n, T_n > t_n | Y_{n-1} = i_{n-1}, S_{n-1} = s_{n-1}, \ldots, Y_0 = i_0, S_0 = 0) = q_{i_{n-1}i_n} e^{-\lambda_{i_{n-1}} t_n}.$$

$\Rightarrow$ Gemeinsame (bedingte) Dichte für $Y_n$ diskret, $T_n$ stetig (bezüglich Produktmaß aus Zählmaß und Lebesgue-Maß):

$$q_{i_{n-1}i_n} \lambda_{i_{n-1}} e^{-\lambda_{i_{n-1}} t_n}.$$

Likelihood $L(\lambda_{ij}|i_0, i_1, \ldots, i_{n-1}, i_n, t_1, \ldots, t_n)$:

$$
\begin{aligned}
L(\lambda_{ij}) &= p_{i_0}(0) \cdot q_{i_0 i_1} \lambda_{i_0} e^{-\lambda_{i_0} t_1} \cdot \ldots \cdot q_{ij} \lambda_i e^{-\lambda_i t_j} \cdot \ldots \cdot q_{i_{n-1} i_n} \lambda_{i_{n-1}} e^{-\lambda_{i_{n-1}} t_n} \\
&= p_{i_0}(0) \prod_{i \in S} \{\exp(-\lambda_i \gamma_i) \cdot \prod_{j \in S, j \neq i} (q_{ij} \lambda_i)^{n_{ij}}\}
\end{aligned}
$$

mit $\gamma_i$ gesamte Verweildauer in $i$ und $n_{ij}$ Gesamtzahl der Übergänge $i \to j$.

Log-Likelihood (ohne $p_{i_0}(0)$):

$$
\begin{aligned}
l(\lambda_{ij}) &= \sum_{i \in S} -\lambda_i \gamma_i + \sum_{i,j \in S, i \neq j} n_{ij} \log(\lambda_{ij}) \\
&= \sum_{i,j; i \neq j} (-\lambda_{ij} \gamma_i + n_{ij} \log(\lambda_{ij})) \\
\frac{\partial l(\lambda_{ij})}{\partial \lambda_{ij}} &= -\gamma_i + \frac{n_{ij}}{\lambda_{ij}} \stackrel{!}{=} 0
\end{aligned}
$$

Damit ergeben sich die ML-Schätzer

$$\begin{aligned}
\hat{\lambda}_{ij} &= \frac{n_{ij}}{\gamma_i} \\
\hat{\lambda}_i &= \frac{n_i}{\gamma_i}, \text{ mit } n_i = \sum_j n_{ij} \\
\hat{q}_{ij} &= \frac{\hat{\lambda}_{ij}}{\hat{\lambda}_i} = \frac{n_{ij}/\gamma_i}{n_i/\gamma_i} = \frac{n_{ij}}{n_i}
\end{aligned}$$

$\hat{\lambda}_{ij}$ ist konsistent für $\lambda_{ij}$ mit asymptotischer Varianz $\frac{n_{ij}}{\gamma_i^2}$.

- Situation 2: Daten $X(t_1) = i_1, \ldots, X(t_n) = i_n \Rightarrow$ Likelihood

$$L(\lambda_{ij}|X(t_1) = i_1, \ldots, X(t_n) = i_n) = p_{i_0}(0) p_{i_1 i_2}(t_2 - t_1) \cdot \ldots \cdot p_{i_{n-1} i_n}(t_n - t_{n-1})$$

Die Übergangswahrscheinlichkeiten müssen dabei aus den Kolmogorov-Gleichungen bzw. über die Spektralzerlegung von $\Lambda$ bestimmt werden.

$\Rightarrow$ Numerisch aufwendig.