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_n, $n \in \mathbb{N}$ sind exponentialverteilt mit einem Parameter λ_i, gegeben $Y_{n-1} = i$. 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 λ_i.

  Beweis: (für V(t))

  Aus der Homogenität des MP folgt

  $f_i(s) := P(V(t) > s \mid X(t) = i) = P(V(0) > s \mid X(0) = i).$

für alle $r, s \geq 0$:

$$\begin{aligned}
f_i(s+r) &= P(V(0) > s+r \mid X(0)=i) \\
&= P(V(0) > s, V(s) > r \mid X(0)=i) \\
&= P(V(s) > r \mid \underbrace{V(0) > s, X(0)=i}_{\Leftrightarrow X(s)=i, X(0)=i}) \cdot P(V(0) > s \mid X(0)=i) \\
&\overset{\text{ME}}{=} P(V(s) > r \mid X(s)=i) \cdot P(V(0) > s \mid X(0)=i) \\
&\overset{\text{Homogenität}}{=} f_i(r) \cdot f_i(s).
\end{aligned}$$

$f_i(s+r) = f_i(r) \cdot f_i(s)$ charakterisiert eindeutig die Exponentialfunktion.

Wegen $0 \leq f_i(s) \leq 1$ folgt $f_i(s) = e^{-\lambda_i s}$ mit $\lambda_i \geq 0 \Rightarrow$ Behauptung.

- 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 $q_{ij} \geq 0, \; \sum_{j \in S} q_{ij} = 1, q_{ii} = \begin{cases} 0 & , i \text{ stabil} \\ 1 & , i \text{ absorbierend} \end{cases}$

  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 λ_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 λ_i und q_{ij} (bzw. der Intensitäten λ_{ij}):

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

Beweis:

1. Es gilt $Y_{n+1} = X(S_{n+1}) = X(S_n + T_{n+1})$

$$\begin{aligned}
\Rightarrow\ & P(Y_{n+1} = j, T_{n+1} > t \mid Y_n = i, \ldots, Y_0 = i_0, S_n = s_n \ldots, S_0 = 0) \\
= \quad & P(X(S_n + T_{n+1}) = j, T_{n+1} > t \mid X(s_n) = i, \ldots, X(0) = i_0, \\
& \qquad\qquad S_n = s_n, \ldots, S_0 = 0) \\
\overset{\text{ME+Hom.}}{=} \quad & P(X(T_1) = j, T_1 > t \mid X(0) = i) \\
= \quad & \underbrace{P(T_1 > t \mid X(0) = i)}_{e^{-\lambda_i t}\text{, Satz zu expon.vtlt. Verweildauern}} P(X(T_1) = j \mid X(0) = i, T_1 > t) \\
\overset{\text{ME}}{=} \quad & e^{-\lambda_i t} \cdot P\{X(t + V(t)) = j \mid X(t) = i\} \\
\overset{\text{Hom.}}{=} \quad & e^{-\lambda_i t} \cdot P\{Y_1 = j \mid Y_0 = i\} \\
= \quad & e^{-\lambda_i t} q_{ij}
\end{aligned}$$

2. Setze $t = 0$ in 1.

3. $$P(T_{n+1} > t \mid Y_n = i, Y_{n+1} = j) = \frac{P(T_{n+1} > t, Y_{n+1} = j \mid Y_n = i)}{P(Y_{n+1} = j \mid Y_n = i)}$$
$$\overset{1.,2.}{=} \frac{q_{ij} e^{-\lambda_i t}}{q_{ij}} = e^{-\lambda_i t}$$

4. Aus 3. durch Induktion.

- 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 λ_{i_n}-Exponentialverteilung realisiert.

- Definition: Regulärer Markov-Prozess

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

1. $p_{ij}(0) = \lim_{h \downarrow 0} p_{ij}(h) = \delta_{ij} = \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases}$
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 Λ sind gleich 0, d.h. für alle $i \in S$ gilt

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

Beweis der Hauptaussage:

Die Wahrscheinlichkeit für mehr als einen Übergang im Zeitraum h ist gleich $o(h)$,

$$\begin{aligned} P(T_{n+1}+T_{n+2} \le h | Y_n = i, Y_{n+1} = j) &\le P(T_{n+1} \le h, T_{n+2} \le h | Y_n = i, Y_{n+1} = j) \\ &= (1-e^{-\lambda_i h})(1-e^{-\lambda_j h}) = (\lambda_i h + o(h))(\lambda_j h + o(h)) = o(h). \end{aligned}$$

$\Rightarrow$ für $i \neq j$:

$$\begin{aligned} p_{ij}(h) &= P(Y_{n+1} = j, T_{n+1} \le h \mid Y_n = i) + \underbrace{o(h)}_{=P(Y_{n+2}=j, T_{n+1}+T_{n+2} \le h | Y_n = i)+\ldots} \\ &\overset{\text{Struktursatz}}{=} q_{ij} - q_{ij} \underbrace{e^{-\lambda_i h}}_{\substack{=1-\lambda_i h + o(h), \\ \text{Taylorentwicklung um } 0}} + o(h) \\ &= q_{ij}\lambda_i h + o(h) \end{aligned}$$

$$\Rightarrow \lambda_{ij} = p'_{ij}(0) = \lim_{h \downarrow 0} \frac{p_{ij}(h) - p_{ij}(0)}{h} = \lim_{h \downarrow 0} \left[q_{ij}\lambda_i + \frac{o(h)}{h} \right] = q_{ij}\lambda_i.$$

Analog:

$$\begin{aligned} p_{ii}(h) &= P(V(t) > h \mid X(t) = i) + o(h) \\ &= e^{-\lambda_i h} + o(h) \overset{\text{Taylor}}{=} 1 - \lambda_i h + o(h) \end{aligned}$$

$$\Rightarrow\ p'_{ii}(0) = \lambda_{ii} = \lim_{h \downarrow 0} \frac{p_{ii}(h) - p_{ii}(0)}{h} = \lim_{h \downarrow 0} \left[\frac{1 - \lambda_i h - 1}{h} + \frac{o(h)}{h} \right] = -\lambda_i.$$

Zeilensumme von Λ ist gleich 0: $q_{ii} = \begin{cases} 0, i \text{ stabil } (0 < \lambda_i < \infty) \\ 1, i \text{ absorbierend } (\lambda_i = 0) \end{cases}$

$$\begin{aligned} \sum_{j \in S} \lambda_{ij} &= \lambda_{ii} + \sum_{j \in S \setminus \{i\}} \lambda_{ij} = -\lambda_i + \sum_{j \neq i} \lambda_i q_{ij} \\ &= \underbrace{\lambda_i}_{\substack{= 0, \\ i \text{ absorbierend}}} \underbrace{(\sum_{j \neq i} q_{ij} - 1)}_{=0, i \text{ stabil}} = 0. \end{aligned}$$

Beispiel: Molekulare Evolution $S = \{A, T, G, C\}$

Jukes-Cantor-Modell: alle Mutationen gleich "wahrscheinlich"

$$\Lambda = \begin{array}{c} \\ A \\ G \\ C \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & G & C & T \end{array} \\ \begin{pmatrix} -3\alpha & \alpha & \alpha & \alpha \\ \alpha & -3\alpha & \alpha & \alpha \\ \alpha & \alpha & -3\alpha & \alpha \\ \alpha & \alpha & \alpha & -3\alpha \end{pmatrix} \end{array} \xrightarrow[i \text{ stabil } (\lambda_i > 0)]{} Q = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

$$\Rightarrow q_{ii} = 0,\ q_{ij} = \frac{\lambda_{ij}}{\lambda_i} = \frac{\lambda_{ij}}{\sum_{k \neq i} \lambda_{ik}},\ j \neq i$$

Kimura-Modell: Unterschiedliche Intensitäten, aber $A \to C = C \to A$ (Symmetrie)

$$\Lambda = \begin{array}{c} \\ A \\ G \\ C \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & G & C & T \end{array} \\ \begin{pmatrix} * & \alpha & \beta & \gamma \\ \alpha & * & \gamma & \beta \\ \beta & \gamma & * & \alpha \\ \gamma & \beta & \alpha & * \end{pmatrix} \end{array} \qquad \begin{aligned} * &= -(\alpha + \beta + \gamma) \\ \theta &= (\alpha, \beta, \gamma) \end{aligned}$$

Beispiel: Geburtsprozess

- Bestimmung der Übergangswahrscheinlichkeiten $P(t)$ aus Λ

  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 λ_i der Verweildauern bzw. die Intensitätsmatrix Λ 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 Λ

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

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

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

Beweis:

Informell, in Matrixdarstellung (vgl. FKO S. 113). Die Beziehung zwischen $p'_{ij}(0)$ und λ_{ij} in Satz lässt sich in Matrizenform durch

$$\lim_{h \downarrow 0} \frac{P(h) - I}{h} = \Lambda$$

darstellen. Dann folgt mit Hilfe der Chapman-Kolmogorov-Gleichungen

$$\frac{P(t+h)-P(t)}{h} = \frac{P(t)P(h)-P(t)}{h} = \frac{P(t)(P(h)-I)}{h}$$

bzw.

$$\frac{P(t+h)-P(t)}{h} = \frac{(P(h)-I)P(t)}{h}.$$

Grenzübergang $h \to 0$ liefert

$$P'(t) = P(t)\Lambda \quad \text{bzw.} \quad P'(t) = \Lambda P(t)$$

mit $P(0) = I$ wegen $p_{ij}(0) = \delta_{ij}$.
Die Lösung ergibt sich aus der Theorie linearer Differentialgleichungssysteme analog zu $f'(t) = \lambda f(t)$ mit $f(0) = 1 \Rightarrow f(t) = e^{\lambda t}$. □

- Beispiel: Markov-Prozess mit 2 Zuständen

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

Eigenwertzerlegung von Λ:

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

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

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

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

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

$$\Rightarrow \quad p'_{00}(t) = +\mu - (\lambda + \mu) p_{00}(t).$$

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. $\lim_{t\to\infty} P(X(t) = j) = \pi_j$ existiert $\forall\ j \in S$, und ist von der Anfangsverteilung unabhängig.
  2. Die Grenzverteilung π 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. $\lim_{t\to\infty} P(t) = \begin{pmatrix} \pi \\ \vdots \\ \pi \\ \vdots \\ \pi \end{pmatrix}$
  4. $\pi = (..., \pi_j, ...)$ ist strikt positiv, d.h. $\pi_j > 0$ für alle j.

5. π ist die einzige stationäre Verteilung von X, d.h. es gilt

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

Plausibilitätsbegründung zu 2.: (vgl. auch FKO S. 116)

Kolmogorov'sche Vorwärtsgleichung:

$$P'(t) = P(t) \cdot \Lambda$$

Existiert $\lim_{t\to\infty} P(t)$ (Aussage 3)

$$\lim_{t\to\infty} P(t) = \begin{pmatrix} - - -\pi - -- \\ - - -\pi - -- \end{pmatrix},$$

dann folgt offensichtlich

$$\lim_{t\to\infty} P'(t) = 0$$

da

$$\frac{d}{dt}\pi_j = \frac{d}{dt}\lim_{t\to\infty} p_{ij}(t) = 0$$

(π_j hängt ja nicht von t ab, da Grenzverteilung).

Insgesamt folgt

$$0 = \lim_{t\to\infty} P'(t) = \lim_{t\to\infty} P(t)\Lambda = \begin{pmatrix} - - - \pi - - - \\ - - - \pi - - - \end{pmatrix} \cdot \Lambda,$$

in jeder Zeile steht also $\pi \cdot \Lambda = 0$, d.h. 2. (a) gilt.

Die Äquivalenz zu 2(b) gilt wegen $\nu = \nu \cdot Q \Leftrightarrow \nu(Q - I) = 0$.

Erweiterung mit $\frac{1}{\lambda_i} \cdot \lambda_i$: $\nu\text{diag}(\frac{\lambda_i}{\lambda_i})(Q - I) = 0 \Leftrightarrow$

$$\underbrace{(\frac{\nu_1}{\lambda_1}, \frac{\nu_2}{\lambda_2}, \ldots)}_{\mu} \cdot \underbrace{\begin{pmatrix} -\lambda_1 & \lambda_1 q_{12} & \cdots \\ \lambda_2 q_{21} & -\lambda_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}}_{\Lambda} = 0$$

2(b) ist *plausibel*: Grenzverteilung ist proportional zu

$$\underbrace{\left(\frac{1}{\lambda_j}\right)}_{\text{durchschnittliche Aufenthaltsdauer in j}} \cdot \underbrace{\nu_j}_{\text{Wahrscheinlichkeit sich in j aufzuhalten (für großes } t)}$$

- Beispiel: Markov-Prozess mit zwei Zuständen

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

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

$$-\mu_1\lambda_1 + \mu_2\lambda_2 = 0 \quad \Leftrightarrow \quad -\mu_1\lambda_1 + (1-\mu_1)\lambda_2 = 0$$
$$\mu_1\lambda_1 - \mu_2\lambda_2 = 0 \quad \Leftrightarrow \quad \lambda_2 - \mu_1(\lambda_1 + \lambda_2) = 0$$

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

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

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

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

Bedingte 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

$$\begin{aligned} \hat{\lambda}_{ij} &= \frac{n_{ij}}{\gamma_i} \\ \hat{\lambda}_i &= \frac{n_i}{\gamma_i}, \text{ mit } n_i = \sum_{j \neq i} n_{ij} \\ \hat{q}_{ij} &= \frac{\hat{\lambda}_{ij}}{\hat{\lambda}_i} = \frac{n_{ij}/\gamma_i}{n_i/\gamma_i} = \frac{n_{ij}}{n_i}, \ j \neq i. \end{aligned}$$

$\hat{\lambda}_{ij}$ ist konsistent für λ_{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 Λ bestimmt werden.

$\Rightarrow$ Numerisch aufwendig.