Q5. Markov-Prozesse in kontinuierlicher Zeit

Gliederung

1. Zeitkontinuierliche Markov-Prozesse
2. Vom Modell zum Markov-Prozess
3. Stationäre Analyse
4. Transiente Analyse
5. Stationäre Analyse von Modellen mit unendlichem Zustandsraum

Literatur (primär)

- C. G. Cassandras, S. Lafortune
Introduction to Discrete Event Systems. Springer 2008
Kap. 7.3, 7.4
- W. J. Stewart
Numerical Solution of Markov Chains. Princeton University Press 1994, Kap. 1, 5
- P. Buchholz et al
Quantitative Systemanalyse mit Markovschen Ketten
Teubner 1994, Kap. 2, 4, 5, 6, 8

Ziele:

- Zeitkontinuierliche Markov-Prozessen definieren können
- Modellierungsmächtigkeit zeitkontinuierlicher Markov-Prozesse kennen lernen
- Übergang vom Modell zur Markov-Prozesse kennen lernen
- Systeme mit Hilfe von Markov-Prozessen bewerten können
- Algorithmen zur Analyse von zeitkontinuierlichen Markov-Prozessen kennen lernen
- Das M/M/1-System einordnen und handhaben können
- Ideen zur Analyse von strukturierten unendlichen Systemen kennen lernen

Q5.1 Zeitkontinuierliche Markov-Prozesse

Wir betrachten Zeit $T=\mathbb{R}_+$ und
endliche/abzählbare Zustandsräume (bei Bedarf $X \subseteq \mathbb{N}$)

Schreibweise $X(t)$ Zustand zum Zeitpunkt $t \in \mathbb{R}_+$

Für einen zeitkontinuierlichen Markov-Prozess gilt

$P[X(t_{k+1})=x_{k+1}|X(t_k)=x_k,\ldots,X(t_0)=x_0] = P[X(t_{k+1})=x_{k+1}|X(t_k)=x_k]$

Für $t_{k+1} \geq t_k \geq \ldots \geq t_0$

d.h. Zukunft hängt nur vom aktuellen Zustand ab!

Beide Markov Bedingungen sind erfüllt:

M1: in der Vergangenheit eingenommene Zustände sind irrelevant

M2: Verweilzeit im aktuellen Zustand ist irrelevant

M2 bedeutet, dass die Verteilung der Restverweilzeit im Zustand unabhängig von der bisherigen Verweilzeit ist (Gedächtnislosigkeit!)

Einschub:

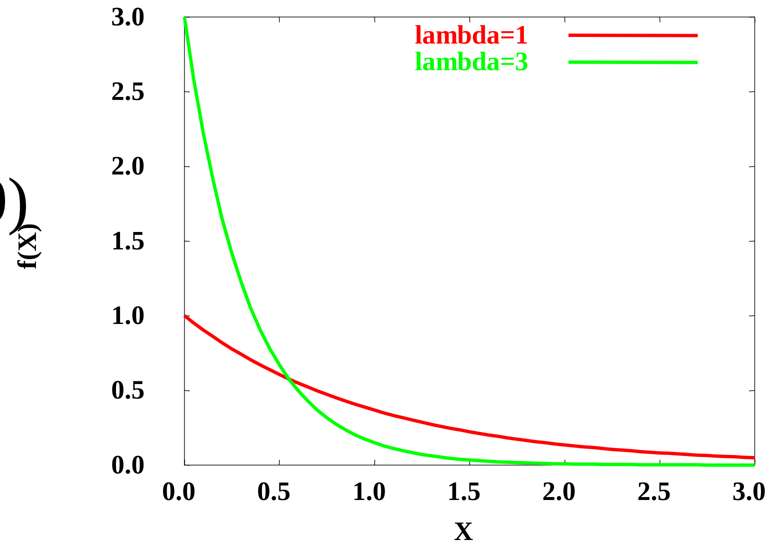
Exponentialverteilung (Parameter $\lambda>0$)

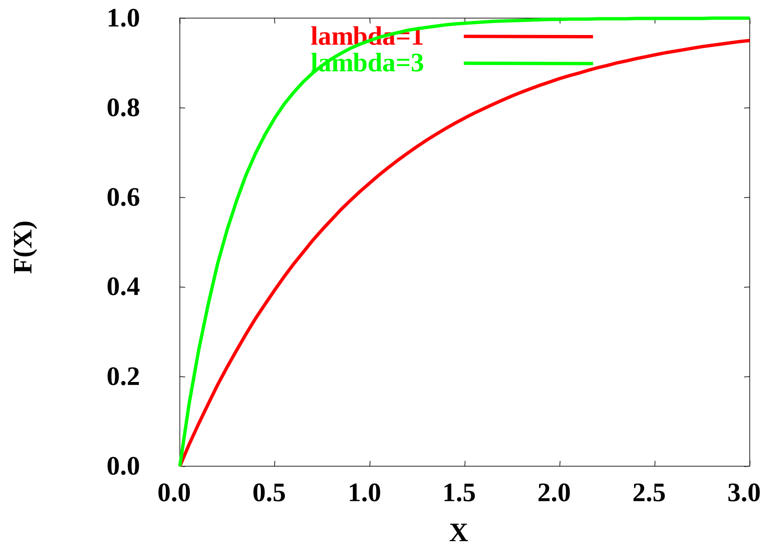
$$f(x)=\begin{cases}\lambda\cdot e^{-\lambda x} & \text{falls}\; x\geq 0\\ 0 & \text{sonst}\end{cases}$$

$$F(x)=\begin{cases}1-e^{-\lambda x} & \text{falls}\; x\geq 0\\ 0 & \text{sonst}\end{cases}$$

- $E(X)=1/\lambda$
- $\sigma^2(X)=1/\lambda^2$

$\Rightarrow VK(X)=1$





Gedächtnislosigkeitseigenschaft der Exponentialverteilung:

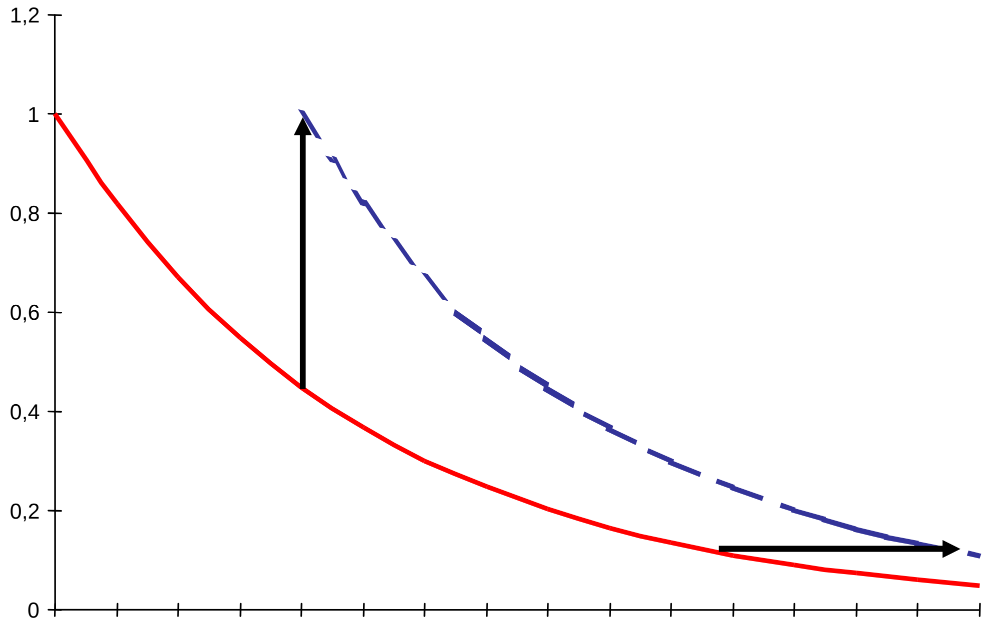
Falls eine ZV X exponentialverteilt ist, so gilt

$P[X > t + s \mid X > t] = P[X > s]$, da

$P[X > t + s \mid X > t] = e^{-\lambda(s+t)} / e^{-\lambda t} = e^{-\lambda s} = P[X > s]$

Dies hat zur Folge, dass die Restzeit einer Exponentialverteilung immer exponentialverteilt mit einer konstanten Rate ist

Die Exponentialverteilung ist die einzige gedächtnislose kontinuierliche Verteilung



Weitere Eigenschaften der Exponentialverteilung:

- Additivität: Das Minimum zweier Exponential-Verteilungen mit Raten λ und μ ist exponentiell-verteilt mit Rate ($\lambda + \mu$)
(Offensichtlich kann dies auf eine beliebige Anzahl von Exponential-Verteilungen erweitert werden)

- Ausdünnung: Wenn Ereignisse mit exponentiell verteilten Eintrittszeiten mit Rate λ durch eine Zufallsentscheidung nur mit Wahrscheinlichkeit p akzeptiert werden, so ist die Zeit zwischen zwei akzeptierten Ereignissen exponentiell verteilt mit Rate $p\lambda$

- Falls Ereignisse mit einer exponentialverteilten Generierungszeit mit Rate λ erzeugt werden, so entspricht die Anzahl Ereignisse, die in einem Intervall [0,t] erzeugt werden, einem Poisson-Prozess mit Rate λt.
D.h. P[k Ereignisse in einem Intervall der Länge t] = $e^{-\lambda t} (\lambda t)^k/k!$

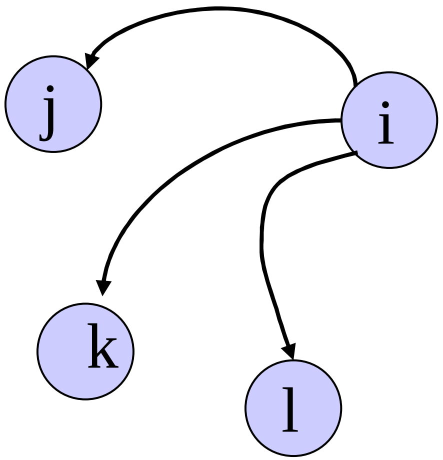
Zurück zu zeitkontinuierlichen Markov-Prozessen:

$p_{ij}(s,t) = P[X(t) = j \mid X(s) = i]$

Wir beschränken uns auf homogene Markov-Prozesse, d.h.:

$p_{ij}(\tau) = P[X(s+\tau) = j \mid X(s) = i]$ für alle $s \geq 0$

Offensichtlich muss $\Sigma_j\, p_{ij}(\tau) = 1$ für alle τ gelten



Verhalten ausgehen von $X_t = i$:

- Im Gegensatz zum zeitdiskreten Fall ist die Anzahl der Transitionen im Intervall $[t,t+\Delta]$ nicht a priori beschränkt
- aber $\lim_{\Delta\to 0} p_{ij}(\Delta) = 0$ für $i \neq j$, falls die Verweilzeit in einem Zustand durch eine Exponentialverteilung mit endlicher Rate beschrieben ist

$\Rightarrow \lim_{\Delta\to 0} p_{ii}(\Delta) = 1$ (Summe der Wahrscheinlichkeiten)

Ziel: Untersuchung einzelner Transitionen

o(Δ) definiert Funktion für die gilt $\lim_{\Delta\to 0} (o(\Delta)/\Delta) = 0$

Transitionsrate ist definiert als $$q_{ij} = \lim_{\Delta\to\infty} \frac{p_{ij}(\Delta)}{\Delta} \text{ für } i \neq j$$

Damit gilt auch $p_{ij}(\Delta) = q_{ij}\Delta + o(\Delta)$

Da die Summe der Wahrscheinlichkeiten für jedes Δ 1 ergeben muss, gilt

$$1 - p_{ii}(\Delta) = \sum_{i\neq j} p_{ij}(\Delta) = \sum_{i\neq j} q_{ij}\Delta + o(\Delta)$$

Bestimmung von q_{ii}:

$$q_{ii} = \lim_{\Delta\to 0} \frac{p_{ii}(\Delta) - 1}{\Delta} = \lim_{\Delta\to 0} \frac{-\sum_{j\neq i} q_{ij}\Delta + o(\Delta)}{\Delta} \quad \Rightarrow \quad q_{ii} = -\sum_{j\neq i} q_{ij} \leq 0$$

Transitionsraten q_{ij} werden einer Matrix **Q** zusammengefasst

Eigenschaften von **Q**:

- Zeilensumme 0
- Nicht negative Elemente außerhalb der Diagonalen

Verhalten des Markov-Prozesses im Zustand i:

- Die Verweilzeit im Zustand ist exponentiell verteilt mit Rate $-q_{ii}$
- Beim Verlasen des Zustandes geht der Prozess mit Wahrscheinlichkeit q_{ij} / $-q_{ii}$ in Zustand j über

Eingebetteter zeitdiskreter Markov-Prozess der Übergangswahrscheinlichkeiten:

- Zustandsübergangswahrscheinlichkeiten $s_{ij} = q_{ij} / -q_{ii}$ für $i \neq j$ und $s_{ii}=0$
- Matrix **S** des eingebetteten diskreten Markov-Prozesses Y

Y_k ist der Zustand nach der k-ten Transition

Verweilzeit im Zustand wird nicht beachtet

Ein kleines Beispiel (Wetter in Dortmund in kontinuierlicher Zeit)

Zustandsraum S = {1, 2, 3}

x = 1: Regen

x = 2: Bedeckt

x = 3: Sonnig

$$\mathbf{Q} = \begin{pmatrix} -0.6 & 0.3 & 0.3 \\ 0.2 & -0.5 & 0.3 \\ 0.2 & 0.2 & -0.4 \end{pmatrix}$$

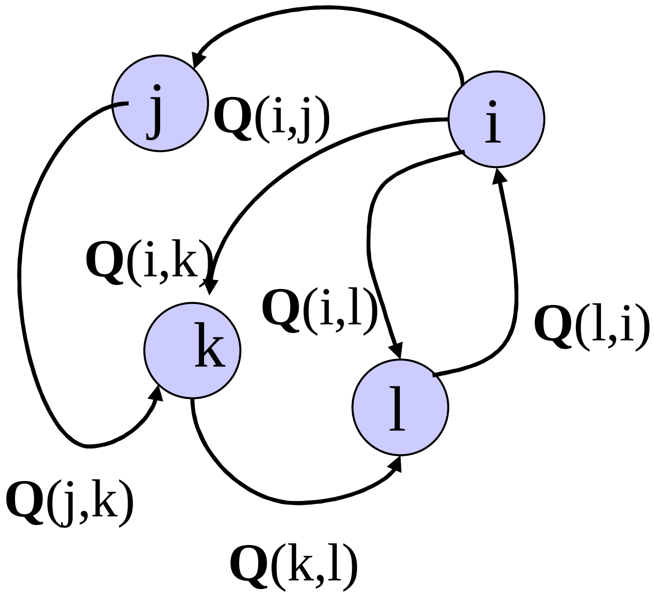
Interpretation: Wenn es bedeckt ist, wird es eine exponentiell verteilte Zeit mit Rate 0.5 bedeckt bleiben (im Mittel also 2 Tage), dann folgt mit Wahrscheinlichkeit 0.4 ein Regentag und mit Wahrscheinlichkeit 0.6 ein Sonnentag

Eingebetteter Markov-Prozess

$$\mathbf{S} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.4 & 0 & 0.6 \\ 0.5 & 0.5 & 0 \end{pmatrix}$$

Interpretation der transienten oder stationären Verteilung des eingebetteten Markov-Prozesses!?

Wie im zeitdiskreten Fall können zeitkontinuierliche Markov-Prozesse mittels gerichteter bewerteter Graphen dargestellt werden



Mit Hilfe des Graphen werden definiert:

- Erreichbarkeit – Unerreichbarkeit von Zuständen
- Abgeschlossene und irreduzible Zustandsmengen

Weitere Eigenschaften können aus dem eingebetteten zeitdiskreten Markov-Prozess abgeleitet werden:

- Rekurrenz von Zuständen: i im eingebetteten Prozess rekurrent $\Rightarrow$ i im zeitkontinuierlichen Prozess rekurrent
- Vorsicht: positive Rekurrenz und Nullrekurrenz übertragen sich genauso wenig wie Periodizität

Wie im zeitdiskreten Fall wird das Verhalten des Markov-Prozesses auch im zeitkontinuierlichen Fall durch die Chapman-Kolmogorov Gleichungen beschrieben

Es gilt dann für homogen Markov-Prozesse gilt:

$$p_{ij}(\tau) = \sum_k p_{ik}(\tau - \alpha) p_{kj}(\alpha) \quad \text{für } 0 \leq \alpha \leq \tau$$

$$\begin{aligned} p_{ij}(t+\Delta) - p_{ij}(t) &= \sum_k p_{ik}(t+\Delta-\alpha)p_{kj}(\alpha) - \sum_k p_{ik}(t-\alpha)p_{kj}(\alpha) \\ &= \sum_k (p_{ik}(t+\Delta-\alpha) - p_{ik}(t-\alpha))p_{kj}(\alpha) \end{aligned}$$

Wenn beide Seiten durch Δ dividiert werden und $\Delta \to 0$, $\alpha \to t$ betrachtet wird, so ergeben sich

Chapman-Kolmogorov Vowärtsgleichungen: $\dfrac{dp_{ij}(t)}{dt} = \sum_k q_{kj} p_{ik}(t)$

Chapman-Kolmogorov Rückwärtsgleichungen: $\dfrac{dp_{ij}(t)}{dt} = \sum_k q_{ik} p_{kj}(t)$

Sei **π**(t) der Vektor der Zustandswahrscheinlichkeiten zum Zeitpunkt t

$$\pi_i(t+\Delta) = \pi_i(t)\left(1 - \sum_{j \neq i} q_{ij}\Delta\right) + \left(\sum_{k \neq i} q_{ki}\pi_k(t)\right)\Delta + o(\Delta)$$

Da die Diagonalelemente von **Q** die negative Zeilensumme beinhalten gilt auch

$$\pi_i(t+\Delta) = \pi_i(t) + \left(\sum_{k} q_{ki}\pi_k(t)\right)\Delta + o(\Delta)$$

und folgende Grenzwertbetrachtung

$$\lim_{\Delta \to 0}\left(\frac{\pi_i(t+\Delta) - \pi_i(t)}{\Delta}\right) = \lim_{\Delta \to 0}\left(\sum_{k} q_{ki}\pi_k(t) + \frac{o(\Delta)}{\Delta}\right)$$

d.h. $\frac{\pi_i(t)}{dt} = \sum_{k} q_{ki}\pi_k(t) \Rightarrow \frac{\pi(t)}{dt} = \pi(t)\mathbf{Q}$

Im stationären Fall ändert sich die Wahrscheinlichkeitsverteilung nicht mehr, d.h. die Ableitung wird **0**.

Q5.2 Vom Modell zum Markov-Prozess

Beschreibung der Modelle üblicherweise in einem Modellierungsformalismus

$\Rightarrow$ Ableitung des Markov-Prozesses

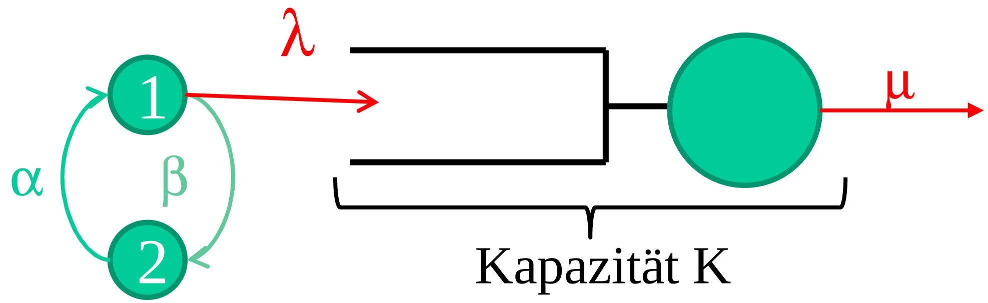
Unterscheidung zeitdiskrete und zeitkontinuierliche Modelle:

- In zeitdiskreten Modellen mit parallelen Aktivitäten, können alle Kombinationen parallel aktivierter Transitionen auftreten
- In zeitkontinuierlichen Modellen treten Ereignisse einzeln auf

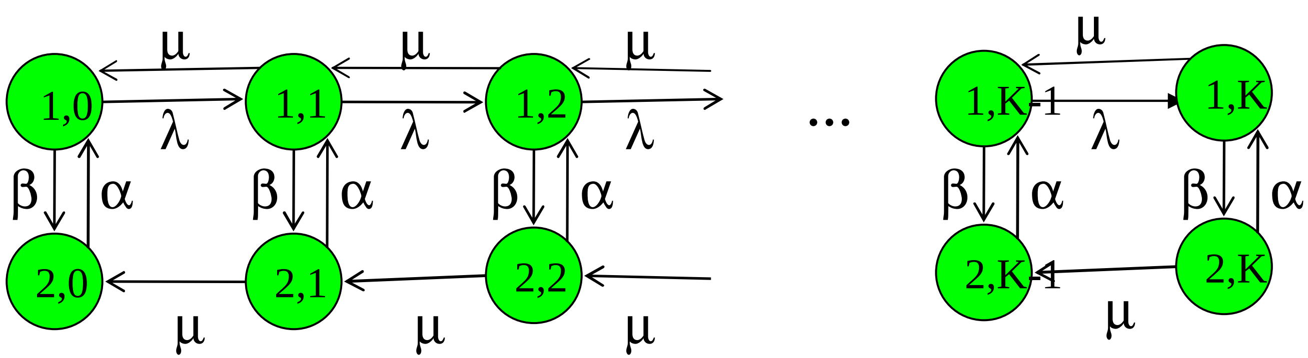
Wir betrachten als Beispiele:

- Warteschlangennetze
- Stochastische Petri-Netze
- Stochastische Prozessalgebren

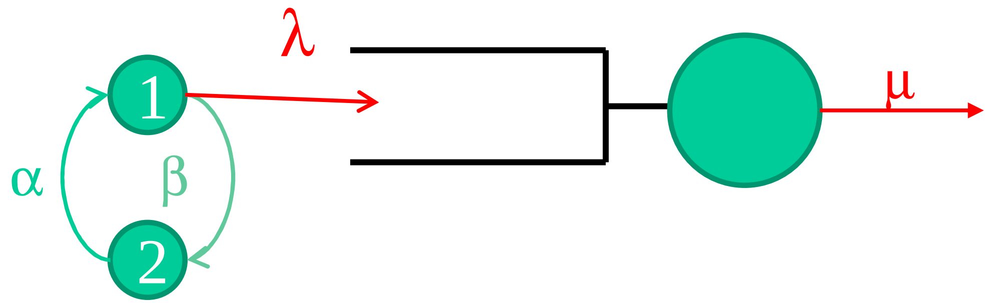
Warteschlangennetz: Station mit unterbrochenem Poisson-Ankunftsprozess (vgl. zeitdiskretes Modell in Q4)



Zustandsraum: S = {(1,0),(2,0),(1,1), …,(1,K),(2,K)}



Leistungsmaße in Warteschlangennetzen:

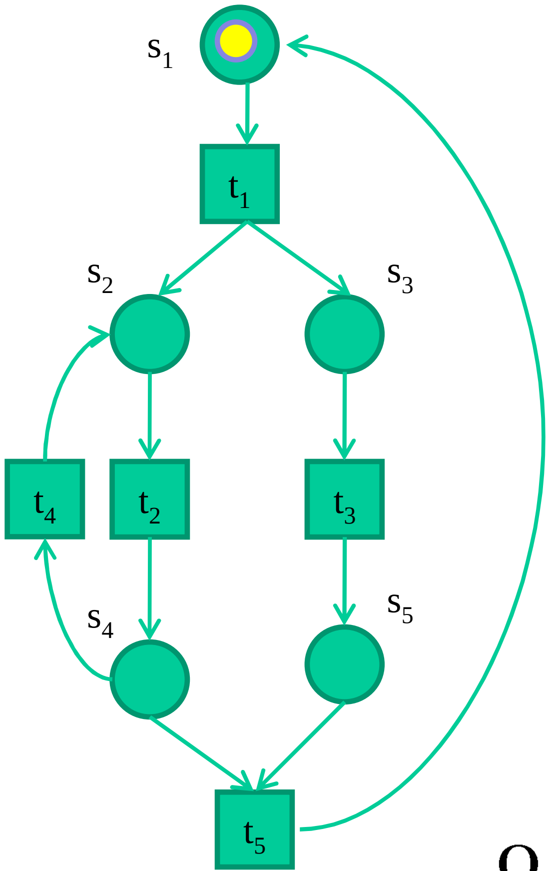


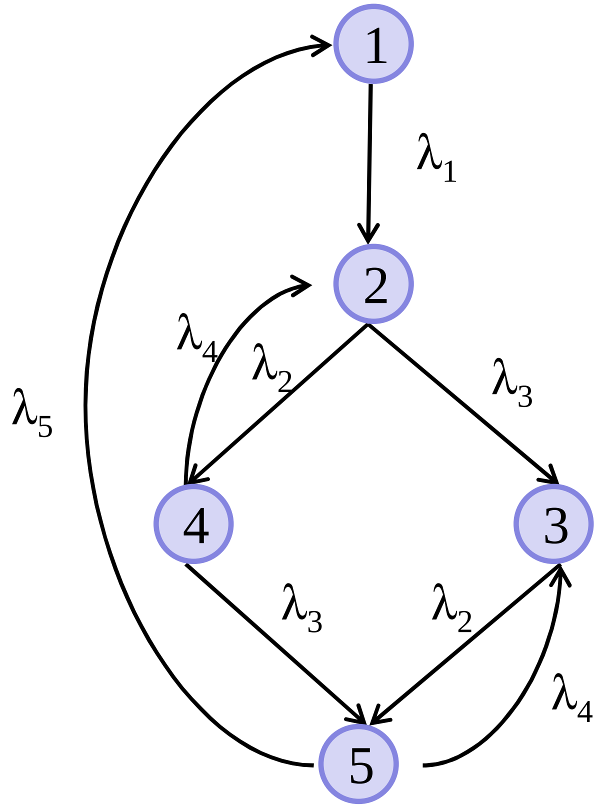
Stationäre Maße werden mit Hilfe des stationären Vektors $\boldsymbol{\pi}$ berechnet (transiente Maße mit Hilfe des transienten Vektors $\boldsymbol{\pi}(t)$)

Beispiele:

- Mittlere Population im Puffer $\Sigma_{k=1\ldots K}\, k(\pi_{1,k} + \pi_{2.k})$
- Auslastung des Bedieners $\Sigma_{k=1\ldots K}\, (\pi_{1,k} + \pi_{2.k})$
- Durchsatz $\Sigma_{k=1\ldots K}\, \mu(\pi_{1,k} + \pi_{2.k})$

Stochastische Petri-Netze





$$\mathbf{Q} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & 0 & 0 \\ 0 & -\lambda_2-\lambda_3 & \lambda_3 & \lambda_4 & 0 \\ 0 & 0 & -\lambda_2 & 0 & \lambda_2 \\ 0 & \lambda_4 & 0 & -\lambda_3-\lambda_4 & \lambda_3 \\ \lambda_5 & 0 & \lambda_4 & 0 & -\lambda_4-\lambda_5 \end{pmatrix}$$

Feuerungsrate $t_i=\lambda_i$

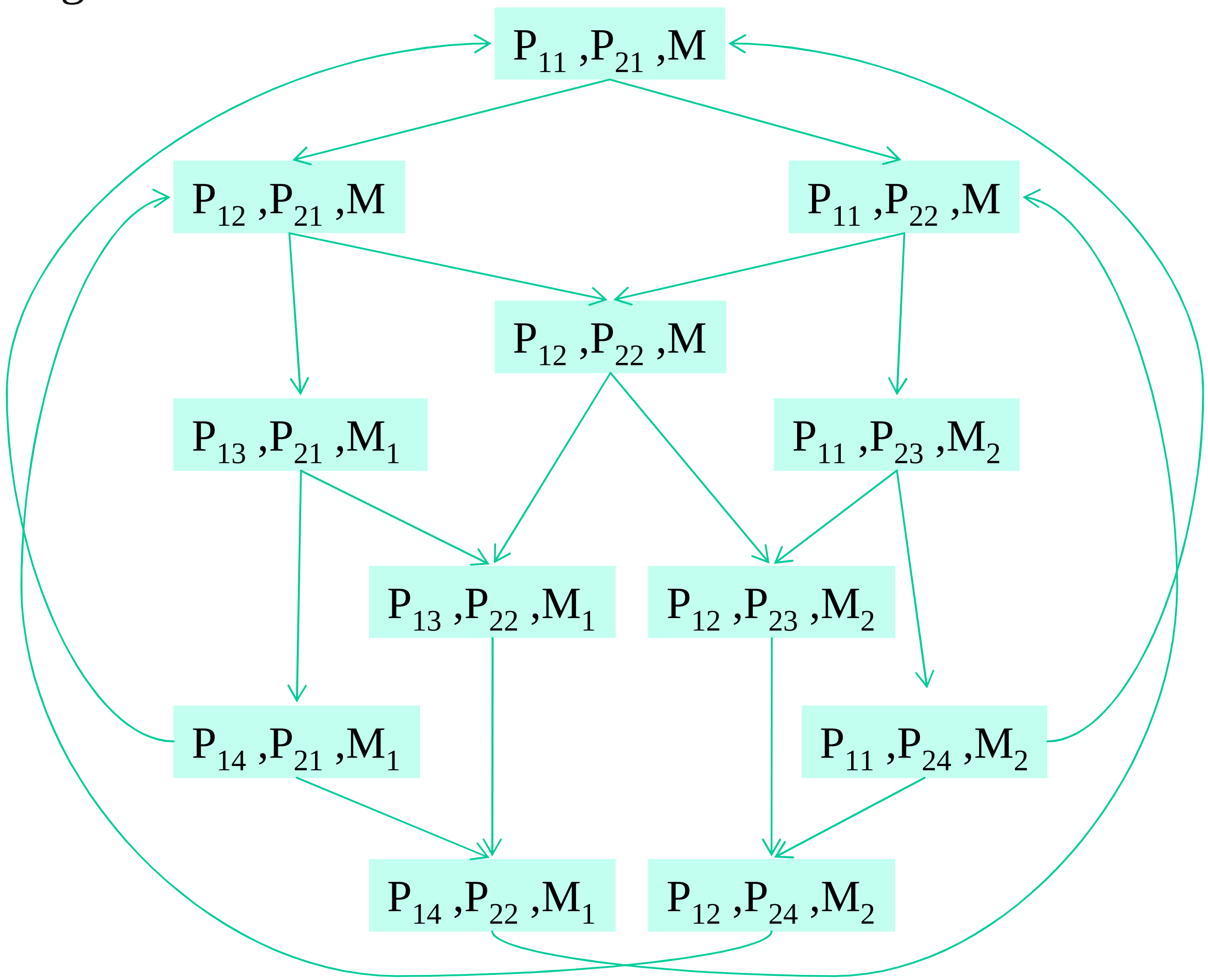
$S=\{s_1,s_2s_3,\ s_2s_5,s_3s_4,s_4s_5\}=\{1,2,3,4,5\}$

Stochastische Prozessalgebren

$P_{11} = (\text{think}, \lambda_1).P_{12}$
$P_{12} = (\text{get}_1, \alpha_1).P_{13}$
$P_{13} = (\text{work}, \mu_1).P_{14}$
$P_{14} = (\text{rel}_1, \beta_1).P_{11}$

$P_{21} = (\text{think}, \lambda_2).P_{22}$
$P_{22} = (\text{get}_2, \alpha_2).P_{23}$
$P_{23} = (\text{work}, \mu_2).P_{24}$
$P_{24} = (\text{rel}_2, \beta_2).P_{21}$

$M = (\text{get}_1, \top).M_1 +$
$\qquad (\text{get}_2, \top).M_2$
$M_1 = (\text{rel}, \top).M$
$M_2 = (\text{rel}, \top).M$

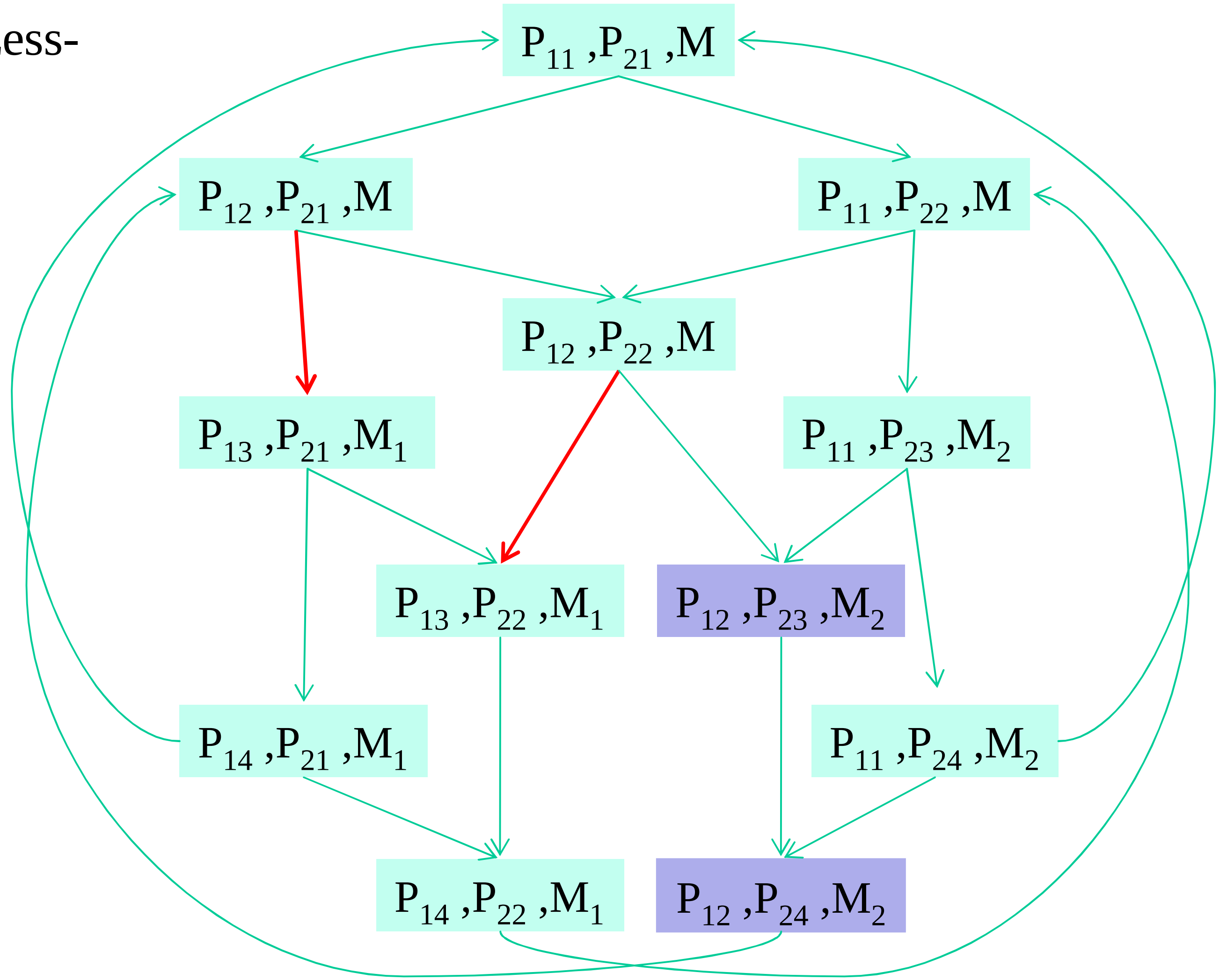


Leistungsmaße in Prozess-Algebren:

- Definition über Zustände von Komponenten oder
- Transitionsbe-schriftungen

Beispiel

- Zustand P_{12} und Ressource M_2
- Durchsatz der mit get_1 beschrifteten Transitionen



Q5.3 Stationäre Analyse

Ziel ist die Bestimmung von $\lim_{t\to\infty}\boldsymbol{\pi}(t)$

Falls der Markov-Prozess stationär ist, gilt $\lim_{t\to\infty}\boldsymbol{\pi}(t) = \boldsymbol{\pi}$

d.h. der Grenzwert konvergiert gegen einen eindeutigen Vektor

Unter Nutzung der Chapman-Kolmogorov Gleichungen gilt dann:

$$\frac{\pi}{dt} = \pi\mathbf{Q} = \mathbf{0} \text{ und } \sum_i \pi_i = 1$$

Bleiben folgende Fragen:

- Wann existiert eine eindeutige stationäre Verteilung?
- Wann ist die stationäre Verteilung unabhängig von $\boldsymbol{\pi}(0)$?
- Wie berechnet man die stationäre Verteilung?
- Gibt es einen Zusammenhang zwischen der stationären Verteilung des zeitkontinuierlichen Markov-Prozesses und des eingebetteten zeitdiskreten Markov-Prozesses?

Existenz der stationären Verteilung:

➢ Für einen endlichen irreduziblen zeitkontinuierlichen Markov-Prozess existiert eine eindeutige stationäre Verteilung, diese ist unabhängig von $\pi(0)$.

➢ Für einen unendlichen irreduziblen zeitkontinuierlichen Markov-Prozess in dem alle Zustände positiv rekurrent sind, existiert eine eindeutige stationäre Verteilung, diese ist unabhängig von $\pi(0)$.

➢ Für einen Markov-Prozess dessen Zustandsraum S sich in Untermengen S_t, $S_{r1}, \ldots, S_{rK}$ disjunkt unterteilen lässt, so dass

- ➢ S_t nur transiente Zustände beinhaltet
- ➢ alle S_{ri} irreduzible Zustandsmengen sind, in denen alle Zustände positiv rekurrent sind

kann eine eindeutige stationäre Verteilung berechnet werden, die von $\pi(0)$ abhängt

Bedingungen wie im zeitdiskreten Fall ohne Periodizität!

Berechnung der stationären Verteilung:

Lösung eines linearen Gleichungssystems

➢Für endliche Zustandsräume Standard
aber Zustandsräume können sehr groß werden (Millionen, Milliarden, …)

➢Im endlichen Fall nur bei spezieller Struktur lösbar (rekursive Berechnung)

Wir beschränken uns erst einmal auf den endlichen Fall

$$\mathbf{Q} = \begin{pmatrix} -0.6 & 0.3 & 0.3 \\ 0.2 & -0.5 & 0.3 \\ 0.2 & 0.2 & -0.4 \end{pmatrix}$$

$$\begin{array}{rclclcl} 0 & = & -0.6\pi_1 & + & 0.3\pi_2 & + & 0.3\pi_3 \\ 0 & = & 0.2\pi_1 & - & 0.5\pi_2 & + & 0.3\pi_3 \\ 0 & = & 0.2\pi_1 & + & 0.2\pi_2 & - & 0.4\pi_3 \\ 1 & = & \pi_1 & + & \pi_2 & + & \pi_3 \end{array}$$

Eine der ersten drei Gleichungen ist redundant!

Resultat lautet $\boldsymbol{\pi}$ = (0.25, 0.3214, 0.4286)

Berechnung der stationären Lösung:

➢Direktes Lösungsverfahren auf Basis der Gauß-Elimination (und Varianten)

Aufwand: Speicher $O(n^2)$ Zeit $O(n^3)$

➢Iteratives Lösungsverfahren berechnet konsekutive Approximationen

Aufwand: Speicher O(nz) (nz: Nichtnuzllelemente in **Q**), Zeit ??

Beispiel: Power-Methode

Sei $\alpha \geq \max_i |\mathbf{Q}_{i,i}|$ und $\mathbf{P} = \mathbf{Q}/\alpha + \mathbf{I}$ (**P** ist eine stochastische Matrix!)

Berechne $\lim_{k\to\infty} \boldsymbol{\pi}(k) = \boldsymbol{\pi}(k-1)\mathbf{P}$ konvergiert gegen $\boldsymbol{\pi}$, falls **P** aperiodisch

Beispiel:

Für α=0.6

$$\mathbf{Q} = \begin{pmatrix} -0.6 & 0.3 & 0.3 \\ 0.2 & -0.5 & 0.3 \\ 0.2 & 0.2 & -0.4 \end{pmatrix} \Rightarrow \mathbf{P} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 1/6 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Power-Methode konvergiert oft sehr langsam, es gibt bessere iterative Methoden!

Zusammenhang zwischen der stationären Verteilung des zeitkontinuierlichen Markov-Prozesses mit Matrix **Q** und des eingebetteten zeitdiskreten Markov-Prozesses mit Matrix **S** für endliche Zustandsräume

- **Q** irreduzibel $\Leftrightarrow$ **S** irreduzibel (**S** sei zusätzlich aperiodisch)
- Sei $\boldsymbol{\pi}$ stationäre Verteilung von **Q** und $\boldsymbol{\theta}$ stationäre Verteilung von **S**

Es gilt $\boldsymbol{\pi}_i = \boldsymbol{\theta}_i / -\mathbf{Q}_{i,i} \left(\Sigma_j\, \boldsymbol{\theta}_j / -\mathbf{Q}_{j,j}\right)^{-1}$

$$\mathbf{Q} = \begin{pmatrix} -0.6 & 0.3 & 0.3 \\ 0.2 & -0.5 & 0.3 \\ 0.2 & 0.2 & -0.4 \end{pmatrix} \qquad \mathbf{S} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.4 & 0 & 0.6 \\ 0.5 & 0.5 & 0 \end{pmatrix}$$

$\boldsymbol{\pi}$ = (0.25, 0.3214, 0.4286)

$\boldsymbol{\theta}$ = (0.3111, 0.3333, 0.3556)

Q5. 4 Transiente Analyse

Berechnung von $\boldsymbol{\pi}(t)$ in Abhängigkeit von $\boldsymbol{\pi}(0)$

Es gilt $\boldsymbol{\pi}_i(t) = \Sigma_j \boldsymbol{\pi}_j(0)P[X(t)=i \mid X(0) = j] = \Sigma_j \pi_j(0)p_{ji}(t)$

In Matrixform $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\mathbf{P}(t)$ mit der Lösung $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)e^{\mathbf{Q}t}$

Approximation durch Taylor-Reihe $e^{Qt} = \sum_{k=0}^{\infty} \frac{(\mathbf{Q}t)^k}{k!}$

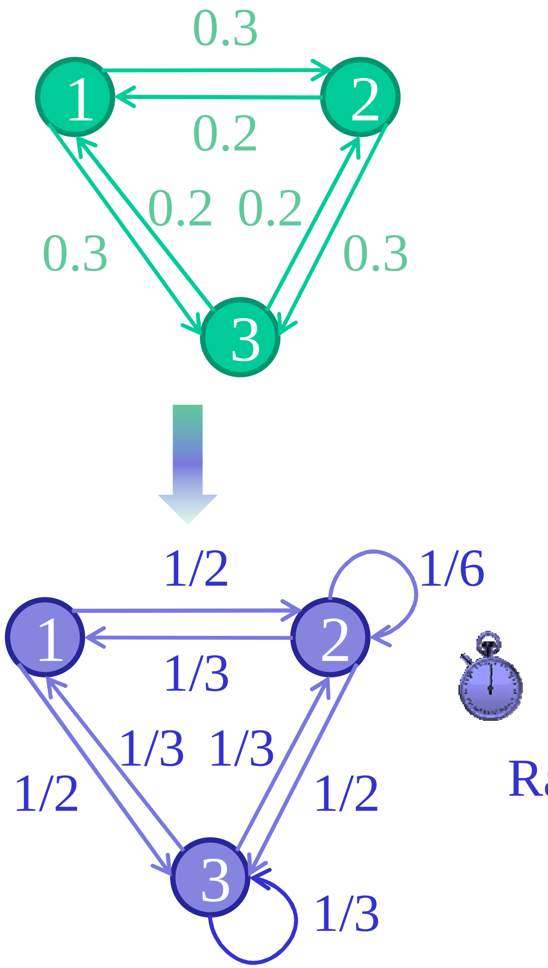
Damit gilt: $\pi(t) \approx \pi(0) \sum_{k=0}^{K} \frac{(\mathbf{Q}t)^k}{k!}$

$$\mathbf{Q} = \begin{pmatrix} -0.6 & 0.3 & 0.3 \\ 0.2 & -0.5 & 0.3 \\ 0.2 & 0.2 & -0.4 \end{pmatrix} \Rightarrow \mathbf{P}(2.0) \approx \begin{pmatrix} 0.401 & 0.276 & 0.323 \\ 0.200 & 0.478 & 0.323 \\ 0.200 & 0.231 & 0.569 \end{pmatrix}$$

Berechnung u.U. in mehreren Schritten $\boldsymbol{\pi}(0), \boldsymbol{\pi}(t1),\ldots,\boldsymbol{\pi}(t_L)=\boldsymbol{\pi}(t)$

Alternative Methode: Randomiserung/Uniformisierung

Zeitkontinuierlicher
Markov-Prozess

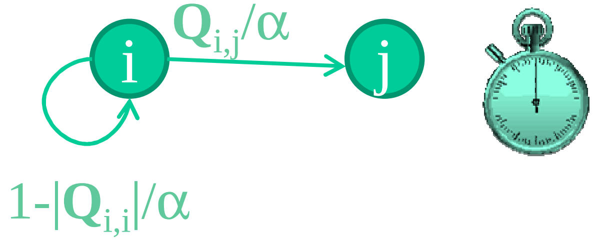


Rate 0.6

Einführung eines Weckers mit exponentiell verteilten Zeiten mit Rate $\alpha \geq \max_i |\mathbf{Q}_{i,i}|$

Falls der Prozess im Zustand i ist und der Wecker abläuft, so akzeptiere Ereignis mit Wahrscheinlichkeit $|\mathbf{Q}_{i,i}| / \alpha$

$\Rightarrow$ Zeit bis zum nächsten Ereignis exponentiell verteilt mit Rate $|\mathbf{Q}_{i,i}|$



Exponentiell verteilt mit Rate α

Transformation des zeitkontinuierlichen Prozesses in einen zeitdiskreten Prozess + einen Poisson-Prozess

$\mathbf{Q} \Rightarrow \mathbf{P} = \mathbf{Q}/\alpha + \mathbf{I}$ (stochastische Matrix)

Es gilt dann $\pi(t) = \pi(0) \sum_{k=0}^{\infty} e^{-\alpha t} \frac{(\alpha t)^k}{k!} \mathbf{P}^k$

Vorteil gegenüber Taylor-Reihe, keine negativen Werte und Berechnung von Fehlerschranken möglich

$$\pi(0) \sum_{k=0}^{K} e^{-\alpha t} \frac{(\alpha t)^k}{k!} \mathbf{P}^k \leq \pi(t) \leq \pi(0) \sum_{k=0}^{K} e^{-\alpha t} \frac{(\alpha t)^k}{k!} \mathbf{P}^k + \epsilon \mathbf{e}$$

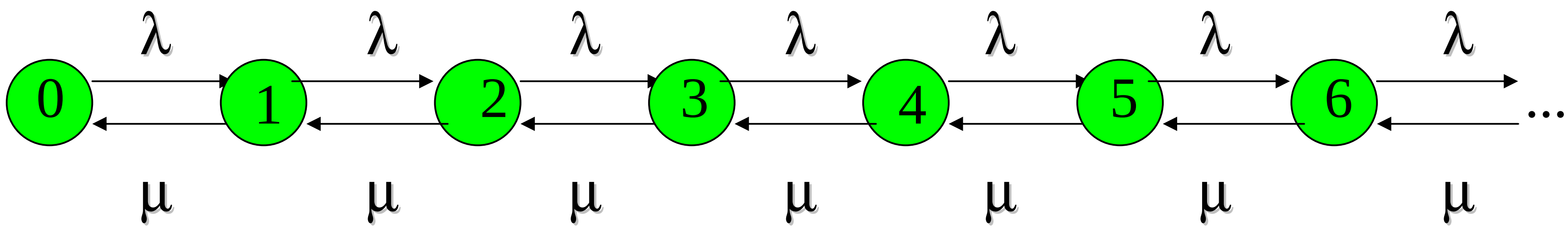
mit $\epsilon = 1 - \sum_{k=0}^{K} e^{-\alpha t} \frac{(\alpha t)^k}{k!}$

Q5.5 Stationäre Analyse von Modellen mit unendlichem Zustandsraum

Lösung eines Gleichungssystems mit unendlich vielen Variablen

$\Rightarrow$ Nur in speziellen Fällen möglich, wenn eine rekursive Struktur vorliegt

Beispiel: Geburts-/Todesprozess
(hier zeitkontinuierlich, zeitdiskret analog zu behandeln)



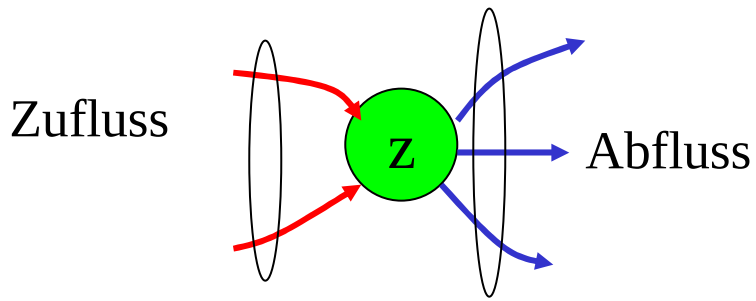
Sei $\rho = \lambda / \mu$

Basis der stationären Analyse von Markov-Prozessen

Wahrscheinlichkeitsfluss: W. im Zustand zu sein · Transitionsrate

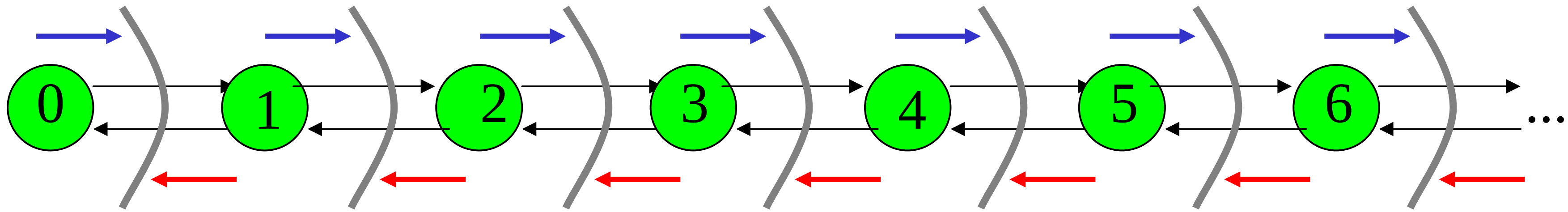
Im stationären Zustand ändert sich die Wahrscheinlichkeitsverteilung nicht ⇒

Fluss aus einem Zustand = Fluss in den Zustand



⇒ Berechnung der stationären Wahrscheinlichkeiten lässt sich auf Lösung eines linearen Gleichungssystems mit rekursiver Struktur reduzieren

Geburts-/Todes-Prozess:



Flussgleichgewicht über jeden Schnitt!

Resultierendes Gleichungssystem

$$\pi_i \cdot \lambda = p_{i+1} \cdot \mu \;\Rightarrow\; \pi_i = \pi_0 \left(\frac{\lambda}{\mu}\right)^i = \rho^i$$

- Falls wir $\boldsymbol{\pi}_0$ kennen, können wir alle $\boldsymbol{\pi}_i$ berechnen

Überlegungen zur Berechnung von $\boldsymbol{\pi}_0$:

$\boldsymbol{\pi}_i$ (i = 0, ..., ∝) definieren eine Wahrscheinlichkeitsverteilung

⇒ Summe der $\boldsymbol{\pi}_i$ muss 1.0 ergeben

Damit gilt: $$\sum_{i=0}^{\infty} \pi_i = \pi_0 \sum_{i=0}^{\infty} \rho^i \;\Rightarrow\; \pi_0 = \frac{1}{\sum_{i=0}^{\infty} \rho^i}$$

Vorausgesetzt die Reihensumme ist endlich!

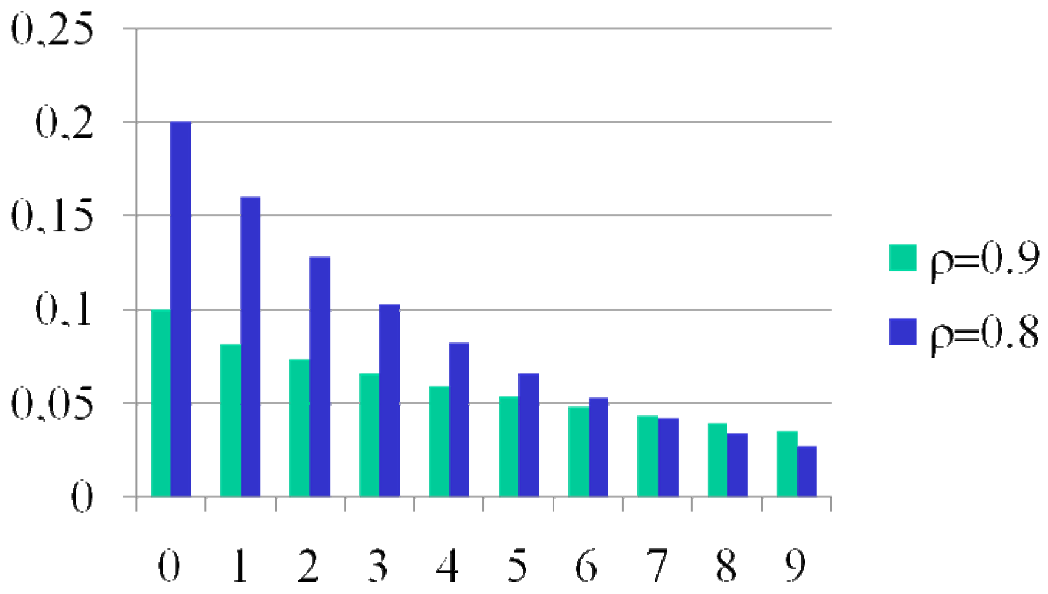
Geometrische Reihe: $$\sum_{i=0}^{\infty} \rho^i = (1-\rho)^{-1} \quad \text{falls } \rho < 1$$

Falls ρ ≥ 1 werden alle Zustände nullrekurrent

⇒alle Zustandswahrscheinlichkeiten werden 0

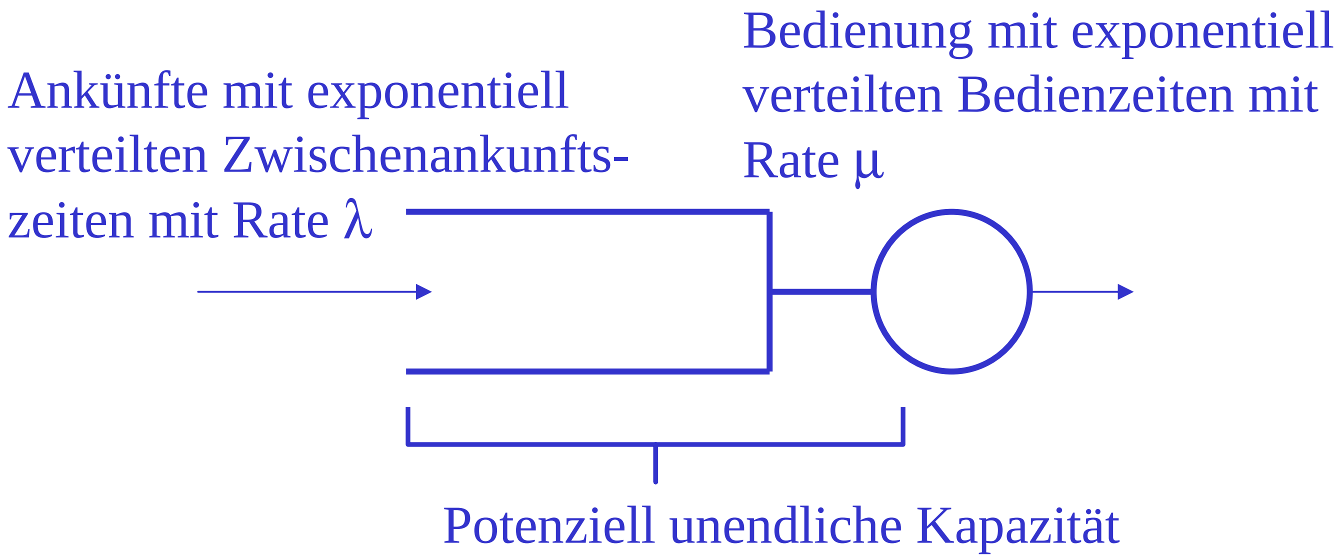
Prozess driftet nach rechts (gegen unendlich)

Für ρ < 1 geometrisch fallende Zustandswahrscheinlichkeiten



Interpretation als einfache Warteschlange

(Details zur Modellklasse im nächsten Kapitel)



Typische Fragestellungen:

- Wie viele Kunden warten im Mittel?

(Population N (inkl. Bediner) bzw. N_w (exkl. Bediener)

- Wie viele Kunden werden im Durchschnitt pro Zeiteinheit bedient? (Durchsatz X)

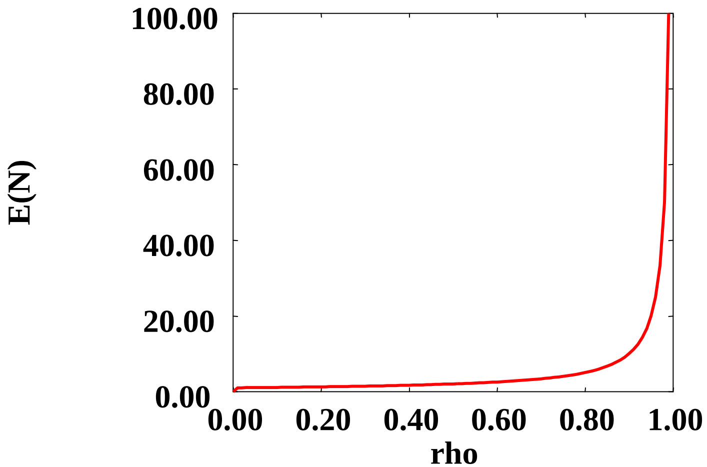
- Wie lange wartet ein Kunde im Mittel? (Wartezeit W)

(N, N_w, X, W sind Zufallsvariablen, wir berechnen Erwartungswerte!)

Auslastung: $\rho = 1 - \pi_0$ Durchsatz $E(X) = \lambda$

Mittlere Population:
$$\begin{aligned} E(N) &= \sum_{i=0}^{\infty} i \cdot \pi_i &&= (1-\rho)\sum_{i=1}^{\infty} i \cdot \rho^i \\ &= (1-\rho)\rho\frac{d}{d\rho}\sum_{i=0}^{\infty}\rho^i &&= (1-\rho)\rho\frac{d}{d\rho}\frac{1}{1-\rho} \\ &= (1-\rho)\rho\frac{1}{(1-\rho)^2} &&= \frac{\rho}{1-\rho} \end{aligned}$$

Mittlere Population:
(nur Warteraum)
$$\begin{aligned} E(N_w) &= \sum_{i=0}^{\infty}(i-1) \cdot \pi_i &&= (1-\rho)\rho\sum_{i=1}^{\infty} i \cdot \rho^i \\ &= \rho E(N) &&= \frac{\rho^2}{1-\rho} \end{aligned}$$



Verweilzeiten später siehe Q6.1

Verallgemeinerung der Ideen

Zur Erinnerung:

Sei $\mathbf{Q}$ irreduzible Matrix eines zeitkontinuierlichen Markov-Prozesses

Sei $\alpha \geq \max_i |\mathbf{Q}_{ii}|$ ($\alpha < \infty$) und $\mathbf{P} = \mathbf{Q}/\alpha + \mathbf{I}$

Sei $\mathbf{P}$ aperiodisch (immer erreichbar durch Wahl von $\alpha > \max_i |\mathbf{Q}_{ii}|$)

Es gilt dann für Vektor $\boldsymbol{\pi}$ mit $\Sigma_i\, \boldsymbol{\pi}_i = 1$: $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0} \Leftrightarrow \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$

$\Rightarrow$Wir können mit $\mathbf{Q}$ oder $\mathbf{P}$ arbeiten

(und nutzen im Folgenden $\mathbf{P}$)

Ziel: Herausarbeiten von Strukturen, die es erlauben auch bei unendlichen Zustandsmengen $\boldsymbol{\pi}$ zu berechnen

Bisher behandelte Matrixstruktur (siehe M/M/1-System)
Geburts-/Todesprozess:

$$\mathbf{P} = \begin{pmatrix} 1-p_{00} & p_{01} & & \\ p_{10} & 1-p_{10}-p_{12} & & \\ & p_{21} & 1-p_{21}-p_{23} & p_{23} \\ & \ddots & \ddots & \ddots \end{pmatrix}$$

Mit der rekursiven Berechnung:

$$\pi_{i+1} = (p_{i+1i})^{-1}\left((1-p_{ii})\pi_i - p_{i-1i}\pi_{i-1}\right), \quad \pi_1 = (p_{10})^{-1}(1-p_{00})\pi_0$$

Wenn π_0 bekannt ist, so können alle π_i berechnet werden!

Möglicher Ansatz:
1. Setze $\pi_0 = 1$
2. Berechne $\pi_1,\ldots,\pi_N$
3. Normalisiere den resultierenden Vektor

Auftretender Fehler proportional zu

$$\sum_{i=N+1}^{\infty} \pi_i \Big/ \sum_{i=0}^{N} \pi_i$$

Verallgemeinerung obere Hessenberg-Matrix

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & \cdots \\ p_{10} & p_{11} & p_{12} & p_{13} & \cdots \\ 0 & p_{21} & p_{22} & p_{23} & \cdots \\ 0 & 0 & p_{32} & p_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Mit der rekursiven Berechnung:

$$\pi_{i+1} = (p_{i+1i})^{-1} \left((1 - p_{ii})\pi_i - \sum_{j=0}^{i-1} p_{ji}\pi_j \right) \quad , \quad \pi_1 = (p_{10})^{-1}(1 - p_{00})\pi_0$$

Rekursive Berechnung, wie auf vorheriger Folie!

Geht es noch weiter?

Ja, wenn sich nicht einzelne Elemente, sondern ganze Untermatrizen wiederholen!

$\Rightarrow$ Matrix-Geometrische/Matrix-Analytische Methoden (hier nur erste Ansätze)

Matrix eines Quasi-Geburts-Todes-Prozesses (QBD):

$$\mathbf{Q} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{B}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{pmatrix}$$

Dimensionen der Matrizen

- $\mathbf{B}_1$: p×p
- $\mathbf{B}_0$: p×m
- **B**2 : m×p
- $\mathbf{A}_0$, $\mathbf{A}_1$ und $\mathbf{A}_2$: m×m

Es gilt (mit **e**=(1,…,1), **0**=(0,…,0)):

- $\mathbf{B}_1\mathbf{e}^T + \mathbf{B}_0\mathbf{e}^T = \mathbf{0}^T$
- $\mathbf{B}_1\mathbf{e}^T + (\mathbf{A}_0 + \mathbf{A}_1)\, \mathbf{e}^T = \mathbf{0}^T$
- $(\mathbf{A}_0 + \mathbf{A}_1 + \mathbf{A}_2)\, \mathbf{e}^T = \mathbf{0}^T$

Sei nun $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \ldots)$ mit

- $\boldsymbol{\pi}_0$ Vektor der Länge p
- $\boldsymbol{\pi}_i$ (i>0) Vektor der Länge m

Ges. Vektor $\boldsymbol{\pi}$ mit $\boldsymbol{\pi}\mathbf{Q}=\mathbf{0}$ und $\boldsymbol{\pi}\mathbf{e}^T = 1$, sofern dieser existiert!
Wir nehmen an, dass **Q** irreduzibel ist

Struktur der Matrix ähnelt Geburts-Todes-Prozess

wenn Skalare durch Matrizen ersetzt werden!

Lösung Geburts-Todes-Prozess: $\boldsymbol{\pi}_i = \rho\boldsymbol{\pi}_{i-1} = \rho^i\boldsymbol{\pi}_0$ und $\boldsymbol{\pi}_0 = 1- \rho$

Geburts-Todes-Prozess $\Rightarrow$ Quasi-Geburts-Todes-Prozess

Ersetze ρ durch Matrix **R** der Dimension $m\times m$ und

Prüfe, was passiert!

Erst einmal für j≥2 (bei j=0 oder 1 u.U. Dimensionsprobleme)

$$\pi_{j-1}\mathbf{A}_0 + \pi_j\mathbf{A}_1 + \pi_{j+1}\mathbf{A}_2 = \mathbf{0} \qquad (*)$$

Wenn eine Matrix R wie angenommen existiert so gilt für j≥2:

$$\pi_j = \pi_{j-1}\mathbf{R} \;\Rightarrow\; \pi_j = \pi_1\mathbf{R}^{j-1}$$

Eingesetzt in (*):

$$\pi_1\mathbf{R}^{j-2}\mathbf{A}_0 + \pi_1\mathbf{R}^{j-1}\mathbf{A}_1 + \pi_1\mathbf{R}^{j}\mathbf{A}_2 = \mathbf{0} \qquad \text{für alle } j\geq 2$$

Damit muss auch gelten:

$$\pi_1\mathbf{R}^{j-2}\left(\mathbf{A}_0 + \mathbf{R}\mathbf{A}_1 + \mathbf{R}^2\mathbf{A}_2\right) = \mathbf{0} \;\Rightarrow\; \mathbf{A}_0 + \mathbf{R}\mathbf{A}_1 + \mathbf{R}^2\mathbf{A}_2 = \mathbf{0}$$

Quadratische Matrixgleichung zur Bestimmung von **R**

Damit $\boldsymbol{\pi}$ existiert , muss gelten $\boldsymbol{\pi}_1\Sigma_{i=0,..}\mathbf{R}^i\ \mathbf{e}^T < \infty$

d.h. der Spektralradius von **R** muss kleiner 1 sein!

Angenommen, wir hätten ein „passendes“ **R** gefunden,

dann muss noch $\boldsymbol{\pi}_0$ berechnet werden!

Es gilt:

$$\begin{aligned} \pi_0 \mathbf{B}_1 + \pi_1 \mathbf{B}_2 &= \mathbf{0} \\ \pi_0 \mathbf{B}_0 + \pi_1 \mathbf{A}_1 + \pi_2 \mathbf{A}_2 &= \mathbf{0} \end{aligned}$$

Einsetzen von $\boldsymbol{\pi}_1\mathbf{R}$ für $\boldsymbol{\pi}_2$ liefert:

$$(\pi_1, \pi_2) \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_0 \\ \mathbf{B}_2 & \mathbf{A}_1 + \mathbf{R}\mathbf{A}_2 \end{pmatrix} = \mathbf{0}$$

Lösung erfordert zusätzliche Normierungsbedingung (Summe der W. 1):

$$1 = \pi_0 \mathbf{e}^T + \pi_1 \sum_{j=0}^{\infty} \mathbf{R}^j \mathbf{e}^T = \pi_0 \mathbf{e}^T + \pi_1 \left(\mathbf{I} - \mathbf{R}\right)^{-1} \mathbf{e}^T$$

Berechnung von **R**:

Viele Verfahren existieren, wir schauen uns nur einen einfachen Algorithmus an

Matrix $\mathbf{A}_1$ besitzt negative Diagonalelemente und $\mathbf{A}_1\mathbf{e}^T \leq \mathbf{0}$

Falls **Q** irreduzibel ist, existiert $(\mathbf{A}_1)^{-1}$

Damit können wir umformen:

$$\mathbf{0} = \mathbf{A}_0 + \mathbf{R}\mathbf{A}_1 + \mathbf{R}^2\mathbf{A}_2 \;\Rightarrow \mathbf{R} = -\left(\mathbf{A}_0 \left(\mathbf{A}_1\right)^{-1} + \mathbf{R}^2\mathbf{A}_2 \left(\mathbf{A}_1\right)^{-1}\right)$$

Fixpunkt-Algorithmus, initialisiere **R**[0]=0 und iteriere

$$\mathbf{R}[k] = -\left(\mathbf{A}_0 \left(\mathbf{A}_1\right)^{-1} + \left(\mathbf{R}[k-1]\right)^2 \mathbf{A}_2 \left(\mathbf{A}_1\right)^{-1}\right)$$

bis $\|\mathbf{R}[k]-\mathbf{R}[k-1]\| < \varepsilon$

Verfahren konvergiert, aber u.U. sehr langsam!