

§40 Stochastische Prozesse

1. Definition (stochastischer Prozess)

Ein stochastischer Prozess (X_t) ist eine Familie von ZV mit $t \in \mathbb{R}$ oder $t \in \mathbb{N}$

Wir denken dabei an t als Zeitparameter, der kontinuierlich oder diskret ist.

Beispiel:

Y_k sei $\mathcal{B}_{n,p_k}$ -verteilt, unabhängig (oder $Y_k \stackrel{\text{i.i.d.}}{\sim} Y$)

$X_n = \sum_{k=1}^n Y_k$, X_n ist ein stochastischer Prozess

X_n hängt nur von X_{n-1} und Y_n ab $X_n = X_{n-1} + Y_n$

2. Definition (Markov-Kette)

Ein stochastischer Prozess (X_n) heißt Markov-Kette, wenn

$$P(X_n = c \mid X_{n-1} = b, X_i \text{ bekannt}, i \leq n-1) \\ = P(X_n = c \mid X_{n-1} = b)$$

Die Verteilung von X_n hängt bei gegebenem $X_{n-1} = b$ nur von X_{n-1} ab und nicht von X_k , $k \leq n-1$.

Wichtigste Klasse:

Markov-Ketten mit endlich vielen (oder diskreten) Werten, d.h. $P(X_n \in \{1, \dots, k\}) = 1$

③ Beispiel:

- Prozessor bearbeitet 1 Aufgabe pro Zeitschritt
- In jedem Zeitschritt kommen neue Aufgaben dazu
- Cache ≤ 2 , andere Aufträge werden ignoriert

repariert eine Aufgabe bearbeiten
weitere Aufgaben dazu
max. insgesamt 2

forever

$A_n \stackrel{\text{st}}{=} A$ sei die Anzahl der zusätzlichen
Aufträge im n -ten Zeitpunkt, sei eine
geometrisch verteilte ZV mit Parameter
 $1-p$, beginnend bei 0.

$$P(A=i) = (1-p) \cdot p^i$$

$$P(A \geq i) = p^i$$

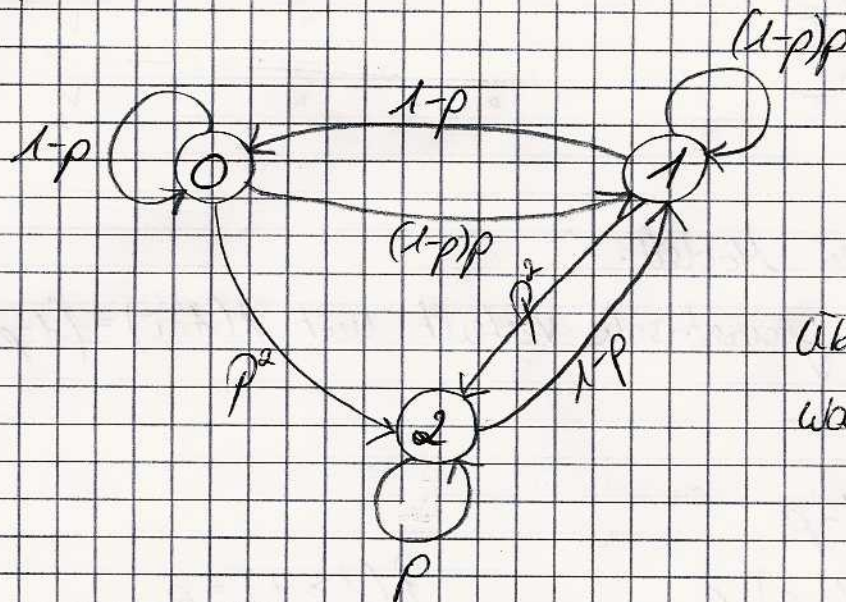
$$E(A) = x \frac{d}{dx} F_A$$

$$F_X(x) = \sum_{i=0}^{\infty} (1-p)p^i x^i = \frac{(1-p)}{(1-px)}$$

$$x \frac{d}{dx} F_A(x) = (1-p) \frac{px}{(1-px)^2}$$

$$E(X) = \sum_{x=1}^{\infty} x \cdot p \cdot (1-p)^{x-1} = \frac{p}{1-p}$$

Graphisches Modell



Übergangs-
wahrscheinlichkeiten

5. Allgemein

Zustände $\{1, \dots, k\}$

$$P_{ij}^n = P(X_n = i \mid X_0 = j)$$

$$M_n = (P_{ij}^n) \in M(k \times k, \mathbb{R}) \quad \text{Matrix der Übergangswahrscheinlichkeiten}$$

Ist $M_n = M$ unabhängig von n , dann haben wir eine
Zeitschritt-unabhängige Markov-Kette.

Im Bsp:

$$M = \begin{pmatrix} 1-p & 1-p & 0 \\ (1-p)p & (1-p)p & 1-p \\ p^2 & p^2 & p \end{pmatrix}$$

$$\pi_0 = \begin{pmatrix} \pi_{01} \\ \pi_{02} \\ \pi_{03} \end{pmatrix}$$

ω -Verteilung
für X_0

Der Vektor $M\pi_0$ gibt dann die Verteilung für X_1 an.
 " $M^2\pi_0$ " " " X_2 "
 " $M^n\pi_0$ " " " X_n "

Zum graphischen Modell:

Zugänge sind geometrisch verteilt mit $P(A=i) = (1-p)p^i$
 $i \in \mathbb{Z}_{\geq 0}$.

$$P(A=0) = 1-p$$

$$P(A=1) = (1-p) \cdot p$$

$$P(A \geq 1) = p$$

$$P(A \geq 2) = \sum_{i=2}^{\infty} (1-p)p^i = p^2$$

6. Simulation

7. Allgemein

(X_n) Markov-Kette mit endlich vielen Zuständen,
 Übergangsmatrix $M = (p_{ij})$.

mit $P(X_n=i | X_{n-1}=j) = P(X_m=i, X_{m-1}=j)$

In der Simulation hatten wir in dem Beispiel
 folgende Fragen bejaht.

Frage 1) Ist π^0 irgendeine Anfangsverteilung,
 gilt $\lim_{n \rightarrow \infty} M^n \pi^0 = \pi$
 wobei π eine stationäre Zufallsverteilung
 ist?
 $\pi = M\pi$

Frage 2) Gilt $\lim_{n \rightarrow \infty} M^n = (\pi, \pi, \pi, \dots, \pi) \in M(k \times k)$?

Bemerkung:

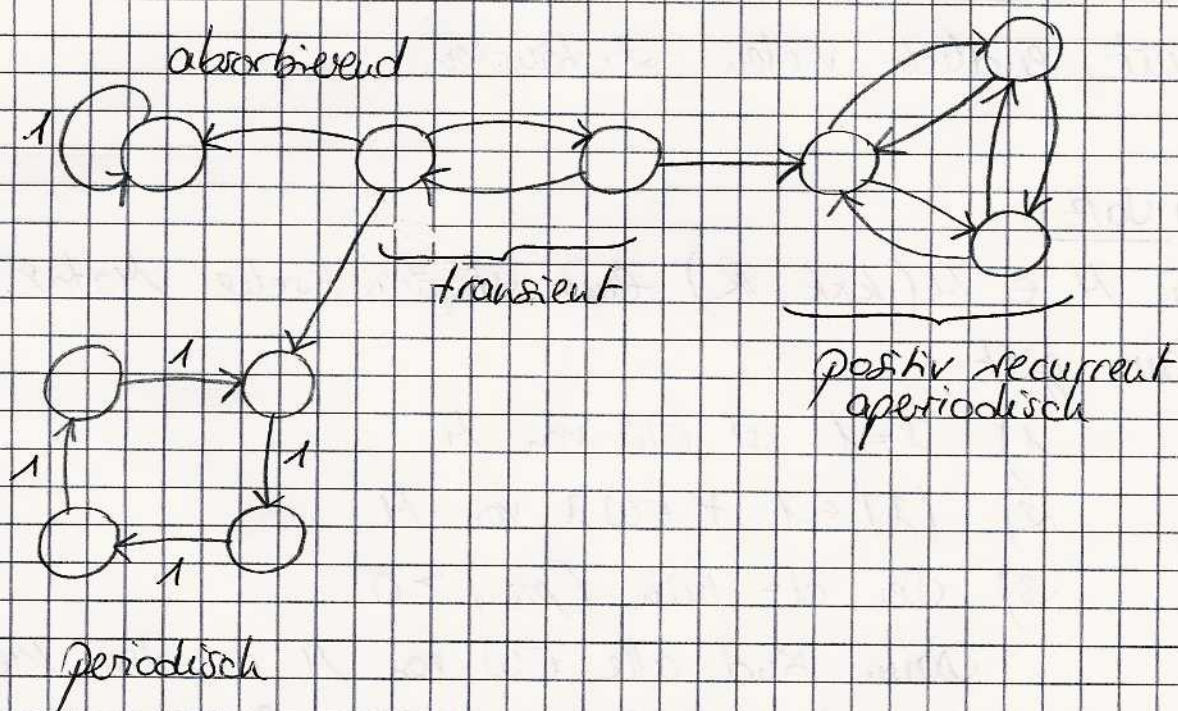
Wenn Antwort Frage 2 Ja!

$\Rightarrow$ Limes Zustand unabhängig von Anfangsverteilung.

$\Rightarrow$ Frage 1 Ja!

8. Beispiel:

Endliche, zeitunabhängige Markov-Kette mit Graphen.



9. Definition (transient, recurrent, absorbierend, period.)

Ein Zustand i heißt

a) transient, wenn $P(X_m \neq i \ \forall m \geq n \mid X_n = i) > 0$

b) recurrent, wenn $P(X_m \neq i \ \forall m \geq n \mid X_n = i) = 0$
(mit Wahrscheinlichkeit 1 kehren wir zurück)

c) periodisch mit Periode l , wenn
 $P(X_{l+n} = i \mid X_n = i) = 1$

d) eine Menge I von Zuständen heißt absorbierend, wenn $P(X_{n+1} \in I \mid X_n = I) = 1$

10. Definition (stochastische Matrix)

Eine Matrix $M = (p_{ij}) \in M(k \times k, \mathbb{R})$ mit $p_{ij} \in [0, 1]$ und $\sum_{j=1}^k p_{ij} = 1$ heißt stochastische Matrix.

Eine solche Matrix beschreibt die Übergangswahrscheinlichkeiten eines Markovschen Prozesses mit endlich vielen Zuständen.

11. Satz

Sei $M \in M(k \times k, \mathbb{R})$ eine stochastische Matrix.

Dann gilt

- 1) $\lambda = 1$ ist EW von M
- 2) $|\lambda| \leq 1 \quad \forall \text{ EW } \lambda \text{ von } M$
- 3) Sei $d = \min(p_{ii}) > 0$

Dann sind alle EW von M in dem Kreis $K = \{\lambda \in \mathbb{C} \mid |\lambda - d| \leq 1 - d\}$ enthalten.

Insbesondere ist $\lambda = 1$ der einzige EW mit $|\lambda| = 1$.

- 4) Ist λ ein EW mit $|\lambda| = 1$, dann hat die Summe der Jordanklocke zum EW λ die Gestalt λE_s , wobei s die Vielfachheit der Nullstelle λ in $\chi_M(t) = \det(M - tE)$ ist.
- 5) Ist λ ein EW mit $|\lambda| = 1$, dann $\exists m > 0$, so dass $\lambda^m = 1$ ist, d.h. λ ist m -te Einheitswurzel.

Beweis:

1) M stochastisch

$\Rightarrow \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^k$ ist EV von ${}^t M$
zum EW $\lambda = 1$.

2) Wir betrachten die Spaltensummennorm
 $\|M\|_s = \max_j \left(\sum_{i=1}^k |p_{ij}| \right)$

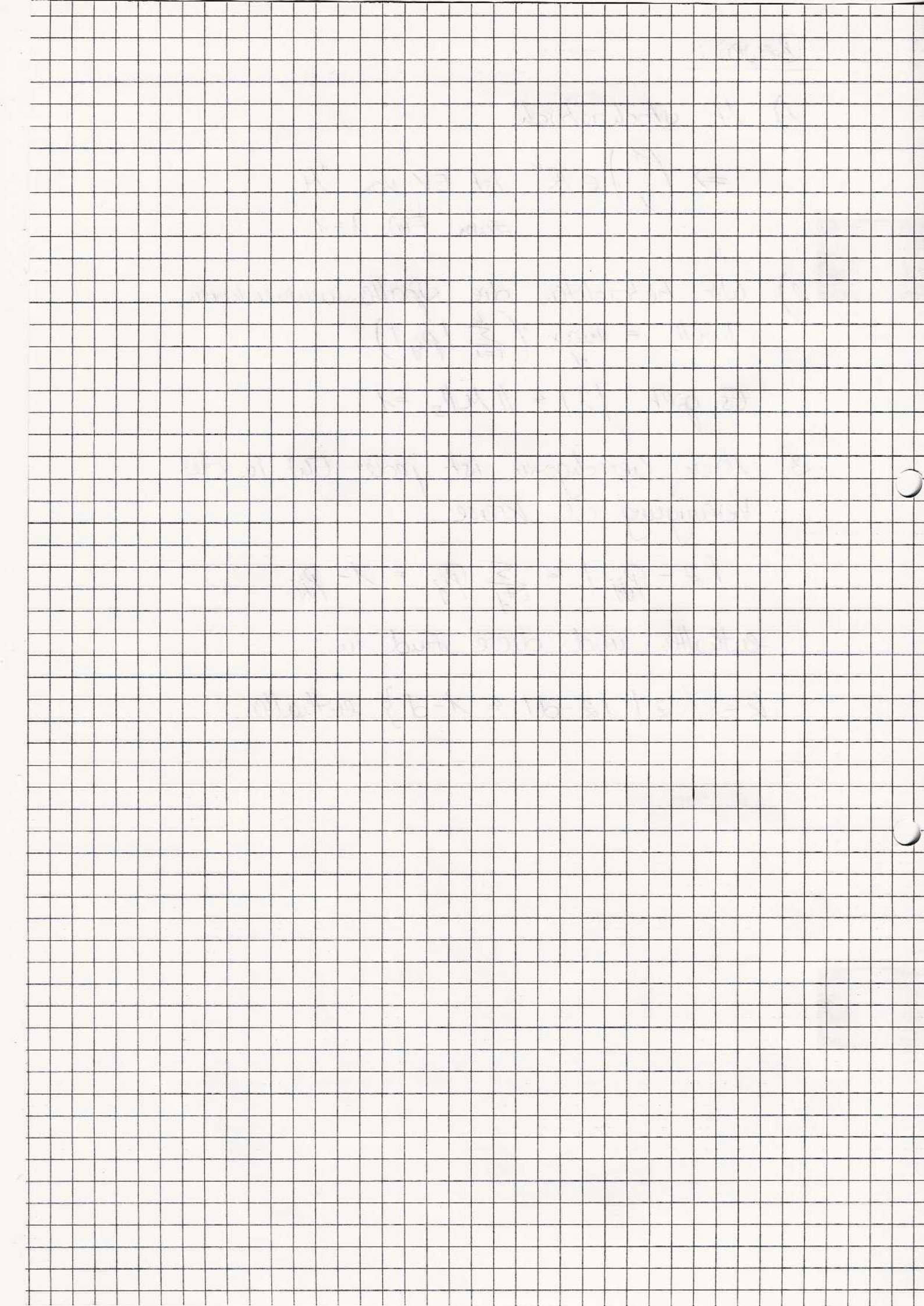
Es gilt $|2| \leq \|M\|_s = 1$

3) Nach Gerschgorin ist jeder EW in der Vereinigung der Kreise

$$|2 - p_{jj}| \leq \sum_{i \neq j} p_{ij} = 1 - p_{jj}$$

enthalten und diese sind in

$K = \{z \mid |2 - d| \leq 1 - d\}$ enthalten.



$$M = (p_{ij}) \text{ mit}$$

$$p_{ij} \geq 0 \quad \sum_{i=1}^k p_{ij} = 1$$

heißt stochastische Matrix

Beschreibt eine Markov-Kette mit k Zuständen.

$$p_{ij} = P(X_{n+1} = i \mid X_n = j)$$

Bsp.: Labyrinth mit Maus

$$\begin{array}{ccccccc} 1 & = & 3 & = & 5 & 7 & < & 9 \\ \parallel & & & & \parallel & \parallel & & \nearrow \\ 2 & = & 4 & = & 6 & = & 8 & = & 10 \end{array}$$

$$p_{ii} = \frac{1}{2}$$

$$p_{ij} = \begin{cases} \frac{1}{2} \cdot \frac{1}{\#(\text{Türen aus } j \text{ heraus})} \\ 0 & \text{keine Tür von } j \text{ nach } i \end{cases}$$

$$M = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & \frac{1}{6} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{6} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{6} & 0 & \frac{1}{2} \end{pmatrix}$$

Satz: Sei M eine stochastische Matrix.

Dann gilt:

- 1) $\lambda = 1$ ist EW
- 2) $|\lambda| \leq 1$ für alle EW λ von M
- 3) Ist $d = \min_{i,j} p_{ij} > 0$, dann sind alle EW von M in dem Kreis

$$K = \{z \in \mathbb{C} \mid |z - d| \leq 1 - d\} \text{ enthalten}$$

Insbesondere ist dann $\lambda = 1$ der einzige EW mit $|\lambda| = 1$.
Dann hat die Summe der Jordankästchen zu λ die Gestalt λE_s , wobei $s = \text{Vielfachheit der Nullstelle } t = \lambda \text{ im charakt. Polynom.}$

- 5) Ist λ EW mit $|\lambda| = 1$, dann $\exists m \geq 1$, so daß $\lambda^m = 1$, d.h. λ ist eine m -te Einheitswurzel.

Beweis:

- 4) Sei J die Jordannormalform von M und T die Übergangsmatrix, $J = T \cdot M \cdot T^{-1}$

Dann gilt

$$J^n = (T M T^{-1})^n = T \cdot M^n \cdot T^{-1}$$

Sei nun B ein Jordankästchen zu λ ,

$$B = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

$\Rightarrow B^n$ ist ein Kästchen von J^n .

$$B^2 = \begin{pmatrix} \lambda^2 & 2\lambda & 1 & 0 \\ & \ddots & \ddots & \ddots \\ 0 & & 2\lambda & \\ & & & \lambda^2 \end{pmatrix}$$

$$B^3 = \begin{pmatrix} \lambda^3 & 3\lambda^2 & 3\lambda & 1 & 0 \\ & \ddots & \ddots & \ddots & \ddots \\ 0 & & 3\lambda^2 & 3\lambda & \\ & & & \lambda^3 & \end{pmatrix}$$

$$B = \lambda \cdot E + N \quad N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}, L \times L \text{ Kästchen, } N^L = 0$$

$$B^n = (\lambda E + N)^n = \sum_{h=0}^n \lambda^{n-h} N^h \binom{n}{h}$$

insbesondere

$$B^n = \begin{pmatrix} \lambda^n & n \cdot \lambda^{n-1} & \dots \\ & \ddots & \ddots \\ 0 & & \lambda^n \end{pmatrix}$$

$$\begin{aligned} n+1 &= |\lambda^n| + n \cdot |\lambda^{n-1}| \quad \text{da } |\lambda|=1 \\ &\leq \|B^n\|_5 \leq \|J^n\|_5 = \|T M^n T^{-1}\|_5 \\ &\leq \|T\|_5 \cdot \|M^n\|_5 \cdot \|T^{-1}\|_5 \\ &\leq \underbrace{\|T\|_5 \cdot \|T^{-1}\|_5}_{\text{Konstante}} \end{aligned}$$

$\Rightarrow$ Blockgröße von $B < 2$

$\Rightarrow B$ 1x1-Matrix

$\Rightarrow 4)$

Beweis zu 3)

Wir benötigen Notation $K = \{1, \dots, k\}$ für $j \in K$.

$M_1(j) = \{i \in K \mid p_{ij} > 0\}$ die Menge der Zustände, die wir von j in einem Schritt erreichen können.

$K_0 \subset K$, $M_1(K_0) = \bigcup_{j \in K_0} M_1(j)$ die Zustände, die von einem Zustand in K_0 in einem Schritt erreicht werden können.

Induktiv: $M_0(K_0) = K_0$,

$$M_t(K_0) = M_t(M_{t-1}(K_0))$$

Die Zustände, die wir von K_0 in t Schritten erreichen können.

$M(K_0) = \bigcup_{t \geq 0} M_t(K_0)$ die Menge aller irgendwie erreichbaren Zustände aus K_0 .

zu 5) Sei jetzt $\lambda \in W$ von M mit $|\lambda| = 1$ und $x = {}^t(x_1, \dots, x_k)$ ein Eigenvektor zu λ von tM .

Sei $y = \max |x_j|$ und $K_0 = \{j \mid |x_j| = y\}$

Dann gilt: $x_j \cdot \lambda = \sum_{i=1}^k x_i p_{ij}$

Sei $j \in K_0$. Dann gilt:

$$y = |x_j| = |x_j \cdot \lambda| = \left| \sum_{i=1}^k x_i p_{ij} \right|$$

$$\leq \sum_{i=1}^k |x_i| \cdot p_{ij} \leq \sum_{i=1}^k y p_{ij} = y$$

und alle $\geq$ -Zeichen in dieser Kette sind $=$ -Zeichen.

$$\Rightarrow a) |x_i| = y \quad \forall i \in M_+(j)$$

$$b) \text{ alle } x_i \text{ mit } i \in M_+(j)$$

stimmen überein, d.h. sind als komplexe Zahlen gleich.

$$x_j \cdot \lambda = \sum_{i=1}^k x_i \cdot p_{ij} = x_0 \cdot \sum_{i=1}^k p_{ij} = x_0$$

für $\lambda \in M_+(j)$

Wir wählen j_0 fest und setzen $R = \{j \mid x_j = x_{j_0}\}$

Dann gilt:

$$\emptyset \neq M_+(R) \ni i : x_i = \lambda x_{j_0}$$

$$\Rightarrow \emptyset \neq M_t(R) \ni i : x_i = \lambda^t x_{j_0}$$

$M_0(R), M_1(R), \dots, M_s(R)$ alle $\neq \emptyset$

$$\bigcup_{s \geq 0} M_s(R) \subset K_0$$

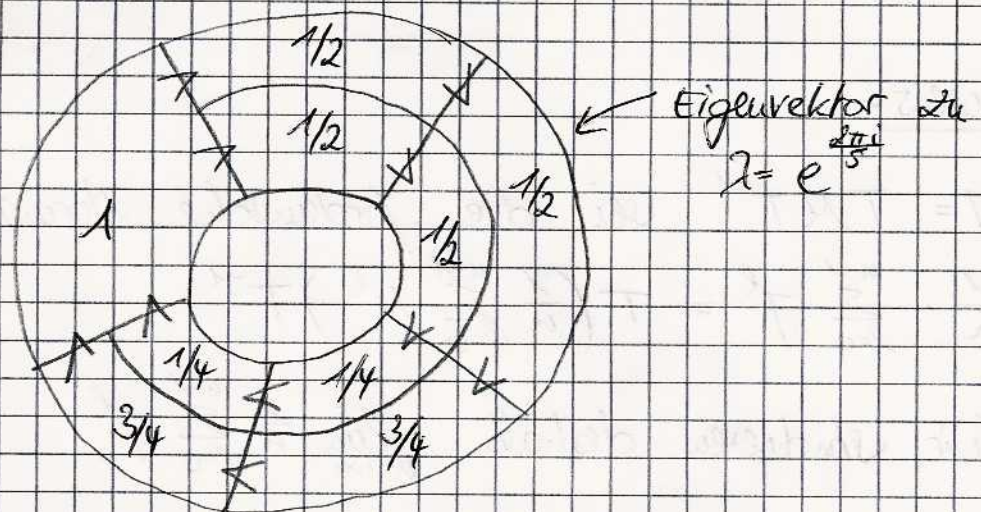
$\Rightarrow \exists s \neq t$, so daß $M_s(R) \cap M_t(R) \neq \emptyset$

$$\Rightarrow x_i = \lambda^s x_j = \lambda^t x_j$$

$$\Rightarrow \lambda^{s-t} = 1$$

$\Rightarrow \lambda$ ist Einheitswurzel

Beispiel: Labyrinth mit wanderndem Maus, $p_{ii} = 0$.



12. Satz (Ergodensatz)

Sei M stochastische Matrix

a) Dann existiert

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=0}^{n-1} M^i \right) = Q$$

Q ist ein Projektor, der M vertauscht:

$$Q^2 = Q = M Q = Q \cdot M$$

Der Rang von Q ist die Vielfachheit des Eigenwertes $\lambda = 1$ von M und Q beschreibt eine Projektion auf $\text{Eig}(M, 1)$.

b) Ist $\lambda=1$ der einzige EW von M mit $|\lambda|=1$, dann gilt:

$$\lim_{n \rightarrow \infty} M^n = Q$$

c) Ist $\lambda=1$ der einzige EW von M mit $|\lambda|=1$ und $\dim \text{Eig}(M, 1) = 1$, dann gilt:

$$Q = \lim_{n \rightarrow \infty} M^n = \begin{pmatrix} z_1 & \dots & z_1 \\ \vdots & & \vdots \\ z_k & \dots & z_k \end{pmatrix} \text{ mit } z = (z_1, \dots, z_k)$$

der eindeutig bestimmte Gleichgewichtszustandsverteilung
 $z_i \geq 0$, $\sum_{i=1}^k z_i = 1$ $z = M_2$

Beweis:

$J = T M T^{-1}$ sei die Jordansche Normalform.

$$\frac{1}{n} \cdot \sum_{l=0}^{n-1} J^l = T \left(\frac{1}{n} \sum_{l=0}^{n-1} M^l \right) T^{-1}$$

Wir studieren deshalb $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=0}^{n-1} J^l$

Sei B ein Jordanblock > 1

$$B = \begin{pmatrix} \lambda & 1 & 0 \\ & \ddots & \vdots \\ 0 & & \lambda \end{pmatrix}, \text{ dann gilt } |\lambda| < 1$$

$$\Rightarrow B^l = \begin{pmatrix} \lambda^l & \binom{l}{1} \lambda^{l-1} & \binom{l}{2} \lambda^{l-2} & \dots & \binom{l}{r-1} \lambda^{l-r+1} \\ & \ddots & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

$$\left| \sum_{l=0}^{n-1} \binom{l}{s} \lambda^{l-s+1} \right| \leq \sum_{l=0}^{n-1} \binom{l}{s} |\lambda|^{l-s+1}$$

$$\leq \sum_{l=0}^{\infty} \binom{l}{s} |\lambda|^{l-s+1} = \frac{1}{s!} \left[\left(\frac{d}{dx} \right)^s \left(\frac{1}{1-x} \right) \right]_{x=|\lambda|} = L_s < \infty$$

$$\Rightarrow \frac{1}{n} \left(\sum_{l=0}^{n-1} B^l \right) \xrightarrow{n \rightarrow \infty} 0 \quad \text{und} \quad \lim_{l \rightarrow \infty} B^l = 0$$

da $\sum_{l=0}^{\infty} B^l$ existiert.

λ Einheitswurzel, etwa $\lambda^m = 1, \lambda \neq 1$.

$$\Rightarrow \sum_{l=0}^{m-1} \lambda^l = 0$$

$$\parallel$$

$$\frac{\lambda^m - 1}{\lambda - 1}$$

$$\Rightarrow \left| \sum_{l=0}^{n-1} \lambda^l \right| \leq \frac{m}{2} \Rightarrow \frac{1}{n} \sum_{l=0}^{n-1} \lambda^l = \frac{m}{2^n} \xrightarrow{n \rightarrow \infty} 0$$

$$\lambda = 1 \quad \frac{1}{n} \sum_{l=0}^{n-1} \lambda^l = 1$$

$$\Rightarrow \frac{1}{n} \sum_{l=0}^{n-1} J^l \xrightarrow{n \rightarrow \infty} \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix}$$

S = Vielfachheit von 1 ab EW von M

$$\Rightarrow Q = T^{-1} \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} T \quad Q^2 = Q,$$

$$Q \cdot M = T^{-1} \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} T T^{-1} J T$$

$$= T^{-1} \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_s & 0 \\ 0 & * \end{pmatrix} T$$

$$= T^{-1} \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} T = Q$$

Ebenso $M \cdot Q = Q$

Bleibt c) einzusehen:

$$Q = \lim_{n \rightarrow \infty} \underbrace{\left(\frac{1}{n} \sum_{\ell=0}^{n-1} M^\ell \right)}_{\text{stochastisch}} \text{ ist stochastisch}$$

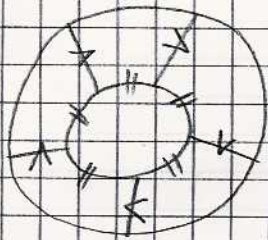
$\text{rang } Q = 1 \Rightarrow \text{Behauptung}$

Markovsche Kette,

$\mathcal{Z} = \{1, \dots, k\}$ Zustände

$$P_{ij} = P(X_{t+1} = i \mid X_t = j)$$

$$P_{ii} = 1/2$$



13) Bemerkung

M stochastische Matrix. Äquivalent sind

$$\dim \left[\bigoplus_{\lambda \text{ mit } |\lambda|=1} \text{Eig}(M, \lambda) \right] = 1$$

und

$$\dim \left[\bigoplus_{\lambda \text{ mit } |\lambda|=1} \text{Eig}(M^n, \lambda) \right] = 1$$

In diesem Fall ist $\lambda=1$ der einzige Eigenwert von M (bzw. M^n), mit $|\lambda|=1$ und $\dim \text{Eig}(M, 1) = 1$

$$\lim_{n \rightarrow \infty} M^n = \begin{pmatrix} z_1 & \dots & z_1 \\ \vdots & & \vdots \\ z_k & & z_k \end{pmatrix}$$

wobei $z = (z_1, \dots, z_k)$

Gleichgewichtsverteilung

$$Mz = z, \quad z_i \geq 0, \quad \sum_{i=1}^k z_i = 1.$$

§ 4.1 Verborgene Markovsche Modelle

Anwendung: Maschinelles Lernen, Spracherkennung, Mustererkennung (Genom)

1. Definition (Hidden-Markov-Modell)

Ein Hidden-Markov-Modell ist ein Tupel

$$\Lambda = (M, B, \pi, v)$$

M stochastische $k \times k$ -Matrix

v Alphabet (Zichensatz)

B $k \times r$ -Matrix (B_{ix}), $B_{ix} \geq 0$, $\sum_x B_{ix} = 1$

π $k \times 1$ -Matrix (Anfangszustandsverteilung)

Dies modelliert folgenden Prozess

1) Wählen Anfangszustand $Z_1 \in \{1, \dots, k\}$ gemäß π zufällig aus

2) Z_t Zustand zum Zeitpunkt t gegeben, wählen Z_{t+1} zufällig gemäß M aus.

3) Sind im Zustand $Z_t = i$, dann wählen wir ein $x \in v$ mit Wahrscheinlichkeit B_{ix} aus.

Resultat:

Eine Zustandsfolge $Z_1, Z_2, \dots, Z_T$

und Zeichenfolge $S = s_1 s_2 \dots s_T$

T : Gesamtschrittzahl

Der Beobachter kann nur $s_1, \dots, s_T$ lesen, die Zustandsfolge bleibt verborgen. (hidden).

② Grundlegende Aufgaben

- I) Gegeben das Modell $\Lambda = (M, B, \pi, V)$
und die Zeichenfolge $S = S_1 \dots S_T$.
Wie groß ist die Wahrscheinlichkeit, dass S von Λ generiert wird?
- II) Gegeben S und Λ , welches ist die
wahrscheinlichste Zustandsfolge $Z_1 \dots Z_T$.
- III) Gegeben S und partielle Information über Λ
etwa: Anzahl der Zustände, der Graph,
Approximationen der Übergangswahrscheinlichkeit $(p_{ij}) = M$.

Ziel: Welches Λ liefert S mit der größten
Wahrscheinlichkeit?

I) $P(S/\Lambda)$ lässt sich schnell auf 2 Methoden berechnen.

③ Vorwärtsmethode

Wir setzen $\alpha_i(i) = P(S_1 \dots S_i / Z_i = i, \Lambda)$

Dann gilt

$$\alpha_1(i) = \pi_i B(i, S_1) \quad (\text{Anfang})$$

$$\alpha_{i+1}(i) = \sum_{j=1}^K B(i, S_{i+1}) p_{ij} \alpha_i(j) \quad (\text{Rekursion})$$

$$P(S/\Lambda) = \sum_{i=1}^K \alpha_i(i) \pi_i \quad (\text{Ende})$$

④ Rückwärtsmethode

$$\beta_t(i) = P(S_{t+1} \dots S_T \mid Z_t = i, \Lambda)$$

Dann gilt

$$\beta_T(i) = 1$$

Und

$$\beta_t(i) = \sum_{j=1}^k p_{ji} \cdot B(j, S_{t+1}) \beta_{t+1}(j)$$

(Ende)

$$P(S/\Lambda) = \sum_{j=1}^k \pi_j \cdot B(j, S_1) \beta_1(j)$$

⑤ Raten der Zustandssequenz

Gegeben S und Λ , wollen $Z_1 \dots Z_T$ finden mit der größten Wahrscheinlichkeit.

Betrachten

$$\gamma_i(t) = P(Z_t = i \mid S, \Lambda)$$

$$= P(S_1 \dots S_t \mid Z_t = i, \Lambda)$$

$$\cdot \cancel{P(S_{t+1} \dots S_T \mid Z_t = i, \Lambda)} \quad \cancel{P(S/\Lambda)}$$

$\leadsto$ Die wahrscheinlichste Zustandsfolge $Z_1 \dots Z_T$ ist die, bei der

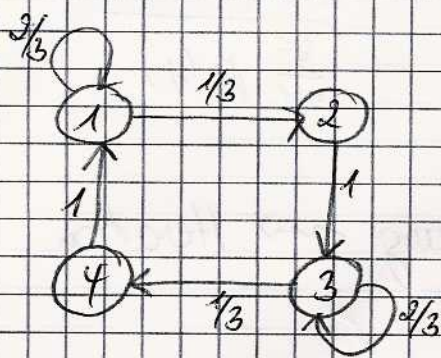
$$\alpha_t(Z_t) \cdot \beta_t(Z_t) = \max_i \{ \alpha_t(i) \beta_t(i) \} \text{ gilt.}$$

Warnung: • Nur lokal maximiert.

• $P(Z_1 \dots Z_T) = 0$, da eventuell Übergang von Z_t nach Z_{t+1} Wahrscheinlichkeit 0 hat

Alternative: Viterbi-Algorithmus

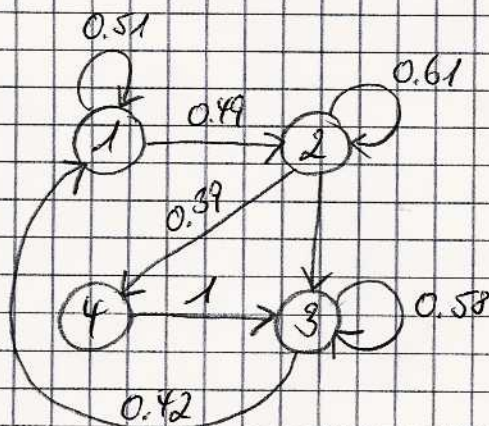
6. III Verbesserung des Modells (Baum-Welsh)



$$\begin{pmatrix} 2/3 & 0 & 0 & 1 \\ 1/3 & 0 & 0 & 0 \\ 0 & 1 & 1/3 & 0 \\ 0 & 0 & 2/3 & 0 \end{pmatrix}$$

Baum-Welsh :

$$\begin{pmatrix} 0.51 & 0 & 0.42 & 0 \\ 0.49 & 0.61 & 0 & 0 \\ 0 & 0 & 0.58 & 1 \\ 0 & 0.39 & 0 & 0 \end{pmatrix}$$



Haben Approximation $\Lambda = (\mu, B, \pi, v)$.

Wollen die Parameter verbessern unter der Beobachtung S .

$$\begin{aligned} \tilde{\xi}_t(i, j) &= P(Z_t = j, Z_{t+1} = i \mid S, \Lambda) \\ &= \frac{\beta_{t+1}(i) b(i, S_{t+1}) p_{ij} \xi_t(j)}{P(S \mid \Lambda)} \end{aligned}$$

Dann gilt $\gamma_t(j) = \sum_{i=1}^K \tilde{\xi}_t(i, j)$

Schätzen neue Matrix μ, B, π .

$$\overline{p}_{ij} = \frac{\sum_{t=1}^{T-1} \sum_j (c_{ij})}{\sum_{t=1}^{T-1} \gamma_t(j)}, \quad \overline{b}(j, x) = \frac{\sum_{t=1, \phi_t=x}^{T-1} \phi_t(j)}{\sum_{t=1}^{T-1} \gamma_t(j)}$$

$$\overline{\pi}_i = \gamma_1(i) \rightarrow \text{Verbesserung} \rightarrow \text{Iteration}$$

Kapitel NUMERIK§ 4.2 Rundungsfehler

Bei Gleitkommadarstellung sind Rundungsfehler unvermeidlich.

1. Beispiel:

$$\begin{pmatrix} 1.00 \cdot 10^{-4} & 1.00 \\ 1.00 & 1.00 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1.00 \\ 2.00 \end{pmatrix}$$

Rechengenauigkeit: 3 Dezimalstellen

korrekte Lösung auf 4 Stellen genau: $x_1 = 1.000$
 $x_2 = 0.999$

1a unserer Genauigkeit: $x_1 = 1.00$
 $x_2 = 1.00$

Gauss-Algorithmus:

$$\left(\begin{array}{cc|c} 1.00 \cdot 10^{-4} & 1.00 & 1.00 \\ 0 & 1.00 \cdot 10^{-4} & 1.00 \cdot 10^{-4} \end{array} \right)$$

Lösung: $x_2 = 1.00$ (korrekt), $x_1 = 0.00$ (falsch!)

2. Gaussalgorithmus mit Spaltenpivotstrategie

a) $A^{(k)} \rightarrow A^{(k+1)}$ k -ter Eliminationsschritt

Wählen ein $p \in \{k, \dots, n\}$, so dß
 $|a_{kp}^{(k)}| \geq |a_{kj}^{(k)}|$ für $j = k, \dots, n$

b) vertauschen p -te mit k -ter Zeile

$$A^{(k)}$$

$$P_k A^{(k)}$$

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

c) $L_k P_k A^{(k)} = A^{(k+1)}$

wobei

$$L_k = \begin{pmatrix} & & & & \\ & & & & \\ & & 1 & & 0 \\ & & l_{k+1,k} & 1 & \\ 0 & l_{n,k} & & & 1 \end{pmatrix}$$

k -te Spalte

offenbar gilt

$$P_k \cdot \hat{L}_j \cdot P_k = \hat{L}_j \quad \text{für } j < k$$

$$\text{mit } |l_{j,k}| \leq 1$$

③ Folgerung

$A \in GL(n, \mathbb{R})$ $\exists P$ Permutationsbasis, eine unipotente untere Dreiecksmatrix

$$L = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ l_{ij} & & 1 \end{pmatrix} \quad \text{mit } |l_{ij}| \leq 1$$

eine obere Dreiecksmatrix $R = \begin{pmatrix} r_{11} & & * \\ & \ddots & \\ 0 & & r_{nn} \end{pmatrix}$,

so d.h.s. $PA = LR$ (LR-Zerlegung)

Beweis:

$$P = P_{n-1} \cdots P_2 \cdot P_1$$

$$\tilde{L}_k = P_{n-1} \cdots P_{k+1} L_k P_{k+1} \cdots P_{n-1}$$

$$\tilde{L} = \tilde{L}_{n-1} \cdots \tilde{L}_1 \quad \text{unipotente untere Dreiecksmatrix}$$

$$R = \tilde{L} P A \quad \leadsto \quad L = \tilde{L}^{-1} \Rightarrow P A = L R$$

Bsp.:

$$P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1.00 & 1.00 \\ 1.00 \cdot 10^{-4} & 1.00 \end{pmatrix} = P_1 \cdot A$$

$$L_1 = \begin{pmatrix} 1 & 0 \\ -1.00 \cdot 10^{-4} & 1 \end{pmatrix} \quad L_1 P_1 A = \underbrace{\begin{pmatrix} 1.00 & 1.00 \\ 0 & 1.00 \end{pmatrix}}_R$$

$$P = P_1, L = \tilde{L}_1^{-1} = \begin{pmatrix} 1 & 0 \\ 1.00 \cdot 10^{-4} & 1 \end{pmatrix}$$

$$\Rightarrow P A = L R$$

A pos. def. symmetrisch, Gaussalgorithmus wird modifiziert, um die Struktur zu erhalten.

④ Satz

a) (rationale Cholesky-Zerlegung)

$A > 0$ sym. pos. def.

Dann $\exists$ unipotente untere Dreiecksmatrix L ,

eine Diagonalmatrix $D = \text{diag}(d_1, \dots, d_n)$, $d_i > 0$,
 so daß $A = L D^t L$

b) Für $\tilde{L} = L \cdot D^{1/2}$, $D^{1/2} = \text{diag}(\sqrt{d_1}, \dots, \sqrt{d_n})$
 gilt dann

$$A = \tilde{L} \cdot {}^t \tilde{L} \quad (\text{Cholesky-Zerlegung})$$

Beweis:

a) $A > 0 \Rightarrow a_{ii} > 0$

$$A = \begin{pmatrix} a_{11} & z^t \\ z & * \end{pmatrix} \quad L_1 = \begin{pmatrix} 1 & & 0 \\ z/a_{11} & 1 & \\ \vdots & & \ddots \\ -z_n/a_{11} & & & 1 \end{pmatrix}$$

$$L_1 A = \begin{pmatrix} a_{11} & {}^t z \\ 0 & * \end{pmatrix} \quad L_1 A {}^t L_1 = \begin{pmatrix} a_{11} & 0 \\ 0 & * \end{pmatrix}$$

Induktion: $\tilde{L} = L_{n-1} L_{n-2} \dots L_1$

$$\tilde{L} A {}^t \tilde{L} = \begin{pmatrix} a_{11} & & 0 \\ & d_2 & \\ 0 & & \ddots \\ & & & d_n \end{pmatrix} = D$$

$$L = \tilde{L}^{-1}, \Rightarrow A = L D^t L$$

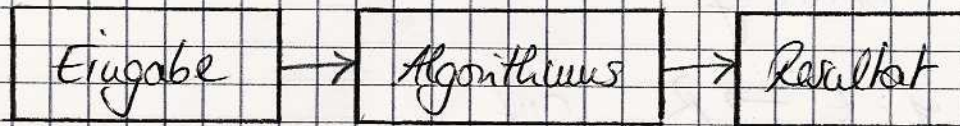
b) $D = D^{1/2} \cdot D^{1/2}$

$$\tilde{L} = L D^{1/2} \Rightarrow {}^t \tilde{L} = {}^t D^{1/2} {}^t L$$

$$\Rightarrow \tilde{L} {}^t \tilde{L} = L D^{1/2} D^{1/2} {}^t L = A$$

5. Fehleranalyse

Die numerische Behandlung eines Problems besteht aus



Fehler im Resultat haben zwei Ursachen:

- a) Fehler in der Eingabe (unvermeidlich)
- b) Fehler im Algorithmus

Die Fehler in a) werden wir durch "Kondition" des Problems messen, Fehler in b) durch die Stabilität des Algorithmus.

Gleitkommazahl

$$fl(x) = \pm a \cdot d^e \quad x \in \mathbb{R}$$

$e \in \{e_{\min}, \dots, e_{\max}\}$ Exponent

$$d \in \mathbb{N}_{\geq 2}$$

$$a = \sum_{j=1}^{\ell} a_j d^j \quad d\text{-adischer Bruch } (< 1)$$

ℓ Schranke

Relativer Fehler

$$\frac{|x - fl(x)|}{|x|} \leq \epsilon_{ps} = \frac{d^{-\ell+1}}{2}$$

float in C

$$\epsilon_{ps} \approx 10^{-7}$$

⑥ Kondition eines Problems

Wir fassen einen Algorithmus für ein Problem als eine Abbildung

$$f: E \rightarrow \mathbb{R} \quad \text{auf} \\ E, \mathbb{R} \subset \mathbb{R}^n, \mathbb{R}^m.$$

Definition: (absolute, relative Kondition)

Die absolute, normweise Kondition von f im Eingabepunkt x ist die kleinste Zahl

$K_{\text{abs}} \geq 0$, so dass

$$\|f(\tilde{x}) - f(x)\| \leq K_{\text{abs}} \cdot \|\tilde{x} - x\| + o(\|\tilde{x} - x\|)$$

für $\tilde{x} \rightarrow x$.

Die relative Kondition ist das kleinste K_{rel} ,

so dass

$$\frac{\|f(\tilde{x}) - f(x)\|}{\|f(x)\|} \leq K_{\text{rel}} \frac{\|\tilde{x} - x\|}{\|x\|} + o\left(\frac{\|\tilde{x} - x\|}{\|x\|}\right)$$

für $\tilde{x} \rightarrow x$

Bemerkung:

Ist $f: E \rightarrow \mathbb{R}$ diffbar in x , dann gilt nach MWS

$$K_{\text{abs}} = \|\mathcal{D}f(x)\|$$

↑
Matrix (norm)

$$K_{\text{rel}} = \frac{\|x\|}{\|f(x)\|} \cdot \|\mathcal{D}f(x)\|$$

⑦. Beispiel Kondition der Addition

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(a, b) = a + b$$

$$Df(a, b) = (1, 1)$$

$$K_{\text{abs}} = 1; \text{ wenn } \|x\|_1 = \sum_{j=1}^n |x_j|$$

$$K_{\text{rel}} = \frac{|a| + |b|}{|a + b|}$$

$$\Rightarrow K_{\text{rel}} \gg 1, \text{ wenn } a \approx -b$$

$$K_{\text{rel}} = 1, \text{ wenn } \operatorname{sgn}(a) = \operatorname{sgn}(b)$$

Bei Subtraktion gleichgroßer Zahlen haben wir Auslöschung

$$\begin{array}{r} 0.123456 * \\ - 0.123345 * \\ \hline 0.000111 \end{array} \quad \begin{array}{l} \} 7 \text{ Stellen genau} \\ \rightarrow 0.111 * 10^{-3} \} 3 \text{ Stellen genau} \end{array}$$

$\Rightarrow$ Subtraktion gleichgroßer Zahlen möglichst vermeiden

⑧. Beispiel

Quadratische Gleichung

$$x^2 - 2px + q = 0$$

$$\text{Lösung: } x_{1,2} = p \pm \sqrt{p^2 - q}$$

Ist eine Lösung nahe bei Null, dann haben wir Auslöschung.

Bessere Formel:

$$x_1 = p + \operatorname{sgn}(p) \sqrt{p^2 - q}$$

$$x_2 = q/x_1$$

keine Austöschung

⑨ Beispiel Kondition von $Ax = b$

a) A fest, b variabel

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$b \mapsto A^{-1} \cdot b$$

diffbar $Df = A^{-1}$

$$\Rightarrow K_{\text{abs}} = \|A^{-1}\|, \quad K_{\text{rel}} = \frac{\|b\|}{\|A^{-1}b\|} \cdot \|A^{-1}\| = \frac{\|Ax\|}{\|x\|} \cdot \|A^{-1}\|$$

b) b fest, $A \in GL(n, \mathbb{R})$ variiert

$$f: GL(n, \mathbb{R}) \rightarrow \mathbb{R}^n$$

$$A \mapsto A^{-1} \cdot b$$

⑩ Lemma: Die Abbildung

$$g: GL(n, \mathbb{R}) \rightarrow GL(n, \mathbb{R}) \subset \mathbb{R}^{n^2}$$
$$A \mapsto g(A) = A^{-1}$$

ist diffbar mit

$$(Dg)(A): \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2} = M(n \times n, \mathbb{R})$$

ψ

$$C \mapsto -A^{-1}CA^{-1}$$

Beweis:

$$(\mathcal{D}g)(A)(C) = \left. \frac{d}{dt} (A + tC)^{-1} \right|_{t=0}$$

Richtungsableitung
von g in A in Richtung C

$$E_n = (A + tC)(A + tC)^{-1}$$

$$\Rightarrow 0 = C(A + tC)^{-1} + (A + tC) \frac{d}{dt} (A + tC)^{-1}$$

$$t=0 :$$

$$0 = C \cdot A^{-1} + A \cdot (\mathcal{D}g)(A)(C)$$

$$\Rightarrow \mathcal{D}g(A)(C) = -A^{-1}CA^{-1}$$

$$\Rightarrow K_{\text{abs}} = \|\mathcal{D}f(A)b\| = \sup_{\|c\|=1} \|A^{-1}CA^{-1}b\|$$

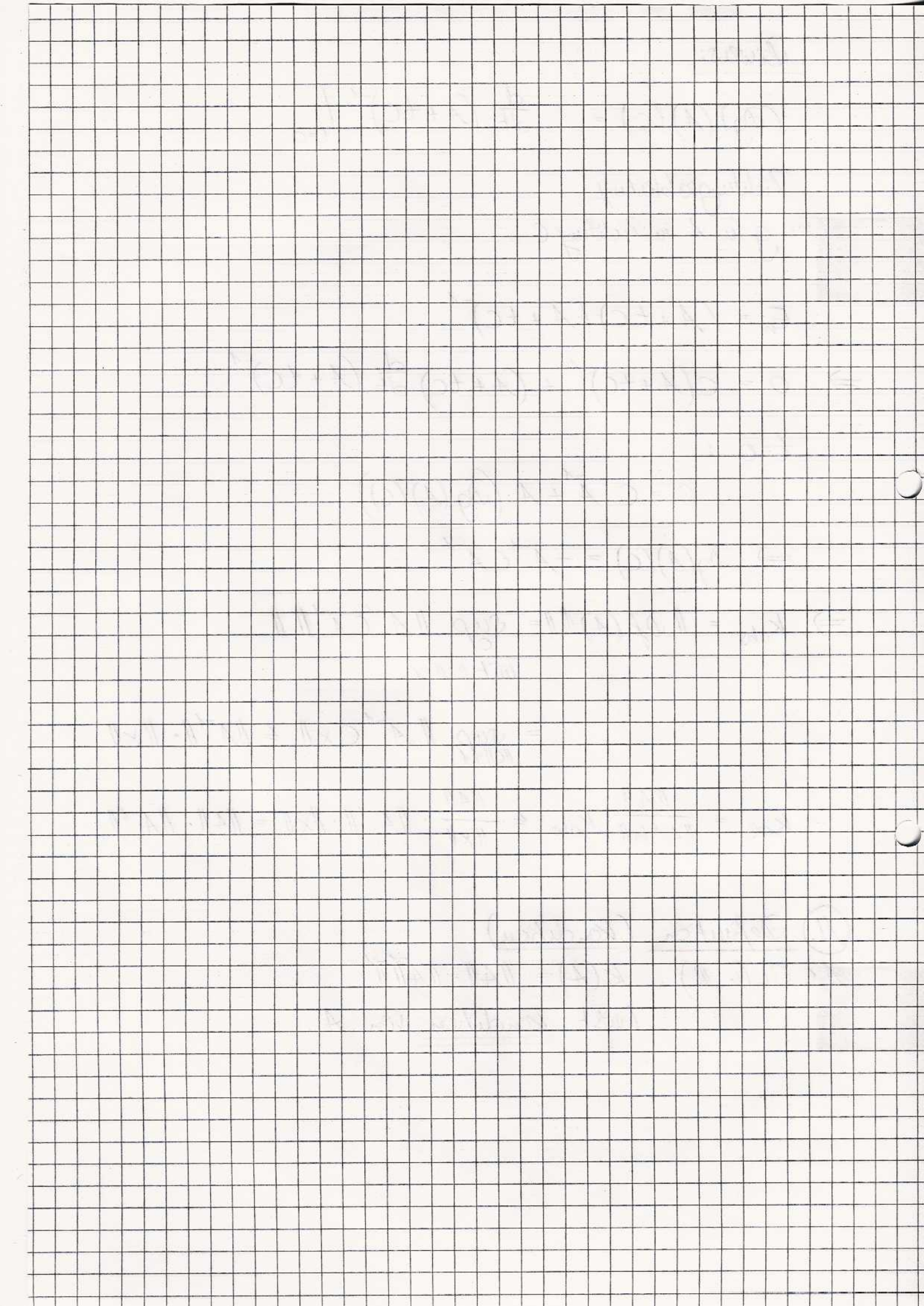
$$= \sup_{\|x\|=1} \|A^{-1}Cx\| \leq \|A^{-1}\| \cdot \|x\|$$

$$K_{\text{rel}} = \frac{\|A\|}{\|A^{-1}b\|} K_{\text{abs}} \leq \frac{\|A\|}{\|x\|} \cdot \|A^{-1}\| \cdot \|x\| = \|A\| \cdot \|A^{-1}\|$$

11. Definition (Kondition)

$$A \in GL(n, \mathbb{R}) \quad K(A) = \|A\| \cdot \|A^{-1}\|$$

heißt Kondition von A .



$$K(A) = \|A\| \cdot \|A^{-1}\|$$

Kondition einer Matrix

$$K_{rel} \leq K(A)$$

$$f(A) := A^{-1}b = x \quad \text{Lösung von } Ax=b, b \text{ fest}$$

Dann misst K_{rel} den unvermeidlichen Fehler durch Runden

$$\frac{\|f(\tilde{A}) - f(A)\|}{\|f(A)\|} \leq K_{rel} \frac{\|\tilde{A} - A\|}{\|A\|}$$

1.2) Bemerkung

1) Für $\|\cdot\| = \|\cdot\|_2$ gilt

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \lambda_{\max}$$

wobei $\lambda_{\max} = \max \{ |\lambda| \mid \lambda \in \text{EW von } A \}$

$$\|A^{-1}\| = \min_{\|x\|=1} \|Ax\| = \lambda_{\min}$$

$$K(A) = \frac{\lambda_{\max}}{\lambda_{\min}} = \max_{\|x\|=1} \|Ax\| / \min_{\|x\|=1} \|Ax\|$$

2) Insbesondere $K(A) \geq 1$

und $K(A) = 1$ gilt genau dann wenn

$$A = a \cdot Q \text{ mit } a \in \mathbb{R}, Q \in O(n)$$

Beweis: $\max_{\|x\|=1} (\|Ax\|) = \min_{\|x\|=1} (\|Ax\|) = a$

$$\Rightarrow \underbrace{\frac{1}{a}}_Q \cdot Ax = 1 \quad \forall x \text{ mit } \|x\|=1$$

so gilt: $\|Qx\| = \|x\| \quad \forall x \in \mathbb{R}^n$

$\Rightarrow Q \in O(n)$ (ist orthogonal)

13) Stabilitätsindikator

Sei $f: E \rightarrow \mathbb{R}$ Algorithmus

$\tilde{f}: E \rightarrow \mathbb{R}$ Gleitkomma-Realisierung

Wir messen die Stabilität von $\tilde{f}$ in Termen, des unvermeidlichen Fehlers, der durch die Kondition beschrieben wird.

Definition: (Stabilitätsindikator von $\tilde{f}$ und x)

Sei $\tilde{f}$ eine Gleitkomma-Realisierung von f .

Der (normweise) Stabilitätsindikator ist die kleinste Zahl $\sigma \geq 0$, so dass

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \leq \sigma \cdot \kappa \cdot \text{eps} + o(\text{eps})$$

für $\text{eps} \rightarrow 0$.

14) Lemma

Für die Elementaroperationen $\circ \in \{+, -, \cdot, /\}$ und deren Gleitkomma-Realisierung $\hat{\circ} \in \{\hat{+}, \hat{-}, \hat{\cdot}, \hat{/}\}$ gilt:

$$\sigma \cdot \kappa \leq 1.$$

Beweis:

$$a, b, \hat{\circ} \in \{\hat{+}, \hat{-}, \hat{\cdot}, \hat{/}\}$$

$$a \hat{\circ} b = a \circ b (1 + \varepsilon) \quad \text{mit } |\varepsilon| \leq \text{eps}$$

Also

$$\frac{|a \hat{\circ} b - a \circ b|}{|a \circ b|} = \frac{|\varepsilon (a \circ b)|}{|a \circ b|} = |\varepsilon| \leq \text{eps} \Rightarrow \sigma \kappa \leq 1$$

Bei der Analyse der Stabilität ist es oft zweckmäßig, den Algorithmus in Schritte zu zerlegen.

$$f = h \circ g : \mathbb{R}^n \xrightarrow{g} \mathbb{R}^l \xrightarrow{h} \mathbb{R}^m$$

15. Lemma

Sei $\tilde{f} = \tilde{h} \circ \tilde{g}$ die Realisierung als Komposition zweier Realisierungen. Dann gilt:

$$T_{\tilde{f}} \cdot K_f \leq T_{\tilde{h}} \cdot K_h + T_{\tilde{g}} \cdot K_g \cdot K_h$$

Beweis:

$$\begin{aligned} \|\tilde{f}(x) - f(x)\| &= \|\tilde{h}(\tilde{g}(x)) - h(g(x))\| \\ &\leq \|\tilde{h}(\tilde{g}(x)) - h(\tilde{g}(x))\| + \|h(\tilde{g}(x)) - h(g(x))\| \\ &\leq T_{\tilde{h}} \cdot K_h \|\tilde{g}(x) - g(x)\| \text{eps} + K_h \cdot \frac{\|\tilde{g}(x) - g(x)\|}{\|g(x)\|} \|h(g(x))\| \\ &\leq T_{\tilde{h}} K_h \|f(x)\| \text{eps} + K_h T_{\tilde{g}} K_g \text{eps} \cdot \|f(x)\| + o(\text{eps}) \end{aligned}$$

$$\Rightarrow \frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \leq (T_{\tilde{h}} \cdot K_h + K_h T_{\tilde{g}} K_g) \cdot \text{eps}$$

→ unvermeidliche Subtraktionen in einem Algorithmus möglichst am Anfang ausführen.

16. Beispiel

Summation $S_n : \mathbb{R}^n \rightarrow \mathbb{R}$
 $x_1, \dots, x_n \mapsto \sum_{i=1}^n x_i$

Naiver Algorithmus ist rekursiv

$$S_n = S_{n-1} \circ \alpha_n, \quad \alpha_n : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$$

$$(x_1, \dots, x_n) \mapsto (x_1 + x_2, x_3, \dots, x_n)$$

$$K_n = K_1, \quad \sigma_n = \sigma_1$$

stimmt mit der Kondition und Stabilitätsindex Addition überein:

Mit Lemma 15 und Rekursion

$$K_j = K_{S_j}, \quad \sigma_j = \sigma_{S_j}$$

$$\sigma_n K_n \leq \sigma_{n-1} K_{n-1} + \sigma_1 K_1 \cdot K_{n-1} \leq (1 + \sigma_{n-1} \cdot K_{n-1}) \leq (1 + \sigma_{n-1}) K_n$$

Nun:

$$K_n = \frac{\sum_{i=1}^n |x_i|}{\left| \sum_{i=1}^n x_i \right|} \geq 1$$

$$K_{n-1} = \frac{|x_1 + x_2| + \sum_{i=3}^n |x_i|}{\left| \sum_{i=1}^n x_i \right|} \leq K_n$$

$$\Rightarrow \sigma_n \leq 1 + \sigma_{n-1}$$

$$\Rightarrow \sigma_n \leq n-1$$

17) Definition (numerisch stabil)

Eine Realisierung $\tilde{f}$ von f heißt numerisch stabil, wenn $\sigma_{\tilde{f}} \leq n$, wobei n die Anzahl der Elementaroperationen von f ist.

M stochastisch, $\lambda=1$ einziger EW von M in $|\lambda|=1$, einfach, $\text{Eig}(M, 1) = \left\langle \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix} \right\rangle \subseteq \sum z_i = 1$

Satz: $\Rightarrow \lim_{n \rightarrow \infty} M^n = \begin{pmatrix} z_1 & \dots & z_1 \\ \vdots & & \vdots \\ z_k & \dots & z_k \end{pmatrix}$

Bei der Komposition von Skalarfunktionen lässt sich das Lemma 15 verbessern.

18. Lemma:

$f = h \circ g$ Komposition Skalarfunktionen $\tilde{f} = \tilde{h} \circ \tilde{g}$

Dann gilt

$$\sigma_f = \frac{\sigma_h}{K_g} + \sigma_g$$

Beweis:

In diesem Fall ist wegen Kettenregel die Kondition der Komposition

$$K_f = \frac{|x| \cdot |f'(x)|}{|f(x)|} \stackrel{\text{Kettenregel}}{=} \frac{|h'(g(x))| |g'(x)|}{|h(g(x))|} \cdot \frac{|g'(x)| |x|}{|g(x)|} \\ = K_h \cdot K_g$$

Lemma 15

$$\sigma_f \cdot K_f = \sigma_f \cdot K_h \cdot K_g \leq \sigma_h \cdot K_h + \sigma_g K_g K_h \\ \Rightarrow \sigma_f \leq \frac{\sigma_h}{K_g} + \sigma_g$$

19. Beispiel

Auswertung von trigonometrischen Polynomen

$$f(x) = \sum_{k=1}^n a_k \cos(kx) + b_k \sin(kx)$$

Wir wollen alle $\cos(kx)$ berechnen.

Verwenden Rekursionsformel

$$\cos(k+1)x = 2 \cos kx \cdot \cos x - \cos(k-1)x$$

Additionstheorem $\Rightarrow \cos(k+1)x = \cos kx \cdot \cos x - \sin kx \cdot \sin x$

$$\cdot \sin x \parallel 0 = \sin kx - \sin(k-1)x \cos x - \cos(k-1)x \sin x$$

$$\cdot \cos x \parallel 0 = \cos kx - \cos(k-1)x \cdot \cos x + \sin(k-1)x \sin x$$

$$\cos(k+1)x = 2 \cos x \cdot \cos kx - \cos kx$$

Wir setzen $C_n = \cos(nx)$

Rekursionsformel

$$C_{k+1} = 2 \cos x C_k - C_{k-1}$$

Stabilität: $g = \cos x$ geht weiterhin ein.

$\Rightarrow$ Faktor $1/k_g$ in der Formel für σ .

$$\frac{1}{k_g} = \frac{|\cos x|}{|\sin x| \cdot |x|} = \frac{1}{\underbrace{x \tan x}_{\approx x^2}} \xrightarrow{x \rightarrow 0} \infty$$

$\Rightarrow$ Diese Rekursion ist schlecht für x nahe 0 (oder π)

$$\begin{aligned} \text{Besser: } \cos(x) &= \cos\left(\frac{x}{2} + \frac{x}{2}\right) = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2} \\ &= 1 - 2 \sin^2 \frac{x}{2} \end{aligned}$$

$$\Delta_{C_k} = C_{k+1} - C_k = \cos(k+1)x - \cos kx$$

Rekursion

$$\begin{cases} \Delta_{C_k} = -4 \sin^2 \frac{x}{2} C_k + \Delta_{C_{k-1}} \\ C_{k+1} = C_k + \Delta_{C_k} \end{cases}$$

$$g(x) = \sin^2\left(\frac{x}{2}\right) \quad g'(x) = \sin \frac{x}{2} \cdot \cos \frac{x}{2}$$

$$\frac{1}{k_g} = \frac{|\sin^2(\frac{x}{2})|}{|x \sin \frac{x}{2} \cos \frac{x}{2}|} = \frac{1}{\frac{\tan \frac{x}{2}}{x}} \xrightarrow{x \rightarrow 0} \frac{1}{2}$$

gut konditioniert
nahe $x=0$