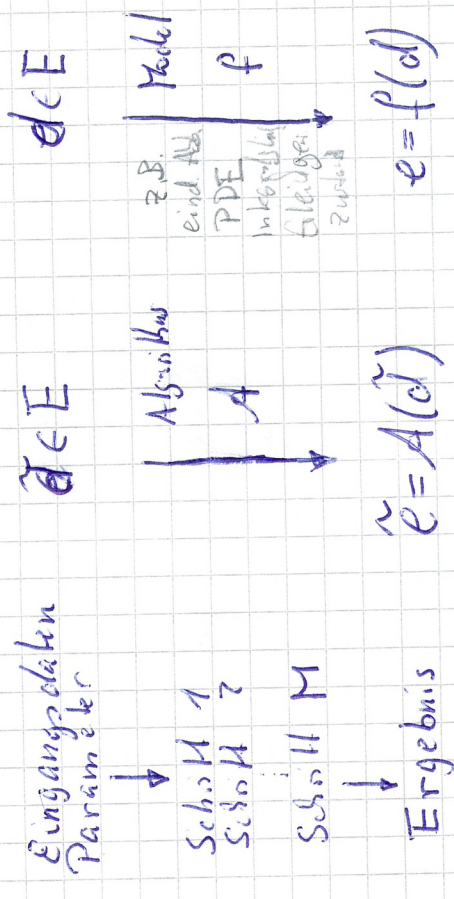


# A. Grundbegriffe und Computermathematik

## A.1 Algorithmen und Fehlerfortpflanzung

Def. Ein Algorithmus ist eine eind. Vordr. bestehend aus endlich vielen Schritten

Zur Lsg eines Problems  
allgemeines Aufgaben:



Bsp.: Berechne den Luftwiderstand eines Autos im ~~Stromungskanal~~

Eingangsdaten:  
Form des Autos  
Form des ~~Stromungskanal~~ Umgebungs  
Geschwindigkeit des Autos

Ergebnis: Luftwiderstand

## Probleme

- die Form des Autos lässt sich nur approximativ beschreiben
- die Form der Umgebung ist u.U. unbekannt
- das Ergebnis lässt sich nicht durch eine Formel berechnen, sondern
- \* die ~~Luft~~ Bewegung der Luft muss in jedem Punkt simuliert werden
- jeder Punkt geht nicht da nur endlich viele Punkte im Computer darstellbar

→ ein Algorithmus kann das Teil in der Bed nur approximativ umsetzen.

## Anforderungen an einen Algorithmus

1. Genauigkeit  $A(d) \approx f(d)$
2. Geschwindigkeit
3. Stabilität  $A(d) \approx \underbrace{A(d)}_{=d}$
4. Einfachheit

## Ein Maß für die Geschwindigkeit

Die Numerische Komplexität beschreibt die Abhängigkeit zwischen der Größe des Eingangsdaten, des Ergebnisses und der Anzahl der Perforationsstellen.

Rechenoperationen - Anzahl der Additionen und Multiplikationen.  
Floating Point Operations.

Man betrachte nun nur die Größenordnung

Bsp:  $\mathbb{F} \rightarrow$  Bildverarbeitung

$N$  - Anzahl der Eingabebilder

$M$  - Anzahl der Ausgabebilder

Dann bedeutet die reine Komplexität

$O(N+M)$  - das Alg. ist linear in  $N$  und  $M$

$O(N^2+M^3)$  - das Alg. ist quadratisch in  $N$  und kubisch in  $M$

$O(N \log N)$  - das Alg. ist stabil sortiert linear in  $N$

$O(2^N)$  - Exponentiel, Traveling Salesman

# Maße für die Genauigkeit

## • absolute Fehler

$|x - \tilde{x}|$  für einfache Zahlen  $x, \tilde{x}$

für Vektoren betrachtet man Normen

$$\underline{x} = (x_1, \dots, x_N), \quad \tilde{\underline{x}} = (\tilde{x}_1, \dots, \tilde{x}_N)$$

$$\|\underline{x} - \tilde{\underline{x}}\|_{\infty} = \max_{i=1, \dots, N} |x_i - \tilde{x}_i| \quad \text{Sup-Norm}$$

$$\|\underline{x} - \tilde{\underline{x}}\|_p = \left( \sum_{i=1}^N |x_i - \tilde{x}_i|^p \right)^{1/p}$$

Bsp.  $p=2 \rightarrow$  Euklidischer Abstand

$p=1 \rightarrow$  Manhattan Abstand

$p=\infty \rightarrow$  Maximum Abstand

## • relatives Fehler

$$\frac{|x - \tilde{x}|}{|x|} \quad \text{bzw.} \quad \frac{\|\underline{x} - \tilde{\underline{x}}\|}{\|\underline{x}\|} \quad x \neq 0$$

$x$  - wahres Wert

$\tilde{x}$  - Approximation

Vorteile gegenüber dem Absolut Fehler:

- unabhängig von der Größe der Einheits
- dimensionslos

Ende  
1. Vorlesung

## Maße für die Stabilität

Def. Ein Model  $e = f(d)$  heißt gut konditioniert, falls kleine Änderungen

in den Daten  $\vec{d} = d + \Delta d$  nur zu kleinen Änderungen im Ergebnis  $f(\vec{d})$  führen, also

$$\frac{\|d - \vec{d}\|}{\|d\|} < \delta \Rightarrow \frac{\|f(d) - f(\vec{d})\|}{\|f(d)\|} < \varepsilon$$

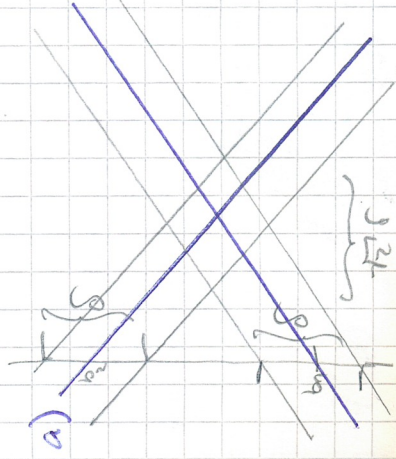
Der maximale Quotient aus  $\varepsilon/\delta$  heißt

relative Kondition  $K$  von  $f$ , also

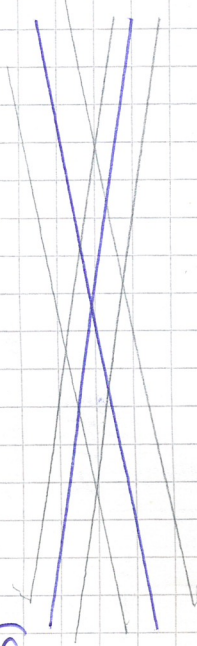
$$K = \sup_{d \neq \vec{d}} \frac{\|f(d) - f(\vec{d})\|}{\|f(d)\|} / \frac{\|d - \vec{d}\|}{\|d\|}$$

Bsp. Bestimme den Schnittpunkt zweier

Geraden



b)



Je parallel die Geraden werden  
desto größer wird das Verhältnis  
zwischen  $\delta$  und  $\epsilon$ , d.h.  $\kappa \rightarrow \infty$   
Für immer schlechtere Geraden.

Def. Ein Algorithmus heißt stabil,  
falls kleine Änderungen in den Eingangs-  
daten zu den gleichen Änderungen in  
den Ausgangsdaten führen wie beim

Zugehörig Modell, d.h.

$$\frac{\|d - \tilde{d}\|}{\|d\|} < \delta \Rightarrow \frac{\|A(d) - A(\tilde{d})\|}{\|A(d)\|} < \kappa \cdot \delta$$

Bsp. Berechne  $f(d) = \frac{1 - \cos d}{d}$

Kondition:

$$\kappa \approx \left| \frac{f'(d) \cdot d}{f(d)} \right|$$

$$f'(d) = \frac{\sin d}{1 - \cos d} = 0$$

↙ eigenwert

$$f'(d) = \frac{d \sin d - 1 + \cos d}{1 - \cos d}$$

das absuchen

für kleine  $d$ :

$$\begin{aligned}\lim_{d \rightarrow 0} k(d) &= \lim_{d \rightarrow 0} \left| \frac{d \sin d}{1 - \cos d} - 1 \right| \\ &= \lim_{d \rightarrow 0} \left| \frac{\sin d + d \cos d}{\sin d} - 1 \right| \\ &= \lim_{d \rightarrow 0} \left| \frac{d \cos d}{\sin d} \right| \\ &= \lim_{d \rightarrow 0} \left| \frac{\cos d - d \sin d}{\cos d} \right| = 1\end{aligned}$$

→ Das Modell (Problem) ist gut konditioniert.

Aber: Direkte Implementierung führt zu einem instabilen Algorithmus

Ende  
2. Vorlesung

Besser: Reihenentwicklung (Taylor)

$$\begin{aligned}\frac{1 - \cos d}{d} &= \frac{1}{d} \left( 1 - \sum_{k=0}^{\infty} \frac{(-1)^k d^{2k}}{(2k)!} \right) \\ &= \frac{1}{d} \left( 1 - \left( 1 - \frac{d^2}{2} + \frac{d^4}{24} - \dots \right) \right)\end{aligned}$$

wir setzen  $A(d) = \frac{d}{2}$  für  $d < \varepsilon$

$$\Rightarrow \left| \frac{1 - \cos d}{d} - \frac{d}{2} \right| \leq \frac{d^3}{24}$$

Leibniz Regel

für  $\varepsilon \rightarrow 0$  ist  $A(d) \approx \frac{1 - \cos d}{d}$

Bsp. Quadratische Gleichung  $x^2 - px + q = 0$

$$f(p, q) = \left( \frac{p - \sqrt{p^2 - 4q}}{2}, \frac{p + \sqrt{p^2 - 4q}}{2} \right), \quad f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

Eine Nullstelle nahe Null haben wir für

$p \gg q$ . Ein naives Berechnen von

$p - \sqrt{p^2 - 4q}$  führt zu großen Fehlern

Besser:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{q}{x_2} \\ x_2 = p + \sqrt{p^2 - 4q} \end{pmatrix}$$

$$\text{Denn } \frac{q}{p + \sqrt{p^2 - 4q}} = \frac{q(p - \sqrt{p^2 - 4q})}{p^2 - p^2 + 4q} = \frac{q(p - \sqrt{p^2 - 4q})}{4q}$$

$$= p - \sqrt{p^2 - 4q}$$



## 1.2. Computeralithmetik

Problem: Es gibt unendlich viele Zahlen.

Der Computer kann aber nur endlich viele Zustände speichern. Daher muß man sich auf eine Zuordnung zwischen einer endlich an Zahlmenge und den verfügbaren Zuständen einigen.

Möglichkeit 1 nicht negative ganze Zahl

$$X = [x_7 x_6 x_5 x_4 x_3 x_2 x_1 x_0]_2 \\ = x_7 2^7 + x_6 2^6 + x_5 2^5 + \dots + x_0 2^0$$

darstellbar  
Bereich  $0 \dots 2^8 - 1$

Möglichkeit 2 ganze Zahlen

$$X = [x_7 x_6 \dots x_0]_2 \\ = -x_7 2^7 + x_6 2^6 + x_5 2^5 + \dots + x_0 2^0$$

darstellb. Bereich:  $-2^{N-1} \dots 2^{N-1} - 1$

Bsp.:  $5 = [00000101]_2$

$-5 = [11111011]_2$

$0 = [0 \dots 0]_2$

Möglichkeit 3: Fixpunktzahlen

$$\begin{aligned}
 x &= [x_7 \cdot x_6 \cdot x_5 \cdot x_4 \cdot x_3 \cdot x_2 \cdot x_1 \cdot x_0]_2^{-7} \\
 &= -x_7 \cdot 2^0 + x_6 \cdot 2^{-1} + x_5 \cdot 2^{-2} + \dots + x_0 \cdot 2^{-7}
 \end{aligned}$$

darstellbarer Bereich:  $2^{-7} \cdot \{ \dots - 2^{N-1} \dots 2^{N-1} - 1 \}$

Bsp:  $\frac{5}{16} = \frac{1}{16} + \frac{4}{16} = 2^{-4} + 2^{-2}$   
 $= [0.0101000]_2$

Möglichkeit 4: Gleitpunktzahlen

$$\begin{aligned}
 x &= \underbrace{[x_4 x_3 x_2 x_1 x_0]}_{\text{Mantisse}} \mid \underbrace{[e_2 e_1 e_0]}_{\text{Exponent}}_2 \\
 &= [x_4 x_3 x_2 x_1 x_0]_2 \cdot 2^{[e_2 e_1 e_0]_2}
 \end{aligned}$$

darstellbarer Bereich:  $-2^3 \dots -2^{-8}$   
 $2^{-8} \dots 2^3 - 1$   
 0

Ende:  
2. Vorlesung

Achtung: einige Zahlen haben mehrere Darst. als Gleitpunktzahl.

Bsp:  $[0.10001000]_2 = [0.010010001]_2$   
 $= [0.00101010]_2$   
 $= [0.00011011]_2$   
 $= \frac{1}{2}$

Ausgeg: normalisierte Gleichkommazahlen

$$X = [\pm 1 \cdot X_3 X_2 X_1 X_0 | e_2 e_1 e_0]_2$$
$$= \pm [1 \cdot X_3 X_2 X_1 X_0]_2 \cdot 2^{[e_2 e_1 e_0]_2}$$

das ist die beste Darstellung da mit den vorzeichen Nachkommastellen

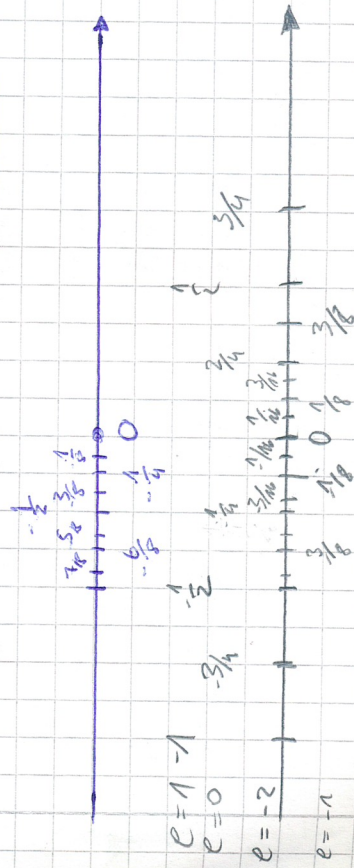
die 0 lässt sich nicht als normalisierte Gleichkommazahl schreiben

→ Annahmen für

$\{e_2 e_1 e_0\} = \text{Dop0}$  → denormalisiertes GPZ inkl. Null

$\{e_2 e_1 e_0\} = [111] \rightarrow \pm \infty, NaN$

Anordnung der Gleichkommazahl



Maximaler Wert 3 → 8 Ziffern  $0.01 = \frac{1}{4}$   
Exponenten für 2:  $-2 - 1 - 0 - 1$   
nicht normalisiert

## Rundungsfehler:

Für  $x \in \mathbb{R}$  bezeichne  $[x]$  die Computerzahl, welche  $x$  am nächsten ist.

Für jede normierte Gleitpunktzahl gilt und für  $x$  im darstellbaren Bereich

$$\frac{|x - [x]|}{|x|} = \frac{(10^{\lfloor x_{n-1} x_0 \rfloor} - 1) \cdot 10^{-n}}{(10^{\lfloor x_{n-1} x_0 \rfloor} - 1) \cdot 10^{-n}} \cdot 2^p$$
$$\leq 2^{-n-1} = \varepsilon$$

Bem

① Die Maschinengenauigkeit ist nur von der Mantissenlänge abhängig.

Für den absoluten Fehler gilt

②  $A = 2^{-n}$  Bei den Gleitkommazahlen

ist der relative Rundungsfehler über den gesamten darstellbaren Bereich gleichmäßig beschränkt - der absolute Rundungsfehler ist umso größer je größer  $|x|$  ist.

Bei Fixpunktzahlen ist der absolute Rundungsfehler über den gesamten darstellbaren Bereich gleichmäßig beschränkt während der

relative Rundungsfehler für einen  $x$

Stoß wird.

$$\textcircled{3} \underbrace{(1-2^{-n-1})}_{1-\varepsilon} x \leq [x] \leq \underbrace{(1+2^{-n-1})}_{1+\varepsilon} x$$

④ Die Maschinengenauigkeit  $\varepsilon$  ist die

kleinste Zahl für die gilt

$$[1+\varepsilon] > 1$$

d.h. für alle  $\gamma$  mit  $[y] < [E]$

$$\text{gilt } [1+\gamma] = 1$$

Internationale Konventionen: IEEE 754

a) einfache Genauigkeit 32 bit

1 bit Vorzeichen  
23 bit Mantisse  
8 bit Exponent

$\varepsilon \approx 10^{-7}$  Überlände häufig

darstellb. Bereich:  
 $10^{-45}$  -  $10^{38}$  Überlände selten

b) doppelte Genauigkeit 64 bit

1 bit Vorzeichen  
52 bit Mantisse  
11 bit Exponent

$\varepsilon \approx 10^{-16}$

P = -1022 - 1023  
darst. Bereich:  $10^{-308}$  -  $10^{308}$

## Grundrechenarten

In Computeralgebra gehen nicht die üblichen Rechenoperationen, sondern

$$X \oplus Y = [X+Y], [X \ominus Y] = [X-Y]$$

$$X \otimes Y = [X \cdot Y], X \oslash Y = [X/Y]$$

Falls es zu Reinen über od. Unterlauf Rand,

d.h. falls das Ergebnis eine normierte

Einheitsformzahl ist gilt

$$\left| \frac{X+Y-(X \otimes Y)}{1+Y} \right| < \varepsilon$$

analog für die anderen Grundrechenarten

→ Die Grundrechenarten sind im Sinne des Computerjournalist sein.

Kondition

$$K = \left| \frac{X+Y-(\tilde{X}+\tilde{Y})}{1+Y} \right| / \left| \frac{X-\tilde{X}}{1+Y} \right|$$
$$= \frac{\left| \frac{X-\tilde{X}}{1+Y} \right|}{\left| \frac{X-\tilde{X}}{1+Y} \right|} = \frac{1}{1+Y}$$

→ Für  $X \approx -Y$  ist die Abschätzung

beliebig konditioniert. → Auslöschung

Bsp. 0 Basis 10, Normierung  $\rightarrow$

$$x = 1.234567 \rightarrow [x] = 1.23457 \rightarrow \varepsilon_x = \left| \frac{x - [x]}{x} \right| \approx 3 \cdot 10^{-6}$$

$$y = -1.23457 \rightarrow [y] = -1.23457 \rightarrow \varepsilon_y = 0$$

$$x+y = -3 \cdot 10^{-6}, [x]+[y] = 0, \rightarrow \varepsilon_{x+y} = \left| \frac{x+y - ([x]+[y])}{x+y} \right| = 1$$

$\rightarrow$  keine exakte Darstellung mehr

$\rightarrow$  keine Rundungsfehler vermeiden

Die übliche arithmetischen Gesetze  
nicht z.P.

$$\bullet x \oplus z = x \quad \forall x \geq 1 \text{ u. } |z| < \varepsilon$$

$$\bullet (x \oplus y) \oplus z \neq x \oplus (y \oplus z)$$

$$\textcircled{2} f(x) = \begin{cases} 1 - \cos x & x \neq 0 \\ 0 & x = 0 \end{cases}$$

### ③ Quadratische Gleichung:

Bestimme NS von

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

$$p, q \geq 0$$

$$f(p, q) = \left( p + \sqrt{p^2 - q} \right) \cdot \mathbb{R}_+^2 \rightarrow \mathbb{R}^2$$

~~Eine~~ Für  $p \gg q$  ist  $\sqrt{p^2 - q} \approx p$ ,

d.h. bei der Berechnung der 2. NS

trifft Auslöschung auf, d.h. das Fehler,

das bei der Wurzelberechnung gemacht

Wird vergrößert.

$$\text{Besser: } \tilde{f}(p, q) = \begin{pmatrix} x_1 = p + \sqrt{p^2 - q} \\ x_2 = \frac{q}{x_1} \end{pmatrix}$$

$$\text{Dann: } \frac{q}{p + \sqrt{p^2 - q}} = \frac{q(p - \sqrt{p^2 - q})}{p^2 - p^2 + q} = p - \sqrt{p^2 - q}$$

→ Satz von Vieta.

④ Summiere die Zahlen  $1, \frac{1}{2}, \frac{1}{3}, \dots \rightarrow \infty$

Harmonische Reihe, Divergiert am Comp

klamm:

klamm:

Note: Summation in Computerarithmetik  
ist nicht assoziativ!



## 2. lineare Gleichungssysteme

### 2.1. Vektorraum und lineare Abb.

$\mathbb{R}^n$  - Vektorraum aller Vektoren  $x = (x_1, \dots, x_n)^t$

Operationen:  $x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)^t$$

$$\lambda x = (\lambda x_1, \lambda x_2, \dots)^t$$

$$\langle x, y \rangle = x^t y = x_1 y_1 + \dots + x_n y_n$$

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\langle x, x \rangle}$$

Veranschaulichung

Basis: Seien  $e_i = (0, \dots, 1, \dots, 0) \in \mathbb{R}^n$   
 $i = 1, \dots, n$

Dann lässt sich jeder Vektor  $x \in \mathbb{R}^n$

Schreiben als  $x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$ .

Def: Eine Abb.  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  heißt

linear, falls für alle  $x, y \in \mathbb{R}^m$  und  $\lambda \in \mathbb{R}$

gilt:  $f(x + \lambda y) = f(x) + \lambda f(y)$ .

Folgerung: Sei  $\underline{x} = x_1 e_1 + x_2 e_2 + \dots + x_m e_m$   
dann gilt:  $f(\underline{x}) = x_1 f(e_1) + \dots + x_m f(e_m)$

Das können wir als Matrix schreiben

$$A = \begin{pmatrix} | & & | \\ f(e_1) & \dots & f(e_m) \\ | & & | \end{pmatrix}$$

Damit gilt:

$$A \underline{x} = \begin{pmatrix} | & & | \\ f(e_1) & \dots & f(e_m) \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

$$= x_1 f(e_1) + \dots + x_m f(e_m) = f(\underline{x})$$

$\Rightarrow$  lineare Abb. können durch Matrizen

dargestellt werden. Umgekehrt definiert jede

Matrix  $B \in \mathbb{R}^{n \times m}$  eine lineare Abb.

$$g: \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad g(\underline{x}) = B \underline{x}$$

Merke: Willst die Matrix du erhalten

schreib die Bilder in der Spalten.

# Beispiele linearer Abb.

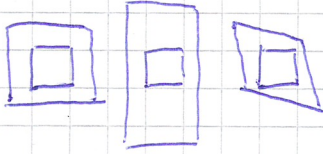
## ① Skalierung

Diagonalmatrix  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{Id}$  Identität

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$A = \begin{pmatrix} \lambda_1 & \mu \\ \mu & \lambda_2 \end{pmatrix}$$



$\Rightarrow$  Symmetrische Matrizen  $A^T = A$

## ② Rotation / Basistransformationen

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

Spiegelung

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$\Rightarrow$  Orthogonale Matrizen  $A^T = A^{-1}$

$\rightarrow A A^T = \text{Id}$   
 $A^T A = \text{Id}$   $\Rightarrow$  alle Zeilen und Spalten sind senkrecht

## ③ Projektion / Einbettung

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

$$A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$A^2 = A$$

#### 4) Ableitungsop.

• als Vektorraum betrachten wir den

Raum aller Polynome vom Grad  $n$

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 x^0$$

$$p'(x) = n \cdot a_n x^{n-1} + a_{n-2}(n-2) x^{n-2} + \dots + a_1$$

• das ist eine lin. Abb.  $\Rightarrow$  wir können

sie als Matrix schreiben.

$$\begin{pmatrix} 0 & \dots & 0 \\ n & & \\ & n-1 & \\ & & \ddots \\ & & & 1 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{pmatrix} = \begin{pmatrix} b_n \\ b_{n-1} \\ \vdots \\ b_0 \end{pmatrix}$$

$$D \cdot \underline{a} = \underline{b}$$

$$\text{mit } p'(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_0 x^0$$

Das ist die Darstellung des Ableitungs-

operator bzgl. einer ganz bestimmten

Basis - der Monombasis - andere Darst.

sind möglich u. sinnvoll,

Beachte:  $D^2$  ist einfach die zweite Ableitung

und  $(D^2 + D) \cdot a$  liefert die

Koeff. für die Summe aus  $e^{2x}$  u.

zweiter Ableit.

# Singularwertzerlegung

Satz: Jede Matrix  $A \in \mathbb{R}^{n \times m}$  lässt

sich als das Produkt

$$A = U \Sigma V^T$$

schreiben wobei:

$U \in \mathbb{R}^{n \times n}$ ,  $V \in \mathbb{R}^{m \times m}$

orthogonale Matrizen

$\Sigma \in \mathbb{R}^{n \times m}$  eine Diagonalmatrix

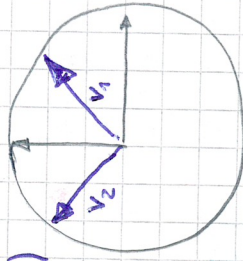
$$\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_m & \\ 0 & & & 0 \end{pmatrix} \text{ od. } \Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}$$

$$V = (v_1, \dots, v_m)$$

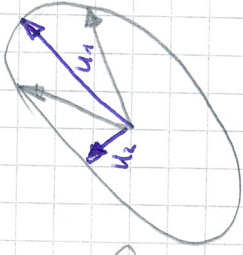
$$U = (u_1, \dots, u_n)$$

für  $n \geq m$

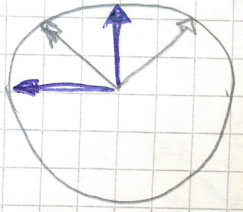
für  $m > n$



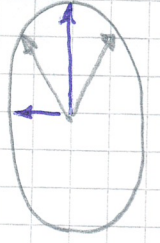
A



V



Σ



Bsp.  $A = \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}$

Wähle  $U = V = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} = \begin{pmatrix} \cos \pi/4 & -\sin \pi/4 \\ \sin \pi/4 & \cos \pi/4 \end{pmatrix}$

und  $\Sigma = \begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix}$

Dann gilt:

$$\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 8 & -4 \\ 8 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 12 & 4 \\ 4 & 12 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}$$

Bem.

- Die Diagonalelemente  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$  heißen Singulärwerte von  $A$ .

- Die Matrizen  $U, V$  sind Basisformen welche  $A$  in eine Störungsform entlang der Koordinaten transportieren

- $U = V \Leftrightarrow A$  ist symmetrisch.

- Da  $U, V$  orthogonal sind, bilden

die ~~Basen~~ Spalten  $U$ -Basen  $\Rightarrow$

links- und rechts-singuläre Vektoren.

Spezialfall

$m = n \Rightarrow$  quadratische Matrix

Def: Sei  $x \in \mathbb{R}^n \setminus \{0\}$ ,  $\lambda \in \mathbb{R}$  mit  $Ax = \lambda x$ . Dann heißt  $\lambda$  EW von  $A$  und  $x$  EV von  $A$ .

Bem: Ist  $A = U \Sigma U^T$  eine sym. Matrix und  $U = \begin{pmatrix} u_1 & \dots & u_m \end{pmatrix}$  die Spalten von  $U$ . Dann gilt:

$$\begin{aligned}
 A u_k &= U \Sigma U^T u_k \\
 &= U \Sigma \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \\ \vdots \end{pmatrix} = \sigma_k u_k \\
 &= U \Sigma \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} = \sigma_k e_k \\
 &= \sigma_k u_k = \sigma_k \begin{pmatrix} u_{1k} \\ u_{2k} \\ \vdots \\ u_{mk} \end{pmatrix} = \sigma_k \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \\ \vdots \end{pmatrix} = \sigma_k e_k \\
 &= \sigma_k u_k
 \end{aligned}$$

$\Rightarrow$  Die Spalten von  $U$  sind EV

Zu den EW  $\sigma_k$

$\Rightarrow$  Ist  $A$  sym. so besitzt  $A$  eine Basis aus Eigenvektoren. Ist  $A$  nicht sym. so muss  $A$  keine einz. EV haben.

## Berechnung von EW und EV

$$\bullet Ax = \lambda x \Leftrightarrow (A - \lambda \text{Id})x = 0$$

$$\Leftrightarrow \det(A - \lambda \text{Id}) = 0$$

$\det(A - \lambda \text{Id})$  ist ein Polynom in  $\lambda$

$\Rightarrow$  Bestimme NS  $\Rightarrow \lambda_1, \dots, \lambda_n$

$\rightarrow$  löse die LGS  $(A - \lambda_i \text{Id})x = 0$

Bsp.  $A = \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}$

$$\det \begin{pmatrix} 6-\lambda & 2 \\ 2 & 6-\lambda \end{pmatrix} = (6-\lambda)^2 - 4$$

$$= 32 - 12\lambda + \lambda^2$$

$$= (\lambda - 8)(\lambda - 4) \Rightarrow \lambda_1 = 8 \quad \lambda_2 = 4$$

EV zu  $\lambda_1 = 8$ :

$$\begin{pmatrix} 6-8 & 2 \\ 2 & 6-8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

EV zu  $\lambda_2 = 4$

$$\begin{pmatrix} 6-4 & 2 \\ 2 & 6-4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$



## Berechnung der Singularwertzerlegung

- ist  $A$  symmetrisch mit EW  $\lambda_1, \dots, \lambda_n$  und EV  $v_1, \dots, v_n > 0$  dann ist  $A = U \Sigma V^T$  mit  $U = V = \begin{pmatrix} | & & | \\ v_1 & & v_n \\ | & & | \end{pmatrix}$  und  $\Sigma = \begin{pmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_n \end{pmatrix}$  die Singularwertzerlegung von  $A$

- ist  $A$  nicht symmetrisch dann ist  $AA^T \in \mathbb{R}^{n \times n}$  eine sym. Matrix, denn

$$AA^T \text{ hat pos. Bl. } (AA^T)_{ij} = \sum_{k=1}^m a_{ik} a_{jk}$$

da:

$$AA^T = Nx$$

$$\rightarrow x^T AA^T x = Nx^T x$$

$$\rightarrow \lambda = \frac{Nx^T x}{x^T x}$$

- $AA^T$  hat eine Basis von EV  $u_1, \dots, u_n \in \mathbb{R}^n$  und pos. EW  $\lambda_1, \dots, \lambda_n \geq 0$

$$\Rightarrow AA^T = U \begin{pmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_n \end{pmatrix} U^T \text{ mit } U = (u_1, \dots, u_n)$$

analoges gilt für  $A = U \Sigma V^T$

$$AA^T = U \Sigma V^T V \Sigma^T U^T$$

$$= U \begin{pmatrix} \sigma_1 & & 0 \\ & \dots & \\ 0 & & \sigma_n \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \dots & \\ 0 & & \sigma_n \end{pmatrix}^T U^T$$

$$= U \begin{pmatrix} \sigma_1^2 & & \\ & \dots & \\ & & \sigma_n^2 \end{pmatrix} U^T$$

$\Rightarrow \sigma_k = \sqrt{\lambda_k}$  sind die Singulärwerte

$u_1, \dots, u_n$  sind die Linkssingulärwerte

$$\Rightarrow V = \Sigma^{-1} U^T A$$

## Wichtige Kennzahlen von Matrizen

Determinante:  $\det A = \prod_{k=1}^{\min(m,n)} \sigma_k$

Beschreibt die Volumenveränderung eines

Einheitswürfels unter  $A$

Spektralnorm:  $\|A\| = \sigma_1$

Beschreibt die maximale Streckung in

einer Richtung

Kondition:  $\kappa(A) = \frac{\sigma_1}{\sigma_n}$

Beschreibt das Verhältnis von maximaler

Streckung zu minimaler Streckung (Staubung)

## Bem.

• Für alle  $x \in \mathbb{R}^m$  gilt:  $\|Ax\|_2 \leq \|A\| \|x\|_2$

↳  $\|A^{-1}\| = \sigma_n^{-1}$

•  $A$  invertierbar  $\Rightarrow m=n$ ,  $\sigma_n \neq 0$

•  $\kappa(A) = \|A\| \|A^{-1}\| = \kappa(A^{-1})$

Bsp.:  $\left\| \begin{pmatrix} 6 & ? \\ 2 & 6 \end{pmatrix} \right\| = 8$

$\kappa \left( \begin{pmatrix} 6 & ? \\ 2 & 6 \end{pmatrix} \right) = 2$

## 2.2. lineare Gleichungssysteme

Wir betrachten das lineare Gleichungssystem  $Ax = b$ ,  $A \in \mathbb{R}^{n \times m}$ ,  $x \in \mathbb{R}^m$ ,  $b \in \mathbb{R}^n$

Sei  $A = U \Sigma V^T$  die Singulärwert-

zerlegung von  $A$ . Dann gilt:

$$U \Sigma V^T x = b \Leftrightarrow \Sigma V^T x = U^T b$$

$$\Leftrightarrow \Sigma y = U^T b, \quad x = Vy$$

da  $U, V$  orthogonal, also invertierbar sind

reicht es das LGS  $\Sigma y = U^T b$  zu lösen

1. Fall  $m = n$  (quadratischer Fall)

$$\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = U^T b =: c$$

- hat genau eine Lsg. falls  $\sigma_k > 0$
- hat keine Lsg. falls  $\sigma_k = 0$  und  $c_k \neq 0$  für ein  $k = 1..n$
- hat unendlich viele Lsg. falls  $\sigma_k = 0 \Rightarrow c_k = 0$

## 2. Fall $n < m$ (unterbestimmtes System)

$$\begin{array}{|c|} \hline \begin{array}{c} \sigma_1 \quad 0 \\ \vdots \quad \vdots \\ 0 \quad \vdots \\ \sigma_n \end{array} \\ \hline \end{array} \quad \begin{array}{c} \left( \begin{array}{c} y_1 \\ \vdots \\ y_m \end{array} \right) = \left( \begin{array}{c} c_1 \\ \vdots \\ c_n \end{array} \right) \end{array}$$

- keine Lsg. falls  $\sigma_k = 0$  und  $c_k \neq 0$  für ein  $k=1..n$

- sonst unendlich viele Lsg.  $y_1, \dots, y_m$  können frei gewählt werden

## 3. Fall $n > m$ (überbestimmtes Fall)

$$\begin{array}{|c|} \hline \begin{array}{c} \sigma_1 \quad 0 \\ \vdots \quad \vdots \\ 0 \quad \vdots \\ \sigma_m \\ \vdots \\ 0 \end{array} \\ \hline \end{array} \quad \begin{array}{c} \left( \begin{array}{c} y_1 \\ \vdots \\ y_m \end{array} \right) = \left( \begin{array}{c} c_1 \\ \vdots \\ c_n \end{array} \right) \end{array}$$

- Eindeut. Lsg. für  $\sigma_k \neq 0, k=1..m$  und  $c_k=0 \forall k > m$

- keine Lsg.  $c_k \neq 0$  für ein  $k > m+1$  od.  $\sigma_k = 0$

- sonst unendlich viele Lsg.

## Die Kondition LGS

Wir betrachten das Modell

Eingangswerten: rechte Seite  $b \in \mathbb{R}^n$

Ausgangswerten:  $x \in \mathbb{R}^n$  mit  $Ax = b$

Satz Sei  $\Delta b \in \mathbb{R}^n$  eine Störung der Eingangswerten und  $x + \Delta x$  die zugehör.

Lsg. also  $A(x + \Delta x) = b + \Delta b$ .

Dann gilt:

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \kappa(A) \frac{\|\Delta b\|_2}{\|b\|_2}$$

$$\begin{aligned} \text{Bew. } \Delta b &= (b + \Delta b) - b \\ &= A(x + \Delta x) - Ax = A(\Delta x) \end{aligned}$$

$$\Rightarrow \Delta x = A^{-1} \Delta b$$

$$\Rightarrow \|\Delta x\|_2 \leq \|A^{-1}\| \|\Delta b\|_2$$

$$\text{außerdem: } \|b\|_2 = \|A\| \|x\|_2 \leq \|A\| \|x\|_2$$

$$\Rightarrow \|x\|_2 \geq \|A\|^{-1} \|b\|_2$$

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \frac{\|A^{-1}\| \|\Delta b\|_2}{\|A\|^{-1} \|b\|_2} = \underbrace{\|A\| \|A^{-1}\|}_{=\kappa(A)} \frac{\|\Delta b\|_2}{\|b\|_2}$$

Merke: Die Kondition der Matrix ist

gerade die Kondition des Rechts:  
d.h. das Zwerh. für  $\kappa(A)$

## 2.3. Direkte Verfahren zum lösen linearer Gleichungssysteme

### Dreiecksmatrizen und gestufte Systeme

Def: Eine Matrix  $U$  heißt obere Dreiecksmatrix, falls  $A = (a_{ij})$  mit  $a_{ij} = 0$  für alle  $i > j$ .  $A$  heißt untere Dreiecksmatrix

falls  $a_{ij} = 0 \forall i < j$ . Ein System

$Ax = b$  heißt gestuft, falls  $A$  obere

od. untere Dreiecksmatrix ist.

o gestufte Systeme sind einfach lösbar.

### Alg zur Lösg. gestufter Systeme

$Ax = b$ ,  $A = \begin{pmatrix} \circ & \nabla \\ \circ & \circ \end{pmatrix}$  obere Dreiecksm.

$\Rightarrow$  Rückwärtssubstitution

für  $i = n-1$ :  $x_i = \frac{1}{a_{ii}} \left( b_i - \sum_{k=i+1}^n a_{ik} x_k \right)$

$Ax = b$ ,  $A = \begin{pmatrix} \nabla & \circ \\ \circ & \circ \end{pmatrix}$  untere Dreiecksm.

$\Rightarrow$  Vorwärtssubstitution

für  $i = 1 \dots n$ :  $x_i = \frac{1}{a_{ii}} \left( b_i - \sum_{k=1}^{i-1} a_{ik} x_k \right)$

Aufwand:  $\mathcal{O}(N^3)$  Operationen

## Eigenschaften von Dreiecksmatrizen

- Das Produkt und die Summe zweier oberer Dreiecksmatrizen ist wieder eine obere Dreiecksmatrix.
- Die inverse einer invertierbaren ob. Dreiecksm. ist wieder eine ob. Dreiecksm.

## Der Gauß-Algorithmus

Idee: Zerlege  $A = L \cdot R$

$L$  - linke, untere Dreiecksmatrix  
 $R$  - rechte, obere Dreiecksmatrix

Bsp:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 1 & 8 & 0 \end{pmatrix}$$

$$\begin{matrix} a_{21} \\ a_{31} \\ a_{32} \\ a_{33} \end{matrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} = \begin{matrix} \\ \\ \\ \\ \end{matrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 8 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

## Algorithmus: LR-Zerlegung

Eingabe:  $A \in \mathbb{R}^{n \times n}$   
 $b \in \mathbb{R}^n$

① Faktoriere  $A = L \cdot R$  durch Gauß-  
elimination

② Löse  $LY = b$  durch Vorwärtsstufen

③ Löse  $Rx = Y$  durch Rückwärtsst.

Bsp (Fortf.)

$$b = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad LY = b$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \Rightarrow Y = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \Rightarrow X = \begin{pmatrix} 1 \\ \frac{1}{2} \\ -4 \end{pmatrix}$$

Bem.:

• Die Komplexität des Gaußverfahrens  
ist  $\mathcal{O}(N^3)$

• Der Vorteil der LR-Zerlegung gegenüber  
dem gew. Gaußalg besteht darin, dass  
das aufwendigste Schritt - die LR-Zerlegung  
unabhängig von der rechte Seite  $b$  ist



• Der kritische Schritt beim Gauß-Verfahren ist die Division durch die Diagonaleinträge (Diese sollten möglichst ungleich Null sein) Durch geschickte Pivotierung d.h. Zeilen und Spalten vert. so dass das Diagonalelement groß wird kann die Rechnung stabil implementiert werden.



• Das Ate hat die kompl.  $\mathcal{O}(N)$

• Das Nte hat die kompl.  $\mathcal{O}(N)$

das den gesamten wert,  $\mathcal{O}(N)$

A stellt die gesamte dar,  $\mathcal{O}(N)$

$$d.h. \sum_{j=1}^n |a_{ij}| < |a_{ii}|$$

• Das Theorem hat die Num. erkl.  
Komplexität  $\mathcal{O}(N)$

1

## QR - Faktorisierung

Bei der QR Zerlegung wird die Matrix  $A = Q \cdot R$  in eine orthogonale Matrix  $Q$  und eine hoch obere Dreiecksmatrix  $R$  zerlegt.

Lösen eines linearen LS mittels einer

QR-Zerlegung:

$$QRx = b \Leftrightarrow Rx = Q^T b$$

$\Rightarrow$  löse mit Rückwärtsiteration

Bem.

• Die QR Zerlegung ist ungefähr doppelt so teuer wie eine LR-Zerlegung

• Numerisch stabil

• Wird  $n. A.$  zur Berechnung der Eigenwerte verwendet

• od. zur Berechnung der EW.

## Positiv Definit Matrixen u. Cholesky Zer.

Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt

- symmetrisch, falls  $a_{ij} = a_{ji}$
  - positiv definit, falls  $x^T A x \geq 0 \forall x$
- Es gilt

- A positiv def.  $\Leftrightarrow$  alle Eigenwerte von A sind nicht negativ

$$\underline{\text{Bem.}}: \Rightarrow A x = \lambda x \Rightarrow x^T A x = \lambda x^T x \geq 0$$

$\Leftrightarrow$  nur  $\lambda \geq 0$  nutzt A hat eine ON-Basis aus Eigenvektoren.

- $B \in \mathbb{R}^{n \times n}$  beliebig  $\Rightarrow A = B B^T$  sym. pd.

- ist A spd.  $\Rightarrow A = B B^T$  für ein geeignetes B

Def. Eine Faktorisierung von  $A \in \mathbb{R}^{n \times n}$

$$A = C C^T, C = (\begin{matrix} \geq 0 \\ > 0 \end{matrix}), C_{ii} > 0$$

heißt Cholesky-Zerlegung.

Bem.: • A sym. pd.  $\Rightarrow$  A besitzt Chol. Zer.

• Aufwand:  $\frac{1}{2}$  Halbmatrix wie LR

• reum. stabil

• nur eine Matrix muß gep. werden

•  $C C^T$  ist auch reum. spd.

## 2.4. Iterative Verfahren zum Lösen LGS

Bsp. Def. Bei einem iterativen Verfahren wird ausgehend von einem Startwert  $x_0^{(0)}$  eine Folge von „Lösungen“

$$x^{(k+1)} = \Phi(x^{(k)}) \quad (k=1, 2, \dots)$$

berechnet. Ziel ist, dass für  $k \rightarrow \infty$

$$x_k^{(k)} \rightarrow x^* \quad (\text{die wahre Lsg.})$$

$$\text{Für } x^* \text{ gilt } x^{(k+1)} = \Phi(x^{(k)})$$

$$x^* = \Phi(x^*)$$

$x^*$  heißt Fixpunkt des Gleiches  $y = \Phi(x)$ .

Bsp.  $\Phi(x) = \frac{x + \frac{a}{x}}{2}$  Heron Verfahren

$$\text{Fixpt.: } x = \frac{x + \frac{a}{x}}{2} \Leftrightarrow 2x = x + \frac{a}{x}$$

$$\Leftrightarrow x^2 = a \Leftrightarrow x = \sqrt{a}$$

Frage: Wann konvergiert ein iteratives Verfahren?

Idee: Transformiere LGS in

Fixpunktgleichung:

$$Ax = b \quad A = H - N$$

$$\Leftrightarrow (H - N)x = b$$

$$\Leftrightarrow Hx = Nx + b$$

$$\Leftrightarrow x = H^{-1}(Nx + b) = \Phi(x)$$

Iterationsvorschrift:

$$x^{(k+1)} = H^{-1}(Nx^{(k)} + b)$$

Welche Zerlegungen  $A = H - N$  sind sinnvoll?

-  $M$  muss einfach invertierbar sein

$\Rightarrow M$  Diagonalmatrix, Bandmatrix,  
Dreiecksmatrix, orthogonale  $M$

Achtung: hier Summe (nicht Produkt)

$\Rightarrow$  Zerlegung einfach

Das Gesamtschrittverfahren (Jacobi-Verf.)

$$A = \begin{pmatrix} \times & & \\ & \times & \\ & & \times \end{pmatrix} - \begin{pmatrix} \times & & \\ & \times & \\ & & \times \end{pmatrix} = \begin{matrix} \text{H} \\ \text{N} \end{matrix}$$

## Iterationsvorschrift:

$$x^{(k+1)} = M^{-1} (N x^{(k)} + b)$$
$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} x_j^{(k)} \right)$$

d.h. im  $k$ -ten Iterationsschritt lösen

wir unabh. voneinander jeweils die  $i$ -te

Zeile nach der  $i$ -ten Unbekannten  $x_i$  auf

und setzen für alle anderen  $x_j$ ,  $j \neq i$

die Werte des vorherigen Iterationsschritts

ein.

Aufwand: pro Iteration  $\mathcal{O}(N^2)$

Bsp.

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & -4 & 1 \\ 0 & -1 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \|x^{(0)} - x^*\|^2 = 3$$

$$x^{(1)} = \begin{pmatrix} 1/2 \\ -4/4 \\ -1/2 \end{pmatrix} \quad \|x^{(1)} - x^*\|^2 = 1/2$$

$$x^{(2)} = \begin{pmatrix} 1/4 \\ 1/4 \\ -1/2 \end{pmatrix} \quad \|x^{(2)} - x^*\|^2 = 1/16$$

$$x^{(3)} = \begin{pmatrix} 1/8 \\ 3/4 \\ -1/2 \end{pmatrix} \quad \|x^{(3)} - x^*\|^2 = 1/16$$



## Das Einzelschrittverfahren (Gauss-Seidel-Verfahren)

$$A = \begin{pmatrix} \triangle^0 \\ \triangle^1 \\ \triangle^2 \\ \triangle^3 \end{pmatrix} - \begin{pmatrix} \triangle^1 \\ \triangle^2 \\ \triangle^3 \\ \triangle^4 \end{pmatrix} = N$$

Iterationsvorschrift:

$$x^{(k+1)} = N^{-1}(N x^{(k)} + b)$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right)$$

d.h. im  $k$ -ten Iterationsschritt wird die

$i$ -te Zeile nach der  $i$ -ten Variable aufgelöst.

Für alle  $x_j$ ,  $j > i$  werden die Werte des

letzten Iterationsschritts eingesetzt. Für  $x_j$

$j < i$  werden die Zwischenergebnisse des  $i$ -ten

Iterationsschritts verwendet.

Anforderung: Analog zu Gauß-Jordan-Verfahren

Frage: Wann konvergiert ein iteratives Verfahren  $x^{(k+1)} = \Phi(x^{(k)})$ ?

### Satz (Banachscher Fixpunktsatz)

Sei  $\Phi$  ein Kontraktion, d.h. es gilt ein  $q < 1$  so, dass für alle  $x, y$  gilt

$$\|\Phi(x) - \Phi(y)\| \leq q \|x - y\|.$$

Dann hat die Fixpunktgleichung

$$x = \Phi(x)$$

genau eine Lösung  $x^*$  und das iterative

Verfahren  $x^{(k+1)} = \Phi(x^{(k)})$  konvergiert für

jeden Vektor  $x^{(0)}$  gegen  $x^*$  für  $k \rightarrow \infty$ .

Es gelten die Abschätzungen:

a)  $\|x^{(k)} - x^*\| \leq q^k \|x^{(k-1)} - x^*\|$

b)  $\|x^{(k)} - x^*\| \leq \frac{q^k}{1-q} \|x^{(1)} - x^{(0)}\|$

c)  $\|x^{(k)} - x^*\| \leq \frac{q}{1-q} \|x^{(k-1)} - x^{(k-1)}\|$

Für das Gesamt bzw. Fixpunktschrittverf.

haben wir  $x^{(k+1)} = \Phi(x^{(k)}) = M^{-1}(b + Nx^{(k)})$

$\Rightarrow \|\Phi(x) - \Phi(y)\| = \|M^{-1}N(x-y)\|$

$\leq \|M^{-1}N\| \|x-y\|$

$\Rightarrow$  Das Einzel bzw. Gesamtschrittverfahren konvergiert, falls  $\|H^{-1}N\| < 1$ . Insbesondere ist die Konvergenz umso schneller je kleiner  $\|H^{-1}N\|$  ist.

Satz: Ist  $A$  streng diagonaldominant, d.h.  $\sum_{i \neq j} |a_{ij}| < a_{ii}$ , dann gilt

- $A$  ist regulär
- $\|(\infty)^{-1} \cdot (0^0)\| < 1$
- Das Gesamtschrittverfahren konvergiert.
- Das Einzelverfahren konvergiert.

Bem. Vorteile iterativer Verfahren

- ① Für große Matrizen ist die Anzahl der nötigen Iterationen kleiner als die Anzahl der Spalten / Zeilen
- ② Die Anzahl der Schritte hängt von der Genauigkeit.

- ③ Die Matrix  $A$  muß nicht invertiert werden. Es reicht die Matrix multiplizieren  $H^{-1}(b + Nx)$  zu implementieren.

## Das Verfahren des skalarwertigen Abstrahierens

In diesem Abschnitt ist  $A$  immer eine

sym. pos. def. Matrix, d.h.  $x^T A x \geq 0 \forall x$

Zu dem LGS  $Ax = b$  betrachten

wir das Minimierungsproblem

$$F(x) \rightarrow \min \text{ mit } F(x) = \frac{1}{2} x^T A x - x^T b$$

Satz: Es gilt

$$\begin{aligned} \nabla F(x) &= \frac{1}{2} (Ax + A^T x) - b \\ &= Ax - b \end{aligned}$$

also  $\nabla F(x) = 0 \iff Ax = b$

ist  $Ax = b$  dann gilt für Bel.  $\Delta x$

$$F(x + \Delta x) = \frac{1}{2} (x + \Delta x)^T A (x + \Delta x) - (x + \Delta x)^T b$$

$$= \underbrace{\frac{1}{2} x^T A x - x^T b}_{F(x)}$$

$$+ \Delta x^T b + \Delta x^T A \Delta x - \Delta x^T b$$

$$= F(x) + \underbrace{\Delta x^T A \Delta x}_{\geq 0} \geq F(x)$$

$\implies$  die Lösung des LGS  $Ax = b$

ist gleichzeitig das Minimum von  $F(x)$ .

Bsp

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix}$$

$$b = \begin{pmatrix} 2 \\ -8 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$F(x) = \frac{1}{2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 2 \\ -8 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

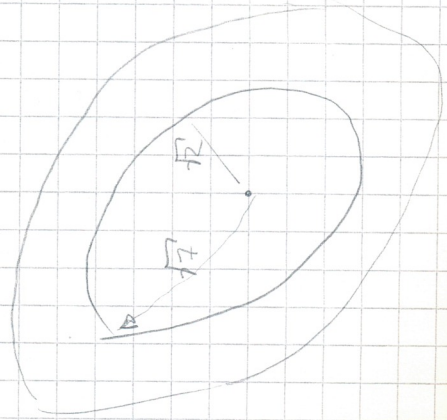
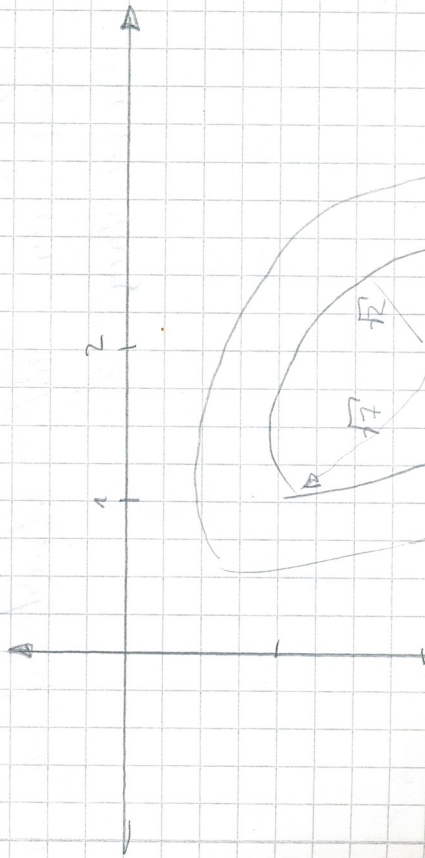
$$= \frac{1}{2} (x_1 \ x_2) \begin{pmatrix} 3x_1 + 2x_2 \\ 2x_1 + 6x_2 \end{pmatrix} - 2x_1 + 8x_2$$

$$= \frac{1}{2} (3x_1^2 + 4x_1x_2 + 6x_2^2) - 2x_1 + 8x_2$$

→ schnell zu integrieren!

$$\text{SVD von } A = U \begin{pmatrix} 7 & \\ & 2 \end{pmatrix} U^T$$

$$\text{mit } U = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix}$$



Ziel: Minimiere die Funktion  $F$  ausgehend von einem Punkt  $x$  entlang einer Richtung  $d$

Iterationsvorschrift:

$$x^{(k+1)} = x^{(k)} + \lambda^{(k)} d^{(k)}$$

$\lambda^{(k)}$  - Schrittweite

$d^{(k)}$  - Suchrichtung

Zwischenfrage: Sei  $d$  eine beliebige Suchrichtung -

was ist die optimale Schrittweite?

$$F(x + \lambda d) \rightarrow \min$$

$$\rightarrow \frac{d}{d\lambda} F(x + \lambda d) =$$

$$F(x + \lambda d) = F(x) + \lambda d^T (Ax - b) + \frac{1}{2} \lambda^2 d^T A d$$

ist eine quadratische Funktion in  $\lambda$

$$\frac{d}{d\lambda} F(x + \lambda d) = d^T (Ax - b) + \lambda d^T A d \stackrel{!}{=} 0$$

$$\Rightarrow \lambda = \frac{d^T r}{d^T A d}$$

das ist die optimale Schrittweite für eine Richtung  $d$ .

Frage 2: Was ist die beste Suchrichtung?

$$\cdot d = -\nabla F(x) = b - Ax = r$$

Alg. des stabilen Abstrakte

Eigenschaft:  $A \in \mathbb{R}^{n \times n}$  stabil  
 $x^{(0)} \in \mathbb{R}^n$  Startwert  
 $\varepsilon > 0$  Zielgenauigkeit

①  $k=0$

②  $r^{(0)} = b - Ax^{(0)}$

③ Sollanzahl  $\|r^{(k)}\| > \varepsilon$

④  $\lambda^{(k)} = \frac{r^{(k)T} r^{(k)}}{r^{(k)T} A r^{(k)}} = \frac{d^{(k)T} r^{(k)}}{d^{(k)T} A d^{(k)}}$

⑤  $x^{(k+1)} = x^{(k)} + \lambda^{(k)} d^{(k)}$

⑥  $r^{(k+1)} = b - Ax^{(k+1)}$

⑦  $k = k+1$

Die Orthogonalität der Suchrichtungen

führt zu einem neg. Zick-Zack-Verhalten.

Bsp.:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}$$



⇒ Die lokale optimale Wahl der Abstrakte  
kann zu langsameren Konvergenz führen.  
Dies hängt mit der Kondition der Matrix ab

Satz: Für die  $k$ -te Iterierte des Gradienten-Verfahrens gilt

$$\|A^{\frac{1}{2}}(X^{(k)} - X^*)\| \leq \left( \frac{\text{cond}(A) - 1}{\text{cond}(A) + 1} \right)^k \|X^{(0)} - X^*\|$$

d.h. das Abklingverf. konvergiert linear mit einer Konstante abhängig von der Kondition der Matrix.

Beob. Sind alle EW von  $A$  gleich, so sind die Höhenlinien-Kurve und das Verfahren konvergiert in einem Schritt.



## Das konjugierte Gradienten Verfahren

Ziel: Finde bessere Suchrichtungen  $d^{(k)}$ ,  
so dass der Zielfunktionswert nicht aufhört,

Man beachte alle  $d^{(k)}$  sind orthogonal zueinander.

Beobachtung: Wir wissen bei jedem Suchschritt  
Schrittweite  $\lambda^{(k)} = \frac{d^{(k)T} r^{(k)}}{d^{(k)T} A d^{(k)}}$  gilt

$$\left( b - A \left( x^{(k-1)} + \lambda^{(k)} d^{(k)} \right) \right)^T d^{(k)} = 0$$

$$\rightarrow d^{(k)T} A \left( x^{(k)} - x^{(k-1)} \right) = 0$$

d.h. die Differenz  $x^{(k)} - x^{(k-1)}$  ist bzgl.  $d^{(k)}$  orthogonal

Skalarprod.  $\langle x^{(k)} - x^{(k-1)}, d^{(k)} \rangle = 0$  senkrecht

Zu der letzten Suchrichtung.

Konjuganz: Wähle alle folg.-d. Suchrichtungen

A-orthogonal zu  $d^{(k)}$

$$d^{(k+1)T} = r^{(k+1)T} - \frac{r^{(k+1)T} A d^{(k)}}{d^{(k)T} A d^{(k)}} d^{(k)}$$

man kann zeigen:  $d^{(k+1)}$  ist A-

orthogonal zu allen vorherigen Suchrichtungen  
 $d^{(0)}, \dots, d^{(k)}$

# Alg. (konjugierte Gradientenverf., CG)

Eingabe:  $A \in \mathbb{R}^{n \times n}$  sym. pos. def.  
 $x^{(0)} \in \mathbb{R}^n$  Startvektor  
 $\varepsilon > 0$  Zielgenauigkeit

- 1  $k=0$
- 2  $r^{(0)} = b - A x^{(0)}$  Residuum / res. Grad
- 3  $d^{(0)} = r^{(0)}$  erste Suchrichtung
- 4 Solange  $\|r^{(k)}\| > \varepsilon$
- 5  $\lambda^{(k)} = \frac{d^{(k)T} r^{(k)}}{d^{(k)T} A d^{(k)}}$  Schrittweite
- 6  $x^{(k+1)} = x^{(k)} + \lambda^{(k)} d^{(k)}$  SD Schritt
- 7  $r^{(k+1)} = b - A x^{(k+1)}$  Residuum
- 8  $d^{(k+1)} = r^{(k+1)} - \frac{r^{(k+1)T} d^{(k)}}{d^{(k)T} A d^{(k)}} d^{(k)}$
- 9  $k = k+1$
- 10 end

Aufwand:  $\mathcal{O}(n^2)$  pro Iterationsschritt.

Bem. • Da die neuen Stützungen  
im  $A$ -orthogonal zu allen vorherigen  
Stützungen ist konvergiert das Verh.  
nach spätestens  $n$ -Schritten.

• Beispiel  $A$  nur  $m$ -veränd. EW

so konvergiert das Verfahren nach  $m$ -Schritten

• Die Konvergenz des CG-Verfahrens  
hängt von der Verteiltheit des EW  
von  $A$  ab.

• Es gilt:

$$\|x^{(k)} - x^*\|_A \leq 2 \left( \frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k \|x^{(0)} - x^*\|_A$$

## 2.5. Über und Unterbestimmte

### Lineare Gleichungssysteme

Bsp.

zu gegebenen Stützstellen  $t_j \in \mathbb{R}$

und Stützwerten  $Y_j$ ,  $j=1, \dots, m$ , na suchen

für das Modell  $y(t) = x_0 + x_1 t$ ,

$Y_j = x_0 + x_1 t_j$  die besten

Parameter  $x_0, x_1 \in \mathbb{R}$

Wir betrachten die Minimierungsaufg.

$$\|r\|_2 = \left( \sum_{j=1}^m (Y_j - (x_0 + x_1 t_j))^2 \right)^{1/2} \rightarrow \min$$

Wir schreiben das als LGS

$$\underbrace{\begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix}}_A \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \approx \underbrace{\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_m \end{pmatrix}}_Y = Y$$

Das lineare GS  $Ax = Y$  ist über-

bestimmt, d.h. keine Lösung

Wir betrachten das Minimierungsproblem

$$\|r\|_2^2 = \|Y - Ax\|_2^2 = \sum (Y_j - (x_0 + x_1 t_j))^2 \rightarrow \min$$

$\Rightarrow$  Methode des kleinsten Quadrats

allgemein wollen wir für  $A \in \mathbb{R}^{m \times n}$ ,  
 $m > n$  und  $Y \in \mathbb{R}^m$  das überbestimmte  
LGS  $Ax = b$  lösen; hier das  
zugehörige kleinste Quadrat Problem

$$\min_{x \in \mathbb{R}^n} \|Ax - Y\|^2 \quad (*)$$

Satz:  $x$  ist Lösung des kleinsten Quadrates  
Problem  $(*) \Leftrightarrow$  wenn  $x$  die Normalen-  
gleichung erfüllt  $A^T Ax = A^T Y$

Bew.

$$A^T Ax = A^T Y \Leftrightarrow A^T (Y - Ax) = 0$$

$$\Leftrightarrow A^T r = 0$$

$\Leftrightarrow r$  ist senkrecht auf allen Spalten von  $A$

$\Leftrightarrow r$  ist orthogonal zu Bildraum von  $A$

$\Leftrightarrow x$  ist Lsg des kleinsten Quadrates  
Problems

Bem. Die Matrix  $A^T A$  ist sym

pos. definit  $\Rightarrow$  das CG-Verfahren ist  
anwendbar.

Wegweise

Bsp. (Fortsetzung Lineare Regression)

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} m & m\bar{t} \\ m\bar{t} & \sum_{j=1}^m t_j^2 \end{pmatrix}$$

wobei  $\bar{t} = \frac{1}{m} \sum_{j=1}^m t_j$  und  $\bar{y} = \frac{1}{m} \sum_{j=1}^m y_j$

$$\begin{pmatrix} m & m\bar{t} \\ m\bar{t} & \sum_{j=1}^m t_j^2 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} m\bar{y} \\ \sum_{j=1}^m t_j y_j \end{pmatrix}$$

$$\Rightarrow x_1 = \frac{\sum_{j=1}^m t_j y_j - m\bar{y}\bar{t}}{\sum_{j=1}^m t_j^2 - m\bar{t}^2}$$

$$x_0 = \bar{y} - x_1 \bar{t}$$

# Unbestimmte Lineare Gleichungssysteme

Sie betrachten das Problem

$$\begin{bmatrix} A \\ x \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

mit  $A \in \mathbb{R}^{m \times n}$ ,  $m < n$

das hier ein Lsg., falls sie ex. existiert, nicht existiert  
ist fügen wir eine zusätzliche Bed. hinzu

$$(*) \quad Ax = b \quad \text{und} \quad \|x\| \rightarrow \min$$

d.h. wir wählen unter allen Lsg. des

LGS das mit minimaler Norm aus.

Satz: Das Problem  $(*)$  hat eine einz.

Lösung, falls das Probl.  $Ax = b$  eine Lsg. hat.

Die ist die Lsg. der Normalgleichung

$$2. \text{ Art: } AA^T y = b, \quad x = A^T y.$$

Bew.  $x^*$  Lsg. der Normalgleichung 2. Art

$$x^* \cdot Ax = b$$

$$\rightarrow \langle x^*, x - x^* \rangle = \langle A^T y^*, x - x^* \rangle$$

$$= \langle y^*, \underbrace{Ax - Ax^*}_{= b - b} \rangle = 0$$

$$\begin{aligned} \Rightarrow \|x\|^2 &= \|x^* + x - x^*\|^2 \\ &= \|x^*\|^2 + \|x - x^*\|^2 \geq \|x^*\|^2 \end{aligned}$$

## Die Pseudoinverse

Sei  $A = U \Sigma V^T$  die SVD von  $A$

$$\text{und } \Sigma^+ = (\sigma_{ii}^+), \quad \sigma_{ii}^+ = \begin{cases} 0 & \sigma_{ii} = 0 \\ \frac{1}{\sigma_{ii}} & \sigma_{ii} \neq 0 \end{cases}$$

$$\text{Dann hei\u00dft} \\ A = V \Sigma^+ U^T$$

die Pseudoinverse von  $A$ . Es gilt

$$\textcircled{1} A A^+ A = A \quad \textcircled{2} A^+ A A^+ = A^+$$

$$\textcircled{3} (A A^+)^T = A A^+$$

$$\textcircled{4} (A^+ A)^T = A^+ A$$

Satz: Die minimum Norm L\u00f6s des LGS

$$Ax = b, \quad \|x\|_2 \rightarrow \min \text{ ist } x = A^+ b.$$

Bew: Zeige  $x = A^+ y$  erf\u00fcllt Normal

gleichung? Ansatz

$$x = A^+ y = A^+ b$$

$$\Rightarrow A A^+ y = A A^+ b = b$$

da  $b$  im Bild von  $A$  liegt.



# 3. Interpolation

## 3.1. Interpolation mit bel. Funktionen

### Interpolation:

geg: Stützstellen  $t_j \in \mathbb{R}$   $j = 1 \dots n$

Stützwerte  $Y_j \in \mathbb{R}$

ges: Funktion  $f$  mit  $f(t_j) = Y_j$

### Anwendungen

- Visualisierung diskreter Daten
- Integration / Differentiation
- Approximation einer komplizierten Fkt. durch einfachere
- Oberflächendesign CAD

Die obige Interpolationsaufgabe ist

nicht sinnvoll gestellt, da es

unendlich viele solche Funktionen gibt.

Wir brauchen zusätzliche Annahmen an  $f$

Hier:  $f$  ist Linearstückerstein, von

$n$  vorgegebenen Fkt  $p_1, \dots, p_n$

d.h.  $f(t) = c_1 p_1(t) + c_2 p_2(t) + \dots + c_n p_n(t)$

durch diese Annahme wird das Inhomogenitätsproblem zu ein LGS

$$c_1 \varphi_1(t_n) + c_2 \varphi_2(t_n) + \dots + c_n \varphi_n(t_n) = Y_n$$

$$c_1 \varphi_1(t_n) + c_2 \varphi_2(t_n) + \dots + c_n \varphi_n(t_n) = Y_n$$

Bzw.

$$\begin{pmatrix} \varphi_1(t_n) & \varphi_2(t_n) & \dots & \varphi_n(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(t_n) & \varphi_2(t_n) & \dots & \varphi_n(t_n) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} Y_n \\ \vdots \\ Y_n \end{pmatrix}$$

### Fragestellungen

- hat das LGS eine LGS? - Eindeutigkeit?
- ist es ein Randwertproblem
- ist  $g: \mathbb{R} \rightarrow \mathbb{R}$  eine "glatte" Fkt mit  $g(t_j) = Y_j$  die stetig ist oder aber  $|g(t) - p(t)|$ ?

Bsp.  $t_j = -1 \ 0 \ 2$   
 $y_j = 1 \ 2 \ 2$

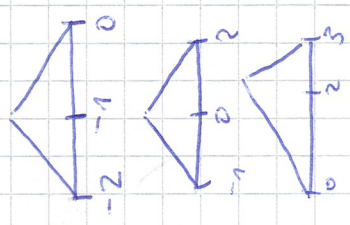
a)  $y_j = 1, t, t^2, f(t) = c_0 + c_1 t + c_2 t^2$

$$\begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

$\Rightarrow f(t) = 2 + \frac{2}{3}t - \frac{1}{3}t^2$

b)  $f_1(t) = \begin{cases} x+2 & x \in [-2, 1] \\ -x & x \in [-1, 0] \\ 0 & \text{sonst} \end{cases}$   
 $f_2(t) = \begin{cases} x+1 & x \in [-1, 0] \\ 1-\frac{x}{2} & x \in [0, 2] \\ 0 & \text{sonst} \end{cases}$   
 $f_3(t) = \begin{cases} x/2 & x \in [0, 2] \\ 3-x & x \in [2, 3] \\ 0 & \text{sonst} \end{cases}$



$$\begin{pmatrix} f_1(t_1) & f_2(t_1) & f_3(t_1) \\ f_1(t_2) & f_2(t_2) & f_3(t_2) \\ f_1(t_3) & f_2(t_3) & f_3(t_3) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

$\Rightarrow c_1 = 1, c_2 = 2, c_3 = 2$

$$\Rightarrow f(t) = \begin{cases} 1 \cdot (x+2) & x \in [-2, -1) \\ -1 \cdot (-x) + 2(x+1) = x+2 & x \in [-1, 0] \\ 2 \left(1 - \frac{x}{2}\right) + \frac{x}{2} = 2 - \frac{x}{2} & x \in [0, 2] \\ 2(3-x) & x \in [2, 3] \\ 0 & \text{sonst} \end{cases}$$

c)  $\varphi_j(t) = e^{-(t-1)^2}$

$$f(t) = c_1 \varphi_1(t) + c_2 \varphi_2(t) + c_3 \varphi_3(t)$$

$$\Rightarrow \underbrace{\begin{pmatrix} 1 & e^{-1} & e^{-9} \\ e^{-1} & 1 & e^{-4} \\ e & e^{-4} & 1 \end{pmatrix}}_A \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

A ist diagonalisierbar, und sym. pos. def.  
 da Expon. Abfall von Dig.

## 4.2. Polynomische Interpolation

Def Sei  $n \in \mathbb{N}$ . Dann heißt

$$p_0 = 1, p_1(t) = t, p_2(t) = t^2, \dots, p_n(t) = t^n$$

die Monombasis des Vektorraumes aller Polynome

$$p_n(t) = c_0 + c_1 t + c_2 t^2 + \dots + c_n t^n$$

von Grad  $n$ .

Das polynomische Interpolationsproblem

besteht zu  $t_0, \dots, t_n$  Stützstellen

$y_0, \dots, y_n$  Stützwerte

finden Polynom  $p_n$  vom Grad  $n$

mit  $p_n(t_j) = y_j$

Satz: Das polynomische Interpolationsproblem

ist für paarweise verschiedene Stützstellen

$t_i \neq t_j$  eindeutig lösbar.

Bew. Existenz  $\rightarrow$  später (Lagrange'scher)

Eindeutigkeit: Seien  $p_n, q_n$  zwei

Polynome vom Grad  $n$  mit  $p_n(t_j) = q_n(t_j) = y_j$

$\Rightarrow p_n - q_n$  ist ein Polynom vom Grad  $n$

mit  $n+1$  Nullen  $\Rightarrow p_n = q_n$

Bem.:

• Komplexität

- Berechnung des LGS:  $\mathcal{O}(n^3)$

- Auswertung des Interpolat.

$$\text{nat. v: } P_n(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_n x^n$$

$$\text{Homog. System: } P_n(x) = c_0 + x(c_1 + x(c_2 + \dots))$$

$\Rightarrow \mathcal{O}(n)$

• Matrix A sehr schnell berechnbar.

$$\text{z.B. } z_j = 100, 101, 102$$

$$\Rightarrow A = \begin{pmatrix} 1 & 100 & 10000 \\ 1 & 101 & 10201 \\ 1 & 102 & 10609 \end{pmatrix}$$

$$\text{cond } A \approx 10^8$$

## Lagrange Interpolation

Def: Zu Stützstellen  $t_j, j=0, \dots, n-1$ , paarweise versch. def. wir das

$k$ -te Lagrangepolynom durch

$$L_k(t) = \frac{(t-t_0)\dots(t-t_{k-1})(t-t_{k+1})\dots(t-t_{n-1})}{(t_k-t_0)(t_k-t_1)\dots(t_k-t_{k-1})(t_k-t_{k+1})\dots(t_k-t_{n-1})}$$
$$= \prod_{\substack{j=0 \\ j \neq k}}^{n-1} \left( \frac{t-t_j}{t_k-t_j} \right)$$

Bem:

$L_k$  ist ein Polynom vom Grad  $k_n$

$$L_k(t_j) = \begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases}$$

für den Ansatz  $P_n = \sum_{j=0}^n c_j L_j$

erhalten wir das LGS

$$\begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} Y_0 \\ \vdots \\ Y_n \end{pmatrix}$$

$$\Rightarrow P_n(t) = Y_0 L_0(t) + Y_1 L_1(t) + \dots + Y_n L_n(t).$$

Der berechnete Polynom  $P_n$  ist genau das selbe wie das über die Newton-Basis konstruiert.

- Cond  $A = 1 \rightarrow$  stabil
- komplexität lösen des LGS  $\mathcal{O}(n)$

### Auswertung des Lagrange Interpolations-

#### Polynoms

naiv:  $\mathcal{O}(n^2)$  je Auswertungspunkt

$$\text{Knotenpolynom: } \ell(t) = \prod_{j=0}^n (t - t_j)$$

- das Knotenpolynom ist ein Polynom vom Grad  $n$  mit NS an den Stützstellen
- es gilt:

$$\ell'(t_k) = \prod_{j \neq k} (t - t_j)$$

$$\rightarrow \ell_k(t) = \frac{\ell(t)}{\ell'(t_k) (t - t_k)}$$

$$\text{mit } \lambda_j = \frac{1}{\ell'(t_j)} \quad \text{gilt}$$

$$P_n(t) = \ell(t) \sum_{j=0}^n \frac{\lambda_j}{t - t_j} \cdot Y_j$$

$\Rightarrow$  komplexität  $\mathcal{O}(n)$

$$\text{wegen } \sum_{j=0}^n \ell_j(t) = 1 \quad \text{folgt}$$

$$\ell_k(t) = \ell(t) \frac{\lambda_k}{t - t_k} \quad \text{folgt}$$



Satz (Barycentric Interpolation Formula)

$$P_n(t) = \sum_{j=0}^n \frac{\lambda_j f_j}{t - t_j} \Bigg/ \sum_{i=0}^n \frac{\lambda_i}{t - t_i} \quad (*)$$

für  $t \neq t_j$

Bem. Formel (\*) erlaubt die schnelle und stabile Auswertung von Lagrangepolynomen.

• Formel (\*) ist Skalierung invariant

Abst. der Knoten  $t_j$

• Formel (\*) ist instabil bei Extrapolation.

## Approximationsgüte

Frage:  $f: [a, b] \rightarrow \mathbb{R}$  ( $n+1$ ) mal stetig diffbar

Stützstellen  $z_j$

Stützwerte  $y_j = f(z_j)$

Interpolationspolynom:  $P_n(x) = P(x) - r_j$

$\rightarrow$  wie groß ist die maximale Fehler

$\max_{x \in [a, b]} |f(x) - P_n(x)|$  ?

Satz:  $|f(x) - P_n(x)| \leq \frac{\max_{x \in [a, b]} |f^{(n+1)}(x)|}{(n+1)!} \|x\|_{\infty}^{n+1}$

Bew. Betrachte  $s \neq z_j$

$$h(x) = f(x) - P_n(x) - \frac{f(x)}{l(x)} (f(s) - P_n(s))$$

$$\Rightarrow h(z_j) = 0 \quad z = z_j, \quad j=0, \dots, n \quad \text{und} \quad t=s$$

$$\Rightarrow \exists \xi \in [a, b] \text{ mit } h^{(n+1)}(\xi) = 0$$

$$\Rightarrow 0 = f^{(n+1)}(\xi) - 0 - \frac{(n+1)!}{l(\xi)} (f(s) - P_n(s))$$

$$\Rightarrow f(s) - P_n(s) = \frac{f^{(n+1)}(\xi)}{(n+1)!} l(s)$$

Bem: Die Approximationsgüte des Polynominterpolation hängt von der Gleichheit von  $f$  und der Lage der Stützstellen  $t_j$  ab, aber nicht von der Wahl der Polynombasis

• Die gleichm. Verteilung der Stützstellen liefert sehr schlechte Approximation.

$$\text{Für } x_j = \frac{j}{n} \quad j=0, \dots, n \quad \text{gilt}$$

$$L(t) = \text{Runge}$$

$\Rightarrow$  das Runge-Polynom hat extreme Ausläufer nahe der Ränder

Bsp: Runge Phänomen  $f(x) = \frac{1}{1+25x^2}$

Frage: Gibt es bessere Stützstellen?  
so dass das Runge-Polynom gleichmäßig approximiert?  
Extremwert?

## Chebyshev Polynome

$$T_n(t) = \cos(n \arccos(t))$$

$$\Rightarrow T_n(\cos \varphi) = \cos(n \varphi)$$

$$T_0(t) = 1$$

$$T_1(t) = t$$

$$T_2(t) = \cos 2\varphi = \cos^2 \varphi - \sin^2 \varphi$$

$$= 2\cos^2 \varphi - 1 = 2t^2 - 1$$

$$T_3(t) = \cos 3\varphi = \cos \varphi \cos 2\varphi - \sin \varphi \sin 2\varphi$$

$$= 4\cos^3 \varphi - 3\cos \varphi = 4t^3 - 3t$$

allgemein:

$$T_{n+1}(t) = 2t T_n(t) - T_{n-1}(t)$$

Eigenschaften:

•  $T_n$  ist ein Polynom vom Grad  $n$

•  $T_n(t) = 2^{n-1} t^n + \dots + n \geq 1$

•  $|T_n(t)| \leq 1$

• Nullstellen:  $t_j = \cos \frac{(2j+1)\pi}{2n}$   $j=0, \dots, n-1$

SAZ: • das Knodepolynom zu  $t_j$  ist

$$L(t) = (t-t_0)(t-t_1) \dots (t-t_{n-1}) \\ = 2^{1-n} T_n(t)$$

Für die Interpolation an Chebyshev-Knoten  $L$

gilt:

$$|f(t) - p_n(t)| \leq \frac{2^{1-n}}{(n+1)!} \max_{s \in F_{1,n}} |f^{(n+1)}(s)|$$

- Interpolation an dem Intervall  $[-1, 1]$
- Chebyshev-Interpolation ist viel besser als Interpolation an gleichm. verteilte Daten.
- die Chebyshev-Knoten sind an den Intervallenden dichter um das Über-

schwingen zu vermeiden:

• Transformation auf bel. Intervall  $[a, b]$ :

$$s \in [a, b] \Rightarrow t = \frac{2s - (a+b)}{b-a} \in [-1, 1]$$

$$t \in [-1, 1] \Rightarrow s = \frac{a+b}{2} + t \frac{b-a}{2}$$

$$\text{Chebyshev-Knoten: } s_j = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{(2j-1)\pi}{2n}$$

Merke:

• Monombasis  $\rightarrow$  schlecht konditioniert  
 $\rightarrow$  instabil

$\rightarrow$  Langsame

Lagrangebasis  $\rightarrow$  gut konditioniert

$\rightarrow$  stabil

$\rightarrow$  schnell

negativ: jedes Interpolationspunkt beeinflusst den Verlauf des ganzen Kurve.