Rechnerarithmetik

Maschinendarstellbare Zahlen M zur Basis B:

$$M = \{x \in \mathbb{R} | x = \pm 0.m_1 m_2 m_3 \dots m_n \cdot B^{\pm e_1 e_2 \dots e_l}\} \cup \{0\}$$

Dabei gilt $m_1 \neq 0$, $m_i, e_i \in \{0,1,\dots,B-1\}$ für $i \neq 0$ und $B \in \mathbb{N}$ $(B > 1)$

Der **Wert** $\hat{\omega}$ einer solchen Zahl ist definiert als

$$\hat{\omega} = \sum_{i=1}^{n} m_i B^{\hat{e}-i}, \qquad \hat{e} = \sum_{i=1}^{l} e_i B^{l-i}$$

x wird als n-stellige Gleitpunktzahl zur Basis B bezeichnet.

Beispiel: $\underbrace{0.3211}_{n=4} \cdot 4^{\underbrace{12}_{l=2}}$

1. $\hat{e} = 1 \cdot 4^1 + 2 \cdot 4^0 = 6$
2. $\hat{\omega} = 3 \cdot 4^5 + 2 \cdot 4^4 + 1 \cdot 4^3 + 1 \cdot 4^2 = 3664$

Gleitpunktzahlen

- Single Precision (32 Bit) $V = 1\,Bit$ $E = 8\,Bit$ $M = 23\,Bit$
- Double Precision (64 Bit) $V = 1\,Bit$ $E = 11\,Bit$ $M = 52\,Bit$

Bei allgemeiner Basis B gilt (**Maschinengenauigkeit** $= eps$)

$$eps := \frac{B}{2} \cdot B^{-n}, \qquad eps_{10} := 5 \cdot 10^{-n}$$

Sie bezeichnet den maximalen relativen Fehler, der durch Rundungen entstehen kann.

$$\left|\frac{rd(x) - x}{x}\right| \leq 5 \cdot 10^{-n} \;\; (da\; x \geq 10^{e-1})$$

Approximations- und Rundungsfehler

Die Maschinenzahlen sind nicht gleichmässig verteilt. Bei jedem Rechner gibt es eine grösste (x_{max}) und kleinste (x_{min}) positive Maschinenzahl.

- $x_{max} = B^{e_{max}} - B^{e_{max}-n} = (1 - B^{-n}) \cdot B^{e_{max}}$
- $x_{min} = B^{e_{min}-1}$

Definition

Gegeben sei eine Näherung $\tilde{x}$ zu einem exakten Wert x

- Absoluter Fehler $|\tilde{x} - x|$
- Relativer Fehler $\left|\frac{\tilde{x}-x}{x}\right| bzw. \frac{|\tilde{x}-x|}{|x|}$ $(x \neq 0)$

Fehlerfortpflanzung bei Funktionsauswertungen / Konditionierung

Näherung für den *absoluten* und *relativen* Fehler bei Funktionsauswertungen

$$\underbrace{|f(\tilde{x}) - f(x)|}_{absoluter\ Fehler\ von\ f(x)} \approx |f'(x)| \cdot \underbrace{|\tilde{x} - x|}_{absoluter\ Fehler\ von\ x}$$

$$\underbrace{\frac{|f(\tilde{x}) - f(x)|}{|f(x)|}}_{relativer\ Fehler\ von\ f(x)} \approx \underbrace{\frac{|f'(x)| \cdot |x|}{|f(x)|}}_{Konditionszahl\ K} \cdot \underbrace{\frac{|\tilde{x} - x|}{|x|}}_{relativer\ Fehler\ von\ x}$$

Den Faktor K nennt man *Konditionszahl.*

$$K := \frac{|f'(x)| \cdot |x|}{|f(x)|}$$

Relative Fehlervergrösserung von x, bei einer Funktionsauswertung von $f(x)$.

- Gut konditionierte Probleme Konditionszahl ist klein (≤ 1)
- Schlecht konditionierte Probleme Konditionszahl ist gross (> 1)

Nullstellenprobleme

Eine Gleichung der Form $F(x) = x$ heisst **Fixpunktgleichung**.

- Ihre Lösungen $\bar{x}$, für die $F(\bar{x}) = \bar{x}$ erfüllt ist, heissen **Fixpunkte**.

Fixpunktiteration

Gegeben sei $F: [a,b] \to \mathbb{R}$, mit $x_0 \in [a,b]$. Die rekursive Folge

$$x_{x+1} \equiv F(x_n), \qquad n = 0, 1, 2, \dots$$

Heisst Fixpunktiteration von F zum Startwert x_0.

Sei $F: [a,b] \to \mathbb{R}$ mit stetiger Ableitung F' und $\bar{x} \in [a,b]$ ein Fixpunkt von F. Dann gilt für die Fixpunktiteration $x_{n+1} = F(x_n)$

- $|F'(\bar{x})| < 1$ x_n konvergiert gegen $\bar{x}$, falls x_0 nahe genug bei $\bar{x}$ liegt *anziehend*
- $|F'(\bar{x})| > 1$ x_n konvergiert für keinen Startwert $x_0 \neq \bar{x}$ *abstossend*

Banachscher Fixpunktsatz

Sei $F: [a,b] \to [a,b]$ und es existiere eine Konstante α, wobei gilt

- α $(0 < \alpha < 1)$: *Lipschitz-Konstante*
- $\forall_{x,y} (x, y \in [a,b])$

$$|F(x) - F(y)| \leq \alpha |x - y|, \qquad \frac{|F(x) - F(y)|}{|x - y|} \leq \alpha$$

Dann gilt

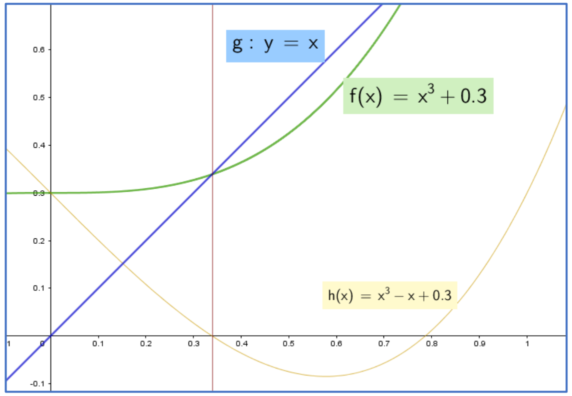
- F hat genau einen Fixpunkt $\bar{x}$ in $[a,b]$
- Die Fixpunktiteration $x_{n+1} = F(x_n)$ konvergiert gegen $\bar{x}$ für alle Startwerte $x_0 \in [a,b]$
- Es gelten die Fehlerabschätzungen
  - $|x_n - \bar{x}| \leq \frac{\alpha^n}{1-\alpha} \cdot |x_1 - x_0|$ a-priori Abschätzung
  - $|x_n - \bar{x}| \leq \frac{\alpha}{1-\alpha} \cdot |x_n - x_{n-1}|$ a-posteriori Abschätzung

Berechne die Nullstellen von $p(x) = x^3 - x + 0.3$

<u>Fixpunktiteration</u>

$$x_{n+1} = F(x_n) = x_n^3 + 0.3$$

$F(x_n)$ steigt stetig an



$F: I \to I$ gilt wenn…

$$F(a) > a, \qquad F(b) < b$$

Alpha bestimmen / überprüfen

$$\alpha = \max_{x \in I} |F'(x)| \leq 1$$

Anzahl Iterationen n berechnen

$$n \geq \frac{\ln\left(\frac{tol \cdot (1 - \alpha)}{|x_1 - x_0|}\right)}{\ln \alpha}$$

Newton-Verfahren

Sukzessive Approximation der Funktionskurve $y = f(x)$ durch Tangenten, deren Schnittpunkt mit der x-Achse problemlos berechnet werden kann.

Lösung ξ der Gleichung $f(x) = 0$ finden.

- Startwert x_0 geeignet wählen (nahe bei ξ)
- Iterationsvorschrift:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Die Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert gegen die Lösung ξ der Gleichung $f(x) = 0$.

$(x_0, x_1, x_2, \dots)$ ist sicher gegeben, wenn im Intervall $[a, b]$, in dem alle Näherungswerte (und die Nullstellen selbst) liegen sollen, die Bedingung

$$\left|\frac{f(x) \cdot f''(x)}{[f'(x)]^2}\right| < 1$$

Erfüllt ist (hinreichende Konvergenzbedingung).

Vereinfachtes Newton-Verfahren

Statt in jedem Schritt $f'(x_n)$ auszurechnen, kann man immer wieder $f'(x_0)$ verwenden.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}$$

Sekantenverfahren

Der Schnittpunkt von Sekanten durch jeweils zwei Punkte $(x_0, f(x_0))$ und $(x_1, f(x_1))$ mit der x –Achse, wird berechnet.

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \cdot f(x_n)$$

Konvergenzgeschwindigkeit

Sei (x_n) eine gegen $\bar{x}$ konvergierende Folge. Dann hat das Verfahren die ***Konvergenzordnung*** $q \geq 1$ wenn es eine Konstante $c > 0$ gibt mit

$$|x_{n+1} - \bar{x}| \leq c \cdot |x_n - \bar{x}|^q$$

Für alle n.

- $q = 1$ lineare Konvergenz verlangt man noch $c < 1$.
- $q = 2$ quadratische Konvergenz

Fehlerabschätzung

Nullstellensatz von Bolzano

Sei $f: [a, b] \to \mathbb{R}$ stetig mit $f(a) \leq 0 \leq f(b)$ oder $f(a) \geq 0 \geq f(b)$. Dann muss f in $[a, b]$ eine Nullstelle besitzen.

Sei x_n also ein iterativ bestimmter Näherungswert einer exakten Nullstelle ξ der stetigen Funktion $F: \mathbb{R} \to \mathbb{R}$ und es gelte für ein vorgegebene Fehlerschranke / Fehlertolerant $\epsilon > 0$

$$f(x_n - \epsilon) \cdot f(x_n + \epsilon) < 0$$

Dann muss gemäss dem Nullstellensatz im offenen Intervall $(x_n - \epsilon, x_n + \epsilon)$ eine Nullstelle ξ liegen und es gilt die Fehlerabschätzung

$$|x_n - \xi| < \epsilon$$

Lineare Gleichungssysteme – Gauss-Algorithmus

Gauss-Algorithmus für ein Gleichungssystem $Ax = b$:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n}, \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$$

Umformung des Gleichungssystems $Ax = b$, in ein äquivalentes Gleichungssystem $\tilde{A}x = b$, so dass die Matrix $\tilde{A}$ als *obere Dreiecksmatrix* vorliegt.

- $z_j \coloneqq z_j - \lambda z_i$ $\quad i < j\ (\lambda \in \mathbb{R})$, z_i ist die i-te Zeile des Gleichungssystems
- $z_i \to z_j$ $\quad$ Vertauschen der i-ten und j-ten Zeile im System

Rekursive Vorschrift für ein Gleichungssystem $\tilde{A}x = b$:

$$x_n = \frac{b_n}{a_{nn}}, x_{n-1} = \frac{b_{n-1} - a_{(n-1)\,n} \cdot x_n}{a_{n-1\,n-1}}, \dots, x_1 = \frac{b_1 - a_{12} \cdot x_2 - \cdots - a_{1n} \cdot x_n}{a_{11}}$$

$$x_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij} \cdot x_j}{a_{ii}}, \qquad i = n, n-1, \dots, 1$$

Fehlerfortpflanzung und Pivotisierung

Für $i = 1, .. n$:

Erzeuge Nullen unterhalb des Diagonalelements in der i-ten Spalte

- Suche das betragsgrösste Element unterhalb der Diagonalen in der i-ten Spalte:
  Wähle k so, dass $|a_{ki}| = \max\{|a_{ji}|\,|\,j = i, \dots n\}$
  $$\begin{cases} falls\ a_{ki} = 0\text{:} & A\ ist\ nicht\ regulär; stop; \\ falls\ a_{ki} \neq 0\text{:} & z_k \leftrightarrow z_i \end{cases}$$
- Eliminationsschritt:
  Für $j = i + 1, \dots, n$ eliminiere das Element a_{ji} durch
  $$z_j \coloneqq z_j - \frac{a_{ji}}{a_{ii}} \cdot z_i$$

Determinante

Gegeben sei eine Matrix A, woraus die obere Dreiecksmatrix $\tilde{A}$ entsteht.

- $\tilde{a}_{ii}$: $\quad$ Diagonalelemente von $\tilde{A}$
- l: $\quad$ Anzahl Zeilenvertauschungen

$$\det(A) = (-1)^l \cdot \det(\tilde{A}) = (-1)^l \cdot \prod_{i=1}^{n} \tilde{a}_{ii}$$

<u>Beispiel</u>

$$\begin{pmatrix} 3 & 5 & 1 \\ 0 & 2 & 2 \\ 6 & 14 & 8 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 1 \\ 0 & 2 & 2 \\ 0 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\det(A) = (3) \cdot (2) \cdot (2) = 12$$

Lineare Gleichungssysteme – LR-Zerlegung

Das ursprüngliche Gleichungssystem $\boldsymbol{Ax} = \boldsymbol{b}$ lautet mit der *LR-Zerlegung*

$$LRx = b \Leftrightarrow Ly = b \text{ und } Rx = y$$

Für eine $n \times n$ Matrix A, gibt es $n \times n$ Matrizen L und R mit den Eigenschaften

- L ist eine *normierte untere Dreiecksmatrix* mit $l_{ii} = 1 \ (i = 1, \dots, n)$
- R ist eine *obere Dreiecksmatrix* mit $r_{ii} \neq 0 \ (i = 1, \dots, n)$
- $A = L \cdot R$ ist die *LR-Zerlegung* von A.

Zerlegung mit Zeilenvertauschung

P_K erhält man aus der Einheitsmatrix I_n durch Vertauschen der *i-ten* und *j-ten* Zeile.

Zeilen-Vertauschungen werden durch $P_1 \dots P_n$ ausgedrückt.

$$P = \prod_{i=1}^{n} P_{n-i+1}$$

Mit dieser Permutationsmatrix erhält man dann als $RL - Zerlegung$

$$PA = LR$$

Das lineare Gleichungssystem $Ax = b$ lässt sich schreiben als $PAx = Pb$ bzw. $LRx = Pb$ und in den zwei Schritten lösen

$$Ly = Pb \rightarrow y = \cdots,$$

$$Rx = y \rightarrow x = \cdots$$

Vertauschung der 1. Und 3. Zeile bei der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \rightarrow A^* = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

$$I^* \cdot A = P_1 \cdot A = A^* = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -3 & -2 \\ 5 & 1 & 4 \end{pmatrix} = LR$$

$$i = 1, j = 2 \rightarrow z_2 = z_2 - \frac{1}{\underbrace{(-1)}_{l_{21}}} \cdot z_1 \rightarrow A_1 = \begin{pmatrix} -1 & 1 & 1 \\ 1 - 1 & -3 + 1 & -1 + 1 \\ 5 & 1 & 4 \end{pmatrix}$$

$$i = 1, j = 3 \rightarrow z_3 = z_3 - \frac{5}{\underbrace{(-1)}_{l_{31}}} \cdot z_1 \rightarrow A_2 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & -1 \\ 5 - 5 & 1 + 5 & 4 + 5 \end{pmatrix}$$

$$i = 2, j = 3 \rightarrow z_3 \equiv z_3 - \frac{6}{\underbrace{(-2)}_{l_{32}}} \cdot z_2 \rightarrow A_3 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 + 0 & 6 - 6 & 9 - 3 \end{pmatrix}$$

$$R = A_3 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 6 \end{pmatrix}$$

Einsetzen in L

$$l_{21} = \frac{1}{-1} = -1, \qquad l_{31} = \frac{5}{-1} = -5, \qquad l_{32} = \frac{6}{-2} = -3$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 5 & -3 & 1 \end{pmatrix}$$

Lineare Gleichungssysteme – QR-Zerlegung

Eine Matrix $Q \in \mathbb{R}^{n\times n}$ heisst orthogonal, wenn $Q^T \cdot Q = I_n$ ist. Dabei ist I_n die $n \times n$ Einheitsmatrix.

Sei $A \in \mathbb{R}^{n\times n}$. Eine *QR-Zerlegung* von A ist eine Darstellung von A als Produkt einer orthogonalen $n \times n$ Matrix Q und einer rechtsoberen $n \times n$ Dreiecksmatrix R

$$A = QR$$

Lösung des Gleichungssystems

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

Die orthogonale $n \times n$ Matrix heisst ***Householder-Matrix***.

$$H := I_n - 2uu^T, \qquad |u| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} = 1$$

Householder-Matrix zum Vektor u

$$u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad \tilde{u} = \frac{u}{|u|} = \frac{1}{\sqrt{14}} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$H = I_n - 2\tilde{u}\tilde{u}^T = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{I_n} - 2 \cdot \frac{1}{14} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$$

Algorithmus zur QR-Zerlegung

$$R := A, \qquad Q := I_n$$

Für $i = 1, \ldots, n-1$:
erzeuge Nullen in R in der *i-ten* Spalte unterhalb der Diagonalen

1. H_i mit $(n-i+1) \times (n-i+1)$ berechnen
2. H_i mit I_{i-1} Block links oben erweitern $\to Q_i$
3. $R := Q_i \cdot R$
4. $Q := Q \cdot Q_i^T$

Ablauf

$$H_1 \cdot A_1 = H_1 \cdot \underbrace{\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}}_{A_1} = \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix} \to \underbrace{\begin{pmatrix} * & * \\ * & * \end{pmatrix}}_{A_2}$$

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \qquad e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

1. $v_1 := a_1 + sign(a_{11}) \cdot |a_1| \cdot e_1$
2. $u_1 := \frac{1}{|v_1|} \cdot v_1$
3. $H_1 := I_n - 2u_1u_1^T = Q_1$

$$H_2 \cdot A_2 = H_2 \cdot \underbrace{\begin{pmatrix} * & * \\ * & * \end{pmatrix}}_{A_2} = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$$

$$Q_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & H_2 & H_2 \\ 0 & H_2 & H_2 \end{pmatrix}$$

$$Q := Q_1^T \cdot Q_2^T, \qquad R := \underbrace{Q_2 \cdot Q_1}_{Q^{-1}} \cdot A$$

Lineare Gleichungssysteme – Fehlerberechnung

Eine Abbildung $\|.\|: \mathbb{R}^n \to \mathbb{R}$ heisst ***Vektornorm***, wenn die folgenden Bedingungen für alle $x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$ erfüllt sind:

- $\|x\| \geq 0$ und $\|x\| = 0 \Leftrightarrow x = 0$
- $\|\lambda x\| = |\lambda| \cdot \|x\|$
- $\|x + y\| \leq \|x\| + \|y\|$ "$Dreiecksgleichung$"

Für Vektoren $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ gibt es die folgenden ***Vektornormen***

- 1-Norm Summennorm $\|x\|_1 = \sum_{i=1}^n |x_i|$
- 2-Norm Euklidische Norm $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$
- ∞-Norm Maximumnorm $\|x\|_\infty = \max\limits_{i=1,\dots,n} |x_i|$

Für eine $n \times n$ Matrix $A \in \mathbb{R}^{n \times n}$ gibt es die folgenden ***Matrixnormen***

- 1-Norm Spaltensummennorm $\|A\|_1 = \max\limits_{j-1,\dots,n} \sum_{i=1}^n |x_i|$
- 2-Norm Spektralnorm $\|A\|_2 = \sqrt{\rho(A^T A)}$
- ∞-Norm Zeilensummennorm $\|A\|_\infty = \max\limits_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$

Abschätzung für Fehlerhafte Matrizen

Sei $\|.\|$ eine Norm, $A, \tilde{A} \in \mathbb{R}^{n \times n}$ eine reguläre $n \times n$ Matrix und $x, \tilde{x}, b, \tilde{b} \in \mathbb{R}^n$ mit $Ax = b$ und $\tilde{A}\tilde{x} = \tilde{b}$. Falls

$$cond(A) \cdot \frac{\|A - \tilde{A}\|}{\|A\|} < 1$$

Dann gilt

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \frac{cond(A)}{1 - cond(A) \cdot \frac{\|A - \tilde{A}\|}{\|A\|}} \cdot \left(\frac{\|A - \tilde{A}\|}{\|A\|} + \frac{\|b - \tilde{b}\|}{\|b\|}\right)$$

Abschätzung für Fehlerhafte Vektoren

Sei $\|.\|$ eine Norm, $A \in \mathbb{R}^{n \times n}$ eine reguläre $n \times n$ Matrix und $x, \tilde{x}, b, \tilde{b} \in \mathbb{R}^n$ mit $Ax = b$ und $A\tilde{x} = \tilde{b}$. Dann gilt für den absoluten und den relativen Fehler in x:

- $\|x - \tilde{x}\| \leq \|A^{-1}\| \cdot \|b - \tilde{b}\|$
- $\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|b - \tilde{b}\|}{\|b\|}$, $falls\ \|b\| \neq 0$

Die Zahl $cond(A) = \|A\| \cdot \|A^{-1}\|$ nennt man Konditionszahl der Matrix A

- $cond(A)$ gross → schlechte Konditionierung

Untersuchen Sie die Fehlerfortpflanzung im linearen Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 2 & 4 \\ 4 & 8.1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 1.5 \end{pmatrix}$$

Für den Fall, dass die rechte Seite von $\tilde{b}$ in jeder Komponente um maximal 0.1 von b abweicht.

$$\|\tilde{b} - b\|_\infty \leq 0.1, \qquad \|A\|_\infty = \max\{2 + 4,\ \ 4 + 8.1\} = 12.1$$

$$\|A^{-1}\|_\infty = \left\|\begin{pmatrix} 40.5 & -20 \\ -20 & 10 \end{pmatrix}\right\|_\infty = 60.5$$

$$cond(A)_\infty = \|A\|_\infty \cdot \|A^{-1}\|_\infty = 12.1 \cdot 60.5 = 732.05$$

$$\|x - \tilde{x}\|_\infty \leq \|A^{-1}\|_\infty \cdot \|b - \tilde{b}\|_\infty \leq 60.5 \cdot 0.1 = \underbrace{6.05}_{absoluter\ Fehler}$$

$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} \leq cond(A)_\infty \cdot \frac{\|b - \tilde{b}\|_\infty}{\|b\|_\infty} \leq 732 \cdot \frac{0.1}{1.5} = \underbrace{48.8}_{relativer\ Fehler}$$

Lineare Gleichungssysteme – Fehlerberechnung und Aufwandschätzung

Aufwandschätzung

Die Anzahl Gleitkommaoperationen werden in Abhängigkeit von n bestimmt.

$$\sum_{i=1}^{n} i = \frac{(n+1)\cdot n}{2} \; und \; \sum_{i=1}^{n} i^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n, \qquad n = Dimension$$

Ein Algorithmus hat die Ordnung $O(n^q)$, wenn $q > 0$ die minimale Zahl ist, für die es eine Konstante $C > 0$ gibt, so dass der Algorithmus für alle $n \in N$ weniger als

<u>Beispiel</u>

Wie viele Gleitkommaoperationen benötigt das Rückwärtseinsetzen gemäss Gauss?

$$x_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij}\cdot x_j}{a_{ii}}, \qquad i = n, n-1, \dots, 1$$

Multiplikation und Division

$$1 + 2 + 3 + \cdots + n = \sum_{i=1}^{n} i = \frac{(n+1)\cdot n}{2}$$

Addition und Subtraktion

$$0 + 1 + 2 + \cdots + n - 1 = \sum_{i=1}^{n-1} i = \frac{(n-1+1)\cdot(n-1)}{2} = \frac{(n-1)\cdot n}{2}$$

Summe beider Operationstypen

$$\frac{n^2}{2} + \frac{n}{2} + \frac{n^2}{2} - \frac{n}{2} = n^2$$

Lineare Gleichungssysteme – Iterative Verfahren

Iterative Verfahren sind effizienter, jedoch kann man keine genauen Lösungen erwarten. Ausgehend von einem Startvektor $x^{(0)}$ berechnet man mittels einer Rechenvorschrift $F: \mathbb{R}^n \to \mathbb{R}^n$ iterativ eine Folge von Vektoren

$$x^{(k+1)} = F\left(x^{(k)}\right) \; mit \; k = 0,1,2,\dots$$

Jacobi-Verfahren / Jacobi-Verfahren und **Gauss-Seidel-Verfahren**

Zu lösen sei $Ax = b$. Die Matrix $A = \left(a_{ij}\right)$ sei zerlegt in der Form $A = L + D + R =$

$$\underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{nn-1} & 0 \end{pmatrix}}_{=:L} + \underbrace{\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{pmatrix}}_{=:D} + \underbrace{\begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}}_{=:R}$$

<u>Beispiel</u> mit *Jacobi*

$$Ax = b, \qquad A = \begin{pmatrix} 8 & 5 & 2 \\ 5 & 9 & 1 \\ 4 & 2 & 7 \end{pmatrix}, \qquad b = \begin{pmatrix} 19 \\ 5 \\ 34 \end{pmatrix}, \qquad x^{(0)} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$$

$$x_1^{(1)} = \frac{1}{8}\left(19 - \sum_{j=1, j\neq 1}^{3} a_{1j} \cdot x_j^{(0)}\right) = \frac{1}{8}(19 - (5 \cdot -1 + 2 \cdot 3)) = \frac{18}{8}$$

$$x_2^{(1)} = \frac{1}{9}\left(5 - \sum_{j=1, j\neq 2}^{3} a_{2j} \cdot x_j^{(0)}\right) = \frac{1}{9}(5 - (5 \cdot 1 + 1 \cdot 3)) = -\frac{1}{3}$$

$$x_3^{(1)} = \frac{1}{7}\left(34 - \sum_{j=1, j\neq 3}^{3} a_{3j} \cdot x_j^{(0)}\right) = \frac{1}{7}(34 - (4 \cdot 1 + 2 \cdot -1)) = \frac{32}{7}$$

Fixpunktiteration gemäss *Jacobi (Gesamtschritt-Verfahren)*:

$$Dx^{(k+1)} = -(L + R)x^{(k)} + b$$

$$x^{(k+1)} = -D^{-1}(L + R)x^{(k)} + D^{-1}b$$

Implementation /Allgemeine Form gemäss *Jacobi*

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

Fixpunktiteration gemäss *Gauss-Seidel (Einzelschritt-Verfahren)*:

$$(D + L)x^{(k+1)} = -Rx^{(k)} + b$$

$$x^{(k+1)} = -(D + L)^{-1} \cdot Rx^{(k)} + (D + L)^{-1} \cdot b$$

Implementation / Allgemeine Form gemäss *Gauss-Seidel*

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

Lineare Gleichungssysteme – Iterative Verfahren

Gegeben sei eine Fixpunktiteration $$x^{(n+1)} = Bx^{(n)} + c =: F\left(x^{(n)}\right)$$	Für das Gesamtschrittverfahren (Jacobi) gilt $B = -D^{-1}(L+R)$ Für das Einzelschrittverfahren (Gauss-Seidel) gilt $B = -(D+L)^{-1}R$
Und $\bar{x} \in \mathbb{R}^n$ sei bezüglich der Norm $\lVert . \rVert$ anziehender Fixpunkt. Dann konvergiert die Fixpunktiteration für alle Startvektoren $x^{(0)} \in \mathbb{R}^n$ gegen $\bar{x}$ und es gelten die Abschätzungen • $\lVert x^{(n)} - \bar{x} \rVert \leq \frac{\lVert B \rVert^n}{1-\lVert B \rVert} \cdot \lVert x^{(1)} - x^{(0)} \rVert$ *a-priori Abschätzung* • $\lVert x^{(n)} - \bar{x} \rVert \leq \frac{\lVert B \rVert}{1-\lVert B \rVert} \cdot \lVert x^{(n)} - x^{(n-1)} \rVert$ *a-posteriori Abschätzung*	Wobei B eine $n \times n$ Matrix ist und $c \in \mathbb{R}^n$. Weiter sei $\lVert . \rVert$ eine der eingeführten Normen und $\bar{x} \in \mathbb{R}^n$ erfülle $\bar{x} = B\bar{x} + c = F(\bar{x})$. Dann heisst • *$\bar{x}$ anziehender Fixpunkt,* *falls* $\lVert B \rVert < 1$ • *$\bar{x}$ abstossender Fixpunkt,* *falls* $\lVert B \rVert > 1$
A ist eine ***diagonaldominante*** Matrix, falls **eines** der beiden folgenden Kriterien gilt • $\textit{für alle } i = 1, \dots, n\colon \lvert a_{ii} \rvert > \sum_{j=1, j\neq i}^{n} \lvert a_{i,j} \rvert$ (Zeilensummenkriterium) • $\textit{für alle } j = 1, \dots, n\colon \lvert a_{jj} \rvert > \sum_{i=1, i\neq j}^{n} \lvert a_{i,j} \rvert$ (Spaltensummenkriterium) <u>Beispiel</u> $$A = \begin{pmatrix} 4 & -1 & 1 \\ -2 & 5 & 1 \\ 1 & -2 & 5 \end{pmatrix} \rightarrow \sum_{j=1, j\neq i}^{n} \lvert a_{ij} \rvert \rightarrow \begin{cases} i = 1 \rightarrow 4 > 2 \\ i = 2 \rightarrow 5 > 3 \\ i = 3 \rightarrow 5 > 3 \end{cases}$$ Fall A *diagonaldominant* ist, konvergiert das Gesamtschrittverfahren (Jacobi) und auch das Einzelschrittverfahren (Gauss-Seidel) für $Ax = b$. Ein notwendiges und hinreichendes Kriterium für Konvergenz ist $$\textit{Spektralradius } \rho(B) < 1$$	

Lineare Gleichungssysteme - Komplexe Zahlen

Die Menge der komplexen Zahlen $\mathbb{C}$ erweitert die Menge der reellen Zahlen $\mathbb{R}$, so dass nun also auch Gleichungen der folgenden Art lösbar werden

$$x^2 + 1 = 0$$

Dafür wird die imaginäre Einheit i mit der folgenden Eigenschaft eingeführt.

$$i^2 = -1$$

Eine komplexe Zahl z ist ein geordnetes Paar(x, y) zweier Zahlen x und y.

$$z = x + iy$$

Die imaginäre Einheit i ist definiert durch

$$i^2 = -1$$

Die Menge der komplexen Zahlen wird mit $\mathbb{C}$ bezeichnet

$$\mathbb{C} = \{z | z = x + iy \; mit \; x, y \in \mathbb{R}\}$$

Die reellen Bestandteile x und y von z werden als Real- und Imaginärteil bezeichnet

- *Realteil* von z $\quad Re(z) = x$
- *Imaginärteil* von z $\quad Im(z) = y$

Die zu z konjugierte komplexe Zahl ist definiert als $z^* = x - iy$. Dies entspricht der an der $x -$ Achse gespiegelten Zahl.

Der Betrag einer komplexen Zahl ist definiert als $|z| = \sqrt{x^2 + y^2} = \sqrt{z \cdot z^*}$. Dies entspricht der Länge des Zeigers.

Darstellungsformen

- Normalform $\quad z = x + iy$
- Trigonometrische Form $\quad z = r(\cos\varphi + i \cdot \sin\varphi)$
- Exponentialform $\quad z = re^{i\varphi}$

$$x = r \cdot \cos\varphi\,, \qquad y = r \cdot \sin\varphi\,, \qquad r = \sqrt{x^2 + y^2}$$

$$\varphi = \arcsin\left(\frac{y}{r}\right) = \arccos\left(\frac{x}{r}\right)$$

$$e^{i\varphi} = \cos\varphi + i \cdot \sin\varphi$$

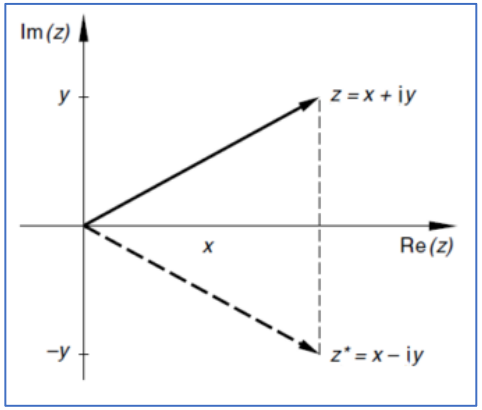
<u>Beispiel</u>

$$z = 3 - 11i$$

$$3 = r \cdot \cos\varphi\,, \qquad 11 = r \cdot \sin\varphi\,, \qquad r = \sqrt{3^2 + 11^2} = \sqrt{130}$$

$$\arcsin\left(\frac{11}{\sqrt{130}}\right) = \varphi = 1.3$$

$$z = \cos(1.3) + i \cdot \sin(1.3)\,, \qquad z = \sqrt{130} \cdot e^{i \cdot 1.3}$$



Lineare Gleichungssysteme - Komplexe Zahlen

Grundrechenarten

Es sei $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$

- Summation $\quad z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$
- Subtraktion $\quad z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2)$

Multiplikation

$$z_1 \cdot z_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

$$z_1 \cdot z_2 = r_1e^{i\varphi_1} \cdot r_2e^{i\varphi_2} = r_1r_2e^{i(\varphi_1+\varphi_2)}$$

Division

$$\frac{z_1}{z_2} = \frac{z_1 \cdot z_2^*}{z_2 \cdot z_2^*} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)}$$

$$= \frac{(x_1x_2 + y_1y_2) + i(y_1x_2 - x_1y_2)}{x_2^2 + y_2^2} = \frac{(x_1x_2 + y_1y_2)}{x_2^2 + y_2^2} + i\frac{(y_1x_2 - x_1y_2)}{x_2^2 + y_2^2}$$

$$\frac{z_1}{z_2} = \frac{r_1e^{i\varphi_1}}{r_2e^{i\varphi_2}} = \frac{r_1}{r_2}e^{i(\varphi_1+\varphi_2)}$$

Potenzieren und Radizieren

Die n-te Potenz einer komplexen Zahl lässt sich einfach berechnen, wenn diese in der trigonometrischen oder der Exponentialform vorliegt (Sei $n \in \mathbb{N}$):

$$z = r \cdot e^{i\varphi} \to z^n = \left(re^{i\varphi}\right)^n = r^ne^{in\varphi} = r^n(\cos(n\varphi) + i \cdot \sin(n\varphi))$$

Fundamentalgesetz der Algebra

Eine algebraische Gleichung n-ten Grades mit komplexen Koeffizienten und Variablen $a_i, z \in \mathbb{C}$

$$a_nz^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0 = 0$$

Besitzt in der Menge $\mathbb{C}$ der komplexen Zahlen genau n Lösungen

Wurzel einer komplexen Zahl

Eine komplexe Zahl z wird als n-te Wurzel von $a \in \mathbb{C}$ bezeichnet, wenn

$$z^n = a \to z = \sqrt[n]{a}$$

Lösungen der algebraischen Gleichung $z^n = a$

$$z^n = a = r_0e^{i\varphi} \; (r_0 > 0; n = 2,3,4, \dots)$$

Besitzt in der Menge $\mathbb{C}$ genau n verschiedene Lösungen (Wurzeln)

$$z_k = r(\cos\varphi_k + i \cdot \sin\varphi_k) = re^{i\varphi_k}$$

$$r = \sqrt[n]{r_0}, \qquad \varphi_k = \frac{\varphi + k \cdot 2\pi}{n}, \qquad (f\ddot{u}r\ k = 0,1,2, \dots, n-1)$$

Die zugehörigen Bildpunkte liegen in der komplexen Zahlenebene auf einem Kreis um den Nullpunkt mit dem Radius $r = \sqrt[n]{r_0}$ und bilden die Ecken eines regelmässigen n-Ecks.

Lineare Gleichungssysteme – Eigenwerte

Es sei $A \in \mathbb{R}^{n \times n}$. $\lambda \in \mathbb{C}$ heisst *Eigenwert* von A, wenn es einen Vektor $x \in \mathbb{C}^n \backslash \{0\}$ gibt mit

$$Ax = \lambda x$$

x heisst dann *Eigenvektor* von A.

Eigenschaften von Eigenwerten

$$Ax - \lambda x = 0 \Leftrightarrow (A - \lambda I_n) \cdot x = 0$$

Die Eigenwerte einer Diagonal- oder eine Dreiecksmatrix sind deren Diagonalelemente.

Polynom und Spur

Es sei $A \in \mathbb{R}^{n \times n}, \lambda \in \mathbb{C}$. Dann gilt

$$\lambda \text{ ist ein Eigenwert von } A \Leftrightarrow \det(A - \lambda I_n) = 0$$

Die Abbildung p ist definiert durch

$$p(\lambda) \rightarrow \det(A - \lambda I_n)$$

Ist ein Polynom vom Grad n und wird *charakteristisches Polynom* von A genannt. Die Eigenwerte von A sind also die Nullstellen des charakteristischen Polynoms. Damit hat A also genau n Eigenwerte, von denen manche mehrfach vorliegen können.

Die Determinante der Matrix A ist gerade das Produkt ihrer Eigenwerte $\lambda_1, \ldots, \lambda_n$. Die Summe der Eigenwerte ist gleich der Summe der Diagonalelemente von A, d.h. gleich der *Spur* (tr) von A:

- $\det(A) = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n$
- $tr(A) = a_{11} + a_{22} + \cdots + a_{nn} = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Ist λ_i ein Eigenwert der regulären Matrix A, so ist der Kehrwert $\frac{1}{\lambda_i}$ ein Eigenwert der inversen Matrix A^{-1}.

Vielfachheit und Spektrum

Es sei $A \in \mathbb{R}^{n \times n}$. Die *Vielfachheit*, mit der λ als Nullstelle des charakteristischen Polynoms von A auftritt, heisst *algebraische Vielfachheit* von λ.

Das *Spektrum* $\sigma(A)$ ist die Menge aller Eigenwerte von A.

<u>Beispiel</u>

Berechne *Spektrum*, *Determinante* und *Spur* von

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

Eigenwerte

$$\lambda_1 = 1, \qquad \lambda_2 = 3, \qquad \lambda_3 = 2$$

Determinante

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 = 6$$

Spur

$$tr(A) = \lambda_1 + \lambda_2 + \lambda_3 = 6$$

Spektrum

$$\sigma(A) = 3$$

Eigenschaften von Eigenvektoren

Seien zwei Eigenvektoren x, y zum selben Eigenwert $\lambda \in \mathbb{C}$ einer Matrix $A \in \mathbb{R}^n \times \mathbb{R}^n$, so ist $x + y$ und auch jedes Vielfach von x ebenfalls ein Eigenvektor zum Eigenwert λ:

$$A(x+y) = Ax + Ay = \lambda x + \lambda = \lambda(x+y)$$

$$A(\mu x) = \mu A x = \mu \lambda x = \lambda \mu x$$

Eigenraum

Sei $\lambda \in \mathbb{C}$ ein Eigenwert von $A \in \mathbb{R}^{n \times n}$. Dann bilden die Eigenvektoren zum Eigenwert λ zusammen mit dem Nullvektor 0 einen Unterraum von $\mathbb{C}^n$, den sogenannten *Eigenraum*.

Der Eigenraum des Eigenwertes λ ist die Lösungsmenge des homogenen LGS

$$(A - \lambda I_n)x = 0$$

Welches nur dann eine nicht-triviale Lösung aufweist, wenn $rg(A - \lambda I_n) < n$.

Die *Dimension* des Eigenraumes von λ wird die *geometrische Vielfachheit* von λ genannt. Sie berechnet sich als

$$n - rg(A - \lambda I_n)$$

Und gibt die Anzahl der lin. Unabhängigen Eigenvektoren zum Eigenwert λ.

Geometrische und algebraische Vielfachheit eines Eigenwerts müssen nicht gleich sein. Die geom. Vielfachheit ist aber stets kleiner oder gleich der algebraischen Vielfachheit.

Beispiel: Berechne *Eigenwerte, Eigenvektoren, Eigenräume*

$$A = \begin{pmatrix} 2 & 5 \\ -1 & -2 \end{pmatrix}, \qquad A - \lambda I_n = \begin{pmatrix} 2-\lambda & 5 \\ -1 & -2-\lambda \end{pmatrix}$$

$$p(\lambda) = \det(A - \lambda I_n) = (2-\lambda)(-2-\lambda) - 5 \cdot -1$$

$$p(\lambda) = -4 + \lambda^2 + 5 = \lambda^2 + 1 = 0$$

$$\lambda^2 = -1 = i^2$$

Eigenwerte

$$\lambda_1 = i, \qquad \lambda_2 = -i$$

Eigenvektor für $\boldsymbol{\lambda_1 = i}$

$$\begin{pmatrix} 2-i & 5 \\ -1 & -2-i \end{pmatrix} \rightarrow \begin{pmatrix} 2-i & 5 \\ 0 & -2-i+\frac{5}{2-i} \end{pmatrix}$$

$$-2-i+\frac{5}{2-i} = (2-i)(-2-i) + 5 = 1 + i^2 = 0$$

$$0 = (2-i) \cdot x_1 + 5 \cdot x_2$$

$$x_1 = -\frac{5x_2}{2-i} \cdot \frac{2+i}{2+i} = -\frac{5 \cdot (2+i)}{4-i^2} = -\frac{10+5i}{5} = -2-i$$

$$x_1 = \begin{pmatrix} -2-i \\ 1 \end{pmatrix}$$

Eigenraum

$$E_{\lambda_1} = \left\{ x \middle| x = \mu = \begin{pmatrix} -2-i \\ 1 \end{pmatrix}, \mu \in \mathbb{R} \right\}$$

$$E_{\lambda_2} = \left\{ x \middle| x = \mu = \begin{pmatrix} -2+i \\ 1 \end{pmatrix}, \mu \in \mathbb{R} \right\}$$

Lineare Gleichungssysteme – Numerische Berechnung von Eigenvektoren und Eigenwerten

Ähnliche Matrizen / Diagonalisierbarkeit

Es seien $A, B \in \mathbb{R}^{n\times n}$ und T eine reguläre Matrix mit … so heissen B und A zueinander *ähnliche Matrizen*.

$$B = T^{-1}AT$$

Im Spezialfall, dass $B = D$ ein Diagonalmatrix ist, also … nennt man A diagonalisierbar.

$$D = T^{-1}AT$$

Eigenwerte und Eigenvektoren ähnlicher / diagonalisierbarer Matrizen

Es seien $A, B \in \mathbb{R}^{n\times n}$ zueinander ähnliche Matrizen. Dann gilt

1. A und B haben dieselben Eigenwerte, inkl. deren algebraische Vielfachheit
2. Ist x ein Eigenvektor zum Eigenwert λ von B, dann ist Tx ein Eigenvektor zum Eigenwert λ von A.
3. Falls A diagonalisierbar ist
   - Diagonalelemente von D sind die Eigenwerte von A
   - Die linear unabhängigen Eigenvektoren von A stehen in den Spalten von T

Der ***Spektralradius*** $p(A)$ einer Matrix $A \in \mathbb{R}^{n\times n}$ ist definiert als

$$p(A) = \max\{|\lambda| | \lambda \text{ } ist\ ein\ Eigenwert\ von\ A \in \mathbb{R}^{n\times n}\}$$

Sei $A \in \mathbb{R}^{n\times n}$ eine diagonalisierbare Matrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ und dem betragsmässig grössten Eigenwert λ_1 mit

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

Vektoriteration / von-Mises-Iteration

So konvergieren für (fast) jeden Startvektor $v^{(0)} \in \mathbb{C}^n$ mit Länge 1 die Folgen

$$v^{(k+1)} = \frac{Av^{(k)}}{\left\|Av^{(k)}\right\|_2}, \qquad \lambda^{(k+1)} = \frac{\left(v^{(k)}\right)^T Av^{(k)}}{(v^{(k)})^T v^{(k)}}$$

Für $k \to \infty$ gegen einen Eigenvektor v zum Eigenwert λ_1 von A (also $v^{(k)} \to v$ und $\lambda^{(k)} \to \lambda_1$)

***QR*-Verfahren**

Sei $A \in R^{n\times n}$

$$A_0 \coloneqq A, \qquad P_0 \coloneqq I_n$$

Für $i = 0, 1, 2, \dots$:

- $A_i \coloneqq Q_i \cdot R_i$ *QR*-Zerlegung von A_i
- $A_{i+1} \coloneqq R_i \cdot Q_i$
- $P_{i+1} \coloneqq P_i \cdot Q_i$