Numerisches Programmieren (IN0019)

Frank R. Schmidt

Winter Semester 2016/2017

3. Lineare Gleichungssysteme 2

Lineare Gleichungen **3**

Runden und Kondition (Wdh.) 4
Parabolische Flugbahn (Beispiel) 5
Parabolische Flugbahn (Beispiel) 6
Einführendes Gleichungssystem 7
Einführendes Gleichungssystem 8
Lösung der Probleme 9
Gauß'sches Eliminationsverfahren 10

Matrizen **11**

Matrix-Schreibweise 12
Matrix-Multiplikation (Wdh.) 13
Spezielle Matrix-Multiplikation (Wdh.) 14
Einheits-Matrix (Wdh.) 15
Invertierbarkeit (Wdh.) 16
Eigenwert (Wdh.) 17
Orthogonales Diagonalisieren (Wdh.) 18

Symmetrische Matrizen (Wdh.) . . . 19
Singulärwert-Zerlegung (1/2) . . . 20
Singulärwert-Zerlegung (2/2) . . . 21

LU-Zerlegung **22**
Eliminationsverfahren (1. Schritt) . . . 23
Eliminationsverfahren (2. Schritt) . . . 24
Eliminationsverfahren (Interpretation) . . . 25
LU-Zerlegung . . . 26
Auslöschungen . . . 27
Pivotisierung . . . 28
Erweiterte LU-Zerlegung . . . 29
Lösen von Gleichungssystemen . . . 30

Kondition **31**
Kondition . . . 32
Orthogonale Matrizen (Kondition) . . . 33
Allgemeine Matrizen (Kondition) . . . 34
Diagonalmatrizen (Kondition) . . . 35
Lineares Gleichungssystem (Kondition) . . . 36

Matrix-Normen **37**
Norm . . . 38
Matrix-Norm . . . 39
Spektral-Norm . . . 40
Frobenius-Norm . . . 41
Zusammenfassung . . . 42

Lineare Gleichungen 3 / 42

Runden und Kondition (Wdh.)

Wenn wir mit Fließkommazahlen arbeiten, werden Rundungen durchgeführt. Jede Rundung kann dargestellt werden als

$$\text{round}(x) = x \cdot (1 + \varepsilon_x) \qquad \|\varepsilon_x\| \leqslant \varepsilon_{\mathbb{M}}$$

An dieser Schreibweise sehen wir, wie sich relative Fehler verändern.

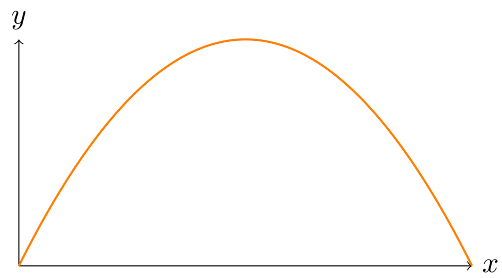
Die Kondition eines Problems $f\colon \mathbb{R}^n \to \mathbb{R}^m$ beschreibt wie relative Fehler verstärkt werden. Wenn $n = m = 1$ und $f\colon \mathbb{R} \to \mathbb{R}$ differenzierbar ist, gilt

$$\kappa = \left| \frac{x \cdot f'(x)}{f(x)} \right|.$$

Im Allgemeinen haben wir

$$\kappa = \frac{\|x\|}{\|f(x)\|} \cdot \limsup_{\delta \to 0} \frac{\|f(x+\delta) - f(x)\|}{\|\delta\|}$$

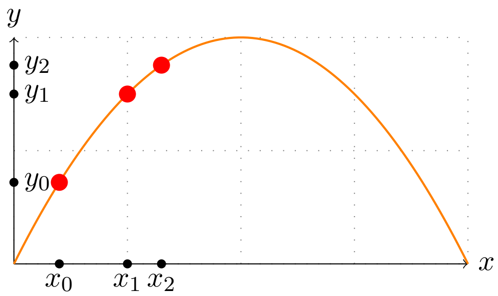
Parabolische Flugbahn (Beispiel)



Wenn wir Reibungen ignorieren, wird die Flugbahn durch eine Parabel $y = f(x) = a \cdot x^2 + b \cdot x + c$ beschrieben.

Wir wollen diese Parabel bestimmen, um **Flugweite**, **Abwurfhöhe** und **Abwurfwinkel** zu bestimmen.

Parabolische Flugbahn (Beispiel)



Um die Parabel zu bestimmen, liegen uns Messwerte zu Flughöhe y_i und Flugweite x_i zu verschiedenen Zeitpunkten zur Verfügung.

$(x_0, y_0) = (0.4, 0.72)$ $(x_1, y_1) = (1.0, 1.5)$ $(x_2, y_2) = (1.3, 1.755)$

Einführendes Gleichungssystem

Das führt zu einem Gleichungssystem in (a, b, c).

$$\begin{array}{ccccccc} x_0^2 \cdot a & + & x_0 \cdot b & + & c & = & y_0 \\ x_1^2 \cdot a & + & x_1 \cdot b & + & c & = & y_1 \\ x_2^2 \cdot a & + & x_2 \cdot b & + & c & = & y_2 \end{array}$$

Für unser konkretes Beispiel bedeutet das

$$\begin{array}{ccccccc} 0.16a & + & 0.4b & + & c & = & 0.720 \\ 1.00a & + & 1.0b & + & c & = & 1.500 \\ 1.69a & + & 1.3b & + & c & = & 1.755 \end{array}$$

Durch Umformen erhält man

$$\begin{array}{ccccccc} a & + & b & + & 1.000c & = & 1.5 \\ & & b & + & 3.500c & = & 2.0 \\ & & & & 0.675c & = & 0.0 \end{array}$$

Einführendes Gleichungssystem

Das Gleichungssystem

$$\begin{array}{ccccccc} a & + & b & + & 1.000c & = & 1.5 \\ & & b & + & 3.500c & = & 2.0 \\ & & & & 0.675c & = & 0.0 \end{array}$$

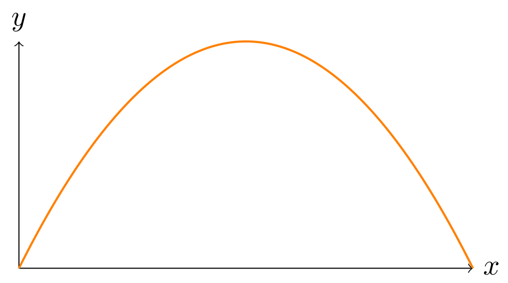
führt zu den Lösungen

$$c = 0 \qquad b = 2 \qquad a = -0.5$$

Die Parabel hat also die Form

$$f(x) = -\frac{1}{2}x^2 + 2x = \frac{1}{2}x(4-x)$$

Lösung der Probleme



Für die Parabel $\frac{1}{2}x(4-x)$ gilt nun:

Flugweite Löse $f(x) = 0$, d.h. `weite=4.0`
Abwurfhöhe Berechne $f(0)$, d.h. `hoehe=0.0`
Abwurfwinkel Berechne $\tan^{-1}(f'(0))$, d.h. `winkel=atan(2.0)` $\approx 63.4^\circ$

Gauß'sches Eliminationsverfahren

Eine allgemeine Methode zum Lösen von linearen Gleichungsssystemen geht auf Gauß zurück:

1. Eliminiere eine Variable nach der anderen.
2. Das führt zu einem *dreieckigen* Gleichungssystem.
3. Dieses System kann durch *Rückwärtseinsetzen* gelöst werden.

Der wesentliche Rechenaufwand besteht darin, das Gleichungssystem in die Dreiecksform zu bringen. Die **Elementaroperationen**, die dabei benutzt werden, sind

1. **Multiplikation einer Zeile** mit einer Zahl ungleich 0.
2. **Addieren von zwei Zeilen**.
3. **Vertauschen** von Zeilen oder Spalten.

Um das Gleichungssystem kompakter zu schreiben, benutzen wir Matrizen.

Matrix-Schreibweise

Anstelle eines Gleichungssystem in der Form von

$$\begin{array}{ccccccc} 0.16a & + & 0.4b & + & c & = & 0.720 \\ 1.00a & + & 1.0b & + & c & = & 1.500 \\ 1.69a & + & 1.3b & + & c & = & 1.755 \end{array}$$

schreiben wir

$$\underbrace{\begin{pmatrix} 0.16 & 0.4 & 1 \\ 1.00 & 1.0 & 1 \\ 1.69 & 1.3 & 1 \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 0.720 \\ 0.500 \\ 1.755 \end{pmatrix}}_{y}$$

Es geht also darum, $A \cdot x = y$ zu lösen.

Matrix-Multiplikation (Wdh.)

Gegeben zwei Matrizen $A \in \mathbb{R}^{m \times r}$ und $B \in \mathbb{R}^{r \times n}$, das Matrixprodukt $C := A \cdot B \in \mathbb{R}^{m \times n}$ ist wie folgt definiert:

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ \mathbf{a_{i1}} & \cdots & \mathbf{a_{ir}} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mr} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & \mathbf{b_{1j}} & \cdots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{r1} & \cdots & \mathbf{b_{rj}} & \cdots & b_{rn} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \mathbf{c_{ij}} & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix}$$

mit

$$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}$$

Spezielle Matrix-Multiplikation (Wdh.)

Schreibt man Vektoren $x \in \mathbb{R}^n$ als Spaltenvektoren, d.h. als $n \times 1$-Matrizen, so wird für eine $m \times n$-Matrix $A \in \mathbb{R}^n$ eine lineare Abbildung $x \mapsto A \cdot x$ definiert.

Für zwei Vektoren $x, y \in \mathbb{R}^n$ wird das **innere Produkt** bzw. **Skalarprodukt** $\langle x, y \rangle$ und das **äußere Produkt** $x \otimes y$ via

$$\langle x, y \rangle := x^\top y = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ z_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$$

$$x \otimes y := x y^\top = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}$$

definiert.

Einheits-Matrix (Wdh.)

Bezüglich der Multiplikation von $n \times n$-Matrizen ist die **Einheitsmatrix**

$$I := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}$$

ein **neutrales Element**, d.h.

$$I \cdot A = A \cdot I = A \qquad \forall A \in \mathbb{R}^{n\times n}$$

Wenn es zu $A \in \mathbb{R}^{n\times n}$ eine Matrix $B \in \mathbb{R}^{n\times n}$ gibt, so dass

$$B \cdot A = A \cdot B = I$$

gilt, so ist B eindeutig und heißt **inverse Matrix zu A**. Wir schreiben A^{-1} für die inverse Matrix von A.

Invertierbarkeit (Wdh.)

Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist nur dann invertierbar (d.h. A^{-1} existiert), wenn eine der folgenden (äquivalenten) Eigenschaften erfüllt ist

- $\det(A) \neq 0$.
- Die Spalten von A bilden eine Basis von $\mathbb{R}^n$.
- Die Spalten von A sind linear unabhängig.
- Der Rang von A ist n.
- Für jeden Vektor $b \in \mathbb{R}^n$ ist $A \cdot x = b$ eindeutig lösbar.

Für die **Determinante** $\det(\cdot)$ gilt

$$\begin{aligned}
\det(I) &= 1 \\
\det(A \cdot B) &= \det(A) \cdot \det(B) \\
\det\left(A^\top\right) &= \det(A) \\
\det\left(A^{-1}\right) &= \det(A)^{-1}
\end{aligned}$$

Eigenwert (Wdh.)

Ein Vektor $v \in \mathbb{R}^n$ heißt **Eigenvektor** zum **Eigenwert** $\lambda \in \mathbb{R}$ einer Matrix $A \in \mathbb{R}^{n \times n}$, wenn gilt:

$$A \cdot v = \lambda \cdot v \qquad\qquad (v \neq 0),$$

d.h. die Richtung v beschreibt die **Fixgerade** der Abbildung $x \mapsto A \cdot x$.

A heißt **diagonalisierbar**, wenn es n linear unabhängige Eigenvektoren $v_1, \ldots, v_n \in \mathbb{R}^n$ zu den n Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ gibt.

Die Eigenvektoren beschreiben eine invertierbare Matrix $V \in \mathbb{R}^{n \times n}$ und die Eigenwerte beschreiben eine Diagonalmatrix $\Lambda \in \mathbb{R}^{n \times n}$:

$$V := \begin{pmatrix} \vdots & & \vdots \\ v_1 & \cdots & v_n \\ \vdots & & \vdots \end{pmatrix} \qquad\qquad \Lambda := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Orthogonales Diagonalisieren (Wdh.)

Aus V und Λ läßt sich A wie folgt wieder herstellen

$$A = V \cdot \Lambda \cdot V^{-1},$$

d.h. A läßt sich durch die Diagonalmatrix Λ darstellen.

Wenn die sogenannte **Eigenbasis** $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis ist, d.h., wenn $\langle v_i, v_j \rangle = 0$ für $i \neq j$ und $\langle v_i, v_i \rangle = 1$, dann ist

$$\left(V \cdot V^\top\right)_{ij} = \sum_{k=1}^{n} v_{ik} v_{jk} = \langle v_i, v_j \rangle = I_{ij}$$

und wir haben für die **Orthogonalmatrix** V gerade

$$V \cdot V^\top = I \qquad \Rightarrow \qquad V^{-1} = V^\top \qquad \Rightarrow \qquad A = V \cdot \Lambda \cdot V^\top$$

Symmetrische Matrizen (Wdh.)

Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt **symmetrisch**, wenn $A = A^\top$ gilt. Sind v_1 und v_2 Eigenvektoren zu den unterschiedlichen Eigenwerten $\lambda_1 \neq \lambda_2$, so gilt

$$\begin{aligned} \lambda_1 \langle v_1, v_2\rangle &= \langle \lambda_1 v_1, v_2\rangle = v_1^\top A^\top v_2 = v_1^\top A v_2 \\ &= \langle v_1, A v_2\rangle = \lambda_2 \langle v_1, v_2\rangle \end{aligned}$$

Da $\lambda_1 \neq \lambda_2$ folgt daraus, dass

$$\langle v_1, v_2\rangle = 0$$

für Eigenvektoren zu unterschiedlichen Eigenwerten, d.h. **symmetrische Matrizen sind orthogonal diagonalisierbar**.

Gilt darüber hinaus, dass A **positiv semi-definit** ist ($\langle x, Ax\rangle \geqslant 0$), so sind alle Eigenwerte nicht-negativ.

Singulärwert-Zerlegung (1/2)

Für jede Matrix $A \in \mathbb{R}^{m\times n}$ sind die Matrizen $A^\top A \in \mathbb{R}^{n\times n}$ und $AA^\top \in \mathbb{R}^{m\times m}$ symmetrisch und positiv semi-definit. Daher gibt es orthogonale Matrizen $U \in \mathbb{R}^{m\times m}$ und $V \in \mathbb{R}^{n\times n}$, so dass

$$AA^\top = U\Sigma_0^2 U^\top \qquad A^\top A = V\Sigma_1^2 V^\top$$

mit nicht-negativen Diagonalmatrizen Σ_0 und Σ_1.
Für jeden Eigenvektor v_i von $A^\top A$ zum Eigenwert $\sigma_{1,i}^2$ gilt

$$AA^\top Av_i = A\sigma_{1,i}^2 v_i = \sigma_{1,i}^2 Av_i.$$

Für $Av_i \neq 0$ ist somit Av_i ein Eigenvektor von $AA^\top$ zum Eigenwert $\sigma_{1,i}^2 = \sigma_{0,i}^2$, d.h.:

$$Av_i \neq 0 \qquad \Rightarrow \qquad u_i = \frac{Av_i}{\sqrt{v_i^\top A^\top Av_i}} = \frac{Av_i}{\sigma_{\cdot,i}}$$

Singulärwert-Zerlegung (2/2)

Falls $Av_i \neq 0$, gilt also für $\sigma_i := \sigma_{\cdot,i}$

$$\sigma_i u_i = Av_i.$$

Diese Eigenschaft gilt auch für $Av_i = 0$ und $\sigma_i = 0$. Somit erhalten wir

$$U\Sigma = AV \qquad \Rightarrow \qquad A = U\Sigma V^\top$$

Diese Darstellung nennt man **Singulärwertzerlegung** und die nicht-negativen Diagonaleinträge von Σ heißen **Singulärwerte**.

Interessanterweise haben A und $A^\top$ die gleichen Singulärwerte:

$$A = U\Sigma V^\top \qquad A^\top = V\Sigma U^\top$$

Eliminationsverfahren (1. Schritt)

Das Eliminationsverfahren verwandelt das Problem im ersten Schritt von

$$\underbrace{\begin{pmatrix} 0.16 & 0.4 & 1 \\ 1.00 & 1.0 & 1 \\ 1.69 & 1.3 & 1 \end{pmatrix}}_{A_0} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} 0.720 \\ 1.500 \\ 1.755 \end{pmatrix}}_{y_0}$$

durch `Zeile_2 -= 6.25*Zeile_1; Zeile_3 -= 10.5625*Zeile_1;` nach

$$\underbrace{\begin{pmatrix} 0.16 & 0.400 & 1.0000 \\ 0.00 & -1.500 & -5.2500 \\ 0.00 & -2.925 & -9.5625 \end{pmatrix}}_{A_1} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} 0.720 \\ -3.000 \\ -5.8500 \end{pmatrix}}_{y_1}$$

Definieren wir $L_1 = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 6.25 & 0 & 0 \\ 10.5625 & 0 & 0 \end{smallmatrix}\right)$, so ist
$A_1 = (I - L_1)A_0$ und $y_1 = (I - L_1)y_0$.

Eliminationsverfahren (2. Schritt)

Das Eliminationsverfahren verwandelt das Problem im zweiten Schritt von

$$\underbrace{\begin{pmatrix} 0.16 & 0.400 & 1.0000 \\ 0.00 & -1.500 & -5.2500 \\ 0.00 & -2.925 & -9.5625 \end{pmatrix}}_{A_1} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} 0.720 \\ -3.000 \\ -5.8500 \end{pmatrix}}_{y_1}$$

durch `Zeile_3 -= 1.95*Zeile_2;` nach

$$\underbrace{\begin{pmatrix} 0.16 & 0.4 & 1.000 \\ 0.00 & -1.5 & -5.250 \\ 0.00 & 0.0 & 0.675 \end{pmatrix}}_{A_2} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} 0.72 \\ -3.00 \\ 0.00 \end{pmatrix}}_{y_2}$$

Definieren wir $L_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1.95 & 0 \end{pmatrix}$, so ist
$A_2 = (I - L_2)A_1$ und $y_2 = (I - L_2)y_1$.

Eliminationsverfahren (Interpretation)

Die Matrix $U := A_n$ ist eine Matrix in **oberer Dreiecksform** (engl. **upper triangular form**) und wir erhalten folgende Darstellung von $A = A_0$:

$$\tilde{L} \cdot A = U \qquad \tilde{L} := (I - L_{n-1}) \cdot \ldots \cdot (I - L_1).$$

Mit $L := \tilde{L}^{-1}$ erhalten wir die sogenannte LU-Zerlegung von A

$$A = \tilde{L}^{-1} \cdot U = L \cdot U.$$

Wegen $L_i \cdot L_j = 0$ erhalten wir

$$(I - L_i) \cdot (I + L_i) = I + L_i - L_i - L_i^2 = I \qquad \Rightarrow \qquad (I - L_i)^{-1} = (I + L_i)$$

und somit

$$\begin{aligned} L &= ((I - L_{n-1}) \ldots (I - L_1))^{-1} = (I + L_1) \ldots (I + L_{n-1}) \\ &= I + L_1 + \ldots + L_{n-1} \end{aligned}$$

LU-Zerlegung

Zu jeder quadratischen Matrix $A \in \mathbb{R}^{n \times n}$ gibt es eine Zerlegung in eine **obere Dreiecksmatrix** $U \in \mathbb{R}^{n \times n}$ und eine untere Dreiecksmatrix $L \in \mathbb{R}^{n \times n}$, so dass

$$A = L \cdot U$$

und die Diagonaleinträge von L alle 1 sind.

Wenn A invertierbar ist, sind die Einträge auf der Diagonalen von U alle von 0 verschieden.

Die LU-Zerlegung entspricht dem Gaußschen Eliminationsverfahren.

Auslöschungen

Die Berechnung der LU-Zerlegung berechnet vor allen Dingen Differenzen von Matrixzeilen. Dies kann zu Auslöschungen führen.

Außerdem kann das Verfahren dazu führen, dass wir eine Division durch 0 durchführen.

Um beide Probleme zu lösen, wollen wir Zeilen so austauschen, dass in der aktiven Spalte das betragsmäßigste größte Element steht (Pivotelement).

Dadurch wird das Verfahren numerisch stabiler.

Vertauschen von Zeilen kann auch durch eine Matrixmultiplikation erreicht werden. Hierzu werden **Permutationsmatrizen** benutzt. Das sind Matrizen die in jeder Zeile und jeder Spalte genau eine 1 besitzen und ansonsten aus Nullen bestehen.

Pivotisierung

Da die Spalten einer Permutationsmatrix $P \in \mathbb{R}^{n \times n}$ eine Orthnormalbasis beschreiben, gilt $P^{-1} = P^\top$. Während $P \cdot A$ Zeilen von A vertauscht, vertauscht $A \cdot P$ die Spalten von A.

Durch das Nutzen von Permutation haben wir nun folgende Darstellung:

$$U = (I - L_{n-1})P_{n-1} \dots (I - L_1)P_1$$

Interessanterweise kann man die Permutation nach rechts durchreichen:

$$P_{i+1}(I - L_i)P_i = P_{i+1}P_i - P_{i+1}L_iP_i$$

$\tilde{L}_i := P_{i+1}L_iP_{i+1}^\top$ besitzt wir L_i nur Einträge in der i-ten Spalte.

$$= P_{i+1}P_i - \tilde{L}_iP_{i+1}P_i = (I - \tilde{L}_i)P_{i+1}P_i$$

Erweiterte LU-Zerlegung

Aus den Beobachtungen haben wir folgende Darstellung

$$U = (I - L_{n-1}) \dots (I - L_1) P A,$$

wobei P die Permutationsmatrix $P = P_{n-1} \dots P_1$ ist.

Zu jeder quadratischen Matrix $A \in \mathbb{R}^{n \times n}$ gibt es eine **Permutationsmatrix** $P \in \mathbb{R}^{n \times n}$, eine **obere Dreiecksmatrix** $U \in \mathbb{R}^{n \times n}$ und eine **untere Dreiecksmatrix** $L \in \mathbb{R}^{n \times n}$, so dass

$$P \cdot A = L \cdot U,$$

wobei die Diagonaleinträge von L alle 1 sind.

Eine weitere Strategie benutzt **Totalpivotsuche**, d.h. es wird das betragsmäßig größte Element in der kompletten Untermatrix gesucht und dann Zeilen- und Spaltenvertauschungen durchgeführt.

Lösen von Gleichungssystemen

Angenommen, wir haben die LU-Zerlegung $L \cdot U = P \cdot A$ der Matrix A. Wie löst man das Gleichungssystem $A \cdot x = b$?

$$
\begin{aligned}
b =& Ax = P^\top P A x = P^\top L U x \\
Pb =& L(Ux)
\end{aligned}
$$

```
b_permut = P*b;
// Solve Ly=b_permut
y[1] = b_permut[1];
for (i=2; i<=n; i++) {
  y[i] = b_permut[i] - \sum_{j=1}^{i-1} L[i][j]*y[j];
}
// Solve Ux=y
x[n] = y[n]/U[n][n];
for (i=n-1; i>0; i--) {
  x[i] = (y[i] - \sum_{j=i+1}^{n} U[i][j]*x[j])/U[i][i];
}
```

Kondition

Die LU-Zerlegung ist einfach zu implementieren. Allerdings werden viele Differenzen berechnet, so dass das Verfahren numerisch instabil ist. Bevor wir andere Verfahren kennen lernen, sollten wir uns erst einmal die Kondition der Matrixmultiplikation ansehen:

Sei die Matrix $A \in \mathbb{R}^{n \times n}$ und der Vektor $x \in \mathbb{R}^n$ gegeben. Die absolute Kondition von $x \mapsto A \cdot x$ ist gerade

$$\kappa_{\mathsf{abs}} = \limsup_{d \to 0} \frac{\|A(x+d) - A(x)\|}{\|d\|} = \max_{d \in \mathbb{R}^n} \frac{\|Ad\|}{\|d\|}$$

Die Kondition von $x \mapsto A \cdot x$ hängt von x ab und wir haben

$$\kappa = \frac{\|x\|}{\|Ax\|} \kappa_{\mathsf{abs}} \leqslant \frac{\max_{y \in \mathbb{R}^n} \frac{\|Ay\|}{\|y\|}}{\min_{y \in \mathbb{R}^n} \frac{\|Ay\|}{\|y\|}} =: \kappa(A)$$

Orthogonale Matrizen (Kondition)

Wenn $A \in \mathbb{R}^{n\times n}$ eine **orthogonale Matrix** ist ($A^\top A = I$), dann gilt

$$\|Ay\|^2 = \langle Ay, Ay\rangle = y^\top A^\top Ay = \|y\|^2$$

Somit gilt für jedes $x \in \mathbb{R}^n$ immer $\frac{\|Ax\|}{\|x\|} = 1$ und wir haben

$$\kappa(A) = 1 \qquad\qquad \forall A, A^\top A = I$$

Ist $A \in \mathbb{R}^{n\times n}$ eine **orthogonal diagonalisierbar Matrix** ($A = U\Lambda U^\top$), so gilt

$$\begin{aligned}\|Ay\|^2 &= y^\top U\Lambda U^\top U\Lambda U^\top y = \left\|\Lambda U^\top y\right\|^2\\ \kappa(A) &= \frac{\max_{y\in\mathbb{R}^n} \frac{\|Ay\|}{\|y\|}}{\min_{y\in\mathbb{R}^n} \frac{\|Ay\|}{\|y\|}} = \frac{\max_{y\in\mathbb{R}^n} \frac{\left\|\Lambda U^\top y\right\|}{\|U^\top y\|}}{\min_{y\in\mathbb{R}^n} \frac{\|\Lambda U^\top y\|}{\|U^\top y\|}} = \kappa(\Lambda)\end{aligned}$$

Allgemeine Matrizen (Kondition)

Betrachten wir zu $A \in \mathbb{R}^{n \times n}$ die Singulärwertzerlegung $A = U\Sigma V^\top$. Dann gilt

$$\|Ay\|^2 = y^\top V\Sigma U^\top U\Sigma V^\top y = y^\top V\Sigma\Sigma V^\top y = \left\|\Sigma V^\top y\right\|^2$$

Damit gilt

$$\kappa(A) = \frac{\max_{y\in\mathbb{R}^n} \frac{\|Ay\|}{\|y\|}}{\min_{y\in\mathbb{R}^n} \frac{\|Ay\|}{\|y\|}} = \frac{\max_{y\in\mathbb{R}^n} \frac{\left\|\Sigma V^\top y\right\|}{\|V^\top y\|}}{\min_{y\in\mathbb{R}^n} \frac{\|\Sigma V^\top y\|}{\|V^\top y\|}} = \kappa(\Sigma)$$

Insgesamt läßt sich die Berechnung der Kondition einer Matrix A also auf die Kondition einer Diagonalmatrix zurückführen. Bei den Diagonaleinträgen handelt es sich entweder um die Eigenwerte einer orthogonal diagonaliserbaren Matrix oder um die Singulärwerte.

Diagonalmatrizen (Kondition)

Sei $\Lambda \in \mathbb{R}^{n\times n}$ eine Diagonalmatrix mit den Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$. Weiter sei $x \in \mathbb{R}^n$ ein Vektor der Länge $\alpha > 0$, d.h. $\sum_{i=1}^n x_i^2 = \alpha^2$.

Bezeichnen wir mit $\lambda_{\max}$ den betragsmäßig größten und mit $\lambda_{\min}$ den betragsmäßig kleinsten Diagonaleintrag, so ist

$$\|\Lambda x\|^2 = \sum_{i=1}^n \lambda_i^2 x_i^2 \begin{cases} \geqslant (\lambda_{\min} \cdot \alpha)^2 \\ \leqslant (\lambda_{\max} \cdot \alpha)^2 \end{cases}$$

und wir erhalten

$$\kappa(\Lambda) = \frac{\lambda_{\max}}{\lambda_{\min}},$$

d.h. die Konditionszahl der Matrixmultiplikation ist durch die Singulärwerte bestimmt und liegt zwischen 1 und ∞ (für singuläre d.h. nicht-invertierbare Matrizen).

Lineares Gleichungssystem (Kondition)

Beim Lösen eines linearen Gleichungssystems $A \cdot x = b$ ist man nicht an der Multiplikation mit A sondern mit A^{-1} interessiert.

Nun haben wir

$$\kappa(A^{-1}) = \frac{\max_{y\in\mathbb{R}^n} \frac{\|A^{-1}y\|}{\|y\|}}{\min_{y\in\mathbb{R}^n} \frac{\|A^{-1}y\|}{\|y\|}} = \frac{\max_{y\in\mathbb{R}^n} \frac{\|y\|}{\|Ay\|}}{\min_{y\in\mathbb{R}^n} \frac{\|y\|}{\|Ay\|}} = \frac{\max_{y\in\mathbb{R}^n} \frac{\|Ay\|}{\|y\|}}{\min_{y\in\mathbb{R}^n} \frac{\|Ay\|}{\|y\|}} = \kappa(A)$$

Somit ist das Lösen eines linearen Gleichungssystems numerisch genauso schwierig wie die Matrix-Vektor-Multiplikation.

Viele Verfahren versuchen die Matrix A als Produkt von Matrizen darzustellen, die eine geringe Konditionszahl haben und darüber hinaus einfach zu invertieren sind. Beispiele hierfür sind *orthogonale Matrizen* und *Diagonalmatrizen*.

Norm

Um Längen von Vektoren zu berechnen nutzen wir üblicherweise die **Euklidische Norm**

$$\begin{aligned} \|\cdot\| : \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \sqrt{\langle x, x \rangle} \end{aligned}$$

Allerdings kann man auch andere Normen definieren.

Definition 1. Sei $\mathbb{V}$ ein $\mathbb{R}$-Vektorraum. Dann heißt eine Funktion $\|\cdot\| : \mathbb{V} \to \mathbb{R}_0^+$ eine **Norm**, wenn Folgendes gilt:

$$\|x\| = 0 \Leftrightarrow x = 0 \qquad \text{(Positiv Definit)}$$

$$\|\alpha \cdot x\| = |\alpha| \cdot \|x\| \qquad \text{(Absolut Homogen)}$$

$$\|x + y\| \leqslant \|x\| + \|y\| \qquad \text{(Dreiecksungleichung)}$$

Matrix-Norm

Wenn man sich die Menge aller Matrizen $A \in \mathbb{R}^{n\times n}$ ansieht, so bildet diese Menge einen $\mathbb{R}$-Vektorraum. So kann man z.B. die Summe von Matrizen bilden und Matrizen mit einer reelen Zahl multiplizieren.

Wählt man eine beliebige Norm $\|\cdot\|_V$ auf $\mathbb{R}^n$, so induziert diese die **Matrixnorm** $\|\cdot\|_M$ auf $\mathbb{R}^{n\times n}$ wie folgt:

$$\|A\|_M := \max_{x\neq 0} \frac{\|Ax\|_V}{\|x\|_V}$$

Für Matrixnormen gelten folgende Ungleichungen

$$\begin{aligned}
\|Ax\| &\leqslant \max_{y\neq 0} \frac{\|Ay\|}{\|y\|} \|x\| = \|A\| \, \|x\| \\
\|AB\| &= \max_{x\neq 0} \frac{\|ABx\|}{\|x\|} = \max_{x\neq 0} \frac{\|ABx\| \, \|Bx\|}{\|Bx\| \, \|x\|} \leqslant \max_{y\neq 0} \frac{\|Ay\|}{\|y\|} \max_{x\neq 0} \frac{\|Bx\|}{\|x\|} \leqslant \|A\| \, \|B\|
\end{aligned}$$

Spektral-Norm

Nutzen wir die von der Euklidischen Norm induzierten Matrixnorm, so erhalten wir die sogenannte **Spektralnorm**

$$\|A\|_2 := \max_{x \neq 0} \frac{\sqrt{\langle Ax, Ax \rangle}}{\sqrt{\langle x, x \rangle}}$$

Wir haben bereits gesehen, dass für orthogonal diagonalisierbare Matrizen A gerade Folgendes gilt

$$\|A\|_2 = \lambda_{\max}(A)$$

Insbesondere haben wir für die Kondition

$$\kappa(A) = \|A\|_2 \left\|A^{-1}\right\|_2$$

Frobenius-Norm

Statt der Spektralnorm kann man eine Matrix einfach als Vektor auffassen und davon die Euklidische Norm berechnen. Dies liefert uns die **Frobeniusnorm**

$$\|A\|_F := \sqrt{\sum_{i,j=1}^{n} (a_{ij})^2}$$

Eine kompaktere Schreibweise für diese Norm ist

$$\|A\|_F = \sqrt{\operatorname{spur}(A^\top A)}$$

Ist A orthogonal diagonalisierbar ($A = U\Lambda U^\top$) so gilt

$$\|A\|_F = \sqrt{\operatorname{spur}(U\Lambda^2 U^\top)} = \sqrt{\operatorname{spur}(\Lambda^2)} = \sqrt{\sum_{i=1}^{n} \lambda_i^2}$$

Zusammenfassung

- Viele Probleme führen zu **linearen Gleichungssystemen (LGS)**.
- Ein LGS kann kompakter dargestellt werden durch **Matrixoperationen**.
- Die Lösung eines LGS kann durch das **Gaußsche Eliminierungsverfahren** erhalten werden.
- In Matrixnotation entspricht dies der **LU-Zerlegung** in eine untere und eine obere Dreiecksmatrix.
- Die Verstärkung des Eingabefehlers beim Lösen eines LGS kann durch die **Konditionszahl** geschätzt werden.
- Die Kondition läßt sich in der **Spektralnorm** ausdrücken.
- **Orthonormalmatrizen** haben die Kondition 1.