

Vorlesungs-Skript

Mathematik für
Wirtschaftsinformatiker

Rüdiger Seydel



Universität zu Köln
Mathematisches Institut

Version 2009

© R. Seydel, Köln 2009

www.mi.uni-koeln.de/~seydel

Vorwort

Dies ist das Skript zu meiner Vorlesung Mathematik für Wirtschaftsinformatiker, gelesen im Wintersemester 2008/2009 und im Sommersemester 2009. Im ersten Semester ist die Vorlesung vierstündig, im zweiten Semester zweistündig. Dazu gehören jeweils Übungen, welche ein wichtiger Teil der Veranstaltung sind.

Das Skript gibt nur den wesentlichen Inhalt der Vorlesung wieder. Zusätzliche Erklärungen und Ergänzungen werden während der Vorlesung gegeben; das Skript ersetzt nicht den Besuch der Vorlesung!

Für Korrekturlesen bin ich Herrn Markus Lücking dankbar.

Köln, Juli 2009

Inhalt

Prolog 1

Kapitel 1: Elemente Praktischer Mathematik

- 1.1 Messen in der Ebene 3
- 1.2 Gleichungen, Ungleichungen 5
- 1.3 Mengen in der Ebene 6
- 1.4 Zahlen 7
- 1.5 Fehler 10

Kapitel 2: Lineare Algebra

- 2.1 Vektoren, Matrizen, Gleichungssysteme 17
- 2.2 Der Algorithmus von Gauß 22
- 2.3 Arbeiten mit Matrizen 23
- 2.4 Vektorräume 26
- 2.5 Dimension und Basis von Vektorräumen 28
- 2.6 Determinanten 31
- 2.7 Eigenwerte 33
- 2.8 Normen 35
- 2.9 Householder Matrizen 37

Kapitel 3: Analysis

- 3.1 Vollständige Induktion 41
- 3.2 Komplexe Zahlen 42
- 3.3 Polynome und Rationale Funktionen 45
- 3.4 Der Euklidische Algorithmus; Kettenoperationen 49
- 3.5 Funktionen 50
- 3.6 Spezielle Funktionen 55
- 3.7 Grenzwert 59
- 3.8 Differentialrechnung 63
- 3.9 Integrale 69
- 3.10 Logarithmus und Exponentialfunktion 73
- 3.11 Reihen und Potenzreihen 74
- 3.12 Funktionen von mehreren Veränderlichen 79

Kapitel 4: Approximation mit Kurven

- 4.1 Approximation und Interpolation 81
- 4.2 Interpolation mit Polynomen 82
- 4.3 Interpolation mit Splines 87
- 4.4 Bézier-Kurven 91
- 4.5 Integration mit Trapezsummen; Extrapolation 93
- 4.6 Diskrete Fourier-Transformation 96
- 4.7 Fast Fourier-Transformation 99
- 4.8 Ausgleichsprobleme, data fit 103

Kapitel 5: Nichtlineare Gleichungssysteme und Iterationen

- 5.1 Lösen einer skalaren Gleichung 107
- 5.2 Zur Konvergenz 108
- 5.3 Das allgemeine Newtonverfahren 110
- 5.4 Approximation der Jacobi-Matrix 111

Kapitel 6: Optimierung

- 6.1 Optimierungsprobleme 113
- 6.2 Methoden der Analysis 114
- 6.3 Lineare Optimierung 116
- 6.4 Konvexe Optimierung 120

Literatur

Als weiterführende Literatur sind unter anderem fast alle Titel mit *Höherer Mathematik* oder auch *Ingenieurmathematik* geeignet. Es ist nicht möglich, der Vielzahl hervorragender Werke gerecht zu werden. Deswegen hier nur exemplarisch drei Werke:

R. Ansorge, H.J. Oberle: Mathematik für Ingenieure I. Wiley-VCH

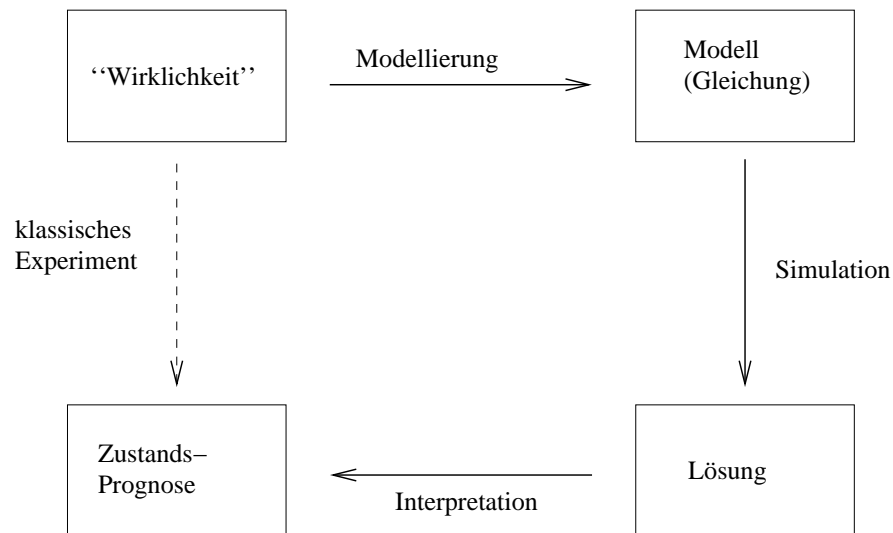
K. Meyberg, P. Vachenauer: Höhere Mathematik. Band 1. Springer-Verlag, Berlin

R. Sauer: Ingenieurmathematik. Springer-Verlag

Eine Formelsammlung ist nützlich.

Prolog

Charakterisierung Wissenschaftlicher Arbeit: Modellierung & Simulation



- ★ Modellierung = Beschreibung von Vorgängen in Natur, Wirtschaft und Technik durch Gleichungen
- ★ Simulation = Anwendung mathematisch/numerischer Methoden zur Lösung der Gleichungen
- “mathware”

Der Doppelschritt “Modellierung & Simulation” ist zwingend!

Beispiele “Experiment unmöglich” da

- zu gefährlich (Umweltrisiken, Wirtschaftsentwicklung)
- zu groß, experimentelle Restriktionen (Ozeanströmungen)
- zu lange Zeiträume (Bevölkerungsentwicklung, Klima)
- zu teuer (Windkanal wird durch Computer-Simulationen ersetzt.)
- existiert nicht (Saurier etc. im Kino)

Mathematik: Sprache unserer wissenschaftlich-technischen Welt, Schlüsseldisziplin der Hochtechnologie!

Kapitel 1 Elemente Praktischer Mathematik

1.1 Messen in der Ebene

die Ebene: der Ort aller mathematischen Erklärungen.

einfache mathematische Objekte:

Punkt, Gerade(nstück), Kurve, Kreis, "Winkel" (Haken), Flächenstück.

A. Messung / Zahlen

Charakterisierung obiger Objekte:

Punkt: Lage (zu einem Referenzpunkt)

Kurve, Geradenstück: Länge, Neigung

Kreis: Durchmesser, Umfang

Schnitt von Geraden: Winkel

Flächenstück: Inhalt

Zur Quantifizierung sind Zahlen erforderlich.

historisch: z.B. 2 Fuß, 3 Fingerbreit: Körpersystem, mit Anzahl der Wiederholungen.

→ "Natürliche Zahlen" 1,2,3,... Bezeichnung: \mathbb{N}

Verhältnisse: z.B. beim Kreis: $\pi := \frac{\text{Umfang}}{\text{Durchmesser}}$

Hinweis: $:=$ meint: die "linke Seite" wird durch die "rechte Seite" definiert. Obiges π definiert die **Kreiszahl**.

Frage: Ist $\pi \in \mathbb{N}$? (Antwort: nein! \mathbb{N} reicht nicht. Hierzu und zu Zahlen vgl. späteren Abschnitt 1.4)

B. Größe eines Winkels

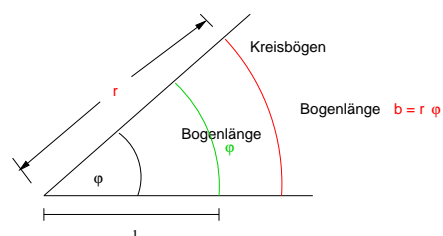
1.) **Messung in Grad:** Teilung des Kreises in 360 Teile (willkürlich).

Ein Teil wird als Grad bezeichnet, 1° ,

Grad/Min/Sek: $1^\circ = 60'$, $1' = 60''$

z.B. $40^\circ 36' 12''$

2.) **im Bogenmaß** (natürliches Maß)



Winkel im Bogenmaß = $\frac{\text{am Radius } r \text{ ausgeschnittene Bogenlänge } b}{r}$

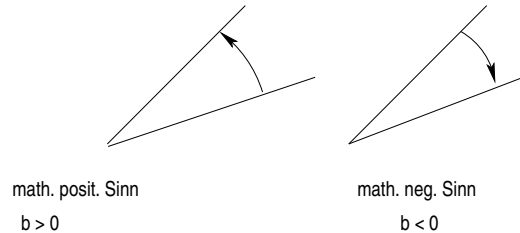
dimensionslos, unabhängig vom Maßstab.

Umfang Vollkreis mit Radius r ist $2r\pi \Rightarrow$ ("daraus folgt")

im Gradmaß	360°	180°	90°	60°	45°	30°	1°
im Bogenmaß	2π	π	$\frac{\pi}{2}$	$\frac{\pi}{3}$	$\frac{\pi}{4}$	$\frac{\pi}{6}$	$\frac{\pi}{180} \approx 0.01745\dots$

(Dezimal-Punkt !)

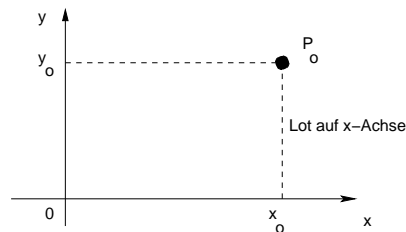
Orientierung:



C. Koordinaten

1.) Kartesische Koordinaten eines Punktes P_0

Vorgabe eines Punktes O und einer (x -)Achse; positive Drehung um 90° um O erzeugt y -Achse (Schnittpunkt O). Lote von P_0 auf die Achsen \rightarrow Fußpunkte bestimmen x - und y -Koordinate x_0, y_0

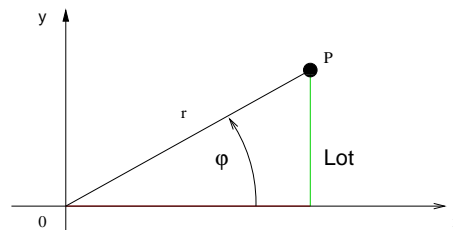


Man schreibt $P_0 = (x_0, y_0)$; $O = (0, 0)$ heißt "Nullpunkt".

Eindeutige Zuordnung zwischen der Ebene und allen Zahlenpaaren.

2.) Polarkoordinaten

Jeder Punkt $\neq O$ kann eindeutig durch den Abstand r zum Nullpunkt und den Winkel zwischen der positiven x -Achse und dieser Geraden dargestellt werden.



Zusammenhang zu kartesischen Koordinaten

$$x = r \cos \varphi \quad , \quad y = r \sin \varphi \quad ,$$

$$r = \sqrt{x^2 + y^2} \quad , \quad \varphi = \begin{cases} \arccos \frac{x}{r} & \text{falls } y \geq 0 \\ 2\pi - \arccos \frac{x}{r} & \text{falls } y < 0 \end{cases}$$

1.2 Gleichungen, Ungleichungen

“=” seit dem späten 16. Jahrhundert

Arten von Gleichungen

1. Identität: z.B. $3 - 4x + 5x = 3 + x$ oder: $1 + 2 + \dots + n = \frac{1}{2}n(n+1)$
2. Abbildungsgleichung: z.B. $y = f(x) := (x+3)^2 - 4$
Jedem x in einem gewissen Bereich wird ein y zugeordnet: “Funktion”.
3. Bestimmungsgleichung: z.B. $x^2 - 4 + 2x = 0$ oder $|x| - x^2 = 0$.
Hier heißt x “Unbekannte”.

Def. Absolutbetrag:

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases}$$

(Fallunterscheidungen notwendig, interessant oft dann, wenn komplizierter Ausdruck $|A(x)|$)

4. Differentialgleichungen:

z.B. $\frac{dy}{dx} = -2y$ (oder $y' = -2y$)

ist Gleichung für eine **Funktion** $y(x)$. In der Gleichung tritt y mit der ersten oder höheren Ableitungen auf.

(Lösung: $y(x) = ce^{-2x}$ für beliebige Konstanten c)

(Probe: ...)

Ungleichungen

Größenvergleich $x < y$ (x ist kleiner y und nicht gleich y)

$x \leq y$ (x ist “kleiner gleich” y) bzw. $x > y$, $x \geq y$

Die Ungleichung $x < y$ gibt eine Abschätzung der Zahl x nach oben durch y an; y ist Schranke (obere) zu x .

Beispiel (für Ungleichungen, Abschätzungen, untere und obere Schranken)

Zeige: π ist keine natürliche Zahl!

Beweis:

umschreibendes Viereck mit Umfang 8 (Figur ...)

inneres 8-Eck mit Umfang d

“Einheitskreis” mit Radius 1 und Umfang c .

Offenbar gilt: $c < 8$ (8 ist obere Schranke zu c)

und: $d < c$ (d ist untere Schranke); leicht zu zeigen: $d > 6$.

Beide Ungleichungen zusammengefaßt:

$$6 < d < c < 8, \quad \text{es folgt:}$$

$$6 < c < 8$$

$$\pi = \frac{c}{2r} = \frac{c}{2} \Rightarrow c = 2\pi \Rightarrow$$

$$6 < 2\pi < 8 \Rightarrow 3 < \pi < 4$$

Also: π ist keine natürliche Zahl.

Def. Intervall: Durch das Ungleichungspaar $a \leq x \leq b$ ist ein “Intervall” definiert als die Menge aller x für welche sowohl $a \leq x$ **und** $x \leq b$ gilt (gleichzeitig!). Das Intervall ist nicht-leer, wenn $a \leq b$.

Mengen-Schreibweise:

$$[a, b] := \{x | a \leq x \leq b\}$$

Bei Einschluß der **Randpunkte** spricht man von **abgeschlossenem** Intervall. Sonst: **offenes** Intervall: $\{x | a < x < b\}$; Bezeichnung auch $]a, b[$ oder (a, b) .

Varianten: "halboffen", $a \leq x < b$, $a < x \leq b$, falls nur ein Endpunkt zum Intervall gehört.

Rechenregeln zu Betrag und Ungleichungen

$$x \leq y, a \leq b \Rightarrow x + a \leq y + b$$

$$x \leq y, 0 \leq a \Rightarrow ax \leq ay$$

$$x \leq y, a < 0 \Rightarrow ax \geq ay \quad !$$

$$x \leq y \Rightarrow -y \leq -x$$

$$0 < x \leq y \Leftrightarrow 0 < \frac{1}{y} \leq \frac{1}{x}$$

$$-|a| \leq a \leq |a|,$$

$$|-a| = |a|,$$

$$|ab| = |a| |b|, \left| \frac{a}{b} \right| = \frac{|a|}{|b|} \quad (\text{falls } b \neq 0)$$

$$|a| \leq |b| \Leftrightarrow -b \leq a \leq b \quad !$$

weiter gilt die "**Dreiecksungleichung**"

$$|a + b| \leq |a| + |b|$$

Beweis:

$$\begin{array}{ccc} -|a| & \leq a & \leq |a| \\ -|b| & \leq b & \leq |b| \end{array}$$

Addition:

$$-(|a| + |b|) \leq \underbrace{a + b}_{(Abk.)=:A} \leq \underbrace{|a| + |b|}_{=:B}$$

Also $-B \leq A \leq B$;

$$\Leftrightarrow |A| \leq B \Leftrightarrow |a + b| \leq |a| + |b|$$

1.3 Mengen in der Ebene

geometrische Bedeutung von Gleichungen und Ungleichungen, wenn sowohl x als auch y variabel sind.

Beispiele:

1.) $a \leq x \leq b$

$\{(x, y) | a \leq x \leq b\}$, d.h. y bleibt frei

(Hinweis: In 3-dim. (x, y, z) -Raum wäre durch $a \leq x \leq b$ eine **Scheibe** definiert.

Also: Angabe (x, y) des einbettenden Raumes wichtig!

2.) $\{(x, y) | a \leq x \leq b, y = -2\}$

- 3.) $x \geq -4$ **und** $0 \leq y \leq 6$ d.h. 3 Ungleichungen gleichzeitig!
 (Hinweis zur Vorgehensweise: zuerst die Gleichungen aufmalen (ergibt potentielle Ränder), dann die richtige Seite feststellen (z.B. mit Testpunkten).
 Halbstreifen als Schnitt von drei Halbebenen.
- 4.) $0 = x - y + 1$ ist eine **Gerade**, $y = x + 1$
- 5.) $y = -x^2 + 1$ **und** $-1 \leq x \leq 3$
- 6.) $y^2 \leq x < 4$ sind 3 Ungleichungen $x < 4$, $-\sqrt{x} \leq y$, $y \leq +\sqrt{x}$
- 7.) $x^2 + y^2 \leq 1$ ($\Leftrightarrow r^2 \leq 1$)
- 8.) $(x - 2)^2 + (y + 1)^2 = 4$, $y \geq -x + 1$
 $x' := x - 2$, $y' := y + 1$
 Allgemein: Kreisgleichung
 $(x - a)^2 + (y - b)^2 = r^2$ ist Kreis mit Mittelpunkt $(x, y) = (a, b)$
 ((Vorsicht: Vorzeichen!))

Schnitt von 2 Geraden

allg. Gerade in der Ebene:

$$ax + by + c = 0$$

Also: **eine** Gleichung in 2 Unbekannten x und y . x und y treten "linear" auf; lineare Gleichung in x und y ;

zwei Geraden in der Ebene:

$$\begin{array}{rcl} a_1x + b_1y + c_1 & = & 0 \quad 1. \text{ Gerade} \\ a_2x + b_2y + c_2 & = & 0 \quad 2. \text{ Gerade} \end{array}$$

2 Möglichkeiten:

- 1.) Die Geraden schneiden sich in einem Punkt
- 2.) Die Geraden sind "parallel" (sie schneiden sich nicht oder sind identisch)

Berechnung des Schnittpunktes

Addiere geeignete Vielfache der Gleichungen.

"geeignet": so dass eine Gleichung mit nur einer Unbekannten entsteht.

allgemeines Verfahren: Gaußscher Eliminations-Algorithmus

1.4 Zahlen

A.) Natürliche Zahlen $1, 2, 3, 4, \dots$

(\mathbb{N} ist "abgeschlossen" bzgl. Addition und Multiplikation.)

Abschließung bzgl. der Subtraktion führt zu:

B.) Ganze Zahlen

$$\begin{aligned} \mathbb{Z} &:= \mathbb{N} \cup 0 \cup \text{negative natürliche Zahlen} \\ &= \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, \dots\} \end{aligned}$$

C.) Rationale Zahlen

$$\mathbb{Q} := \{ \text{Brüche} \} = \left\{ \frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{N} \right\}$$

Division (außer durch 0) in \mathbb{Q} unbeschränkt ausführbar.

Satz: \mathbb{Q} ist abzählbar! Denn: alle $q \in \mathbb{Q}$ sind enthalten im Schema...

D.) Irrationale Zahlen = Zahlen, die nicht rational sind.

Satz: Es gibt irrationale Zahlen!

“**Indirekter Beweis**”: Annahme des Gegenteils;

hieraus Herleitung eines Widerspruchs zu den Annahmen.

Prinzip ($A \Rightarrow B$) \Leftrightarrow ($\neg B \Rightarrow \neg A$)

Hier: Annahme $s = \frac{m}{n}$, ohne Einschränkung m und n teilerfremd (sonst kürzen)

(vgl. Übung)

E.) Reelle Zahlen

Exponenten-Schreibweise ($n \in \mathbb{N}$): $a^n = \underbrace{a \cdot a \cdot \dots \cdot a}_{n \text{ mal}}$

$$a^{-n} = \frac{1}{a^n} \quad , \quad a^0 = 1$$

Motiv./Bsp.: $3.907 = 3 + \frac{9}{10} + \frac{7}{1000} = 3 \cdot 10^0 + 9 \cdot 10^{-1} + 7 \cdot 10^{-3}$ ist “endlicher”
Dezimalbruch, d.h. bricht nach endlich vielen Stellen ab.

reelle Zahlen:

$$\begin{aligned} \mathbb{R} &:= \{ \text{alle endlichen und unendlichen Dezimalbrüche} \} \\ &= \{ p_1 10^{k-1} + p_2 10^{k-2} + \dots + p_k 10^0 + q_1 10^{-1} + q_2 10^{-2} + \dots \mid p_i, q_i \in \{0, 1, 2, \dots, 9\} \} \\ &= \{ p_1 p_2 \dots p_k \cdot q_1 q_2 q_3 \dots \} \text{ “Positionssystem”} \end{aligned}$$

Beschränkung auf $p, q \leq 9$ wegen der Eindeutigkeit der Darstellung.

geometrisch auf der Zahlengerade:

Intervallschachtelung: bricht man eine Zahl nach der m -ten Stelle hinter Dezimalpunkt ab, so liegt die Zahl im Intervall der Breite 10^{-m} . Hinzufügen der nächsten Ziffer ($(m+1)$ -te Ziffer) \rightarrow Breite des Intervalles 10^{-m-1} u.s.w.

Intervall-Längen werden beliebig klein. \rightarrow konvergiert gegen Punkt.

Damit äquivalente Definition von \mathbb{R} : alle Punkte auf der Zahlengerade $\rightarrow \mathbb{R}$.

Bei den Dezimalzahlen ist 10 die “Basis”. Die reellen Zahlen können auch mit jeder anderen Basis aufgebaut werden.

F.) Komplexe Zahlen

Der Zahlenbereich \mathbb{R} kann erweitert werden zu den Komplexen Zahlen, in denen auch $i := \sqrt{-1}$ etc. möglich ist. (\rightarrow später in Abschnitt 3.2)

G.) Maschinenzahlen

Def.: Gleitpunktdarstellung (floating-point representation):

$$x = \sigma \cdot m \cdot B^E$$

σ : Vorzeichen

m : Mantisse

B : Basis

E : Exponent

Def.: Normalisierte Gleitpunktzahl ($\neq 0$): Gleitpunktzahl mit Eigenschaft $B^{-1} \leq m < 1$.

Darstellung der Mantisse:

$$m = d_1 B^{-1} + d_2 B^{-2} + d_3 B^{-3} + \dots$$

mit **Ziffern (digits)** $d_i \in \{0, 1, \dots, B-1\}$ und $d_1 \neq 0$ für normalisierte Zahl.

Beispiel: 18.5 als Dezimalzahl [= $(18.5)_{10}$], umrechnen in Dualzahl:

$$\begin{aligned} 18.5 &= 16 + 2 + \frac{1}{2} \\ &= \mathbf{1} \cdot 2^4 + \mathbf{0} \cdot 2^3 + \mathbf{0} \cdot 2^2 + \mathbf{1} \cdot 2^1 + \mathbf{0} \cdot 2^0 + \mathbf{1} \cdot 2^{-1} \\ &\quad \uparrow \quad \quad \uparrow \quad \quad \quad \quad \quad \quad \uparrow \\ &\quad \text{Ziffern} \quad \quad \quad \quad \quad \quad \text{Binärpunkt} \\ &= \mathbf{10010.1} \end{aligned}$$

Schreibweise mit Basis gekennzeichnet:

$$(18.5)_{10} = (10010.1)_2$$

normalisiert:

$$0.185 \cdot 10^2 = 0.100101 \cdot 2^5$$

Bemerkung: Bei diesem Beispiel wurden sowohl bei $B = 10$ als auch bei $B = 2$ nur endlich viele Ziffern benötigt. Im Allgemeinen gibt es **keine** derartige Entsprechung!

Im Rechner stehen nur endlich viele Stellen zur Verfügung!

Definition: Die Zahlen, die in einem Rechner dargestellt werden können, heißen **Maschinenzahlen**.

“Genauigkeit” misst letztlich den **Abstand von Zahlen**.

Frage: Was ist der kleinste Abstand zwischen zwei Maschinenzahlen?

Gedankenexperiment: mit 53 bit-Mantisse

$$\begin{aligned} 0.1111 \dots 11 &= 1/2 + \dots + 2^{-53} = 1 - 2^{-53} = 1 - 2^{-k} \\ 1.0000 \dots 00 &= 1 \end{aligned}$$

(Die Binär-Ziffern heißen **bit** = **binary digit**.)

Definition: Es sei k die Anzahl der bit der Mantisse. Die Zahl 2^{-k} heißt **relative Maschinengenauigkeit** des Binär-Rechners.

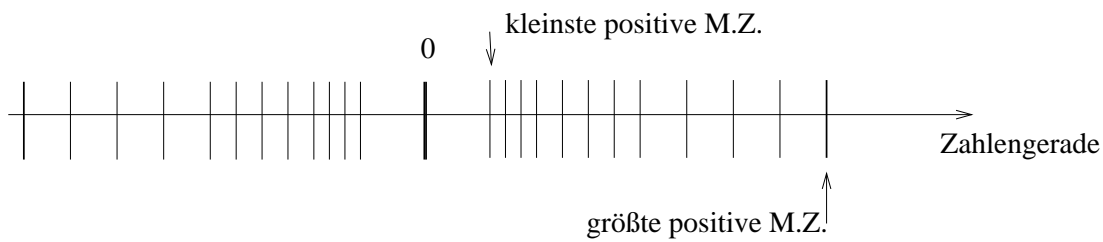
Bezeichnung: ϵ_{mach} (oder *eps*, oder ϵ_0) **Bedeutung:** ϵ_{mach} ist der kleinste relative Abstand zwischen zwei Maschinenzahlen, und damit **unterste Schranke für alle mit einem Rechner erreichbaren (rel.) Genauigkeiten!**

Beispiel: von oben:

$$\begin{aligned} \epsilon_{\text{mach}} &= 2^{-53} = \exp(-53 \log 2) \\ &= 1.11 \dots * 10^{-16} \end{aligned}$$

d.h., bei $k = 53$ ist eine relative Genauigkeit von maximal ca. 16 dezimalen Stellen möglich.

Die Maschinenzahlen sind ungleichmäßig verteilt. Schematisch:



Jede Zahl (Eingabe oder Resultat einer Rechnung) muss im Rechner durch eine Maschinenzahl ersetzt werden:

$$x \rightarrow fl(x)$$

Der Einfluss der dabei entstehenden Fehler kann groß sein.

1.5 Fehler

A.) relativer / absoluter Fehler

Motivation: Berechne $f(x) = \frac{1 - \cos x}{x}$ für $x = 0.0001$

Taschenrechner: $\bar{f}(x) = 0.00004900000000\dots$

exakt: $f(x) = 0.00004999999995\dots$

Definition/Bezeichnung: Es bezeichne \bar{x} eine Näherung zum exakten x (oder \bar{f} zu f , usw.)

$x - \bar{x}$ (oder $\bar{x} - x$ oder $|x - \bar{x}|$) heißt **absoluter Fehler**.

$\frac{x - \bar{x}}{x}$ (oder ...) heißt **relativer Fehler**.

Praxis: x (exakt) ist unbekannt, bekannt ist nur \bar{x} , und eine Schätzung ξ für den absoluten Fehler. Dann: Schätzung für den relativen Fehler ist $|\frac{\xi}{\bar{x}}|$

Hinweis: Nur die Größenordnung des Fehlers ist von Interesse.

obiges Beispiel: maximal 2 Stellen sind exakt.

$$|f - \bar{f}| = 0.00000099\dots \approx 10^{-6}$$

$$\left| \frac{f - \bar{f}}{f} \right| = 2 \cdot 10^{-2}$$

Sprechweise (Dezimalsystem):

“Genauigkeit von j (dezimalen) Stellen” heißt:
 $|\text{relativer Fehler}| < 10^{-j}$

(im Sinne von maximalem ganzzahligem j)

j “significant digits”

Beispiel: Äquatorumfang

$$\bar{x} = 40000000 \text{ m}$$

$$x = 40075161 \text{ m}$$

\Rightarrow relativer Fehler $\frac{75}{40075} \approx 0.0019 \approx 10^{-3}$, d.h. knapp 3 Stellen genau. Absoluter Fehler hier: 75161, gibt keinen Hinweis auf Genauigkeit. Schlimmer noch: Übergang zu anderer Skalierung (z.B. km) ändert nicht die Genauigkeit, und auch nicht den relativen Fehler, leider aber den absoluten Fehler.

Merke:

Der absolute Fehler ist nicht skalierungsunabhängig und kann leicht manipuliert werden. Zur Beschreibung der Genauigkeit ist i.A. **(nur) der relative Fehler von Bedeutung.**

Problem: Grenzen des relativen Fehlers, wenn exaktes Resultat = 0.**B.) Erzeugen einer Maschinenzahl** $x \rightarrow fl(x)$ **Beispiel:** ($k = 4$ Dezimalstellen)

$$\pi = 3.1415926 \dots$$

↓ Normalisieren

$$0.31415926 \cdot 10^1$$

$$0.3141 \cdot 10^1$$

Abschneiden
(round to zero)

$$0.3142 \cdot 10^1$$

Runden
(round to nearest)**benachbarte
Maschinenzahlen****Satz 1:**

$$\left| \frac{fl(x) - x}{x} \right| \leq \epsilon_{\text{mach}}$$

(falls kein
overflow/underflow)(folgt aus Definition von ϵ_{mach})Dabei gilt $\epsilon_{\text{mach}} = B^{1-k}$ beim Abbrechen und $\epsilon_{\text{mach}} = \frac{1}{2}B^{1-k}$ beim Runden.**Kommentar:** Bereits bei der Eingabe muss man einen Fehler der Größe ϵ_{mach} erwarten!**C.) Grundrechenarten:** + − ∗ ÷ *Symbol* : ∘**Satz 2:**

Es seien x und y Maschinenzahlen. (!)
Falls kein overflow/underflow:
Der relative Fehler bei jeder der 4 Grundrechenarten
ist durch ϵ_{mach} beschränkt.

(folgt aus Satz 1: zuerst $x \circ y$ exakt als Zwischenresultat, danach $fl(x \circ y)$)

Das heißt

$$\left| \frac{fl(x \circ y) - (x \circ y)}{x \circ y} \right| \leq \epsilon_{\text{mach}}$$

Im Rechner gelten einige grundlegende mathematische Gesetze nicht! So gibt es Maschinenzahlen x und $y \neq 0$ mit $fl(x + y) = x$. (z.B. $x = 1$, $|y| < \epsilon_{\text{mach}}$).

Die Gesetze $(a + b) + c = a + (b + c)$ (*Assoziativität*) und $a(b + c) = ab + ac$ (*Distributivität*) **gelten im Rechner i.A. nicht.**

Merke:

Algebraisch äquivalente Umformungen können im Rechner zu verschiedenen Resultaten führen!

D.) Spezielle Funktionen Auch bei der Berechnung der speziellen Funktion wie z.B.

$$\sqrt{x}, \sin x, e^x, \arctan x$$

werden im Rechner durch Anwendung scharfsinniger Algorithmen der Mathematik meist die gleichen Genauigkeiten (wie oben ϵ_{mach}) erreicht.

Gelegentlich ist allerdings nur der absolute Fehler beschränkt.

Zum Beispiel ist für $x \approx 1$ die Version $\ln(1 + x)$ (mit $x \approx 0$) besser als $\ln(x)$.

E.) Algorithmen

Im Allgemeinen ist jede eingegebene Zahl x_1, x_2, \dots, x_n fehlerbehaftet. (zumindest Rundungsfehler; Fehler meist noch größer wegen Mess- und Modellfehlern!)

Die Resultate y_1, y_2, \dots, y_m hängen von der Eingabe ab:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \xrightarrow[\varphi]{\text{Zuordnung}} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Frage: Wie pflanzen sich Fehler in der Eingabe (oder in den Zwischenschritten) bis zum Resultat fort?

Beispiel: Berechne $a^2 - b^2$

Eingabe $x_1 = a, x_2 = b$, d.h. $n = 2, y(= y_1) = x_1^2 - x_2^2$, d.h. $m = 1$

1. Möglichkeit: $y = x_1^2 - x_2^2 =: \varphi(x_1, x_2)$
2. Möglichkeit: $y = (x_1 - x_2) \cdot (x_1 + x_2) =: \tilde{\varphi}(x_1, x_2)$

Zahlenbeispiel: $a = .3237$, $b = .3134$ ($B = 10$, $k = 4$)

1. Möglichkeit liefert: $0.\underline{6580} * 10^{-2}$
2. Möglichkeit liefert: $0.\underline{6562} * 10^{-2}$
- exaktes Resultat: $0.656213 * 10^{-2}$

Definition: Ein **Algorithmus** ist eine der Reihenfolge nach eindeutig festgelegte Sequenz von endlich oder unendlich vielen “elementaren” Operationen.

Frage: Was sind “gute” Algorithmen, wie lässt sich die “Güte” verbessern?

Vorbemerkungen:

1. Eine bequeme äquivalente Schreibweise für den relativen Fehler

$$\epsilon = \frac{\bar{x} - x}{x}$$

ist $\bar{x} = x(1 + \epsilon)$.

2. Δ als Symbol für absoluten Fehler: z.B. $\Delta x = \bar{x} - x$
3. Vernachlässigen kleiner Terme:

$$a + b\epsilon + c\epsilon^2 \doteq a + b\epsilon$$

wenn nur $|\epsilon|$ klein genug ist.

(Schreibweise \doteq für “in 1. Näherung”)

analog z.B.

$$a + b\epsilon + c\epsilon_1\epsilon_2 \doteq a + b\epsilon$$

Schreibweise für Terme höherer Ordnung:

$$O(\epsilon^2) \quad , \quad \text{oder T.h.O.}$$

F.) Auslöschung

Analyse der Subtraktion positiver Zahlen: $y = x_1 - x_2$

$$\bar{y} = fl(\bar{x}_1 - \bar{x}_2) \text{ mit } \bar{x}_1 = x_1(1 + \epsilon_1)$$

$$\bar{x}_2 = x_2(1 + \epsilon_2)$$

ϵ_1, ϵ_2 sind die relativen Fehler von $\bar{x}_1, \bar{x}_2, \epsilon_3$ für Subtraktion.

$|\epsilon_3|$ ist klein (Satz 2 in A.3), aber ϵ_1 und ϵ_2 können größer sein!

$$\begin{aligned} \bar{y} &= (x_1(1 + \epsilon_1) - x_2(1 + \epsilon_2))(1 + \epsilon_3) \\ &= (x_1 - x_2 + x_1\epsilon_1 - x_2\epsilon_2)(1 + \epsilon_3) \\ &= \underbrace{x_1 - x_2}_y + x_1\epsilon_1 - x_2\epsilon_2 + (x_1 - x_2)\epsilon_3 + O(\epsilon_i\epsilon_{\text{mach}}) \\ &\doteq y(1 + \epsilon_3 + \frac{x_1}{x_1 - x_2}\epsilon_1 - \frac{x_2}{x_1 - x_2}\epsilon_2) = y(1 + \epsilon_{\text{sub}}) \end{aligned}$$

Der relative Fehler im Resultat der Subtraktion, verursacht durch Eingangsfehler ϵ_1, ϵ_2 ist

$$\frac{1}{x_1 - x_2}(x_1\epsilon_1 - x_2\epsilon_2)$$

also abhängig von x_1 und x_2 .

Folgerung: Wenn $x_1 \approx x_2$, dann ist wenigstens einer der Faktoren $|\frac{x_1}{x_1-x_2}|, |\frac{x_2}{x_1-x_2}|$ sehr groß und verstärkt ϵ_1, ϵ_2 .

Damit ist der relative Fehler im Ergebnis der Subtraktion $\gg \epsilon_{\text{mach}}$

Bezeichnung: Bei Subtraktion von 2 Zahlen x_1, x_2 gleichen Vorzeichens (Addition bei verschiedenen Vorzeichen!) passiert **Auslöschung**, wenn $x_1 \approx x_2$.

Merke:

Versuche, gefährliche Subtraktion durch Umformung zu vermeiden!

Beispiel: $1 - \cos x$ für $x \approx 0$. modellhafte Überlegung:

$$\begin{aligned}
 x_1 = 1 & : 1.0000000000000 \quad \text{mit } \epsilon_1 = 0 \\
 -x_2 = -\cos x, \text{ z.B.} & : -\underbrace{0.9999998391234}_{\substack{\text{korrekt} & \text{falsch}}} \quad \text{mit } \epsilon_2 = 10^{-10} \\
 \hline
 & = .0000001608766 \quad \epsilon_{x_1-x_2} = 10^{-4} \\
 \text{normalisiert} & = 0.\underbrace{1608766\dots\dots}_{\substack{\text{korrekt "Schmutz"} \\ \text{völlig falsch!} \\ \text{(wird als Resultat mit ausgedruckt und} \\ \text{leider häufig geglaubt.)}}}
 \end{aligned}$$

Abhilfe: einige Terme der Taylorentwicklung von $\cos x$ (Übung!).

G.) Kondition

Vorbemerkung: "Linearisierung" als Approximation einer Funktion, d.h. Ersetzen der Funktion durch Tangente. (Figur ...)

$$\begin{aligned}
 & \text{Linearisierung/Taylorentwicklung} \\
 \text{Allgemein: } \left. \begin{aligned} y &= \varphi(x) \\ \bar{y} &= \varphi(\bar{x}) \end{aligned} \right\} \bar{y} = \varphi(\bar{x}) = \varphi(x + \Delta x) = \varphi(x) + \frac{d\varphi(x)}{dx} \Delta x + \text{T.h.O.} \\
 & \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \downarrow \\
 & \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad y \\
 \Rightarrow \bar{y} - y &= \varphi(\bar{x}) - \varphi(x) = \frac{d\varphi(x)}{dx} \cdot (\bar{x} - x) + \underbrace{O((\bar{x} - x)^2)}_{\text{T.h.O.}}
 \end{aligned}$$

Hinweis: $\frac{d\varphi(x)}{dx}$ ist eine symbolische Schreibweise für die erste Ableitung.

Kurzfassung für den skalaren Fall ($n = m = 1$): $\Delta y \doteq \frac{d\varphi(x)}{dx} \Delta x$. Das beschreibt die **Verstärkung** des *absoluten* Eingangsfehlers Δx .

Es folgt für den *relativen* Eingangsfehler

$$\frac{\Delta y}{y} \doteq \frac{d\varphi(x)}{dx} \frac{x}{y} \frac{\Delta x}{x} \tag{*}$$

Die **Konditionszahl des relativen Fehlers** ist $\frac{d\varphi(x)}{dx} \frac{x}{y}$, diejenige des absoluten Fehlers ist $\frac{d\varphi(x)}{dx}$.

Definition: Ein Problem heißt **schlecht konditioniert**, wenn die |Konditionszahlen| “groß” sind, sonst **gut konditioniert**. (“groß” heißt z.B. > 10).

Für den Vektorfall

$$\begin{cases} y_1 = \varphi_1(x_1, \dots, x_n) \\ \vdots \\ y_m = \varphi_m(x_1, \dots, x_n) \end{cases}$$

heißt das in erster Näherung:

$$\begin{cases} \Delta y_1 \doteq \frac{\partial \varphi_1(x)}{\partial x_1} \Delta x_1 + \dots + \frac{\partial \varphi_1(x)}{\partial x_n} \Delta x_n \\ \vdots \\ \Delta y_m \doteq \frac{\partial \varphi_m(x)}{\partial x_1} \Delta x_1 + \dots + \frac{\partial \varphi_m(x)}{\partial x_n} \Delta x_n \end{cases} \quad (**)$$

Dabei steht ∂ als Symbol für (erste) partielle Ableitungen (siehe z.B. Abschnitt 3.12). Die hier benötigten Vektoren, Matrizen, partiellen Ableitungen, Taylorentwicklung werden in den folgenden beiden Kapiteln der Vorlesung eingeführt und diskutiert.

Beispiel: $y = \varphi(x_1, x_2)$, d.h. $m = 1$, $n = 2$

$\frac{\partial \varphi(x_1, x_2)}{\partial x_1}$ ist die Ableitung nach x_1 , wobei für diese Ableitung x_2 als Konstante angesehen wird;

$\frac{\partial \varphi(x_1, x_2)}{\partial x_2}$ ist die Ableitung nach x_2 , wobei für diese Ableitung x_1 als Konstante angesehen wird.

(**) in Summenschreibweise:

$$\Delta y_i \doteq \sum_{j=1}^n \underbrace{\frac{\partial \varphi_i(x)}{\partial x_j}}_{\text{messen Einfluss des absoluten Fehlers } \Delta x_j \text{ auf } y_i} \Delta x_j$$

messen Einfluss des absoluten Fehlers Δx_j auf y_i

Relativer Fehler analog wie in (*) (für x_j , $y_i \neq 0$):

$$\frac{\Delta y_i}{y_i} \doteq \sum_{j=1}^n \underbrace{\frac{\partial \varphi_i(x)}{\partial x_j} \frac{x_j}{y_i}}_{\text{“Konditionszahlen” des relativen Fehlers}} \frac{\Delta x_j}{x_j}$$

Die Konditionszahlen des relativen Fehlers geben an, wie stark der (relative) **Eingangsfehler**

$$\epsilon_{x_j} = \frac{\Delta x_j}{x_j}$$

den (relativen) Resultatfehler

$$\epsilon_{y_i} = \frac{\Delta y_i}{y_i}$$

beeinflusst. (Fehler von Zwischenoperationen sind hierin noch nicht enthalten!)

Beispiel: (Übungen, sowie:)

\sqrt{x} ist gut konditioniert! Denn für $y = \varphi(x) = \sqrt{x}$ ergibt sich

$$\frac{d\varphi(x)}{dx} \cdot \frac{x}{y} = \frac{1}{2\sqrt{x}} \frac{x}{\sqrt{x}} = \frac{1}{2}$$

Merke: Die “Kondition” $\frac{d\varphi}{dx}$ bzw. $\frac{d\varphi}{dx} \frac{x}{y}$ hängt **nicht** vom Algorithmus ab, der das Problem lösen soll: Der Beitrag ungenauer Eingabedaten lässt sich durch keine noch so scharfsinnige Wahl eines Algorithmus beeinflussen!

(zu Rundungsfehlern siehe Abschnitt 2.9)

Kapitel 2 Lineare Algebra

2.1 Vektoren, Matrizen, Gleichungssysteme

Ein rechteckiges Schema, in welches in “Zeilen” und “Spalten” Variablen oder Zahlen eingetragen sind, heißt **Matrix**.

A.) Beispiele

Beispiel 1

Lineares Gleichungssystem, z.B.

$$\begin{cases} 7x - 3y = -5 \\ 2x + 4y = 3. \end{cases}$$

Die Koeffizienten, mit denen die Unbekannten x und y in die Gleichungen eingehen, lassen sich in einer (2×2) -Matrix darstellen:

$$\begin{pmatrix} 7 & -3 \\ 2 & 4 \end{pmatrix}$$

(2 Zeilen, 2 Spalten)

Spezialfälle: Eine Matrix, die nur aus einer Zeile besteht, heißt **Zeilenvektor**. Eine Matrix, die nur aus einer Spalte besteht, heißt **Spaltenvektor**.

Konvention: Wenn man von “Vektor” spricht, meint man Spaltenvektor.

obiges Beispiel:

$\begin{pmatrix} -5 \\ 3 \end{pmatrix}$ ist der Vektor der “rechten Seite”

$\begin{pmatrix} x \\ y \end{pmatrix}$ ist der Vektor der Unbekannten

Beispiel 2 (vgl. Kondition im Abschnitt 1.5G)

$$\begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \varphi_m}{\partial x_1} & \cdots & \frac{\partial \varphi_m}{\partial x_n} \end{pmatrix}$$

Mit m Zeilen und n Spalten ist dies eine $m \times n$ -Matrix.

Die drei Teile eines Gleichungssystems:

- 1.) Vektor der Unbekannten
- 2.) Matrix der Koeffizienten (Faktoren der Unbekannten)
- 3.) Vektor der “rechten Seite”

Hinweis: Die Reihenfolge der Unbekannten und Gleichungen spiegelt sich in der Koeffizienten-Matrix bzw. im Vektor der rechten Seite wider!

Bem.: Ein Term wie $7x$ ist “**linear**” in x , Terme wie x^2 , $\sin x$, $\frac{1}{x}$ sind **nichtlinear** in x . Allgemein ist ein additiver Term

$$\alpha x^1 \quad (= \alpha x)$$

ein linearer Term (linear in x).

Bsp.: xy^2 ist linear in x und nichtlinear in y , und xy ist nichtlinear, wenn x und y Unbekannte sind.

Die Gleichungen von Bsp. 1 und Bsp. 2 sind Systeme linearer Gleichungen (linear in allen Variablen, hier x und y).

B.) Allgemeine Schreibweise

1.) Vektor der Unbekannten:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} =: \vec{x} =: x$$

Die Anzahl n der Unbekannten heißt die “**Dimension**” des Gleichungssystems. Im Beispiel 1 gilt $n = 2$.

2.) Matrix der Koeffizienten:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} =: A$$

a_{ij} ist das “Element” in der i -ten Zeile, j -ten Spalte. (ij ist hier kein Produkt, sondern zwei nebeneinandergestellte Indices)

3.) Vektor der “rechten Seite”:

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} =: \vec{b} =: b$$

Definition: Ein **Skalar** ist eine einzelne Zahl, kein Vektor.

(schöner) Brauch:

Schreibe Matrizen als große Buchstaben (A, B, C, \dots) und Vektoren als kleine Buchstaben (x, y, c, b, \dots), evtl. mit “Pfeil” (\vec{x}, \vec{b}), und Skalare mit griechischen Buchstaben ($\alpha, \beta, \gamma, \dots$).

Bezeichnung: Die Elemente eines Vektors heißen **Komponenten**.

Hierarchie: Skalar, Vektor, Matrix (0-dim. / 1-dim. / 2-dim. “Feld”)

C.) Rechnen mit Vektoren

Es seien $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ zwei Vektoren.

Addition: $a + b = \vec{a} + \vec{b} := \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}$, also **komponentenweise**.

analog: Subtraktion $\vec{a} - \vec{b}$

Vektor * Skalar Es sei α ein Skalar.

$\alpha a = \alpha \vec{a} := \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_r \end{pmatrix}$, also wiederum komponentenweise

Vektor * Vektor : geht das ?

Definition: Skalarprodukt $\vec{a} \cdot \vec{b} := a_1 b_1 + a_2 b_2 + \dots + a_n b_n$

klar: Das Skalarprodukt ist ein Skalar!

andere Bezeichnungen: $a^t b$, $\langle a, b \rangle$, ...

1. Hinweis: Division durch einen Vektor gibt es nicht!
2. Hinweis: Aufwand: $n + (n - 1) = 2n - 1 = O(n)$ Operationen
3. Hinweis: Für den Spezialfall $n = 3$ gibt es auch ein anderes Produkt von Vektoren, das einen Vektor liefert ("Vektorprodukt").

Aus jedem Zeilenvektor entsteht ein Spaltenvektor durch "Kippen", und umgekehrt. Diesen Prozess des "Kippens" nennt man **Transponieren**. Schreibweise z.B.

\vec{a}' , \vec{a}^T , \vec{a}^t sind Zeilenvektoren;

\vec{a} (ohne Transponierungsmarkierung) ist immer ein Spaltenvektor.

D.) Längen und Winkel

geometrische Bedeutung von Vektoren: Jedem Vektor u entspricht ein "Pfeil" (gerichtetes Geradenstück) mit Fußpunkt im Koordinatenursprung und Spitze im Punkt mit den Koordinaten u (Illustration für $n = 2$ in der (x_1, x_2) -Ebene ...). So macht es Sinn, nach Längen von Vektoren und Winkeln zwischen Vektoren zu fragen.

Definition: (Euklidische) Länge eines Vektors $\|v\| := \sqrt{v^t v}$ ("Norm", Bezeichnung auch $\|v\|_2$ im Unterschied zu anderen Normen)

Definition: Einheitsvektor $\|u\| = 1$ Beispiele: $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v := \frac{u}{\|u\|}$

Definition: Winkel φ zwischen v und w :

$$\cos \varphi = \frac{v^t w}{\|v\| \|w\|}$$

Kriterium für aufeinander senkrecht stehende Vektoren ($\varphi = \pi/2$):

$$u^t v = 0 \Leftrightarrow u \text{ und } v \text{ orthogonal}$$

Schwarz'sche Ungleichung $|v^t w| \leq \|v\| \|w\|$

E.) Matrizen sind aus Vektoren aufgebaut

Jede Matrix kann man sich aus Zeilen- oder Spaltenvektoren aufgebaut denken:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = (\vec{s}_1 \quad \vec{s}_2 \quad \dots \quad \vec{s}_n) = \begin{pmatrix} \vec{z}_1^t \\ \vec{z}_2^t \\ \vdots \\ \vec{z}_n^t \end{pmatrix}$$

n Spaltenvektoren

n Zeilenvektoren

z.B. $\vec{s}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{n2} \end{pmatrix}$ (nur der "Zeilenindex läuft"), oder $\vec{z}_6^t = (a_{61}, a_{62}, \dots, a_{6n})$.

F.) lineares Gleichungssystem in Matrixschreibweise:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \tag{1}$$

Das sind n skalare Gleichungen.

Nach obigen Rechenregeln gibt es **zwei äquivalente** Schreibweisen:

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} x_2 + \dots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{2}$$

(eine Vektorgleichung; Summe von Spaltenvektoren), oder, da jede skalare Gleichung in (1) ein Skalarprodukt beinhaltet:

$$\begin{aligned} (a_{11}, a_{12}, \dots, a_{1n}) \cdot \vec{x} &= b_1 \\ (a_{21}, a_{22}, \dots, a_{2n}) \cdot \vec{x} &= b_2 \\ &\vdots \\ (a_{n1}, a_{n2}, \dots, a_{nn}) \cdot \vec{x} &= b_n \end{aligned} \tag{3}$$

(n skalare Gleichungen)

Definiton: Matrix * Spaltenvektor

$$Ax := \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} x_1 + \dots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} x_n$$

(also über Version (2), oder, äquivalent, über Version (3))

Mit diesen Konventionen lassen sich lineare Gleichungssysteme so schreiben: $A\vec{x} = \vec{b}$ oder

$$Ax = b \quad (4)$$

(oder andere Buchstaben, oder Angabe von Zahlen wie im Übungsblatt)

Die "Matrix-Vektor-Schreibweise" (4) dominiert die Literatur. Inhaltlich identisch wie (1), (2) oder (3). In Literatur meist Fettdruck statt Pfeilen

$$\mathbf{Ax} = \mathbf{b}$$

(Fettdruck für Matrizen und Vektoren, Normaldruck für Skalare).

weiteres Beispiel für Matrix * Vektor (vgl. Abschnitt 1.5G):

$$\begin{pmatrix} \Delta y \\ \vdots \\ \Delta y \end{pmatrix} \doteq \underbrace{\begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1} & \dots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \varphi_m}{\partial x_1} & \dots & \frac{\partial \varphi_m}{\partial x_n} \end{pmatrix}}_{=: D\varphi} \begin{pmatrix} \Delta x \\ \vdots \\ \Delta x \end{pmatrix}$$

Definition: transponierte Matrix: Vertauscht man die Zeilen mit den Spalten von A , dann erhält man A^t .

Beispiele:

$$1.) A = \begin{pmatrix} 2 & 5 \\ 3 & -1 \\ 0 & 1 \end{pmatrix}, \quad A^t = \begin{pmatrix} 2 & 3 & 0 \\ 5 & -1 & 1 \end{pmatrix}$$

2.) Vektor $a \Rightarrow a^t$ ist Zeile

Definition: symmetrische Matrix, wenn $A = A^t$ (d.h. $a_{ij} = a_{ji}$ für alle i, j)

Definition: Nullmatrix 0: enthält nur Nullen

Definition: Einheitsmatrix:

$$I := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

2.2 Der Algorithmus von Gauß

Gegeben: n skalare, lineare Gleichungen für n Unbekannte x_1, \dots, x_n :

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \quad \text{oder } Ax = b$$

Die i -te Gleichung lautet $\sum_{j=1}^n a_{ij}x_j = b_i$

Idee: Addiere geeignete Vielfache von Zeile k auf Zeile i , derart, dass Nullen entstehen. Strategie der Wahl von i, k , so dass ein mit $Ax = b$ **äquivalentes** Gleichungssystem $\tilde{A}x = \tilde{b}$ entsteht, bei dem \tilde{A} obere "Dreiecksmatrix" ist. Dies geschieht i.A. "vorwärts", vgl. Algorithmus unten. Resultat dieser **ersten Phase**:

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot \\ & & & & & \cdot \\ & & & & & & 0 & \cdot & \cdot \\ & & & & & & & & \cdot \end{pmatrix} x = \tilde{b}$$

Danach 2. Phase: Rückwärtselimination:

$$x_n = \frac{\tilde{b}_n}{\tilde{a}_{nn}} \Rightarrow x_{n-1} = \frac{1}{\tilde{a}_{n-1,n-1}} (\tilde{b}_{n-1} - x_n \tilde{a}_{n-1,n})$$

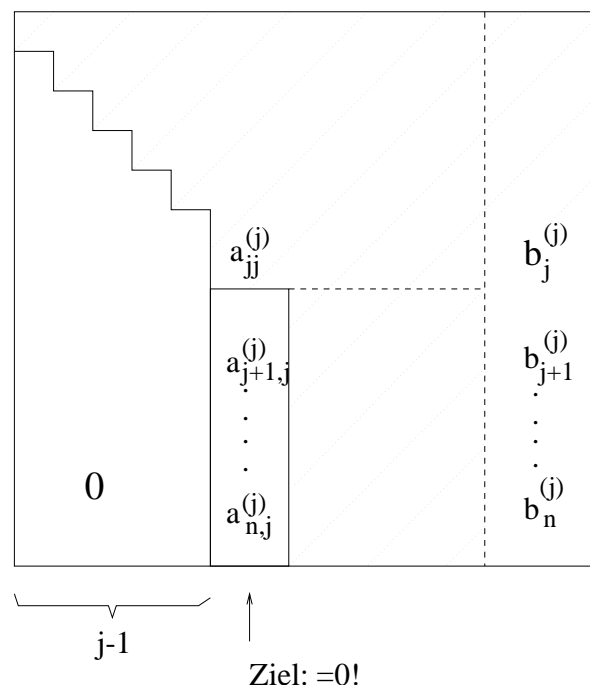
u.s.w. erhält man x_{n-2}, \dots , bis x_2, x_1

Kosten der ersten Phase: $O(n^3)$ Operation, s.u.

Kosten der zweiten Phase: in jeder Zeile i.W. ein Skalarprodukt (von zunehmender Länge), also jeweils $O(n)$ Operationen. Macht zusammen $O(n^2)$ Operationen für die Rückwärtselimination.

Für die Zusammenfassung lasse die Tilde (\sim) weg. Die Restmatrizen werden jeweils verändert; aktuelle Version mit hochgestellten Indices:

j -ter Hauptschritt:



Algorithmus

Für $k = 1, 2, \dots, n - 1$:
 Annahme $a_{kk} \neq 0$ (anderenfalls Zeilenvertauschung von Zeile k mit Zeile "unterhalb": *Pivotstrategie*)

für $i = k + 1, k + 2, \dots, n$:
für $j = k + 1, k + 2, \dots, n$:

$$a_{ij} := a_{ij} - \frac{a_{ik}}{a_{kk}} a_{kj}$$

$$b_i := b_i - \frac{a_{ik}}{a_{kk}} b_k$$

$$x_n := \frac{b_n}{a_{nn}}$$

Für $i = n - 1, n - 2, \dots, 1$:

$$x_i := (b_i - \sum_{j=i+1}^n a_{ij} x_j) / a_{ii}$$

Hinweis: Die ursprüngliche Matrix A wird dabei überschrieben.

Pivotstrategie: Wichtige Klassen von Matrizen brauchen keine Pivotstrategie, da $a_{kk} \neq 0$ jeweils garantiert werden kann. Solche Matrizen sind z.B.

- (1) die diagonal-dominanten Matrizen ($|a_{kk}| > \sum_{i \neq k}^n |a_{ki}|$ für alle k)
- (2) die positiv-definiten Matrizen ($A^T = A$ und $x^T A x > 0$ für $x \neq 0$)

Kosten: Drei ineinandergeschachtelte Schleifen deuten auf $O(n^3)$ hin. Konkret: $\frac{2}{3}n^3 + O(n^2)$ Operationen (+, -, ·, /). Ein kompaktes Programm ist möglich mit $a_{i,n+1} := b_i$, d.h. Anhängen von b an A . Es können auch mehrere Vektoren b gleichzeitig bearbeitet werden.

Warnung: Für $n > 3$ **nicht** die Cramersche Regel benutzen, da viel zu teuer bzw. nicht durchführbar!

2.3 Arbeiten mit Matrizen

Bezeichnung der Elemente der Matrizen A, B, C :

$$A = ((a_{ij})) , \quad B = ((b_{ij})) , \quad C = ((c_{ij}))$$

A.) Addition von Matrizen

erfolgt komponentenweise: $C = A + B$ mit

$$c_{ij} := a_{ij} + b_{ij} \quad \text{für alle } i, j .$$

Voraussetzung: Die Matrizen müssen gleich groß sein. Analog: Subtraktion.

Regeln:

$$\begin{aligned} A + B &= B + A \\ (A + B) + C &= A + (B + C) \end{aligned}$$

B.) Multiplikation von Matrizen

Matrizen A und B müssen für das Produkt AB wie folgt zusammenpassen:

$$A \text{ sei } m \times n \quad \text{und} \quad B \text{ sei } n \times l .$$

Dann ist $C := AB$ definiert über

$$c_{ij} := \sum_{k=1}^n a_{ik} b_{kj} \quad \text{für } i = 1, \dots, m, \quad j = 1, \dots, l.$$

Also: Skalarprodukt der i -ten Zeile der Matrix A mit der j -ten Spalte von B .

Vorsicht mit der Bezeichnung: In dieser Formel steht A für den linken Faktor und B für den rechten Faktor. (gegebenenfalls an andere Bezeichnungen anpassen)

Die Produktmatrix C ist also $m \times l$.

Spezialfall $m = l = 1$: Zeile * Vektor, d.h. Skalarprodukt, z.B. $u^t v$

Rechenregeln:

$$A(BC) = (AB)C = ABC \quad (\text{Klammern hier unnötig})$$

$$C(A + B) = CA + CB$$

$$(AB)^t = B^t A^t$$

aber: i.A. $AB \neq BA$!

möglich ist:

Matrix * Spaltenvektor (z.B. Ax)

Zeilenvektor * Matrix (z.B. $y^t A$)

$y^t Ax$ ist Skalar

nicht möglich:

Matrix * Zeile, Spalte * Matrix,

Division durch Matrix.

C.) Inverse Matrix

Voraussetzung: A ist quadratisch: $n \times n$.

Falls es eine Matrix B gibt mit $AB = I$, dann heißt B **inverse Matrix** von A .

Bezeichnung: A^{-1}

Dann gilt

$$AA^{-1} = I = A^{-1}A$$

Vorsicht: Die Inverse existiert nicht zu jeder Matrix! Matrizen, zu denen es keine Inverse gibt, heißen **singulär**.

Notwendiges und hinreichendes **Kriterium für Singularität von A** :

Es gibt Vektor $x \neq 0$ mit $Ax = 0$.

Die **nichtsingulären** Matrizen heißen auch **regulär**.

Spezialfall 2×2 Matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$ad - bc$ heißt die **Determinante** von $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

A invertierbar \Leftrightarrow Determinante $\neq 0$.

Rechenregeln:

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A^{-1})^t = (A^t)^{-1}$$

Falls A^{-1} existiert (A nichtsingulär), dann muss A n Pivotelemente $\neq 0$ haben.

D.) Permutationsmatrizen

Definition: Eine $n \times n$ -Matrix P heißt **Permutationsmatrix**, wenn in jeder Zeile und in jeder Spalte von P genau eine 1 steht und sonst nur Nullen.

(Beispiel einer Permutationsmatrix...)

Anwendung::

PA vertauscht Zeilen von A ,

AP vertauscht Spalten von A .

E.) Orthogonale Matrizen

Eine quadratische Matrix Q mit der Eigenschaft $Q^t Q = I$ heißt **orthogonal**. Es gilt also $Q^t = Q^{-1}$. Die Spalten von Q sind "orthonormal", d.h. orthogonal mit Einheitslänge.

Beispiel 1: Drehmatrix $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$

Die Multiplikation

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

bewirkt eine **Drehung** des Vektors $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ um den Winkel φ im mathematisch positiven Sinn (d.h. gegen den Uhrzeigersinn).

Illustration am Beispiel mit $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\varphi = \frac{\pi}{2}$.

Beispiel 2: Jede Permutationsmatrix, weil die Spalten orthonormal sind.

Beispiel 3: Reflexionen, Householder-Matrizen, vgl. §2.9, z.B.

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

Qx vermittelt eine Spiegelung an der Ebene $x_3 = 0$.

F.) LU-Zerlegung: Falls A faktorisiert ist in $A = LU$,

$$\text{wobei } L = \begin{pmatrix} 1 & & & 0 \\ \cdot & 1 & & \\ \cdot & \cdot & 1 & \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, U = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

dann ist das Lösen von $Ax = b$ einfach: Wegen

$$Ax = \underbrace{LUx}_{=y} = b = Ly$$

1. Schritt: Löse $Ly = b$: "Vorwärtssubstitution"
2. Schritt: Löse $Ux = y$: "Rückwärtselimination"

(Häufige Bezeichnung R statt U , dann also LR -Zerlegung.)

Frage: Existiert eine Zerlegung $A = LU$ immer?

Antwort: Nur wenn A nichtsingulär (also “regulär”) ist und man Zeilenvertauschungen von A zulässt. Zur Formalisierung benutze Permutationsmatrizen.

Satz: Für jede nichtsinguläre (= reguläre) Matrix A gibt es eine Permutationsmatrix P sowie Dreiecksmatrizen L und U von obigem Typ, so dass

$$PA = LU$$

Ausführung der LU -Zerlegung ist **äquivalent** zum Gaußschen Algorithmus.

G.) Qualität der Lösung

Definition: Residuum: Es sei \bar{x} eine berechnete “Lösung”, dann heißt der Vektor $r := b - A\bar{x}$ das Residuum von $Ax = b$.

Das Residuum ist durch Zeilenskalierung $DAx = Db$ beliebig manipulierbar (D : Diagonalmatrix). Denn: neues Residuum $= Db - DA\bar{x} = Dr$.

Warnung: kleines Residuum $\not\Rightarrow$ genaue Lösung

2.4 Vektorräume

$$\mathbb{R}^n := \{ \text{Spaltenvektoren mit } n \text{ Komponenten} \}$$

Definition: Ein **Vektorraum** ist eine Menge, deren Elemente durch zwei Operationen verknüpft werden können, hier Addition und Multiplikation. Dabei gelten die üblichen (Vektor-)Rechenregeln, und Addition $u + v$ und Multiplikation mit Skalaren αu führen nicht aus dem Raum heraus.

D.h. das Ergebnis gehört zu dem gleichen Raum. “Reeller” Vektorraum, weil $\alpha \in \mathbb{R}$.

Beispiele:

- 1.) \mathbb{R}^n
- 2.) der Raum aller reellen Funktionen.

Unterräume von Vektorräumen enthalten immer die 0, $u + v$ und αu .

Beispiele:

$$1.) \left\{ \left(\begin{array}{c} x_1 \\ \vdots \\ x_{n-1} \\ 0 \end{array} \right) \mid x_i \in \mathbb{R} \right\}$$

- 2.) **Spaltenraum** einer Matrix (“Bild einer Matrix”)

$$:= \{ Ax \mid x \in \mathbb{R}^n \}$$

$$= \{ \text{Linearkombinationen der Spalten von } A \}$$

Nachweis: $A \cdot 0 = 0$, u, v im Spaltenraum $\Rightarrow \exists x, y$ mit $u = Ax$, $v = Ay$ (usw.)

2a) speziell mit $A = \begin{pmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{pmatrix}$

Ebene, aufgespannt von den Spalten: $u = \alpha \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix}$

Bemerkung: Wenn A eine $m \times n$ -Matrix ist (m Zeilen, n Spalten) und $x \in \mathbb{R}^n$, dann ist $Ax \in \mathbb{R}^m$. Das Ergebnis Ax kann man als "Bild" von x ansehen. (Illustration Urbildraum, Bildraum ...) Wie das Beispiel 2a) mit $m = 3$ zeigt, muss der Bildraum nicht der "volle" \mathbb{R}^m sein.

3.) **Kern** von A

$$\text{Kern}(A) := \{\text{Lösung } x \text{ von } Ax = 0\}$$

Für alle invertierbaren Matrizen ist dies nur der Nullvektor.

$\text{Kern}(A)$ ist Unterraum, **denn:** Es gelte $Ax = 0$ und $Ay = 0$.

$$\text{dann folgt } \begin{cases} A(x+y) & = Ax + Ay = 0 \\ A(\alpha x) & = \alpha Ax = 0 \end{cases}$$

3a) $x + 2y + 3z = 0$ im \mathbb{R}^3 beschreibt **Ebene** durch Koordinaten-Ursprung.

$$A = (1 \quad 2 \quad 3)$$

$x + 2y + 3z = 6$ ist auch Ebene, aber kein Unterraum, weil die 0 nicht enthalten ist.

Zurück zu $Ax = b$: Frage: Zählen identische Gleichungen oder $0 = 0$ mit?

Definition: Rang einer Matrix ist die Anzahl ihrer Pivotelemente ($\neq 0$).

Bezeichnung: r

Definition: Treppenmatrix, auch "Zeilenstufenform"

$$U = \begin{pmatrix} * & & & & & & & & & \\ 0 & * & * & & & & & & & \\ 0 & 0 & 0 & 0 & * & & & & & \\ 0 & 0 & 0 & 0 & 0 & * & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Diese Beispiel-(5×9)-Matrix hat 4 Pivotelemente. * bedeutet: Zahl $\neq 0$; der Rest der Zeilen ist mit beliebigen Zahlen gefüllt.

Auftreten: Wie beim Gaußschen Algorithmus gezeigt, lassen sich Treppenmatrizen durch Zeilenoperationen aus jeder Matrix A erzeugen! (einschließlich Zeilenvertauschungen)

Beispiel: Bei obiger 5×9 -Matrix U ist $r = 4$.

$$\text{Spaltenraum: } \left\{ \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ 0 \end{pmatrix} \right\} \subset \mathbb{R}^5$$

Kern(U): "Pivotvariablen" sind diejenigen x_j , bei denen Pivotelemente stehen, also hier x_1, x_2, x_5, x_6 . "Freie Variablen": alle anderen, hier also:

$$x_3, x_4, x_7, x_8, x_9 .$$

Die freien Variablen können beliebig gesetzt werden! (Z.B. mit Nullen und Einsen.) Anschließend erhält man die Pivotvariablen durch Lösen von $Ux = 0$.

Jede Setzung der freien Variablen führt auf eine "spezielle Lösung". Z.B. setze willkürlich $x_3 = 1, x_4 = x_7 = x_8 = x_9 = 0$.

Vereinbarung: In jeder speziellen Lösung der Gleichung $Ax = 0$ bzw. $Ux = 0$ wird eine der freien Variablen auf 1, die anderen auf 0 gesetzt.

Zahlenbeispiel [Strang] $U = \begin{pmatrix} 1 & 5 & 7 \\ 0 & 0 & 9 \end{pmatrix}$

freie Variable: x_2 (z.B. setze $x_2 = 1$)

allgemeine Lösung von $Ux = 0$:

$$\begin{aligned} x_3 = 0 &\Rightarrow x_1 + 5x_2 + 0 = 0 \\ \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} -5x_2 \\ x_2 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} -5 \\ 1 \\ 0 \end{pmatrix} \quad (\text{Gerade im } (x, y, z) \text{ - Raum}) \end{aligned}$$

spezielle Lösung: $\begin{pmatrix} -5 \\ 1 \\ 0 \end{pmatrix}$

Bei jeder $m \times n$ -Matrix gilt $\text{Rang} \leq m$ **und** $\text{Rang} \leq n$.

Man sagt, der Rang ist "voll", wenn er maximal ist, d.h. $r = \min(n, m)$.

Lösungen von $Ax = b$:

Für $Ax = 0$ gibt es $n - \text{Rang}(A)$ viele spezielle Lösungen.

Homogene Lösung x_h löst $Ax = 0$. ($Ax = 0$ heißt auch "homogenes Gleichungssystem".)

Eine partikuläre Lösung x_p erhält man, wenn alle freien Variablen = 0 gesetzt werden: $Ax_p = b$.

allgemeine Lösung: $x = x_h + x_p$ (denn: $Ax = 0 + Ax_p = b$)

2.5 Dimension und Basis von Vektorräumen

Definition: Die Spalten einer Matrix heißen **linear unabhängig**, wenn $x = 0$ die einzige Lösung von $Ax = 0$ ist. (d.h. A ist regulär)

Noch allgemeiner (nicht nur bei Matrizen):

n **Vektoren** $\vec{v}_1, \dots, \vec{v}_n$ heißen linear unabhängig, wenn

$$x_1 \vec{v}_1 + \dots + x_n \vec{v}_n = \vec{0} \Rightarrow x_1 = \dots = x_n = 0 .$$

(D.h. der Nullvektor lässt sich nur durch die triviale Kombination $0 \cdot \vec{v}_1 + \dots + 0 \cdot \vec{v}_n$ darstellen.)

Anderenfalls heißen die Vektoren **linear abhängig**.

Beispiel im \mathbb{R}^2 :

$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sind linear unabhängig

$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1.01 \end{pmatrix}$ sind linear unabhängig

$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ sind linear abhängig

$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ sind linear abhängig

Im \mathbb{R}^n gibt es maximal n linear unabhängige Vektoren.

Beispiel: Sind die Vektoren $v_1 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 5 \\ 6 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ linear unabhängig?

$$\rightarrow x_1 v_1 + x_2 v_2 + x_3 v_3 = 0$$

$$\Leftrightarrow x_1 \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 5 \\ 6 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \underbrace{\begin{pmatrix} 1 & 5 & 3 \\ -2 & 6 & 2 \\ 3 & -1 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

allgemeine Lösung von $Ax = 0$:

$$x = \begin{pmatrix} -\frac{1}{2}c \\ -\frac{1}{2}c \\ c \end{pmatrix}, \quad c \in \mathbb{R} \text{ beliebig}$$

Insbesondere ist $x_1 = -\frac{1}{2}, x_2 = -\frac{1}{2}, x_3 = 1$ eine Lösung.

$\Rightarrow v_1, v_2, v_3$ sind linear abhängig.

Definition: Eine Menge von Vektoren **erzeugt** einen Raum, wenn die Menge der Linearkombinationen den Raum genau ausfüllt. (auch "aufspannen" statt "erzeugen")

Menge der **Linearkombinationen** von n Vektoren $\vec{v}_1, \dots, \vec{v}_n$:

$$\left\{ \sum_{i=1}^n \alpha_i \vec{v}_i \mid \alpha_i \in \mathbb{R} \right\}$$

Bezeichnung: $\text{span}(\vec{v}_1, \dots, \vec{v}_n)$ oder "lineare Hülle"

Beispiel: Der Spaltenraum einer Matrix A wird von den Spalten erzeugt.

Defintion: Eine **Basis** eines Vektorraumes ist eine Menge von Vektoren mit den *beiden* Eigenschaften:

- 1.) Die Vektoren sind linear unabhängig.
- 2.) Die Vektoren erzeugen den Vektorraum.

Beispiel: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ bilden die sogenannte "Standard-Basis" des \mathbb{R}^3 (analog \mathbb{R}^n)

Beispiel: Die Spalten jeder invertierbaren $n \times n$ -Matrix bilden eine Basis des \mathbb{R}^n .

Sind $\vec{v}_1, \dots, \vec{v}_m$ und $\vec{w}_1, \dots, \vec{w}_n$ zwei Basen desselben Vektorraumes, so gilt $m = n$.

Defintion: Die **Dimension eines Vektorraumes** ist durch die Anzahl der Vektoren in jeder Basis des Raums definiert.

Beispiel: \mathbb{R}^n ist n -dimensional.

Beispiel: $\{1, x, x^2\}$ ist Basis des Raumes aller Polynome vom Grad ≤ 2

$$\{\alpha_0 + \alpha_1 x + \alpha_2 x^2 \mid \alpha_i \in \mathbb{R}\}$$

Beispiel: Der Kern von A hat die Dimension $n - \text{Rang}(A)$, und die speziellen Lösungen bilden eine Basis von $\text{Kern}(A)$.

Beispiel: Matrix vom Rang 1:

$$A = uv^t = \text{Spalte} * \text{Zeile}$$

Wichtige Eigenschaften:

Satz: Es sei A eine reelle $m \times n$ Matrix (m Zeilen, n Spalten). Dann gelten die Eigenschaften:

- 1.) $\text{Rang}(A) = \text{Rang}(A^t)$
- 2.) $\dim(\text{Bild}(A)) + \dim(\text{Kern}(A)) = n$
- 3.) $\dim(\text{Bild}(A^t)) + \dim(\text{Kern}(A^t)) = m$
- 4.) $\text{Rang}(A) = \dim(\text{Bild}(A))$

Beispiel: lineares Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 6 \\ x_2 + 2x_4 &= 3 \\ 3x_1 - x_4 &= 2 \\ x_1 + 3x_2 + 3x_3 + 2x_4 &= 9 \end{aligned}$$

$$\Leftrightarrow Ax = b \text{ mit } A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 0 & -1 \\ 1 & 3 & 3 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 3 \\ 2 \\ 9 \end{pmatrix}$$

Anwendung des Gaußschen Algorithmus liefert:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 6 \\ 0 & 1 & 0 & 2 & 3 \\ 3 & 0 & 0 & -1 & 2 \\ 1 & 3 & 3 & 2 & 9 \end{array} \right) \rightarrow \dots \rightarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 6 \\ 0 & 1 & 0 & 2 & 3 \\ 0 & 0 & -9 & 11 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$\rightarrow \text{Rang}(A) = 3,$

$\text{Kern}(A) = \{x \in \mathbb{R}^4 \mid x_4 = a, x_3 = \frac{11}{9}a, x_2 = -2a, x_1 = \frac{1}{3}a, a \in \mathbb{R}\}$

$$= \left\{ a \cdot \begin{pmatrix} \frac{1}{3} \\ -2 \\ \frac{11}{9} \\ 1 \end{pmatrix} \mid a \in \mathbb{R} \right\} = \text{span} \left(\begin{pmatrix} \frac{1}{3} \\ -2 \\ \frac{11}{9} \\ 1 \end{pmatrix} \right)$$

\rightarrow Lösungsmenge des homogenen Gleichungssystems

x_4 ist freie Variable, $\dim(\text{Kern}(A)) = 1.$

partikuläre Lösung von $Ax = b$:

$$x_p = \begin{pmatrix} \frac{2}{3} \\ 3 \\ -\frac{2}{9} \\ 0 \end{pmatrix}$$

\Rightarrow allgemeine Lösung von $Ax = b$:

$$x = x_h + x_p = \begin{pmatrix} \frac{2}{3} + \frac{1}{3}c \\ 3 - 2c \\ -\frac{2}{9} + \frac{11}{9}c \\ c \end{pmatrix}, c \in \mathbb{R}$$

2.6 Determinanten

- $A \in \mathbb{R}^{2 \times 2}$: $\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$

Schema: siehe Vorlesung

- $A \in \mathbb{R}^{3 \times 3}$:

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{aligned}$$

Schema: siehe Vorlesung (Regel von Sarrus)

Definition: Sei $A \in \mathbb{R}^{n \times n}$. Der **Minor** M_{ij} des Elements a_{ij} ist als Determinante derjenigen Untermatrix definiert, die sich durch Streichen der i -ten Zeile und j -ten Spalte aus A ergibt. Die Zahl $C_{ij} := (-1)^{i+j} M_{ij}$ heißt **Kofaktor** des Elements a_{ij} .

Beispiel:

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

$$M_{32} = \det \begin{pmatrix} 3 & -4 \\ 2 & 6 \end{pmatrix} = 26, \quad C_{32} = (-1)^{3+2} M_{32} = -26$$

Vorzeichen für C_{ij} aus Schema:

$$\begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & & & & \end{pmatrix}$$

Satz: Sei $A \in \mathbb{R}^{n \times n}$. Dann gilt für jedes $1 \leq i \leq n$ und jedes $1 \leq j \leq n$:

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \dots + a_{nj}C_{nj} = \sum_{k=1}^n a_{kj}C_{kj}$$

(Entwicklung nach der j -ten Spalte)

und

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \dots + a_{in}C_{in} = \sum_{k=1}^n a_{ik}C_{ik}$$

(Entwicklung nach der i -ten Zeile)

Beispiel: Entwicklung nach der 1. Zeile

$$\det \begin{pmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{pmatrix} = 3 \cdot \det \begin{pmatrix} -4 & 3 \\ 4 & -2 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} -2 & 3 \\ 5 & -2 \end{pmatrix} + 0 \cdot \det \begin{pmatrix} -2 & -4 \\ 5 & 4 \end{pmatrix}$$

$$= \dots = -1$$

Beispiel: Entwicklung nach der 1. Spalte

$$\det \begin{pmatrix} 0 & 2 & -1 & 1 \\ 1 & 3 & 4 & 2 \\ 0 & -1 & 1 & -3 \\ 1 & 0 & 3 & -2 \end{pmatrix} = -1 \cdot \det \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -3 \\ 0 & 3 & -2 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 2 & -1 & 1 \\ 3 & 4 & 2 \\ -1 & 1 & -3 \end{pmatrix}$$

$$= \dots = 15$$

Satz: Es seien $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$. Dann sind folgende Aussagen äquivalent:

- (1) $Ax = b$ ist eindeutig lösbar.
- (2) A ist regulär.
- (3) $\text{Kern}(A) = \{0\}$.
- (4) $\det(A) \neq 0$.

Satz:

- (1) $\det(AB) = \det(A) \det(B)$
- (2) $\det(A^T) = \det(A)$

Orthogonale Matrizen Q haben die Determinante $+1$ oder -1 , denn:

$$1 = \det(I) = \det(Q^T Q) = \det(Q^T) \det(Q) = \det(Q) \det(Q) = (\det(Q))^2$$

2.7 Eigenwerte

A $n \times n$ Matrix, $A = ((a_{ij}))_{i=1, \dots, n, j=1, \dots, n}$

x n -Vektor (Komponenten evtl. komplexe Zahlen; vgl. späteren Abschnitt 3.2)

Definition: Falls es ein $x \neq 0$ gibt mit einem λ , so dass

$$Ax = \lambda x$$

gilt, dann heißt x **Eigenvektor** und λ **Eigenwert** (wiederum evtl. komplexe Zahlen).

Erste Eigenschaften:

- (1) A hat höchstens n verschiedene Eigenwerte.
- (2) Wenn A symmetrisch ($a_{kj} = a_{jk}$) ist, dann sind die Eigenwerte reell.
- (3) Eigenvektoren zu verschiedenen Eigenwerten sind **linear unabhängig**.
(Hinweis: Jeder n -Vektor y kann eindeutig durch n linear unabhängige Vektoren $x^{(1)}, \dots, x^{(n)}$ dargestellt werden: $y = \sum_{k=1}^n c_k x^{(k)}$. Insbesondere wenn A n verschiedene Eigenwerte hat, können die $x^{(k)}$ als Eigenvektoren gewählt werden.)
- (4) Eigenwerte sind Nullstellen des **charakteristischen Polynoms**

$$\varphi(\lambda) := \det(A - \lambda I),$$

wobei I die Einheitsmatrix ist.

[denn:

$$\begin{aligned} \exists x \neq 0 : Ax = \lambda x &\Leftrightarrow \exists x \neq 0 : (A - \lambda I)x = 0 \\ &\Leftrightarrow \det(A - \lambda I) = 0 \end{aligned}$$

Definition: Mehrfacher Eigenwert: λ_k ist ein mehrfacher Eigenwert, wenn es eine mehrfache Nullstelle des charakteristischen Polynoms $\varphi(\lambda)$ ist.

(entsprechend: "einfacher Eigenwert")

Beispiel: $A = \begin{pmatrix} -7 & 13 & -16 \\ 13 & -10 & 13 \\ -16 & 13 & -7 \end{pmatrix}$

charakteristisches Polynom $\varphi(\lambda) =$

$$\det \begin{pmatrix} -7 - \lambda & 13 & -16 \\ 13 & -10 - \lambda & 13 \\ -16 & 13 & -7 - \lambda \end{pmatrix}$$

(entwickelt nach erster Zeile:)

$$\begin{aligned} &= - (7 + \lambda) \{ (10 + \lambda)(7 + \lambda) - 13^2 \} - 13 \{ -13(7 + \lambda) + 13 \cdot 16 \} - 16 \{ 13^2 - 16(10 + \lambda) \} \\ &= \dots = -\lambda^3 - 24\lambda^2 + 405\lambda - 972 \\ &= -(\lambda - 3)(\lambda - 9)(\lambda + 36) \\ \Rightarrow \lambda_1 &= 3, \quad \lambda_2 = 9, \quad \lambda_3 = -36 \quad \text{sind die Eigenwerte} \end{aligned}$$

Berechnung der Eigenvektoren über lineare Gleichungssysteme:

z.B. für λ_3 :

$$(A - \lambda_3 I)x = (A + 36I)x = \begin{pmatrix} 29 & 13 & -16 \\ 13 & 26 & 13 \\ -16 & 13 & 29 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \stackrel{!}{=} 0$$

Bemerkung:

- (1) Wegen der Singularität von $(A - \lambda I)$ ist wenigstens eine der drei skalaren Gleichungen redundant.
- (2) Falls x Eigenvektor ist, dann auch αx für alle Faktoren $\alpha \neq 0$. Also ist x so normierbar, dass eine Komponente = 1 ist.

obiges Beispiel: Versuch mit 1. Komponente:

$$x_1 = 0 \Rightarrow x_2 = x_3 = 0,$$

also: $x_1 \neq 0$. Setze $x_1 = 1$.

Es bleibt:

$$\left. \begin{aligned} 13x_2 - 16x_3 &= -29 \\ 26x_2 + 13x_3 &= -13 \\ 13x_2 + 29x_3 &= 16 \end{aligned} \right\} \rightarrow 3. \text{ Gleich.} - 1. \text{ Gleich.}$$

$$45x_3 = 45 \Rightarrow x_3 = 1 \Rightarrow x_2 = \frac{-13}{13} = -1$$

Also:

$$x^{(3)} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } \lambda_3 = 36;$$

analoges Vorgehen für λ_1, λ_2 .

2.8 Normen

Abstand von skalaren Zahlen $|x - \bar{x}|$;

Ziel: Verallgemeinern auf Vektoren und Matrizen.

A.) Matrix-Normen:

Vektor-Normen $\|x\|$ haben die “Norm-Eigenschaften”

$$(V1) \quad \|x\| \geq 0$$

$$(V2) \quad \|x\| = 0 \Leftrightarrow x = 0$$

$$(V3) \quad \|\alpha x\| = |\alpha| \|x\| \text{ für Skalar } \alpha$$

$$(V4) \quad \|x + y\| \leq \|x\| + \|y\| \text{ (Dreiecksungleichung)}$$

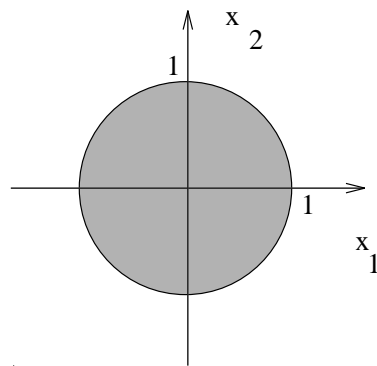
(geometrische Bedeutung der Dreiecksungleichung im \mathbb{R}^2 ...)

Beispiele:

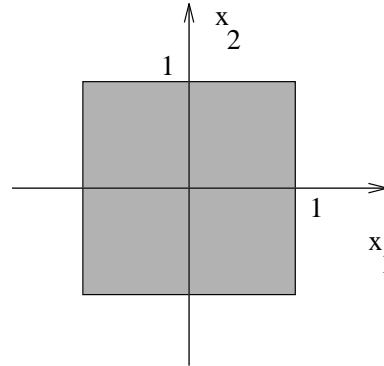
$$\|x\|_2 := \sqrt{\sum_{i=1}^n x_i^2} \quad \text{“Euklidische Norm”, Länge, “2-Norm”}$$

$$\text{(Folgerung: } \|x\|_2^2 = x^T x \text{)}$$

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i| \quad \text{“Maximum-Norm”, “Unendlich-Norm”}$$



$$\|x\|_2 \leq 1$$



$$\|x\|_\infty \leq 1$$

(Erklärung der Indices 2, ∞ über p -Norm...)

Die zur Vektornorm $\|x\|$ passende “**natürliche**” **Matrixnorm** ist definiert als

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

(Von der ursprünglichen Definition als “supremum” auch der Name *least upper bound*; heißt auch “lub-Norm”.) Es gilt wegen (V3)

$$\|A\| = \max_{\|y\|=1} \|Ay\|, \quad \text{weil } y := \frac{x}{\|x\|}.$$

Die natürliche Matrixnorm hat die vier Eigenschaften analog zu den Vektornormen, und zusätzlich die Eigenschaften

$$(M5) \quad \|AB\| \leq \|A\| \|B\| \quad \text{(Submultiplikativität)}$$

$$(M6) \quad \|Ax\| \leq \|A\| \|x\| \quad \text{(Verträglichkeit)}$$

$$(M7) \quad \|I\| = 1.$$

Vorbemerkungen zur Herleitung von $\|A\|_2$:

$Bx = \lambda x$ Eigenwertproblem mit $x \neq 0$ (x Eigenvektor, λ Eigenwert)

$Bx = \lambda x \Rightarrow \lambda = \frac{x^T Bx}{x^T x}$. Dieser Quotient heißt Rayleigh-Quotient.

$B := A^T A$ ist symmetrisch. (weil wegen $(CD)^T = D^T C^T$ gilt: $B^T = B$)

B symmetrisch \Rightarrow max. Rayleigh-Quotient ist der maximale Eigenwert

Beispiele:

$$\|A\|_2 := \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max \sqrt{\frac{x^T A^T A x}{x^T x}} = \sqrt{\lambda_{\max}(A^T A)} = \sqrt{\text{größter Eigenwert von } A^T A}$$

Folgerung: $\|H\|_2 = 1$ für orthogonale Matrizen H . (H orthogonal $\Leftrightarrow H^T = H^{-1}$)

$$\|A\|_\infty := \max_{\|y\|_\infty=1} \|Ay\|_\infty = \max_i \sum_{k=1}^n |a_{ik}| = \text{größte Zeilensumme}$$

B.) Störungen in b :

Im Folgenden sei die Matrixnorm immer die zur Vektornorm passende.

Frage: Was ist die Auswirkung Δx auf die Lösung x von $Ax = b$, wenn b um Δb verändert wird?

$$A(x + \Delta x) = b + \Delta b \quad \Rightarrow \quad A\Delta x = \Delta b$$

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

$$\text{andererseits: } b = Ax \Rightarrow \|b\| \leq \|A\| \|x\| \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

$$\text{zusammen: } \frac{\|\Delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}$$

Definition: Konditionszahl einer Matrix

$$\text{cond}(A) := \|A\| \|A^{-1}\|$$

Also ist der Einfluss relativer Fehler in b auf den relativen Fehler in x durch den Faktor $\text{cond}(A)$ beschränkt. (relative Fehler in obigem Sinne, nicht komponentenweise)

Es gilt: $\text{cond}(A) \geq 1$, denn

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = \text{cond}A$$

Also: Für eine **gut-konditionierte** Matrix gilt: $\text{cond}(A)$ "nahe" bei 1 (z.B. < 10).

C.) Änderungen in A :

$$A \rightarrow A + \Delta A$$

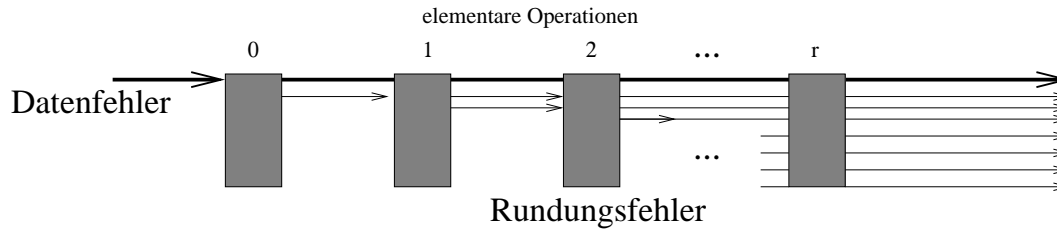
Man kann zeigen: Falls die Störung ΔA klein ist im Sinne $\|A^{-1}\| \|\Delta A\| < 1$, dann gilt

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \text{cond}(A) \frac{\|\Delta A\|}{\|A\|}} \frac{\|\Delta A\|}{\|A\|},$$

also ist auch hier die Konditionszahl $\text{cond}(A)$ entscheidend.

2.9 Householder-Matrizen

Wie wirken sich Rundungsfehler aus? Die Rundungsfehler jedes Zwischenschrittes wirken sich auf den gesamten restlichen Algorithmus aus!



Wenn φ_i die einzelnen Operationen sind, kann ein Algorithmus $y = \varphi(x)$ mit r Operationen dargestellt werden als

$$x \rightarrow \varphi_0(x) \rightarrow \varphi_1(\varphi_0(x)) \rightarrow \dots \rightarrow y = \varphi(x)$$

Die jeweiligen Rundungsfehler können als **Eingangsfehler** des **Rest-Algorithmus** angesehen werden. Sie werden **mit der Kondition der jeweiligen Restabbildung verstärkt**. (vgl. Abschnitt 2.8 C.)

Definition: Ein Algorithmus I heißt **stabiler** als Algorithmus II, wenn der Rundungsfehlereinfluss bei Algorithmus I kleiner ist.

Ein Algorithmus heißt **numerisch stabil** oder **gutartig**, wenn die Rundungsfehler das Ergebnis nicht stärker als die Eingangsfehler beeinflussen.

Frage: Gibt es eine bessere Alternative zur Lösung von $Ax = b$ als den Gaußschen Algorithmus?

(Bemerkung zur Bezeichnung: Bei $Ax = b$ sind A und b die Daten und x ist das Resultat.)

Situation beim Gauß-Algorithmus: Die Transformation der Matrix A auf Dreiecksform U mit Zeilenoperationen erzeugt eine Kette von Zwischenmatrizen:

$$A =: A^{(0)} \rightarrow A^{(1)} \rightarrow A^{(2)} \rightarrow \dots \rightarrow A^{(n-1)} = U.$$

Rundungsfehler der Zwischenschritte (bei der Berechnung von $A^{(j)}$) wirken sich als Störungen ΔA auf das Ergebnis aus, sie werden mit den Faktoren $\text{cond}(A^{(j)})$ verstärkt.

Forderung/gesucht: Alternative zum Gauß-Algorithmus derart, dass

- (1) $\text{cond}(A^{(j)}) = \text{cond}(A^{(j-1)})$ für alle j
- (2) $A^{(n-1)}$ = obere Dreiecksmatrix.

[(1) \Rightarrow Verfahren ist gutartig, d.h. stabil für *alle* Matrizen A .] Zur Ausführung verwende geeignete Matrizen $H^{(j)}$ und $A^{(j)} := H^{(j)} A^{(j-1)}$.

1. Idee: Man nehme **orthogonale** Matrizen H (d.h. $H^T H = I$). Folgerung für $\|\cdot\|_2$:

$$\begin{aligned} \|A\|_2 &= \|H^T H A\|_2 \leq \underbrace{\|H^T\|_2}_{=1} \|H A\|_2 \leq \|H\|_2 \|A\|_2 = \|A\|_2 \\ &\Rightarrow \|H A\|_2 = \|A\|_2. \quad \text{Analog: } \|A^{-1}\|_2 = \|A^{-1} H^T\|_2. \\ &\Rightarrow \text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \|H A\|_2 \underbrace{\|A^{-1} H^T\|_2}_{(H A)^{-1}} = \text{cond}(H A) \end{aligned}$$

$A^{(j)} = H^{(j)} A^{(j-1)} \Rightarrow$ Forderung (1) ist erfüllt.

2. Idee: Definition der Householder-Matrix: $H := I - 2ww^t$ mit beliebigem Einheitsvektor w , d.h. $w^t w = 1$

zur Erinnerung:

ww^t ist Matrix (vom Rang 1)!

$$(AB)^t = B^t A^t$$

Eigenschaften von H :

1. H ist **symmetrisch**, denn: $H^t = I^t - 2(ww^t)^t = I - 2ww^t = H$
2. H ist **orthogonal**, denn: $H^t H = H H = I - 2ww^t - 2ww^t + 4w \underbrace{w^t w}_{=1} w^t = I$
3. $y = Hx$ ist **Spiegelung** an der (Hyper-)Ebene, die senkrecht zu w ist und den Nullpunkt enthält. (\rightarrow Illustration...)
4. $\|Hx\|_2 = \|x\|_2$ (denn $\|Hx\|_2 = \sqrt{(Hx)^t Hx} = \dots$)

Hinweis: $H = I - \frac{uu^t}{\kappa}$ mit $\kappa = \frac{1}{2}u^t u$, u beliebig $\neq 0$ ist Householder-Matrix.

3. Idee: Man wähle w so, dass $Ha \stackrel{!}{=} \gamma e_1 = \begin{pmatrix} \gamma \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

(Anwendung z.B. a die 1. Spalte von $A = A^{(0)}$)

Konstruktion von w (bzw. u): Eigenschaft 4. $\Rightarrow \|Ha\|_2^2 = \gamma^2 = \|a\|_2^2$, also $\gamma = \pm \|a\|_2$. Es folgt

$$\pm \|a\|_2 e_1 \stackrel{!}{=} Ha = a - \frac{1}{\kappa} u(u^t a)$$

$$\Rightarrow u \frac{u^t a}{\kappa} = a \mp \|a\|_2 e_1$$

und also

$$u = \text{const} \cdot (a \pm \|a\|_2 e_1).$$

Da bei u nur die Richtung wichtig ist, setze $\text{const} := 1$. Wähle Vorzeichen so, dass keine Auslöschung auftritt.

Resultat: für gegebenen Vektor $a \neq 0$ setze

$$u := a + \frac{a_1}{|a_1|} \|a\|_2 e_1 = \begin{pmatrix} a_1 + \frac{a_1}{|a_1|} \|a\|_2 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

(falls $a_1 = 0$, setze $\frac{a_1}{|a_1|} = 1$)

$$\kappa := \|a\|_2 (\|a\|_2 + |a_1|)$$

$$H := I - \frac{uu^t}{\kappa}$$

Anwendung auf $Ax = b$: Es sei a die erste Spalte von A , $H^{(1)}$ wie oben konstruiert.

$$\left(A^{(1)} \mid b^{(1)} \right) := H^{(1)} \left(A^{(0)} \mid b^{(0)} \right) = \left(\begin{array}{c|cccc} * & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots \\ \hline & & & & \text{Restmatrix} \end{array} \right)$$

2. Schritt

$$T^{(2)} := \left(\begin{array}{c|ccc} 1 & 0 & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \hline & & & H^{(2)} \end{array} \right)$$

wobei $H^{(2)}$ diejenige Householder-Matrix der Größe $(n - 1) \times (n - 1)$ ist, welche die 1. Spalte der Restmatrix “behandelt” wie oben erklärt. $(A^{(2)}|b^{(2)}) = T^{(2)}(A^{(1)}|b^{(1)})$ usw.

$$T^{(j)} := \left(\begin{array}{c|ccc} 1 & 0 & \cdots & \cdots \\ 0 & 1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \hline & & & H^{(j)} \end{array} \right)$$

Die $T^{(j)}$ sind orthogonal und symmetrisch. Allgemein:

$$\left(A^{(j)} \mid b^{(j)} \right) := T^{(j)} \left(A^{(j-1)} \mid b^{(j-1)} \right)$$

am Ende:

$$\begin{aligned} (R=)U &= T^{(n-1)} \cdot \dots \cdot T^{(2)} H^{(1)} A \\ &\Rightarrow \underbrace{H^{(1)} T^{(2)} \cdot \dots \cdot T^{(n-1)}}_{=: Q \text{ orthogonal und vollbesetzt}} U = A \end{aligned}$$

Diese Zerlegung heißt auch **QR-Zerlegung**, $A = QR$. Wegen $Ax = QRx = b$ gilt

$$Rx = Q^T b = b^{(n-1)}.$$

Lösen von $Rx = b^{(n-1)}$ wie beim Gauß-Algorithmus.

Praxis: Keine Matrizen $H^{(j)}$ speichern, sondern nur die Vektoren u ; sie passen in das untere Dreieck hinein (siehe Pfeil), bis auf die Diagonalelemente.

$$\left(\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \cdot & \cdot \\ & & & & & & \cdot \end{array} \right)$$

Aufwand der QR-Zerlegung: $O(\frac{4}{3}n^3)$ Operationen, also zweimal so teuer wie die Gauß-Elimination.

Kapitel 3 Analysis

3.1 Vollständige Induktion

Häufig in der Mathematik:

Aussagen, die für eine beliebige natürliche Zahl gelten.

Beispiel: $1^3 + 2^3 + \dots + n^3 = \left[\frac{n(n+1)}{2}\right]^2$ ist die Aussage $\mathcal{A}(n)$ für beliebige $n \in \mathbb{N}$.

Beweisverfahren hierfür: **“Vollständige Induktion”**

- Schritt:** Beweis der Aussage für eine kleinste natürliche Zahl, meist für $n = 1$ (“Induktionsanfang” - “Verankerung”). D.h. Nachprüfen von $\mathcal{A}(1)$.
- Schritt:** Nachweis der Gültigkeit für $n + 1$, falls Aussage für n richtig ist. (“Induktionsschluss”) Schluss von n auf $n + 1$.

Folgerung aus beiden Schritten:

gilt für $n = 1$, für $n = 2, n = 3, \dots$ gilt also für alle $n \in \mathbb{N}$.

obiges Beispiel:

1. Schritt: $n = 1$: $1^3 = \left[\frac{1 \cdot (1+1)}{2}\right]^2$: O.K.

2. Schritt:

$$\begin{aligned} & \{1^3 + 2^3 + \dots + n^3\} + (n+1)^3 \\ & \text{durch Benutzung der Induktionsannahme } \mathcal{A}(n) \\ & = \left\{ \left[\frac{n(n+1)}{2}\right]^2 \right\} + (n+1)^3 \\ & = \left(\frac{n+1}{2}\right)^2 \cdot \underbrace{\left[n^2 + 4(n+1)\right]}_{(n+2)^2} = \left[\frac{(n+1)(n+2)}{2}\right]^2 \end{aligned}$$

das ist die Aussage $\mathcal{A}(n+1)$

Bemerkungen:

- kein konstruktiver Beweis
- häufiger Fehler: in der Aussage statt n einfach $n + 1$ setzen; das ist kein Beweis.

Weitere Formeln, die sich z.B. mit vollständiger Induktion beweisen lassen:

$$1 + 2 + 3 + \dots + n = \frac{1}{2}n(n+1)$$

$$1 + x + x^2 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x} \quad (\text{abgebrochene geometrische Reihe})$$

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1} \quad \text{für die Binomialkoeffizienten}$$

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \quad \text{mit } \binom{n}{0} := 1; \quad \text{Fakultät } n! := 1 \cdot 2 \cdot \dots \cdot n$$

3.2 Komplexe Zahlen

In Abschnitt 1.4 hatten wir die Erweiterung der Zahlenbereiche

$$\mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q} \rightarrow \mathbb{R}$$

(natürliche Zahlen \rightarrow ganze Zahlen \rightarrow rationale Zahlen \rightarrow reelle Zahlen) betrachtet.

Erneute Erweiterung des Zahlenbereichs:

Die Gleichung $x^2 + 1 = 0$ (z.B.) besitzt im Bereich der reellen Zahlen keine Lösung. Wir führen neue, künstliche Zahlen ein, welche diese Gleichung (per definitionem) lösen sollen. Zunächst ist formal/symbolisch

$$x = \pm\sqrt{-1} \quad (\text{sinnlos im Bereich der reellen Zahlen})$$

neues Symbol

$$x = \pm i$$

dann ist

$$i^2 + 1 = 0$$

$$(-i)^2 + 1 = 0,$$

woraus

$$i^2 = (-i)^2 = -1$$

$$\text{symbolisch: } i = \sqrt{-1}$$

Definition: Sind a und b zwei beliebige reelle Zahlen, so definiert das Zahlenpaar (a, b) die **komplexe Zahl** $z = a + ib$

a heißt Realteil von z , $a = \text{Re}\{z\}$

b heißt Imaginärteil von z , $b = \text{Im}\{z\}$, $b \in \mathbb{R}$

i heißt imaginäre Einheit. Es gilt:

$$i^1 = i$$

$$i^2 = -1$$

$$i^3 = ii^2 = -i$$

$$i^4 = i^2i^2 = 1$$

$$i^5 = i$$

$$i^6 = -1$$

$$i^7 = -i$$

$$i^8 = 1$$

Rechenregeln für komplexe Zahlen $z_1 = a_1 + ib_1$ und $z_2 = a_2 + ib_2$:

Addition: $z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)$

Subtraktion: $z_1 - z_2 = (a_1 - a_2) + i(b_1 - b_2)$

Multiplikation:

$$\begin{aligned} z_1 z_2 &= a_1 a_2 + a_1 i b_2 + i b_1 a_2 + i^2 b_1 b_2 \\ &= (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1) \end{aligned}$$

Division:

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{a_1 + ib_1}{a_2 + ib_2} = \frac{(a_1 + ib_1)(a_2 - ib_2)}{(a_2 + ib_2)(a_2 - ib_2)} \\ &= \frac{a_1 a_2 + b_1 b_2 + i(a_2 b_1 - a_1 b_2)}{a_2^2 + b_2^2}; \quad a_2^2 + b_2^2 > 0 \end{aligned}$$

Falls in $z = a + ib$ der Imaginärteil $b = 0$ ist, dann $z = a$ (Reduktion der komplexen Zahlen auf reelle Zahlen). D.h. die reellen Zahlen sind in den komplexen Zahlen enthalten.

Falls $a = 0$, dann $z = ib$ (Reduktion auf rein imaginäre Zahlen).

Speziell ist i eine komplexe Zahl ($a = 0, b = 1$).

Definition Für $z = a + ib$ heißt $\bar{z} := a - ib$ **konjugiert** zu z . Es gilt

$$z\bar{z} = a^2 + b^2 \geq 0.$$

($z\bar{z}$ ist also reell.)

Gleichung zwischen komplexen Zahlen:

$$z_1 = z_2$$

ist äquivalent zu **zwei reellen** Gleichungen:

$$a_1 + ib_1 = a_2 + ib_2 \Leftrightarrow \begin{cases} a_1 = a_2 & \text{und} \\ b_1 = b_2 \end{cases}$$

Insbesondere:

$$z = a + ib = 0 \text{ (komplexe Null)} \Leftrightarrow a = 0 \text{ und } b = 0$$

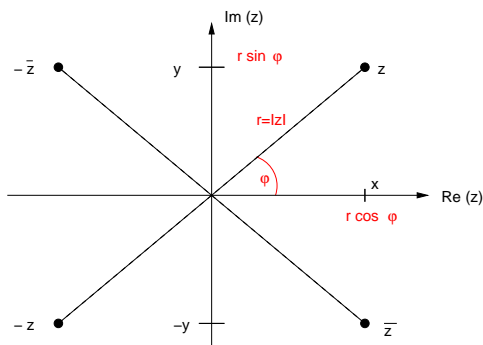
Symbol für Menge der komplexen Zahlen: $\mathbb{C} := \{z = a + ib \mid a, b \in \mathbb{R}\}$

Geometrische Darstellung der komplexen Zahlen

$$z = x + iy, \quad x \in \mathbb{R}, y \in \mathbb{R}$$

$$z \leftrightarrow \text{reelles Zahlenpaar } (x, y) \leftrightarrow \text{Punkt in Ebene}$$

Definition: (Gaußsche) **Zahlenebene** der komplexen Zahlen: eindeutige Zuordnung zwischen allen komplexen Zahlen z und allen Punkten (x, y) der Ebene durch $x = \operatorname{Re}(z)$, $y = \operatorname{Im}(z)$.



Folgerung:

- 1.) $\bar{z} = x - iy$ entsteht aus z durch Spiegelung an x -Achse.
- 2.) $-z$ durch Spiegelung am Nullpunkt.
- 3.) komplexe Zahlen auch mit Polarkoordinaten darstellbar:

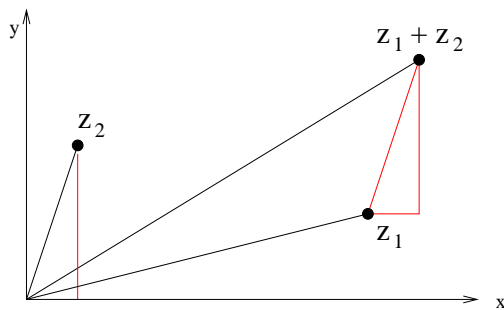
$$z = r(\cos \varphi + i \sin \varphi)$$

Bezeichnung: $|z| := r = \sqrt{x^2 + y^2}$ ist der **Betrag** einer komplexen Zahl.

$$\arg(z) = \operatorname{arc}(z) = \varphi = \arccos \frac{x}{r} = \arcsin \frac{y}{r} = \arctan \frac{y}{x}$$

(je nach Vorzeichen; vgl. Abschnitt 1.1C oder Formelsammlung) ist das **Argument** oder der **Arcus** von z .

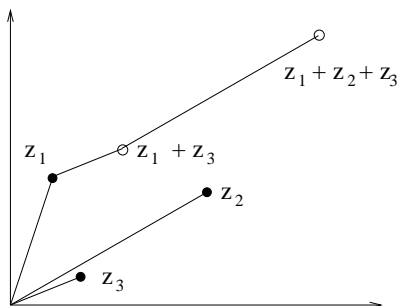
Deutung der Rechenregeln: Addition: Parallelogramm: $z_1 + z_2$ durch **Parallelverschiebung** von z_2 oder z_1



Subtraktion: $z_1 - z_2 = z_1 + (-z_2)$

Vektoren lassen sich **parallel verschieben**. Vereinbarung: Alle durch Parallelverschiebung ineinander überführbare Vektoren (d.h. gleiche Länge und Richtung) sind identisch. Der **“Ortsvektor”** ist der spezielle Vektor, dessen Fußpunkt in 0 ist.

Also: Die Addition von komplexen Zahlen ist geometrisch als Hintereinandersetzen von Vektoren deutbar.



(geometrische Darstellung der) **Multiplikation:**

$$z_1 = r_1(\cos \varphi_1 + i \sin \varphi_1), \quad z_2 = r_2(\cos \varphi_2 + i \sin \varphi_2)$$

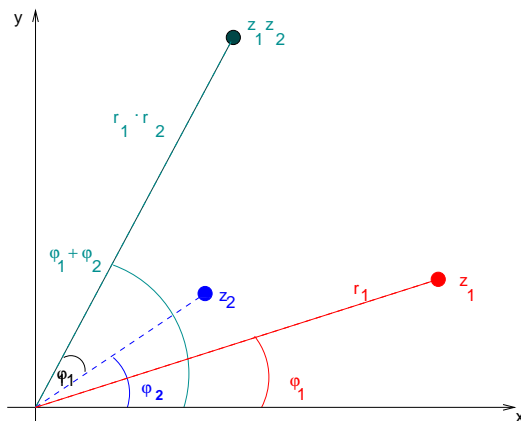
$$\begin{aligned} z_1 z_2 &= r_1(\cos \varphi_1 + i \sin \varphi_1) \cdot r_2(\cos \varphi_2 + i \sin \varphi_2) \\ &= r_1 r_2 \cdot [\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 + i(\sin \varphi_1 \cos \varphi_2 + \sin \varphi_2 \cos \varphi_1)] \\ &= r_1 r_2 \cdot [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)] \end{aligned}$$

Es folgt:

$$|z_1 z_2| = r_1 r_2 = |z_1| |z_2|$$

$$\arg(z_1 z_2) = \varphi_1 + \varphi_2 = \arg(z_1) + \arg(z_2)$$

D.h. bei der Multiplikation werden die Beträge multipliziert und die Winkel addiert:



Hinweis: Insbesondere bei grafischen Darstellungen macht ein Winkel nur Sinn im Bereich $0 \leq \varphi \leq 2\pi$, darüberhinaus fängt der Winkel wieder bei 0 an. Deswegen wird bei $\varphi_1 + \varphi_2 > 2\pi$ die Zahl 2π subtrahiert.

Spezialfälle: $|z^2| = |z|^2$, $\arg(z^2) = 2 \arg(z)$; $|z^m| = |z|^m$, $\arg(z^m) = m \arg(z)$

zur Division: Für Quotienten ergibt sich:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} \cdot [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)], \quad (r_2 \neq 0)$$

also:

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}, \quad \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2)$$

Anwendung/Folgerungen für $z = i$: wegen $|i| = 1$, $\arg(i) = \frac{\pi}{2}$ vermittelt iz eine **Drehung** um $+\frac{\pi}{2}$ bzw. 90° , denn $iz = -y + ix$.

Dreiecksungleichung gilt auch in \mathbb{C} :

$$\left| |z_1| - |z_2| \right| \leq |z_1 \pm z_2| \leq |z_1| + |z_2|$$

3.3 Polynome und Rationale Funktionen

A.) Polynome (im Reellen)

Definition: Polynom vom Grad m ist die Funktion

$$a_0 + a_1x + a_2x^2 + \dots + a_mx^m = \sum_{k=0}^m a_k x^k, \quad \text{mit } a_m \neq 0$$

Polynome sind von größter Wichtigkeit. In Abschnitt 2.7 traten sie auf als charakteristisches Polynom von Matrizen. Wichtig sind Polynome auch zur Approximation von Funktionen, z.B. $\cos x \approx 1 - \frac{x^2}{2}$ für $x \approx 0$ (vgl. u.a. Abschnitt 3.11).

Bezeichnung: $P_m(x) := \sum_{k=0}^m a_k x^k$. Die a_k heißen "Koeffizienten".

Gleichheit von Polynomen:

$$\sum_{k=0}^m a_k x^k = \sum_{k=0}^m b_k x^k \quad \text{für alle } x \quad \Leftrightarrow \quad a_k = b_k \quad \text{für alle } k$$

(Beweis: $x = 0 \Rightarrow a_0 = b_0$. Dividiere durch x . x können beliebig klein werden $\Rightarrow a_1 = b_1 \Rightarrow \dots$)

Folgerung: Koeffizientenvergleich

z.B.: $x^3 + ax^2 + bx + c =$

[Suche Beziehung zwischen (angenommenen) Nullstellen x_1, x_2, x_3 . Falls solche Nullstellen existieren (s.u.), ist dies von der Form]

$$= (x - x_1)(x - x_2)(x - x_3)$$

Ausmultiplizieren $\Rightarrow \dots = x^3 - (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_2x_3 + x_3x_1)x - x_1x_2x_3$

Koeffizientenvergleich \Rightarrow

$$\left. \begin{array}{l} a = -x_1 - x_2 - x_3 \\ b = x_1x_2 + x_2x_3 + x_3x_1 \\ c = -x_1x_2x_3 \end{array} \right\} \text{ Satz von Vieta, hier für } m = 3$$

Auswertung von Polynomen mit "Horner-Schema"

$$P_m(x) = (\dots((a_m \cdot x + a_{m-1}) \cdot x + a_{m-2}) \cdot x + \dots + a_1) \cdot x + a_0 .$$

(m Multiplikationen, m Additionen)

B.) Nullstellen von Polynomen

$P_m(x) = 0$ heißt eine "algebraische Gleichung", und Nullstellen heißen auch "Wurzeln".

Annahme: $P_m(x) = 0$ für $x = x_1$, also: $P_m(x_1) = 0$

$$\begin{aligned} \Rightarrow P_m(x) &= P_m(x) - P_m(x_1) = a_0 + a_1x + \dots + a_mx^m - a_0 - a_1x_1 - \dots - a_mx_1^m \\ &= a_1(x - x_1) + a_2(x^2 - x_1^2) + \dots + a_m(x^m - x_1^m) \\ &= a_1(x - x_1) + a_2(x - x_1)(x + x_1) + a_3(x - x_1)(x^2 + xx_1 + x_1^2) + \\ &\quad + \dots + a_m(x - x_1)(x^{m-1} + x^{m-2}x_1 + \dots) \\ &= (x - x_1) \cdot \{a_1 + a_2(x + x_1) + a_3(x^2 + xx_1 + x_1^2) + \dots\} \\ &= (x - x_1) \{c_0 + c_1x + c_2x^2 + \dots + c_{m-1}x^{m-1}\} \text{ mit } c_{m-1} = a_m \neq 0 \end{aligned}$$

D.h. das Polynom $P_m(x)$ kann ohne Rest durch $x - x_k$ dividiert werden, wenn x_k Nullstelle ist.

Definition: Der Faktor $x - x_k$ heißt **Linearfaktor**.

Die Berechnung von $P_{m-1}(x) = \frac{P_m(x)}{x-x_1}$ heißt **Abdividieren**.

Falls auch P_{m-1} eine Nullstelle hat, für $x = x_2$, dann kann der Linearfaktor $(x - x_2)$ abgespalten werden, und es entsteht $P_{m-2}(x)$.

u.s.w. ...

Dieser Prozess kann fortgeführt werden, solange das jeweils entstehende Polynom niederen Grades eine Nullstelle hat, längstens also bis

$$P_m(x) = (x - x_1)(x - x_2) \cdot \dots \cdot (x - x_m) \cdot a_m$$

Damit haben wir gezeigt: Ein Polynom m -ten Grades kann höchstens m Nullstellen haben. Also gilt:

Satz: Hat ein Polynom $P_m(x)$ mehr als m Nullstellen, so muss $P_m(x) \equiv 0$ gelten, also $a_0 = a_1 = \dots = a_m = 0$.

x_1 heißt **k -fache Nullstelle**, wenn $P_n(x)$ ohne Rest durch $(x - x_1)^k$ teilbar ist, nicht aber durch $(x - x_1)^{k+1}$.

C.) Zur Existenz von Lösungen algebraischer Gleichungen

Über den reellen Zahlen, d.h. $x \in \mathbb{R}$, ist eine Zerlegung von $P_m(x)$ in Linearfaktoren nicht immer möglich.

Beispiel: $1 + x^2 = 0$ hat *keine* Lösung für $x \in \mathbb{R}$.

Aber: Über den komplexen Zahlen, d.h. $x \in \mathbb{C}$, ist jede algebraische Gleichung lösbar!

Fundamentalsatz der Algebra: Ein Polynom vom Grad m hat genau m Wurzeln $\in \mathbb{C}$, wobei k -fache Nullstellen k mal gezählt werden.

Beispiel:

$$\begin{aligned}x^4 + 1 = 0 & \text{ keine Lösung in } \mathbb{R} . \\x^4 + 1 = 0 & \text{ in } \mathbb{C} :\end{aligned}$$

Erste Überlegung: $(z^2)^2 = -1 \Rightarrow z^2 = \pm i$ ergibt 4 Lösungen mit \sqrt{i} .

Was ist \sqrt{i} ?

andere Überlegung mit dem Argument "Drehung":

$$zzzz = -1$$

Die Zahl $z_1 := \frac{1}{\sqrt{2}}(1 + i)$ mit $|z_1| = 1$, $\varphi = \text{arc}(z_1) = \pi/4$ ergibt $z_1^2 = i$, $z_1^3 = \frac{1}{\sqrt{2}}(-1 + i)$, $z_1^4 = -1$ mit 2φ , 3φ , 4φ . Für $4\varphi = \pi$ wird -1 erreicht. Analog:

$$\begin{aligned}z_2 &:= \frac{1}{\sqrt{2}}(-1 + i) & \text{ mit } \text{arc}(z_2) &= \frac{3}{4}\pi \\z_3 &:= \frac{1}{\sqrt{2}}(-1 - i) & \text{ mit } \text{arc}(z_3) &= \frac{5}{4}\pi \\z_4 &:= \frac{1}{\sqrt{2}}(1 - i) & \text{ mit } \text{arc}(z_4) &= \frac{7}{4}\pi\end{aligned}$$

Illustration: jeweils Drehungen; alle 4 Nullstellen liegen auf dem Einheitskreis;

$$\begin{aligned}\text{arc}(z_2^4) &= 3\pi \text{ entspricht } \pi \\ \text{arc}(z_3^4) &= 5\pi \text{ entspricht } \pi \\ \text{arc}(z_4^4) &= 7\pi \text{ entspricht } \pi\end{aligned}$$

Für Winkel $> 2\pi$ werden ganzzahlige Vielfache von 2π abgezogen.

Satz: Wenn eine algebraische Gleichung reelle Koeffizienten hat und $z_1 = x_1 + iy_1$ Wurzel ist, dann ist auch $\bar{z}_1 = x_1 - iy_1$ Wurzel.

Anwendung: Eigenwerte von reellen Matrizen als Nullstellen des charakteristischen Polynoms. Die Linearfaktoren $(z - z_1)$ und $(z - \bar{z}_1)$ können jeweils zu einem quadratischen Faktor mit reellen Koeffizienten zusammengefasst werden.

obiges Beispiel: $z_4 = \bar{z}_1$, $z_3 = \bar{z}_2$, und

$$z^4 + 1 = (z - z_1)(z - \bar{z}_1)(z - z_2)(z - \bar{z}_2)$$

Produkt der ersten beiden Linearfaktoren:

$$\begin{aligned}& \left(z - \frac{1}{\sqrt{2}}(1 + i)\right)\left(z - \frac{1}{\sqrt{2}}(1 - i)\right) \\ &= \left(z - \frac{1}{\sqrt{2}}\right)^2 - \frac{i^2}{2} \\ &= \left(z^2 - \frac{2}{\sqrt{2}}z + \frac{1}{2} + \frac{1}{2}\right) \\ &= (z^2 - \sqrt{2}z + 1)\end{aligned}$$

insgesamt

$$z^4 + 1 = (z^2 - \sqrt{2}z + 1)(z^2 + \sqrt{2}z + 1)$$

Zusammenfassend: Die folgende *reelle* Darstellung ist möglich:

Jedes Polynom mit reellen Koeffizienten lässt sich in reelle lineare und reelle quadratische Faktoren (bzw. ihre Potenzen) zerlegen. Dabei haben die quadratischen Faktoren keine reelle Nullstelle.

D.) Rationale Funktionen

Quotienten zweier Polynome

$$f(x) = \frac{P_m(x)}{Q_n(x)} = \frac{a_m x^m + \dots + a_1 x + a_0}{b_n x^n + \dots + b_1 x + b_0}, \quad a_m \neq 0, b_n \neq 0$$

heißen **rationale Funktionen**.

Polynome ergeben sich als Spezialfall für $n = 0$ und heißen auch **“ganze” rationale Funktionen**. $n > 0$: **“gebrochen” rationale Funktionen**.

Ohne Einschränkung: P_m und Q_n seien teilerfremd. Anderenfalls Zähler und Nenner durch gemeinsames Teilerpolynom kürzen.

Methode: Euklidischer Algorithmus (\rightarrow Abschnitt 3.4)

Nach Kürzen: keine gemeinsamen Nullstellen mehr. Die Nullstellen des Zählers sind die Nullstellen von $f(x)$. Die Nullstellen des Nenners sind die Unendlichkeitsstellen. (“vertikale Asymptoten” oder “Polstellen”)

asymptotisches Verhalten (für $x \rightarrow +\infty$ und $x \rightarrow -\infty$)

$$f(x) = \frac{P_m(x)}{Q_n(x)} = \frac{a_m x^m}{b_n x^n} \cdot \frac{(1 + \frac{a_{m-1}}{a_m} \frac{1}{x} + \dots + \frac{a_0}{a_m} \frac{1}{x^m})}{1 + \frac{b_{n-1}}{b_n} \frac{1}{x} + \dots + \frac{b_0}{b_n} \frac{1}{x^n}}$$

Der zweite Faktor geht 1 für $|x| \rightarrow \infty$. Der erste Faktor beschreibt das asymptotische Verhalten.

Fälle:

$$\begin{cases} m < n & \Rightarrow f(x) \rightarrow 0 \\ m = n & \Rightarrow f(x) \rightarrow \frac{a_m}{b_n} \\ m > n & \Rightarrow f(x) \rightarrow \infty \end{cases}$$

Weitere Untersuchungen zu $m > n$:

Durch **Ausdividieren** ist stets eine Zerlegung in Summe von einer **ganzzrationalen Funktion** plus einer **gebrochenen rationalen Funktion** möglich.

Beispiel:

$$\frac{2x^3 + 5x^2 + x - 7}{x^2 + x - 2} = 2x + 3 + \frac{2x - 1}{x^2 + x - 2}$$

Durch $2x + 3$ ist eine *schräge* Asymptote definiert. Der Bruch auf der rechten Seite hat als Nenner $(x - 1)(x + 2)$ und damit gibt es senkrechte Asymptoten bei $x_1 = 1$ und $x_2 = -2$. Illustration siehe Vorlesung.

3.4 Der Euklidische Algorithmus; Kettenoperationen

In diesem Abschnitt: Fast alle Zahlen seien $\in \mathbb{N}$, z.B. $a, b, c, s \in \mathbb{N}$.

a und b heißen **Teiler** von c , wenn $c = ab$

weitere Anwendung: Polynomdivision.

Kettendivision: $a > b$, teile a durch b ...

$$\begin{array}{rcl} a & = & bs_0 + r_1 & \text{(mit } r_1 < b) \\ b & = & r_1s_1 + r_2 & \text{(mit } r_2 < r_1) \\ r_1 & = & r_2s_2 + r_3 & \\ & & \vdots & \vdots \\ r_{n-3} & = & r_{n-2}s_{n-2} + r_{n-1} & \\ r_{n-2} & = & r_{n-1}s_{n-1} + r_n & \\ r_{n-1} & = & r_ns_n \quad (+0) & \end{array}$$

Rückwärts-Folgerung:

$$\begin{aligned} r_n \text{ teilt } r_{n-1} &\Rightarrow r_n \text{ teilt } r_{n-2} \Rightarrow \dots \\ &\Rightarrow r_n \text{ teilt } b \text{ und } r_n \text{ teilt } a \end{aligned}$$

r_n ist also gemeinsamer Teiler von a und b .

Frage: Gibt es einen größeren Teiler?

Vorwärts-Folgerung:

$$\begin{aligned} t \text{ teile } a \text{ und } b &\Rightarrow t \text{ teilt } r_1 \Rightarrow t \text{ teilt } r_2 \Rightarrow \\ \dots &\Rightarrow t \text{ teilt } r_n \Rightarrow t \leq r_n \end{aligned}$$

Also: Die mit obiger Kettendivision berechnete Zahl r_n ist der **größte gemeinsame Teiler** von a und b .

Euklidischer Algorithmus für natürliche Zahlen (iterierte Division, Kettendivision)

$$r_0 := a, r_1 := b,$$

für $j = 1, 2, \dots$:

$$r_{j-1} = r_j \cdot s_j + r_{j+1} \quad \text{mit } 0 \leq r_{j+1} < r_j$$

Der Algorithmus bricht ab bei k mit $r_{k+1} = 0$. Resultat:

$$r_k = \text{ggT}(a, b)$$

Wegen $\text{ggT}(a, b) \cdot \text{kgV}(a, b) = a \cdot b$ gilt:

$$\text{kgV}(a, b) = \frac{a \cdot b}{r_k}$$

(vgl. Vorlesung über Informatik)

Obige Kettendivision ist auf Polynome anwendbar.

Beispiel:

$$f(x) = \frac{P_3(x)}{Q_3(x)} = \frac{x^3 - x^2 - x + 1}{x^3 + 2x^2 - x - 2}$$

Was ist der größte gemeinsame Teiler?

Bei der Polynomdivision kommt es auf Zahlenfaktoren nicht an.

$$P_3 : Q_3 = 1 \text{ mit Rest } r_1 = -3x^2 + 3 = -3(x^2 - 1)$$

Also:

$$P_3 = Q_3 \cdot 1 - 3x^2 + 3$$

Der Rest kann durch Multiplikation mit beliebiger Zahl vereinfacht werden zu \tilde{r}_1 . (hier: $\tilde{r}_1 = x^2 - 1$)

$$Q_3 = (x^2 - 1) \cdot (x + 2) + 0$$

Folgerung: $x^2 - 1$ ist größtes gemeinsames Teilerpolynom.

Kürzen \Rightarrow

$$\frac{P_3(x)}{Q_3(x)} = \frac{x - 1}{x + 2}$$

Es bleibt eine senkrechte Asymptote für $x = -2$. In der ursprünglichen Version hatte der Nenner die weiteren Nullstellen $x = +1$ und $x = -1$. Da diese beiden auch Nullstellen des Zählers waren, ist $f(x)$ durch das Kürzen an den Stellen $x = \pm 1$ "repariert".

3.5 Funktionen

A.) Der Funktionsbegriff

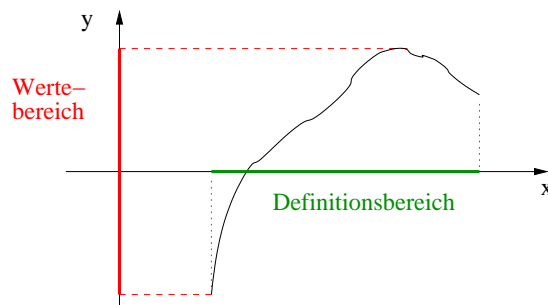
Wird jeder Zahl $x \in \mathbb{R}$ eines vorgegebenen Bereiches eindeutig eine Zahl $y \in \mathbb{R}$ zugeordnet, so heißt y eine **Funktion** der "Veränderlichen" x . Schreibweise

$$y = f(x) \quad (\text{oder } y = y(x)),$$

dabei: y : "abhängige" Variable, x : "unabhängige" Variable, "Argument der Funktion"

Vorsicht: auch andere Bezeichnungen! z.B. $\psi = \sin \varphi$ oder $r = \cos t$.

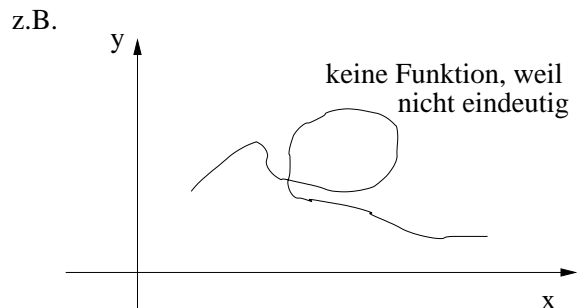
Die Menge derjenigen x , für die $f(x)$ definiert ist, heißt **Definitionsbereich**. Die Menge der zugehörigen Werte $f(x)$, für x im Definitionsbereich, heißt **Wertebereich**. Fasst man x und y als Koordinaten auf, dann ist eine graphische Darstellung möglich. Jedem Paar x, y entspricht ein Punkt in der (x, y) -Ebene. Für alle x im Definitionsbereich beschreibt der Punkt $(x, y(x))$ typischerweise ein Kurvenstück. Die Punktmenge $\{(x, f(x)) | x \in \text{Definitionsbereich}\}$ heißt **Graph** der Funktion $f(x)$.



Die Definitionsbereiche sind meist

"offene" Intervalle	$a < x < b$	oder
"abgeschlossene" Intervalle	$a \leq x \leq b$	oder
"halboffene" Intervalle	$a \leq x < b$	oder $a < x \leq b$
"unendliche" Intervalle	$-\infty < x < \infty$	oder
"einseitig unendliche" Intervalle	z.B. $a \leq x < \infty$	

Bemerkung: Nicht jede Kurve entspricht einer Funktion!



Beispiele:

$$f(x) = \sin x \quad \text{für } -\infty < x < \infty$$

$$f(x) = +\sqrt{x} \quad \text{für } 0 \leq x < \infty$$

$$f(x) = +\sqrt{r^2 - x^2} \quad \text{für } -r \leq x \leq r$$

$$f(x) = \begin{cases} 1 & \text{für } x > 0 \\ -1 & \text{für } x < 0 \end{cases}$$

$$f(x) = x \quad \text{für } x \in \mathbb{N} \text{ (kein Kurvenstück, sondern diskrete Punkte)}$$

$$f(x) = \log x \quad \text{für } x > 0$$

Hintereinanderausführung oder “Verkettung” von Funktionen: Falls der Wertebereich von f im Definitionsbereich von g enthalten ist, dann ist

$$h(x) = g(f(x))$$

eine Funktion. Beispiel: $h(x) := \sqrt{1 + x^2}$, mit $f(x) = 1 + x^2$ und $g(y) = \sqrt{y}$.

B.) Umkehrfunktionen, Monotonie

Wenn auch jedem Funktionswert $y = f(x)$ *eindeutig* ein Urbild x entspricht, dann kann x als Funktion von y betrachtet werden:

$$x = g(y).$$

Dann heißt $x = g(y)$ die **Umkehrfunktion** von $y = f(x)$.

Bezeichnet man bei der Umkehrfunktion $x = g(y)$ die unabhängige Veränderliche mit x und die abhängige Veränderliche mit y , dann spricht man von der inversen Funktion $y = g(x) = f^{-1}(x)$. (Vorsicht: Die symbolische Schreibweise f^{-1} bedeutet hier *nicht* Division, sondern Inversion einer Funktion.)

Umkehrfunktionen gibt es also, wenn in *beiden* Richtungen $x \leftrightarrow y$ Eindeutigkeit vorliegt.

Zuordnung $x \rightarrow y$ eindeutig \Leftrightarrow jede Parallele zur y -Achse wird von $f(x)$ genau einmal geschnitten.

Zuordnung $y \rightarrow x$ eindeutig \Leftrightarrow jede Parallele zur x -Achse wird von $g(y)$ genau einmal geschnitten.

Hinreichend für die Existenz einer Umkehrfunktion ist strenge **Monotonie** von f .

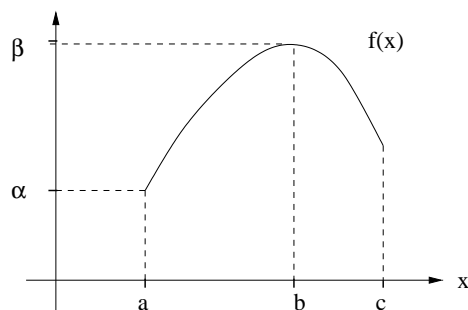
Definition:

$f(x)$ (streng) **monoton wachsend** auf einem Intervall, wenn $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$ für alle x_1, x_2 in dem Intervall.

$f(x)$ **monoton fallend** auf Intervall, wenn $x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$

$f(x)$ **monoton** auf einem Intervall, wenn f auf dem gesamten Intervall entweder monoton wachsend oder monoton fallend ist (**ausschließlich**).

Beispiel



auf

$a \leq x \leq b$: f monoton wachsend

$b \leq x \leq c$: f monoton fallend

$a \leq x \leq c$: f ist **nicht** monoton

$\alpha \leq y \leq \beta$: keine eindeutige Zuordnung $y \rightarrow x$

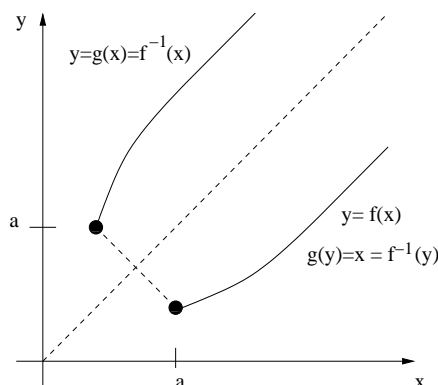
also: Für $f(x)$ auf $a \leq x \leq c$ gibt es keine Umkehrfunktion! Umkehrfunktionen existieren aber stückweise, d.h. für Teilintervalle. Die genaue Angabe der jeweiligen (Teil-)Intervalle ist also bei Monotonie-Angaben wichtig.

Definition: Eine Funktion $f(x)$, für die auch die Zuordnung $y \rightarrow x$ eindeutig ist, heißt **eindeutig** oder **injektiv**.

Zusammen: Die folgenden **Kriterien** für eine Funktion $f(x)$ sind äquivalent zur Injektivität:

- (a) Die Gleichung $f(x) = y$ hat für alle y höchstens eine Lösung.
- (b) für alle $x_1, x_2 \in$ Definitionsbereich: $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$
- (c) für alle $x_1, x_2 \in$ Definitionsbereich: $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$
- (d) f ist über dem Definitionsbereich umkehrbar.

Geometrische **Konstruktion:** Die Kurven $y = f(x)$ und $y = f^{-1}(x)$ gehen durch Spiegelung an der Winkelhalbierenden auseinander hervor.



Nach Definition der Umkehrfunktion gilt die Identität $f(f^{-1}(y)) = y$ für alle $y \in$ Wertebereich, und $f^{-1}(f(x)) = x$ für alle $x \in$ Definitionsbereich.

Implizite Darstellung von Funktionen:

Bisher: nach y **aufgelöste** Gleichungen $y = f(x)$.

Beispiele: $y = \sin x$ oder $y = \frac{1}{1-x}$.

Falls nach x aufgelöst: $x = g(y)$ oder $x = f^{-1}(y)$.

Derartige Darstellungen von Funktionen nennt man **explizit**.

Alternative: Funktionen können **implizit** definiert werden durch eine Gleichung $F(x, y) = 0$. Durch Auflösen der Gleichung $F(x, y) = 0$ nach x oder nach y können *in vielen Fällen* explizite Funktionen hergeleitet werden. Umgekehrt kann eine explizite Darstellung $y = f(x)$ immer als $F(x, y) = 0$ geschrieben werden: $F = y - f(x)$.

Beispiel: $F(x, y) = x^2 + y^2 - 1 = 0$

Auflösen z.B. nach y liefert *zwei* Funktionen:

$$y_1(x) := +\sqrt{1-x^2}$$

$$y_2(x) := -\sqrt{1-x^2}$$

Stückweise existieren Umkehrfunktionen:

$$x_1(y) := +\sqrt{1-y^2} \quad , \quad x_2(y) := -\sqrt{1-y^2}$$

(Abbildungen vgl. Vorlesung)

Beachte: Eine Gleichung $F(x, y) = 0$ definiert implizit entweder

- (1) *keine* Funktion (Beispiel $F(x, y) = x^2 + y^2 + 1 = 0$), oder
- (2) eine *eindeutige* Funktion (Beispiel $F(x, y) = y - e^x = 0$), oder
- (3) *mehrere* Funktionen (Beispiel Kreis).

Parameterdarstellung

Eine weitere Alternative zur Definition von Kurven (außer mit expliziten oder impliziten Funktionen) ist die **Parameterdarstellung**: Dabei werden zwei Funktionen $x(t)$ und $y(t)$ als Komponenten zu einem Vektor

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

zusammengefasst. Die den beiden Komponenten gemeinsame unabhängige Variable t heißt hier **Parameter**.

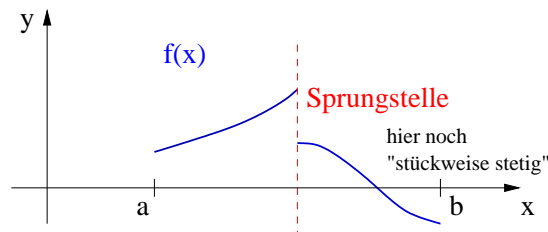
Beispiel: $\begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix}$ beschreibt für wachsende t den Weg eines Punktes in der (x, y) -Ebene, der einen Kreis mit Radius r entlangwandert. Für $0 \leq t \leq 2\pi$ wird dieser Kreis genau einmal durchlaufen. (vgl. Polarkoordinaten in Abschnitt 1.1B.)

C.) Stetigkeit

Anschaulich: Eine Funktion ist “**stetig**”, wenn man sie zeichnen kann, ohne den Stift abzusetzen.

Beispiele für unstetige Funktionen:

1.)

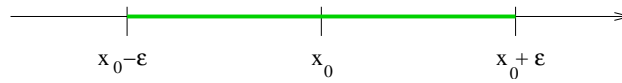


2.) Rationale Funktionen mit Nullstellen des Nenners: Pol-Verhalten, vgl. Abschnitt 3.3D.

3.)

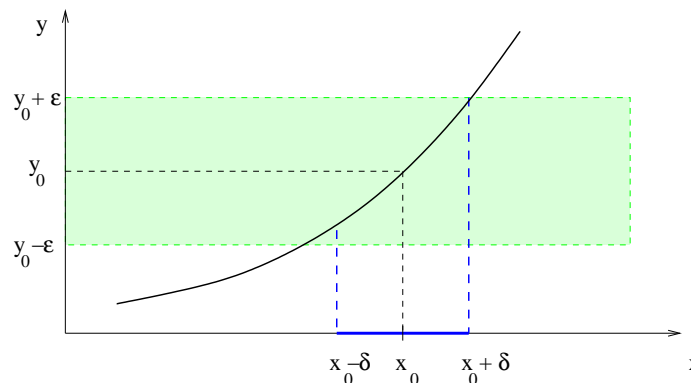
$$y = \begin{cases} 0 & \text{für rationale } x \\ 1 & \text{für irrationale } x \end{cases}$$

Definition: “Umgebung” Eine ϵ -Umgebung von x_0 ist die Menge aller $x \in \mathbb{R}$ mit $|x - x_0| < \epsilon$, für $\epsilon > 0$.



meist: *kleine* Umgebungen, d.h. ϵ ist ein kleiner “Radius”.

Stetigkeit: Die Funktionswerte $f(x)$ in einer Umgebung von x_0 ändern sich beliebig wenig, wenn die Umgebung genügend klein ist.



Für jeden beliebig schmalen Streifen um $y_0 = f(x_0)$ gibt es eine Umgebung um x_0 , so dass der Graph von $f(x)$ für das Intervall $x_0 - \delta \leq x \leq x_0 + \delta$ in dem horizontalen Streifen liegt.

formale Definition: $f(x)$ ist **stetig an der Stelle** x_0 , wenn man zu jedem $\epsilon > 0$ eine Zahl $\delta > 0$ angeben kann, so dass

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon .$$

Und: f ist stetig in einem Intervall, wenn $f(x)$ für $\forall x \in \text{Intervall}$ stetig ist.

Bemerkung: δ hängt i.A. ab von ϵ und von x_0 : $\delta = \delta(\epsilon, x_0)$

Bedeutung stetiger Funktionen: Für sie gelten wichtige Eigenschaften!

1.) **Zwischenwertsatz:** Ist $f(x)$ stetig für $a \leq x \leq b$, und $f(a) < c < f(b)$, so gibt es ein x^* in $a < x^* < b$ mit $f(x^*) = c$.

(analog $f(a) > c > f(b)$)

Weitere **Eigenschaften** stetiger Funktionen:

Für jede auf einem **abgeschlossenen** Intervall $a \leq x \leq b$ stetige Funktion $f(x)$ gilt:

- 2.) f ist auf dem Intervall **beschränkt**, d.h. es gibt eine Schranke K , so dass $|f(x)| < K$ für $a \leq x \leq b$.
- 3.) f nimmt auf $a \leq x \leq b$ ein Maximum und ein Minimum an, d.h.

$$\exists x_1, x_2 \text{ in } a \leq x \leq b \text{ so dass } f(x_1) \leq f(x) \leq f(x_2) \text{ für alle } a \leq x \leq b.$$

Zur Wichtigkeit der Abgeschlossenheit: $f(x) = \frac{1}{x}$ auf dem offenen Intervall $0 < x < \infty$ ist nicht beschränkt.

Für allgemeine Intervalle gilt:

Sind f und g stetig, dann sind auch $f \pm g$, $\alpha \cdot f$, $f \cdot g$ stetig, sowie $\frac{f}{g}$ in jedem Teilintervall von x , auf dem $g(x) \neq 0$, und ebenso die Hintereinanderausführung $h(x) := f(g(x))$ stetig.

Beispiel: $p(x) = a_n x^n + \dots + a_1 x + a_0$ (Polynom) ist stetig.

Praktische Abkürzungen (zur Erinnerung):

\exists für: "es existieren", z.B. $\exists x^*$

\forall für: "Für alle", z.B. $\forall x$

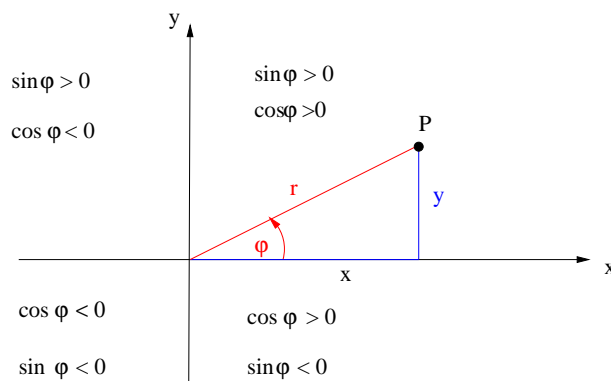
3.6 Spezielle Funktionen

A.) Kreisfunktionen

(auch **Trigonometrische Funktionen**)

\sin , \cos , \tan , \cot

(in der Geometrie zunächst nur Seitenverhältnisse im rechtwinkligem Dreieck, dann Ausdehnung über den Kreis, schließlich für ganz \mathbb{R} .)



Definitionen:

$$\sin \varphi := \frac{y}{r} \quad (\text{"Gegenkathete zu Hypotenuse"}) \quad y = r \sin \varphi$$

$$\cos \varphi := \frac{x}{r} \quad (\text{"Ankathete zu Hypotenuse"}) \quad x = r \cos \varphi$$

$$\tan \varphi := \frac{\sin \varphi}{\cos \varphi} = \frac{y}{x} \quad (\text{"Steigung"})$$

$$\cot \varphi := \frac{1}{\tan \varphi} = \frac{\cos \varphi}{\sin \varphi} = \frac{x}{y}$$

(Für Illustrationen sei auf die Vorlesung verwiesen, oder auf Literatur wie Formelsammlungen.)

Vorzeichen von $\cos \varphi =$ Vorzeichen von x

Vorzeichen von $\sin \varphi =$ Vorzeichen von y

Nach Erweiterung der trigonometrischen Funktionen auf den Kreis als nächster Schritt periodische Fortsetzung auf ganz \mathbb{R} :

$$\begin{aligned}\cos(\varphi + n \cdot 2\pi) &= \cos \varphi & \text{für } \forall n \in \mathbb{Z} \\ \sin(\varphi + n \cdot 2\pi) &= \sin \varphi & \text{für } \forall n \in \mathbb{Z} .\end{aligned}$$

Damit sind \sin, \cos überall definiert. Bezeichnung für die unabhängige Variable: φ oder t oder x (anderes x als oben!): z.B. $\sin x, \cos x$, für $x \in \mathbb{R}$.

einige **Eigenschaften der Kreisfunktionen**

$$\begin{aligned}\cos(-\varphi) &= \cos \varphi & , & \quad \cos(\pi - \varphi) = -\cos \varphi \\ \sin(-\varphi) &= -\sin \varphi & , & \quad \sin(\pi - \varphi) = \sin \varphi \\ \cos(\varphi + \pi) &= -\cos \varphi & , & \quad \cos\left(\frac{\pi}{2} - \varphi\right) = \sin \varphi \\ \sin(\varphi + \pi) &= -\sin \varphi & , & \quad \sin\left(\frac{\pi}{2} - \varphi\right) = \cos \varphi , \\ \cos\left(\varphi + \frac{\pi}{2}\right) &= -\sin \varphi \\ \sin\left(\varphi + \frac{\pi}{2}\right) &= +\cos \varphi & , & \quad \cos^2 \varphi + \sin^2 \varphi = 1 \text{ (Pythagoras) .}\end{aligned}$$

$\cos x$ und $\sin x$ sind stetig für alle x ;

$$|\cos x| \leq 1 \quad , \quad |\sin x| \leq 1$$

$$\cos x = 0 \text{ für } x = (2n + 1)\frac{\pi}{2} \quad , \quad n \in \mathbb{Z}$$

$$\sin x = 0 \text{ für } x = n \cdot \pi \quad , \quad n \in \mathbb{Z} .$$

(Additionstheoreme, z.B. $\sin(x + y) = \dots$: s.u. bzw. Formelsammlung)

$$\left. \begin{aligned}\tan(\varphi \pm n\pi) &= \tan \varphi \\ \cot(\varphi \pm n\pi) &= \cot \varphi\end{aligned} \right\} n = 0, 1, 2 \dots$$

$$\tan(-\varphi) = -\tan \varphi$$

$$\cot(-\varphi) = -\cot \varphi$$

$\tan x$ und $\cot x$ sind nur **stückweise stetig!**

Exkurs: Bezeichnungen/Definitionen

- * Eine Funktion $f(x)$ heißt **periodisch mit Periode** p , wenn für jedes x und jedes $n \in \mathbb{Z}$ gilt:

$$f(x + np) = f(x)$$

Beispiele:

$$\sin x, \cos x : \text{ Periode } 2\pi$$

$$\tan x, \cot x : \text{ Periode } \pi$$

- * Eine Funktion $f(x)$ heißt **gerade**, wenn $f(-x) = f(x)$
Beispiele: $\cos x, x^2$. geometrische Bedeutung: Symmetrie zur y -Achse.
- * Eine Funktion $f(x)$ heißt **ungerade**, wenn $f(-x) = -f(x)$
Beispiele: $\sin x, \tan x, \cot x, x^3$. geometrische Bedeutung: Symmetrie zum Nullpunkt.

Weitere Beispiele für **gerade** Funktionen:

$$f(x) := \sin |x|, \text{ denn: } f(-x) = \sin |-x| = \sin |x| = f(x)$$

$$f(x) := |\sin x|, \text{ denn: } f(-x) = |\sin(-x)| = |-\sin x| = |\sin x| = f(x)$$

B.) Weitere Formeln zu trigonometrischen Funktionen

Multiplikation von komplexen Zahlen (vgl. Abschnitt 3.2) z_1 und $z_2 \in \mathbb{C}$ mit $|z_1| = |z_2| = 1$:

$$(\cos \varphi + i \sin \varphi) \cdot (\cos \psi + i \sin \psi) = \cos(\varphi + \psi) + i \sin(\varphi + \psi)$$

Andererseits, mit Ausmultiplizieren erhalte

$$(\cos \varphi \cos \psi - \sin \varphi \sin \psi) + i(\sin \varphi \cos \psi + \cos \varphi \sin \psi)$$

Folgerung mit "Koeffizientenvergleich":

$$\begin{aligned} \sin(\varphi + \psi) &= \sin \varphi \cos \psi + \cos \varphi \sin \psi \\ \cos(\varphi + \psi) &= \cos \varphi \cos \psi - \sin \varphi \sin \psi \end{aligned}$$

Diese Formeln heißen **Additionstheoreme**.

Hieraus folgen andere wichtige Formeln, siehe Formelsammlung!

z.B. für

$$\begin{aligned} \sin(\varphi - \psi), \quad \cos(\varphi - \psi), \quad \tan(\varphi \pm \psi); \\ \sin(2\varphi), \quad \cos(2\varphi), \quad \sin \alpha \pm \sin \beta, \quad \cos \alpha \pm \cos \beta \end{aligned}$$

C.) Anwendungen trigonometrischer Funktionen:

- 1.) überall im Alltag (vom Regenbogen zum Farbfernsehen)
- 2.) Lösungen einfacher Schwingungsgleichungen
- 3.) Fourier: Für periodische Funktionen $y(t)$ mit Periode T

$$y(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k \frac{2\pi}{T} t) + b_k \sin(k \frac{2\pi}{T} t))$$

Fourier-Koeffizienten:

$$a_k = \frac{2}{T} \int_0^T y(t) \cos(k \frac{2\pi}{T} t) dt$$

$$b_k = \frac{2}{T} \int_0^T y(t) \sin(k \frac{2\pi}{T} t) dt$$

$y \longrightarrow a_k, b_k$: Fourier-Analyse

$a_k, b_k \longrightarrow y$: Fourier-Synthese

D.) Potenzen und Wurzeln von reellen Zahlen (≥ 0)

Schulstoff: $x^0 := 1$ für $x \neq 0$, $x^n := \underbrace{x \cdot x \cdot \dots \cdot x}_{n\text{-mal}}$, $x^{-n} := \frac{1}{x^n}$

$$\Rightarrow x^n x^m = x^{n+m}, \quad x^{mn} = (x^m)^n = (x^n)^m \quad \text{für } n, m \in \mathbb{Z}, x \in \mathbb{R}$$

Es seien $x > 0$, $y > 0$ reelle Zahlen, $n \in \mathbb{Z}$.

Die Funktion $x = y^n$ ist für $x \geq 0$ monoton und stetig \Rightarrow Dann existiert Umkehrfunktion (vgl. Abschnitt 3.4.B). D.h. für jedes $x > 0$ gibt es genau ein $y > 0$ so dass $x = y^n$.

Schreibweisen: $y = (x)^{1/n} = x^{1/n} = \sqrt[n]{x}$ “**n-te Wurzel**”

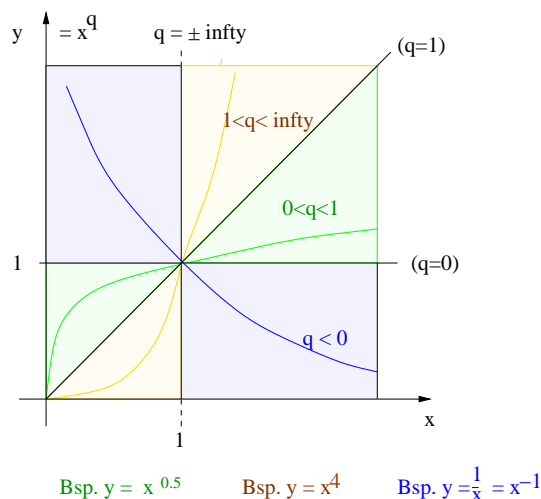
Sei nun $q \in \mathbb{Q}$, $q = \frac{m}{n}$. Dann gilt für $x \geq 0$:

$$x^q := x^{\frac{m}{n}} = \sqrt[n]{x^m} = (\sqrt[n]{x})^m$$

$$\Rightarrow x^q x^p = x^{q+p} \quad \text{auch für rationale Exponenten.}$$

Potenzen mit reellen Exponenten: \rightarrow SPÄTER

Funktionen $y = x^q$: alle gehen durch den Punkt (1, 1).



Die Kurven von $y = x^q$, $x \geq 0$, $q \in \mathbb{Q}$ “überdecken” den Quadranten.

E.) Weitere spezielle Funktionen

Umkehrfunktionen der trigonometrischen Funktionen:

$$\arcsin, \quad \arccos, \quad \arctan, \quad \operatorname{arccot}$$

vgl. Formelsammlungen.

Exponentialfunktion $\exp(x) := e^x$ und $\log(x)$ werden in Abschnitt 3.10 formal definiert. Sie sind Umkehrfunktionen zueinander (Verlauf und Achsenschnittpunkte: vgl. Figur)

weitere wichtige Eigenschaften:

$$e^{x_1} e^{x_2} = e^{x_1+x_2} \quad ; \quad \log x_1 + \log x_2 = \log(x_1 x_2)$$

F.) Kurvendiskussionen

zu diskutieren: $y = f(x)$

Programm:

- 1.) Feststellung des Definitionsbereiches und des Wertebereiches
- 2.) Bestimmung der Achsenschnittpunkte $x = \alpha_i$, wo

$$f(\alpha_i) = 0,$$

sowie $f(0)$.

3.) Bei rationalen Funktionen:

Teiler-Fremdheit herstellen (Euklidischer Algorithmus), und bei $m \geq n$ Ausdividieren

4.) Asymptotisches Verhalten für $x \rightarrow \pm\infty$; Bestimmung von Unendlichkeitsstellen $x = \beta_i$

$$(f(x) \rightarrow \pm\infty \quad \text{für} \quad y \rightarrow \beta_i)$$

mit Bestimmung des Vorzeichens von $f(x)$ für β_i^+, β_i^-

5.) Auffinden evtl. Symmetrien (f gerade oder ungerade?)

6.) Monotonie, Maxima, Minima, Wendepunkte (Schule, bzw. Abschnitt 3.8)

7.) evtl. weitere Kurvenpunkte berechnen.

8.) (qualitative) Skizze

Bemerkung: Es sind nicht immer *alle* Punkte des obigen Programms sinnvoll oder machbar.

3.7 Grenzwert

A.) Folgen und Reihen

Eine Folge von Zahlen $a_1, a_2, a_3, \dots, a_\nu, \dots$ nennt man **(Zahlen-)Folge**. Dargestellt auf der Zahlengerade entspricht dies einer **Punktfolge**.

Beispiele:

(a) $1, 2, 3, 4, \dots$ d.h. $a_\nu = \nu$

(b) $1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots$ d.h. $a_\nu = (-1)^{\nu-1} \frac{1}{\nu}$

(c) $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \dots$ d.h. $a_\nu = \frac{\nu}{1+\nu}$

(d) $0.2, 0.22, 0.222, \dots$ d.h. $a_\nu = \sum_{k=1}^{\nu} 2 \cdot 10^{-k}$

Bemerkung: Der Index ν kann z.B. auch bei 0 loslaufen. Abkürzende Bezeichnung für Folgen: $\{a_\nu\}$ oder $\{a_\nu\}_{\nu=1}^{\infty}$ oder evtl. $\{a_\nu\}_{\nu=0}^{\infty}$

Definition: Eine Folge heißt **beschränkt**, wenn es eine Zahl M gibt, so dass $|a_\nu| \leq M$ für alle ν ; oder äquivalent: $M_1 \leq a_\nu \leq M_2$ (M_1 = "untere Schranke", M_2 = "obere Schranke")

M "Schranke"

obige **Beispiele:**

(b), (c) beschränkt durch jede Zahl $M \geq 1$

(d) z.B. $M = \frac{2}{9}$ (kleinste aller oberen Schranken)

(a) unbeschränkt

Definition: Monotonie einer Folge: Eine Folge heißt monoton, wenn eine der folgenden Bedingungen für alle ν zutrifft:

$a_{\nu+1} > a_\nu$: monoton wachsend

$a_{\nu+1} < a_\nu$: monoton abnehmend oder fallend

$a_{\nu+1} \geq a_\nu$: monoton nicht abnehmend

$a_{\nu+1} \leq a_\nu$: monoton nicht wachsend

Beispiel:

(a), (c), (d): monoton wachsend

(b): nicht monoton (aber $|a_\nu|$ ist monoton abnehmend)**Definition:** h heißt **Häufungspunkt** oder **Häufungsstelle** der Folge $\{a_\nu\}$, wenn in jeder noch so kleinen Umgebung von h unendliche viele a_ν liegen.d.h.: für alle $\varepsilon > 0$ gilt $|h - a_\nu| < \varepsilon$ für unendlich viele a_ν .obige **Beispiele:**(b) : Häufungsstelle: $h = 0$ (c) : Häufungsstelle: $h = 1$ (d) : Häufungsstelle: $h = 2/9$ Das Beispiel $(-1)^{\nu-1} \frac{\nu}{1+\nu}$ hat 2 Häufungsstellen: $h = +1$ und $h = -1$.**Satz [Bolzano & Weierstraß]**

Jede beschränkte Folge hat mindestens eine Häufungsstelle.

(unendliche Folgen, d.h. $\nu \rightarrow \infty$)*Beweis:* $|a_\nu| \leq M$ definiert ein Intervall I_1 der Länge $2M$ mit unendlich vielen a_ν .Intervallhalbierung \Rightarrow in wenigstens einem Teilintervall liegen ∞ -viele a_ν : Intervall I_2 . (usw.) \rightarrow "Intervallschachtelung": $I_1 \supset I_2 \supset I_3 \supset \dots$, Breite der I_k geht gegen 0.In jedem I_k liegen ∞ -viele a_ν . I_k konvergieren gegen einen Punkt, der die Voraussetzungen des Häufungspunktes erfüllt.**Definition: Reihe:** Es sei $\{a_\nu\}$ eine Zahlenfolge. Die Folge der **Partialsommen** $\{s_\nu\}$,definiert durch $s_1 := a_1$, $s_2 := a_1 + a_2, \dots, s_\nu := \sum_{k=1}^{\nu} a_k, \dots$, heißt unendliche Reihe.Bezeichnung: $\sum_{\nu=1}^{\infty} a_\nu$.**B.) Konvergenz von Folgen****Definition:** a heißt **Grenzwert** der Folge $\{a_\nu\}$, wenn in jeder noch so kleinen Umgebung von a **alle** a_ν **mit Ausnahme von endlich vielen** liegen.D.h. für alle $\varepsilon > 0$ gibt es $N = N(\varepsilon) \in \mathbb{N}$, so dass $|a - a_\nu| < \varepsilon$ für alle $\nu > N(\varepsilon)$.

Folgerungen:

- 1.) Jeder Grenzwert ist Häufungspunkt.
- 2.) Falls 2 Häufungspunkte existieren, gibt es keinen Grenzwert.
- 3.) Also ist nicht jeder Häufungspunkt auch Grenzwert. (Beispiel: $1, -1, 1, -1, \dots$)

Definition: Hat eine Folge $\{a_\nu\}$ einen Grenzwert a , so heißt sie **konvergent**.Bezeichnung: $\lim_{\nu \rightarrow \infty} a_\nu = a$ ("limes")oder: $a_\nu \rightarrow a$ für $\nu \rightarrow \infty$ **Definition:** Nichtkonvergente Folgen heißen **divergent**.**Konvergenzkriterium von Cauchy**Eine Folge $\{a_\nu\}$ ist "dann und nur dann" (d.h. \Leftrightarrow) konvergent, wenn für alle hinreichend großen $n \in \mathbb{N}$ und beliebige $m \in \mathbb{N}$ die Differenzbeträge $|a_{n+m} - a_n|$ beliebig klein werden.

D.h. äquivalent zur Konvergenz der $\{a_\nu\}$ ist:

$$\forall \varepsilon > 0 \quad \exists N(\varepsilon) \text{ so dass } \forall n > N(\varepsilon) \text{ und } \forall m \in \mathbb{N} \quad \text{gilt } |a_{m+n} - a_n| < \varepsilon$$

Bemerkung: Der Grenzwert a wird hier nicht explizit benötigt!

Beispiele:

$$1.) \quad 0.2, 0.22, \dots, a_\nu = \sum_{k=1}^{\nu} 2 \cdot 10^{-k} \Rightarrow$$

$$a_{m+n} - a_n = \underbrace{\sum_{k=n+1}^{n+m} 2 \cdot 10^{-k}}_{m \text{ Terme}} = 2 \cdot 10^{-(n+1)} \cdot \left(\underbrace{1 + \frac{1}{10} + \frac{1}{100} + \dots + 10^{-(m-1)}}_{\text{Mit } x := \frac{1}{10} \text{ ist dies eine Partialsumme } 1+x+x^2+\dots+x^{m-1}=:s_m} \right)$$

Herleitung einer Formel für s_m :

$$\begin{aligned} x s_m &= x + x^2 + \dots + x^m \Rightarrow \\ x s_m - s_m &= x^m - 1 \Rightarrow s_m = \frac{1 - x^m}{1 - x} \end{aligned}$$

Es folgt für das Beispiel:

$$\begin{aligned} s_m &< \frac{1}{1 - \frac{1}{10}} = \frac{10}{9} \Rightarrow \\ a_{m+n} - a_n &< 2 \cdot 10^{-n} \frac{10}{9} < 10^{-n} \end{aligned}$$

Für genügend großes n wird dies kleiner als jedes $\varepsilon > 0$.

2.) $s_m := 1 + x + x^2 + \dots + x^{m-1}$ heißt **endliche geometrische Reihe**.

$$s_m = \frac{1-x^m}{1-x}$$

Für $|x| < 1 \Rightarrow x^m \rightarrow 0$ für $m \rightarrow \infty$

$$\Rightarrow s_m \rightarrow \frac{1}{1-x} \text{ für } m \rightarrow \infty$$

$$\Rightarrow \sum_{\nu=0}^{\infty} x^\nu = \frac{1}{1-x} \text{ für } |x| < 1$$

heißt (unendliche) **geometrische Reihe**.

Regeln

$$a_\nu \rightarrow a, b_\nu \rightarrow b \Rightarrow \begin{cases} (a_\nu \pm b_\nu) \rightarrow (a \pm b) \\ (a_\nu \cdot b_\nu) \rightarrow (a \cdot b) \\ \left(\frac{a_\nu}{b_\nu}\right) \rightarrow \left(\frac{a}{b}\right) \end{cases} \quad \text{falls } b_\nu \neq 0 \text{ und } b \neq 0$$

$$a_\nu \rightarrow a \Rightarrow |a_\nu| \rightarrow |a|$$

$$a_\nu \rightarrow a \Rightarrow \alpha a_\nu \rightarrow \alpha a$$

$$a_\nu \rightarrow a, b_\nu \rightarrow b, a_\nu < b_\nu \Rightarrow a \leq b \quad !!$$

$$\text{(Beispiel: } a_\nu = \frac{1}{(\nu+1)^2}, b_\nu = \frac{1}{\nu+1}, a = b = 0)$$

einfaches hinreichendes Kriterium: $\{a_\nu\}$ monoton und beschränkt \Rightarrow Konvergenz

C.) Grenzwert einer Funktion; Stetigkeit

Definition: g ist **Grenzwert der Funktion** $f(x)$ **an der Stelle** $x = x_0$, wenn für **jede** Folge $x_\nu \rightarrow x_0$, für die $f(x)$ definiert ist, die Folge $\{f(x_\nu)\}$ gegen g konvergiert.

Schreibweise: $\lim_{x \rightarrow x_0} f(x) = g$.

→ alternative Definition der **Stetigkeit** von f in x_0 ($g = f(x_0)$):

Die Funktion $f(x)$ ist an $x = x_0$ **stetig**, wenn

$$f(x) \rightarrow f(x_0) \text{ für } x \rightarrow x_0$$

(im Sinne: für *alle* Folgen $x_\nu \rightarrow x_0$)

Folgerung: f stetig $\Rightarrow f$ und Limes vertauschbar:

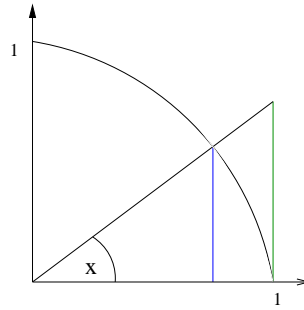
$$\lim_{x \rightarrow x_0} f(x) = f(\lim_{x \rightarrow x_0} x) \quad (= f(x_0))$$

z.B. $\sqrt{\lim a_\nu} = \lim \sqrt{a_\nu}$

Beispiel:

$$f(x) = \begin{cases} 1 & \text{für } x = 0 \\ \frac{\sin x}{x} & \text{für } x \neq 0 \end{cases} \quad \text{ist stetig, denn:}$$

zu zeigen ist: $\frac{\sin x}{x} \rightarrow 1$ für $x \rightarrow 0$.



Kreis mit Radius 1. x im Bogenmaß \Rightarrow Fläche des Kreissektors ist $\frac{x}{2}$.
Vergleich mit den Flächen des “inneren” und des “äußeren” Dreiecks:

$$\frac{1}{2} \sin x \cos x < \frac{x}{2} < \frac{1}{2} \tan x \quad \Rightarrow \quad \cos x < \frac{x}{\sin x} < \frac{1}{\cos x}$$

Also: $x \rightarrow 0 \Rightarrow \cos x \rightarrow 1$

3.8 Differentialrechnung

A.) Differentialquotient

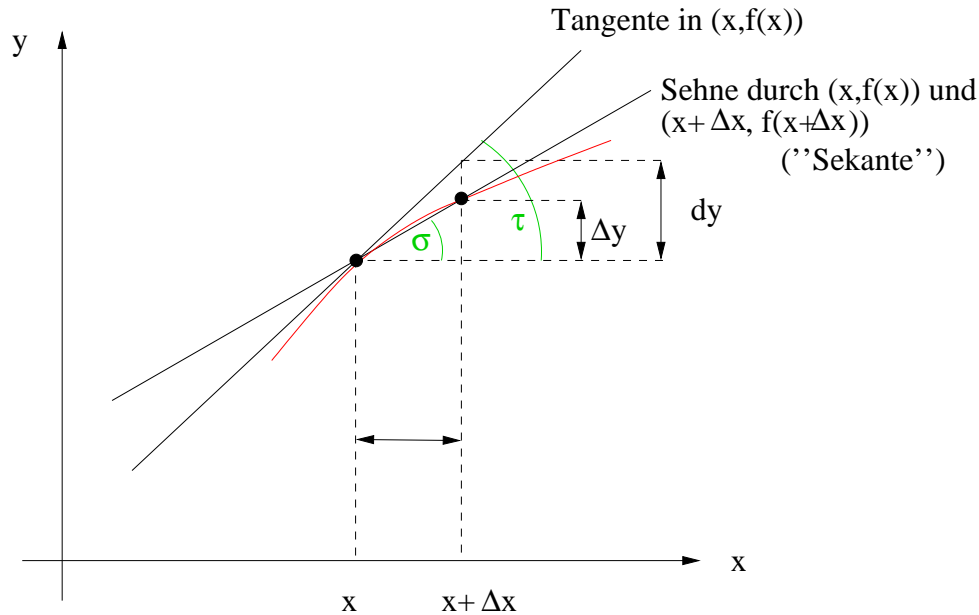
$y = f(x)$ Funktion. Es sei $|h|$ eine kleine Änderung von x .

Wie verhält sich $f(x + h)$?

praktische Bezeichnung: $h = \Delta x$ ("Inkrement". Dies ist **kein** Produkt, sondern Symbol für "kleine" Änderungen.)

Definition: Differenzenquotient

$$\frac{\Delta y}{\Delta x} := \frac{f(x + \Delta x) - f(x)}{\Delta x} = \tan \sigma$$



Bemerkung: $f(x)$ stetig $\Rightarrow \Delta y \rightarrow 0$ für $\Delta x \rightarrow 0$

Definition: Falls der Grenzwert $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ "existiert" (d.h. $\frac{\Delta y}{\Delta x}$ konvergiert für $\Delta x \rightarrow 0$ gegen einen endlichen Grenzwert), nennen wir ihn **Differentialquotient**, oder **Ableitung**, und bezeichnen ihn mit

$$y'(x) \quad \text{oder} \quad \frac{dy}{dx}$$

geometrisch: Die Sekante strebt gegen die Tangente mit Steigung

$$y'(x) = \tan \tau .$$

Beispiele:

1.)

$$\begin{aligned} y = f(x) &= a + bx + cx^2 \quad \Rightarrow \\ \Delta y &= a + b(x + \Delta x) + c(x + \Delta x)^2 - a - bx - cx^2 \\ &= b\Delta x + c\Delta x^2 + 2c\Delta x \cdot x \quad \Rightarrow \\ \frac{\Delta y}{\Delta x} &= b + 2cx + c\Delta x \quad \Rightarrow \quad \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = b + 2cx \end{aligned}$$

$$2.) y = \sin x$$

$$\Delta y = \sin(x + \Delta x) - \sin x = 2 \sin \frac{\Delta x}{2} \cos(x + \frac{\Delta x}{2})$$

$$\Rightarrow \frac{\Delta y}{\Delta x} = \frac{\sin \frac{\Delta x}{2}}{\frac{\Delta x}{2}} \cdot \cos(x + \frac{\Delta x}{2}) \longrightarrow \cos x \text{ für } \Delta x \rightarrow 0,$$

$$\text{da } \lim_{\xi \rightarrow 0} \frac{\sin \xi}{\xi} = 1.$$

B.) Differenzierbare Funktionen

Falls $\lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$ existiert (d.h. Konvergenz für $\Delta x \rightarrow 0$), nennen wir f in x **differenzierbar**.

Folgerung: $\lim_{\Delta x \rightarrow 0} f(x + \Delta x) = f(x)$, d.h. f ist stetig.

Also ist Differenzierbarkeit eine stärkere Forderung:

$$f \text{ differenzierbar in } x \Rightarrow f \text{ stetig in } x$$

Gilt die Umkehrung? Nein!

1. Beispiel: $y = \sqrt{x}$ an $x = 0$:

Differenzenquotient:

$$\frac{\sqrt{0 + \Delta x} - \sqrt{0}}{\Delta x} = \frac{\sqrt{\Delta x}}{\Delta x} = \frac{1}{\sqrt{\Delta x}} \longrightarrow \infty \text{ für } \Delta x \rightarrow 0^+$$

also in $x = 0$ nicht differenzierbar (Bezeichnung 0^+ : siehe unten).

für $x > 0$:

$$\frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x} = \frac{\Delta x}{\Delta x(\sqrt{x + \Delta x} + \sqrt{x})} \longrightarrow \frac{1}{2\sqrt{x}} \text{ für } \Delta x \rightarrow 0$$

d.h. differenzierbar in $x > 0$.

2. Beispiel: $f(x) = |x|$ an $x = 0$:

\Rightarrow Differenzenquotient

$$\frac{|0 + \Delta x| - |0|}{\Delta x} = \frac{|\Delta x|}{\Delta x} = \begin{cases} +1 & \text{für } \Delta x \rightarrow 0^+ \\ -1 & \text{für } \Delta x \rightarrow 0^- \end{cases}$$

Hinter dem $\Delta x \rightarrow 0$ können wir uns eine beliebige Folge $x_\nu \rightarrow 0$ vorstellen. Da also nicht für alle Folgen $x_\nu \rightarrow 0$ ein einheitlicher Grenzwert existiert, ist $|x|$ in $x = 0$ nicht differenzierbar.

Aber: Man kann **einseitige** Grenzwerte definieren, wenn man in der Definition zu Beginn von Abschnitt 3.7C das "jede Folge $x_\nu \rightarrow x_0$ " einschränkt durch $x_\nu < x_0$ oder $x_\nu > x_0$. Schreibweise:

$$\lim_{x \rightarrow x_0^-} (\dots) \quad \text{bzw.} \quad \lim_{x \rightarrow x_0^+} (\dots)$$

Bei Beispiel 2 gibt es sowohl einen linksseitigen wie einen rechtsseitigen Grenzwert.

Bemerkung: Es gibt sogar stetige Funktionen, die *nirgends* differenzierbar sind (z.B. Aktienkurse).

Die meisten praktisch auftretenden Funktionen sind stückweise differenzierbar. D.h. $f(x)$ ist differenzierbar bis auf endlich viele Ausnahmepunkte $x_1, x_2, x_3, \dots, x_n$.

(Illustration an Figur)

Allgemein gilt: $f(x)$ nicht stetig an $x_0 \Rightarrow f(x)$ nicht differenzierbar an x_0 .

C.) Höhere Ableitung (für Funktionen $y(x)$)

$y'(x)$ ist wiederum Funktion, $v(x) := y'(x)$

$v(x)$ existiert für diejenigen x , für die $y(x)$ differenzierbar ist.

Auch für $v(x) = y'(x)$ können der Differenzen- und der Differentialquotient betrachtet werden. Existiert letzterer, so nennen wir $v'(x) = \frac{dv}{dx} = \frac{dy'(x)}{dx}$ die **2. Ableitung** von $y(x)$.

Schreibweisen: $y''(x)$, $\frac{d^2y}{dx^2}$

(Diese "2" bedeutet keine Quadratur! $\frac{d^2}{dx^2}$ ist Symbol für die 2. Ableitung nach x .)

Definition: $y(x)$ heißt **glatt** (oder stetig differenzierbar), wenn $y(x)$ differenzierbar ist und $y'(x)$ stetig ist.

(Dann hat $y(x)$ keinen "Knick"; die Tangente an die Kurve variiert stetig.)

[Hinweis für Spezialisten: Es gibt Funktionen, welche in x differenzierbar sind, aber $y'(x)$ ist *nicht* stetig. (z.B. $y(x) = x^2 \sin \frac{1}{x}$ für $x \neq 0$, $y(0) := 0$)]

Auch $y''(x)$ ist Funktion, deren Ableitung analog definiert werden kann $\longrightarrow y'''(x)$ (3. Ableitung).

u.s.w. n -te Ableitung

Bezeichnung $y^{(n)} = \frac{d^n y}{dx^n} := \frac{d}{dx} y^{(n-1)}$ ("rekursiv" definiert).

praktische

Definition: Eine Funktion $f(x)$ heißt **C^k -Funktion** auf $[a, b]$, wenn $f(x)$ auf $[a, b]$ **k -mal "stetig differenzierbar"** ist. D.h. die Ableitungen $f', \dots, f^{(k)}$ existieren und $f^{(k)}$ ist stetig für $a \leq x \leq b$.

d.h. $y(x) \ C^k$ -Funktion $\Leftrightarrow y(x), y'(x), \dots, y^{(k)}(x)$ stetig.

C^0 bezeichnet die Menge aller stetigen Funktionen.

$f \in C^0$ heißt: f ist stetig.

$f \in C^2$ heißt: f, f' und f'' existieren und sind stetig.

$f(x)$ ist "glatt", wenn f wenigstens C^1 ist.

D.) Ableitungsregeln

elementare Regeln:

$$y(x) = cu(x) \text{ mit } c \text{ konstant} \Rightarrow y' = cu'$$

$$y(x) = u(x) \pm v(x) \Rightarrow y' = u' \pm v'$$

$$y(x) = u(x)v(x) \Rightarrow y' = u'v + uv'$$

$$y = \frac{u(x)}{v(x)} \Rightarrow y' = \frac{u'v - uv'}{v^2}$$

$y = y(x)$, Umkehrfunktion $x = x(y) \Rightarrow \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{y'}$ falls $y' \neq 0$ (d.h. monoton!)

Kettenregel: Hintereinanderausführung von Funktionen

$$y = y(x), \quad x = x(t), \quad y(x(t)) \Rightarrow \frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

Beispiel: $h(x) = (x^4 + 6x + 5)^3$ Anwenden der Kettenregel:

Definition: $f(x) = x^4 + 6x + 5$ und $g(x) = x^3 \Rightarrow h(x) = (f(x))^3 = g(f(x))$

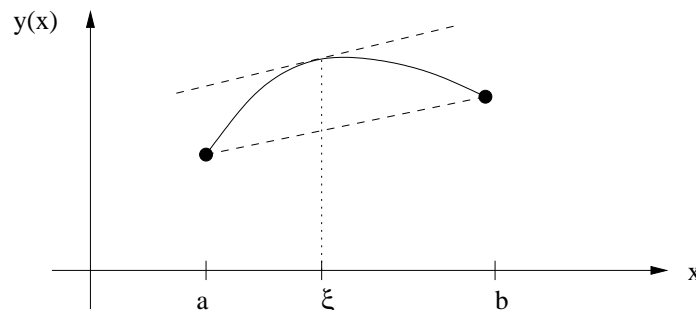
$$\begin{aligned} h'(x) &= \frac{dg(f(x))}{df} \cdot \frac{df(x)}{dx} = 3(f(x))^2(4x^3 + 6) \\ &= 3(4x^3 + 6)(x^4 + 6x + 5)^2 \end{aligned}$$

(“Nachdifferenzieren”. Hinweis zu dem Beispiel: Ergebnis kann man sofort hinschreiben.)

Beispiel: $f(x) = \sin(ax + b)$

$$f'(x) = \cos(ax + b) \cdot a = a \cos(ax + b) \quad (a \text{ vom Nachdifferenzieren})$$

E.) Mittelwertsatz



Mittelwertsatz der Differentialrechnung (für eine Funktion):

Falls $y(x)$ stetig in $a \leq x \leq b$ und differenzierbar in $a < x < b$ ist, **dann** gibt es (mindestens) ein ξ in $a < \xi < b$, so dass $y'(\xi) = \frac{y(b) - y(a)}{b - a}$.

Folgerung für Fehlerrechnung

$y = f(x)$, x : Eingangsgröße, y : Ausgangsgröße.

Es sei Δx Fehler in x , was ist Δy ? Mit Mittelwertsatz, falls Voraussetzung erfüllt ist: $\Delta y = f'(\xi)\Delta x \Rightarrow |\Delta y| \leq M|\Delta x|$, wobei M Schranke für $f'(x)$ im jeweiligen x -Bereich ist.

Folgerungen bzgl. Monotonie

Es sei $y'(x) > 0$ in $a \leq x \leq b$ und $x_1 < x_2$ beliebig im Intervall.

Mittelwertsatz $\Rightarrow \exists \xi$ mit $x_1 < \xi < x_2$, so dass $\frac{y(x_2) - y(x_1)}{x_2 - x_1} = y'(\xi) > 0 \Rightarrow y(x_2) > y(x_1)$

Also gilt der **Satz:** (Die Voraussetzungen des Mittelwertsatzes seien erfüllt.) $y'(x) > 0 \Rightarrow y$ streng monoton wachsend (jeweils auf dem ganzen Intervall)

analog: $y'(x) < 0 \Rightarrow y$ monoton fallend.

(mit jeweiliger Intervall-Angabe)

abgeschwächt:

$$\begin{aligned} y'(x) \geq 0 &\Leftrightarrow \text{monoton wachsend} \\ y'(x) = 0 &\Leftrightarrow y \text{ ist konstant} \end{aligned}$$

F.) Regeln nach l'Hospital

Für Ausdrücke in unbestimmter Form

Beispiel: $\lim_{x \rightarrow 0} \frac{\sin x}{x}$ ist von der Form " $\frac{0}{0}$ "

Schreibe $f(x)$ für den Zähler und $g(x)$ für den Nenner.

Es gelte entweder

Voraussetzung (1): $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ (d.h. von der Form " $\frac{0}{0}$ ")

Voraussetzung (2): $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \infty$ (d.h. Form " $\frac{\infty}{\infty}$ ")

Bei jeder der beiden Voraussetzungen gilt die Aussage

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

(sofern der rechte Grenzwert existiert).

Beispiel: $\lim_{x \rightarrow 0} \frac{\sin x}{x} \stackrel{\text{Ho}}{=} \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1$ (Typ $\frac{0}{0}$)

Ausdrücke der Form " $0 \cdot \infty$ " und " $\infty - \infty$ " lassen sich in obige Ausdrücke umformen!

Beispiel:

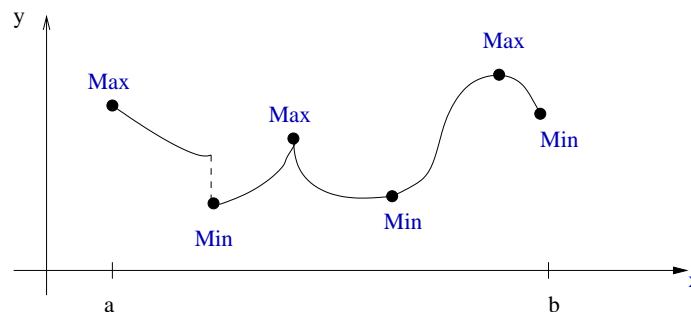
$$\lim_{x \rightarrow 0} (x \cot x) \quad (\text{Typ } "0 \cdot \infty") = \lim_{x \rightarrow 0} \frac{x}{\tan x} \stackrel{\text{Ho}}{=} \lim_{x \rightarrow 0} \frac{1}{\cos^2 x} = 1 \quad (\text{Typ } \frac{0}{0})$$

G.) Maxima und Minima

Gegeben: Funktion $y = f(x)$ auf einem Intervall $I : a \leq x \leq b$.

Definition: $f(x_0)$ ist ein **Maximum** von $f(x)$, wenn es eine Umgebung U um x_0 gibt, so dass für alle $x \neq x_0$ in I und in U gilt: $f(x) < f(x_0)$.

Analog: **Minimum**. (Beide heißen Extremum)



Satz: Es sei x_0 im **Inneren** eines Teil-Intervalls, auf dem $f(x)$ differenzierbar ist, und $f(x_0)$ ein Extremum. **Dann** gilt $f'(x_0) = 0$.

Bemerkung: Der Satz sagt nichts über Extrema an Intervallenden und Stellen, an welchen $f(x)$ nicht glatt ist.

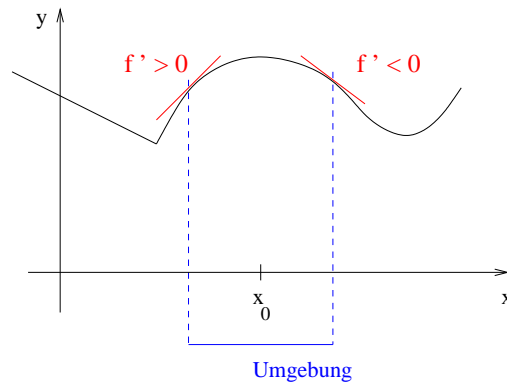
Folgerung: Kandidaten für Extrema in I sind

- 1.) die Randpunkte des Intervalls,
- 2.) die Punkte x , an denen f nicht glatt ist,
- 3.) die Stellen x mit $f'(x) = 0$.

Test, ob Maximum oder Minimum:

Es sei x in einer (kleinen) Umgebung von x_0 .

1. Kriterium: Gilt $f'(x) > 0$ für $x < x_0$ und $f'(x) < 0$ für $x > 0$, dann ist $f(x_0)$ ein lokales Maximum.



Gilt $f'(x) < 0$ für $x < x_0$ und $f'(x) > 0$ für $x > x_0$, dann ist $f(x_0)$ ein lokales Minimum. (Dieses Kriterium kann an Randextrema angepasst werden.)

2. Kriterium: Es sei x_0 im Inneren und $f'(x_0) = 0$. Für $f \in C^2$ gilt:

$$\begin{aligned} f''(x_0) < 0 &\Rightarrow f(x_0) \text{ lokales Maximum} \\ f''(x_0) > 0 &\Rightarrow f(x_0) \text{ lokales Minimum} \end{aligned}$$

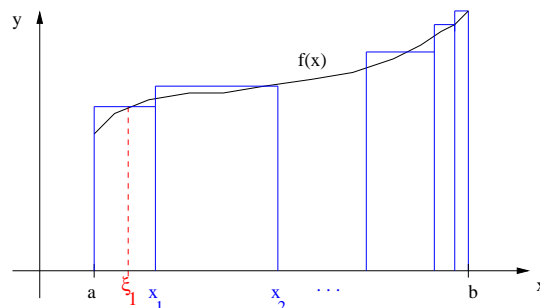
Zur **Kurvendiskussion** (vgl. Abschnitt 3.6F) hinzufügen:

Stetigkeit und Differenzierbarkeit prüfen, eventuell Monotonie-Bereiche feststellen.

3.9 Integrale

A.) Bestimmtes Integral; Flächeninhalt

Es sei $f(x)$ stetig auf $a \leq x \leq b$.



Für $f > 0$ betrachten wir den Flächeninhalt, der durch

$$y = 0, \quad x = a, \quad x = b, \quad y = f(x) \quad (*)$$

begrenzt wird.

Näherung durch Fläche von **Rechtecken**:

Einteilung des Intervalles $[a, b]$ in n Teilintervalle durch Zwischenpunkte

$$a := x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n := b$$

Dies definiert parallele Streifen der Breiten

$$x_i - x_{i-1}$$

Das i -te Rechteck habe die Länge $f(\xi_i)$ für ein ξ_i in $x_{i-1} \leq x \leq x_i$.

⇒ Gesamtfläche der Rechtecke:

$$S_n := \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$$

Für $n = 1, 2, 3, \dots$ ist S_n eine **Folge** von Zahlen. Was ist, wenn $n \rightarrow \infty$ und gleichzeitig $x_i - x_{i-1} \rightarrow 0$?

(z.B. $x_i = a + i \frac{b-a}{n}$)

Satz: Für alle stetigen $f(x)$ (auch negativ) hat die Folge S_n einen Grenzwert, egal wie die Teilintervalle und die ξ_i darin gewählt werden.

Definiton: $\int_a^b f(x) dx := \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$ (heißt auch "Riemann-Integral")

Der *Beweis* des Satzes verwendet

$$m_i := \text{Min von } f \text{ auf } x_{i-1} \leq x \leq x_i$$

$$M_i := \text{Max von } f \text{ auf } x_{i-1} \leq x \leq x_i$$

und

$$\text{"Obersumme" } \overline{S}_n := \sum_{i=1}^n M_i(x_i - x_{i-1})$$

$$\text{"Untersumme" } \underline{S}_n := \sum_{i=1}^n m_i(x_i - x_{i-1})$$

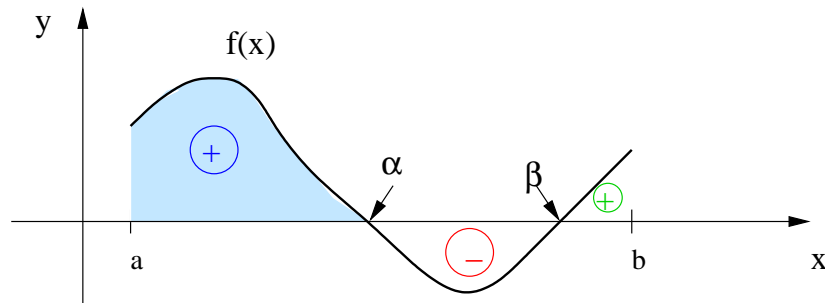
Klar ist $\underline{S}_n \leq S_n \leq \overline{S}_n$. Es folgen Konvergenz-Betrachtungen für $n \rightarrow \infty$, die hier nicht durchgeführt werden.

Bemerkung:

- 1.) $\int_a^b f(x)dx$ heißt **bestimmtes Integral**, a und b heißen *Integrationsgrenzen*, $f(x)$ heißt **Integrand**.
- 2.) Die "Integrationsvariable" ist intern und kann beliebig bezeichnet werden:

$$\int_a^b f(x)dx = \int_a^b f(\alpha)d\alpha = \dots$$

- 3.) Das obige bestimmte Integral ist eine **Zahl**, keine Funktion.
- 4.) $\int_a^b f(x)dx$ ist nur dann der (positive) Flächeninhalt des Bereiches (*), wenn $f(x) \geq 0$ für $a \leq x \leq b$. Falls $f(x) \leq 0$ für alle $a \leq x \leq b$ wäre der Flächeninhalt $-\int_a^b f(x)dx$. Bei verschiedenen Vorzeichen von f : Nullstellen aufsuchen und Teilintegrale berechnen!



$$\text{Gesamtfläche} = \int_a^\alpha f(x)dx - \int_\alpha^\beta f(x)dx + \int_\beta^b f(x)dx$$

- 5.) Aus der Definition des Integrals folgen diverse einfache Regeln, wie

$$\int_a^b 1 dx = b - a, \text{ usw. (s.u.)}$$

B.) Mittelwertsatz der Integralrechnung

Es seien f und g stetig und $g(x) \geq 0$ auf $a \leq x \leq b$. Dann \exists ein ξ mit $a \leq \xi \leq b$, so dass

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx$$

Speziell für $g = 1$: $\int_a^b f(x)dx = f(\xi)(b - a)$

Integrations-Regeln

$$\int_a^b dx = b - a$$

$$f(x) \leq g(x) \text{ für } a \leq x \leq b \Rightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx \quad (*)$$

(Spezialfälle $f = 0$ oder $g = 0$!)

$$\alpha \text{ konstant} \quad \Rightarrow \quad \int_a^b \alpha y(x) dx = \alpha \int_a^b y(x) dx$$

$$\int_a^b (f(x) \pm g(x)) dx = \int_a^b f(x) dx \pm \int_a^b g(x) dx$$

(“Linearität des Integrations-Operators”)

$$\int_a^b y(x) dx + \int_b^c y(x) dx = \int_a^c y(x) dx \quad \text{für } a \leq b \leq c$$

$$\int_a^b y(x) dx = - \int_b^a y(x) dx$$

$$m \leq f(x) \leq M \quad \underset{(*)}{\Rightarrow} \quad m(b-a) \leq \int_a^b f(x) dx \leq M(b-a)$$

(rechtes “ \leq ”: wegen $g := M$ und $\int_a^b dx = b-a$)

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx \quad (\text{für } a \leq b)$$

(folgt mit $-|f(x)| \leq f(x) \leq |f(x)|$ aus $(*)$)

C.) Unbestimmtes Integral und Differentiation

jetzt: Integration als “Umkehrung” der Differentiation. In

$$\int_a^b f(t) dt$$

fasse b als variabel auf: $F(b) := \int_a^b f(t) dt$.

schreibe $b \rightarrow x$, d.h.

$$F(x) := \int_a^x f(t) dt$$

mit Mittelwertsatz:

$$\begin{aligned} F(x + \Delta x) - F(x) &= \int_a^{x+\Delta x} f(t) dt - \int_a^x f(t) dt = \\ &= \int_x^{x+\Delta x} f(t) dt = f(\xi) \int_x^{x+\Delta x} dt = f(\xi) \Delta x \\ \Rightarrow \quad \frac{F(x + \Delta x) - F(x)}{\Delta x} &= f(\xi) \quad \text{mit } x \leq \xi \leq x + \Delta x \end{aligned}$$

Das ist der

Hauptsatz der Differential- und Integralrechnung: Die Integralfunktion $F(x)$ einer stetigen Funktion $f(x)$ ist differenzierbar, und es gilt

$$\frac{dF(x)}{dx} = \frac{d}{dx} \left(\int_a^x f(t) dt \right) = f(x)$$

und

$$\int_a^b f(x) dx = F(b) - F(a).$$

(zum 2. Teil der Aussage:)

Übergang zu einer anderen Konstanten \tilde{a} statt a ändert $F(x)$ nur um einen konstanten Wert.

Man schreibt deshalb auch

$$\int f(x) dx = F(x) + \text{const.}$$

Bezeichnung: $\int f(x) dx$ heißt **unbestimmtes Integral** und ist eine Funktion! Jede Funktion $F(x)$ mit $\frac{dF(x)}{dx} = f(x)$ heißt **Stammfunktion** von $f(x)$.

D.) Uneigentliche Integrale

Beispiel:

$$\int_a^\infty f(x) dx := \lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

Integrale mit unendlichen Integrationsintervallen oder Unendlichkeitsstellen der Integranden heißen **uneigentlich**.

Hinweis: Integrations-*Methoden* wie partielle Integration oder Substitution: Schulstoff bzw. Literatur.

3.10 Logarithmus und Exponentialfunktionen

A.) Der natürliche Logarithmus

Integriere $\frac{1}{x}$: Die Stammfunktion ist:

Definition: $\ln x := \int_1^x \frac{1}{t} dt$ für $x > 0$ heißt “natürlicher Logarithmus”.

Folgerungen:

- 1.) Definition \Rightarrow Ableitung $\frac{d}{dx}(\ln x) = \frac{1}{x}$ für $x > 0$.
- 2.) $\ln x$ ist nur für $x > 0$ definiert.

im Folgenden seien $x, x_1, x_2 > 0$

3.)

$$\begin{aligned} \ln(x_1 \cdot x_2) &= \int_1^{x_1 x_2} \frac{dt}{t} = \int_1^{x_1} \frac{dt}{t} + \int_{x_1}^{x_1 x_2} \frac{du}{u} \quad (\text{Substitution: } u = x_1 t, \frac{du}{u} = \frac{x_1 dt}{x_1 t}) \\ &= \ln x_1 + \ln x_2 \end{aligned}$$

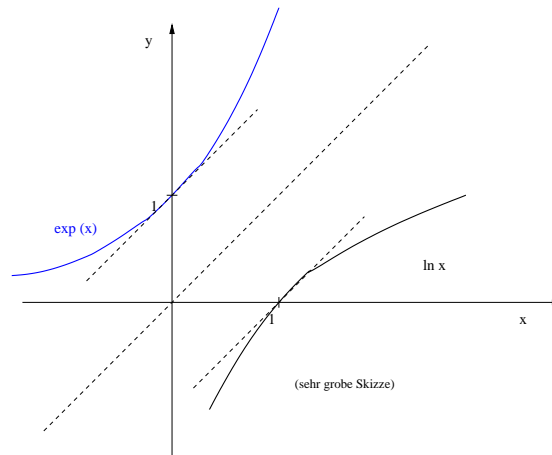
4.)

$$\ln\left(\frac{1}{x}\right) = \int_1^{1/x} \frac{dt}{t} \quad (\text{Substitution } u = \frac{1}{t}) = \int_1^x u \left(-\frac{1}{u^2}\right) du = -\int_1^x \frac{dt}{t} = -\ln x$$

5.) $\ln(x^q) = q \ln x$ für $x > 0$ und rationales q (folgt aus 3.), 4.)) (gilt auch für $q \in \mathbb{R}$)

B.) Exponentialfunktionen

$\ln x$ streng monoton \Rightarrow Umkehrfunktion existiert, nenne sie $\exp(x)$



$$x = \ln y \Leftrightarrow y = \exp(x)$$

Besondere Zahl: $e := \exp(1) = 2.71828\dots$

In Abschnitt 3.6D: x^q mit $q = \frac{m}{n}$ rational.

Setze speziell $x = e$.

Was ist e^q ? Aus den \ln -Eigenschaften folgt

$$\ln e^q = q \ln e = q$$

Umkehrfunktion \exp anwenden ergibt:

$$\exp(q) \equiv e^q \quad (\text{Synonyme})$$

Als Umkehrung von $\ln x$ ist e^x jetzt auch für *nicht*rationale x definiert.

Eigenschaften

1.) $e^0 = 1$; negative x -Achse ist waagerechte Asymptote;

$$e^x \geq 1 + x; \quad e^x > 0$$

2.) Ableitung: $x = \ln y \Rightarrow \frac{dx}{dy} = \frac{1}{y} = \frac{1}{e^x}$, also

$$\frac{dy}{dx} = \frac{d}{dx} e^x = e^x$$

Die Exponentialfunktion ist die einzige Funktion mit der Eigenschaft $y' = y$.

3.) $e^{x+y} = e^x e^y$; $e^{-x} = \frac{1}{e^x}$

4.) e^{-x} ist exponentiell abfallender Prozeß.

5.) Wegen 3.) ($e^{an\Delta t} = e^{a(n-1)\Delta t} \cdot e^{a\Delta t}$) gilt für exponentielles Wachstum/Abnahme: Der Wert nach gleichen Zeitabständen ergibt sich durch Multiplikation mit immer dem gleichen problemabhängigen Faktor.

Halbwertszeit: dasjenige Δt , für das $e^{a\Delta t} = \frac{1}{2}$.

C.) Allgemeine Exponentialfunktionen und Logarithmen

Analog wie e^x auch a^x mit $a > 0$.

$$\ln(a^x) = x \ln a \quad \Rightarrow \quad \exp(\ln a^x) = a^x = e^{x \ln a}$$

D.h. a^x wird (zum Ausrechnen) auf \exp und \ln zurückgeführt. Mit $y := a^x$ ist $\ln y = x \ln a$. Es folgt

$$x = \frac{\ln y}{\ln a} \quad \text{Umkehrfunktion für } 0 < a < 1 \text{ und } 1 < a$$

Definition/Bezeichnung: ${}^a \log x$, oder

$$\log_a x := \frac{\ln x}{\ln a}$$

heißt "**Logarithmus zur Basis a**" für $x > 0$, $a > 0$, $a \neq 1$

Speziell $a = e$: ${}^e \log x = \ln x$

$a = 10$: ${}^{10} \log x =$ **Briggscher Logarithmus**

immer gilt: ${}^a \log a = 1$, ${}^a \log 1 = 0$,

Rechenregeln wie beim natürlichen Logarithmus.

3.11 Reihen und Potenzreihen

Motivation: $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \dots$

A.) Konvergenzkriterien

Zur Erinnerung: $s_n = \sum_{\nu=1}^n a_\nu$ Partialsummen

Definition: $\sum_{\nu=1}^{\infty} a_{\nu} := \lim_{n \rightarrow \infty} s_n$ heißt **Reihe**.

Beispiel: **geometrische Reihe** $1 + x + x^2 + \dots = \sum_{\nu=0}^{\infty} x^{\nu} = \frac{1}{1-x}$ für $|x| < 1$.

Die geometrische Reihe divergiert für $|x| \geq 1$!

Definition: alternierende Reihe: a_{ν} sind abwechselnd positiv und negativ.

Beispiel: geometrische Reihe für $x = -\frac{1}{2}$:

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} \pm \dots = \frac{1}{1 + \frac{1}{2}} = \frac{2}{3}$$

Definition: Reihe heißt **absolut konvergent**, wenn $\sum_{\nu=1}^{\infty} |a_{\nu}|$ konvergiert.

Folgerungen aus Cauchys Konvergenzkriterium (Abschnitt 3.7B):

1.) Reihe konvergiert absolut \Rightarrow Reihe konvergiert.

zum Beweis: $|a_{n+1} + \dots + a_{n+m}| \leq |a_{n+1}| + \dots + |a_{n+m}|$

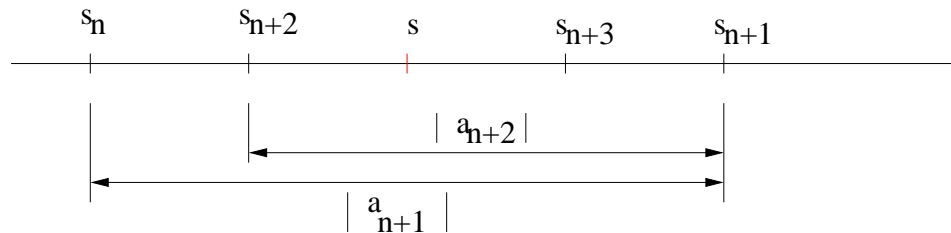
2.) Reihe konvergiert $\Rightarrow a_{\nu} \rightarrow 0$ für $\nu \rightarrow \infty$

Warnung: Umkehrung gilt nicht! z.B. divergiert $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$ ("harmonische Reihe").

3.) Reihe sei **alternierend** und es gelte für ein N und $\forall \nu \geq N$ die Ungleichung $|a_{\nu}| > |a_{\nu+1}|$. Dann gilt:

$$\text{Reihe konvergiert} \Leftrightarrow a_{\nu} \rightarrow 0 \quad \text{für} \quad \nu \rightarrow \infty$$

Beweis: s_n bilden Intervall-Schachtelung.



B.) Rechenregeln

Die Reihen $\sum_{\nu}^{\infty} a_{\nu}$ und $\sum_{\nu}^{\infty} b_{\nu}$ seien konvergent. Dann gilt:

$$\sum_{\nu}^{\infty} a_{\nu} \pm \sum_{\nu}^{\infty} b_{\nu} = \sum_{\nu}^{\infty} (a_{\nu} \pm b_{\nu})$$

$$\alpha \sum_{\nu}^{\infty} a_{\nu} = \sum_{\nu}^{\infty} \alpha a_{\nu}$$

Warnung: Klammern setzen oder fortlassen oder Umordnen ist **i.A. nicht** erlaubt!!

Beispiel: $\sum_{\nu=1}^{\infty} (-1)^{\nu+1} \frac{1}{\nu} = \ln 2$ (\rightarrow später: vergleiche Taylorentwicklungen)

Umordnen: $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} - \frac{1}{6} + \dots = \frac{3}{2} \ln 2$

C.) Konvergenzbedingungen durch Reihenvergleich

Eine Reihe $\sum_{\nu} b_{\nu}$ heißt **Majorante** von $\sum_{\nu} a_{\nu}$, wenn $0 \leq |a_k| \leq b_k$ für $k \geq N$. Dann gilt:

Majoranten-Kriterium:

$$(1) \sum_{\nu} b_{\nu} \text{ konvergent} \Rightarrow \sum_{\nu} a_{\nu} \text{ konvergiert absolut}$$

$$(2) \sum_{\nu} |a_{\nu}| = \infty \Rightarrow \sum_{\nu} b_{\nu} = \infty$$

Anwendung: Vergleich mit bekannter Reihe.

Folgerung ist

Quotienten-Kriterium: Falls $|\frac{a_{\nu+1}}{a_{\nu}}| \leq q < 1$ ist $\forall \nu > N$, dann konvergiert Reihe absolut.

Wurzelkriterium: Falls $\sqrt[\nu]{|a_{\nu}|} \leq q < 1 \quad \forall \nu > N$, dann absolute Konvergenz.

Bemerkungen:

- 1.) Jede **endliche** Teilsumme $a_0 + a_1 + \dots + a_N$ spielt für Konvergenz oder Divergenz keine Rolle.
- 2.) Insbesondere liegt Konvergenz vor, falls:

$$\lim_{\nu \rightarrow \infty} \left| \frac{a_{\nu+1}}{a_{\nu}} \right| = q < 1$$

$$\text{bzw.} \quad \lim_{\nu \rightarrow \infty} \sqrt[\nu]{|a_{\nu}|} = q < 1$$

- 3.) Falls $\lim_{\nu \rightarrow \infty} \left| \frac{a_{\nu+1}}{a_{\nu}} \right| > 1$, dann Divergenz.

Beispiel: $\sum_{\nu=0}^{\infty} \frac{x^{\nu}}{\nu!} = 1 + x + \frac{x^2}{2!} + \frac{x^2}{3!} + \dots$

Quotientenkriterium: $\frac{a_{\nu+1}}{a_{\nu}} = \frac{\frac{x^{\nu+1}}{(\nu+1)!}}{\frac{x^{\nu}}{\nu!}} = \frac{x}{\nu+1} \rightarrow 0$ für $\nu \rightarrow \infty \Rightarrow$ konvergiert für alle x

Beispiel: $\sum_{\nu=1}^{\infty} \frac{1}{\nu}$:

Quotientenkriterium: $\frac{a_{\nu+1}}{a_{\nu}} = \frac{\nu}{\nu+1} \rightarrow 1$, Kriterium **nicht anwendbar** da $q = 1$, d.h. $|\frac{a_{\nu+1}}{a_{\nu}}| < 1$ genügt **nicht** für Konvergenz.

D.) Potenzreihen

Falls bei einer Reihe $\sum_{\nu=0}^{\infty} b_{\nu}$ die Koeffizienten Funktionen $b_{\nu} = f_{\nu}(x)$ sind, dann heißt die Reihe **Funktionsreihe**.

Wichtigstes Beispiel:

$$f_{\nu}(x) = a_{\nu} x^{\nu}$$

Eine solche Reihe

$$\sum_{\nu=0}^{\infty} a_{\nu} x^{\nu} = a_0 + a_1 x + a_2 x^2 + \dots$$

heißt **Potenzreihe**.

Alle ihre Teilsummen sind Polynome.

Konvergenz:

$$\frac{b_{\nu+1}}{b_{\nu}} = \frac{a_{\nu+1} x^{\nu+1}}{a_{\nu} x^{\nu}} = x \frac{a_{\nu+1}}{a_{\nu}}$$

hinreichend: $\lim_{\nu \rightarrow \infty} |x| \left| \frac{a_{\nu+1}}{a_{\nu}} \right| < 1$

äquivalent: $|x| \lim_{\nu} \left| \frac{a_{\nu+1}}{a_{\nu}} \right| < 1$

D.h. für alle x mit $|x| < \lim_{\nu \rightarrow \infty} \left| \frac{a_{\nu}}{a_{\nu+1}} \right|$ konvergiert Potenzreihe absolut.

Definition: $\rho := \lim_{\nu \rightarrow \infty} \left| \frac{a_{\nu}}{a_{\nu+1}} \right|$ heißt **Konvergenzradius** der Potenzreihe (falls $a_{\nu} \neq 0$).

Zusammen:

$$\begin{aligned} |x| < \rho &\Rightarrow \text{absolute Konvergenz} \\ |x| > \rho &\Rightarrow \text{Divergenz} \end{aligned}$$

Der Konvergenzradius ρ kann auch 0 oder ∞ sein;

Fall $\rho = 0$: Konvergenz nur für $x = 0$

Fall $\rho = \infty$: Konvergenz für alle x

Für $|x| = \rho \neq 0$ Divergenz oder Konvergenz möglich; dann gibt es keine allgemeine Aussage.

allgemeine Potenzreihe: $\sum_{\nu}^{\infty} a_{\nu} (x - \alpha)^{\nu}$

Konvergenzbereich $|x - \alpha| < \rho$ (Inneres eines Kreises um $x = \alpha$ mit Radius ρ)

E.) Die Taylor-Formel

Approximation von Funktionen durch Polynome

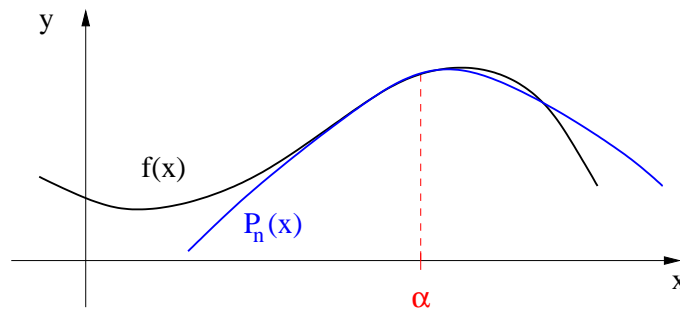
(Approximation = Näherung)

gegeben: Funktion $f(x)$ auf Intervall.

gesucht: Polynom $a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n =: P_n(x)$ mit Koeffizienten a_i derart, dass auf dem Intervall gilt $f(x) \approx P_n(x)$.

Ziel: Das Polynom soll für $x = \alpha$ mit f in allen Ableitungen bis zur n -ten Ordnung übereinstimmen. D.h.

$$\frac{d^i}{dx^i} P_n(\alpha) = \frac{d^i}{dx^i} f(\alpha) \quad \text{für } i = 0, 1, \dots, n \quad (*)$$



Lösung: Taylor-Polynom:

$$\begin{aligned} P_n(x) &= \sum_{k=0}^n \frac{f^{(k)}(\alpha)}{k!} (x - \alpha)^k \\ &= f(\alpha) + \frac{f'(\alpha)}{1!} (x - \alpha) + \frac{f''(\alpha)}{2!} (x - \alpha)^2 + \dots + \frac{f^{(n)}(\alpha)}{n!} (x - \alpha)^n \end{aligned}$$

Probe: (*) stimmt. $(0! := 1)$

Fehler beim Taylor-Polynom

Frage: Wie groß ist der Fehler $f(x) - P_n(x)$?

Bezeichnung: $R_n(x, \alpha) := f(x) - P_n(x)$ heißt **Restglied**.

Es gilt für $(n + 1)$ -mal stetig differenzierbare Funktionen f :

$$R_n(x, \alpha) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - \alpha)^{n+1}$$

für ξ zwischen x und α

Folgerung: Der Fehler $f(x) - P_n(x)$ ist umso kleiner, je näher x an α liegt.

Beispiel:

$$f(x) = e^x, \quad \alpha = 0. \quad \Rightarrow \quad f^{(i)}(\alpha) = 1 \quad \forall i \quad \text{und:}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!} x^{n+1}$$

F.) Taylor-Entwicklung

Annahme: $f(x)$ für $x = \alpha$ beliebig oft stetig differenzierbar. \Rightarrow Taylor-Formel für beliebige n gültig.

Es gilt:

$$\text{Restglied } R_n(x, \alpha) \longrightarrow 0 \text{ für } n \rightarrow \infty \text{ und } |x - \alpha| < \rho$$

\Updownarrow

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!} (x - \alpha)^k \text{ konvergiert gegen } f(x) \text{ für } |x - \alpha| < \rho$$

Definition: Die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!} (x - \alpha)^k$$

heißt **Taylorreihe** von $f(x)$ mit **Entwicklungspunkt** $x = \alpha$.

Man sagt, $f(x)$ läßt sich um (den Entwicklungspunkt) α **in eine Taylorreihe entwickeln**.

oder: $f(x)$ wird in Potenzreihe entwickelt. Wird kein Entwicklungspunkt α angegeben,

so ist meist $\alpha = 0$ gemeint: $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$.

Beispiel: $f(x) = \sin x$ in Potenzreihe, $\alpha = 0$

$$\begin{array}{ll} f'(x) = \cos x & f'(0) = 1 \\ f''(x) = -\sin x & f''(0) = 0 \\ f'''(x) = -\cos x & f'''(0) = -1 \\ f^{(4)}(x) = \sin x & f^{(4)}(0) = 0 \end{array}$$

(danach periodisch)

allgemein für $f(x) = \sin x$: $f^{(2k)}(0) = 0$; $f^{(2k+1)}(0) = (-1)^k$

$$|R_n| \leq \frac{1}{(n+1)!} x^{n+1} \longrightarrow 0 \quad \text{für } n \rightarrow \infty \quad \text{für beliebige } x$$

Zusammen:

$$\sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} \pm \dots \quad \text{für alle } x$$

3.12 Funktionen von mehreren Veränderlichen

$f(x_1, x_2, \dots, x_n)$ ist eine reellwertige Funktion von n Veränderlichen x_1, x_2, \dots, x_n . Der Definitionsbereich ist der \mathbb{R}^n oder ein Teil vom \mathbb{R}^n .

Beispiel: $f(x_1, x_2)$ kann gedeutet werden als Höhe über der (x_1, x_2) -Ebene. (Illustration: Höhenlinien $f(x_1, x_2) = c$)

Schreibweise x für (x_1, \dots, x_n)

Partielle Ableitung erster Ordnung

$$\frac{\partial f(x)}{\partial x_k} = \frac{\partial f(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_n)}{\partial x_k}$$

entspricht der gewöhnlichen ersten Ableitung von

$$g(t) := f(x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_n),$$

wobei beim Differenzieren nach x_k die anderen Veränderlichen $x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n$ als konstant angesehen werden.

Beispiel:

$$f(x_1, x_2) = x_1 x_2^2$$

$$\implies \frac{\partial f}{\partial x_1} = x_2^2 \quad , \quad \frac{\partial f}{\partial x_2} = x_1 2x_2$$

Definition: Der Vektor aller 1. partiellen Ableitungen heißt **Gradient**:

$$\text{grad}f(x) := \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)$$

(als Spaltenvektor)

Maximum oder Minimum von f :

Notwendig für Extremum im Inneren des Definitionsbereiches von f ist

$$\text{grad}f(x) = 0 .$$

Diese Vektorgleichung ist ein System von n Gleichungen.

(Analogie zu $f'(x) = 0$ in Abschnitt 3.8G)

Analog zum eindimensionalen Fall gibt es Taylorentwicklungen.

Beispiel: Kondition (zu Abschnitt 1.5 G)

$$\begin{aligned} \Delta f &\doteq \frac{\partial f(x)}{\partial x_1} \Delta x_1 + \dots + \frac{\partial f(x)}{\partial x_n} \Delta x_n \\ &= (\text{grad}f(x))^t \Delta x \end{aligned}$$

wobei Δx der Vektor mit den Komponenten $\Delta x_1, \dots, \Delta x_n$ ist.

Partielle Ableitungen höherer Ordnung: analog!

Bezeichnungen z.B.

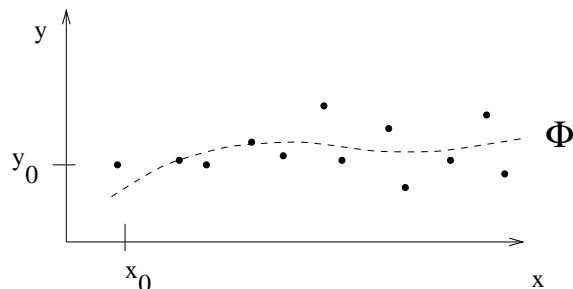
$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} \text{ für die zweite partielle Ableitung nach } x_1$$

Anwendung: Abschnitt 6.2

Kapitel 4 Approximation mit Kurven

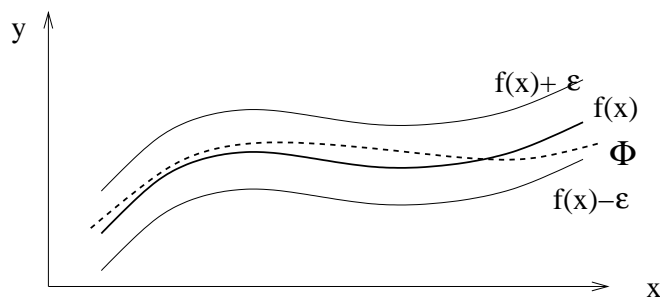
4.1 Approximation und Interpolation

Problem 1: gegeben: Punkte / Zahlenpaare (x_i, y_i) , $i = 0, 1, \dots, n$ (z.B. aus Messungen oder Berechnungen)



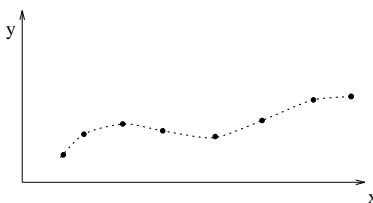
Frage: Welche Funktion $\Phi(x)$ passt sich “am besten” den Punkten an?

Problem 2: gegeben: Funktion $f(x)$, die nicht elementar auswertbar ist (z.B. $\sin x$)



Aufgabe: Finde (einfachere) Ersatzfunktion Φ , so dass $\Phi(x) \approx f(x)$, d.h. $|\Phi - f| < \epsilon$ für eine vorgegebene Fehlerschranke ϵ

Problem 3: gegeben: Punkte (x_i, y_i) . Aufgabe: Φ soll die Punkte **interpolieren**, d.h. $\Phi(x_i) = y_i$ für alle i .



Problem 4: gegeben: Integrationsproblem (“Quadratur”)

$$\int_a^b f(x) dx$$

Gesucht: $\Phi(x)$, so dass

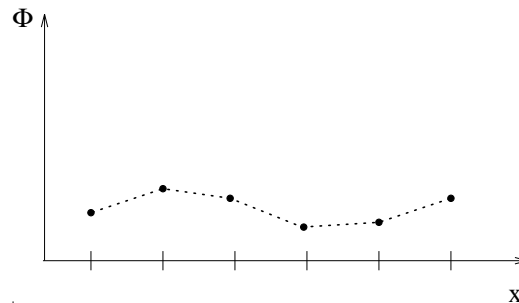
- 1.) $\int_a^b \Phi(x) dx \approx \int_a^b f(x) dx$ (Genauigkeit)
- 2.) $\int_a^b \Phi(x) dx$ einfach auszuwerten (Effizienz)

Problem 5: Φ soll Lösungsfunktionen von Differentialgleichungen approximieren \rightarrow Literatur

Das Resultat einer Approximation hängt entscheidend von dem Typ von Funktion $\Phi(x)$ ab, den wir zulassen werden ("Funktionenklasse").

Beispiele für Funktionenklassen:

- 1.) Polynome $\Phi(x) = p_n(x) = a_0 + a_1x + \dots + a_nx^n$ mit Koeffizienten a_0, \dots, a_n
(Beispiele: Taylor-Polynome)
- 2.) Streckenzug



(d.h. intervallweise Polynome 1. Grades)

- 3.) trigonometrische Funktionen

$$\Phi(x) = a_0 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$$

(Beispiel: Fourier-Reihen)

4.2 Interpolation mit Polynomen

zugelassene Funktionenklasse für Φ

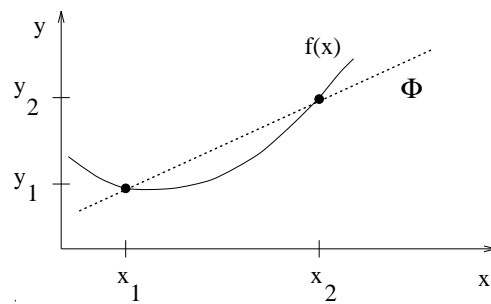
$$p_n(x) = a_0 + a_1x + \dots + a_nx^n$$

$$\begin{array}{ll} \text{speziell } n = 1 : \text{ Gerade} & a_0 + a_1x \\ n = 2 : \text{ Parabel} & a_0 + a_1x + a_2x^2 \end{array}$$

A. Lineare Interpolation

nur Geraden zugelassen: $\Phi = a_0 + a_1x$

gegeben: 2 "Stützpunkte" $(x_1, y_1), (x_2, y_2)$.



Lösung ist

$$\Phi(x) = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

$$= \frac{f(x_1)(x_2 - x) + f(x_2)(x - x_1)}{x_2 - x_1}$$

(früher häufig angewandt, um Zwischenwerte bei Tabellen zu approximieren.)

klassische “*Interpolation*”: für ein x mit $x_1 < x < x_2$ } berechne $\Phi(x)$
 “*Extrapolation*”: für ein x außerhalb } als Näherung
 zu $f(x)$

B. Lagrange-Polynome

Gegeben: $n + 1$ Stützpunkte

$(x_0, y_0), \dots, (x_n, y_n)$, x_i verschieden, d.h. $x_i \neq x_j$ für $i \neq j$

Definition: *Lagrange-Polynom*:

$$L_k(x) := \frac{(x - x_0) \cdot \dots \cdot (x - x_{k-1}) \cdot (x - x_{k+1}) \cdot \dots \cdot (x - x_n)}{(x_k - x_0) \cdot \dots \cdot (x_k - x_{k-1}) \cdot (x_k - x_{k+1}) \cdot \dots \cdot (x_k - x_n)}$$

Eigenschaften:

$$L_k(x_k) = 1$$

$$L_k(x_i) = 0 \text{ für } i \neq k$$

Polynom n -ten Grades

Folgerung:

$$\Phi(x) := L_0(x)y_0 + L_1(x)y_1 + \dots + L_n(x)y_n$$

interpoliert: $\Phi(x_k) = y_k$ für $k = 0, \dots, n$

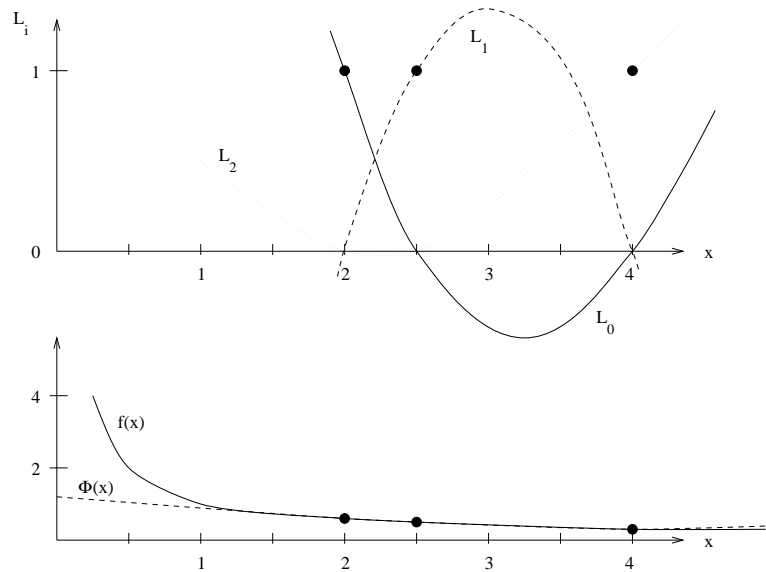
Beispiel: $x_0 = 2$, $x_1 = 2.5$, $x_2 = 4$, $y_i = \frac{1}{x_i}$

$$L_0(x) = \frac{(x - 2.5)(x - 4)}{(2 - 2.5)(2 - 4)} = x^2 - 6.5x + 10$$

$$L_1(x) = -\frac{4}{3}(x^2 - 6x + 8)$$

$$L_2(x) = \frac{1}{3}\left(x^2 - \frac{9}{2}x + 5\right)$$

$$\begin{aligned}\Phi(x) &= \frac{1}{2}L_0 + \frac{2}{5}L_1 + \frac{1}{4}L_2 = \\ &= 0.05x^2 - 0.425x + 1.15\end{aligned}$$



C. Eigenschaften

- 1.) Es seien x_0, x_1, \dots, x_n voneinander verschieden. Dann **gibt es genau ein** Polynom p_n vom **Grad** $\leq n$, das interpoliert: $p_n(x_i) = y_i$ (Beweis: Annahme es gibt zwei...)
- 2.) **Fehler:**
Annahme: Die Punkte "liegen auf" f (d.h. $f(x_i) = y_i \quad \forall i$).
Frage: Wie "weit" ist p_n von f entfernt?

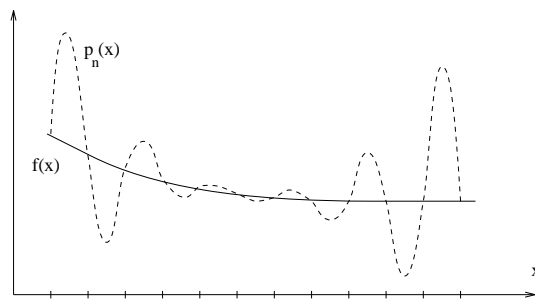
$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)(x - x_1) \cdot \dots \cdot (x - x_n)$$

hierbei ist ξ im Intervall, welches x, x_0, \dots, x_n umfasst. Formel brauchbar, wenn $f^{(n+1)}$ wenigstens in der Größenordnung bekannt ist oder abschätzbar: $|f^{(n+1)}(\xi)| \leq C$. (Beispiel $f(x) = \sin x \rightarrow C = 1$).

- 3.) gute Eigenschaften: einfache Theorie und Konstruktion

sehr schlechte Eigenschaften:

- 4.) Außerhalb von x_0, \dots, x_n steigt Fehler stark an \Rightarrow Extrapolation ist i.a. nicht sinnvoll.
- 5.) n groß: schlecht konditioniert ($p_n(x)$ hängt empfindlich von Werten y_i ab.)
- 6.) n groß und (beliebig) vorgegebene (z.B. äquidistante) x_i : starke Oszillationen zum Rand des Intervalls



(Abhilfe möglich, wenn x_i "optimal" gelegt werden können, d.h. zunehmend dichter zu den "Rändern" x_0 und x_n . → *Chebyshev-nodes*, Čebyšev-Polynom)

Empfehlung:

Benutze Polynom-Interpolation nur bei kleinem Grad! (sagen wir $n < 6$), d.h. bei **wenig** Stützpunkten.

D. Rekursionsformel von Neville

zur Auswertung eines Interpolations-Polynoms $p_n(x)$ für einen vorgegebenen Wert von x .

Ziel/Idee: Konstruiere Polynom höheren Grades billig aus Polynomen niederen Grades.

Es sei $Q_{i,j}(x)$ das Polynom j -ten Grades (für $i \geq j$), welches die Stützpunkte $(x_{i-j}, y_{i-j}), \dots, (x_i, y_i)$ interpoliert.

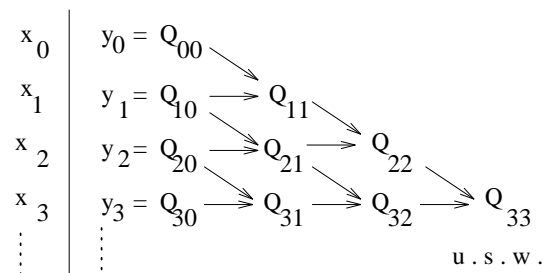
Dann gilt die **Rekursion**

$$Q_{i,0} := y_i \quad (\text{Startwerte}), \quad i = 0, 1, 2, \dots$$

$$Q_{i,j}(x) := Q_{i,j-1} + \frac{Q_{i,j-1} - Q_{i-1,j-1}}{\frac{x-x_{i-j}}{x-x_i} - 1}$$

für $1 \leq j \leq i$, und $i = 1, 2, \dots$

Schema/Anordnung im Tableau:



Beispiel: $(x_i, y_i) = (0, 1), (1, 3), (3, 2)$. Wert des Interpolations-Polynomes für $x = 2$?

i	x_i	y_i
0	0	$1 = Q_{00}$
1	1	$3 = Q_{10}$
2	3	$2 = Q_{20}$

(die Nenner in Kreisen)

Praxis: Für anderes x von vorne beginnen. Nur die letzte Zeile des Tableaus speichern. Stützstellen brauchen nicht geordnet zu sein, Hinzunahme weiterer Stützpunkte/Zeilen problemlos.

E. Konstruktion der Koeffizienten

Anstatt $p_n(x) = a_0 + a_1x + \dots + a_nx^n$ besser der Ansatz von Newton:

$$p_n(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) + \dots + b_n(x - x_0)(x - x_1) \cdot \dots \cdot (x - x_{n-1})$$

Vorteil gegenüber dem bisherigen Ansatz $p_n(x) = a_0 + a_1x + \dots + a_nx^n$:
 Koeffizienten b_i folgen mit geringerem Aufwand:

$$\begin{aligned}
 y_0 = p_n(x_0) &= b_0 + 0 && \rightarrow b_0 = y_0 \\
 y_1 = p_n(x_1) &= b_0 + b_1(x_1 - x_0) + 0 && \rightarrow b_1 = \frac{y_1 - y_0}{x_1 - x_0} \\
 y_2 = p_n(x_2) &= b_0 + b_1(x_2 - x_0) + b_2(x_2 - x_0)(x_2 - x_1) && \rightarrow b_2 = \frac{\dots}{\dots} \\
 &\vdots &&
 \end{aligned}$$

jeweils nur **eine** Unbekannte pro Gleichung, rekursives Vorgehen!

Wegen der Eindeutigkeit der Interpolation sind beide Interpolations-Polynome **identisch**. (i.a. $a_i \neq b_i$!)

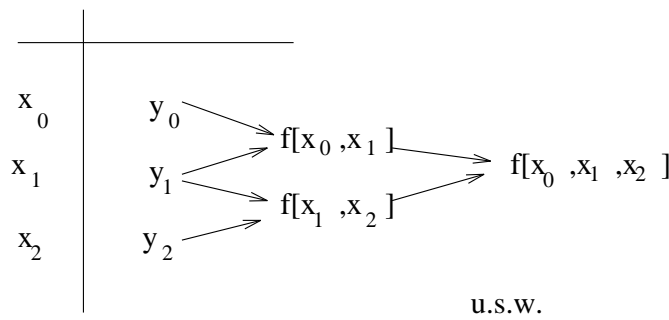
Rekursion und Definition für “dividierte Differenzen”:

$$\begin{aligned}
 f[x_i] &:= y_i && \text{(Startwerte)} \\
 f[x_i, x_{i+1}, \dots, x_{i+k-1}, x_{i+k}] &:= \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}
 \end{aligned}$$

Damit gilt:

$$\begin{aligned}
 b_0 &= f[x_0], & b_1 &= f[x_0, x_1] && \text{(s.o.),} \\
 b_2 &= f[x_0, x_1, x_2], \dots, & b_n &= f[x_0, x_1, \dots, x_n]
 \end{aligned}$$

Tableau für einfaches Rechnen:



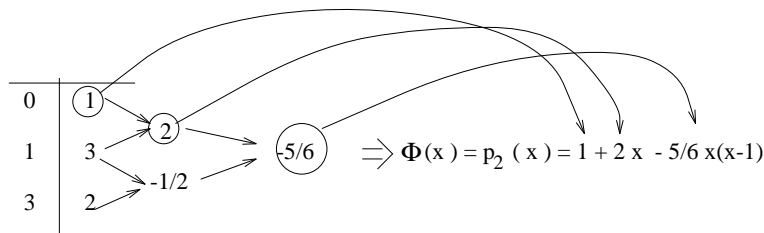
d.h.: Koeffizienten b_i befinden sich in der oberen Diagonale.

Anwendung der Rekursion: Die Koeffizienten des Polynoms werden benötigt, wenn z.B. ein Interpolations-Polynom für *viele* x auszuwerten ist.

Auswertung dann mit Horner-Schema:

$$p(x) = (((b_n(x - x_{n-1}) + b_{n-1})(x - x_{n-2}) + \dots + b_1)(x - x_0) + b_0$$

Beispiel: (von oben)



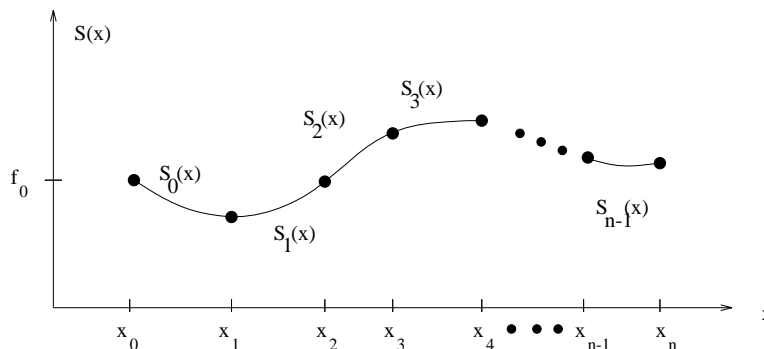
4.3 Interpolation mit Splines

Polynominterpolation von § 4.2: brauchbar nur bei geringer Anzahl von Punkten.

gesucht: Interpolation, auch wenn *viele* Punkte $(x_0, f_0), \dots, (x_n, f_n)$ vorgegeben sind!

A. Definition des Splines

Idee:



Berechne in jedem Teilintervall ein eigenes Polynom derart, dass die Polynom-Stücke an den "Nahtstellen" x_j möglichst **glatt** zusammenstoßen.

Wähle Polynome von jeweils 3. Grad, d.h.

Ansatz für “kubischen Spline”:

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3$$

für $x_j \leq x < x_{j+1}$ und $j = 0, 1, \dots, n - 1$

In jedem Teilintervall benötige a_j, b_j, c_j, d_j , also insgesamt $4n$ Unbekannte. Der Gesamt-Spline $S(x)$ besteht stückweise aus $S_0(x), S_1(x), \dots, S_{n-1}(x)$.

Forderungen an $S(x)$:

		Anzahl der Bedingungen
(1) interpoliert: $j = 0, 1, \dots, n - 1$	$S_j(x_j) = f_j$, dazu $S_{n-1}(x_n) = f_n$	$n + 1$
(2) ist stetig: ($j = 1, \dots, n - 1$)	$S_j(x_j) = S_{j-1}(x_j)$	$n - 1$
(3) 1. Ableitung ist stetig:	$S'_j(x_j) = S'_{j-1}(x_j)$	$n - 1$
(4) 2. Ableitung ist stetig:	$S''_j(x_j) = S''_{j-1}(x_j)$	$n - 1$
		$4n - 2$

Brauche $4n$ Gleichungen; noch 2 Bedingungen frei.

Wichtige Beispiele für zwei noch fehlende “Randbedingungen”:

(5a) *freier Rand*: $S''_0(x_0) = 0, S''_{n-1}(x_n) = 0$, “natürlicher Spline” (Mechanik: Krümmung an Endpunkten = 0)

oder

(5b) *eingespannter Rand*: $S'_0(x_0) = f'_0, S'_{n-1}(x_n) = f'_n$ für vorgeschriebene Werte f'_0, f'_n ; eventuell vorher Näherungen für f'_0, f'_n berechnen. (Z.B. am linken Rand: Ein “kurzes” Polynom $P_{\text{links}}(x)$ ist so zu berechnen, dass es die “linkesten” drei oder vier Punkte interpoliert. Dann $y'_0 := P'_{\text{links}}(x_0)$. Analog am rechten Rand.)

oder

(5c) *periodische Bedingungen*: (Vor.: $f_0 = f_n$): $S'_0(x_0) = S'_{n-1}(x_n), S''_0(x_0) = S''_{n-1}(x_n)$

oder

(5d) *not a knot*: Stetigkeit von S''' an x_1 und x_{n-1} .

Die obigen $4n$ Bedingungen (1)–(5) definieren eindeutig die Konstanten a_j, b_j, c_j, d_j , und damit den kubischen Spline.

B. Berechnung des kubischen Splines

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3$$

Interpolation (1) $\Rightarrow S_j(x_j) = a_j = f_j$, d.h. a_j sind bekannt!

Bezeichnung: $h_j := x_j - x_{j-1}$

$$(2) \Rightarrow f_j = S_j(x_j) = S_{j-1}(x_j) = f_{j-1} + b_{j-1}h_j + c_{j-1}h_j^2 + d_{j-1}h_j^3 \quad (2')$$

$$(3) \Rightarrow S'_j(x_j) = b_j = S'_{j-1}(x_j) = b_{j-1} + 2c_{j-1}h_j + 3d_{j-1}h_j^2 \quad (3')$$

$$(4) \Rightarrow c_j = c_{j-1} + 3d_{j-1}h_j \quad (\text{gekürzt mit } 2) \quad (4')$$

Elimination der d_j aus (4') und b_j führt auf

$$\boxed{h_j c_{j-1} + 2(h_j + h_{j+1})c_j + h_{j+1}c_{j+1} = 3\frac{f_{j+1} - f_j}{h_{j+1}} - 3\frac{f_j - f_{j-1}}{h_j}} \quad (*)$$

für $j = 1, 2, \dots, n - 1$

Dies ist ein **System von $n - 1$ linearen Gleichungen** für die $n + 1$ Unbekannten

$$c_j = \frac{S_j''(x_j)}{2} \quad (\text{wird gleich wieder mit 2 multipliziert})$$

Die Randbedingungen können durch zwei zusätzliche Gleichungen definiert werden.

Abkürzungen: für $j = 1, 2, \dots, n - 1$

$$M_j := S_j''(x_j)$$

$$\lambda_j := \frac{h_{j+1}}{h_j + h_{j+1}}$$

$$\mu_j := \frac{h_j}{h_j + h_{j+1}} = 1 - \lambda_j$$

$$r_j := \frac{6}{h_j + h_{j+1}} \cdot \left(\frac{f_{j+1} - f_j}{h_{j+1}} - \frac{f_j - f_{j-1}}{h_j} \right)$$

und für Fall (5a) ($c_0 = c_n = M_0 := M_n := 0$) zusätzlich :

$$\lambda_0 := 0, \quad r_0 := 0, \quad \mu_n := 0, \quad r_n := 0$$

Damit lautet (*) mit Randbedingungen in Matrix-Schreibweise:

$$\begin{pmatrix} 2 & \lambda_0 & & & 0 \\ \mu_1 & 2 & \lambda_1 & & \\ & \ddots & \ddots & \ddots & \\ & & \mu_{n-1} & 2 & \lambda_{n-1} \\ 0 & & & \mu_n & 2 \end{pmatrix} \begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_{n-1} \\ M_n \end{pmatrix} = \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_{n-1} \\ r_n \end{pmatrix}$$

Die Matrix ist **tridiagonal**, die Lösung mit dem Gaußschen Verfahren ist hier extrem billig (Aufwand $O(n)$ statt $O(n^3)$).

Algorithmus (natürlicher kubischer Spline):

Eingabe: $(x_0, f_0), (x_1, f_1), \dots, (x_n, f_n)$

berechne $h_j, \lambda_j, \mu_j, r_j$

löse Gleichungssystem mit Vorwärts-/Rückwärts-Schleife, ergibt M_j

Koeffizienten aus

$$a_j = y_j = f_j$$

$$c_j = M_j/2$$

$$d_j = \frac{M_{j+1} - M_j}{6h_{j+1}} \quad (\text{aus (4')})$$

$$b_j = \frac{f_{j+1} - f_j}{h_{j+1}} - \frac{h_{j+1}}{6}(2M_j + M_{j+1}) \quad (\text{aus (2')})$$

Hinweise:

- 1.) keine Matrix speichern, nur Vektoren!
- 2.) Für (5a) hätte ein System der Größe $n - 1$ ausgereicht, weil $c_0 = c_n = 0$; in obigem größeren Rahmen lässt sich mit anderen $\lambda_0, \mu_n, r_0, r_n$ auch (5b) leicht realisieren.
- 3.) Aus der Struktur der Matrix folgt ihre Nichtsingularität. Es folgt, dass der Spline eindeutig definiert ist.

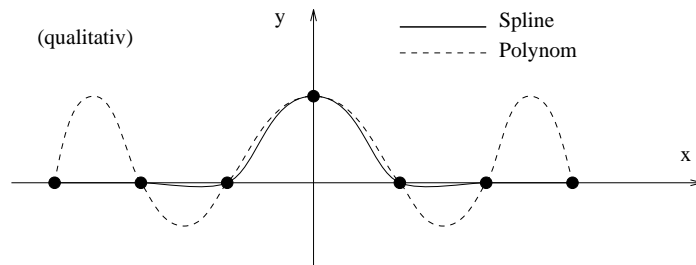
C. Weitere Eigenschaften

Glattheit: nach Konstruktion $S \in C^2[x_0, x_n]$, und glatt auch im folgenden Sinn: Es gilt

$$\int_{x_0}^{x_n} (S''(x))^2 dx \leq \int_{x_0}^{x_n} (f''(x))^2 dx$$

für **alle** beliebigen anderen interpolierenden Funktionen $f(x)$, welche 2-mal stetig differenzierbar sind *und* die Randbedingung erfüllen.

Folgerung: Splines haben geringe "Krümmung" und oszillieren deswegen relativ wenig.



Ergänzungen zur Vorlesung:

- 1.) Vorwärts- / Rückwärtsschleife:

$$\left\{ \begin{array}{l} \delta_i \text{ "neue" Diagonalelemente, } \delta_0 := 2, \tilde{r}_0 := r_0, \tilde{r}_i \text{ "neue" rechte Seite:} \\ \text{für } i = 0, 1, \dots, n-1: \\ \delta_{i+1} = 2 - \lambda_i \frac{\mu_{i+1}}{\delta_i}, \quad \tilde{r}_{i+1} = r_{i+1} - \tilde{r}_i \frac{\mu_{i+1}}{\delta_i} \\ M_n := \frac{\tilde{r}_n}{\delta_n} \\ \text{für } k = n-1, n-2, \dots, 0: \\ M_k = \frac{1}{\delta_k} (\tilde{r}_k - \lambda_k M_{k+1}) \end{array} \right.$$

- 2.) eingespannter Rand:

$$\lambda_0 := 1, \quad r_0 := \frac{6}{h_1} \left(\frac{f_1 - f_0}{h_1} - f'_0 \right),$$

$$\mu_n := 1, \quad r_n := \frac{6}{h_n} \left(f'_n - \frac{f_n - f_{n-1}}{h_n} \right)$$

4.4 Bézier-Kurven

1962 Bézier (Renault); Casteljau (Citroën)

Bernstein-Polynome: für allgemeines Intervall $a \leq x \leq b$:

$$B_i^n(x; a, b) := \binom{n}{i} \frac{(b-x)^{n-i}(x-a)^i}{(b-a)^n}$$

für $i = 0, 1, \dots, n$ ist i -tes Bernstein-Polynom vom Grad n .

(Binomialkoeffizient $\binom{n}{i} = \frac{n!}{i!(n-i)!}$)

oder speziell für $0 \leq t \leq 1$:

$$B_i^n(t) := B_i^n(t; 0, 1) = \binom{n}{i} (1-t)^{n-i} t^i$$

Beispiel $n = 2$:

$$B_0^2(t) = (1-t)^2$$

$$B_1^2(t) = 2(1-t)t$$

$$B_2^2(t) = t^2$$

Eigenschaften für $B_i^n(t)$:

- (1) $\sum_{i=0}^n B_i^n(t) = 1$ ("Zerlegung der 1")
- (2) $t = 0$ ist i -fache Nullstelle ($i > 0$), $t = 1$ ist $(n-i)$ -fache Nullstelle ($i < n$)
- (3) $B_i^n(t) \geq 0$ für $0 \leq t \leq 1$
- (4) $B_i^n(t)$ hat genau ein Maximum, bei $t = \frac{i}{n}$
- (5) Rekursion: $B_i^n(t) = tB_{i-1}^{n-1}(t) + (1-t)B_i^{n-1}(t)$

Vektorwertiges Polynom: Es seien a_0, a_1, \dots, a_n Vektoren im \mathbb{R}^d (CAD: $d = 2$ oder $d = 3$), dann ist $P(t) := \sum_{i=0}^n a_i t^i$ ein vektorwertiges Polynom ("polynomiale Kurve"). Für z.B. $0 \leq t \leq 1$ beschreibt $P(t)$ ein Kurvenstück im \mathbb{R}^d , dabei ist t der Parameter der Kurve.

Es seien p_0, p_1, \dots, p_n Punkte im \mathbb{R}^d .

Bernstein-Bézier-Kurven:

$$P(t) := \sum_{i=0}^n p_i B_i^n(t)$$

Die Punkte p_0, p_1, \dots, p_n heißen Bézier-Punkte oder **Kontrollpunkte**. Der Streckenzug $p_0 \leftrightarrow p_1 \leftrightarrow \dots \leftrightarrow p_n$ heißt **B-B-Polygon**.

Eigenschaften der B-B-Kurven:

- (a) $P(0) = p_0, \quad P(1) = p_n$

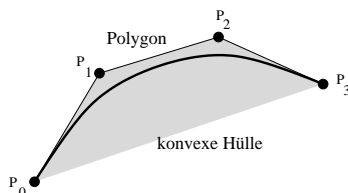
- (b) In den Endpunkten sind die B-B-Kurven tangential zum Polygon durch die Kontrollpunkte. (zeige hierzu $P'(0)$ =Steigung der Gerade $p_0 \rightarrow p_1$, analog $P'(1)$)
- (c) Die Kurve liegt in der durch p_0, p_1, \dots, p_n aufgespannten *konvexen Hülle*

$$\left\{ \sum_{i=0}^n s_i p_i \text{ mit } s_i \geq 0, \sum_{i=0}^n s_i = 1 \right\}$$

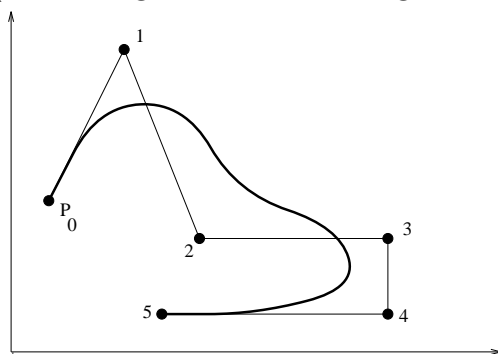
(folgt mit $s_i = B_i^n(t)$ aus (3) und (1) "Zerlegung der 1".) [konvexe Menge: Mit je zwei Punkten ist auch ihre Geraden-Verbindung enthalten.]

- (d) Die Kurve geht "relativ nahe" an den Kontrollpunkten vorbei. (folgt aus (4): für $t = i/n$ ist $s_i = B_i^n$ maximal, die anderen s_k (für $k \neq i$) eher klein)

Beispiel (im \mathbf{R}^2): (Reihenfolge der Numerierung ist wichtig!)



Beispiel (qualitativ): (Reihenfolge der Numerierung ist wichtig!)



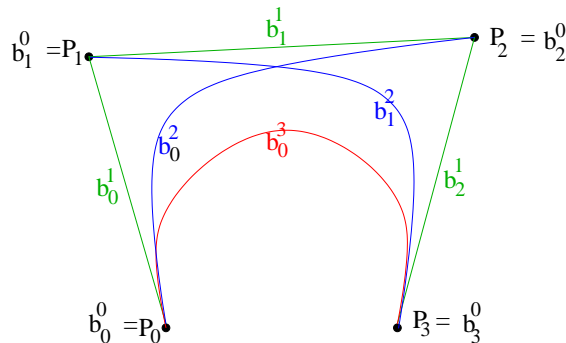
Berechnung: erfolgt rekursiv. Hierzu werden die vektorwertigen **Teilpolynome** b_i^k definiert:

$$b_i^k(t) := \sum_{j=0}^k p_{i+j} B_j^k(t).$$

Es gehen die Punkte $p_i, p_{i+1}, \dots, p_{i+k}$ ein. Man zeigt leicht

$$b_i^k(t) = (1-t)b_i^{k-1} + t b_{i+1}^{k-1} \quad \text{und} \quad b_0^n(t) = P(t).$$

Anordnung der b_i^k in Tableau (Einträge sind Vektoren). Literatur: Algorithmus von Casteljau, startet mit $b_i^0 = p_i$.



Hinweis: Mit **B-Splines** können ähnliche Kurven erzeugt werden. (“B” ← **B**asis)

4.5 Integration mit Trapezsummen; Extrapolation

Aufgabe: Berechne $\int_a^b f(x)dx$, wobei a, b und f gegeben, f genügend glatt auf $a \leq x \leq b$.

1. Idee: Approximiere $f(x)$ durch interpolierendes Polynom $p_n(x)$. Dann

$$\begin{aligned} \int_a^b f(x)dx &\approx \int_a^b p_n(x)dx = \int_a^b \sum_{i=0}^n f(x_i)L_i(x)dx = \\ &= \sum_{i=0}^n f(x_i) \underbrace{\int_a^b L_i(x)dx}_{c_i} \end{aligned}$$

hängen nicht von f ab.

Falls x_i äquidistant mit

$$h = x_{j+1} - x_j = \frac{b-a}{n}, \quad x_j := a + jh; \quad \text{dann } c_i = hc_i$$

(L_i : Lagrange Polynome; Letzteres wird mit der Substitution $x = a + ht$ gezeigt.)

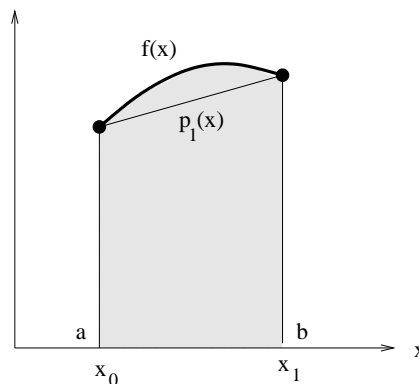
Die “Gewichte” c_i können in der Literatur gefunden werden, bzw. in den folgenden Beispielen.

Zusammenfassung: Es gelten die *Newton-Cotes-Formeln*:

$$\int_a^b f(x)dx \approx \int_a^b p_n(x)dx = h \sum_{i=0}^n c_i f(x_i)$$

1. Beispiel: Trapezregel

$n = 1$, d.h. ersetze f durch Gerade $p_1(x)$



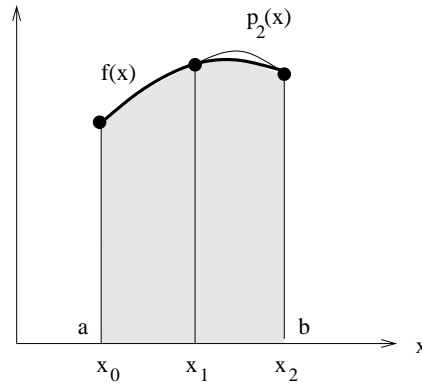
$$c_0 = c_1 = \frac{1}{2}$$

$$\begin{aligned} \int_a^b f(x)dx &\approx \frac{x_1 - x_0}{2} (f(x_0) + f(x_1)) \\ &= \text{Fläche des Trapezes} \end{aligned}$$

2. Beispiel: “Simpsons Regel”, “Keplersche Faßregel”

$n = 2$, d.h. ersetze f durch Parabel p_2

Definiere $h := \frac{b-a}{n} = \frac{b-a}{2}$, $x_0 = a$, $x_1 = \frac{b+a}{2}$, $x_2 = b$



Es gilt: $c_0 = \frac{1}{3}$, $c_1 = \frac{4}{3}$, $c_2 = \frac{1}{3}$. Es folgt:

$$\int_a^b f(x)dx \approx \frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2))$$

= Fläche unter der Parabel

absoluter Fehler:

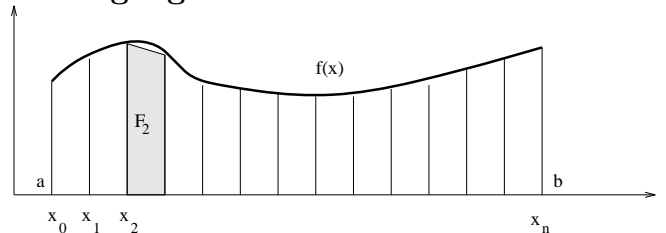
Trapezregel: $\frac{h^3}{12}f''(\xi) = O(h^3)$ für ein ξ im Intervall.

Simpson: $O(h^5)$

Man erhält diese Fehlerordnungen durch Integration des Fehlers $f(x) - p(x)$ (vgl. § 4.2 C) und Anwendung des Mittelwertsatzes.

Genauigkeitssteigerung nicht durch Polynom höherer Ordnung, sondern mit

2. Idee: stückweiser Zugang



$$h := \frac{b-a}{n}, \quad x_i = a + ih$$

Fläche des i -ten Trapezes: $F_i = h \frac{f(a + ih) + f(a + (i + 1)h)}{2}$ führt auf **Trapez-Summe**

$T(h) := F_0 + F_1 + \dots + F_{n-1}$, also:

$$T(h) = h \left[\frac{f(a)}{2} + f(a+h) + f(a+2h) + \dots + f(b-h) + \frac{f(b)}{2} \right]$$

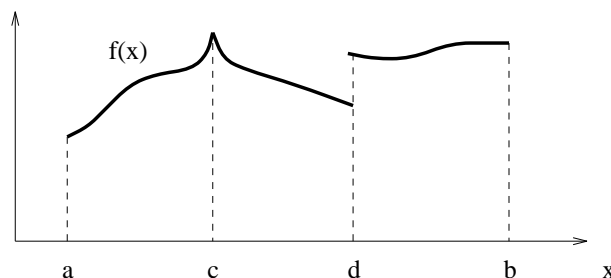
Fehler:

$$T(h) - \int_a^b f(x)dx = \frac{b-a}{12}h^2 f''(\xi) = O(h^2)$$

für $a < \xi < b$ und f zweimal stetig differenzierbar auf $a \leq x \leq b$ (d.h. $f \in C^2[a, b]$)
 d.h. eine Ordnung der Trapezregel geht durch das Aufsummieren verloren.

Hinweis: Wenn $f(x)$ nur stückweise glatt ist, dann aufspalten in Teilintervalle und

getrennt integrieren: $\int_a^b = \int_a^c + \int_c^d + \int_d^b$



3. Idee: (Asymptotische) Entwicklung des Fehlers: Man kann für genügend glattes f beweisen:

$$(*) \left\{ \begin{array}{l} T(h) = \tau_0 + \tau_1 h^2 + \tau_2 h^4 + \dots + \tau_m h^{2m} + R_{m+1} h^{2m+2} \\ \text{wobei } \tau_0 = \int_a^b f(x)dx \text{ der exakte Wert ist,} \\ \tau_1, \tau_2 \dots \text{ von } h \text{ unabhängig} \end{array} \right.$$

(Nachweis vgl. Literatur.)

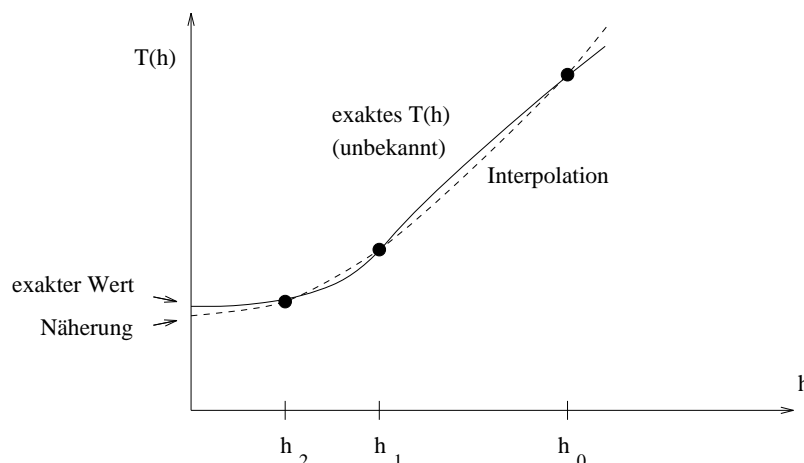
(*) heißt auch "quadratische" Entwicklung, da Polynom bzw. Reihe in h^2 ; Restglied R_{m+1} ist beschränkt; enthält obige Fehlerformel für $m = 0$. Wegen (*) verhält sich $T(h)$ für kleine h wie ein Polynom in h^2 .

4. Idee: Extrapolation (Richardson 1927, Romberg 1955)

Ziel: Berechnung von $T(0) := \lim_{h \rightarrow 0} T(h)$, dies ist der exakte Wert des Integrals. Kann aber wegen $n \rightarrow \infty$ nicht direkt ausgewertet werden (wäre auch nicht effizient).

möglich: $T(h)$ kann ausgewertet werden für einige ausgewählte $h > 0$, z.B. $T(h_0)$, $T(h_1)$, $T(h_2)$, eventuell auch mehr.

Idee: Lege durch $T(h_0)$, $T(h_1)$, ... interpolierendes Polynom $\tilde{T}(h)$ und werte $\tilde{T}(0)$ aus (einfach!), und nehme $\tilde{T}(0)$ als Näherung zu $T(0) = \int_a^b f(x)dx$.



(Da $\tilde{T}(h)$ Polynom in h^2 ist, ist die "Extrapolation" auf $h = 0$ problemlos.)

Algorithmus “Romberg Integration”:

- (1) Lege “Schrittweitenfolge” h_i fest, z.B. Halbierungsfolge $h_0 := b - a, h_1 = \frac{h_0}{2}, \dots, h_i = \frac{h_{i-1}}{2}, \dots$
- (2) Berechne Trapezsummen

$$T_{i0} := T(h_i) \text{ für } i = 0, 1, \dots, m$$

- (3) Mit Neville-Schema, welches sich hier schreiben lässt als

$$T_{ik} := T_{i,k-1} + \frac{T_{i,k-1} - T_{i-1,k-1}}{\left(\frac{h_{i-k}}{h_i}\right)^2 - 1}$$

berechne den Wert $\tilde{T}(0)$ des Interpolationspolynoms für $h = 0$ als Näherung zum Integral.

Praxis: Mit $x := h^2, x_i := h_i^2$ braucht Neville-Programm nicht geändert zu werden.
Eingabe: $h_i^2, T(h_i)$. $[T(h), \text{ nicht } T(h^2)!!]$

Bei Halbierungsfolge rechentechnische Vergünstigungen:

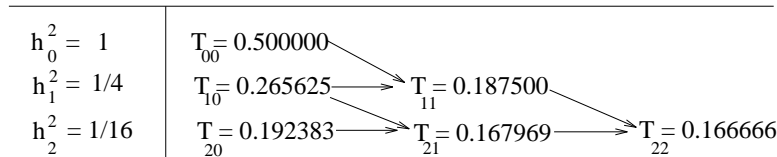
$$T(h_i) = T(h_{i-1}) \cdot \frac{1}{2} + h_i[f(a + h_i) + f(a + 3h_i) + \dots];$$

die Faktoren $\frac{1}{\left(\frac{h_{i-k}}{h_i}\right)^2 - 1}$ sind 1/3, 1/15, 1/63, 1/255 ... (dezimal abzuspeichern)

Beispiel:

$$\int_0^1 x^5 dx, \quad h_0 = 1, \quad h_1 = \frac{1}{2}, \quad h_2 = \frac{1}{4}$$

(6-stellige Rechnung)



Fehler: $O(h_0^2 h_1^2 h_2^2)$

4.6 Diskrete Fourier-Transformation

gegeben: Daten einer Zeitreihe.

gesucht: Struktur, insbesondere zyklische Strukturen. Dazu benötigte Frequenzen mit ihren Amplituten.

A. Grundlagen

\mathbb{C} : komplexe Zahlen. Es sei i die imaginäre Einheit, vgl. Abschnitt 3.2.

Setzt man in der Potenzreihe (vgl. Abschnitt 3.11) für e^x statt des $x \in \mathbb{R}$ die rein imaginäre Veränderliche $i\varphi$ ein, so ergibt sich

$$e^{i\varphi} = \underbrace{\left(1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} \mp \dots\right)}_{\cos \varphi} + i \underbrace{\left(\varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} \mp \dots\right)}_{\sin \varphi}$$

Das ist die **Eulersche Formel**

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

Für $0 \leq \varphi \leq 2\pi$ beschreibt $e^{i\varphi}$ demnach den Einheitskreis in der komplexen Ebene \mathbb{C} .

Wegen den Eigenschaften der Multiplikation in \mathbb{C} (vgl. Abschnitte 3.2, 3.3C) gilt zum Beispiel für die Gleichung

$$(e^{i\varphi})^4 = e^{i4\varphi} = 1,$$

dass $\varphi = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ die Lösungen sind. D.h. mit $\varphi_j := \frac{2\pi}{4}j$ für $j = 0, 1, 2, 3$ sind die $e^{i\varphi_j}$ die vier vierten Einheitswurzeln der komplexen 1.

Allgemein: $e^{i\frac{2\pi}{n}}$ ist n -te Einheitswurzel der komplexen 1.

Abkürzung: $z_j := e^{i\frac{2\pi}{n}j}$, $j = 0, 1, \dots, n-1$, sind alle n -ten Wurzeln der 1.

Für diese z_j gilt:

$$z_j^\nu = z_\nu^j$$

sowie **Orthogonalität:**

$$\sum_{\nu=0}^{n-1} z_\nu^j z_\nu^{-l} = \begin{cases} n & \text{für } l = j \\ 0 & \text{für } l \neq j \end{cases} \text{ für } l, j \in \{0, \dots, n-1\}$$

Beweis der Orthogonalität:

$$z_\nu^n = 1 \text{ für alle } \nu \Rightarrow 0 = z_\nu^n - 1 = \underbrace{(z_\nu - 1)}_{\neq 0 \text{ für } \nu \neq 0, \nu \neq \pm n, \dots} (z_\nu^{n-1} + z_\nu^{n-2} + \dots + z_\nu + 1)$$

$$\Rightarrow \sum_{j=0}^{n-1} z_\nu^j = \begin{cases} n & \text{für } \nu = 0, \pm n, \dots \\ 0 & \text{für } \nu \neq 0, \pm n, \dots \end{cases} \quad (\text{denn dann } z_\nu = 1)$$

$$\Rightarrow \sum_{j=0}^{n-1} z_j^k z_j^{-l} = \sum_{j=0}^{n-1} z_j^{k-l} = \sum_{j=0}^{n-1} z_{k-l}^j = \begin{cases} n & \text{für } k-l = 0 \\ 0 & \text{für } k-l \neq 0 \end{cases}$$

B. Komplexe Diskrete Fourier-Transformation:

Satz: Es seien $g_0, \dots, g_{n-1}, c_0, \dots, c_{n-1} \in \mathbb{C}$. Dann gilt:

$$c_\nu = \frac{1}{n} \sum_{j=0}^{n-1} g_j e^{-i\nu j \frac{2\pi}{n}} \iff g_j = \sum_{\nu=0}^{n-1} c_\nu e^{i\nu j \frac{2\pi}{n}}$$

(1)

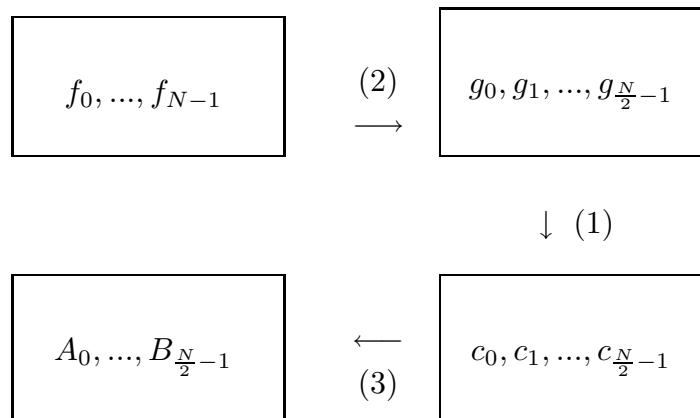
Die Formeln in (1) vermitteln die **komplexe diskrete Fourier-Transformation** zwischen den jeweils n komplexen Zahlen c_ν und g_ν .

Beweis von (1): (eine Richtung \Leftarrow , die andere analog)

$$\begin{aligned}
 \sum_j g_j z_j^{-\nu} &= \sum_{j=0}^{n-1} (c_0 z_0^j + \dots + c_\nu z_\nu^j + \dots + c_{n-1} z_{n-1}^j) \cdot z_j^{-\nu} \\
 &= \sum_j (c_0 z_j^0 + \dots + c_\nu z_j^\nu + \dots + c_{n-1} z_j^{n-1}) z_j^{-\nu} \\
 &= c_0 \underbrace{\sum_j z_j^0 z_j^{-\nu}}_{=0} + \dots + c_\nu \underbrace{\sum_j z_j^\nu z_j^{-\nu}}_{=n} + \dots + c_{n-1} \underbrace{\sum_j z_j^{n-1} z_j^{-\nu}}_{=0} \\
 &= c_\nu \cdot n \quad (\text{wegen Lemma})
 \end{aligned}$$

C. Übergang zwischen reeller und komplexer Version:

Gegeben f_0, f_1, \dots, f_{N-1} reelle Daten (z.B. Zeitreihe) (N gerade). Gesucht: reelle Fourier Koeffizienten $A_0, A_1, \dots, A_{\frac{N}{2}}, B_1, \dots, B_{\frac{N}{2}-1}$. (Diese entsprechen den stetigen Fourier-Koeffizienten a_k, b_k von Abschnitt 3.6C.) Statt einer direkten Transformation (unten in (4)) "Umweg" über $n := \frac{N}{2}$ komplexe Zahlen:



(oder bei Bedarf andersherum)

Transformation im Datenraum: g_j werden künstlich gesetzt aus aufeinanderfolgenden f -Werten:

$$g_j := f_{2j} + i f_{2j+1} \text{ für } j = 0, \dots, n-1$$

(2)

Übergang zwischen den Koeffizienten: Es sei $B_0 := B_n := 0$ und $c_n := c_0$. Dann gilt:

$$A_\nu - i B_\nu = \frac{1}{2}(c_\nu + \bar{c}_{n-\nu}) + \frac{1}{2i}(c_\nu - \bar{c}_{n-\nu})e^{-i\nu\frac{\pi}{n}} \text{ für } \nu = 0, 1, \dots, n$$

(3)

(Das “Quer” in \bar{c} bedeutet die konjugiert-komplexe Version.) Algorithmus: rechte Seite ausrechnen, nach Real- und Imaginärteil trennen und A_ν , $-B_\nu$ ablesen.

Ergänzung: Für die Koeffizienten der reellen diskreten Fourier-Transformation gilt der direkte Zusammenhang

$$A_\nu = \frac{2}{N} \sum_{j=0}^{N-1} f_j \cos(\nu j \frac{2\pi}{N}), \quad B_\nu = \frac{2}{N} \sum_{j=0}^{N-1} f_j \sin(\nu j \frac{2\pi}{N})$$

(4)

D. Zwei Anwendungen

Es sei $f(t)$ periodisch mit Periode T und stückweise stetig. $\omega := \frac{2\pi}{T}$

Interpolation: Das **trigonometrische Polynom**

$$f_n(t) := \frac{A_0}{2} + \sum_{\nu=1}^{n-1} (A_\nu \cos \nu \omega t + B_\nu \sin \nu \omega t) + \frac{A_n}{2} \cos n \omega t$$

interpoliert f an den Stützstellen

$$t_j := j \frac{T}{N} = j \frac{T}{2n}, \quad j = 0, 1, \dots, 2n.$$

Trigonometrisches Ausgleichsproblem: Das trigonometrische Polynom

$$f_m(t) := \frac{A_0}{2} + \sum_{\nu=1}^m (A_\nu \cos \nu \omega t + B_\nu \sin \nu \omega t)$$

für $m < \frac{N}{2} = n$ **minimiert**

$$\sum_{j=1}^N (f_m(t_j) - f(t_j))^2.$$

($f(t_j)$) können äquidistant erhobene Daten sein, f muss nicht bekannt sein.)

4.7 Fast Fourier-Transformation

zur Berechnung von Summen der Form

$$c_\nu = \frac{1}{n} \cdot \sum_{j=0}^{n-1} g_j e^{-i\nu j \frac{2\pi}{n}} \quad \text{für } \nu = 0, \dots, n-1, \quad g_\nu \in \mathbb{C}$$

A. Aufwand

Aufwand bei naivem Vorgehen: pro ν n trigonometrische Funktionen (Exponential-Funktionen auswerten), und n Multiplikationen, dazu $n - 1$ Additionen. Das sind für alle Koeffizienten **$O(n^2)$ Operationen.**

n meist sehr groß! (z.B. $n = 10000$ oder viel mehr)

Idee 1965 Cooley & Tukey, ähnlich früher Gauß, oder Runge:

Fast Fourier-Transformation (FFT): Aufwand **$O(n \log_2 n)$ Operationen**

n	$n \log_2 n$	Ersparnis-Faktor $\frac{n}{\log_2 n}$
10^3	$\approx 10 \cdot 10^3$	100
10^4	$\approx 13 \cdot 10^4$	700
10^5	$\approx 17 \cdot 10^5$	6000

Annahme: n ist Zweierpotenz, d.h. $n = 2^p$

d.h. Folge $\frac{n}{2}, \frac{n}{4}, \frac{n}{8}, \dots$ ist wohldefiniert und besteht aus natürlichen Zahlen. Abkürzung: $m := \frac{n}{2}$. Für jedes ν gilt:

$$\begin{aligned}
 c_\nu &= \frac{1}{n} \sum_{k=0}^{m-1} g_{2k} e^{-i\nu 2k \frac{2\pi}{n}} + \frac{1}{n} \sum_{k=0}^{m-1} g_{2k+1} e^{-i\nu (2k+1) \frac{2\pi}{n}} \\
 &= \underbrace{\frac{1}{n} \sum_{k=0}^{m-1} g_{2k} e^{-i\nu k \frac{2\pi}{m}}}_{\downarrow} + \underbrace{\left(e^{-i\nu \frac{2\pi}{n}} \right)}_{\substack{w^\nu \text{ mit Abkürzung} \\ \boxed{w := e^{-i \frac{2\pi}{n}}} }} \star \underbrace{\frac{1}{n} \sum_{k=0}^{m-1} g_{2k+1} e^{-i\nu k \frac{2\pi}{m}}}_{\downarrow}
 \end{aligned}$$

Zwei Summen von der gleichen FFT-Struktur

Jede Summe hat $2^{p-1} = m$ Terme. An der 2. Summe ist eine Multiplikation \star erforderlich (mit $e^{-i\nu \frac{2\pi}{n}}$), für alle c_ν insgesamt m viele (nur m , nicht n ; s.u.).

Jede der beiden FFT-Summen kann analog reduziert werden; es treten weitere $2 \cdot \frac{m}{2}$ Multiplikationen auf.

u.s.w.:

Mit m Multiplikationen $\star \rightarrow 2^1 = 2$ Summen mit je 2^{p-1} Termen.

Mit $2 \cdot \frac{m}{2}$ Multiplikationen $\rightarrow 2^2$ Summen mit je 2^{p-2} Termen.

Mit $4 \cdot \frac{m}{4}$ Multiplikationen $\rightarrow 2^3$ Summen mit je 2^{p-3} Termen.

\vdots

Mit m Multiplikationen $\rightarrow 2^p$ "Summen" mit je **1 Term.**

D.h. nach p Reduktionsschritten bleibt in den Summen nichts mehr zu tun!

$p = \log_2 n$ Reduktionen mit je $m = \frac{n}{2}$ Multiplikationen ergibt insgesamt $O(n \log_2 n)$ Operationen.

B. Analyse

Analyse eines Reduktionsschrittes: Die Summen

$$\sum_{k=0}^{m-1} g_{\square} \cdot e^{-i\nu k \frac{2\pi}{m}}$$

sind periodisch mit Periode m und brauchen deshalb für c_m, \dots, c_{n-1} nicht neu berechnet zu werden. (\square ist hier Platzhalter für $2k$ bzw. $2k + 1$)

Es gilt nämlich (ν durch $m + j$ ersetzt):

$$\begin{aligned}
 c_{m+j} &= \frac{1}{n} \sum_{k=0}^{m-1} g_{2k} e^{-i(m+j)k \frac{2\pi}{m}} + \underbrace{\left(e^{-i(m+j) \frac{2\pi}{n}} \right)}_{(\downarrow^2)} \frac{1}{n} \sum_{k=0}^{m-1} g_{2k+1} e^{-i(m+j)k \frac{2\pi}{m}} \\
 &\quad \left. \begin{array}{l} (\downarrow^1) \quad e^{-imk \frac{2\pi}{m}} = e^{-ik2\pi} = +1 \\ (\downarrow^2) \quad e^{-im \frac{2\pi}{n}} = e^{-i\pi} = -1 \end{array} \right\} \implies \\
 c_{m+j} &= \frac{1}{n} \sum_{k=0}^{m-1} g_{2k} e^{-ijk \frac{2\pi}{m}} - \underbrace{\left(e^{-ij \frac{2\pi}{n}} \right)}_{w^j} \star \frac{1}{n} \sum_{k=0}^{m-1} g_{2k+1} e^{-ijk \frac{2\pi}{m}}
 \end{aligned}$$

Das bedeutet Aufspalten der g_{\square} nach geraden und ungeraden Nummern (*der 1. Schritt*). Definiere zwei m -Vektoren y' und y'' für $j = 0, 1, \dots, m - 1$ über ihre Komponenten y_j

$$y'_j := \frac{1}{n} \sum_{k=0}^{m-1} g_{2k} e^{-ijk \frac{2\pi}{m}}, \quad y''_j := \frac{1}{n} \sum_{k=0}^{m-1} g_{2k+1} e^{-ijk \frac{2\pi}{m}}$$

D.h. wegen $n = 2m$ sind $y' = \frac{1}{2} \text{FFT}(g_0, g_2, g_4, \dots)$ und $y'' = \frac{1}{2} \text{FFT}(g_1, g_3, g_5, \dots)$ FFTs der halben Größe (*2. Schritt*).

Dann gilt (*der 3. Schritt*):

$ \begin{aligned} c_j &= y'_j + w^j \star y''_j \\ c_{m+j} &= y'_j - (w^j y''_j) \end{aligned} $	}	$j = 0, 1, \dots, m-1$	Aufwand: m Multiplikationen $w^j \star y''_j$, $m + m = n$ Additionen/Subtraktionen.
---	---	------------------------	--

(Dazu kommt die rekursive FFT-Berechnung von y', y'' , s.o.)

Zusammenfassung/rekursive Definition von FFT_n aus $\text{FFT}_{n/2}$:

$(m = n/2)$

$$\begin{array}{ccc}
 \left(\begin{array}{c} g_0 \\ g_1 \\ \vdots \\ g_{n-1} \end{array} \right) & \begin{array}{c} \nearrow \\ \searrow \end{array} & \begin{array}{c} \left(\begin{array}{c} g_0 \\ g_2 \\ \vdots \\ g_1 \\ g_3 \\ \vdots \end{array} \right) \\ \rightarrow y' = \frac{1}{2} \text{FFT}_m(g_0, g_2, \dots) \rightarrow \\ \rightarrow y'' = \frac{1}{2} \text{FFT}_m(g_1, g_3, \dots) \rightarrow \end{array}
 \end{array}$$

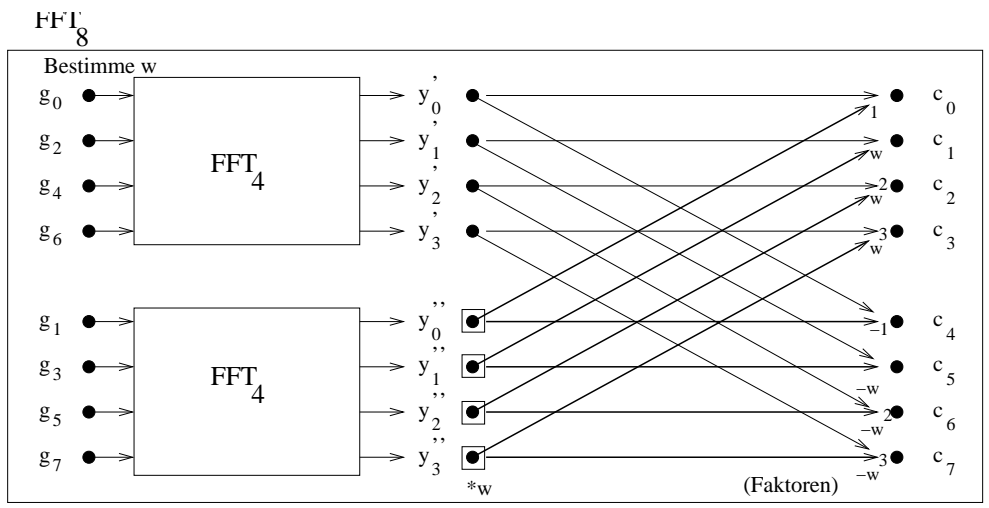
1. Schritt

2. Schritt

$$\begin{pmatrix} y'_j + w^j \star y''_j \\ \dots\dots\dots \\ y'_j - w^j y''_j \end{pmatrix} \rightarrow \begin{pmatrix} c_0 \\ \vdots \\ c_{m-1} \\ \dots\dots\dots \\ c_m \\ \vdots \\ c_{n-1} \end{pmatrix} =: \text{FFT}_n(g_0, g_1, \dots, g_{n-1})$$

3. Schritt

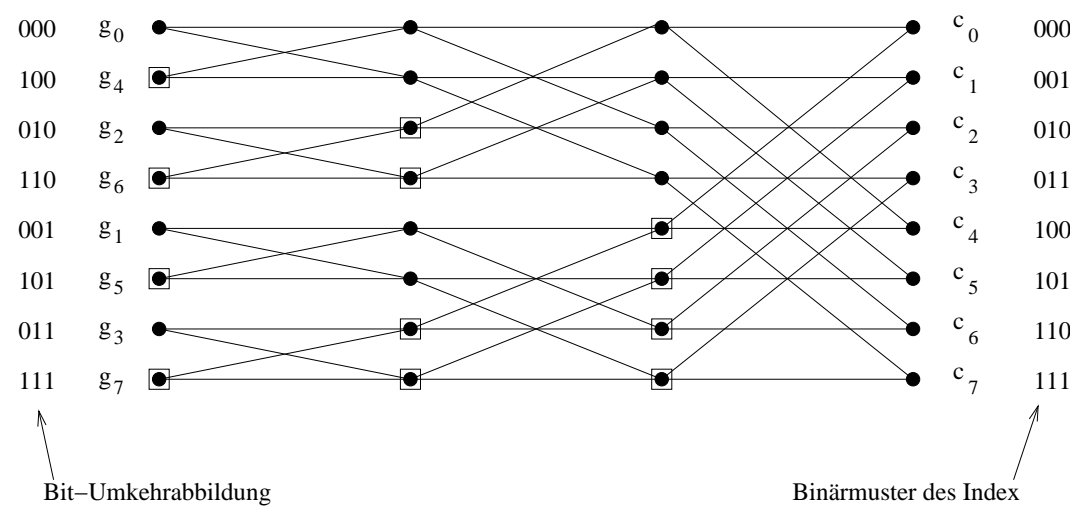
Illustration eines Rekursions-Schrittes für $n = 8$:



Ersetze rekursiv die FFT-Boxen durch analoge Schemata! (d.h. der linke Teil der Illustration ist vorläufig.) Der Input von FFT_4 wird wiederum nach geraden und ungeraden Positionen sortiert ...

Es ist eine einfache Buchhaltung (ohne Umspeichern) möglich, wenn die g_j gleich "richtig" angeordnet werden! Hierzu verwende die **Bit-Umkehrabbildung**, vgl. Illustration. Es werden dann jeweils benachbarte Paare von Untervektoren kombiniert, ausgehend von Zweiergruppen, dann Viergruppen etc.

Illustration $n = 8$:



Praktischer Hinweis: Es wird nur **ein** komplexer n -Vektor benötigt als Arbeitsvektor; zu Beginn zu besetzen mit

$$(g_0, g_4, g_2, \dots, g_7)$$

(Beispiel $n = 8$). Am Ende enthält der Vektor die gesuchten Koeffizienten. Die komplexen Zahlen w^j können als Wurzeln der komplexen 1 mit Winkelhalbierung berechnet werden oder durch Rekursionen der trigonometrischen Funktionen (vgl. [Stoer I, Paragraph 1.4]). Falls n keine Zweier-Potenz ist: Es gibt analoge FFTs, die mit Primfaktor-Zerlegung arbeiten.

Programm in den “Recipes” [Press et al.]

4.8 Ausgleichsprobleme, data fit

Mess-Reihe (t_k, y_k) , $k = 1, \dots, m$

Man nimmt ein Verteilungsgesetz an (“Modell”):

$$y = g(t) \text{ mit noch zu bestimmenden Parametern } x_1, \dots, x_n,$$

also:

$$g(t; x_1, \dots, x_n)$$

Beispiel: $g(t) = x_1 + x_2 t + x_3 t^2$

Annahme: mehr Daten als Parameter, d.h. $m > n$.

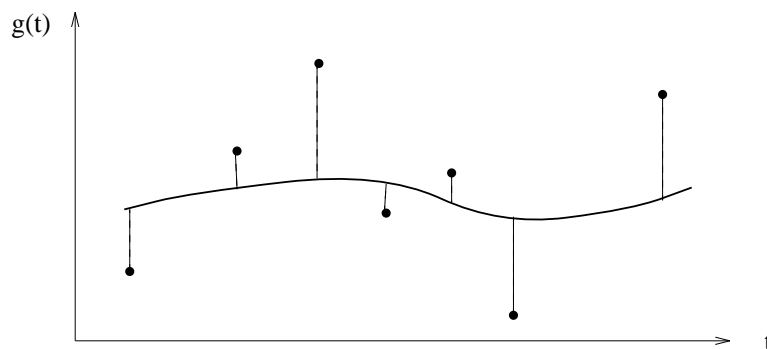
Gesucht: x_1, \dots, x_n , so dass

$$\begin{cases} y_1 = g(t_1; x_1, \dots, x_n) =: f_1(x_1, \dots, x_n) \\ \vdots \\ y_m = g(t_m; x_1, \dots, x_n) =: f_m(x_1, \dots, x_n) \end{cases}$$

“möglichst gut” erfüllt sind.

Wegen $m > n$ ist dies ein “überbestimmtes” Gleichungssystem.

Schreibe die Messdaten y_k nun als Vektor y , also Vektor-Schreibweise: $y = f(x)$



Spezialfall $g(t) = x_1 + x_2 t$: **Regressions-Gerade.**

[Fall $m < n$: Es gibt ∞ -viele Lösungen. \rightarrow Kapitel Optimierung.]

Methode der kleinsten Quadrate, "Gauß-Approximation":

minimiere über x_1, \dots, x_n die Funktion

$$\phi(x) := \sum_{k=1}^m (y_k - f_k(x_1, \dots, x_n))^2$$

Für den minimierenden Vektor x müssen alle partiellen Ableitungen 1. Ordnung von ϕ verschwinden, vergleiche Abschnitt 3.12. Also gilt mit der Kettenregel

$$2 \sum_{k=1}^m (y_k - f_k(x)) \frac{\partial f_k(x)}{\partial x_i} = 0 \quad \text{für } i = 1, \dots, n$$

Annahme: g linear in x_1, \dots, x_n (nicht unbedingt in t)!

Dann gibt es eine Matrix A mit

$$f(x) = Ax, \quad A \text{ (} m \times n \text{) Matrix, obiges Beispiel: } A = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ \vdots & \vdots & \vdots \end{pmatrix}$$

und $\phi(x) := (y - Ax)^T (y - Ax) = \|y - Ax\|_2^2$ ist zu minimieren.

Es gilt

$$\phi(x) := x^T A^T A x - 2x^T A^T y + y^T y,$$

und man kann zeigen: $\text{grad } \phi = 2A^T A x - 2A^T y$.

gradient $[\phi(x)] = 0 \Rightarrow \boxed{A^T A x = A^T y}$ "Normalengleichung"

des linearen Problems (notwendige Bedingung für Minimum)

Der **Gedanke**, das Gleichungssystem

$$(A^T A)x = A^T y \quad (*)$$

numerisch zu lösen, ist **nicht gut**: Der Übergang auf (*) **verschlechtert** die Kondition (im Vergleich zum ursprünglichen Minimierungsproblem)!

Gute Methode: n Householder Transformationen $H^{(j)}$ direkt auf das überbestimmte Gleichungssystem $Ax = y$ anwenden, so dass A und y simultan transformiert werden auf

$$A^{(n)} = \begin{pmatrix} \overbrace{\begin{matrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix}}^n \\ 0 \\ \hline 0 \end{pmatrix} \begin{matrix} \leftarrow n \\ \\ \leftarrow m-n \end{matrix}, \quad y^{(n)} = \begin{pmatrix} h_1 \\ - \\ h_2 \end{pmatrix} \leftarrow n$$

(Zu Householder-Matrizen vgl. Abschnitt 2.9.)

Die Matrix $H := H^{(n)} \cdot \dots \cdot H^{(1)}$ ist orthogonal.

Es folgt die Beziehung (für alle x)

$$\|y - Ax\|_2 = \|H(y - Ax)\|_2 = \|y^{(n)} - A^{(n)}x\|_2 = \left\| \begin{pmatrix} h_1 - Ux \\ h_2 \end{pmatrix} \right\|_2 \geq \|h_2\|_2$$

Also: $\|y - Ax\|_2$ wird minimal für die Lösung x von $Ux = h_1$

(äquivalent zur Minimierung von $\|y - Ax\|_2^2$)

Methode:

1. Berechne U und h_1 mit Householder-Transformationen;
2. erhalte x aus $Ux = h_1$ durch Rückwärtselimination.

Hinweis: Nichtlineare Ausgleichsprobleme werden linearisiert und iterativ gelöst.

Kapitel 5 Nichtlineare Gleichungssysteme und Iterationen

Wir betrachten das System

$$f(x) = 0$$

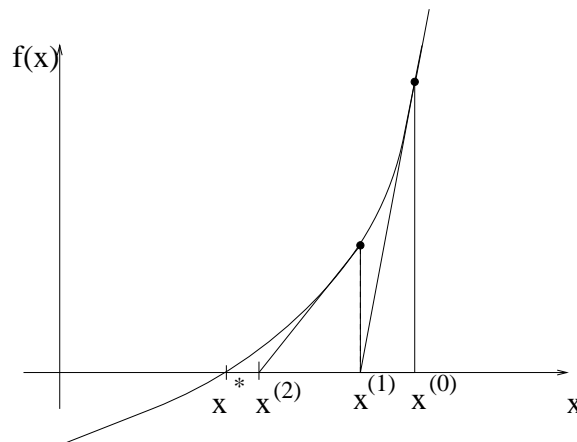
von n skalaren Gleichungen $f_i(x_1, \dots, x_n) = 0$, $i = 1, \dots, n$.

Gesucht: Nullstelle x^* von $f(x) = 0$. Es sei $x^{(0)}$ eine Näherung zu x^* , oder einfach ein "Startvektor".

5.1 Lösen einer skalaren Gleichung

Zunächst $n = 1$. Für skalare Gleichungen gibt es viele Verfahren. Z.B. Bisektion: Intervall auf x -Achse, welches genau die Nullstelle enthält \rightarrow fortgesetzte Intervall-Halbierung...

Newton-Verfahren:



Erhalte bessere Näherung durch Anlegen der Tangente an $(x^{(0)}, f(x^{(0)}))$, und durch Aufsuchen der **Nullstelle der Tangente** mit $f'(x^{(0)}) = \frac{f(x^{(0)})}{x^{(0)} - x^{(1)}}$; die Auflösung ergibt

$$x^{(1)} = x^{(0)} - f(x^{(0)})/f'(x^{(0)}) .$$

Wiederholung: Tangente an $(x^{(1)}, f(x^{(1)})) \rightarrow x^{(2)}$ u.s.w. $x^{(3)}, x^{(4)}, \dots$

Dieses Verfahren ist das klassische **Newton-Verfahren**.

[*Herleiten einer Iterationsfolge* $x^{(1)}, x^{(2)}, \dots$ aus abgebrochener Taylor-Entwicklung:

$$0 = f(x^*) = f(x) + (x^* - x)f'(x) + T.h.O. ,$$

also $0 = f(x) + (x^* - x)f'(x)$, mit geeigneter Interpretation von x und x^* .]

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}$$

für $k = 0, 1, 2, \dots$

Die Iteration ist von der Form

$$x^{(k+1)} = \Phi(x^{(k)}), \quad \text{mit } \Phi(x) = x - \frac{f(x)}{f'(x)}$$

“**Fixpunkt-Iteration**”. Es gilt die Fixpunkt-Gleichung $\Phi(x^*) = x^*$.

5.2 Zur Konvergenz

Beispiel: Newton-Verfahren zur Wurzelberechnung:

$$\begin{aligned} x^* = \sqrt{a} &\longrightarrow f(x) = x^2 - a \\ \Rightarrow x^{(k+1)} &= x^{(k)} - \frac{(x^{(k)})^2 - a}{2x^{(k)}} \end{aligned}$$

Also

$$\Phi(x) = \frac{1}{2} \left(x + \frac{a}{x} \right)$$

$\sqrt{2}$? $x^{(0)} = 1, \quad a = 2$

k	$x^{(k)}$
0	1
1	1.5
2	1.416666666
3	1.414215686
4	1.414213562

Wir beobachten “lokal” (d.h. in der Nähe von x^*) eine Verdoppelung der korrekten Stellenzahl in jedem Schritt. Man sagt deshalb auch, dass die Konvergenz des Newton-Verfahrens **quadratisch** ist.

Definition: Konvergenzordnung $p \geq 1$ der Folge $x^{(k)}$ gegen x^* , wenn es ein C gibt, so dass für alle k

$$\|x^{(k+1)} - x^*\| \leq C \|x^{(k)} - x^*\|^p$$

mit $0 < C$ für $p > 1$, und $0 < C < 1$ für $p = 1$.

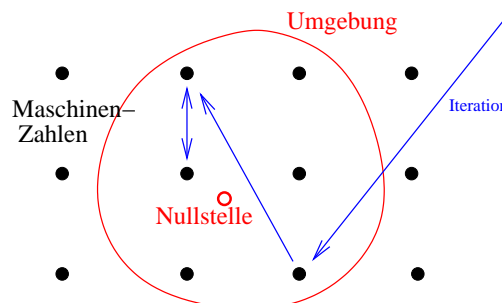
Folgerung: $x^{(k)} \rightarrow x^*$ für $k \rightarrow \infty$.

Newton-Verfahren: $p = 2$.

$p = 1$: “lineare Konvergenz” (langsam!)

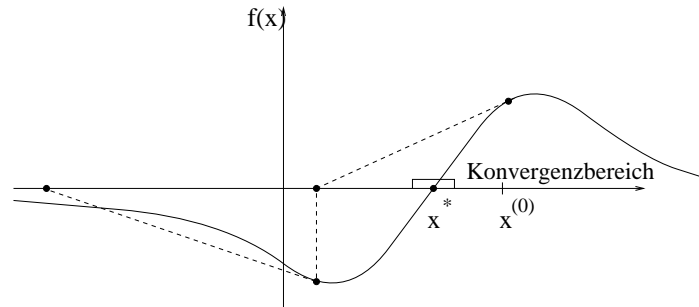
“Superlineare Konvergenz” wenn $C = C_k \rightarrow 0$ für $k \rightarrow \infty$.

Praxis: Im Rechner ist $\|x^{(k+1)} - x^{(k)}\|$ *keine* Nullfolge! Grund: Es stehen nur endlich viele Maschinenzahlen zu Verfügung, besonders wenige in einer Umgebung $\mathcal{U}(x^*)$. Am Ende i.A. periodisches Verhalten.



Probleme:

- (1) “Wann” Konvergenz ?
- (2) Wie schnell?
- (3) Übertragbarkeit auf Vektorfunktionen
- (4) Technische Probleme: Abbruchkriterium

Beispiel für Divergenz

Divergenz: für das gewählte $x^{(0)}$. Es gibt ein kleines Intervall um x^* mit Konvergenz.
Wie “trifft” man hinein?

Dämpfung des Newtonverfahrens

$$x^{(k+1)} = x^{(k)} - \lambda \frac{f(x^{(k)})}{f'(x^{(k)})}$$

mit “Dämpfungsfaktor” λ derart, dass die “Testfunktion”

$$T(x) := (f(x))^2 \quad \text{minimal wird.}$$

Hintergrund: $T(x^*) = (f(x^*))^2 = 0$, und $T(x) > 0$ für $x \neq x^*$.

Für $\lambda = 1$ ist das klassische Newtonverfahren enthalten.

Das Minimum bzw. ein optimales λ in jedem Iterationsschritt ist i.a. schwer zu finden, deshalb Ersatzmethode (hier nur Grundprinzip):

Gedämpftes Newton-Verfahren:

$\lambda = 1$, $k = 0$

(1) $k \rightarrow k + 1$

Auswerten von $f(x^{(k)})/f'(x^{(k)})$ [= $x^{(k)} - x^{(k+1)}$ = Newton-Schritt]

(2) berechne $\bar{x} := x^{(k)} - \lambda \frac{f(x^{(k)})}{f'(x^{(k)})}$ [Versuchskandidat]

falls $T(\bar{x}) \geq T(x^{(k)})$ und $\lambda \geq \lambda_{min}$, dann $\lambda := \frac{1}{2}\lambda$; goto (2) [λ -Halbierung]

falls $T(\bar{x}) \geq T(x^{(k)})$ und $\lambda < \lambda_{min}$: STOP “no convergence” [Fehlaustrag]

$x^{(k+1)} := \bar{x}$ [$T(\bar{x}) < T(x^{(k)})$, \bar{x} wird akzeptiert]

falls $\lambda \leq \frac{1}{2}$ setze $\lambda := 2\lambda$ [λ möglichst in die Nähe von 1]

Konvergenzabfrage und -Ausgang

goto (1)

(enthält Divergenz-Abbruch-Kriterium!) z.B. $\lambda_{min} = 0.01$

Die Dämpfung beinhaltet keine Garantie für Konvergenz!

Abbruchkriterium, Prüfen der Konvergenz

vorgebene Fehlerschranke sei ϵ (z.B. $\epsilon = 10^{-6}$)

Möglichkeit $|f(x^{(k+1)})| < \epsilon$ genügt i.A. nicht, da dieses Kriterium eher von der Skalierung von f abhängt als vom Fehler.

Möglichkeit $|x^{(k+1)} - x^{(k)}| < \epsilon$ genügt i.A. auch nicht, da dieses Kriterium nur einen *Hinweis* auf den absoluten Fehler $|x^{(k+1)} - x^*|$ darstellt; es ist wiederum skalierungsabhängig. Besser ist ein Maß für den relativen Fehler.

Professionelle Programme verlangen für den “Konvergenz-Ausgang” das Unterschreiten *mehrerer* Fehlerschranken, z.B.

$$|f(x^{(k+1)})| < \epsilon \quad \text{und} \quad |x^{(k+1)} - x^{(k)}| < \epsilon \cdot |x^{(k)}|$$

5.3 Das allgemeine Newtonverfahren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad f(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0$$

Linearisierung [bzw. Taylorentwicklung um $x^{(0)}$]:

$$0 = f(x^*) \approx f(x^{(0)}) + \underbrace{Df(x^{(0)})}_{\text{“Jacobi-Matrix”}}(x^* - x^{(0)}) \quad (*)$$

$$Df(x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Aus (*) erhalte ein lineares Gleichungssystem

$$Df(x^{(0)}) (x^{(1)} - x^{(0)}) = -f(x^{(0)})$$

für den ersten Korrekturvektor $\Delta x = x^{(1)} - x^{(0)}$. Analog wie im Fall $n = 1$ erhalten wir das allgemeine (gedämpfte) Newton-Verfahren:

Newton-Algorithmus

für $k = 0, 1, 2, \dots$:

Schritt 1: Berechne oder approximiere $Df(x^{(k)})$

Schritt 2: Löse das lineare Gleichungssystem

$$Df(x^{(k)}) \Delta x = -f(x^{(k)})$$

z.B. mit Gauß Algorithmus $\rightarrow \Delta x$

Schritt 3: Dämpfungs-Strategie:

Ermittle λ derart, dass

$$T(x^{(k)} + \lambda \Delta x) < T(x^{(k)}),$$

wo $T(x) := \sum_{i=1}^n (f_i(x))^2$

Schritt 4: $x^{(k+1)} = x^{(k)} + \lambda \Delta x$

$k \rightarrow k + 1$

goto Schritt 1

Beispiel:

$$f(x) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} (x_2 - 2)^2 - x_1 + 1 \\ 3x_1x_2^3 - 3x_1 + 1 \end{pmatrix} = 0$$

geometrisch: Nullstelle x^* ist in (x_1, x_2) -Ebene Schnittpunkt der Kurven, die durch $f_1(x_1, x_2) = 0$ und $f_2(x_1, x_2) = 0$ definiert sind.

Jacobi-Matrix:

$$\begin{pmatrix} -1 & 2(x_2 - 2) \\ 3x_2^3 - 3 & 9x_1x_2^2 \end{pmatrix}$$

1. Iteration für Startpunkt $x^{(0)} := (2, 1)$: Gleichungssystem für den Newton-Schritt:

$$\begin{pmatrix} -1 & -2 \\ 0 & 18 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -f(x^{(0)})$$

Lösung: $\Delta x_2 = -1/18 \Rightarrow \Delta x_1 = -2\Delta x_2 = 1/9 \Rightarrow$

$$x^{(1)} = x^{(0)} + \Delta x = \begin{pmatrix} 2 + 1/9 \\ 1 - 1/18 \end{pmatrix} \approx \begin{pmatrix} 2.11 \\ 0.944 \end{pmatrix}$$

Skizze: Schnitt von Parabel $x_1 = 1 + (x_2 - 2)^2$ mit der "hyperbel-ähnlichen" Kurve $3x_1(1 - x_2^3) = 1$.

5.4 Approximation der Jacobi-Matrix

Die exakte Berechnung von $\partial f_i / \partial x_j$ ist meist zu schwierig, oft sogar unmöglich!

Motivation im skalaren Fall:

1. Möglichkeit:

"numerische Approximation"
(im engeren Sinn)

$$f'(x^{(k)}) \approx \frac{f(x^{(k)} + h) - f(x^{(k)})}{h}$$

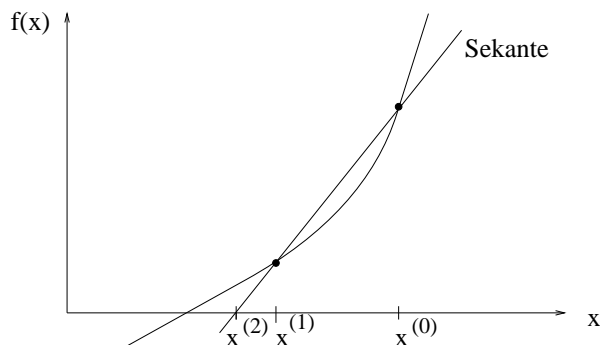
für kleines $|h|$, z.B. $h = 10^{-4}$
(nicht zu klein wegen
Auslöschung !)

2. Möglichkeit

"Sekantenverfahren"

$$f'(x^{(k)}) \approx \frac{f(x^{(k)}) - f(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}$$

hier kein Extra-Funktions-
Aufruf; vergleiche Figur.
Genauigkeit hängt vom
Iterationsverlauf ab.



Numerische Differentiation im Vektor-Fall:

(Es geht um Schritt 1 des Newton-Algorithmus.)

Es sei $v^{(j)}$ der Vektor der j -ten Spalte der Jacobi-Matrix Df ,

$$v^{(j)} := \begin{pmatrix} \frac{\partial f_1}{\partial x_j} \\ \vdots \\ \frac{\partial f_n}{\partial x_j} \end{pmatrix} = \frac{\partial f(x)}{\partial x_j}$$

Er wird ersetzt durch die Näherung

$$\frac{f(x_1, \dots, x_{j-1}, x_j + h, x_{j+1}, \dots, x_n) - f(x)}{h}$$

Also: spaltenweises Vorgehen. Für jede Spalte von Df ist **eine** Auswertung von f notwendig, insgesamt also n Auswertungen von f zuzüglich der von $f(x)$.**Algorithmus:** $(f(x)$ sei schon berechnet und als F gespeichert.) $\eta := 10^{-4}$ (z.B.)Für $j = 1, \dots, n$:

$$\left| \begin{array}{l} h = \eta \cdot \max\{1, |x_j|\} \cdot \text{sign}(x_j) \\ x_j := x_j + h \\ \text{werte } f \text{ aus} \rightarrow \text{Vektor } G \\ v^{(j)} := \frac{1}{h}(G - F) \\ x_j := x_j - h \quad [\text{Restauration von } x] \end{array} \right.$$

Hinweise:

1. Auch eine Verallgemeinerung der 2. Möglichkeit ist möglich: "Rang-1 Verfahren", "Broyden Approximation". Solche Verfahren benötigen mehr Iterationen als Newton; jede Iteration ist aber billiger.
2. Im \mathbb{R}^1 ist eine Kombination von Bisektion, Sekante und inverser Interpolation sinnvoll: Alg.v.Dekker-Brent, MATLAB fzero.

Kapitel 6 Optimierung

6.1 Optimierungsprobleme

Bezeichnungen:

$$x \in \mathbb{R}^n \quad , \quad f(x) \text{ reellwertig, ebenso} \\ g_1(x), \dots, g_m(x) \quad , \\ h_1(x), \dots, h_l(x)$$

Mit der Forderung $x \in \mathcal{M}$ für geeignetes $\mathcal{M} \subset \mathbb{R}^n$ kann z.B. auf alle ganzzahligen Vektoren eingeschränkt werden. Dieses Kapitel: $\mathcal{M} = \mathbb{R}^n$.

mathematische Optimierungsaufgabe:

gesucht ist Vektor x so dass

$$\left\{ \begin{array}{l} f(x) \rightarrow \text{Max (oder Min)} \\ g_i(x) = 0 \quad , \quad 1 \leq i \leq m \\ h_j(x) \leq 0 \quad , \quad 1 \leq j \leq l \\ x \in \mathcal{M} \end{array} \right. \quad , \quad (\text{OA})$$

(“endlich-dimensionale” Optimierung, weil Optimierung im \mathbb{R}^n)

Vektor-Schreibweise

$$g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} \quad , \quad h(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_l(x) \end{pmatrix} \quad ,$$

0 ist passender Vektor, und \leq ist komponentenweise zu verstehen.

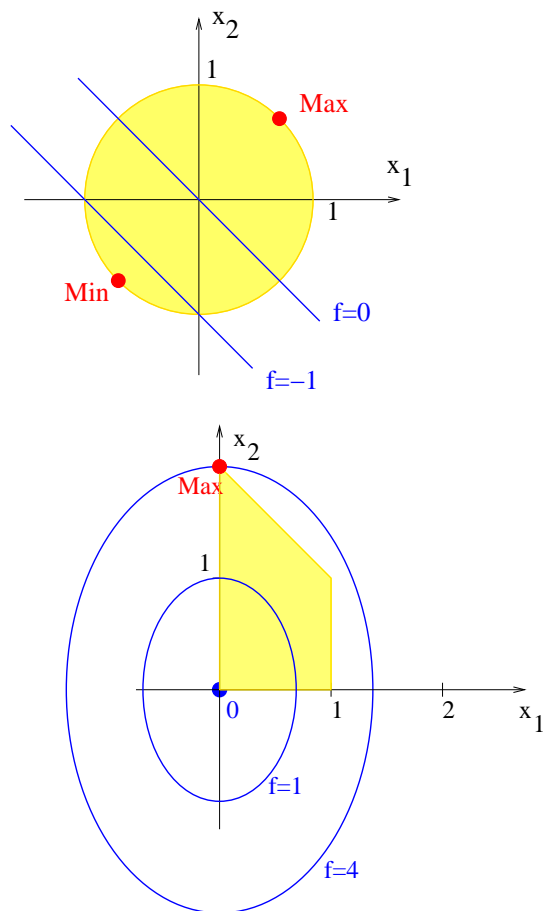
dann: $f(x) \rightarrow \text{Max/Min}$ unter den Nebenbedingungen $g(x) = 0$, $h(x) \leq 0$, $x \in \mathcal{M}$.

f heißt *Zielfunktion*.

Ein x welches die Nebenbedingungen (N.B.) erfüllt, heißt *zulässig*.

Beispiele

- (1) $f(x_1, x_2) := x_1 + x_2$ (d.h. linear), ohne Nebenbedingungen.
OA hat keine Lösung, weil f unbeschränkt.
- (2) f wie eben, unter der N.B. $x_1^2 + x_2^2 \leq 1$. (Einheitskreis-Scheibe)
OA hat Lösung.
 $h(x_1, x_2) := x_1^2 + x_2^2 - 1$ (vergleiche Figur)
- (3) f wie eben, unter $x_1^2 + x_2^2 \leq 1$ und $x_1 - x_2 + 3 = 0$.
OA hat keine Lösung, weil keine zulässigen Punkte existieren.
- (4) $f(x_1, x_2) := 2x_1^2 + x_2^2$ (d.h. nichtlinear, speziell: quadratisch) unter den N.B.
 $x_i \geq 0 \quad (i = 1, 2)$
 $x_1 \leq 1$
 $x_2 \leq 2 - x_1$
(vergleiche Figur)



6.2 Methoden der Analysis

A. Freie Optimierung

Suche Min (Max) von $f(x)$ *ohne* Nebenbedingung!

(Annahmen: f sei C^2 , $x \in \mathbb{R}^n$)

Am Extremum gilt $\text{grad}f(x) = 0$, d.h.

$$\frac{\partial f(x)}{\partial x_1} = \dots = \frac{\partial f(x)}{\partial x_n} = 0$$

Der Gradient-Vektor $F(x) := \text{grad}f(x)$ wird oft mit f_x bezeichnet.

Die Komponenten von F sind also

$$F_i = \frac{\partial f(x)}{\partial x_i}$$

Dann kann das **Newton-Verfahren** verwendet werden zur Berechnung einer Nullstelle von F . Benötige dazu die Jacobi-Matrix von F :

$$\begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

Das ist die Matrix aller zweiten partiellen Ableitungen von f , auch Hesse-Matrix benannt, oft bezeichnet mit f_{xx} .

VORSICHT: Hier bezeichnet f die zu minimierende Funktion, während F die Funktion ist, deren Nullstelle gesucht wird!

Es resultiert das folgende **Iterative Verfahren** für $k = 0, 1, 2, \dots$: ($DF \cdot \Delta x = -F$)

$$f_{xx}(x^{(k)})\Delta x = -f_x(x^{(k)})$$

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

für geeigneten Startvektor $x^{(0)}$. Der Vektor der jeweiligen ‘Suchrichtung’ Δx ergibt sich also aus der Newton-Methode für $F(x) = 0$.

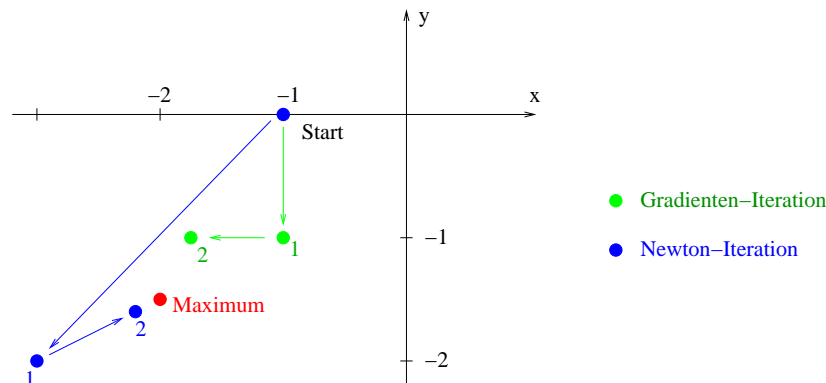
Beispiel: ($n = 2$, hier mit Bezeichnung $x_1 \rightarrow x$, $x_2 \rightarrow y$) Suche Maximum von

$$f(x, y) = \frac{1}{3}x^3 - 2y^2 + 2xy - x - 2y .$$

Das Newton-Gleichungssystem zur iterativen Berechnung einer Nullstelle von F ist hier

$$\begin{pmatrix} 2x^{(k)} & 2 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = - \begin{pmatrix} (x^{(k)})^2 + 2y^{(k)} - 1 \\ -4y^{(k)} + 2x^{(k)} - 2 \end{pmatrix}$$

Startet man mit Vektor $(x^{(0)}, y^{(0)}) = (-1, 0)$ und verwendet die Newton-Richtungen $(\Delta x, \Delta y)$, dann ergibt sich $(x^{(1)}, y^{(1)}) = (-3, -2)$, und $(x^{(2)}, y^{(2)}) = (-2.2, -1.6)$. (Hinweis: exaktes Maximum für $(-2, -1.5)$; vgl. Figur)



Allgemeinere Betrachtung:

Ein Vektor v heißt **Abstiegsrichtung** im Punkt x , wenn es ein $\varepsilon > 0$ gibt mit

$$f(x + tv) < f(x) \quad \text{für } 0 < t < \varepsilon .$$

Beispiel: negativer Gradient: $v = -\text{grad}f(x)$

Wenn ein Maximum gesucht wird, dann *Aufstiegsrichtung*, z.B. $+\text{grad}f$.

Der Gradient ist diejenige Richtung, in der f maximal ansteigt! Und $-\text{grad}f$ ist die Richtung des maximalen Gefälles!

Algorithmus (erläutert für Minimierung):

- 1.) Konstruiere in $x^{(k)}$ eine Abstiegsrichtung $v^{(k)}$.
- 2.) Führe ein-dimensionale Minimierung durch in Richtung $v^{(k)}$. Bei dieser “line search” wird das Minimum von $f(x^{(k)} + tv^{(k)})$ näherungsweise ermittelt. Wenn das Minimum dieser *line search* für ein $t = t_k$ erreicht wird, dann setze $x^{(k+1)} = x^{(k)} + t_k v^{(k)}$.

Man nennt diese iterative Minimierung **Gradientenverfahren**, wenn $v = -\text{grad } f$.

obiges Beispiel: (gleicher Startvektor) Wegen $\text{grad}f(x^{(0)}) = \begin{pmatrix} 0 \\ -4 \end{pmatrix}$ ist $v^{(0)} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ eine Aufstiegsrichtung. Zu *maximieren* ist also in Richtung

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} - t \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -t \end{pmatrix}$$

für $t > 0$. Eingesetzt in f :

$$f(-1, -t) = \frac{2}{3} - 2t^2 + 4t$$

hat Maximum für $t = 1$. Also $(x^{(1)}, y^{(1)}) = (-1, -1)$. (Die nächste Iteration ergibt $(-\sqrt{3}, -1)$, vgl. Figur)

Hinweis: Die Newton-Richtung ist i.A. “schneller” als die Gradienten-Richtung, weil sie auch Informationen der zweiten Ableitung berücksichtigt.

B. Optimierung mit Gleichungen als Nebenbedingungen

$f(x) \rightarrow \text{Min}$ unter N.B. $g_i(x) = 0$, $i = 1, \dots, m$ (g sei wie f differenzierbar) mit “Lagrange-Multiplikatoren” $\lambda_1, \dots, \lambda_m$ (Vektor λ)

$$L(x, \lambda) := f(x) + \sum_{j=1}^m \lambda_j g_j(x)$$

$n + m$ Unbekannte $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$.

Satz: x^* minimiere das Optimierungsproblem. Dann gibt es λ , so dass $\text{grad } [L(x^*, \lambda)] = 0$. (Das sind $n + m$ Gleichungen.)

C. Konvexe Optimierung

vgl. Abschnitt 6.4.

6.3 Lineare Optimierung

f, g, h seien alle linear in x :

$$\begin{aligned} f(x) &= c_1 x_1 + \dots + c_n x_n &= c^T x \\ g_i(x) &= d_{i1} x_1 + \dots + d_{in} x_n - r_i &, i = 1, \dots, m \\ h_j(x) &= b_{j1} x_1 + \dots + b_{jn} x_n - s_j &, j = 1, \dots, l \end{aligned}$$

Matrix-Vektor-Schreibweise: Mit

$$\begin{aligned} (c_i) &\rightarrow c && , \quad n \text{ Vektor} , \text{ analog } r \in \mathbb{R}^m, s \in \mathbb{R}^l \\ ((d_{ik})) &\rightarrow D && , \quad m \times n \text{ Matrix} \\ ((b_{jk})) &\rightarrow B && , \quad l \times n \text{ Matrix} \end{aligned}$$

ergibt sich aus dem Problem von § 6.1 speziell das *Lineare Programm* (LP)

$$\left\{ \begin{array}{l} f(x) = c^T x \rightarrow \text{Min (oder Max)} \\ Dx = r \\ Bx \leq s \end{array} \right. \quad \text{(LP)}$$

Unter den diversen möglichen Formen von Problemen linearer Optimierung wird ein Typ als **Standardform** oder **Normalform** bezeichnet:

$$\begin{aligned} f(x) = c^T x &\rightarrow \text{Min (oder Max)} \\ \text{unter den N.B. } Ax = b &\text{ und } x \geq 0 \end{aligned} \quad \text{(NF)}$$

(Vektoren x, c gegenüber (LP) modifiziert, ebenso wie die Dimension n !)

“Tricks” zur Überführung von (LP) in Normalform (NF):

- 1.) $Bx \leq s \Leftrightarrow y := s - Bx \geq 0$
 y heißt Schlupf-Variable (*slack variable*)
 Füge y als weitere Variable und $Bx + y = s$ als weitere Gleichung hinzu.
- 2.) Realisierung “freier” x_i (im Sinne $x_i < 0$ ist erlaubt): Definiere

$$\begin{aligned} x_i^+ &:= \max\{x_i, 0\} && \geq 0 \\ x_i^- &:= \max\{-x_i, 0\} && \geq 0 \end{aligned}$$

Dann gilt: $x_i = x_i^+ - x_i^-$

Füge x_i^+ und x_i^- als neue Variable hinzu, nehme x_i heraus, und ersetze in den Gleichungen $x_i = x_i^+ - x_i^-$.

eventuell noch

- 3.) $f(x) \rightarrow \text{Max} \Leftrightarrow -f(x) \rightarrow \text{Min}$

Beispiel

$$\begin{array}{ll} \text{(LP)} & \rightarrow \text{(NF)} \\ x_1 + 2x_3 \leq 30 & \rightarrow x_1 + 2x_3 + y_1 = 30 \\ 2x_2 - 3x_4 \leq 0 & \rightarrow 2x_2 - 3x_4 + y_2 = 0 \\ x_2 - x_3 + 2x_4 \geq 1 & \rightarrow x_2 - x_3 + 2x_4 - y_3 = 1 \\ x_1 + x_2 + x_3 + x_4 = 4 & \rightarrow \text{bleibt} \\ x_i \geq 0 & \rightarrow \text{bleibt, hinzu kommt } y_i \geq 0 \end{array}$$

neue Variable für die Optimierung:

$$x_1, \dots, x_4, y_1, y_2, y_3 \quad , \quad \text{alle } \geq 0 \quad (\text{nenne sie etwa } \tilde{x}).$$

Dann gilt $A\tilde{x} = b$ mit $\tilde{x} \in \mathbb{R}^7$ und

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & -3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 2 & 0 & 0 & -1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} , \quad b = \begin{pmatrix} 30 \\ 0 \\ 1 \\ 4 \end{pmatrix}$$

Die k -te Spalte von A ergibt sich aus den Koeffizienten zu x_k in (NF). Die 4 Gleichungen im \mathbb{R}^7 (falls überhaupt lösbar) definieren eine Menge, welche mindestens die Dimension 3 hat. (siehe unten) Die Menge ist nicht leer, denn zum Beispiel mit $x_4 = 4$, $x_1 = x_2 = x_3 = 0$ sind die Ungleichungen erfüllt, d.h.

$$\tilde{x} = (0, 0, 0, 4, 30, 12, 7)^T$$

erfüllt die Gleichungen und ist also zulässiger Punkt.

Allgemeine Analyse der Menge der zulässigen Punkte

Hierzu betrachte ohne Einschränkung nur die Standardform (NF). A sei $m \times n$ Matrix, $n > m$, $Ax = b$, $x \geq 0$. (D.h. n ist anders als in (LP). Das x hier entspricht dem \tilde{x} .)

D.h. $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$; m ist die Anzahl der Gleichungen. Die Menge der zulässigen Punkte, also die Lösungen von $Ax = b$ mit $x \geq 0$, ist ein Polyeder (*polyhedron*) [1]. Polyeder sind konvex (sofern nicht leer).

$Ax = b$ bedeutet: Der Vektor b ist Linearkombination der Spalten von A . Wenn $Ax = b$ lösbar ist, dann ist die Menge der Lösungen ein Teilraum des \mathbb{R}^n von der Dimension

$$n - \text{Rang}(A),$$

siehe auch Abschnitt 2.5.

Die Funktion $c^T x$ nimmt ihr Minimum (Max), falls es existiert, in mindestens einem endlichen Eckpunkt des Polyeders an.

Die Ecken heißen **Basislösungen**. Ein zulässiger Punkt x_B ist Basislösung genau dann, wenn diejenigen Spalten von A , die zu positiven Komponenten von x_B gehören, linear unabhängig sind.

Folgerungen für $\text{Rang}(A) = m$:

- ⇒ Es gibt m linear unabhängige Spalten.
- ⇒ Maximal m Komponenten von x_B sind positiv.
- ⇒ Mindestens $n - m$ Komponenten von x_B sind = 0.

Setzt man $n - m$ Komponenten von x zu Null, so definiert das $n - m$ Gleichungen/Ebenen $x_i = 0$. Zusammen mit den m Gleichungen von $Ax = b$ sind dies insgesamt n Gleichungen/Ebenen, diese definieren i.A. einen Eckpunkt des Polyeders.

Wechsel der Ecke:

Gibt man eines derjenigen $x_i = 0$ "frei", etwa x_k , welches nun erhöht wird ($x_k > 0$), dann läuft man eine Kante des Polyeders entlang, bis eine andere Komponente der freien Komponenten von x (z.B. x_j) Null wird. Dann hat man eine andere Ecke erreicht, und einen **Austausch-Schritt** vollzogen.

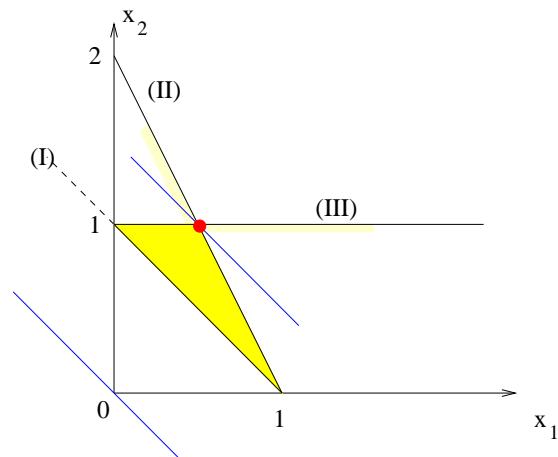
Da die Zielfunktion $c^T x$ im nicht-degenerierten Fall ihr Optimum in einer Ecke annimmt, geht es bei der linearen Optimierung um die Konstruktion geeigneter Austausch-Schritte.

[1] Ein Polyeder ist definiert als Lösungsmenge einer endlichen Anzahl von linearen Gleichungen und Ungleichungen.

Beispiel:

$$\begin{aligned} x_2 &\geq 1 - x_1 && (I) \\ x_2 &\leq 2 - 2x_1 && (II) \\ x_2 &\leq 1 && (III) \end{aligned}$$

Zielfunktion z.B. $-x_1 - x_2 = \text{Min!}$ (graphische Lösung...)



Die zugehörige Normalform ist gegeben durch

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

Also: $n = 5$, $m = 3$. Es gibt mehrere Möglichkeiten, 2 ($= n - m$) der 5 Komponenten zu Null zu setzen. Wählen wir als Null-Kandidaten die Schlupf-Variablen x_3, x_4, x_5 aus, so ergeben sich 3 Möglichkeiten und also 3 Eckpunkte. Diese haben die Koordinaten

$$(1, 0, \mathbf{0}, \mathbf{0}, 1)$$

$$\left(\frac{1}{2}, 1, \frac{1}{2}, \mathbf{0}, \mathbf{0}\right)$$

$$(0, 1, \mathbf{0}, 1, \mathbf{0})$$

(Die beiden fettgedruckten Nullen sind jeweils gesetzt, und die übrigen 3 Komponenten ergeben sich aus $Ax = b$.)

Algorithmus (Simplex-Verfahren): Das Simplex-Verfahren läuft am Rand des Polyeders entlang der Kanten von einer Ecke zur nächsten, so dass sich der Wert der Zielfunktion $c^T x$ verbessert.

Dies bedeutet jeweils einen planmäßigen Austausch der Basislösung derart, dass eine Null-Komponente positiv wird und gleichzeitig eine positive Komponente im Vektor x Null wird. Ziel dabei: minimieren (oder max.) von $c^T x$.

Ausführung des Simplex-Verfahrens: vgl. Vorlesung oder Literatur zu Operations Research.

6.4 Konvexe Optimierung

(Vorgehen ähnlich wie in Abschnitt 6.2B)

Konvexe Mengen

Für $x, y \in \mathbb{R}^n$ ist

$$\theta x + (1 - \theta)y$$

eine Gerade durch x und y (für $\theta \in \mathbb{R}$).

Definition: Eine Menge C heißt *konvex*, wenn für alle $x, y \in C$ und alle θ mit $0 \leq \theta \leq 1$ gilt:

$$\theta x + (1 - \theta)y \in C.$$

Vgl. Abschnitt 4.4: Die konvexe Hülle von $p_1, \dots, p_k \in \mathbb{R}^n$ ist

$$\left\{ \sum_{i=1}^k s_i p_i \quad \text{für } s_i \geq 0 \forall i, \sum_{i=1}^k s_i = 1 \right\}.$$

Konvexe Funktionen

Der Definitionsbereich D einer Funktion f sei konvex. f heißt dann *konvex*, wenn für alle $x, y \in D$ und θ mit $0 \leq \theta \leq 1$ gilt

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

Konvexes Optimierungsproblem

(KO)

$f(x) \rightarrow \text{Min.}$ unter N.B.

$h_j(x) \leq 0 \quad (j = 1, \dots, l),$

wobei f, h_1, \dots, h_l alle konvex seien.

Vorteile dieser Annahmen:

1. Jedes lokale Minimum einer konvexen Funktion f ist auch globales Minimum.
2. Die Menge, welche durch $h_j(x) \leq 0 \quad (j = 1, \dots, l)$ definiert ist, ist konvex (sofern nicht leer).

Beispiele

- Beispiel (4) in Abschnitt 6.1
- Methode der kleinsten Quadrate, vgl. Abschnitt 4.8 (dort $f = \phi$; keine N.B.)
- Wenn A eine symmetrische und positiv-definite Matrix ($x^T A x > 0$ für $x \neq 0$) ist, dann ist die Funktion $x^T A x$ konvex.
- Lineare Programmierung: wichtiger Spezialfall, vgl. Abschnitt 6.3

Einfacher Fall von (KO): Wenn für ein x gilt $\text{grad } f(x) = 0$ und N.B. ist erfüllt, dann ist $x^* := x$ das Minimum. (vgl. Abschnitt 6.2)

Ansonsten: Anwenden der Karush-Kuhn-Tucker-Bedingungen.

Karush-Kuhn-Tucker-Bedingungen (KKT, manchmal auch nur Kuhn-Tucker-Bedingungen genannt) Betrachte (KO) mit der zusätzlichen Voraussetzung, dass die Funktionen f, h_1, \dots, h_l alle differenzierbar sind.

x^* sei ein Minimum. Dies ist äquivalent zur Aussage:

Es gibt l Lagrange-Multiplikatoren $y_j \geq 0$ ($j = 1, \dots, l$),
so dass

$$\frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^l y_j \frac{\partial h_j(x^*)}{\partial x_i} = 0 \quad \text{für } i = 1, \dots, n$$

$$\sum_{j=1}^l y_j h_j(x^*) = 0$$

Wegen $h \leq 0$ und $y \geq 0$ ist letztere Gleichung äquivalent zu l skalaren Gleichungen $y_j h_j(x^*) = 0$, für $j = 1, \dots, l$. Also bestehen die KKT-Bedingungen aus $n+l$ Gleichungen für die $n+l$ Unbekannten $x_1, \dots, x_n, y_1, \dots, y_l$. In Vektorschreibweise lassen sich die KKT-Bedingungen kurz schreiben

$$y \geq 0, \quad \text{grad}f + (Dh)^T y = 0, \quad y^T h = 0$$

dabei bezeichnet Dh die Jacobi-Matrix des Vektors h ; $(Dh)^T y = \sum_j y_j \text{grad} g_j$.

Beispiel (4) in Abschnitt 6.1:

$$f(x_1, x_2) := 2x_1^2 + x_2^2 = \max!$$

$$h_1(x_1, x_2) := -x_1 \leq 0$$

$$h_2(x_1, x_2) := -x_2 \leq 0$$

$$h_3(x_1, x_2) := x_1 - 1 \leq 0$$

$$h_4(x_1, x_2) := x_1 + x_2 - 2 \leq 0$$

Die Voraussetzungen sind erfüllt; $n = 2$ und $l = 4$.

Die ersten zwei ($i = 1, 2$) Gleichungen von KKT sind:

$$4x_1 + y_1 \cdot (-1) + y_2 \cdot 0 + y_3 \cdot 1 + y_4 \cdot 1 = 0$$

$$2x_2 + y_1 \cdot 0 + y_2 \cdot (-1) + y_3 \cdot 0 + y_4 \cdot 1 = 0$$

Die weiteren 4 Gleichungen sind klar. Als Lösung ergibt sich

$$(x_1, \dots, x_n, y_1, \dots, y_l) = (2, 0, 0, -8, 0, -8).$$