



Luca Lontos  
3. April 2022

## Inhaltsverzeichnis

<b>1</b>	<b>Interpolation</b>	<b>3</b>
1.1	Polynominterpolation	3
1.1.1	Lagrange	3
1.1.2	Newton	3
1.1.3	Fehlerabschätzung	3
1.2	Spline-Interpolation	4
1.2.1	Linear	4
1.2.2	Kubisch	4
<b>2</b>	<b>Numerische Integration</b>	<b>6</b>
2.1	Newton-Cotes-Quadratur	6
2.1.1	geschlossen	6
2.1.2	offen	6
2.2	Summierte Newton-Cotes	7
<b>3</b>	<b>Numerische Behandlung von Anfangswertproblemen gewöhnlicher Differentialgleichungen</b>	<b>8</b>
3.1	Einführung	8
3.1.1	Verfahren	8
3.2	Steife Differentialgleichungen	9
3.2.1	Stabilitätsgebiete einiger Verfahren	10
<b>4</b>	<b>Lineare Gleichungssysteme</b>	<b>11</b>
4.1	Problemstellung	11
4.2	Gaußsche Eliminationsverfahren, Dreieckszerlegung einer Matrix	11
4.2.1	Lösung gestaffelter Gleichungssysteme	11
4.2.2	Gaußsches Eliminationsverfahren	11
4.2.3	Pivotstrategie	11
4.2.4	LR-Zerlegung	11
4.2.5	Matrizenklassen, die keine Pivotsuche erfordern	12
4.3	Das Cholesky Verfahren	12
4.4	Fehlerabschätzung und Rundungsfehlereinfluss	12
4.4.1	Fehlerabschätzung für gestörte Gleichungssysteme	12
<b>5</b>	<b>Nichtlineare Gleichungssysteme</b>	<b>14</b>
5.2	Das Newton-Verfahren	14
5.2.1	Herleitung	14
5.2.2	Superlineare und quadratische lokale Konvergenz des Newton-Verfahrens	14
5.2.3	Globalisierung des Newton-Verfahrens	14
<b>6</b>	<b>Verfahren zur Eigenwert- und Eigenvektorberechnung</b>	<b>15</b>
6.1	Eigenwertprobleme	15
6.1.1	Grundlagen	15
6.1.3	Grundkonzepte numerischer Verfahren	15
6.1.4	Störungstheorie für Eigenwertprobleme	15

6.2	Die Vektoriteration . . . . .	16
6.2.1	Definition und Eigenschaften der Vektoriteration . . . . .	16
6.2.2	Die Vektoriteration nach v. Mises und Wielandt . . . . .	17
6.3	Das QR-Verfahren . . . . .	17
<b>7</b>	<b>Grundbegriffe der Statistik und Wahrscheinlichkeitstheorie</b>	<b>18</b>
7.1	Messreihen . . . . .	18
7.2	Lage- und Streumaßzahlen . . . . .	18
7.2.1	Lagemaßzahlen . . . . .	18
7.2.2	Streuungsmaße . . . . .	18
7.2.3	Zweidimensionale Messreihen . . . . .	19
7.2.4	Regressionsgerade . . . . .	19
7.3	Zufallsexperimente und Wahrscheinlichkeit . . . . .	19
7.3.1	Zufallsexperimente . . . . .	19
7.3.2	Wahrscheinlichkeit . . . . .	20
7.3.3	Elementare Formeln der Kombinatorik . . . . .	20
7.4	Bedingte Wahrscheinlichkeit, Unabhängigkeiten . . . . .	20
7.4.1	Bedingte Wahrscheinlichkeit . . . . .	20
7.4.2	Unabhängigkeit . . . . .	21
7.5	Zufallsvariablen und Verteilungsfunktion . . . . .	21
7.5.1	Beispiele für diskrete Verteilungen . . . . .	21
7.5.2	Beispiele für Stetige Verteilungen . . . . .	22
7.6	Erwartungswert und Varianz . . . . .	23
7.6.1	Rechenregeln . . . . .	23
7.7	Gesetz der großen Zahlen, zentraler Grenzwertsatz . . . . .	24
7.7.1	Das schwache Gesetz der großen Zahlen . . . . .	24
7.7.2	Zentraler Grenzwertsatz . . . . .	24
7.8	Testverteilungen und Quantilapproximationen . . . . .	25
7.8.1	Wichtige Anwendungsbeispiele . . . . .	26
<b>8</b>	<b>Schätzverfahren und Konfidenzintervalle</b>	<b>26</b>
8.1	Grundlagen zu Schätzverfahren . . . . .	26
8.2	Maximum-Likelihood-Schätzer . . . . .	27
8.2.1	Kochrezept für Berechnung . . . . .	27
8.3	Konfidenzintervalle . . . . .	28
8.3.1	Konstruktion von Konfidenzintervallen . . . . .	28
<b>9</b>	<b>Tests bei Normalverteilungsannahmen</b>	<b>29</b>
9.1	Grundlagen . . . . .	29
9.2	Wichtige Tests bei Normalverteilungsannahme . . . . .	29
9.3	Verteilungstests . . . . .	29

# 1 Interpolation

## 1.1 Polynominterpolation

$n$  = Anzahl Stützstellen - 1

### 1.1.1 Lagrange

$$p_n(x) = \sum_{k=0}^n y_k L_{k,n}(x) \quad \text{mit} \quad L_{k,n}(x) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j}$$

Trick effiziente/ übersichtliche Berechnung

k/j	$x_0$	$x_1$	$\dots$	$x_n$	$L_{k,n}$
$x_0$	$\backslash$	$\frac{x - x_1}{x_0 - x_1}$	$\dots$	$\frac{x - x_n}{x_0 - x_n}$	$\prod$ diese Zeile
$x_1$	$\frac{x - x_0}{x_1 - x_0}$	$\backslash$	$\dots$	$\frac{x - x_n}{x_1 - x_n}$	$\prod$ diese Zeile
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$x_n$	$\frac{x - x_0}{x_n - x_0}$	$\frac{x - x_1}{x_n - x_1}$	$\dots$	$\backslash$	$\prod$ diese Zeile

**Satz 1.1.1** es gibt genau ein Polynom  $p(x)$  vom Grad  $\leq n$ , das die Interpolationsbedingungen erfüllt, nämlich  $p_n(x)$

### 1.1.2 Newton

$$p_n(x) = \gamma_0 + \sum_{i=1}^n \gamma_i (x - x_0) \dots (x - x_{i-1}), \quad \gamma_i = f_{[x_0, \dots, x_i]}$$

$$f_{[x_j, \dots, x_{j+k}]} = \frac{f_{[x_{j+1}, \dots, x_{j+k}]} - f_{[x_j, \dots, x_{j+k-1}]}}{x_{j+k} - x_j} = \frac{f_{\text{unten}} - f_{\text{oben}}}{x_{\text{unten}} - x_{\text{oben}}}$$

Die Oberste Zeile ab dem Strich sind dann die  $\gamma_i$

$$\begin{array}{c|l} x_0 & f_{[x_0]} = y_0 \searrow \\ & f_{[x_0, x_1]} \searrow \\ x_1 & f_{[x_1]} = y_1 \swarrow \quad f_{[x_0, x_1, x_2]} \\ & f_{[x_1, x_2]} \swarrow \\ x_2 & f_{[x_2]} = y_2 \nearrow \\ \vdots & \end{array}$$

Hinweis: Man kann Zeilen vertauschen. Das heißt wenn viele  $f_{[x]} = 0$  lohnt es sich meist diese nach oben zu tauschen.

### 1.1.3 Fehlerabschätzung

**Satz 1.1.3** Voraussetzungen für max error Korollar:  $f$  ist  $n+1$  mal stetig diffbar

**Korollar 1.1.4** Wenn Satz 1.1.3 gilt:

Äquidistant:

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

Tschebyschev-Abszissen:

$$\max_{x \in [a, b]} |f(x) - p_n(x)| \leq \max_{x \in [a, b]} \frac{|f^{(n+1)}(x)|}{(n+1)!} \left( \frac{b-a}{2} \right)^{n+1} 2^{-n}$$

### Inverse Interpolation:

Sei  $f : [a, b] \rightarrow \mathbb{R}$  bijektiv. Sind dann alle  $(x_i, y_i), y_i = f(x_i)$  Stützpunkte von  $f$  dann sind  $(y_i, x_i)$  Stützpunkte für  $f^{-1}$  und eine Approximation kann durch Interpolation dieser gewonnen werden.

Hinweis:

$f$  ist auf Wertebereich von  $x$  stetig differenzierbar und aufgrund der Ableitung streng monoton wachsend  $\Rightarrow$  injektiv

Die Grenzen von  $x$  müssen genau auf die Grenzen von  $y$  abbilden, da dann (wegen monoton wachsend) die gesamte Zielmenge getroffen wird  $\Rightarrow$  surjektiv

$\Rightarrow$  bijektiv

---

## 1.2 Spline-Interpolation

---

**Definition 1.2.1** Eine Splinefunktion der Ordnung  $k$  zur Zerlegung  $\Delta$  ist eine Funktion  $s : [a, b] \rightarrow \mathbb{R}$  mit folgenden Eigenschaften:

- es gilt  $s \in C^{k-1}([a, b])$  (stetig diffbar)
- $s$  stimmt auf jedem Intervall  $[x_i, x_{i+1}]$

### Spline-Interpolation:

Zu einer Zerlegung  $\Delta = \{x_i : a = x_0 < x_1 < \dots < x_n = b\}$  und werten  $y_i \in \mathbb{R}, i = 0, \dots, n$  bestimme  $s \in S_{\Delta, k}$  mit

$$s(x_i) = y_i, i = 0, \dots, n$$

---

### 1.2.1 Linear

---

$$s(x) = s_i(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i} y_i + \frac{x - x_i}{x_{i+1} - x_i} y_{i+1} \quad \forall x \in [x_i, x_{i+1}]$$

mit Hilfsknoten  $x_{-1} < a$  und  $x_{n+1} > b$  gilt dann

$$\varphi_i(x) = \begin{cases} 0 & \text{falls } x \leq x_{i-1} \\ \frac{x - x_{i-1}}{x_i - x_{i-1}} & \text{falls } x \in [x_{i-1}, x_i] \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & \text{falls } x \in [x_i, x_{i+1}] \\ 0 & \text{falls } x > x_{i+1} \end{cases}$$
$$s(x) = \sum_{i=0}^n y_i \varphi_i(x), x \in [a, b]$$

**Satz 1.2.2** Zu einer Zerlegung  $\Delta = \{x_i : a = x_0 < x_1 < \dots < x_n = b\}$  von  $[a, b]$  und Werten  $y_i$  existiert genau ein interpolierender Spline

### Satz 1.2.3

$$\max_{x \in [a, b]} |f(x) - s(x)| \leq \frac{1}{8} \max_{x \in [a, b]} |f''(x)| h_{\max}^2$$

---

### 1.2.2 Kubisch

---

$$s_i(x) = \frac{1}{6} \left( \frac{(x_{i+1} - x)^3}{x_{i+1} - x_i} M_i + \frac{(x - x_i)^3}{x_{i+1} - x_i} M_{i+1} \right) + c_i(x - x_i) + d_i$$

mit  $h_i = x_{i+1} - x_i$

$$d_i = y_i - \frac{h_i^2}{6} M_i \quad c_i = \frac{y_{i+1} - y_i}{h_i} - \frac{h_i}{6} (M_{i+1} - M_i)$$

**Natürliche Randbedingungen:**

$$M_0 = M_n = 0 \quad b_0 = b_n = 0 \\ \lambda_0 = \lambda_n = 0 \quad \mu_0 = \mu_n = 1$$

**Hermite Randbedingungen:**

$$b_0 = \frac{y_1 - y_0}{h_0} - f'(a) \quad b_n = f'(b) - \frac{y_n - y_{n-1}}{h_{n-1}} \\ \lambda_0 = \frac{h_0}{3} \quad \mu_0 = \frac{h_0}{3} \quad \lambda_n = \frac{h_{n-1}}{6} \quad \mu_n = \frac{h_{n-1}}{3}$$

$$\begin{pmatrix} \mu_0 & \lambda_0 & & & \\ \frac{h_0}{6} & \frac{h_0+h_1}{3} & \frac{h_1}{6} & & \\ & \frac{h_1}{6} & \frac{h_1+h_2}{3} & \frac{h_2}{6} & \\ & & \dots & \dots & \dots \\ & & & \lambda_n & \mu_n \end{pmatrix} \begin{pmatrix} M_0 \\ M_1 \\ M_2 \\ \dots \\ M_n \end{pmatrix} = \begin{pmatrix} b_0 \\ \frac{y_2-y_1}{h_1} - \frac{y_1-y_0}{h_0} \\ \frac{y_3-y_2}{h_2} - \frac{y_2-y_1}{h_1} \\ \dots \\ b_n \end{pmatrix}$$

**Satz 1.2.6** Fehler Natürlich

$$|f(x) - s(x)| \leq \frac{h_{max}}{h_{min}} \sup_{\xi \in [a,b]} |f^{(4)}(\xi)| h_{max}^4$$

$$|f^{(k)}(x) - s^{(k)}(x)| \leq \frac{2h_{max}}{h_{min}} \sup_{\xi \in [a,b]} |f^{(4)}(\xi)| h_{max}^{4-k}$$

**Satz 1.2.7** Fehler Hermit

$$|f(x) - s(x)| \leq \frac{5}{384} \sup_{\xi \in [a,b]} |f^{(4)}(\xi)| h_{max}^4$$

$$|f^{(k)}(x) - s^{(k)}(x)| \leq \frac{2h_{max}}{h_{min}} \sup_{\xi \in [a,b]} |f^{(4)}(\xi)| h_{max}^{4-k}$$

---

## 2 Numerische Integration

---

### 2.1 Newton-Cotes-Quadratur

---

#### 2.1.1 geschlossen

---

**Definition 2.1.1** Eine Integrationsformel  $J(f) = \sum_{i=0}^n \beta_i f(x_i)$  heißt exakt vom Grad  $n$  falls sie alle Polynome bis mindestens vom Grad  $n$  exakt integriert

Allgemeiner Fehler

$$\int_a^b \|f(x) - p_n(x)\| dx \leq \frac{\|f^{(n+1)}(\xi)\|}{(n+1)!} (b-a)^{n+2} = \max_{x \in [a,b]} \frac{|f^{(n+1)}(x)|}{(n+1)!} (b-a)^{n+2}$$

Berechnung

$$x_i = a + ih \qquad i = 0, \dots, n, \qquad h = \frac{b-a}{n}$$

$$I_n(f) = h \sum_{i=0}^n \alpha_{i,n} f(x_i)$$

$n$	$h$	$\alpha_{i,n}$				$E_n(f)$	Name
1	$b-a$	$\frac{1}{2}$	$\frac{1}{2}$			$-\frac{f^{(2)}(\xi)}{12}h^3$	Trapezregel
2	$\frac{b-a}{2}$	$\frac{1}{3}$	$\frac{4}{3}$	$\frac{1}{3}$		$-\frac{f^{(4)}(\xi)}{90}h^5$	Simpson-Regel
3	$\frac{b-a}{3}$	$\frac{3}{8}$	$\frac{9}{8}$	$\frac{9}{8}$	$\frac{3}{8}$	$-\frac{3f^{(4)}(\xi)}{80}h^5$	3/8-Regel
4	$\frac{b-a}{4}$	$\frac{14}{45}$	$\frac{64}{45}$	$\frac{24}{45}$	$\frac{64}{45}$	$-\frac{8f^{(6)}(\xi)}{945}h^7$	Milne-Regel

---

#### 2.1.2 offen

---

Berechnung

$$x_i = a + ih \qquad i = 1, \dots, n+1, \qquad h = \frac{b-a}{n+2}$$

$$\tilde{I}_n(f) = h \sum_{i=1}^{n+1} \tilde{\alpha}_{i,n} f(x_i)$$

$n$	$h$	$\tilde{\alpha}_{i,n}$			$\tilde{E}_n(f)$	Name
0	$\frac{b-a}{2}$	2			$\frac{f^{(2)}(\xi)}{3} h^3$	Rechteckregel
1	$\frac{b-a}{3}$	$\frac{3}{2}$	$\frac{3}{2}$		$\frac{3f^{(2)}(\xi)}{4} h^3$	Rechteckregel
2	$\frac{b-a}{4}$	$\frac{8}{3}$	$-\frac{4}{3}$	$\frac{8}{3}$	$\frac{28f^{(4)}(\xi)}{90} h^5$	Rechteckregel

---

## 2.2 Summierte Newton-Cotes

---

Newton Cotes nur auf kleinen Intervallen präzise in kleinen Intervallen

Idee: Aufteilen in Teilintervalle  $m$

$$N = n \cdot m$$

$$x_i = a + ih$$

$$H = \frac{b-a}{m}$$

$$i = 0, \dots, N$$

$$h = \frac{b-a}{N}$$

### Summierte Trapezregel

(geschlossen,  $n = 1$ ,  $h = \frac{b-a}{m}$ )

$$S_N^{(1)}(f) = \frac{h}{2} \sum_{j=0}^{m-1} (f(x_j) + f(x_{j+1}))$$

$$\text{Fehler: } R_N^{(1)}(f) = -\frac{f''(\xi)}{12}(b-a)h^2$$

### Summierte Simpson-Regel

(geschlossen,  $n = 2$ ,  $h = \frac{b-a}{2m}$ )

$$S_N^{(2)}(f) = \frac{h}{3} \sum_{j=0}^{m-1} (f(x_{2j}) + 4f(x_{2j+1}) + f(x_{2j+2}))$$

$$\text{Fehler: } R_N^{(2)}(f) = -\frac{f^{(4)}(\xi)}{180}(b-a)h^4$$

### Summierte Rechteck-Regel

(offen,  $n = 0$ ,  $2m = N$ ,  $h = \frac{b-a}{N}$ )

$$\tilde{S}_N^{(0)} = 2h \sum_{j=1}^m f(x_{2j-1})$$

$$\text{Fehler: } \tilde{R}_N^{(0)}(f) = \frac{f''(\xi)}{6}(b-a)h^2$$

---

### 3 Numerische Behandlung von Anfangswertproblemen gewöhnlicher Differentialgleichungen

---

#### 3.1 Einführung

---

$t_j$  sind die Stützpunkte. Also untere Grenze + Schrittweite ( $a + j * h$ )  
 $y'(t) = f(t, y(t))$

---

##### 3.1.1 Verfahren

---

**Expliziter Euler**

$$u_{j+1} := u_j + hf(t_j, u_j)$$

**Impliziter Euler**

$$u_{j+1} := u_j + hf(t_{j+1}, u_{j+1})$$

**Implizite Trapezregel**

$$u_{j+1} := u_j + \frac{h}{2}(f(t_j, u_j) + f(t_{j+1}, u_{j+1}))$$

**Verfahren von Heun, erstes Runge-Kutta-Verfahren 2. Ordnung**

$$u_{j+1} := u_j + \frac{h}{2}(f(t_j, u_j) + f(t_{j+1}, u_j + hf(t_j, u_j)))$$

kann man auch schreiben als:

$$u_{j+1} := u_j + \frac{h}{2}(k_1 + k_2)$$

mit  $k_1 = f(t_j, u_j)$ ,  $k_2 = f(t_{j+1}, u_j + hk_1)$

**Modifizierter Euler, zweites Runge-Kutta-Verfahren 2. Ordnung**

$$u_{j+1} = u_j + hk_2$$

mit  $k_1 = f(t_j, u_j)$ ,  $k_2 = f(t_j + \frac{h}{2}, u_j + \frac{h}{2}k_1)$

**Klassisches Runge-Kutta-Verfahren 4. Ordnung (RK4)**

$$u_{j+1} := u_j + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$k_1 := f(t_j, u_j)$$

$$k_2 := f(t_j + \frac{h}{2}, u_j + \frac{h}{2}k_1)$$

$$k_3 := f(t_j + \frac{h}{2}, u_j + \frac{h}{2}k_2)$$

$$k_4 := f(t_{j+1}, u_j + hk_3)$$

**r-stufiges explizites Runge-Kutta-Verfahren**

$$k_i(t, u, h) = k_i := f(t + \gamma_i h, u + h \sum_{j=1}^{i-1} \alpha_{ij} k_j), i = 1 \dots r$$

$$\phi(t, h; u) = \sum_{i=1}^r \beta_i k_i$$

**Butcher – Schema**

$\gamma_1$	0				
$\gamma_2$	$\alpha_{21}$	0			
$\gamma_3$	$\alpha_{31}$	$\alpha_{32}$	0		
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\ddots$	
$\gamma_r$	$\alpha_{r1}$	$\dots$	$\dots$	$\alpha_{r,r-1}$	0
	$\beta_1$	$\beta_2$	$\dots$	$\beta_{r-1}$	$\beta_r$



---

**Satz 3.1.7 Konsistenzordnungen  $p$** 

$p = 1$  falls

$$\sum_{i=1}^r \beta_i = 1$$

$p = 2$  falls zusätzlich

$$\sum_{i=1}^r \beta_i \gamma_i = \frac{1}{2}$$

$p = 3$  falls zusätzlich

$$\sum_{i=1}^r \beta_i \gamma_i^2 = \frac{1}{3}$$

$$\sum_{i,j=1}^r \beta_i \alpha_{ij} \gamma_j = \frac{1}{6}$$

$p = 4$  falls zusätzlich

$$\sum_{i=1}^r \beta_i \gamma_i^3 = \frac{1}{4},$$
$$\sum_{i,j=1}^r \beta_i \alpha_{ij} \gamma_j^2 = \frac{1}{12}$$

$$\sum_{i,j=1}^r \beta_i \gamma_i \alpha_{ij} \gamma_j = \frac{1}{8}$$
$$\sum_{i,j,k=1}^r \beta_i \alpha_{ij} \alpha_{jk} \gamma_k = \frac{1}{24}$$

---

**3.2 Steife Differentialgleichungen**

---

**Definition 3.2.2** ein Anfangswertproblem

$$y'(t) = Ay(t) + c$$

$$y(a) = y_0$$

heißt steif, wenn die Realteile der Eigenwerte von  $A$  nichtpositiv sind und  $A$  Eigenwerte mit  $\operatorname{Re}(\lambda) \ll -1$  und Eigenwerte  $\lambda_i$  mit schwach negativen Realteil besitzt.

**Definition 3.2.3** Ein Verfahren heißt  $A$ -Stabil wenn seine Anwendung auf das Modellproblem

$$y' = \lambda y, \quad y(0) = 1, \quad \text{mit } \lambda \in \mathbb{C}, \quad \operatorname{Re}(\lambda) < 0$$

mit der Lösung

$$y(t) = e^{\lambda t}$$

für jede schrittweite  $h > 0$  eine Folge  $\{u_j\}_{j \in \mathbb{N}_0}$  produziert mit

$$|u_{j+1}| \leq |u_j|, \quad \forall j \geq 0$$

bei Einschrittverfahren gilt bei Anwendung auf das Modellproblem die Beziehung

$$u_{j+1} = R(q)u_j \text{ mit } q = \lambda h$$

**Definition 3.2.5** Man nennt  $R$  die Stabilitätsfunktion des Einschrittverfahrens. Die Menge

$$S = \{q \in \mathbb{C} : |R(q)| \leq 1\}$$

heißt Stabilitätsgebiet des Einschrittverfahrens

**Definition 3.2.6** Ein Verfahren heißt  $L$ -Stabil, wenn es  $A$ -Stabil ist und die Stabilitätsfunktion zudem erfüllt

$$\lim_{q \rightarrow -\infty} R(q) = 0$$

---

### 3.2.1 Stabilitätsgebiete einiger Verfahren

---

#### Expliziter Euler

$$\begin{aligned}u_{j+1} &= (1 + \lambda h)u_j \\R(q) &= 1 + q \\S &= \{q \in \mathbb{C} : |1 + q| \geq 1\}\end{aligned}$$

Nicht A-Stabil (gilt für alle Expliziter RKs)

#### Impliziter Euler

$$\begin{aligned}u_{j+1} &= \frac{1}{1 - \lambda h}u_j \\R(q) &= \frac{1}{1 - q}; q \neq 1 \\S &= \{q \in \mathbb{C} : |1 - q| \geq 1\} \supset \{q \in \mathbb{C} : \operatorname{Re}(q) < 0\}\end{aligned}$$

A und L Stabil

#### Implizite Trapezregel

$$\begin{aligned}u_{j+1} &= \frac{1 + \lambda h/2}{1 - \lambda h/2}u_j \\R(q) &= \frac{1 + q/2}{1 - q/2}; q \neq 2 \\S &= \{q \in \mathbb{C} : \operatorname{Re}(q) \leq 0\}\end{aligned}$$

A aber nicht L Stabil

#### Implizite Runge-Kutta-Verfahren

$$k_i = k_i(t, u, h) := f(t + \gamma_i h, u + h \sum_{l=1}^r \alpha_{il} k_l), \quad i = 1, \dots, r$$

$$\phi(t, h; u) = \sum_{i=1}^r \beta_i k_i$$

#### Butcher-Schema

$\gamma_1$	$\alpha_{11}$	$\dots$	$\dots$	$\alpha_{1,r-1}$	$\alpha_{1,r}$
$\gamma_2$	$\alpha_{21}$	$\dots$	$\dots$	$\alpha_{2,r-1}$	$\alpha_{2,r}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\gamma_r$	$\alpha_{r1}$	$\dots$	$\dots$	$\alpha_{r,r-1}$	$\alpha_{r,r}$
	$\beta_1$	$\beta_2$	$\dots$	$\beta_{r-1}$	$\beta_r$

$$u_j + 1 = (1 + q\beta^T(I - qA)^{-1}\mathbf{1})u_j$$

$$R(q) = 1 + q\beta^T(I - qA)^{-1}\mathbf{1} = \frac{\det(I - qA + q\mathbf{1}\beta^T)}{\det(I - qA)}$$

---

## 4 Lineare Gleichungssysteme

---

### 4.1 Problemstellung

---

Lineares Gleichungssystem: Gesucht ist eine Lösung  $x$  von

$$Ax = b$$

mit

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \dots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

**Definition 4.1.1** Das LGS hat eine Lösung g.d.w.

$$\text{rang}(a) = \text{rang}(A, b)$$

---

### 4.2 Gaußsche Eliminationsverfahren, Dreieckszerlegung einer Matrix

---

#### 4.2.1 Lösung gestaffelter Gleichungssysteme

---

Subsection Damit Nummerierung stimmt

---

#### 4.2.2 Gaußsches Eliminationsverfahren

---

1. Wähle ein Pivotelement  $a_{rk}^{(k)} \neq 0, k \leq r \leq n$  vertausche Zeile  $k$  und  $r \rightsquigarrow (\tilde{A}^{(k)}, \tilde{b}^{(k)})$
2. Für  $i = k + 1, \dots, n$ : Subtrahiere das  $l_{ik}$ -fache mit

$$l_{ik} = \frac{\tilde{a}_{ik}^{(k)}}{\tilde{a}_{kk}^{(k)}}$$

der  $k$ -ten Gleichung von der  $i$ -ten Gleichung

---

#### 4.2.3 Pivotstrategie

---

- Spaltenpivot: wähle  $k \leq r \leq n$  mit

$$|a_{rk}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|$$

- Vollständige Pivotsuche: Bestimme  $k \leq r \leq n, k \leq s \leq n$  mit

$$|a_{rs}^{(k)}| = \max_{k \leq i, j \leq n} |a_{ij}^{(k)}|$$

---

#### 4.2.4 LR-Zerlegung

---

Finden einer Zerlegung von  $A$  der Form

$$LR = PA(Q)$$

Gaußsche Eliminationsverfahren liefert dies "gratis" mit

$$R = A^{(n)} \quad c = b^{(n)} \quad L = I + L^{(n)}$$

um  $P$  zu erhalten fängt man mit  $I$  an und tauscht die Zeilen jedes mal wenn man in  $A$  auch die Zeilen tauscht,  $Q$  das selbe nur mit Spalten (nur bei Vollständige Pivotsuche)

nachdem Zerlegung gefunden kann einfach  $Ax = \tilde{b}$  gelöst werden

- Löse  $Lz = P\tilde{b}$
- Löse  $Ry = z$
- Lösung:  $x = Qy$

---

#### 4.2.5 Matrizenklassen, die keine Pivotsuche erfordern

---

- $A = A^T$  ist symmetrisch positiv definit, also

$$x^T A x > 0 \forall x \in \mathbb{R} \setminus \{0\}$$

- $A$  ist strikt diagonaldominant, d.h.,

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|, i = 1, \dots, n$$

- $A$  ist M-Matrix, dh. es gilt

$$a_{ii} > 0, i = 1, \dots, n$$

$$a_{ij} \leq 0, i \neq j$$

$$D^{-1}(A - D) \text{ hat lauter Eigenwerte vom Betrag } < 1, D = \text{diag}(a_{11}, \dots, a_{nn})$$

---

#### 4.3 Das Cholesky Verfahren

---

**Definition 4.3.1** eine reelle Matrix heißt positiv definit falls

$$A = A^T, \quad x^T A x > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

positiv semi definit falls

$$A = A^T, \quad x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

**Satz 4.3.2** Es sei  $A$  positiv definit, Dann gibt es genau eine untere Dreiecksmatrix  $L$  mit positiven Diagonaleinträgen  $l_{ii} > 0$ , so dass

$$LL^T = A \text{ (Cholesky-Zerlegung)}$$

Ferner besitzt  $A$  eine eindeutige Dreieckszerlegung

$$\tilde{L}\tilde{R} = A,$$

wobei  $\tilde{L} = LD^{-1}$ ,  $\tilde{R} = DL^T$

**Satz 4.3.3** Cholesky-Verfahren zur Berechnung der Zerlegung  $LL^T = A$

Für  $j = 1, \dots, n$

$$l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2}$$

Falls Wurzel nicht existiert, STOPP,  $A$  nicht definit

Für  $i = j + 1, \dots, n$ :

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}}{l_{jj}}$$

---

#### 4.4 Fehlerabschätzung und Rundungsfehlereinfluss

---

---

##### 4.4.1 Fehlerabschätzung für gestörte Gleichungssysteme

---

**Definition 4.4.1** Eine Vektornorm auf  $\mathbb{R}^n$  ist eine Abbildung  $x \in \mathbb{R}^n \mapsto \|x\| \in [0, \infty[$  mit den Eigenschaften

- $\|x\| = 0$  nur für  $x = 0$
- $\|ax\| = |a| \|x\|$  für alle  $a \in \mathbb{R}$  und alle  $x \in \mathbb{R}^n$
- $\|x + y\| \leq \|x\| + \|y\|$  für alle  $x, y \in \mathbb{R}^n$  (Dreiecksungleichung)

**Definition 4.4.2** Eine Vektornorm induziert eine Matrixnorm, diese haben die Eigenschaften

- a)  $\|A\| = 0$  nur für  $A = 0$
- b)  $\|aA\| = |a| \|A\|$  für alle  $a \in \mathbb{R}$  und alle  $A \in \mathbb{R}^{n \times n}$
- c)  $\|A + B\| \leq \|A\| + \|B\|$  für alle  $A, B \in \mathbb{R}^{n \times n}$  (Dreiecksungleichung)
- d)  $\|Ax\| \leq \|A\| \|x\|$  für alle  $x \in \mathbb{R}^n$  und alle  $A \in \mathbb{R}^{n \times n}$  (Verträglichkeitsbedingung)
- e)  $\|AB\| \leq \|A\| \|B\|$  für alle  $A, B \in \mathbb{R}^{n \times n}$  (Submultiplikativität)

Beispiele hierfür sind:

$\ x\ _2 = \sqrt{x^T x}$	induziert	$\ A\ _2 = \sqrt{\lambda_{\max}(A^T A)}$	
$\ x\ _1 = \sum_{i=1}^n  x_i $	induziert	$\ A\ _1 = \max_{j=1, \dots, n} \sum_{i=1}^n  a_{ij} $	(Spaltensummennorm)
$\ x\ _\infty = \max_{i=1, \dots, n}  x_i $	induziert	$\ A\ _\infty = \max_{i=1, \dots, n} \sum_{j=1}^n  a_{ij} $	(Zeilensummennorm)

**Definition 4.4.3** Sei  $A \in \mathbb{R}^{n \times n}$  invertierbar und sein  $\|\cdot\|$  eine induzierte Matrixnorm. Dann heißt die Zahl  $\text{cond}(A) = \|A\| \|A^{-1}\|$  die Konditionszahl von  $A$  bezüglich der Matrixnorm

**Satz 4.4.4** (Störeinfluss von Matrix und rechter Seite). Sei  $A \in \mathbb{R}^{n \times n}$  invertierbar,  $b, \Delta b \in \mathbb{R}^n, b \neq 0$  und  $\Delta A \in \mathbb{R}^{n \times n}$  mit  $\|\Delta A\| < 1/\|A^{-1}\|$  mit einer beliebigen durch eine Norm  $\|\cdot\|$  auf  $\mathbb{R}^n$  induzierten Matrixnorm  $\|\cdot\|$ . Ist  $x$  die Lösung von

$$Ax = b$$

und  $\tilde{x}$  die Lösung von

$$(A + \Delta A)\tilde{x} = b + \Delta b$$

dann gilt

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

---

## 5 Nichtlineare Gleichungssysteme

---

### 5.2 Das Newton-Verfahren

---

#### 5.2.1 Herleitung

---

Satz 5.2.1

$$\begin{aligned}x^{(k+1)} &= x^{(k)} - J_F(x^{(k)})^{-1} F(x^{(k)}) \\x^{(k+1)} &= x^{(k)} + s^{(k)} \\s^{(k)} &\text{ Lösung von } J_F(x^{(k)}) s^{(k)} = -F(x^{(k)})\end{aligned}$$

Satz 5.2.2 Algorithmus für Lokales Newton-Verfahren

1. Falls  $F(x^{(k)}) = 0$  STOPP mit Ergebnis  $x^{(k)}$
2. Berechne  $s^{(k)}$
3.  $x^{k+1} = x^{(k)} + s^{(k)}$

---

#### 5.2.2 Superlineare und quadratische lokale Konvergenz des Newton-Verfahrens

---

Satz 5.2.3 (Schnelle Lokale Konvergenz des Newton-Verfahrens) Sei  $F$  stetig diffbar und sei  $\tilde{x} \in \mathbb{R}$  ein Punkt mit  $F(\tilde{x}) = 0$  und  $F'(\tilde{x})$  nichtsingulär. Dann gibt es  $\delta > 0$ , so dass folgende Aussagen gelten:

- i)  $\tilde{x}$  ist die einzige Nullstelle in der  $\delta$ -Kugel

$$B_\delta(\tilde{x}) := \{x \in \mathbb{R}^n : \|x - \tilde{x}\|_2 < \delta\}$$

- ii) Für alle  $x_0 \in B_\delta(\tilde{x})$  Terminiert der Algorithmus 5.2.2 entweder mit  $x^{(k)} = \tilde{x}$  oder erzeugt eine Folge  $(x^{(k)}) \subset B_\delta(\tilde{x})$ , die superlinear gegen  $\tilde{x}$  konvergiert d. h.

$$\lim_{k \rightarrow \infty} x^{(k)} = \tilde{x} \quad \text{wobei} \quad \|x_{k+1} - \tilde{x}\|_2 \leq v_k \|x^{(k)} - \tilde{x}\|_2$$

mit einer Nullfolge  $v_k \searrow 0$

- iii) ist  $F'$  Lipschitz-stetig auf  $B_\delta(\tilde{x})$  mit Konstante  $L$  dann konvergiert  $(x_k)$  sogar quadratisch gegen  $\tilde{x}$ , d.h.

$$\lim_{k \rightarrow \infty} x^{(k)} = \tilde{x} \quad \text{wobei} \quad \|x_{k+1} - \tilde{x}\|_2 \leq C \|x^{(k)} - \tilde{x}\|_2^2$$

wobei für  $\delta > 0$  klein genug  $C = L \|F'(\tilde{x})^{-1}\|_2$  gewählt werden kann. Automatisch wenn zwei mal diffbar in  $B$

---

#### 5.2.3 Globalisierung des Newton-Verfahrens

---

Schrittweitenwahl nach Armijo:

sei  $\delta \in ]0, \frac{1}{2}[$  (z.B.  $10^{-3}$ ) fest gegeben. Wähle das Größte  $\sigma_k \in \{1, \frac{1}{2}, \frac{1}{4}, \dots\}$  mit

$$\|F(x^{(k)} + \sigma_k s^{(k)})\|_2^2 \leq \|F(x^{(k)})\|_2^2 - 2\delta\sigma_k \|F(x^{(k)})\|_2^2$$

Satz 5.2.4 Algo für das Globalisierte Newton-Verfahren:

1. Falls  $F(x^{(k)}) = 0$  STOPP mit Ergebnis  $x^{(k)}$
2. Berechne  $s^{(k)}$
3. Bestimme  $\sigma_k$  nach Armijo
4.  $x_{k+1} = x^{(k)} + \sigma_k s^{(k)}$

---

## 6 Verfahren zur Eigenwert- und Eigenvektorberechnung

---

### 6.1 Eigenwertprobleme

---

#### 6.1.1 Grundlagen

---

**Definition 6.1.1** Eine Zahl  $\lambda \in \mathbb{C}$  heißt Eigenwert einer Matrix  $A \in \mathbb{C}^{n \times n}$ , wenn es einen Vektor  $x \in \mathbb{C}^n$  gibt mit

$$Ax = \lambda x$$

Jeder solche Vektor  $x \in \mathbb{C}^n$  heißt (Rechts-)Eigenvektor zum Eigenwert  $\lambda$ . Die Menge  $\sigma(A)$  aller Eigenwerte von  $A$  heißt Spektrum von  $A$ . Der Unterraum

$$\text{Eig}_A(\lambda) := \{x \in \mathbb{C}^n : (A - \lambda I)x = 0\}$$

Ist der Eigenraum von  $A$  zum Eigenwert  $\lambda$ . Seine Dimension

$$\gamma(\lambda) := \dim \text{Eig}_A(\lambda) = n - \text{Rang}(A - \lambda I)$$

ist die geometrische Vielfachheit von  $\lambda$  und gibt die Maximalzahl linear unabhängigen Eigenvektoren zu  $\lambda$  an.  $\lambda$  ist EW zu  $A$  wenn:

$$\chi(\lambda) := \det(A - \lambda I) = 0$$

also wenn  $\lambda$  Nullstelle des charakteristischen Polynoms  $\chi(\mu)$  von  $A$  ist

$$\begin{aligned}\chi(\mu) &= (-1)^n \mu^n + (-1)^{n-1} \mu^{n-1} \text{Spur}(A) + \dots + \det(A) \\ &= (-1)^n (\mu - \lambda_1)^{\nu_1} \dots (\mu - \lambda_k)^{\nu_k}\end{aligned}$$

man nennt  $\nu(\lambda_i) = \nu_i$  die algebraische Vielfachheit von  $\lambda_i$ . Es gilt:

$$\gamma(\lambda_i) \leq \nu(\lambda_i)$$

**Definition 6.1.2** Sei  $A \in \mathbb{C}^{n \times n}$  beliebig:

- a) Ist  $\lambda$  EW von  $A$ , so ist  $\lambda$  EW von  $A^T$  und  $\bar{\lambda}$  EW von  $A^H := \bar{A}^T$
- b) für jede nichtsinguläre Matrix  $T \in \mathbb{C}^{n \times n}$  hat die zu  $A$  ähnliche Matrix  $B := T^{-1}AT$ , das selbe charakteristische Polynom und die selben Eigenwerte wie  $A$ . Ist  $x$  Eigenvektor von  $A$ , so ist  $y := T^{-1}x$  Eigenvektor von  $B$
- c) Ist  $A$  hermitisch also  $A^H = A$ , dann hat  $A$  lauter reelle Eigenwerte. Ist  $A$  unitär, also  $A^H = A^{-1}$ , so gilt  $|\lambda| = 1$  für jeden Eigenwert  $\lambda$

---

#### 6.1.3 Grundkonzepte numerischer Verfahren

---

##### Vektoriteration

$$x^{(k+1)} = \frac{Bx^{(k)}}{\|Bx^{(k)}\|}, \text{ mit } k = 0, 1, \dots$$

**Ähnlichkeitstransformation auf einfachere Gestalt** Nach Definition 6.1.2 bleiben die Eigenwerte einer Matrix  $A$  bei einer Ähnlichkeitstransformation erhalten. Daher liegt nahe  $A$  durch

$$A^{(0)} := A \rightarrow A^{(1)} \rightarrow \dots \rightarrow A^{(k+1)} = T_k^{-1}A^{(k)}T_k$$

in eine einfachere Form zu überführen

---

#### 6.1.4 Störungstheorie für Eigenwertprobleme

---

**Satz 6.1.3** Bezeichnet  $\lambda_i(A)$ ,  $i = 1, \dots, n$  die angeordneten EWs einer Matrix  $A \in \mathbb{C}^{n \times n}$ . Dann sind die Abbildungen

$$A \in \mathbb{C}^{n \times n} \mapsto \lambda_i(A)$$

stetig. Eigenwerte hängen also stetig von der Matrix ab

**Satz 6.1.4** Es sei  $A = (a_{ij}) \in \mathbb{C}^{n \times n}$  beliebig

a)

$$\sigma(A) \subset \bigcup_{i=1}^n K_i$$

$$K_i = \left\{ \mu \in \mathbb{C} : |\mu - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}| \text{ (Zeilensumme ohne Diagonale)} \right\}$$

b) ist die Vereinigung  $G_1$  von  $k$  Gershgorin-Kreisen disjunkt von der Vereinigung  $G_2$  der restlichen  $n-k$  Gershgorin-Kreise, dann enthält  $G_1$  genau  $k$  EWs und  $G_2$  genau  $n-k$  EWs von  $A$

**Satz 6.1.5** (Bauer/Fike). Es sei  $A \in \mathbb{C}^{n \times n}$  diagonalisierbar, also

$$T^{-1}AT = \text{diag}(\lambda_1, \dots, \lambda_n) =: D$$

Dann gilt für jede Matrix  $\Delta A \in \mathbb{C}^{n \times n}$

$$\forall \mu \in \sigma(A + \Delta A) : \min_{i=1, \dots, n} |\mu - \lambda_i| \leq \text{cond}_2(T) \|\Delta A\|_2$$

mit  $\text{cond}_2(T) := \|T\|_2 \|T^{-1}\|_2$ ;  $\text{cond}_2(T) := 1$  für  $A$  hermitisch

## 6.2 Die Vektoriteration

### 6.2.1 Definition und Eigenschaften der Vektoriteration

**Definition 6.2.1** für eine Matrix  $B \in \mathbb{C}^{n \times n}$  ist die Vektoriteration gegeben durch

$$z^{(k+1)} = \frac{1}{\|Bz^{(k)}\|} Bz^{(k)}, \text{ mit } k = 0, 1, \dots$$

mit einem Startvektor  $z^{(0)} \in \mathbb{C}^n \setminus \{0\}$

mit geeigneter Wahl von  $B$  erhält man so Näherungen an einen Eigenvektor zu  $\lambda$ . Eine Eigenwertnäherung für  $\lambda$  erhält man dann durch den **Rayleighquotienten**

$$R(z^{(k)}, B) = \frac{(z^{(k)})^H B z^{(k)}}{(z^{(k)})^H z^{(k)}}$$

**Satz 6.2.2** Es sei  $B \in \mathbb{C}^{n \times n}$  diagonalisierbar mit EWs  $\lambda_1, \dots, \lambda_n$

$$\lambda_1, \dots, \lambda_r, |\lambda_r| > |\lambda_{r+1}| \geq \dots \geq |\lambda_n| \quad (1)$$

mit  $r < n$ . Falls der Startvektor  $z^{(0)}$  einen Anteil in  $\text{Eig}_B(\lambda_1)$  besitzt, gilt für die Vektoriteration

$$R(z^{(k)}, B) = \frac{(z^{(k)})^H B z^{(k)}}{(z^{(k)})^H z^{(k)}} = \lambda_1 + O(q^k) \text{ für } k \rightarrow \infty, q = \frac{|\lambda_{r+1}|}{|\lambda_1|} < 1$$

zudem gilt

$$z^{(k)} = \frac{\lambda_1^k}{|\lambda_1|^k} \frac{x_1}{\|x_1\|} + O(q^k), k \geq 1$$

wobei  $x_1$  den Anteil von  $z^{(0)}$  in  $\text{Eig}_B(\lambda_1)$  bezeichnet



---

## 6.2.2 Die Vektoriteration nach v. Mises und Wielandt

---

Sei  $A \in \mathbb{C}^{n \times n}$  gegeben

### Einfache Vektoriteration nach von Mises

Erhält man durch die Wahl von  $B = A$  **Inverse Vektoriteration von Wielandt** für  $\mu \neq \lambda_j$  hat die Matrix  $B = (A - \mu I)^{-1}$  die EWs

$$\mu_i = \frac{1}{\lambda_i - \mu}$$

$$z^{(k+1)} = \frac{\hat{z}^{(k+1)}}{\|\hat{z}^{(k+1)}\|} \text{ mit } \hat{z}^{(k+1)} = (A - \mu I)^{-1} z^{(k)}$$

Lösbar durch

$$(A - \mu I) \hat{z}^{(k+1)} = z^{(k)}$$

hat im Falle

$$q := \max_{1 \leq i \leq n, \lambda_i \neq \lambda_j} \frac{|\lambda_j - \mu|}{|\lambda_i - \mu|} < 1$$

die Konvergenzeigenschaften

$$\begin{aligned} R(z^{(k)}, (A - \mu I)^{-1}) &= \frac{(z^{(k)})^H \hat{z}^{(k+1)}}{(z^{(k)})^H z^{(k)}} = \frac{1}{\lambda_i - \mu} + O(q^k) \\ z^{(k)} &= \frac{|\lambda_j - \mu|^k}{(\lambda_j - \mu)^k} \frac{x_j}{\|x_j\|} + O(q^k) \end{aligned}$$

wobei  $x_j$  den Anteil von  $z^{(0)}$  in  $\text{Eig}_A(\lambda_j) = \text{Eig}_{(A - \mu I)^{-1}}(1/(\lambda_j - \mu))$ . Ist  $A$  dabei hermitisch, so erfüllt der Rayleigh-Quotient

$$R(z^{(k)}, (A - \mu I)^{-1}) = \frac{(z^{(k)})^H \hat{z}^{(k+1)}}{(z^{(k)})^H z^{(k)}} = \frac{1}{\lambda_i - \mu} + O(q^{2k})$$

---

## 6.3 Das QR-Verfahren

---

### Definition 6.3.1 Algorithmus QR-Verfahren

Sei  $A \in \mathbb{C}^{n \times n}$  eine gegebene Matrix

- Setze  $A^{(1)} := A$
- für  $l = 1, 2, \dots$ : Berechne

$$\begin{aligned} A^{(l)} &:= Q_l R_l, \\ A^{(l+1)} &:= R_l Q_l \end{aligned}$$

$$Q \in \mathbb{C}^{n \times n} \text{ unitär}$$

$$R_l \in \mathbb{C}^{n \times n} \text{ obere Dreiecksmatrix,}$$

---

## 7 Grundbegriffe der Statistik und Wahrscheinlichkeitstheorie

---

### 7.1 Messreihen

---

- quantitativ diskret: Merkmale sind Ganze zahlen
- quantitativ stetig: Die Merkmale sind reelle Zahlen

**Definition 7.1.1** Zu  $x_1, x_2, \dots, x_n$  ist  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$  die geordnete Messreihe mit  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ . Die empirische Verteilungsfunktion hierzu ist

$$F(z; x_1, x_2, \dots, x_n) = F(z) = \frac{\text{Zahl der } x_i \text{ mit } x_i \leq z}{n}$$

Mit  $r-1$  Zahlen  $a_1 < a_2 < \dots < a_{r-1}$  entsteht eine Unterteilung in  $r$  Klassen.

$$\mathbb{R} = ]-\infty, a_1] \cup ]a_1, a_2] \cup \dots \cup ]a_{r-2}, a_{r-1}] \cup ]a_{r-1}, \infty[$$

Man erhält die relativen Klassenhäufigkeiten für diese durch

$$F(a_1), F(a_2) - F(a_1), \dots, F(a_{r-1}) - F(a_{r-2}), 1 - F(a_{r-1})$$

Für Histogramm wähle  $a_0 < \min(a_1, x_{(1)})$ ,  $a_r > \max(a_{r-1}, x_{(n)})$ . Für jede Klassenhäufigkeit: Teile durch Klassenbreite und zeichne in Diagramm

---

### 7.2 Lage- und Streumaßzahlen

---

#### 7.2.1 Lagemaßzahlen

---

Arithmetisches Mittel:

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

Median:

$$\tilde{x} = \begin{cases} x_{(\frac{n}{2})} & n \text{ wenn gerade} \\ x_{(\frac{n+1}{2})} & n \text{ wenn ungerade} \end{cases} \quad (2)$$

p-Quantil ( $0 < p < 1$ ):

$$x_p = \begin{cases} x_{(np)} & n \text{ wenn } np \text{ ganzzahlig} \\ x_{(\lfloor np \rfloor + 1)} & n \text{ wenn } np \text{ nicht ganzzahlig} \end{cases} \quad (3)$$

$\alpha$  gestutztes Mittel ( $0 < \alpha < 0.5$ ):

$$\bar{x}_\alpha = \frac{1}{n - 2k}(x_{(k+1)} + \dots + x_{(n-k)}), k = \lfloor n\alpha \rfloor$$

---

#### 7.2.2 Streuungsmaße

---

Empirische Varianz oder empirische Stichprobenvarianz:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Empirische Streuung:

$$s = \sqrt{s^2}$$

Spannweite:

$$v = x_{(n)} - x_{(1)}$$

Quartilabstand:

$$q = x_{0.75} - x_{0.25}$$

---

### 7.2.3 Zweidimensionale Messreihen

---

Arithmetisches Mittel:

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

$$\bar{y} = \frac{1}{n}(y_1 + y_2 + \dots + y_n)$$

Empirische Varianz:

$$\begin{aligned}s_x^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \frac{1}{n-1} \left( \sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)\end{aligned}$$

$$\begin{aligned}s_y^2 &= \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \\ &= \frac{1}{n-1} \left( \sum_{i=1}^n y_i^2 - n\bar{y}^2 \right)\end{aligned}$$

Empirische Streuung:

$$s_x = \sqrt{s_x^2}$$

$$s_y = \sqrt{s_y^2}$$

Empirische Kovarianz

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \left( \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \right)$$

Empirischer Korrelationskoeffizient

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

---

### 7.2.4 Regressionsgerade

---

$$y = ax + b$$

$$a = \frac{s_{xy}}{s_x^2}$$

$$b = \bar{y} - a\bar{x}$$

Residuen

$$r_i = y_i - ax_i - b, i = 1, \dots, n$$

Residuenquadrat

$$\sum_{i=1}^n r_i^2 = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- Für  $r_{xy} > 0$  streng monoton fallende Gerade
- Für  $r_{xy} < 0$  streng monoton steigende Gerade
- Für  $r_{xy} = 0$  horizontale Gerade

---

## 7.3 Zufallsexperimente und Wahrscheinlichkeit

---

### 7.3.1 Zufallsexperimente

---

**Definition 7.3.1**  $\Omega$  heißt Ergebnismenge, seine Elemente  $\omega$  Ergebnisse. Teilmengen  $A \subseteq \Omega$  heißen Ereignisse. Ein Ereignis  $A \subseteq \Omega$  tritt ein, falls ein Ergebnis  $\omega \in A$  beobachtet wird.

**Definition 7.3.3** • Das aus zwei Ereignissen  $A$  und  $B$  zusammengesetzte Ereignis  $A \cup B$  tritt ein, falls ein Ergebnis  $\omega : \omega \in A \vee \omega \in B$  beobachtet wird

- Das Ereignis  $A \cap B$  tritt ein, falls ein Ergebnis  $\omega : \omega \in A \wedge \omega \in B$  beobachtet wird ( $A$  und  $B$  wird gleichzeitig beobachtet)
- $A^c = \Omega \setminus A$  heißt zu  $A$  komplementäres Ereignis
- Zwei Ereignisse  $A$  und  $B$  heißen unvereinbar, falls  $A \cap B = \emptyset$
- $\emptyset$  heißt unmögliches Ereignis und  $\Omega$  sicheres Ereignis
- die einelementigen Mengen  $\{\omega\}$  von  $\Omega$  heißen Elementarereignisse

---

### 7.3.2 Wahrscheinlichkeit

---

**Definition 7.3.4** Ein System  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$  von Ereignissen heißt  $\sigma$ -Algebra (oder Ereignissystem) wenn gilt:

- a)  $\Omega \in \mathcal{A}$
- b) Falls  $A \in \mathcal{A}$ , dann auch  $A^c \in \mathcal{A}$
- c) mit jeder Folge  $A_1, A_2, \dots \in \mathcal{A}$  gilt auch  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

**Definition 7.3.6** Eine Abbildung  $P : \mathcal{A} \rightarrow \mathbb{R}$  heißt Wahrscheinlichkeitsmaß, wenn sie Folgenden Axiomen von Komogorov genügt:

- a)  $P(A) \geq 0$  für  $A \in \mathcal{A}$
- b)  $P(\Omega) = 1$
- c)  $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_i P(A_i)$  für paarweise unvereinbare  $A_1, A_2, \dots \in \mathcal{A}$

Für beliebige Ereignisse  $A \subseteq \Omega$  mit Elementzahl  $\# A$  gilt

$$P(A) = \frac{\text{Elementzahl von } A}{n} = \frac{\# A}{\# \Omega}$$

---

### 7.3.3 Elementare Formeln der Kombinatorik

---

**Geordnete Probe mit Wiederholungen**

$$n^k$$

Bsp: k mal würfeln

**Geordnete Probe ohne Wiederholungen**

$$n(n-1)(n-2)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

Bsp: k autos parken auf n Parkplätzen

Im Falle  $k = n$  **Permutation der Menge  $\Omega$**

$$n!$$

**Ungeordnete Probe ohne Wiederholungen**

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Bsp: Lotto

$k_1$  aus  $N_1, k_2$  aus  $N_2$  in  $k$  aus  $N$

$$\frac{\binom{N_1}{k_1} \binom{N_2}{k_2}}{\binom{N}{k}}$$

mit  $N = N_1 + N_2$  und  $k = k_1 + k_2$

---

## 7.4 Bedingte Wahrscheinlichkeit, Unabhängigkeiten

---

### 7.4.1 Bedingte Wahrscheinlichkeit

---

**Bedingte Wahrscheinlichkeit von A unter der Bedingung B**

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

**Regel von der vollständigen Wahrscheinlichkeit**

$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$$

**Formel von Bayes**

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^n P(A_k)P(B|A_k)} = \frac{P(A_i)P(B|A_i)}{P(B)}$$

**Multiplikationsformel**

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap \dots \cap A_{n-1})$$

---

## 7.4.2 Unabhängigkeit

---

**Definition 7.4.3** Zwei Ereignisse  $A$  und  $B$  heißen unabhängig, falls

$$P(A \cap B) = P(A)P(B)$$

Ereignisse  $A_1, \dots, A_n$  heißen vollständig unabhängig, falls für alle  $\{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$  gilt

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$$

---

## 7.5 Zufallsvariablen und Verteilungsfunktion

---

**Definition 7.5.1** Eine Zufallsvariable ist eine Abbildung

$$X : \Omega \rightarrow \mathbb{R}$$

mit der Eigenschaft, dass für jedes Intervall  $I \subseteq \mathbb{R}$  die Urbildmenge

$$A = \{\omega \in \Omega : X(\omega) \in I\}$$

zum Ereignissystem  $\mathcal{A}$  gehört. Die Wahrscheinlichkeit dieses Ereignisses "X nimmt Werte im Intervall I an" bezeichnet man abkürzend mit  $P(X \in I)$  und schreibt

$$P(A \leq X \leq B), P(X \leq x), \quad P(|X - a| < b), P(X = b), \quad usw$$

**Definition 7.5.3** Sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable. Die Abbildung  $F : \mathbb{R} \rightarrow \mathbb{R}$

$$F(x) = P(X \leq x), x \in \mathbb{R}$$

heißt Verteilungsfunktion der Zufallsvariable  $X$

Verteilungsfunktionen sind monoton wachsende Funktionen mit

$$F(-\infty) = 0 \quad F(\infty) = 1 \quad F(x+) = F(x) \quad \forall x \in \mathbb{R}$$

Zudem:

$$\begin{aligned} P(X = a) &= F(a) - F(a-) \\ P(a < X \leq b) &= F(b) - F(a) \\ P(a \leq X < b) &= F(b-) - F(a-) \\ P(a \leq X \leq b) &= F(b) - F(a-) \\ P(x > a) &= 1 - F(a) \end{aligned}$$

Eine Zufallsvariable heißt stetig verteilt mit der Dichte  $f$  wenn ihre Verteilungsfunktion  $F$  durch

$$F(x) = \int_{-\infty}^x f(t)dt, x \in \mathbb{R}$$

gegeben ist.

---

### 7.5.1 Beispiele für diskrete Verteilungen

---

#### Geometrische Verteilung

---

Es sei  $0 < p < 1$  eine Zufallsvariable  $X$  mit dem Wertebereich  $\mathbb{N}$  heißt geometrisch verteilt mit dem Parameter  $p$  falls

$$P(X = i) = (1 - p)^{i-1}p, i = 1, 2, \dots$$

**Anwendung:**

Wird ein Zufallsexperiment, bei dem ein bestimmtes Ereignis mit Wahrscheinlichkeit  $p$  eintritt, so lange unabhängig wiederholt, bis zum ersten Mal dieses Ereignis eintritt, dann kann die Anzahl der dazu benötigten Versuche durch eine geometrisch verteilte Zufallsvariable modelliert werden („Warten auf den ersten Erfolg“).

---

## Binominalverteilung

---

Sei  $n \in \mathbb{N}$  und  $0 < p < 1$ . Eine Zufallsvariable  $X$  mit dem Wertebereich  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  heißt binominalverteilt mit Parametern  $n$  und  $p$ , kurz  $B(n, p)$ -verteilt falls

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, i = 0, 1, 2, \dots, n$$

### Anwendung:

Wird ein Zufallsexperiment, bei dem ein bestimmtes Ereignis mit Wahrscheinlichkeit  $p$  eintritt,  $n$ -mal unabhängig wiederholt, und dabei gezählt, wie oft dieses Ereignis eintritt, so kann diese zufällige Anzahl als  $B(n, p)$ -verteilte Zufallsvariable  $X$  beschrieben werden („Anzahl der Erfolge bei  $n$  Versuchen“).

---

## Poissonverteilung

---

Sei  $\lambda > 0$ . Eine Zufallsvariable  $X$  mit dem Wertebereich  $\mathbb{N}_0$  heißt Poisson-verteilt falls gilt

$$P(X = i) = \frac{\lambda^i}{i!} e^{-\lambda}, i = 0, 1, 2, \dots$$

Sie eignet sich zur Modellierung von Zählergebnissen folgenden Typs: In einer Telefonzentrale wird die Anzahl der innerhalb von 10 Minuten eingehenden Anrufe gezählt.  $\lambda$  gibt die „mittlere Anzahl“ der eingehenden Anrufe an.

$$E(X) = \lambda$$

---

## 7.5.2 Beispiele für Stetige Verteilungen

---

---

### Rechteckverteilung

---

Es sei  $a < b$ . Eine Stetig verteilte Zufallsvariable mit der Dichte

$$f(t) = \begin{cases} \frac{1}{b-a}, & a \leq t \leq b \\ 0 & \text{sonst} \end{cases}$$

heißt rechteckverteilt im Intervall  $[a, b]$ , kurz  $R(a, b)$ -verteilt. Für die Verteilungsfunktion ergibt sich

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & x \geq b \end{cases}$$

---

### Exponentialverteilung

---

Sei  $\lambda > 0$ . Eine Stetig verteilte Zufallsvariable  $X$  mit der Dichte

$$f(t) = \begin{cases} 0, & t < 0 \\ \lambda e^{-\lambda t} & t \geq 0 \end{cases}$$

heißt exponentialverteilt mit Parameter  $\lambda$ , kurz  $Ex(\lambda)$ . Für die Verteilungsfunktion ergibt sich

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

$$E(X) = \frac{1}{\lambda}$$

---

## Normalverteilung

---

Es seien  $\mu \in \mathbb{R}$  (Erwartungswert) und  $\sigma > 0$  (Standardabweichung). ( $\sigma^2 = \text{Varianz}$ )

Eine stetig verteilte Zufallsvariable mit der Dichte

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}, t \in \mathbb{R}$$

heißt normalverteilt mit Parameter  $\mu$  und  $\sigma^2$  kurz:  $N(\mu, \sigma^2)$ . Im Falle  $\mu = 0, \sigma^2 = 1$  spricht man von einer Standard-Normalverteilung mit der Verteilungsfunktion

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt, x \geq 0 \qquad \Phi(x) = P(X \leq x)$$

(Kann aus Tabelle Entnommen werden)

Es gilt:

$$\Phi(0) = \frac{1}{2}, \qquad \Phi(-x) = 1 - \Phi(x) \qquad x \geq 0$$

Allgemein ist die Verteilungsfunktion

$$F_{\mu, \sigma^2}(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

---

## 7.6 Erwartungswert und Varianz

---

Ist  $X$  eine diskret verteilte Zufallsvariable mit Werten  $x_1, x_2, \dots$  so heißt

$$E(X) = \sum_i x_i P(X = x_i)$$

Erwartungswert von  $X$  falls  $\sum_i |x_i| P(X = x_i) < \infty$ . Entspricht arithmetischem Mittel, wenn Wahrscheinlichkeit für alle  $x_i$  gleich ist  
Ist  $X$  eine stetig verteilte Zufallsvariable mit Dichte  $f$  so heißt

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Erwartungswert von  $X$ , falls  $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$

Ist  $h : \mathbb{R} \rightarrow \mathbb{R}$  eine stückweise stetige Funktion dann hat die Zufallsvariable  $h(X)$  für eine diskret verteilte Zufallsvariable  $X$  den Erwartungswert

$$E(h(X)) = E(X) = \sum_i h(x_i) P(X = x_i)$$

Ist  $X$  stetig verteilt mit Dichte  $f$ , dann hat  $h(X)$  den Erwartungswert

$$E(h(X)) = \int_{-\infty}^{\infty} h(x) f(x) dx$$

Der Erwartungswert der quadratischen Abweichung der Zufallsvariablen  $X$  zu  $E(X)$  heißt Varianz von  $X$

$$\text{Var}(X) = E([X - E(X)]^2) = E(X^2) - (E(X))^2$$

Die Standardabweichung ist definiert durch  $\sqrt{\text{Var}(X)}$

---

### 7.6.1 Rechenregeln

---

Es gilt

$$\begin{aligned} E(aX + b) &= aE(X) + b \\ E(h_1(X) + h_2(X)) &= E(h_1(X)) + E(h_2(X)) \end{aligned}$$

für Eine Zufallsvariable  $X, a, b \in \mathbb{R}$  und  $h_1, h_2$  stückweise Stetige Funktionen

Allgemeiner gilt für Zufallsvariablen  $X_1, \dots, X_n$  und  $a_1, \dots, a_n, b \in \mathbb{R}$ :

$$E(a_1 X_1 + \dots + a_n X_n + b) = a_1 E(X_1) + \dots + a_n E(X_n) + b$$

Der Erwartungswert ist also linear

Weiterhin gilt

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Verteilung	$E(X)$	$Var(X)$
$N(\mu, \sigma^2)$	$\mu$	$\sigma^2$
$Ex(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$B(n, p)$	$np$	$np(1-p)$
Geom	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	$\lambda$	$\lambda$
Rechteck	$\frac{a+b}{2}$	$\frac{1}{12}(b-a)$

### Tschebyschevsche Ungleichung

Es gilt

$$P(|X - E(X)| \geq c) \leq \frac{Var(X)}{c^2}, c > 0$$

**Definition 7.6.3 (Unabhängigkeit)** Seien  $X_1, X_2, \dots, X_n$  Zufallsvariablen mit Verteilungsfunktionen  $F_1, F_2, \dots, F_n$  die gemeinsame Verteilungsfunktion von  $X_1, X_2, \dots, X_n$  ist definiert durch

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n), (x_1, \dots, x_n) \in \mathbb{R}^n$$

Die Zufallsvariablen heißen unabhängig, wenn für alle  $(x_1, \dots, x_n) \in \mathbb{R}^n$  die Ereignisse

$$\{X_1 \leq x_1\}, \dots, \{X_n \leq x_n\}$$

vollständig unabhängig sind, also

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = P(X_1 \leq x_1) \dots P(X_n \leq x_n)$$

oder kurz

$$F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n)$$

**Satz 7.6.4** Die Zufallsvariablen  $X_1, X_2, \dots, X_n$  seien unabhängig. Es gilt

$$Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n)$$

## 7.7 Gesetz der großen Zahlen, zentraler Grenzwertsatz

### 7.7.1 Das schwache gesetz der großen Zahlen

**Satz 7.7.1** Ist  $X_1, X_2, \dots$  eine Folge unabhängiger identisch verteilter Zufallsvariablen mit  $E(X_i) = \mu$ ,  $Var(X_i) = \sigma^2$  dann gilt

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \epsilon\right) = 0, \forall \epsilon > 0$$

### 7.7.2 Zentraler Grenzwertsatz

**Satz 7.7.2** Ist  $X_1, X_2, \dots$  eine Folge unabhängiger Zufallsvariablen mit

$$E(X_i) = \mu_i \quad Var(X_i) = \sigma_i^2 \quad i = 1, 2, \dots$$

so gilt unter schwachen zusätzlichen Voraussetzungen, z.B dass die  $X_i$  identisch verteilt sind:

$$\lim_{n \rightarrow \infty} P\left(\frac{X_1 + \dots + X_n - (\mu_1 + \dots + \mu_n)}{\sqrt{\sigma_1^2 + \dots + \sigma_n^2}} \leq y\right) = \Phi(y) \quad \forall y \in \mathbb{R}$$

Ein Arithmetisches Mittel  $\bar{X}_{(n)}$  ist für großes  $n$  also Näherungsweise  $N(\mu, \sigma^2)$  verteilt, wobei

$$\mu = \frac{1}{n} E(X_1 + \dots + X_n) = \frac{1}{n} (\mu_1 + \dots + \mu_n) \quad \sigma^2 = \frac{1}{n^2} Var(X_1 + \dots + X_n) = \frac{1}{n^2} (\sigma_1^2 + \dots + \sigma_n^2)$$



---

## Wahrscheinlichkeitstheoretisches Modell für Messreihen

---

Empirische Verteilungsfunktion

$$F_n(z; x_1, x_2, \dots, x_n) = \frac{\text{Zahl der } x_i \text{ mit } x_i \leq z}{n}$$

**Satz 7.7.4** (Zentralsatz der Statistik) Sei  $X_1, X_2, \dots$  eine Folge unabhängiger identisch verteilter Zufallsvariablen mit der Verteilungsfunktion  $F$  und bezeichne

$$D_n(X_1, \dots, X_n) = \sup_{z \in \mathbb{R}} |F_n(z; X_1, \dots, X_n) - F(z)|$$

die Zufällige Maximalabweichung zwischen empirischer und

wahrer

Verteilungsfunktion, Dann gilt

$$P(\lim_{n \rightarrow \infty} D_n(X_1, \dots, X_n) = 0) = 1$$

$D_n(X_1, \dots, X_n)$  konvergiert also mit Wahrscheinlichkeit 1 gegen 0

---

## 7.8 Testverteilungen und Quantilapproximationen

---

Seien  $Z_1, \dots, Z_n$  unabhängige identisch  $N(0,1)$ -Verteilte Zufallsgrößen.

**$\chi_r^2$ -Verteilung:**

Es sei  $r \in \{1, \dots, n\}$ . Eine Zufallsvariable  $X$  heißt  $\chi_r^2$ -verteilt, falls sie die Verteilungsfunktion

$$F(x) = P(Z_1^2 + \dots + Z_r^2 \leq x), x \in \mathbb{R}$$

besitzt

**$t_r$ -Verteilung:**

es sei  $r \in \{1, \dots, n-1\}$ . Eine Zufallsvariable  $X$  heißt  $t_r$ -verteilt, falls sie die Verteilungsfunktion

$$F(x) = P\left(\frac{Z_{r+1}}{\sqrt{(Z_1^2 + \dots + Z_r^2)/r}} \leq x\right), x \in \mathbb{R}$$

besitzt  **$F_{r,s}$ -Verteilung**

Es sei  $r, s \in \{1, \dots, n-1\}$  mit  $r+s \leq n$ . Eine Zufallsvariable  $X$  heißt  $F$ -verteilt mit  $r$  und  $s$  Freiheitsgraden, falls sie die Verteilungsfunktion

$$F(x) = P\left(\frac{(Z_1^2 + \dots + Z_r^2)/r}{(Z_{r+1}^2 + \dots + Z_{r+s}^2)/s} \leq x\right), x \in \mathbb{R}$$

besitzt

Die Dichte dieser Verteilungen können unter Verwendung der Gamma-Funktion

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt, x > 0$$

und der Beta-Funktion

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt, \alpha, \beta > 0$$

(Don't ask me how)

**Bezeichnungen für Quantile**

Allgemein  $F(x_p) = p$ .

$u_p$	p-Quantil der $N(0,1)$ -Verteilung
$t_{r;p}$	p-Quantil der $t_r$ -Verteilung
$\chi_{r;p}^2$	p-Quantil der $\chi_r^2$ -Verteilung
$F_{r,s;p}$	p-Quantil der $F_{r,s}$ -Verteilung

---

### 7.8.1 Wichtige Anwendungsbeispiele

---

**Satz 7.8.1** Es seien  $X_1, \dots, X_n$  unabhängige, identisch  $N(\mu, \sigma^2)$ -verteilte Zufallsvariable. Dann gilt

- $\bar{X}_{(n)}$  ist  $N(\mu, \sigma^2/n)$ -verteilt
- $\frac{n-1}{\sigma^2} S_{(n)}^2$  ist  $\chi_{n-1}^2$ -verteilt
- $\bar{X}_{(n)}$  und  $S_{(n)}^2$  sind unabhängig
- $\sqrt{n} \frac{\bar{X}_{(n)} - \mu}{\sqrt{S_{(n)}^2}}$  ist  $t_{n-1}$ -verteilt

---

## 8 Schätzverfahren und Konfidenzintervalle

---

### 8.1 Grundlagen zu Schätzverfahren

---

**Definition 8.1.2** Ein Schätzverfahren oder eine Schätzfunktion oder kurz Schätzer ist eine Abbildung

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

Sie ordnet einer Messreihe  $x_1, \dots, x_n$  einen Schätzwert  $T_n(x_1, \dots, x_n)$  (Zufallsvariable) für den unbekannten Wert  $\tau(\theta)$  zu. Die Zufallsvariable  $T_n(X_1, \dots, X_n)$  heißt Schätzvariable.

**Definition 8.1.3** ein Schätzer  $T_n : \mathbb{R}^n \rightarrow \mathbb{R}$  heißt erwartungstreu für  $\tau : \Theta \rightarrow \mathbb{R}$ , falls gilt

$$E_\theta(T_n(X_1, \dots, X_n)) = \tau(\theta) \text{ für alle } \theta \in \Theta$$

**Definition 8.1.4** •  $\tau$  sei gegeben durch  $\tau(\theta) = E_\theta(X) = \mu$ . Das arithmetische Mittel  $\bar{X}_{(n)}$  ist ein erwartungstreuer Schätzer für  $\mu$

- $\tau$  sei gegeben durch  $\tau(\theta) = \text{Var}_\theta(X) = \sigma^2$ . Die Stichprobenvarianz  $S_{(n)}^2$  ist ein erwartungstreuer Schätzer für  $\sigma^2$

**Definition 8.1.5** Eine Folge von Schätzern  $T_n : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $n = 1, 2, \dots$  heißt asymptotisch erwartungstreu für  $\tau$  falls gilt

$$\lim_{n \rightarrow \infty} E_\theta(T_n(X_1, \dots, X_n)) = \tau(\theta) \text{ für alle } \theta \in \Theta$$

**Mittlerer quadratischer Fehler (mean squared error):**

$$MSE_\theta(T) := E_\theta((T - \tau(\theta))^2)$$

Es gilt

$$T \text{ erwartungstreu} \Rightarrow MSE_\theta(T) = \text{Var}_\theta(T)$$

$T_1$  ist effizienter als  $T_2$  wenn gilt

$$MSE_\theta(T_1) \leq MSE_\theta(T_2)$$

**Definition 8.1.6** Eine Folge von Schätzern  $T_1, T_2, \dots$  heißt konsistent für  $\tau$  wenn für alle  $\varepsilon > 0$  und alle  $\theta \in \Theta$  gilt

$$\lim_{n \rightarrow \infty} P_\theta(|T_n(X_1, \dots, X_n) - \tau(\theta)| > \varepsilon) = 0$$

Sie heißt konsistent im quadratischen Mittel für  $\tau$  wenn für alle  $\theta \in \Theta$  gilt

$$\lim_{n \rightarrow \infty} MSE_\theta(T_n) = 0$$

**Satz 8.1.7** Ist  $T_1, T_2, \dots$  eine Folge von Schätzern, die erwartungstreu für  $\tau$  sind und gilt

$$\lim_{n \rightarrow \infty} \text{Var}_\theta(T_n(X_1, \dots, X_n)) = 0 \text{ für alle } \theta \in \Theta$$

dann ist die Folge von Schätzern konsistent für  $\tau$

**Satz 8.1.8** Ist  $T_1, T_2, \dots$  eine Folge von Schätzern, die konsistent im quadratischen Mittel für  $\tau$  ist, dann ist die Folge von Schätzern konsistent für  $\tau$

---

## 8.2 Maximum-Likelihood-Schätzer

---

Bei gegebener Verteilungsklasse  $F_\theta, \theta \in \Theta$ , lassen sich Schätzer für den Parameter  $\theta$  oft mit der Maximum-Likelihood-Methode gewinnen. Sind die zugrundeliegenden Zufallsvariablen  $X_1, \dots, X_n$  stetig mit einer **Dichte** verteilt, so hängt diese ebenfalls von den Parametern ab:

$$f_\theta(x), x \in \mathbb{R}.$$

Wir definieren hier  $\mathbb{X} = \mathbb{R}$ . Im Fall diskreter Zufallsvariablen  $X$ , bzw.  $X_1, \dots, X_n$  definieren wir

$$f_\theta(x) = P_\theta(X = x) \text{ für alle } x \text{ aus dem Wertebereich } \mathbb{X} \text{ von } X.$$

**Definition 8.2.1** Für eine Messreihe  $x_1, \dots, x_n$  heißt die Funktion  $L(\cdot; x_1, \dots, x_n)$  mit

$$L(\theta; x_1, \dots, x_n) = f_\theta(x_1)f_\theta(x_2) \dots f_\theta(x_n)$$

die zu  $x_1, \dots, x_n$  gehörige Likelihood-Funktion

Ein Parameterwert

$$\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$$

mit

$$L(\hat{\theta}; x_1, \dots, x_n) \geq L(\theta; x_1, \dots, x_n) \text{ für alle } \theta \in \Theta$$

heißt Maximum-Likelihood-Schätzwert für  $\Theta$ . Existiert dieser zu jeder möglichen Messreihe, dann heißt

$$T_n : \mathbb{X}^n \rightarrow \Theta, T_n(x_1, \dots, x_n) = \hat{\theta}(x_1, \dots, x_n)$$

Maximum-Likelihood-Schätzer.

---

### 8.2.1 Kochrezept für Berechnung

---

$f_\theta(x)$  entspricht der Dichtefunktion. Bei diskreten Werten siehe oben.

$$L(\theta; x_1, \dots, x_n) = f_\theta(x_1)f_\theta(x_2) \dots f_\theta(x_n)$$

$$\ln(L(\theta; x_1, \dots, x_n)) = \sum_{i=1}^n \ln(f_\theta(x_i))$$

Teilweise kann man  $f_\theta$  durch den  $\ln$  nochmal in Summen aufteilen, da dann alle Summanden ohne  $\theta$  bei der Ableitung rausfliegen.

$$\frac{\partial}{\partial \theta} \ln(L(\theta; x_1, \dots, x_n)) = \frac{\partial}{\partial \theta} \sum_{i=1}^n \ln(f_\theta(x_i)) \stackrel{!}{=} 0$$

Nun nach  $\theta$  umstellen. Dies ist dann die eindeutige Nullstelle

$$\hat{\theta}(x_1, \dots, x_n) = \dots$$

Um zu wissen, dass es ein Maximum ist, muss man noch schauen ob die 2. Ableitung  $< 0$  ist

$$\frac{\partial^2}{\partial \theta^2} \ln(L(\theta; x_1, \dots, x_n)) \stackrel{!}{<} 0$$

---

### 8.3 Konfidenzintervalle

---

**Definition 8.3.1** Sei  $0 < \alpha < 1$ . Das zufällige Intervall  $I(X_1, \dots, X_n)$  heißt Konfidenzintervall für  $\tau(\theta)$  zum Konfidenzniveau  $1 - \alpha$ , falls gilt

$$P_\theta(U(X_1, \dots, X_n) \leq \tau(\theta) \leq O(X_1, \dots, X_n)) \geq 1 - \alpha \text{ für alle } \theta \in \Theta$$

Das zu einer bestimmten Messreihe  $x_1, \dots, x_n$  gehörige Intervall

$$I(x_1, \dots, x_n) = [U(X_1, \dots, X_n), O(X_1, \dots, X_n)]$$

heißt konkretes Schätzintervall für  $\tau(\Theta)$

---

#### 8.3.1 Konstruktion von Konfidenzintervallen

---

Im Folgenden

$$F_\theta(x) = F_{(\mu, \sigma^2)}(x) = \Phi\left(\frac{x - \mu}{\sigma}\right) \quad \bar{X}_{(n)} = \frac{1}{n} \sum_{i=1}^n X_i \quad S_{(n)}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_{(n)})^2$$

Konfidenzintervalle zum Niveau  $1 - \alpha$ :

---

##### Konfidenzintervall für $\mu$ bei bekannter Varianz $\sigma^2 = \sigma_0^2$

---

Hier ist  $\Theta = \{(\mu, \sigma_0^2) : \mu \in \mathbb{R}\}$  und  $\tau(\theta) = \mu$ . Das Konfidenzintervall für  $\mu$  zum Niveau  $1 - \alpha$  lautet

$$I(X_1, \dots, X_n) = \left[ \bar{X}_{(n)} - u_{1-\alpha/2} \frac{\sigma_0}{\sqrt{n}}, \bar{X}_{(n)} + u_{1-\alpha/2} \frac{\sigma_0}{\sqrt{n}} \right]$$

---

##### Konfidenzintervall für $\mu$ bei unbekannter Varianz $\sigma^2$

---

Hier ist  $\Theta = \{(\mu, \sigma_0^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$  und  $\tau(\theta) = \mu$ . Das Konfidenzintervall für  $\mu$  lautet

$$I(X_1, \dots, X_n) = \left[ \bar{X}_{(n)} - t_{n-1; 1-\alpha/2} \sqrt{\frac{S_{(n)}^2}{n}}, \bar{X}_{(n)} + t_{n-1; 1-\alpha/2} \sqrt{\frac{S_{(n)}^2}{n}} \right]$$

---

##### Konfidenzintervall für $\sigma^2$ bei bekanntem Erwartungswert $\mu = \mu_0$

---

Hier ist  $\Theta = \{(\mu, \sigma_0^2) : \sigma^2 > 0\}$  und  $\tau(\theta) = \sigma^2$ . Das Konfidenzintervall für  $\sigma^2$  lautet

$$I(X_1, \dots, X_n) = \left[ \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\chi_{n-1; 1-\alpha/2}^2}, \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\chi_{n-1; \alpha/2}^2} \right]$$

---

##### Konfidenzintervall für $\sigma^2$ bei unbekanntem Erwartungswert $\mu$

---

Hier ist  $\Theta = \{(\mu, \sigma_0^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$  und  $\tau(\theta) = \sigma^2$ . Das Konfidenzintervall für  $\sigma^2$  lautet

$$I(X_1, \dots, X_n) = \left[ \frac{(n-1)S_{(n)}^2}{\chi_{n-1; 1-\alpha/2}^2}, \frac{(n-1)S_{(n)}^2}{\chi_{n-1; \alpha/2}^2} \right]$$

---

## 9 Tests bei Normalverteilungsannahmen

---

### 9.1 Grundlagen

---

#### Konstruktionsprinzip für Test zum Niveau $\alpha$

---

Tests lassen sich nach dem folgenden allgemeinen Prinzip konstruieren

1. Verteilungsannahme formulieren
2. Nullhypothese  $H_0$  formulieren
3. Testgröße  $T$  wählen und ihre Verteilung unter  $H_0$  bestimmen
4.  $I \subseteq \mathbb{R}$  so wählen, dass unter  $H_0$  gilt  $P(T(X_1, \dots, X_n) \in I) \leq \alpha$

---

### 9.2 Wichtige Tests bei Normalverteilungsannahme

---

#### Gauß-Test

---

1.  $X_1, \dots, X_n$  unabhängig, identisch  $N(\mu, \sigma_0^2)$ -verteilt  $\sigma_0^2$  bekannt
2. a)  $H_0 : \mu = \mu_0$ , b)  $H_0 : \mu \leq \mu_0$ , c)  $H_0 : \mu \geq \mu_0$
3. Testgröße

$$T(X_1, \dots, X_n) = \frac{\sqrt{n}}{\sigma_0} (\bar{X}_{(n)} - \mu_0)$$

4. Ablehnung Falls
  - a)  $|T| > u_{1-\alpha/2}$ , b)  $T > u_{1-\alpha}$ , c)  $T < u_\alpha$

---

#### t-Test

---

1.  $X_1, \dots, X_n$  unabhängig, identisch  $N(\mu, \sigma^2)$ -verteilt  $\sigma^2$  unbekannt
2. a)  $H_0 : \mu = \mu_0$ , b)  $H_0 : \mu \leq \mu_0$ , c)  $H_0 : \mu \geq \mu_0$
3. Testgröße

$$T(X_1, \dots, X_n) = \sqrt{n} \frac{\bar{X}_{(n)} - \mu_0}{\sqrt{S_{(n)}^2}}$$

4. Ablehnung Falls
  - a)  $|T| > t_{n-1; 1-\alpha/2}$ , b)  $T > t_{n-1; 1-\alpha}$ , c)  $T < t_{n-1; \alpha}$

---

#### $\chi^2$ -Streuungstest

---

1.  $X_1, \dots, X_n$  unabhängig, identisch  $N(\mu, \sigma^2)$ -verteilt  $\mu, \sigma^2$  unbekannt
2. a)  $H_0 : \sigma^2 = \sigma_0^2$ , b)  $H_0 : \sigma^2 \leq \sigma_0^2$ , c)  $H_0 : \sigma^2 \geq \sigma_0^2$
3. Testgröße

$$T(X_1, \dots, X_n) = \frac{(n-1)}{\sigma_0^2} S_{(n)}^2$$

4. Ablehnung falls
  - a)  $T < \chi_{n-1; \alpha/2}^2$  oder  $T > \chi_{n-1; 1-\alpha/2}^2$ , b)  $T > \chi_{n-1; 1-\alpha}^2$ , c)  $T < \chi_{n-1; \alpha}^2$

---

### 9.3 Verteilungstests

---

Könnt ich auch nicht kürzer zusammenfassen, einfach im Skript schauen