in G bleiben

Gruppe: Gegeben sei eine Menge G. Gegeben sei ferner eine Verknüpfung $*$ auf G, d.h. eine Abbildung von $G \times G$ in G.

↳ Die Menge G zusammen mit der Verknüpfung $= (G, *)$, heisst Gruppe wenn folgende Axiome erfüllt sind

1. Assoziativgesetz: $(a * b) * c = a * (b * c)$ $\boxed{\forall a, b, c \in G}$

2. Es gibt ein neutrales Element $e \in G$ bezüglich der Verknüpfung, d.h., dass es gelten soll:

$\boxed{\text{für jedes } a \text{ soll ein } a^{-1} \text{ existieren}}$

2.1 $e * a = a$

2.2 $a * a^{-1} = e$ wobei a^{-1} = inverses Element von a

Kommutative Gruppe, wenn es ausserdem gilt.

3. Kommutativgesetz: $a * b = b * a$ $\boxed{\forall a, b \in G}$

Vektorraum: Eine Menge V, die mit einer Addition $\dot{+}$: $\boxed{V \times V \to V}$ und einer Multiplikation mit Skalaren $\otimes$: $\boxed{\mathbb{R} \times V \to V}$ versehen ist, heisst reeller Vektorraum falls folgende Bedingungen erfüllt sind:

1. $(V, \dot{+})$ ist eine kommutative Gruppe

2.
$$(\lambda + \mu) \otimes \underline{v} = \lambda \otimes \underline{v} \dot{+} \mu \otimes \underline{v}$$
$$\lambda \otimes (\underline{v} \dot{+} \underline{w}) = \lambda \otimes \underline{v} \dot{+} \lambda \otimes \underline{w}$$
$$\lambda \otimes (\mu \otimes \underline{v}) = (\lambda \mu) \otimes \underline{v}$$
$$1 \otimes \underline{v} = \underline{v}$$
$\forall \underline{v}, \underline{w} \in V \wedge \lambda, \mu \in \mathbb{R}$

↳ hier nie vergessen auch $\underline{v} \dot{+} \underline{w} \in V$ und $\lambda \otimes \underline{v} \in V$ zu prüfen!

Lineare Hülle (Linearkombination der Vektoren $\in$ Menge X):

1. $X \subset V \Rightarrow \text{lin } X \subset V$
2. $X \subset \text{lin } X$ $\quad \forall X \subset V$
3. $X \subset Y \Rightarrow \text{lin } X \subset \text{lin } Y$ $\quad \forall X, Y \subset V$
4. $\text{lin}(\text{lin } X) = \text{lin } X$ $\quad \forall X \subset V$

Unterraum: Seien V, W Vektorräume die durch die selbe $\dot{+}$ $\wedge \otimes$ definiert sind mit $W \subset V$, dann genügt damit W = Unterraum zu zeigen, dass:

1. $\underline{0} \in W$
2. $\underline{v} \dot{+} \underline{w} \in W$ $\quad \forall \underline{v}, \underline{w} \in W$
3. $\lambda \otimes \underline{v} \in W$ $\quad \forall \lambda \in \mathbb{R} \wedge \underline{v} \in W$

wobei A = {Menge von Vektoren von V}

<u>Erzeugendensystem:</u> Sei: $A \subset V$. A heisst Erzeugendensystem von V, wenn sich jedes Element von V als Linearkombination von Elementen in A darstellen lässt, d.h. wenn $V = \text{lin } A$.

↳ Es gilt ausserdem: $A \subset \text{lin } B \Rightarrow \text{lin } A \subset \text{lin } B$ so wenn $\text{lin } A = V$ gilt auch, dass $\text{lin } B = V$, wodurch auch B Erzeugendensystem von V ist.

! Seien V = Vektorraum $\wedge$ $A \subset V$. Dann ist A <u>genau dann</u> lin. unabhängig, wenn jedes $\underline{v} \in V$ auf <u>höchstens eine Weise</u> als lin. Kombination von A dargestellt werden kann.

<u>Basis:</u> Ist V = Vektorraum $\wedge$ B = lin. unabhängiges EZS von V

↳ Dann heisst B = Basis von V.

⇒ Basis wenn jedes $\underline{v} \in V$ sich auf <u>genau eine Weise</u> dargestellt werden kann.

<u>Komponente:</u> Ist $B = \{\underline{v}^i \mid i \in I = \{1, \dots m\}\}$ eine Basis von V und hat $\underline{v} \in V$:

$$\underline{v} = \sum_{i=1}^{m} \lambda_i \underline{v}^i$$

dann heisst λ_i die i-te Komponente (Koordinate) von $\underline{v}$ bezüglich der Basis B.

<u>Basisergänzungssatz:</u> Besitzt V ein endliches EZS (z.B. A) und ist $C \subset V$ lin. unabhängig, so existiert ein B mit $C \subset B \subset A$ die eine Basis von V ist. (d.h. man kann C zu einer endlichen Basis B von V ergänzen)

<u>Austauschsatz von Steinitz:</u> Sei B eine Basis von V und habe $\underline{w} \in V$, die Darstellung: $\underline{w} = \sum_{i=1}^{n} \lambda_i \underline{v}^i$ mit $\lambda_j \neq 0$ für $\underline{1}$ j, dann kann man $\underline{w}$ mit $\underline{v}^j$ für diese j umtauschen und B mit $\underline{w}$ anstatt $\underline{v}^j$ bleibt noch eine Basis.

Dies ist weil 1 Vektor sich genau auf 1 Weise darstellen lässt

↳ somit wenn $\underline{w}$ und $\underline{v}^j$ lin. abhängig dann $\underline{w}$ und $\underline{v}^i$ für alle $i \neq j$ lin. unabhängig da $\underline{v}^j$ und $\underline{v}^i$ lin. unabhängig.

! Es ist nun klar, dass Folgendes gilt:

„Sind $A \wedge B$ zwei Basen eines (endlichdimensionalen) Vektorraums V, so enthalten $A \wedge B$ gleich viele Elemente."

und „Hat ein V.raum eine Basis mit n Elemente und ist C eine lin. unabhängige Teilmenge von V mit n Elementen, so ist C eine Basis von V."

Matrix: $A = A^T \Rightarrow$ Matrix symmetrisch ; $A = -A^T \Rightarrow$ schiefsymmetrisch.

Rechengesetze für Matrizen:

1. $(A^T)^T = A$ 2. $(A+B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$ 4. $(AB)C = A(BC)$
5. $A(B+C) = AB + AC$
6. $(A+B)C = AC + BC$

7. $\underbrace{IA = AI}_{\text{Bei } I \text{ gilt Kommutativgesetz.}} = A$

8. $A^0 = I$ 9. $A^n = A^{n-1}A$
10. $A^{m+n} = A^m \cdot A^n$
11. $AA^{-1} = I = A^{-1}A$
12. $(A^{-1})^{-1} = A$ 13. $(AB)^{-1} = B^{-1}A^{-1}$
14. $(A^T)^{-1} = (A^{-1})^T$

Zeilen- und Spaltenrang: $\text{lin}\{A_{1\cdot}, A_{2\cdot}, \dots, A_{m\cdot}\}$ = Zeilenraum von A
$\text{lin}\{A_{\cdot 1}, A_{\cdot 2}, \dots, A_{\cdot n}\}$ = Spaltenraum von A

! Diese müssen gleich sein.

Dim Zeilenraum von A = Anzahl unabhängige Zeile = Zeilenrang von A
Dim Spaltenraum von A = Anzahl unabhängige Spalten = Spaltenrang von A

↳ es folgt rg A = Sp.rang A = Zeil.rang A mit $A(m \times n)$ muss also gelten:

$$\text{rg}\, A \leq m \quad \wedge \quad \text{rg}\, A \leq n \quad \text{und} \quad \boxed{\text{rg}\, A = \text{rg}\, A^T}$$

Lösbarkeit eines LGS: LGS mit $A\underline{x} = \underline{b}$
⇒ LGS genau lösbar wenn $\text{rg}[A, \underline{b}] = \text{rg}\, A$ oder wenn $\text{rg}\, A = n$
⇒ Ist nun LGS lösbar und $\underline{x}^0$ eine spezielle Lsg. von $A\underline{x} = \underline{b}$, so ist die Lösungsmenge $M = \underline{x}^0 + L$, wobei $L = \{\underline{v} \in \mathbb{R}^n \mid A\underline{v} = \underline{0}\}$

! L ist somit die Lösungsmenge des homogenen LGS und ist ein Unterraum von $\mathbb{R}^m$ und es gilt $\dim L = n - \text{rg}\, A$

! Es gilt $\text{rg}\, A = m$ dann und nur dann wenn $A\underline{x} = \underline{b}$ für jede rechte Seite $\underline{b} \in \mathbb{R}^m$ lösbar ist.

Lineare Abbildung Definition: Gegeben $V \wedge W$ = Vektorräume, φ = Abbildung von V in W.
↳ Die Abbildung heisst linear wenn für $\underline{v}^1, \underline{v}^2 \in V \wedge \lambda_1, \lambda_2 \in \mathbb{R}$ stets gilt:

$$\varphi(\lambda_1 \underline{v}^1 + \lambda_2 \underline{v}^2) = \lambda_1 \varphi(\underline{v}^1) + \lambda_2 \varphi(\underline{v}^2)$$

⇒ lin. Abbildung wird auch Homomorphismus genannt.

Elementare Eigenschaften linearer Abbildungen die gelten müssen (hier $\varphi: V \to W$)

1. $\varphi(\underline{0}_V) = \underline{0}_W$ 2. $\varphi(\underline{v}^1 - \underline{v}^2) = \varphi(\underline{v}^1) - \varphi(\underline{v}^2)$
3. $\varphi(\sum_{i=1}^m \lambda_i \underline{v}^i) = \sum_{i=1}^m \lambda_i \varphi(\underline{v}^i)$ 4. wenn $\underline{v}^i$ $i=1,\dots,m$ lin. abhängig, dann auch $\varphi(\underline{v}^i)$ lin. abhängig.
5. Wenn $\varphi(\underline{v}^i)$ $i=1,\dots,m$ lin. unabhängig, dann auch $\underline{v}^i$ lin. unabhängig.
6. $(\lambda\varphi)(\underline{v}) = \lambda\varphi(\underline{v})$ 7. $(\varphi + \psi)(\underline{v}) = \varphi(\underline{v}) + \psi(\underline{v})$, wobei $\varphi \wedge \psi$ = lin. Abb.
8. Für das Hintereinanderführen von zwei Abbildungen gilt: $(\varphi \circ \psi)(\underline{v}) = \varphi(\psi(\underline{v}))$

Definitionen: 1. $\varphi(V)$ = Bild der Abbildung φ 2. $\ker \varphi = \{\underline{v} \in W \mid \varphi(\underline{v}) = \underline{0}\}$
3. $\varphi(\underline{v}) = \underline{w}$ ← lineare Gleichung

Es gilt nun für $\varphi: V \to W$: $M\{\underline{v} \in V \mid \varphi(\underline{v}) = \underline{w}\}$:

1. M ist nichtleer genau dann, wenn $\underline{w} \in \varphi(V)$
2. Ist $\underline{v}^0$ irgendeine Lösung der Gleichung, so ist M als $M = \underline{v}^0 + \ker \varphi$ darstellbar.

! Sei $f: V \to W$. Dann ist: 1. $f(V)$ = Unterraum von W
2. $\ker f$ = Unterraum von V

<u>Weitere Wichtige Eigenschaften von Abbildungen:</u> (hier $f: V \to W$)

1. gilt für $\underline{v}^1 \neq \underline{v}^2$, $f(\underline{v}^1) \neq f(\underline{v}^2)$, so heisst f injektiv.
2. ist f injektiv und linear so heisst f regulär
3. Gilt $W = f(V)$ so heisst f surjektiv.
4. Ist f injektiv $\wedge$ surjektiv so heisst f bijektiv.
   ↳ Ist f bijektiv und linear so heisst f Isomorphismus zwischen V und W.
5. Sei $f: V \to W$ eine injektive lin. Abbildung mit der Inversen f^{-1}. Dann ist $f^{-1}: f(V) \subset W \to V$ wiederum eine injektive Abbildung

<u>Wenn $f(V)$ endlich dimensional.</u> ($f: V \to W$)

Ist $\{\underline{v}^1, \dots, \underline{v}^k\}$ Basis von $\ker f$ sowie $\{\underline{w}^1, \dots, \underline{w}^r\}$ = Basis von W gegeben und sind $\underline{v}^1, \dots, \underline{v}^r \in V$ mit:

$$f(\underline{v}^i) = \underline{w}^i \quad i = 1, \dots, r$$

dann gilt:
1. $k = n - r$ (Dimension V)
2. $B = \{\underline{v}^1, \dots, \underline{v}^k, \underline{v}^1, \dots, \underline{v}^r\}$ = Basis von V
3. $\dim V = \dim \ker f + \dim f(V) \Rightarrow \dim f(V) \leq \dim V$
4. f ist injektiv genau dann, wenn $\dim f(V) = \dim V$ oder anders gesagt wenn $\dim \ker = 0$ d.h. $\ker f \{\underline{0}\}$
5. Ist $f: V \to W$ bijektiv, so ist W auch endlich dimensional und es gilt $\dim V = \dim W$

<u>Daraus ergibt sich für Matrizen:</u>

1. $\ker f$ = Lösungsmenge des LGS $A\underline{x} = \underline{0}$
2. Abbildung injektiv nur dann, wenn LGS nur die triviale Lösung $\{\underline{0}\}$ hat.
3. $f(V)$ = Spaltenraum von A
4. Abbildung genau dann injektiv wenn $\dim f(V) = n$
   ⇒ also kann die Abbildung, die Matrix, nur injektiv sein wenn $m \geq n$ ist
5. Vereinigt man eine Basis des Lösungsmenge des LGS $A\underline{x} = \underline{0}$ (d.h. der Kern von f) mit je einer speziellen Lösung der inhomogene LGS $A\underline{v} = \underline{w}^i$ bezüglich einer Basis des Spaltenraums von A (= Bild von f) $\{\underline{w}^1, \dots, \underline{w}^r\}$ dann erhält man eine Basis $\{\underline{v}^1, \dots, \underline{v}^k, \underline{v}^1, \dots, \underline{v}^r\}$ von V

! Somit gilt die bekannte Formel: $\dim L = n - \operatorname{rg} A$

<u>Prinzip der lineare Fortsetzung:</u> Hat V eine endliche Basis $\{\underline{v}^1, \dots, \underline{v}^n\}$ und ist $f: V \to W$ linear, so ist die Abbildung f durch die Bilder der Basiselemente $f(\underline{v}^i)$ $i = 1, \dots$ bereits vollständig bestimmt. → siehe auch satz Bild von lin. abhängige Vektoren ist lin. abhängig

Darüber hinaus gilt: Seien V und W Vektorräume und zumindest V endlichdimensional. Ist $\{\underline{v}^1, \dots, \underline{v}^n\}$ eine Basis von V und $\{\underline{w}^1, \dots, \underline{w}^n\} \subset W$, dann gibt es <u>genau eine Abbildung</u> $f: V \to W$ mit:

$$f(\underline{v}^i) = \underline{w}^i \quad \text{für alle} \quad i \in \{1, \dots, n\}$$

<u>Lineare Abbildung zwischen endlichdimensionalen Abbildungen:</u>

Sei $\{\underline{v}^1, \dots, \underline{v}^n\}$ eine Basis von V und $\{\underline{w}^1, \dots, \underline{w}^m\}$ eine Basis von W.
↳ Sei $f: V \to W$ linear. Die Bilder der Basisvektoren sind nun gegeben als:

$$f(\underline{v}^j) = \sum_{i=1}^{m} a_{ij}\, \underline{w}^i \quad , \quad j = 1, \dots, n$$

! Nach dem Prinzip der linearen Fortsetzung ist durch die Bilder der Basisvektoren in V die lineare Abbildung f schon vollständig bestimmt.

↳ Mit anderen Worten ist also die Abbildung f durch die Matrix:

$$A = \begin{array}{c} \begin{array}{cccc} f(\underline{v}^1) & f(\underline{v}^2) & \cdots & f(\underline{v}^n) \end{array} & \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{c} \leftarrow \underline{w}^1 \\ \leftarrow \underline{w}^2 \\ \vdots \\ \underline{w}^m \end{array} \end{array}$$

bereits vollständig bestimmt. ↳ die j-te Spalte von A ist also der Komponentenvektor / Koeffizientenvektor von $f(\underline{v}^j)$ bezüglich der gegebenen Basis W.

! Sei nun $\underline{v} \in V$ beliebig und habe die Komponenten x_j bezüglich der Basisvektoren $\underline{v}^j \; \forall j$, also:

$$\underline{v} = \sum_{j=1}^{n} x_j \underline{v}^j$$

Dann folgt, da f linear:

$$f(\underline{v}) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij}\, \underline{w}^i \right) = \sum_{i=1}^{m} y_i\, \underline{w}^i$$

$$\text{wobei} \quad y_i = \sum_{j=1}^{n} a_{ij}\, x_j \qquad i = 1, \dots, m$$

sehr wichtig

! $\boxed{\text{Fazit:}}$ Bei gegebenen Basen in V und W ist jeder linearen Abbildung $f: V \to W$ eine Matrix A zugeordnet, deren j-te Spalte der Komponentenvektor von $f(\underline{v}^j)$ ist und die eine lineare Abbildung vom Raum $\mathbb{R}^n$ in den Raum $\mathbb{R}^m$ vermittelt.

$A^W \times B = BA$

Wichtig: Die Matrix A hängt von den vorgegebenen Basen in V und W ab.

Aus was oben zu den Beziehung lin. Abbildung $\Leftrightarrow$ Matrix gesagt wurde ist nun klar, dass:

1. Isomorphismus, d.h. Abbildung bijektiv, wenn

$$rg\ A = m = n$$

2. Für zwei Abbildungen $\varphi: V \to W \wedge \psi: V \to W$ gilt:

2.1 für $\lambda\varphi = \lambda A_\varphi$ 2.2 $\varphi + \psi = A_\varphi + A_\psi$

3. $\dim V = n$, $\dim W = m$, $\dim X = r$; $\varphi: V \to W \wedge \psi: W \to X$ mit $A_\varphi\ (m \times n)$ und $A_\psi\ (r \times m)$

dann gehört zu $\psi \circ \varphi$ $A_\psi \cdot A_\varphi$.

↳ Die Definition des Matrizenprodukts, wie wir sie vorgenommen haben, ist also gerechtfertigt durch die Hintereinanderausführung von linearen Ausbildungen.

4. $rg(\varphi) = rg\ A = \dim \varphi(V)$

Zusammenfassend gilt also Folgendes: Seien V und W: Vektorräume mit $\dim V = n \wedge \dim W = m$ und $\varphi: V \to W$ = lin. Abbildung und A Matrix die die lin. Abbildung beschreibt, dann

sind folgende Bedingungen gleichwertig: wobei $rg\ \varphi = \dim \varphi(V)$

$$\varphi \text{ injektiv} = \ker \varphi \{0\} = \dim \ker \varphi = 0 = rg\ \varphi = n = rg\ A = n$$

Ferner sind folgende Bedingungen gleichwertig

$$\varphi \text{ surjektiv (d.h. } W = \varphi(V)) = rg(\varphi) = m \quad (rg(\varphi) = \dim \varphi(V))$$
$$= rg\ A = m$$

Schliesslich sind folgenden Bedingungen gleichwertig:

φ ist bijektiv $= \dim V = \dim W \wedge \varphi$ ist injektiv
$\dim V = \dim W \wedge \varphi$ ist surjektiv.

↳ d.h. es gilt: φ injektiv $\Leftrightarrow rg\ A = n$
φ surjektiv $\Leftrightarrow rg\ A = m$
φ bijektiv $\Leftrightarrow rg\ A = m = n$.

Wichtig: Es gilt für eine beliebige lin. Abbildung $f: \mathbb{R}^n \to \mathbb{R}^m$ sowie Isomorphismen $h: \mathbb{R}^m \to \mathbb{R}^m$ $\wedge$ $\psi: \mathbb{R}^n \to \mathbb{R}^n$:

$$rg(h \circ f) = rg\, f \quad \text{und} \quad rg(f \circ \psi) = rg\, f$$

Koordinatentransformation:

Definition Koor. transformation: Sei V = Vektorraum mit $\dim V = n$ und es seien $B = \{\underline{p}^1, \dots, \underline{p}^n\}$ $\wedge$ $C = (\underline{q}^1, \dots, \underline{q}^n)$ Basen von V.

↳ Die lin. Abb. $f: V \to V$ die durch $\boxed{f(\underline{p}^i) = \underline{q}^i}$ gegeben ist, heisst Koordinatentransformation.

↳ Die Matrix T die diese Abbildung beschreibt heisst Transformationsmatrix bezüglich der Basis B.

Es muss also gelten: $Q = P \cdot T$

und somit gilt $P = Q \cdot T^{-1}$

! Daraus folgt unmittelbar, dass die Matrix S, d.h. die Transformationsmatrix bezüglich der Basis C die Inverse der Matrix T sein muss, d.h.

$$S = T^{-1}$$

Sei nun $\underline{v} \in \mathbb{R}^n$ beliebig und seien $\underline{x}$ und $\underline{y}$ die Komponentenvektoren von $\underline{v}$ bezüglich B bzw. C, d.h.:

$$\underline{v} = \sum_{i=1}^{n} x_i \underline{p}^i = \sum_{i=1}^{n} y_i \underline{q}^i$$

Somit gilt die Matrixdarstellung:

$P = Q\,S$

$S = Q^{-1} P = T^{-1}$

$$\underline{v} = P \cdot \underline{x} = Q \cdot \underline{y}$$

wodurch unmittelbar folgt:

$$\boxed{\underline{y} = Q^{-1} P \underline{x} = T^{-1} \underline{x}}$$

äquivalent zu $S\underline{x}$

also man transformiert die Koeffizienten für B in Koeffizienten für C

$$\boxed{\begin{aligned} &Q \cdot S\underline{x} \\ &= P\underline{x} \end{aligned}}$$

$= Q T^{-1} \underline{x}$

Skalarprodukt und Norm

Skalarprodukt ist eine Verknüpfung die zwei Vektoren eines Vektorraums auf solche Weise eine reelle Zahl zuordnet, dass man einen Winkel zwischen diesen Vektoren betrachten kann.

<u>Definition:</u> Ist in einem Vektorraum V zu je zwei Vektoren $\underline{v}, \underline{w} \in V$ <u>eindeutig</u> eine reelle Zahl $\langle \underline{v}, \underline{w} \rangle$ mit folgende Eigenschaften:

für Verständnis betrachte Beispiel $v_1 w_1 + v_2 w_2$

1. $\langle \underline{v}, \underline{w} \rangle = \langle \underline{w}, \underline{v} \rangle$
2. $\langle \lambda \underline{v}, \underline{w} \rangle = \lambda \langle \underline{v}, \underline{w} \rangle$ für beliebige $\lambda \in \mathbb{R}$
3. $\langle \underline{u} + \underline{v}, \underline{w} \rangle = \langle \underline{u}, \underline{w} \rangle + \langle \underline{v}, \underline{w} \rangle$ für beliebige $\underline{u} \in V$
4. $\langle \underline{v}, \underline{v} \rangle > 0$ falls $\underline{v} \neq \underline{0}$

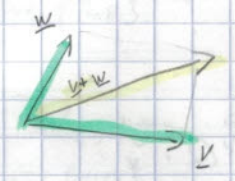
zugeordnet, so heisst $\langle \underline{v}, \underline{w} \rangle$ Skalarprodukt der Vektoren $\underline{v} \wedge \underline{w}$.

<u>Orthogonalität:</u> Ist V ein Vektorraum mit Skalarprodukt $\langle \cdot , \cdot \rangle$, dann heissen zwei Vektoren $\underline{v} \neq \underline{0} \wedge \underline{w} \neq \underline{0}$ aus V orthogonal zueinander falls $\langle \underline{v}, \underline{w} \rangle = 0$ gilt.
Wir schreiben dann $\underline{v} \perp \underline{w}$.

<u>Definition Norm:</u> Ist V ein beliebiger Vektorraum, dann heisst eine auf V definierte $\boxed{\text{reellwertige Funktion}}$ $\underline{v} \in V \mapsto \|\underline{v}\| \in \mathbb{R}$ eine Norm auf V und die Zahl $\|\underline{v}\|$ Norm von $\underline{v}$, falls folgende Eigenschaften erfüllt sind:

1. $\|\underline{v}\| > 0 \quad \forall \underline{v} \in V \setminus \{\underline{0}\}$ (Definitheit)
2. $\|\lambda \underline{v}\| = |\lambda| \, \|\underline{v}\|$ für alle $\lambda \in \mathbb{R} \wedge \underline{v} \in V$ (Homogenität)
3. $\|\underline{v} + \underline{w}\| \leq \|\underline{v}\| + \|\underline{w}\|$ für alle $\underline{v}, \underline{w} \in V$ (Dreiecksungleichung)

↳ da



Sei V ein beliebiger Vektorraum mit einem Skalarprodukt $\langle \cdot , \cdot \rangle$, dann heisst für $\underline{v} \in V$

$$\|\underline{v}\| = \sqrt{\langle \underline{v}, \underline{v} \rangle}$$

die durch das Skalarprodukt $\langle \cdot , \cdot \rangle$ induzierte Norm von $\underline{v}$.

<u>Cauchy-Schwarz'sche Ungleichung:</u> Sei V ein beliebiger Vektorraum mit einem Skalarprodukt $\langle \cdot , \cdot \rangle$ und dadurch induzierter Norm $\|\cdot\|$. Dann gilt für beliebige $\underline{v}, \underline{w} \in V$ die Cauchy-Schwarz'sche Ungleichung:

$$|\langle \underline{v}, \underline{w} \rangle| \leq \|\underline{v}\| \cdot \|\underline{w}\|$$

Zu den wichtigsten Normen zählen:

euklidische Norm: $\|\underline{x}\|_2 := \sqrt{\underline{x}^T \cdot \underline{x}}$

Maximumnorm: $\|\underline{x}\|_\infty := \max_{1 \leq j \leq n} |x_j|$

Summennorm: $\|\underline{x}\|_1 := \sum_{j=1}^{n} |x_j|$

! Sei V ein beliebiger Vektorraum mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ und dadurch induzierter Norm $\| \|$.
Dann gilt für $\underline{v}, \underline{w} \in V$ in der Cauchy-Schwarz'schen Ungleichung die Gleichheit, d.h.

$$|\langle \underline{v}, \underline{w} \rangle| = \|\underline{v}\| \cdot \|\underline{w}\| \quad \Rightarrow$$

dann gilt nämlich für die Norm: $\sqrt{\underline{v}^T \underline{v}} = \underline{v}^T \underline{v} = \langle \underline{v}, \underline{v} \rangle$

genau dann, wenn $\{\underline{v}, \underline{w}\}$ linear abhängig sind.

! Die Gleichheit in der Dreiecksungleichung:

$$\|\underline{v} + \underline{w}\| = \|\underline{v}\| + \|\underline{w}\|$$

gilt hingegen genau wenn $\{\underline{v}, \underline{w}\}$ positiv linear abhängig sind d.h. wenn die zwei Vektoren in der gleiche Richtung laufen.

↳ d.h. es muss also ein $\lambda \geq 0$ existieren, so dass $\underline{v} = \lambda \underline{w}$ oder $\underline{w} = \lambda \underline{v}$

Orthonormalsystem und Orthonormalbasis:

Sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und dadurch induzierter Norm $\| \cdot \|$. Ist $I = \{1, \dots, n\}$, dann heisst die Menge B der Form:

$$B = \{\underline{v}^i \mid i \in \underline{I}\} \subset V \quad \text{mit } \|\underline{v}^i\| = 1 \quad \forall i \in I \quad \text{und} \quad \langle \underline{v}^i, \underline{v}^j \rangle = 0 \quad \forall i, j \in \underline{I}: i \neq j$$

ein Orthonormalsystem (ONS) in V. Ist B überdies eine Basis von V, so heisst B Orthonormalbasis von V.

Es gilt nun: Sei B eine Orthonormalbasis so gilt für ein beliebiges $\underline{w} \in V$:

$$\underline{w} = \sum_{j=1}^{n} x_j \underline{v}^j$$

da aus ONS folgt $\langle \underline{v}^j, \underline{v}^i \rangle = 0$

$$\Rightarrow \langle \underline{w}, \underline{v}^i \rangle = \left\langle \sum_{j=1}^{n} x_j \underline{v}^j, \underline{v}^i \right\rangle = \sum_{j=1}^{n} x_j \langle \underline{v}^j, \underline{v}^i \rangle = x_i \|\underline{v}^i\|^2$$

somit gilt da $\|\underline{v}^i\| = 1$ $\quad \boxed{x_i = \langle \underline{w}, \underline{v}^i \rangle}$

Lin. Approximationsprobleme:

Das Approximationsproblem kann wie folgt aufgeschrieben werden:

$$\min_{\underline{v} \in W} \| \underline{v} - \underline{w} \|$$

Sei nun V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und dadurch induzierter Norm $\| \cdot \|$.
Seien $\underline{v} \in V$ sowie $\mathcal{B} = \{ \underline{w}^1, \dots, \underline{w}^n \}$ ein ONS in V und $W = \text{lin } \mathcal{B}$. Dann hat das lineare Approximationsproblem die eindeutige Lösung $\underline{w}^* = \underline{w}(\underline{v})$ mit:

$$\boxed{\underline{w}(\underline{v}) = \sum_{j=1}^{n} \underbrace{\langle \underline{v}, \underline{w}^j \rangle}_{\text{Koordinaten von } \underline{w}^j} \underline{w}^j}$$

$\cdot\, \underline{w}^j \Rightarrow$ Punkt der am nähesten ist.

Gram-Schmidt'sche Orthonormierung - geometrische Idee:

↳ Somit wissen wir nun wie man ein ONS konstruieren kann:

gegeben $A = \{ \underline{w}^1, \dots, \underline{w}^n \} \subset V$ lin. unabhängig.

Schritt 1: Normiere $\underline{w}^1$ → es ergibt sich $\underline{v}^1 = \frac{1}{\|\underline{w}^1\|} \cdot \underline{w}^1$

Schritt 2: Projiziere $\underline{w}^2$ auf $\underline{v}^1$ bilde das Vektor $\underline{w}^2 - w(\underline{w}^2)$ und normiere dies.
(für Projektion gilt Formel oben $w(\underline{v})$)

Schritt 3: Tue dies bis alle Vektoren von A in eine ONS überführt werden.

4. Mache die Probe, dass die Vektoren orthogonal aufeinander stehen.

! Beachte: In jedem endlich-dimensionalen Vektorraum mit Skalarprodukt existiert eine Orthonormalbasis, man kann sie z.B. mit Hilfe des Orthonormalverfahrens von Gram und Schmidt (siehe oben) aus einer gegebenen Basis konstruieren.

Orthogonales Komplement

Sei V ein Vektorraum mit Skalarprodukt $\langle .,. \rangle$ und W ein Unterraum von V.

↳ Dann heisst:

$$W^{\perp} = \{ \underline{v} \in V \mid \langle \underline{v}, \underline{w} \rangle = 0 \text{ für alle } \underline{w} \in W \}$$

das orthogonale Komplement von W.

! Die Menge $W^{\perp}$ ist wieder ein Unterraum von V. Jeder Vektor $\underline{v} \in V$ lässt sich somit darstellen als:

$$\underline{v} = \underbrace{\underline{w}(\underline{v})}_{\substack{\text{Vektor}\\ \text{aus } W\\ \text{orthogonal}\\ \text{zu } \underline{v}}} + \underbrace{(\underline{v} - \underline{w}(\underline{v}))}_{\substack{\text{Vektor aus}\\ W^{\perp}\text{, d.h.}\\ \text{Vektor orthogonal}\\ \text{zu } \underline{w} \in W}}$$

mit $\underline{w}(\underline{v}) \in W$
und $(\underline{v} - \underline{w}(\underline{v})) \in W^{\perp}$

Eigenwerte und Eigenvektor Definition:

Sei A eine quadratische Matrix der Ordnung n.

↳ Gilt dann für eine reelle Zahl λ und einen Vektor $\underline{x} \in \mathbb{R}^n$ mit $\underline{x} \neq \underline{0}$ die Gleichung

$$A\underline{x} = \lambda \underline{x}$$

so heisst $\underline{x}$ = Eigenvektor zum Eigenwert λ
λ = Eigenwert.

! Zu jedem Eigenwert existiert eine Menge von Eigenvektoren da

$$A \cdot \alpha \underline{x} = \lambda \cdot \alpha \underline{x} \qquad \text{mit } \alpha \in \mathbb{R}$$

⇒ d.h. $\alpha \underline{x}$ ist auch Eigenvektor von λ wenn $\underline{x}$ Eigenvektor zum Eigenwert λ ist.

⇒ Die Menge von Eigenvektoren zum Eigenwert λ wird als Eigenraum definiert.

Eigenraum: Seien $\lambda \in \mathbb{R}$ und $A \in M(n,n)$ gegeben und

$$\text{Eig}(A;\lambda) := \{ \underline{x} \mid A\underline{x} = \lambda \underline{x} \}$$

⇒ Offenbar ist stets $\underline{o} \in \text{Eig}(A;\lambda)$ und <u>$\text{Eig}(A;\lambda)$ ist ein Unterraum von $\mathbb{R}^n$</u>

! Nach Definition ist nun λ Eigenwert von A genau dann, wenn $\dim \text{Eig}(A;\lambda) \geq 1$.

↳ Da wenn $\dim \text{Eig}(A;\lambda) = 0$, dann ist nur $\underline{o}$ zugelassen was nach Voraussetzung Eigenvektor muss unterschiedlich von $\underline{o}$ sein zu keinem Eigenwert führt.

⇒ $\text{Eig}(A;\lambda)$ heisst nun $\boxed{\text{Eigenraum}}$ bezüglich λ.

↳ da durch Umformung gilt:

$$\text{Eig}(A;\lambda) = \{ \underline{x} \mid (A - \lambda I)\underline{x} = \underline{o} \}$$

gilt nun $\dim \text{Eig}(A,\lambda) = n - \text{rg}(A - \lambda I)$

Charakteristisches Polynom:

Sei $A \in M(n,n)$. Die Zahl

$$P(t) = \det(A - tI)$$

ist ein Polynom n-ten Grades und heisst charakteristisches Polynom.

! Das charakteristische Polynom einer Diagonalmatrix Λ mit Einträgen λ_i auf der Hauptdiagonalen hat somit nun nach Determinantengesetzen die Form:

$$\det(\Lambda - tI) = \underbrace{\prod_{i=1}^{n}}_{\text{Produkt}} (\lambda_i - t)$$

↳ Als elementare Schlussfolgerung von:

$$A^T - tI = (A - tI)^T \overset{\text{dass}}{\Rightarrow} \overbrace{\det(A - tI)^T = \det(A - tI)}^{\text{siehe Determinantengesetze}}$$

⇒ $\boxed{\text{Somit haben } A^T \text{ und } A \text{ das gleiche charakteristische Polynom}}$

Wichtig: Es gilt nun:

Die Zahl $\lambda \in \mathbb{R}$ ist genau dann Eigenwert der Matrix $A \in M(n,n)$, wenn es gilt:

$$\det(A - \lambda I) = 0 \quad \Rightarrow \text{ da hier } \dim \mathrm{Eig}(A;\lambda) \geq 1$$

→ d.h., wenn λ eine Nullstelle des charakteristischen Polynom ist.

da $\mathrm{rg}(A-\lambda I) <$

Beachte: Die Determinante ist gegenüber Ähnlichkeitstransformationen invariant, d.h.

→ sind $A, B \in M(n,n)$ ähnlich, so gilt $\det A = \det B$.

! Das charakteristische Polynom und damit auch die Eigenwerte einer Matrix sind gegenüber Ähnlichkeitstransformationen invariant.

$\boxed{\text{Diagonalisierbar:}}$ Eine Matrix $A \in M(n,n)$ heisst diagonalisierbar, wenn eine zu A ähnliche Diagonalmatrix $\Lambda \in M(n,n)$ existiert.

! A ist genau dann diagonalisierbar, wenn es (nicht notwendig verschiedene) Eigenwerte $\lambda_1, \dots, \lambda_n$ <u>und</u> eine Basis des $\mathbb{R}^n$ aus zugehörigen Eigenvektoren $\underline{x}^1, \dots, \underline{x}^n$ gibt, so dass:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = T^{-1} A T \quad \text{mit } T = [\underline{x}^1, \dots, \underline{x}^n]$$

gilt.

Symmetrische Matrizen = stets diagonalisierbar !!

Beachte: <u>Für diagonalisierbare Matrizen</u> $\underset{A \in M(n,n)}{\forall}$ folgt sofort, dass

$$\det A = \lambda_1 \cdot \ldots \cdot \lambda_n$$

Ausserdem muss gelten:

$$\lambda_1 + \ldots + \lambda_n = \text{spur } A := a_{11} + a_{22} + \ldots + a_{nn}$$

Benutze dies für Überprüfung der korrekten Berechnung der Eigenwerte.

! Besonders einfach lässt sich A^k berechnen, wenn A diagonalisierbar ist.

Es gilt nämlich:

$$\Lambda = T^{-1} A T \qquad A = T \Lambda T^{-1}$$

somit gilt $A^2 = A A = T \Lambda T^{-1} T \Lambda T^{-1} = T \Lambda^2 T^{-1}$

und für $A^k = T \Lambda^k T^{-1}$

! Die Eigenwerte von A^k sind somit $(\lambda_j)^k \; \forall j$, die Eigenräume sind hingegen die gleichen wie die von A

Sei A eine symmetrische Matrix (d.h. $A^T = A$) n-reihige Matrix. Sind λ und μ zwei voneinander verschiedene Eigenwerte von A mit den zugehörigen Eigenvektoren $\underline{x}$ und $\underline{y}$, so sind $\underline{x}$ und $\underline{y}$ orthogonal.

! Das charakteristische Polynom einer n-reihigen symmetrischen Matrix A hat, mit ihrer Vielfachheit gezählt (d.h. es können auch gleiche Eigenwerte gleich sein) genau n reelle Nullstellen und es gibt zu diesen Eigenwerten ein Orthonormalsystem von n zugehörigen Eigenvektoren, die dann eine Orthonormalbasis des $\mathbb{R}^n$ bilden.

↳ A ist folglich diagonalisierbar, d.h., ähnlich zur Diagonalmatrix aus den Eigenwerten.

↳ die Ähnlichkeitstransformation erfolgt dann mit der Orthogonalmatrix T (mit also $T^T = T^{-1}$)

Definitheit von Matrizen:

Eine n-reihige symmetrische Matrix A heisst:

1. positiv definit wenn $\underline{v}^T A \underline{v} > 0$ für alle $\underline{v} \in \mathbb{R}^n \setminus \{\underline{o}\}$ gilt
2. positiv semidefinit wenn $\underline{v}^T A \underline{v} \geq 0$ für alle $\underline{v} \in \mathbb{R}^n$ gilt
3. negativ definit wenn $\underline{v}^T A \underline{v} < 0$ für alle $\underline{v} \in \mathbb{R}^n \setminus \{\underline{o}\}$ gilt
4. negativ semidefinit wenn $\underline{v}^T A \underline{v} \leq 0$ für alle $\underline{v} \in \mathbb{R}^n$ gilt
5. indefinit wenn es keine dieser Eigenschaften gilt.

! Offenbar ist A negativ definit genau dann, wenn $-A$ positiv definit ist.

↳ Analog ist A negativ semidefinit genau dann, wenn $-A$ positiv semidefinit ist.

Seien A = symmetrische Matrix und B = reguläre Matrix. Beide von Ordnung n.

↳ dann gilt:

1. A ist pos. definit genau dann, wenn $B^T A B$ positiv definit ist.
2. A ist pos. semidefinit genau dann, wenn $B^T A B$ positiv semidefinit ist.

Somit gilt wegen
T = Transformationsmatrix von A
↳ sie ist Orthogonal

$$\Lambda = T^T A T \quad \Rightarrow \quad A = T \Lambda T^T$$

$$\underline{y} = T^T \underline{x}$$

$$\Rightarrow \quad 0 < \underline{x} A \underline{x} = \underline{x} T \Lambda T^T \underline{x} = \underline{y}^T \Lambda \underline{y}$$

$$0 < \sum_{\nu=1}^{n} \lambda_\nu y_\nu^2$$

Analog gilt für $>$, $\geq$, $\leq$.

Somit gilt: Eine n-reihige symmetrische Matrix mit Eigenwerten $\lambda_1, \dots, \lambda_n$ ist:

1. positiv definit genau dann, wenn $\lambda_j > 0 \quad \forall j \in \{1, \dots, n\}$
2. positiv semidefinit genau dann, wenn $\lambda_j \geq 0 \quad \forall j \in \{1, \dots, n\}$
3. negativ definit genau dann, wenn $\lambda_j < 0 \quad \forall j \in \{1, \dots, n\}$
4. negativ semidefinit genau dann, wenn $\lambda_j \leq 0 \quad \forall j \in \{1, \dots, n\}$

Es folgt somit Matrizen die positiv definit oder negativ definit sind haben vollen Rang $\Rightarrow$ regulär.

↳ Somit folgt, dass eine positiv semidefinite (bzw. negativ semidefinite) symmetrische Matrix genau dann positiv definit (negativ definit) ist, wenn sie regulär ist.

$\Rightarrow$ Eine symmetrische Matrix ist genau dann indefinit wenn sie zwei von Null verschiedene Eigenwerte mit unterschiedlichen Vorzeichen besitzt.

Ein analoges Vorgehen um auf Definitheit von n-reihige symmetrische Matrizen zu schliessen ist:

A positiv definit genau dann, wenn alle Hauptabschnittsdeterminanten $\det A^{[k]}$ von A positiv sind.

A ist negativ definit genau dann, wenn die Hauptabschnittsdeterminanten $\det A^{[k]}$ von A positiv im Falle gerader Ordnung von $A^{[k]}$ aber negativ im Falle von ungerader Ordnung von $A^{[k]}$ sind.

Analysis

Beschränktheit: Eine Menge $M \subset \mathbb{R}$ heisst nach oben beschränkt (bzw. nach unten beschränkt), wenn ein $\bar{k} \in \mathbb{R}$ (bzw. ein $\underline{k} \in \mathbb{R}$) existiert, so dass $x \leq \bar{k}$ (bzw. $x \geq \underline{k}$) für alle $x \in M$ gilt.

Eine nach oben und nach unten beschränkte Menge $M \subset \mathbb{R}$ heisst beschränkt.

! Jede nach oben beschränkte Menge besitzt eine kleinste obere Schranke = Supremum von M

Jede nach unten beschränkte Menge $M \subset \mathbb{R}$ besitzt eine grösste untere Schranke = Infimum von M.

! Falls nun $\sup M \in M$ → dann spricht man von Maximum von M.
Falls $\inf M \in M$ → dann spricht man von Minimum von M.

Konvergenz reeller Folgen: Eine Folge $\{a_k\} \in \mathbb{R}$ hat den Grenzwert (auch Limes) $a^* \in \mathbb{R}$, falls zu jedem $\varepsilon > 0$ ein $m \in \mathbb{N}$ existiert, so dass $|a_k - a^*| < \varepsilon$ $\boxed{\text{für alle } k \geq m}$.

↳ Man schreibt $a^* = \lim_{k \to \infty} a_k$ und sagt kurz $\{a_k\}$ ist konvergent.

! Eine Folge $\{a_k\} \subset \mathbb{R}$ heisst monoton fallend falls $a_{k+1} \leq a_k$, monoton wachsend falls $a_{k+1} \geq a_k$ (im Fall von streng monoton $<, >$).

! ↳ Jede monotone, beschränkte Folge ist auch konvergent.

Satz von Bolzano-Weierstrass: Jede beschränkte Folge $\{a_k\} \subset \mathbb{R}$ hat mindestens einen Häufungspunkt, d.h. es existiert eine unendliche Teilfolge von $\{a_k\}$ die konvergiert.

Normäquivalenz: Sind $\|\cdot\|$ und $\|\cdot\|'$ irgend zwei unterschiedliche Normen im $\mathbb{R}^n$ so existieren Zahlen $\alpha > 0$ und $\beta > 0$, so dass für alle $x \in \mathbb{R}^n$ gilt: $\alpha \|x\| \leq \|x\|' \leq \beta \|x\|$

Definitionen: Sei nun $\|\cdot\|$ irgendeine Norm im $\mathbb{R}^n$ und $z \in \mathbb{R}^n$.

z = Mittelpunkt

Die Menge $B^\circ(z, r) := \{x \in \mathbb{R}^n \mid \|x - z\| < r\}$ $(r > 0)$ heisst offene Kugel um z mit Radius r oder offene r-Umgebung.

Die Menge $B(z, r) := \{x \in \mathbb{R}^n \mid \|x - z\| \leq r\}$ heisst abgeschlossene Kugel um $z \in \mathbb{R}^n$ mit Radius r.

! Eine Teilmenge $U \subset \mathbb{R}^n$ heisst Umgebung von z falls ein $\varepsilon > 0$ $\exists$, so dass $B^\circ(z, \varepsilon) \subset U$

Ein Punkt $z \in M$ einer Teilmenge M von $\mathbb{R}^n$ heisst **innerer Punkt** von M falls eine Umgebung U von z existiert, die in M enthalten ist.

↳ Die Menge aller innerer Punkte $=$ int M

❗ Eine Teilmenge $M \subset \mathbb{R}^n$ heisst offen falls alle Punkt von M innere Punkt von M sind, also falls $M =$ int M gilt.

Eine Teilmenge $M \subset \mathbb{R}^n$ heisst abgeschlossen, falls die Komplementmenge $\mathbb{R}^n \setminus M$ offen ist.

Eine Teilmenge $M \subset \mathbb{R}^n$ heisst beschränkt, falls es ein $x \in \mathbb{R}^n$ und ein $r > 0$ gibt, so dass $M \subset B(x, r)$ gilt.

❗ $\cup$ und $\cap$ endlich vieler offener Menge ergibt = offene Menge
$\cup$ und $\cap$ endlich vieler abgeschlossener Menge ergibt = abgeschlossene Menge

$\cup$ ∞-vieler offener Menge = offene Menge
$\cap$ ∞-vieler abgeschlossene Menge = abgeschlossene Menge

$\cap$ ∞-vieler offener Menge muss nicht immer offen sein
$\cup$ ∞-vieler abgeschlossene Menge muss nicht immer abgeschlossen sein.

abgeschlossen

↳ z.B. $\cup$ der einelementige Menge $\left\{\frac{1}{k}\right\}$ $k \in \mathbb{N}$

↳ hier gilt $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots + \ldots + \ldots$ $\lim_{k \to \infty} a_k = \infty$

Konvergenz

Man sagt die Folge $\{a_k\} \subset \mathbb{R}^n$ konvergiert gegen den Punkt $a^* \in \mathbb{R}^n$, falls zu jedem $\varepsilon > 0$ die Kugel $B^0(a^*, \varepsilon)$ fast alle Elemente der Folge $\{x_k\}$ enthält, d.h. falls zu jedem $\varepsilon > 0$ ein m existiert, so dass:

$$\| a_k - a^* \| < \varepsilon \quad \text{für alle} \quad k \geq m$$

Dabei heisst a^* Limes oder Grenzelement der Folge $\{a_k\}$ und man schreibt:

$$a^* = \lim_{k \to \infty} a_k$$

❗ Mit anderen Worten: Die Folge $\{a_k\}$ konvergiert genau dann gegen a, wenn $\| a_k - a^* \| \to 0$ für $k \to \infty$

Sehr wichtig

❗ Eine Menge $M \subset \mathbb{R}^n$ ist abgeschlossen genau dann, wenn für jede konvergente Folge $\{x_k\}_{k \in \mathbb{N}} \subset M$ ihr Limes auch zu M gehört.

Stetigkeit von Funktionen: Eine Funktion $f: D \subset \mathbb{R}^n \to \mathbb{R}^m$ heisst <u>stetig im Punkt</u> $z \in D$ falls: $\lim_{x \to z} f(x) = f(z)$

d.h. falls zu jeder Folge $\{x_k\}_{k \in N} \subset D$ mit $\lim_{k \to \infty} x_k = z$ gilt:

$$\boxed{\lim_{k \to \infty} f(x_k) = f(z)}$$

! f heisst <u>stetig auf D</u> falls f in jedem Punkt $z \in D$ stetig ist.

! Ist f eine Vektorfunktion (also $m > 1$) bestehend aus den Komponenten $f_1, \dots, f_m$ mit $f_j: D \subset \mathbb{R}^n \to \mathbb{R}$ für alle j von 1 bis m, <u>dann bedeutet Stetigkeit von f, dass jede Komponente stetig sein muss.</u>

Verknüpfungen: Verknüpfungen von stetige Funktionen sind stetig. Bei der Verknüpfung f/g muss die Voraussetzung $(g(x))^2 > 0$ erfüllt sein.

! Seien $f: \mathbb{R}^n \to \mathbb{R}$ stetig und a = Konstante $\in \mathbb{R}$, dann gilt für die Meng

$\{x \in \mathbb{R}^n \mid f(x) < a\}$ - offene Menge

$\{x \in \mathbb{R}^n \mid f(x) \le a\}$ ∧ $\{x \in \mathbb{R}^n \mid f(x) = a\}$ = abgeschlossene Menge

Kompaktheit: Definition: Sei $M \subset \mathbb{R}^n$. Dann heisst M <u>kompakt</u>, wenn aus jeder Folge $\{x_k\}_{k \in N} \subset M$ eine konvergente Teilfolge $\{x_{k_\nu}\}_{\nu \in N}$ ausgewählt werden kann, deren Limes x^* ebenfalls zu M gehört.

⇒ Sei $M \subset \mathbb{R}^n$. Dann ist M genau dann kompakt, wenn M beschränkt und abgeschlossen ist.

! Sei nun M kompakt, dann muss auch $f(M)$ kompakt sein.

Extremwertsatz von Weierstrass:

Seien M eine nichtleere kompakte Teilmenge des $\mathbb{R}^n$ und $f: M \to \mathbb{R}$ stetig. Dann gibt es ein $x^* \in M$ und ein $x_* \in M$, so dass $f(x^*) = \max\{f(x) \mid x \in M\}$ und $f(x_*) = \min\{f(x) \mid x \in M\}$, d.h. die Funktion nimmt auf M sowohl ihr Maximum als auch ihr Minimum an.

Differentialrechnung im $\mathbb{R}^n$

Partiell differenzierbar: Die Funktion $x \mapsto f(x)$ heisst im Punkt x^0 partiell differenzierbar nach x_i, falls der Limes $f_{x_i}(x^0) = \lim_{t \to 0} \frac{f(x_0 + t \cdot e^i) - f(x_0)}{t}$ existiert. Dabei ist e^i der i-te Einheitsvektor im $\mathbb{R}^n$.

Gradient: Als Gradient von f im Punkt $x^0 = (x_1^0, \dots, x_n^0)^T$ bezeichnen wir den Vektor:

$$\nabla f(x^0) = \begin{pmatrix} f_{x_1}(x^0) \\ \vdots \\ f_{x_n}(x^0) \end{pmatrix}$$

! Ist f partiell differenzierbar nach x_i auf D (d.h. in allen Punkten $x \in D$), so heisst die Funktion

$x \in D \to f_{x_i}(x)$ die partielle Ableitungsfunktion von f nach x_i.

Hesse-Matrix: Wenn in x^0 zweite partielle Ableitungen von f nach x_i und x_j für alle $i, j \in \{1, \dots, n\}$ existieren, fasst man sie in der Hesse Matrix:

$$\nabla^2 f(x^0) = \left(f_{x_i x_j}(x^0) \right)_{i,j = 1,n}$$ zusammen.

Satz von Schwarz: Sei f auf einer offenen Umgebung U von $x^0 \in \mathbb{R}^n$ nach x_i und x_j partiell differenzierbar und existiere $f_{x_i x_j}(x)$ für alle $x \in U$.
Ist nun die Funktion $f_{x_i x_j}(\cdot)$ stetig auf U, dann existiert auch $f_{x_j x_i}(x)$ und es gilt:

$$\boxed{f_{x_i x_j}(x) = f_{x_j x_i}(x) \quad \forall x \in U}$$

Differenzierbarkeit: Seien $f: D \subset \mathbb{R}^n \to \mathbb{R}$ eine reellwertige Funktion, D eine offene Menge und $x^0 \in D$.

f heisst im Punkt x^0 differenzierbar (genauer total differenzierbar) falls:

$$\boxed{f(x^0 + u) = f(x^0) + \underbrace{\nabla f(x^0)^T u}_{\text{Differential}} + \underbrace{o(u)}_{\text{Fehler}}}$$

für $x^0 + u \in D$ gilt (dies ist die Linearisierungsformel von Weierstrass) wobei es gelten muss:

$$o(0) = 0 \quad \text{und} \quad \frac{o(u)}{\|u\|} \to 0 \quad \text{falls} \quad \|u\| \to 0$$

Anders gesagt muss zu jedem $\epsilon > 0$ ein $\delta > 0$ geben, so dass gilt:

$$| f(x^0 + u) - f(x^0) - \nabla f(x^0)^T u | \leq \epsilon \| u \|$$

für alle $u \in \mathbb{R}^n$ mit $\| u \| < \delta$

Differenzierbarkeit bei Vektorfunktionen: Sei nun $\underline{f} = (f_1, \ldots, f_m)^T$ eine Vektorfunktion, also $f: D \in \mathbb{R}^n \to \mathbb{R}^m$ und seien wieder D eine offene Menge und $x^0 \in D$.

f heisst nun im Punkt x^0 differenzierbar, falls die Komponentenfunktionen f_i, $i = 1, \ldots, m$ differenzierbar sind.

Fassen wir die Gradienten zeilenweise zu einer Matrix $Df(x^0)$ zusammen, so hat diese die Form:

Jacobi-Matrix:
$$\boxed{Df(x^0) = \begin{pmatrix} \nabla f_1(x^0)^T \\ \nabla f_2(x^0)^T \\ \vdots \\ \nabla f_m(x^0)^T \end{pmatrix}}$$

!

und somit gilt nun: $\boxed{f(x^0 + u) = f(x^0) + Df(x^0) \cdot u + o(u)}$

für $x^0 + u \in D$ wobei $o(u) = (o_1(u), \ldots, o_m(u))$ nun den Vektor der Fehlerfunktionen o_i für die Funktionen f_i ist, es gilt also:

$$o(u) = 0 \quad \text{und} \quad \frac{o(u)}{\|u\|} \to 0 \quad \text{wenn } \|u\| \to 0$$

Eine Funktion mit dieser Eigenschaft von $o(\cdot)$ nennt man o-Typ-Funktion

! Die lineare Abbildung $L: \mathbb{R}^n \to \mathbb{R}^m$ mit $L(u) := Df(x^0)u$ heisst (Fréchet)-Ableitung von f im Punkt x^0.

Ableitung der Linearkombination von differenzierbaren Funktionen:

Elementar aus den Definitionen folgt: sind $f, g: D \subset \mathbb{R}^n \to \mathbb{R}^m$ differenzierbar in $x^0 \in D$, D offen, so ist auch $\lambda f + \mu g$ mit $\lambda, \mu \in \mathbb{R}$ differenzierbar in x^0, es gilt:

$$D(\lambda f + \mu g)(x^0) = \lambda Df(x^0) + \mu Dg(x^0)$$

Totales Differential und Tangentialebene reellwertiger Funktionen:

Für: $f: D \subset \mathbb{R}^n \to \mathbb{R}$, D offen, $x^0 \in D$

$$\boxed{\text{Lineare Approximation: } l(x) = f(x^0) + \nabla f(x^0)^T \cdot \overbrace{(x - x^0)}^{u}}$$

Im Falle von $n = 2$ gilt also:

$$l(x,y) = f(x^0, y^0) + f_x(x^0, y^0)\,(x - x^0) + f_y(x^0, y^0)\,(y - y^0)$$

$$\boxed{\text{totale Differential von } f \text{ in } x^0\text{: } df = \nabla f(x^0)^T\, dx}$$

z.B. also für $n = 2$

$$df = f_{x_1}(x_1^0, x_2^0) \cdot dx_1 + f_{x_2}(x_1^0, x_2^0) \cdot dx_2$$

Tangentialhyperebene: Führt man nun die neue Variable $x_{n+1} = l(x)$ ein und setzt man für $u = x - x^0$ ergibt sich die Tangentialhyperebene:

$$x_{n+1} = f(x^0) + \nabla f(x^0)^T \cdot (x - x^0)$$

z.B. für $n = 2$

$$x_3 = f(x^0) + f_{x_1}(x^0) \cdot (x_1 - x_1^0) + f_{x_2}(x^0)\,(x_2 - x_2^0)$$

Definitionen:

Eine Funktion $f: D \subset \mathbb{R}^n \to \mathbb{R}$, die nach allen Variablen im Punkt x^0 stetige partielle Ableitungen hat, heisst (auf D) stetig differenzierbar im x^0 oder stetig partiell differenzierbar in x^0 nach allen Variablen.

Eine Funktion $f: D \subset \mathbb{R}^n \to \mathbb{R}$, die nach allen Variablen auf D stetige partielle Ableitungen hat, heisst stetig differenzierbar oder stetig partiell differenzierbar nach alle Variablen.

Analoges Konzept bei zweimal stetig differenzierbar in x^0 und zweimal stetig differenzierbar.

verallgemeinerte Kettenregel:

Seien $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^m$ offene Mengen sowie $g: X \to \mathbb{R}^m$ und $f: Y \to \mathbb{R}^q$ Funktionen mit $g(X) \subset Y$.

Sind g in $x^0 \in X$ differenzierbar und f in $y^0 = g(x^0)$ differenzierbar, dann ist auch

$$h := f \circ g : X \to \mathbb{R}^q, \quad h(x) := f(g(x)), \ x \in X$$

in x^0 differenzierbar.

und es gilt:

$$\underbrace{\overset{[q,n]}{Dh(x^0)}}_{\text{Jacobi-Matrix von } h \text{ im Punkt } x^0} = \underbrace{\underbrace{\overset{[q,m]}{Df(g(x^0))}}_{\text{Aussere Ableitung}}}_{\text{Jacobi-Matrix von } f \text{ im Punkt } y^0 = g(x^0)} \cdot \underbrace{\underbrace{\overset{[m,n]}{Dg(x^0)}}_{\text{Innere Ableitung}}}_{\text{Jacobi-Matrix von } g \text{ im Punkt } x^0}$$

Spezialfälle: 1. $h(t) = f(x(t), y(t))$ mit $x, y: X \subset \mathbb{R} \to \mathbb{R}$

Kettenregel dann:

$$h'(t_0) = f_x(x(t_0), y(t_0)) \cdot x'(t_0) + f_y(x(t_0), y(t_0)) \cdot y'(t_0)$$

2. Verallgemeinerung von 1.

Aus $h(t) = f(g_1(t), \dots, g_m(t))$ $\quad g_i: X \subset \mathbb{R} \to \mathbb{R}^m \ \forall i$

$f: \mathbb{R}^m \to \mathbb{R}$

Ergibt sich sofort:

$$Dh(t_0) = h'(t_0) = \nabla f(g(t_0))^T \, Dg(t_0)$$

3. Setzt man in 2. $g(t) := y^0 + tv$, $t \in \mathbb{R}$ ($y^0, v \in \mathbb{R}^m$ gegeben) ein so ist

$$h(t) = f(y^0 + tv)$$

und es ist

$$h'(t_0) = \nabla f(y^0 + t_0 v)^T \, v$$

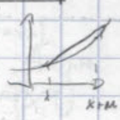
Verallgemeinerte Produktregel: Ist $D \subset \mathbb{R}^n$ eine offene Menge und sind $f, g: D \to \mathbb{R}$ differenzierbar, so ist auch $h(x) := f(x)\, g(x)$, $x \in D$, differenzierbar und es gilt:

$$\boxed{\nabla h(x) = g(x) \nabla f(x) + f(x) \nabla g(x)}$$

Mittelwertsatz der Differentialrechnung reeller Funktionen:

Sei $f: I \subset \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion über einem Intervall I. Dann gibt es zu beliebigen Punkten $x,\ x+u \in I$ ein $\theta \in (0,1)$ so dass:

$$\boxed{f(x+u) = f(x) + \underbrace{u \cdot f'(x+\theta u)}_{u \,\cdot\, \text{Steigung Sekante}}}$$



Mittelwertsatz der Differenzialrechnung in n Veränderlichen:

Sei $D \subset \mathbb{R}^n$ offen und $f: D \to \mathbb{R}$ stetig differenzierbar. Seien $x \in D$ und $u \in \mathbb{R}^n$ derart, dass $x + tu \in D$ für alle $t \in [0,1]$ erfüllt ist. Dann existiert ein $\theta \in (0,1)$, so dass:

$$\boxed{f(x+u) = f(x) + u^T \nabla f(x+\theta u)}$$

Satz von Taylor für reelle Funktionen:

Sei $f: I \subset \mathbb{R} \to \mathbb{R}$ ein $(n+1)$-mal differenzierbare Funktion und $a \in I$. Dann gibt es zu jedem $x \in I \setminus \{a\}$ ein ξ im offenen Intervall zwischen a und x, so dass gilt:

$$f(x) = P_n(x) + R_n(x)$$

wobei: $P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$ und $R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$

Es gilt also:

$$\boxed{f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{(n+1)}}$$

$P_n(x)$ heisst, dabei = n-tes Taylor-Polynom von f.

$R_n(x)$ heisst = Lagrange-Form des Restglieds der Taylor-Formel n-ter Ordnung

Taylor-Formel für Funktionen in n Variablen mit Restglied 2. Ordnung:

Sei $D \subset \mathbb{R}^n$ offen und $f: D \to \mathbb{R}$ zweimal stetig differenzierbar. Seien $x \in D$ und $u \in \mathbb{R}^n$ derart, dass $x + tu \in D$ für alle $t \in [0,1]$ erfüllt ist. Dann existiert ein $\theta \in (0,1)$, so dass gilt:

$$\boxed{f(x+u) = f(x) + u^T \nabla f(x) + \frac{1}{2} u^T \nabla^2 f(x+\theta u)\, u}$$

Notwendige Bedingung 1. Ordnung für lokale Extrema:

Es seien $D \subset \mathbb{R}^n$ offen und $f: D \to \mathbb{R}$ stetig differenzierbar. Hat f an der Stelle $x^0 \in D$ ein lokales Extremum, d.h. existiert ein $\varepsilon > 0$, so dass:

$$(i)\ f(x) \leq f(x^0) \qquad (ii)\ f(x) \geq f(x^0) \qquad \text{für alle } B^0(x^0, \varepsilon)$$

dann gilt:

$$\boxed{\nabla f(x^0) = 0}$$

Strenges lokales Min. / Max.:

Sei f eine auf einer offenen Menge D definierte reellwertige Funktion. Man sagt f hat an der Stelle $x^0 \in D$ ein strenges lokales Minimum, falls ein $\varepsilon > 0$ existiert, so dass:

$$f(x^0) < f(x) \qquad \forall x \in B^0(x^0, \varepsilon), \quad x \neq x^0$$

Analog spricht man von strenges lokales Maximum wenn:

$$f(x^0) > f(x) \qquad \forall x \in B^0(x^0, \varepsilon) \text{ mit } x \neq x^0$$

Hinreichende Optimalitätsbedingung 2. Ordnung:

Es seien $D \subset \mathbb{R}^n$ offen, $f: D \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $x^0 \in D$ ein Punkt mit $\nabla f(x^0) = 0$.

↳ So gelten folgende Kriterien:

1. Ist die Hesse-Matrix $\nabla^2 f(x^0)$ positiv definit, so hat f an der Stelle x^0 ein strenges lokales Minimum.
2. Ist die Hesse-Matrix $\nabla^2 f(x^0)$ negativ definit, so hat f an der Stelle x^0 ein strenges lokales Maximum.
3. Ist die Hesse-Matrix $\nabla^2 f(x^0)$ indefinit, so besitzt f an der Stelle x^0 ein Sattelpunkt.

Notwendige Optimalitätsbedingung 2. Ordnung:

Sei $D \subset \mathbb{R}^n$ offen, $f: D \to \mathbb{R}$ 2-mal stetig differenzierbar

1. Wenn $x^0 \in D$ ein lokales Minimum von f ist, dann muss gelten:
$$\nabla^2 f(x^0) = \text{positiv semidefinit.}$$
2. Wenn $x^0 \in D$ ein lokales Maximum von f ist, dann muss gelten:
$$\nabla^2 f(x^0) = \text{negativ semidefinit.}$$

Konvexität und Konkavität

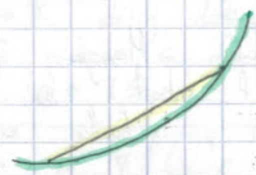
Eine **Menge** $X \subset \mathbb{R}^n$ heisst **konvex**, falls mit zwei Punkten aus X auch ihre Verbindungsstrecke in X liegt, d.h.,

$$\boxed{x, y \in X, \quad \lambda \in (0,1) \quad \Rightarrow \quad \lambda x + (1-\lambda) y \in X}$$

Eine **Funktion** ist **konvex**, falls gilt:

$$\boxed{x, y \in X, \quad \lambda \in (0,1) \quad \Rightarrow \quad f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)}$$

↳ d.h. falls — > —

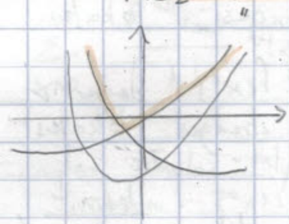


$g : X \to \mathbb{R}$ heisst **konkav**, falls $-g$ konvex ist.

Eigenschaften konvexer Mengen / Funktionen

1. Verknüpfung konvexe Mengen $\Rightarrow$ konvexe Menge.
2. Sind f und g konvexe Funktionen auf einer konvexe Menge $X \subset \mathbb{R}^n$ und μ eine positive reelle Zahl, so sind auch die Funktionen $f+g$ und μf konvex.
3. Sind $X \subset \mathbb{R}^n$ eine konvexe Menge und $f_i : X \to \mathbb{R}$, $i \in I$ (I beliebige Indexmenge), konvexe Funktionen, so ist auch die Funktion $f(x) := \sup_{i \in I} f_i(x)$ konvex.

   ↳ d.h.



4. Ist $\|\cdot\|$ irgendeine Norm im $\mathbb{R}^n$, so ist die Funktion $x \in \mathbb{R}^n \to \|x\|$ konvex.

Konvexitäts- und Konkavitätskriterium für reelle Funktionen:

Sei $D \subseteq \mathbb{R}$ ein Intervall und $f: D \to \mathbb{R}$ differenzierbar, so sind die folgende Aussagen paarweise zueinander äquivalent:

1. f ist konvex über D
2. $f(\varepsilon) \geq f(x) + f'(x)(\varepsilon - x)$ für alle $\varepsilon, x \in D$ ⟸ (folgt da Tangente immer unterhalb Funktion)
3. f' ist monoton steigend auf D (d.h., wenn f zweimal differenzierbar ist, gilt $f''(x) \geq 0$ für alle $x \in D$)

und

1. f konkav über D
2. $f(\varepsilon) \leq f(x) + f'(\varepsilon - x)$ für alle $\varepsilon, x \in D$
3. f' ist monoton fallend auf D (d.h., wenn f zweimal differenzierbar ist, gilt $f''(x) \leq 0$ für alle $x \in D$)

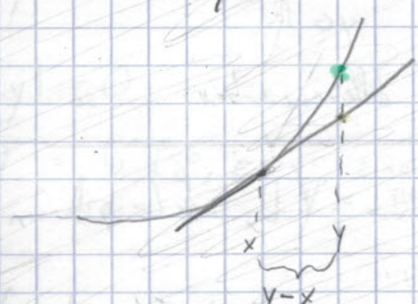
→ Es folgt nun:

1. Seien $f: D \subset \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion und X eine nichtleere konvexe Teilmenge von D. Dann gilt:

   1. f ist konvex auf X genau dann wenn für alle Punkte $x, y \in X$ die Ungleichung

   $$f(y) \geq f(x) + (y - x)^T \nabla f(x) \quad \text{gilt}$$

   ↳ d.h.



   2. Hat f sogar stetige zweite partielle Ableitungen nach allen Variablen so ist f auf X konvex genau dann, wenn für jeden Punkt $x \in X$ die Hesse-Matrix $\nabla^2 f(x)$ <u>positiv semidefinit</u> ist.

Und analog für Konkavität:

$\Rightarrow$ 1. f ist genau dann konkav wenn für alle Punkte $x, y \in X$ die Ungleichung:

$$f(y) \leq f(x) + (y-x)^T \nabla f(x) \quad \text{gilt}$$

2. Hat f stetige zweite partielle Ableitungen nach allen Variablen, so ist f auf X konkav genau dann, wenn für jeden Punkt $x \in X$ die Hesse-Matrix $\nabla^2 f(x)$ <u>negativ semidefinit</u> ist.

Notwendiges und hinreichende Optimalitätskriterium erster Ordnung:

Seien $D \subset \mathbb{R}^n$ eine offene konvexe Menge und $f: D \to \mathbb{R}$ eine konvexe, stetig differenzierbare Funktion. Dann gilt:

1. f besitzt in $x^0 \in D$ ein globales Minimum genau dann, wenn $\nabla f(x^0) = 0$
2. Die konkave Funktion $g = -f$ besitzt in $x^0 \in D$ ein globales Maximum genau dann, wenn $\nabla g(x^0) = 0$

<u>Implizite Funktionen:</u>

Gegeben sei ein Vektorraum: $F(x,y) = \begin{pmatrix} F_1(x,y) \\ \vdots \\ F_m(x,y) \end{pmatrix}$

mit $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$

Wenn x, y stetig differenzierbar, so gilt:

$$DF(x,y) = \left(D_x F(x,y) \, , \, D_y F(x,y) \right)$$

wobei

$$D_x F(x,y) = \begin{pmatrix} \nabla_x F_1(x,y) \\ \vdots \\ \nabla_x F_m(x,y) \end{pmatrix} \in m \times n \text{ Matrix}$$

$$D_y F(x,y) = \begin{pmatrix} \nabla_y F_1(x,y) \\ \vdots \\ \nabla_y F_m(x,y) \end{pmatrix} \in m \times m \text{ Matrix}$$

In dieser Sprache gilt dann somit in einem Punkt (x^0, y^0) mit $F(x^0, y^0) = 0$, dass:

$$F(x^0 + u, y^0 + v) = \underbrace{\cancel{F(x^0, y^0)}}_{\text{da } = 0} + D_x F(x^0, y^0) \cdot u + D_y F(x^0, y^0) \cdot v + o(u, v)$$

$$\Rightarrow \boxed{F(x^0 + u, y^0 + v) = D_x F(x^0, y^0) \cdot u + D_y F(x^0, y^0) \cdot v + o(u,v)}$$

! Satz über implizite Funktionen:

Seien $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^m$ offene Mengen und

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_m \end{pmatrix} : X \times Y \to \mathbb{R}^m$$

stetig differenzierbar. Sei $(x^0, y^0) \in X \times Y$ ein Punkt mit $F(x^0, y^0) = 0$, so dass die (m, m) Matrix $D_y F(x^0, y^0)$ invertierbar ist

Dann gelten folgende Aussagen:

1. Es gibt eine ε-Umgebung von x^0 in $\mathbb{R}^n$ und eine δ-Umgebung von y^0 in $\mathbb{R}^m$, so dass die Gleichung

$$F(x, y) = 0$$

für jedes $x \in U := B^0(x^0, \varepsilon)$ eine Lösung $y = g(x)$ hat, die in $V := B^0(y^0, \delta)$ eindeutig ist.

↳ Speziell gilt $y^0 = g(x^0)$

2. Die in (1) definierte Funktion $g: U \to V$ ist stetig auf U und differenzierbar in x^0 und es gilt:

$$Dg(x^0) = -\left(D_y F(x^0, y^0)\right)^{-1} D_x F(x^0, y^0)$$

! Der Satz über implizite Funktionen liefert also die Existenz von Umgebungen U und V und einer auf U erklärten Vektorfunktion $y = g(x)$, so dass für alle $(x, y) \in U \times V$ die Äquivalenz $F(x,y) = 0 \Leftrightarrow y = g(x)$ gilt.

↳ Wenn also Eigenschaft (1) erfüllt ist so sagt man, dass durch die Gleichung $F(x,y) = 0$ auf $U \times V$ implizit eine Funktion $y = g(x)$ definiert ist.

Unter den Voraussetzungen von Satz über implizite Funktionen existieren eine Umgebung $U' \subset U$ von x^0 und $V' \subset V$ von y^0 so dass $D_y(x,y)$ für alle $(x,y) \in U' \times V'$ invertierbar ist und die Funktion $y = g(x)$ folgende Eigenschaft hat:

$$D_g(x) = \left(-D_y F(x, g(x))\right)^{-1} \cdot D_x F(x, g(x)) \quad \forall x \in U'$$

Satz über Umkehrfunktionen:

Sei $\varepsilon^0 = f(y^0)$ und $Df(y^0)$ invertierbar.

→ Somit Satz implizite Funktion erfüllt da es auch gilt $F(\varepsilon, y) = 0$

↳ Somit $\varepsilon^0 = f(y^0)$ eindeutig.

↳ Es folgt also f bijektiv und somit:

$$\boxed{f^{-1}(\varepsilon) = y \quad \text{und} \quad \varepsilon = f(y)}$$

Ausserdem gilt: $\boxed{Df^{-1}(\varepsilon) = (Df(y))^{-1}}$

! Da $f^{-1}(\varepsilon) = y$

↳ und somit

$Df^{-1}(\varepsilon) = Dy$
$\quad = (Df(y))$

Optimierungsprobleme unter Nebenbedingungen.

Nichtlineare Optimierungsprobleme

hier beschäftigen wir uns mit der folgenden Standardaufgabe:

$$\min f(x) \quad \text{bezüglich} \quad g_i(x) \leq 0 \quad i = 1, \ldots, m$$
$$h_j(x) = 0 \quad j = 1, \ldots, r$$

Notation und Begriffe:

Wir schreiben unsere Minimierungsaufgabe kurz:

$$\min \{ f(x) \mid x \in M \} \quad \text{wobei} \quad M := \{ x \in \mathbb{R}^n \mid g(x) \leq 0 \ , \ h(x) = 0 \}$$

M heisst dabei zulässiger Bereich oder Restriktionsbereich.
Die Elemente von M heissen zulässige Punkte.

· Genügt $x^0 \in M$ nun die Bedingung:

$$f(x^0) \leq f(x) \quad \forall x \in M$$

→ so heisst x^0 globaler Minimalpunkt der Aufgabe.

Gilt dagegen für $x^0 \in M$: $\quad f(x^0) \leq f(x) \quad \forall x \in M \cap B^\circ(x^0, \varepsilon)$

→ so heisst x^0 lokaler Minimalpunkt.

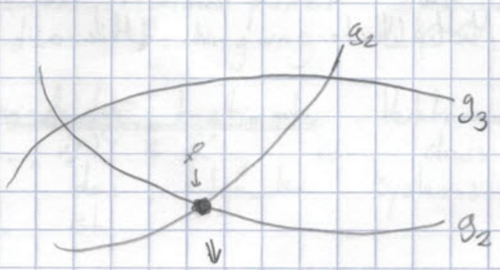
· Ist nun $x^0 \in M$, so heisst:

$$\boxed{I_{akt}(x^0) = \{ i \in \{1, \ldots, m\} \mid g_i(x^0) = 0 \}}$$

Indexmenge der in x^0 aktiven Ungleichungsrestriktionen und die entsprechende Restriktion $g_i(x) \leq 0$ heisst aktiv in x^0.

↳ Die Menge $I_{akt}(x^0)$ kann auch leer sein, dann sind alle Ungleichungen im Punkt x^0 strikt erfüllt (da sie aus M).

Für Verständnis:



hier sind $g_1 \wedge g_2$ aktiv, da mit x^0 $g_1 = 0$ und $g_2 = 0$ gilt.

! Gilt in $x^0 \in M$, dass die Menge der Gradienten:

$$\{ \nabla g_i(x^0),\ i \in I_{akt}(x^0) \} \cup \{ \nabla h_j(x^0),\ j = 1, \dots, r \}$$

linear unabhängig ist, dann sagt man, in x^0 sei die Bedingung LICQ (Linear Independence Constraint Qualification) erfüllt.

Bemerkung: Statt Minimierungs- könnte man auch Maximierungsprobleme betrachten, statt Ungleichungen in $\leq$-Form auch Ungleichungen in $\geq$-Form, denn es gilt stets:

$$\max \{ f(x) \mid x \in M \} \iff -\min \{ -f(x) \mid x \in M \}$$

sowie

$$g_i(x) \geq 0 \iff -g_i(x) \leq 0$$

Lagrange-Bedingungen: Notwendige Optimalitätsbedingungen erster Ordnung für Gleichungsrestriktionen:

Für stetig differenzierbare Funktionen $f: \mathbb{R}^n \to \mathbb{R}$ und $h_1, \dots, h_r: \mathbb{R}^n \to \mathbb{R}$ betrachte man das nichtlineare Optimierungsproblem:

$$\text{Min. } f(x) \quad \text{bezüglich} \quad h_j(x) = 0 \quad \text{mit } j = 1, \dots, r$$

Ist nun x^0 ein Minimalpunkt der Aufgabe, so dass die LICQ Bedingung erfüllt ist, d.h.

$\nabla h_j(x^0)$ lin. unabhängig sind.

$\Rightarrow$ dann existiert ein eindeutig bestimmter Vektor $v^0 = (v_1^0, \dots, v_r^0)^T \in \mathbb{R}^r$, so dass das Paar von Vektoren (x^0, v^0) den Bedingungen:

n-Gleichungen $\longrightarrow$ $\nabla f(x) + \sum_{j=1} v_j \nabla h_j(x) = 0$

und $h_j(x) = 0, \quad j = 1, \dots, r$

genügt.

v_j : Lagrange Multiplikator (in Mikro λ_j)

Sind alle Restriktionen $h_1, \ldots, h_r$ affin-linear, das heisst wenn sie für gewisse Vektoren $0 \neq A_{j.} \in \mathbb{R}^n$ und gewisse Zahlen $b_j \in \mathbb{R}$ von der Form:

$$h_j(x) = A_{j.} x - b_j \qquad j = 1, \ldots, r \qquad (\text{d.h. } \nabla h_j(x) = A_{j.}^T \; \forall x)$$

sind, dann muss die LICQ Bedingung nicht vorausgesetzt werden.

↳ Daraus folgt:

Für stetig differenzierbare Funktionen $f: \mathbb{R}^n \to \mathbb{R}$ sowie $A \in M(r,n)$ und $b \in \mathbb{R}^r$ betrachte man:

$$\text{Min. } f(x) \quad \text{bezüglich} \quad Ax - b = 0$$

Ist x^0 ein lokaler Minimalpunkt der Aufgabe, dann existiert mindestens ein Vektor $v = (v_1^0, \ldots, v_r^0) \in \mathbb{R}^r$, so dass das Paar von Vektoren (x^0, v^0) den Bedingungen:

$$\nabla f(x) + A^T v = 0 \qquad\qquad Ax - b = 0$$

genügt. ! Beachte: hier Eindeutigkeit von v^0 geht verloren.

Zurück zur Optimierungsproblem mit Ungleichungen:

Offenbar ist ein zulässiger Punkt x^0 der Aufgabe genau dann ein lokaler Minimalpunkt, wenn x^0 auch lokaler Minimalpunkt der Aufgabe:

$$\text{Min. } f(x) \quad \text{bezüglich} \quad g_i(x) \leq 0 \quad i \in I_{akt} \quad \wedge \quad h_j(x) = 0$$

$\Rightarrow$ d.h. im Punkt x^0 inaktive Restriktionen spielen bei Optimalitätskriterien keine Rolle

Kuhn-Tucker Bedingung für Ungleichungs- und Gleichungsrestriktionen:

Es sei x^0 ein lokaler Minimalpunkt unserer Optimierungsaufgabe und x^0 erfülle die Bedingung LICQ.

! Dann existieren eindeutig bestimmte Vektoren $u^0 = (u_1^0, \ldots, u_m^0)^T \in \mathbb{R}^m$ und $v^0 = (v_1^0, \ldots, v_r^0)^T \in \mathbb{R}^r$, so dass (x^0, u^0, v^0) ein Lösungsvektor des folgenden Systems von Gleichungen und Ungleichungen ist:

$$1.\quad \nabla f(x) + \sum_{j=1}^{r} v_j \nabla h_j(x) + \sum_{i=1}^{m} u_i \nabla g_i(x) = 0$$

$$2.\quad v_j \in \mathbb{R}, \quad h_j(x) = 0 \quad (j = 1, \dots, r)$$

$$3.\quad u_i \geq 0, \quad g_i(x) \leq 0, \quad u_i g_i(x) = 0 \quad (i = 1, \dots, m)$$

! d.h. bei aktiven Gleichungen muss $u_i > 0$ sein. Bei unaktiven Gleichungen muss $u_i = 0$ sein.

! Die Kuhn-Tucker-Bedingungen gelten auch wenn keine Gleichungen auftreten ⇒ 2. wegen

! Sind alle Restriktionen g_i und h_i affin-linear, kann man auf die Bedingung LICQ verzichten.

↳ Es geht dabei wiederum die Eindeutigkeit des optimalen Paars von Multiplikatoren (u^0, v^0) verloren.

Der Satz lautet nun:

Gegeben sind stetig diffbar Funktionen $f: \mathbb{R}^n \to \mathbb{R}$) sowie $A(r,n)$, $b \in \mathbb{R}^r$ und $B(m,n)$ und $d \in \mathbb{R}^m$. Man betrachte das Problem:

$$\text{Min}_x\ f(x) \quad \text{bezüglich} \quad Ax - b = 0, \quad Bx - d \leq 0$$

Ist x^0 ein lokaler Minimalpunkt, dann existiert mindestens ein Paar von Vektoren $(u^0, v^0) \in \mathbb{R}^n \times \mathbb{R}^r$, so dass (x^0, u^0, v^0) den Bedingungen:

$$1.\quad \nabla f(x) + A^T v + B^T u = 0$$

$$2.\quad Ax - b = 0$$

$$3.\quad Bx - d \leq 0, \quad u \geq 0, \quad u^T(Bx - d) = 0$$

genügt.

Kuhn - Tucker - Bedingungen bei konvexen Aufgaben:

Optimierungsaufgabe:

min. $f(x)$ bezüglich $g_i(x) \leq 0$ $i=1, \ldots, m$ $h_j(x) = 0$ $j=1,\ldots,r$

wobei:

$f: \mathbb{R}^n \to \mathbb{R}$ und $g_i: \mathbb{R}^n \to \mathbb{R}$ stetig diffbar und konvex.

$h_j: \mathbb{R}^n \to \mathbb{R}$ affin - linear

→ In diesem Fall heisst das Optimierungsproblem: konvexes Optimierungsproblem.

↳ Erfüllt nun hier ein Tripel von Vektoren (x^0, u^0, v^0) die Kuhn - Tucker - Bedingungen so ist x^0 ein globaler Minimalpunkt.

↳ hier keine LICQ Bedingung vorausgesetzt.

Setzt man LICQ im Punkt x^0 voraus oder sind alle Nebenbedingungen linear, dann sind die Kuhn-Tucker-Bedingungen ein notwendiges und hinreichendes Kritum dafür dass, x^0 globaler Minimalpunkt der konvexen Aufgabe ist.

Komplexe Zahlen

$$\boxed{i^2 = -1}$$

Eine komplexe Zahl ist definiert durch: $\boxed{z = x + iy}$

wobei x = Realteil y = Imaginärteil z = komplexe Zahl

Addition: $\boxed{(x+iy) + (x'+iy') = (x+x') + i(y+y')}$

Multiplikation: $\boxed{(x+iy) \cdot (x'+iy') = (x \cdot x' - y \cdot y') + i(x' \cdot y + x \cdot y')}$

Rechengesetze:

1. $(\mathbb{C}, +)$ ist eine kommutative Gruppe mit:

$0 + i0$ = Neutralelement

inverses Element: $-x - iy$

2. $(\mathbb{C} \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe mit:

$1 + i0$ = Neutralelement

inverses Element: $\frac{1}{x+iy} \cdot \frac{(x-iy)}{(x-iy)} = \frac{x-iy}{x^2+y^2}$

3. Es gilt das Distributivgesetz, d.h. für beliebige $z_1, z_2, z_3 \in \mathbb{C}$:

$$z_1 \cdot (z_2 + z_3) = (z_1 \cdot z_2) + (z_1 \cdot z_3)$$

Es gilt $0 \cdot z = 0$ und somit gelten auch Assoziativ- und Kommutativgesetz der Multiplikation sowie $1 \cdot z = z$ für alle $z \in \mathbb{C}$

Konjungiert komplexe Zahl und Absolutbetrag:

Konjungiert komplexe Zahl zu $z = x + iy$ ⊖ $\overline{z} = x - iy$

Betrag von z: $|z| = \sqrt{x^2 + y^2}$

Damit

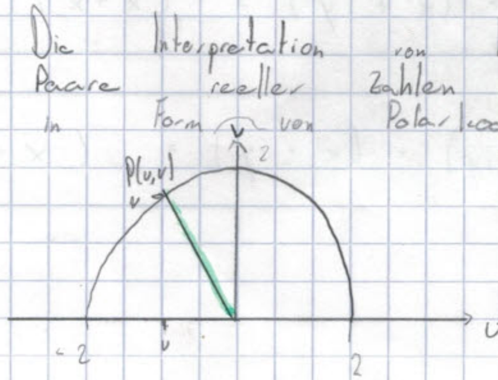
$$z \cdot \overline{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2$$

und

$$z^{-1} = \frac{1}{z} = \frac{\overline{z}}{|z|^2}$$

$$z^2 = -1 \Rightarrow 2 \text{ Lösungen: } -i \wedge i$$

Polarkoordinaten: Die Interpretation von komplexe Zahlen als Paare reeller Zahlen (siehe Def. z) legt es nahe, komplexe Zahlen in Form von Polarkoordinaten darzustellen:



— = Radius.

! Somit folgt unmittelbar aus Pythagoras:

$$\boxed{r^2 = u^2 + v^2}$$

Und aus Sinus und Kosinus Definitionen:

$$\boxed{u = \cos(\varphi) \cdot r} \quad \text{und} \quad \boxed{v = \sin(\varphi) \cdot r}$$

Somit folgt

$$\boxed{\begin{aligned} u^2 + v^2 &= \cos^2(\varphi)\, r^2 + \sin^2(\varphi) \cdot r^2 \\ &= r^2 \underbrace{\left(\cos^2(\varphi) + \sin^2(\varphi)\right)}_{1} = r^2 \end{aligned}}$$

Es folgt: Für jede komplexe Zahl $z = u + iv$ gilt:

$$\boxed{z = \cos(\varphi) \cdot r + i \sin(\varphi) \cdot r = r\left(\cos(\varphi) + i \sin(\varphi)\right)}$$

Und aus $r^2 = u^2 + v^2$ folgt $r = \sqrt{u^2 + v^2}$

$$\Rightarrow \boxed{r = |z| \geq 0}$$

Für die Multiplikation zweier komplexe Zahlen gilt nun:

$$\zeta = s\left(\cos(w) + i \sin(w)\right) \qquad z = r\left(\cos(\varphi) + i \sin(\varphi)\right)$$

$$\begin{aligned} \zeta \cdot z = r s \Big(&\cos(w)\cos(\varphi) - \sin(w)\sin(\varphi)\Big) \\ &+ i\Big(\cos(w)\sin(\varphi) + \sin(w)\cos(\varphi)\Big) \end{aligned}$$

$\Rightarrow$ <u>Aus Additionstheorem:</u> $\boxed{\zeta z = rs\left(\cos(w + \varphi) + i \sin(w + \varphi)\right)}$

Daraus folgt für Potenzen

$$\boxed{z^n = r^n\left(\cos(n\varphi) + i\sin(n\varphi)\right)}$$

$$\boxed{\begin{aligned} z^{-n} &= \frac{1}{z^n} = \frac{\overline{z}^n}{|z^n|^2} = \frac{r^n\left(\cos(n\varphi) - i\sin(n\varphi)\right)}{r^{2n}} \\ &= r^{-n}\left(\cos(-n\varphi) + i\sin(-n\varphi)\right) \end{aligned}}$$

Für n-te Wurzeln gilt hingegen:

hier gilt für $z = r\left(\cos(\varphi) + i\sin(\varphi)\right)$ und n $n \in \mathbb{N}$ und

$$\zeta = z^{1/n} \quad \text{mit} \quad \zeta = s\left(\cos(w) + i\sin(w)\right)$$

$$\Leftrightarrow \quad z = \zeta^n = s^n\left(\left(\cos(n w) + i\sin(n \cdot w)\right)\right)$$

Somit es wird gesucht eine Darstellung mit:

(i) $r = s^n$ (ii) $\cos(\varphi) = \cos(n \cdot w)$ $\sin(\varphi) = \sin(n \cdot w)$

(ii) Muss gelten $\forall$ $\varphi + 2k\pi$ da sin und cos periodisch.

Somit gilt für $z = \zeta$

$$\boxed{z^{\frac{1}{n}} = r^{\frac{1}{n}}\left(\cos\left(\frac{\varphi + 2k\pi}{n}\right) + i\sin\left(\frac{\varphi + 2k\pi}{n}\right)\right)}$$

! Analog bestimmt man n Lösungen von $z^n := z^m$

! Damit besitzt $z = r(\cos(\varphi) + i\sin(\varphi$ n-Verschiedene n-te Wurzeln.
↳ da $k = 0, \dots, n-1$

wichtig!

Fundamentalsätze der Algebra:

Eine Funktion des Typs: $P(z) = \sum_{k=0}^{n} a_k z^k$ $\quad a_k, z \in \mathbb{C} \wedge a_n \neq 0$

heisst komplexes Polynom der Ordnung n (auch n-ten Grades).

Fundamentalsatz der Algebra: Jedes komplexe Polynom der Ordnung $n \geq 1$ besitzt wenigstens eine Nullstelle.

↳ Folgerung: Jedes komplexes Polynom P der Ordnung $n \geq 1$ hat eine Darstellung der Form:

$$P(z) = a_n \cdot (z-\zeta_1)(z-\zeta_2)\cdots(z-\zeta_n)$$

wobei die ζ_i die (eventuell mehrfachen) Nullstellen von P sind.

Konjungiert komplexe Nullstellen bei reellen Koffizienten:

Wenn a_k = reell, dann gilt:

$$P = \sum (a_k z^k = 0 \quad \text{mit} \quad z = r(\cos(\varphi) + i\sin(\varphi))$$

und also $\sum a_k \cdot r^k \cos(k\varphi) + i \sum a_k r^k \sin(k\varphi) = 0$

somit muss gelten: $\sum a_k r^k \cos(k\varphi) = 0 = \sum a_k r^k (-k\varphi)$

und: $i \sum a_k r^k \sin(-k\varphi) = 0 = -i \sum a_k r^k (k\varphi)$

Es folgt somit:

Auch die Konjugierte von z: $\bar{z} = r(\cos(\varphi) - i\sin(\varphi))$

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

ist eine Nullstelle von P, sofern z eine Nullstelle ist.

$\boxed{\text{Potenzreihen und Eulersche Formel.}}$

Für reelle φ folgt aus den Potenzreihendarstellungen von e^z, $\cos(z)$ und $\sin(z)$ die berühmte:

$$\boxed{\text{Eulersche Formel} := \quad e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)}$$

Analog gilt die:

$$\boxed{\text{Eulersche Formel in komplexer Form:} \quad e^{iz} = \cos(z) + i\sin(z)}$$

$\Rightarrow$ Daraus folgt sofort mit $\varphi = \pi$:

$$e^{i\pi} = \cos(\pi) - i\sin(\pi) = -1 + 0$$

$$\boxed{e^{i\pi} + 1 = 0} \quad !$$

Aus eulersche Formel folgt ausserdem:

$$\text{mit} \quad z = r\underbrace{(\cos(\varphi) + i\sin(\varphi))}_{e^{i\varphi}} = \boxed{r \cdot e^{i\varphi}}$$

Ausserdem gilt die Periodizität:

$$\boxed{\begin{aligned} e^{i(\varphi + 2k\pi)} &= \cos(\varphi + 2k\pi) + i\sin(\varphi + 2k\pi) \\ &= \cos(\varphi) + i\sin(\varphi) = e^{i\varphi} \end{aligned}}$$

Ausserdem gilt:

$$e^z = e^{x+iy} = e^x \cdot e^{iy} = e^x \cdot e^{i(y+2k\pi)} = e^{(x+iy)+2k\pi i}$$

$\Rightarrow$ d.h. es gilt $\boxed{e^z = e^{z+2k\pi i}}$

Es gilt Periodizität auch für die Umkehrfunktion:

$$z = \ln \xi \quad , \quad \text{wobei} \quad \xi = e^z$$

denn: $$\xi = e^{x+iy} = e^x (\cos(y) + i \sin(y))$$
$$= e^x (\cos(y + 2k\pi) + i \sin(y + 2k\pi))$$
$$= e^x e^{i(y+2k\pi)} = e^{z+2k\pi i}$$

also folgt auch $$\boxed{z + 2k\pi i = \ln \xi}$$

Eigenschaften von cos(z) und sin(z):

$$e^{-iz} = \cos(z) - \sin(z)$$

$$e^{iz} = \cos(z) + \sin(z)$$

$$\Rightarrow \boxed{\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \qquad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}}$$

folgt $\Rightarrow$ unmittelbar.

$$\boxed{\cos(-z) = \cos(z) \qquad \sin(-z) = -\sin(z)}$$

Andere einfache Beweise die man durchführen kann sind:

$$\boxed{\begin{aligned} &\sin^2(z) + \cos^2(z) = 1 \\ &\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2 \\ &\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2 \end{aligned}}$$

Ausserdem gilt: $$(e^z)^n = e^z \cdot \ldots \cdot e^z = e^{nz}$$

! $$\boxed{(e^{ix})^n = e^{inx} \quad \text{liefert dann} \quad (\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)}$$