Mathespiele
Grundlagen

Offene Fragen

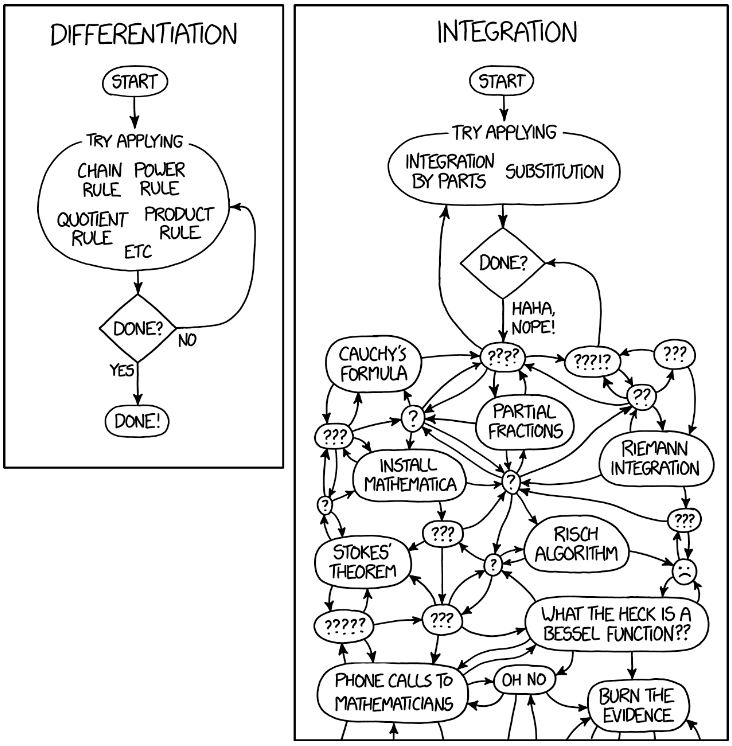
Inhalt: Algebraische Strukturen, Determinanten und Eigenwerte, Grenzwerte, Integration, normierte Räume, Skalarprodukte, Stetigkeit und Differenzierbarkeit, Vektorräume und lineare Abbildungen.

Lernziel: Auf offen gestellte Wissensfragen flüssig, ausführlich und mit geeignetem Fachvokabular zu antworten, ohne dabei vom Thema abzuweichen.

Spielbeschreibung: Dieses Spiel simuliert die mündliche Prüfungssituation. Ein Spieler ist der Prüfer, der andere der Prüfling. Bestenfalls gibt es noch einen Beobachter, der das Wissen und Verhalten der beiden Personen analysiert und seine Beobachtungen den anderen beiden schildert. In Zweierteams wird die Beobachterrolle vom Prüfer übernommen.
Nachdem der Prüfer eine offene Frage gestellt hat, präsentiert der Prüfling sein gesamtes Wissen zum Thema, was durchaus einige Minuten beansprucht. Der Prüfer achtet dabei, dass die Antworten des Prüflings korrekt sind. Bei Fehlern und Abweichung vom Thema unterbricht der Prüfer den Prüfling und unterstützt ihn mit leitenden Fragen und Tipps.
Sobald eine Karte vollständig behandelt wurde, ist es sehr wichtig, die Präsentation zu analysieren um daraus zu lernen. Besonders viel Wert wird dabei auf korrekten Inhalt, verständliche Sprache und Wortwahl, Verhalten sowie Körperhaltung und Stimme gelegt. Zuerst reflektiert der Prüfling laut sein Verhalten, danach erklärt der Beobachter bzw. der Prüfer, wie der Prüfling auf ihn wirkte. Falls Wissenslücken auftauchten, wird das Thema mit dem Skript nochmals gemeinsam erarbeitet.

Was wissen Sie über die Taylorentwicklung?

Nennen Sie alle Ihnen bekannten Matrixklassen, deren Definition und ihre Besonderheiten.



xkcd.com/2117

Feedback, Korrekturen und Ideen bitte an
`philipp.wittmann@tum.de`
oder `maxim.baumgaertel@tum.de`

A, B sind jeweils Elemente der Klasse.

Klasse	Definition
GL_n	$\exists A^{-1}\colon AA^{-1} = I$
SL_n	$A \in \mathrm{GL}_n \wedge \det A = 1$
U_n	$A\overline{A}^T = I (\Rightarrow A \in \mathrm{GL}_n)$
SU_n	$A \in \mathrm{U}_n \wedge \det A = 1 (\Rightarrow A \in \mathrm{SL}_n)$
normal	$\overline{A}^T A = A\overline{A}^T$
diagonalisierbar	$\exists$ ähnliche Diagonalmatrix
selbstadjungiert	$\overline{A}^T = A$ (hermitsch) bzw. $A^T = A$ (symm)
antisymm	$\overline{A}^T = -A$ (bzw. $A^T = -A$)
diagonal	alles 0 außer Diagonaleinträge

statt unitär kann hier auch orthogonal betrachtet werden (O_n statt U_n und SO_n statt SU_n)

Bei der Taylorentwicklung geht es darum, eine k-mal stetig differenzierbare Funktion f lokal um einen **Entwicklungspunkt** x mit einem Polynom $T_k f$ vom Grad $\leq k$ **anzunähern**. Dabei soll der **Fehler**

$$r(h) = f(x+h) - T_k f(x+h)$$

für $h \to 0$ schneller gegen 0 gehen als h^k es tut, in Zeichen

$$\lim_{h\to 0} \frac{r(h)}{h^k} = 0$$

oder $r(h) = o(h^k)$. Für $k = 1$ wissen wir bereits, dass man dies mit der **Ableitung** von f erreichen kann, genauer:

$$T_1 f(x+h) = f(x) + f'(x)h$$

Formel für $T_k f$ im Fall, dass f eine **Potenzreihe** ist, begründen: Koeffizienten c_m von f sind gerade

$$c_m = \frac{f^{(m)}(x)}{m!}.$$

Satz von Taylor (mit Formel für das Restglied) formulieren, Warnung, dass $T_k f \to f$ nicht immer gilt, aussprechen.

Auf den mehrdimensionalen Fall eingehen, etwa Formel für $T_2 f$ aufschreiben:

$$T_2 f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T \nabla^2 f(x) h.$$

Was können Sie über die Hessematrix sagen?

Was können Sie über Differenzierbarkeit sagen?

Was wissen Sie über Determinanten?

Was wissen Sie über
(absolute) Konvergenz von Reihen?

Im Eindimensionalen beschreibt die Ableitung $f'(x)$ die Steigung der Tangente am Graphen von f im Punkt x. Geht man über zu Skalarfelder im Mehrdimensionalen, so definiert man die **Richtungsableitung** von f in Richtung $v \neq 0$ als $\partial_v f(x) = g'(0)$, wobei $g(t) = f(x+tv)$. Existieren $\partial_1 f, \ldots, \partial_n f$, so heißt f **partiell differenzierbar** und der **Gradient** $\nabla f = (\partial_1 f, \ldots, \partial_n f)$ zeigt in Richtung des steilsten Anstiegs (hier könnte man den Graphen mit Tangenten skizzieren).

Die Ableitung $\dot{\gamma}(t)$ einer Kurve γ zur Zeit t kann man als Geschwindigkeitsvektor interpretieren. Dieser liegt tangential zur Spur der Kurve und zeigt in die „Durchlaufrichtung" von γ (hier könnte man die Spur mit Tangentialvektoren zeichnen).

Ganz allgemein ist die **totale Ableitung** einer Funktion $f\colon \mathbb{R}^n \to \mathbb{R}^m$ im Punkt x definiert als eine lineare Abbildung $\mathrm{D}f(x)\colon \mathbb{R}^n \to \mathbb{R}^m$, die f in der Nähe von x approximiert. Genauer gesagt erfüllt sie

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - \mathrm{D}f(x)(h)\|}{\|h\|} = 0$$

und ist dadurch eindeutig bestimmt. Falls ein Skalarfeld f bzw. eine Kurve γ total differenzierbar ist, gilt $\mathrm{D}f = (\nabla f)^T$ bzw. $\mathrm{D}\gamma = \dot{\gamma}$.

Aus stetiger Differenzierbarkeit folgt totale, woraus wiederum partielle (und außerdem Stetigkeit) folgt. Alle anderen Implikationen sind im Allgemeinen nicht gültig.

(Danach: Auf wichtige Sätze eingehen, etwa den **Mittelwertsatz**.)

Die Hessematrix $\nabla^2 f$ eines Skalarfelds $f\colon \mathbb{R}^n \to \mathbb{R}$ ist die Jacobi-Matrix des Gradienten ∇f (falls alles existiert) und beschreibt geometrisch die **Krümmung des Graphen** von f. Ist f zweimal stetig differenzierbar, so folgt aus dem **Satz von Schwarz** ($\partial_i\partial_j f = \partial_j\partial_i f$ für alle $i,j = 1,\ldots,n$), dass $\nabla^2 f$ **symmetrisch** ist. Insbesondere ist $\nabla^2 f$ genau dann positiv/negativ definit, wenn ihre Eigenwerte größer bzw. kleiner 0 sind (entsprechendes, wenn man definit durch semidefinit ersetzt).

Mithilfe der Hessematrix und der Taylorentwicklung kann man **hinreichende Bedingungen** für Extrema herleiten: Ist etwa für einen Punkt x im Definitionsgebiet $\nabla^2 f(x)$ positiv definit und $\nabla f(x) = 0$, so folgt

$$f(x+h) = f(x) + \frac{1}{2}h^T \nabla^2 f(x) h + r(h)$$

und unter Verwendung der positiven Definitheit und

$$\frac{r(h)}{\|h\|^2} \xrightarrow{h\to 0} 0$$

folgt schließlich, dass es ein $\delta > 0$ gibt mit $f(x+h) > f(x)$ für alle h mit $\|h\| < \delta$. Der Punkt x ist also ein striktes lokales Minimum. Für Maxima kann man ähnliche Argumente verwenden, wenn man $\nabla^2 f(x)$ negativ definit fordert.

Eine Reihe

$$\sum_{k=0}^{\infty} a_k \tag{$*$}$$

heißt konvergent, wenn die Folge (s_n) der **Partialsummen** $s_n = a_1 + \cdots + a_n$ konvergiert, und absolut konvergent, wenn die Folge (S_n) mit $S_n = |a_1| + \cdots + |a_n|$ konvergiert (das ist genau dann der Fall, wenn die monotone Folge (S_n) beschränkt ist). Aus absoluter Konvergenz folgt Konvergenz und es gilt die Ungleichung

$$\left|\sum_{k=0}^{\infty} a_k\right| \leq \sum_{k=0}^{\infty} |a_k|.$$

Notwendig für die Konvergenz von ($*$) ist, dass $a_k \to 0$. Das **Cauchy-Kriterium** und das **Leibniz-Kriterium** liefern eine äquivalente bzw. eine hinreichende Bedingung dafür, ob eine Reihe konvergiert. Mit ihnen kann man etwa schließen, dass (a) die harmonische Reihe divergiert und (b) die alternierende harmonische Reihe konvergiert.

Das **Majorantenkriterium**, das **Quotientenkriterium** sowie das **Wurzelkriterium** liefern Bedingungen dafür, wann eine Reihe absolut konvergiert. Beispielsweise kann man mit dem Quotientenkriterium zeigen, dass die Exponentialreihe einen unendlichen Konvergenzradius hat.

Sind zwei Reihen $\mathbf{a}$ und $\mathbf{b}$ absolut konvergent, so sind auch $\mathbf{a} + \lambda\mathbf{b}$ (gliederweise Addition und skalare Multiplikation) sowie das **Cauchy-Produkt** $\mathbf{a} \cdot \mathbf{b}$ absolut konvergent. Damit bildet die Menge aller absolut konvergenten Reihen einen Vektorraum.

(Ab hier: $A, B \in \mathbb{K}^{n\times n}$ beliebig)

Die Determinante ist ein **Polynom** $\mathbb{K}^{n\times n} \to \mathbb{K}$. Damit ist sie unendlich oft stetig differenzierbar (also auch stetig).

Der **Determinantenmultiplikationssatz** besagt $\det(AB) = \det(A)\det(B)$. Außerdem gilt

$$\det(A) = \det(A^T). \tag{:-)}$$

Geometrisch beschreibt die Determinante das ***n*-dimensionale Volumen** des von den Zeilen- oder Spaltenvektoren aufgespannten Spats, siehe (:-). Da die Spalten von A den Bildern der Basisvektoren entsprechen bedeutet eine Determinante von 0, dass das Bild der Basisvektoren ein n-dimensionales Volumen von Null hat und damit ebendiese Bildvektoren einen Unterraum aufspannen, der eine kleinere Dimension als n hat. Da eine lineare Abbildung von einem höher in einen niederdimensionalen Raum nicht injektiv sein kann ist sie auch nicht bijektiv und daraus folgt die berühmte Aussage

$$A \text{ invertierbar} \iff \det(A) \neq 0$$

Mithilfe der Determinante lassen sich außerdem **Eigenwerte bestimmen** (Eigenwerte sind die $\lambda \in \mathbb{K}$ für die die lineare Abbildung $(\lambda I - A)$ nicht invertierbar ist). Deshalb lassen sich durch bestimmen der Nullstellen des charakteristischen Polynoms $\chi_A(\lambda) = \det(\lambda I - A)$ die Eigenwerte bestimmen.

(…danach auf die Regeln zur Bestimmung der Determinante eingehen)

Was wissen Sie über die gleichmäßige Konvergenz von Funktionenfolgen?

Was können Sie mir zur Ähnlichkeit von Matrizen sagen?

Erläutern sie den Zusammenhang zwischen Matrizen und linearen Abbildungen.

Was können Sie über die Konvergenz und den Grenzwert der Potenzreihe

$$\sum_{k=0}^{\infty} X^k \qquad (*)$$

sagen?

(X in $\mathbb{R}$ oder $\mathbb{C}$ oder $\mathbb{R}^{n\times n}$...)

(Ab hier: $A, B \in \mathbb{K}^{n\times n}$)
Eine Matrix A heißt ähnlich zu B genau dann, falls $\exists S \in \mathrm{GL}_n(\mathbb{K})\colon A = S^{-1}BS$. Es bildet damit eine Äquivalenzrelation geschrieben $A \sim B$ (Ab hier: S dementsprechend.)
Wir können zeigen, dass ähnliche Matrizen denselben Rang haben.
Zweitens können wir zeigen, dass ähnliche Matrizen dieselben Eigenwerte haben:

$$A \text{ ähnlich zu } B \iff \exists S \in \mathrm{GL}_n(\mathbb{K})\colon A = S^{-1}BS$$
$$\chi_A(\lambda) = \det(\lambda I - S^{-1}BS) = \det(S^{-1}(\lambda I - B)S) = \chi_B(\lambda)$$
$$\iff \chi_A = \chi_B \implies \text{Eigenwerte von } A = \text{Eigenwerte von } B$$

In der Äquivalenzklasse von A versuchen wir einen möglichst guten Vertreter B zu wählen. Sinnvoll bedeutet z.B. dass man B^n mit wenig Aufwand berechnen kann, denn damit lässt sich auch $A^n = (S^{-1}BS)^n = S^{-1}B^nS$ gut berechnen.
Die einfachste Berechnung einer Matrixpotenz funktioniert bei einer Diagonalmatrix. Eine Matrix A heißt **diagonalisierbar**, falls $A \sim D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ Aus vorher gezeigtem folgt, dass $\lambda_1, \ldots, \lambda_n$ die Eigenwerten von A sind.
Aber nicht alle Matrizen sind diagonalisierbar. Dafür ist aber jede Matrix über $\mathbb{K} = \mathbb{C}$ ähnlich zu einer Jordannormalform.
Der Spektralsatz besagt, dass es für $\mathbb{K} = \mathbb{C}$ eine Diagonalisierung $U \in \mathrm{U}_n\colon A = \overline{U}^T DU$ mit $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ aus Eigenwerten von A gibt, falls A normal ($A\overline{A}^T = \overline{A}^T A$) ist. Falls $\mathbb{K} = \mathbb{R}$ gilt ein ähnliches Resultat mit O statt U und D eine Matrix mit Drehkästchen.

Eine Folge $(f_n\colon D \to \mathbb{R})_{n\in\mathbb{N}}$ heißt gleichmäßig konvergent gegen eine Funktion $f\colon D \to \mathbb{R}$, wenn

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{N} \quad \forall\, n \geq N \quad \forall\, x \in D\colon\ |f_n(x) - f(x)| < \varepsilon.$$

Gleichbedeutend damit ist die Bedingung (das **Supremum** ist die kleinste obere Schranke!)

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{N} \quad \forall\, n \geq N\colon\ \|f_n - f\|_\infty \leq \varepsilon.$$

Sind zum Beispiel alle beteiligten Funktionen Regelfunktionen, so ist gleichmäßige Konvergenz gerade die Konvergenz im normierten Raum $(R[a,b], \|\cdot\|_\infty)$. Das Regelintegral

$$I\colon R[a,b] \to \mathbb{R},\ I(f) = \int_a^b f(x)dx$$

ist eine bezüglich dieser Norm stetige Abbildung, d.h. es gilt $I(f) = \lim I(f_n)$, wenn $f_n \to f$ gleichmäßig (nachprüfen muss man die Stetigkeit von I nur für $f = 0$, aus der Linearität folgt dann die Stetigkeit überall).
Wenn alle f_n stetig sind und $f_n \to f$ gleichmäßig, so ist auch f stetig. Diese Schlussfolgerung ist im Allgemeinen nicht zulässig, wenn $f_n \to f$ nur punktweise, wie man am Standardbeispiel $f_n\colon [0,1] \to \mathbb{R},\ f_n(x) = x^n$ leicht erkennt.
Aus dem Hauptsatz folgt außerdem: Gelten (a) $f_n \in C[a,b]$ differenzierbar in (a,b), (b) f_n' stetig fortsetzbar auf $[a,b]$, (c) $f_n' \to g$ gleichmäßig und (d) $(f_n(x_0))$ konvergent für ein $x_0 \in [a,b]$, so gibt es ein $f \in C[a,b]$ mit $f_n \to f$ gleichmäßig und $f_n' = g$ in (a,b).

In $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$: Für $z \in \mathbb{K}$ mit $|z| \neq 1$ gilt

$$S_n = 1 + |z| + \cdots + |z|^n = \frac{1 - |z|^{n+1}}{1 - |z|},$$

und für $|z| < 1$ folgt hieraus

$$S_n \xrightarrow{n\to\infty} \frac{1}{1-|z|}. \qquad (**)$$

Die **geometrische Reihe** konvergiert also absolut für alle z mit $|z| < 1$ (das hätte man auch mit dem Quotientenkriterium sehen können) und ist gleich der rechten Seite in $(**)$. Tatsächlich ist 1 auch der Konvergenzradius der geometrischen Reihe, da (z^k) für $|z| \geq 1$ keine Nullfolge ist.
In $\mathbb{R}^{n\times n}$: Um von Konvergenz sprechen zu können, benötigt man eine Norm, zum Beispiel die durch eine Norm $\|\cdot\|$ auf dem $\mathbb{R}^n$ induzierte **Operatornorm**: $\|X\| = \sup_{\|v\|=1} \|Xv\|$. Für sie gilt $\|XY\| \leq \|X\|\|Y\|$, woraus per Induktion $\|X^k\| \leq \|X\|^k$ folgt. Insbesondere ist die Reihe $(*)$ (sie heißt jetzt **Neumannsche Reihe**) absolut konvergent, wenn $\|X\| < 1$. In diesem Fall gilt wie oben

$$(I - X)\sum_{k=0}^{n} X^k = I - X^{n+1} \xrightarrow{n\to\infty} I,$$

also ist $I - X$ invertierbar mit

$$(I - X)^{-1} = \sum_{k=0}^{\infty} X^k.$$

Ab hier: V, W K-Vektorräume, $\dim(V) = n$, $\dim(W) = m$, $B = \{b_1, \ldots, b_n\}$ Basis von V und $C = \{c_1, \ldots, c_m\}$ Basis von W
Da jeder Vektorraum eine Basis hat, lässt sich eine lineare Abbildung ϕ (linear: $\forall a \in K\ \forall x, y \in V\colon\ \phi(ax+y) = a\phi(x) + \phi(y)$) als Linearkombination der Bilder der Basisvektoren schreiben. Damit gilt also für eine Abbildung $\phi\colon V \to W$

$$\phi\left(\sum_{k=1}^{n} \lambda_k b_k\right) = \sum_{k=1}^{n} \lambda_k \phi(b_k) = \sum_{k=1}^{n} \lambda_k \sum_{j=1}^{m} \mu_{j,k} c_j,$$

wenn wir $\phi(b_k) = \mu_{1,k}c_1 + \cdots + \mu_{m,k}c_m$ schreiben. Mithilfe der Definition von Matrizen lässt sich die Notation erheblich erleichtern. Mit der Darstellungsmatrix

$$D_\phi = \begin{pmatrix} \mu_{1,1} & \cdots & \mu_{n,1} \\ \vdots & & \vdots \\ \mu_{m,1} & \ldots & \mu_{m,n} \end{pmatrix}$$

und der Definition einer Matrix-Vektor-Multiplikation kann man dann jede lineare Abbildung eindeutig als Matrix schreiben ($\forall v \in V\colon\ \phi(v) = D_\phi \cdot v$).
Da die Komposition von linearen Abbildungen auch wieder eine lineare Abbildung ist können wir ähnlich dazu ein Matrixprodukt definieren sodass das Produkt dieser Matrizen der Verkettung entspricht.
Wählen wir für einen der Räume eine andere Basis, z.B. B' für V, so können wir das diese Basisvektoren durch eine Linearkombination aus den bisherigen in B darstellen. Damit haben wir eine bijektive lineare Abbildung $V \to V$, die wir mithilfe einer Matrix darstellen können. Wir nennen eine solche Matrix die **Basiswechselmatrix** $S_{B,B'}$. Mit $S_{B',B} \cdot D_\phi \cdot S_{B,B'}$ wird …

Was haben denn Integrieren und Differenzieren miteinander zu tun?

Nennen Sie Beispiele von Vektorräumen und deren Dimension und geben Sie wenn möglich eine Norm und eine Basis an.

Erklären und vergleichen Sie die verschiedenen Stetigkeitsbegriffe, die Sie kennen.

Sei f eine stetige oder stetig differenzierbare Funktion $\mathbb{R}^n \to \mathbb{R}^n$ mit $n \geq 1$. Was kann man denn (unter welchen Voraussetzungen) über die Inverse f^{-1} sagen?

Raum	dim	Basis	Norm
K^n	n	$\{e_1,\dots,e_n\}$ mit $e_k := (0,\dots,0,1,0\dots,0)^T$ mit der 1 an der k-ten Stelle	p-Norm
$c_0,\ c,\ \ell^p$	∞	$\{e_k \mid k\in\mathbb{N}\}$	p-Norm
$K[x]$	∞	$\{1,x,x^2,\dots\}$	$\sum_{k=0}^{\infty}\lvert a_k\rvert$
$\mathscr{C}^n(\Omega)$	∞	–	
$L^p(\Omega)$	∞	–	p-Norm
$T(\Omega),\ R(\Omega)$	∞	–	∞-Norm

$c,\ c_0$ Raum der konvergenten und Null-Folgen

ℓ^p, L^p Raum der Folgen/Funktionen, für die die Reihe bzw. das Integral in der p-Norm existiert. (ℓ^1 ist der Raum der absolut konvergenten Folgen)

ℓ^∞, L^∞ Raum der beschränkten Folgen/Funktionen

$T,\ R$ Raum der Treppen- und Regelfunktionen
Außerdem sind $c_0 \subset c \subset \ell^1 \subset l^2 \subset \cdots \subset l^\infty$ Untervektorräume (mit $\|\cdot\|_\infty$ der einzigen Norm für alle).

$\mathscr{C}^\infty \subset \cdots \subset \mathscr{C}^2 \subset \mathscr{C}^1 \subset \mathscr{C}^0 \subset R \subset L^1 \subset \cdots \subset L^\infty$ sind ebenso Untervektorräume (mit $\|\cdot\|_\infty$ der einzigen Norm für alle). T lässt sich nicht in die Kette einordnen, aber $T \subset L^1$ ist ein Unterraum.

Formal: Die Verbindung zwischen Integration und Differentiation beschreibt der **Hauptsatz der Integral und Differentialrechnung**: Ist $f:[a,b]\to\mathbb{R}$ stetig und F eine Stammfunktion von f (d.h. $F\in C[a,b]$ differenzierbar in (a,b) mit $F'=f$), so gilt

$$\int_a^b f(x)dx = F(b)-F(a). \qquad (*)$$

Der Schlüssel, warum dies gilt, liegt im **Mittelwertsatz für Integrale**. Wendet man ihn auf den Differenzenquotienten der Funktion $F_a(x)=\int_a^x f(x)dx$ an, erhält man nach einigen Umformungen und Grenzwertübergang $F'(x)=f(x)$. Ferner unterscheiden sich alle Stammfunktionen F von f nur um eine Konstante von F_a, also gilt $(*)$ nach Definition von F_a.

Intuitiv: Sei $f:[a,b]\to\mathbb{R}$ eine Regelfunktion und $F:[a,b]\to\mathbb{R}$ eine Funktion mit $F(x)=$ "Fläche unterhalb des Graphen von f zwischen a und x". Wenn man nun die rechte Flächengrenze x um einen sehr kleinen Wert dx verschiebt, so verändert sich dadurch die Fläche selbst ungefähr um $dF=f(x)dx$ (das ist das neue Rechteck mit Höhe h und Breite dx, was durch Verschieben der rechten Flächengrenze hinzugekommen ist). Hieraus folgt, dass

$$F'(x)=\frac{\text{“Änderung in den Funktionswerten”}}{\text{“Änderung in den Argumenten”}}=\frac{dF}{dx}=f(x).$$

Im Fall, dass f linear ist, ist f genau dann invertierbar, wenn etwa $\ker f=\{0\}$ (hierzu gibt es noch viele andere äquivalente Bedingungen). Die Inverse kann man durch Lösen von n Gleichungssystemen finden.

Betrachten wir nun den Fall $n=1$: Ist $f\colon [a,b]\to\mathbb{R}$ stetig und streng monoton, so ist $f\colon [a,b]\to[f(a),f(b)]$ bijektiv und f^{-1} ist ebenfalls stetig. Ist f zusätzlich differenzierbar in einem Punkt $x_0\in(a,b)$ mit $f'(x_0)\neq 0$, so ist auch f^{-1} differenzierbar in $y_0=f(x_0)$ und es gilt $(f^{-1})'(y_0)=f'(x_0)^{-1}$. Man kann also grob zusammengefasst sagen, dass unter bestimmten Voraussetzungen f^{-1} den "Glattheitsgrad" von f übernimmt.

Obiges Resultat kann man auch "lokal" formulieren: Sei $f\colon \mathbb{R}\to\mathbb{R}$ stetig differenzierbar, $x_0\in\mathbb{R}$ mit $f'(x_0)\neq 0$. Dann ist $f'(x_0)>0$ oder $f'(x_0)<0$ und aus der Stetigkeit von f' folgt, dass es $\delta>0$ gibt mit $f'(x)>0$ oder $f'(x)<0$ für alle $x\in(x_0-\delta,x_0+\delta)=:I$. f ist also streng monoton und stetig in $\bar{I}$ sowie differenzierbar in I, und aus Vorherigem folgt, dass auch $f^{-1}:f(I)\to I$ differenzierbar in I ist mit $(f^{-1})'(f(x))=f'(x)^{-1}$ für alle $x\in I$ (die Ableitung ist also sogar stetig).

Diese lokale Version lässt sich auch ins Mehrdimensionale übertragen, indem man $f'(x_0)\neq 0$ durch $\det \mathrm{D}f(x_0)\neq 0$ ersetzt. Die Intervalle $I\ni x_0$ und $f(I)\ni f(x_0)$ werden zu offene Mengen U und V um x_0 und $f(x_0)$. Man erhält den **Satz über inverse Funktionen**, für dessen Beweis man den Fixpunktsatz von Banach benötigt.

Sei $f\colon D\to\mathbb{R}$ eine Funktion. Sie ist punktweise stetig in D, wenn sie das **ε-δ-Kriterium**

$$\forall\, a\in D \quad \forall\, \varepsilon>0 \quad \exists\, \delta>0 \quad \forall\, x\in D: \qquad (*)$$
$$|x-a|<\delta \;\Rightarrow\; |f(x)-f(a)|<\varepsilon$$

erfüllt. Will man, dass f gleichmäßig stetig ist, so muss der erste Quantor ($\forall\ a\in D$) in $(*)$ ganz nach hinten verschoben werden. Das bedeutet, dass zu gegebenem ε ein für alle Punkte a gemeinsames δ gewählt werden kann, sodass das obige Kriterium erfüllt ist, oder anschaulicher: Man kann ein ε-δ-Rechteck finden, sodass der Graph von f dieses Rechteck immer an den Seiten und nicht oben oder unten verlässt, und zwar unabhängig davon, an welchem Punkt a man dieses Rechteck legt.

f heißt Lipschitz-stetig, wenn es eine **Lipschitz-Konstante** $L>0$ gibt mit

$$|f(x)-f(a)|\le L|x-a| \quad \forall\, x,a\in D.$$

Aus Lipschitz-Stetigkeit folgt gleichmäßige Stetigkeit (man wählt $\delta=\varepsilon/L$ und beachtet die Reihenfolge der Quantoren), und aus gleichmäßiger Stetigkeit folgt punktweise Stetigkeit.

Beispielsweise ist eine differenzierbare Funktion $f:\mathbb{R}\to\mathbb{R}$ mit beschränkter Ableitung Lipschitz-stetig: Seien $x,a\in D$ beliebig, $x\neq a$. Setze $L=\sup|f'(x)|$. Nach dem Mittelwertsatz der Differentialrechnung gilt

$$\left|\frac{f(x)-f(a)}{x-a}\right|=|f'(x)|\le L,$$

daraus folgt die Behauptung.

Welche Möglichkeiten kennen Sie um festzustellen, dass eine Matrix invertierbar ist?

Was wissen Sie über Fixpunkte?

Was wissen Sie über Kompaktheit?

Was wissen Sie über Skalarprodukte?

Sei (X,d) ein metrischer Raum. Ein $x \in X$ heißt Fixpunkt der Funktion $f\colon X \to X$, wenn $f(x) = x$. Man kann Fixpunkte mit sogenannten **Fixpunktiterationen** finden: Sei $x_0 \in X$ und $x_{n+1} := T(x_n)$ für $n \in \mathbb{N}$. Falls f stetig ist und die so definierte Folge (x_n) konvergiert, dann ist ihr Grenzwert x ein Fixpunkt von f, denn es gilt

$$\begin{aligned} d(f(x),x) &\leq d(f(x),x_{n+1}) + d(x_{n+1},x) \\ &= d(f(x),f(x_n)) + d(x_{n+1},x) \xrightarrow{n\to\infty} 0. \end{aligned}$$

Es bleibt die Frage, wann eine solche Fixpunktiteration konvergiert und ob der Fixpunkt einer Funktion eindeutig bestimmt ist. Der **Fixpunktsatz von Banach** (es gibt noch viele weitere Sätze zu diesem Thema, z.B. der Satz von Brouwer) beantwortet diese Frage:

Ist (X,d) ein **vollständiger** metrischer Raum und $f\colon X \to X$ eine **Kontraktion**, so hat f genau einen Fixpunkt und eine Fixpunktiteration mit beliebigem Anfangswert $x_0 \in X$ konvergiert gegen diesen.

Man benötigt den Fixpunktsatz von Banach etwa, um den Satz über inverse Funktionen zu beweisen.

(Ab hier sei $A \in \mathbb{K}^{n\times n}$)

Es gelten die folgenden Äquivalenzen:

- $\exists B, C \in \mathbb{K}^{n\times n}\colon AB = BC = I$ (Links- und Rechtsinverse)
- $\exists A^{-1} \in \mathbb{K}^{n\times n}\colon AA^{-1} = A^{-1}A = I$ (invertierbar)
- $\det(A) \neq 0$
- $\operatorname{rg}(A) = n$ (regulär)
- $\ker(A) = \{0\}$ ($Ax = 0 \implies x = 0$) (injektiv)
- $\forall b \in \mathbb{K}^n\ \exists x \in \mathbb{K}^n\colon Ax = b$ (surjektiv)
- $\forall b \in \mathbb{K}^n\ \exists! x \in \mathbb{K}^n\colon Ax = b$ (bijektiv)
- Zeilen oder Spalten von A linear unabhäningig (injektiv)
- Zeilen oder Spalten von A erzeugen $\mathbb{K}^n$ (surjektiv)
- Zeilen oder Spalten von A bilden Basis von $\mathbb{K}^n$ (bijektiv)
- A^T ist invertierbar
- 0 ist kein Eigenwert von A.

Mindestens 5 Punkte sollten genannt und die Zusammenhänge erklärt werden.

Um die Inverse zu bestimmen gibt es folgende Formel:

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

(adj bezeichnet die Adjunkte Matrix). Sie ist nur auf $\det(A) \neq 0$ definiert, was den invertierbaren Matrizen entspricht.

Formal ist ein Skalarprodukt eine Abbildung $V \times V \to \mathbb{R}_{\geq 0},\ (x,y) \mapsto \langle x,y\rangle$ über einem $\mathbb{K}$-Vektorraum V, welche **positiv definit**, **bilinear** und **hermitesch** im Fall $\mathbb{K} = \mathbb{C}$ ist bzw. **symmetrisch** im Fall $\mathbb{K} = \mathbb{R}$ ist. Ein Vektorraum mit Skalarprodukt heißt **unitär** mit $\mathbb{K} = \mathbb{C}$ bzw. **euklidisch** mit $\mathbb{K} = \mathbb{R}$. Typische Skalarprodukte für Vektorräume sind:

$$\begin{array}{lll} \mathbb{K}^n & \langle x,y\rangle := \overline{x}^T y & \text{bzw. } x^T y \\ \ell^2 & \langle (x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}\rangle := \displaystyle\sum_{n=0}^{\infty} \overline{x_n} y_n & \text{bzw. } \displaystyle\sum_{n=0}^{\infty} x_n y_n \\ L^2 & \langle f,g\rangle := \displaystyle\int \overline{f} g & \text{bzw. } \displaystyle\int fg \end{array}$$

Aufgrund der positiven Definitheit induziert jedes Skalarprodukt eine **Norm**: $\|x\| := \sqrt{\langle x,x\rangle}$, welche die **Cauchy-Schwarz-Ungleichung** $|\langle x,y\rangle| \leq \|x\|\|y\|$ erfüllt. Im euklidischen Fall lässt sich der Winkel φ zwischen zwei Vektoren x und y durch $\langle x,y\rangle = \|x\|\|y\| \cos\varphi$ bestimmen.

Vektoren heißen **orthogonal**, falls $\langle x,y\rangle = 0$. In diesem Fall gilt mit der induzierten Norm der Satz des Pythagoras: $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. Eine Menge von paarweise orthogonalen Vektoren S heißt Orthogonalsystem und ist linear unabhängig falls $0 \notin S$, also Basis eines Untervektorraums. Sind alle Vektoren in S normiert, so spricht man von einem Orthonormalsystem. Für $\{x_1,\ldots,x_n\} \subset V$ beliebig lässt sich mithilfe des Schmidtschen Orthogonalisierungsverfahrens eine Orthonormalbasis für $\operatorname{span}\{x_1,\ldots,x_n\}$ ermitteln.

Für eine lineare Abbildungen S ist S^* die adjungierte Abbildung, falls $\forall x,y \in V : \langle Sx,y\rangle = \langle x, S^*y\rangle$. Ist S eine Matrix so gilt $S^* = \overline{S}^T$ bzw. $S^* = S^T$. S heißt selbstadjungiert, falls $S = S^*$.

Eine Teilmenge K eines normierten Raums $(X, \|\cdot\|)$ heißt kompakt, wenn jede Folge $(x_k)_{k\in\mathbb{N}}$ in K eine konvergente Teilfolge mit Grenzwert in K hat. Äquivalent hierzu ist die folgende Definition: K heißt kompakt, wenn für jede Familie $(U_i)_{i\in I}$ von offenen Mengen mit

$$K \subset \bigcup_{i\in I} U_i$$

es ein $n \in \mathbb{N}$ und Indizes $i_1,\ldots,i_n \in I$ gibt mit

$$K \subset \bigcup_{k=1}^{n} U_{i_k}.$$

Jede kompakte Menge ist abgeschlossen und beschränkt. Die Umkehrung gilt nach dem **Satz von Bolzano/Weierstraß** im Spezialfall $X = \mathbb{R}^n$. Allgemein ist die Implikation „beschränkt und abgeschlossen $\Rightarrow$ kompakt“ jedoch falsch: Im Raum $C[0,1]$ der stetigen Funktionen, versehen mit der Supremumsnorm, ist die Einheitskugel $K = \{f \in C[0,1] : \|f\|_\infty \leq 1\}$ zwar beschränkt und abgeschlossen, aber nicht kompakt, wie man an der Folge $(f_k)_{k\in\mathbb{N}}$ mit $f_k : [0,1] \to \mathbb{R},\ f_k(x) = x^k$ erkennt. Sie konvergiert punktweise gegen eine unstetige Grenzfunktion. Die Konvergenz kann also nicht gleichmäßig sein, damit gibt es auch keine bezüglich $\|\cdot\|_\infty$ konvergente Teilfolge.

Wenn man eine stetige, reellwertige Funktion f auf einem kompakten Definitionsbereich K betrachtet, kann man einige Aussagen über sie machen. Zum Beispiel ist dann auch das Bild $f(K)$ kompakt, und f nimmt auf K sowohl **Maximum** als auch **Minimum** an. Ferner ist f auf K auch gleichmäßig stetig.