Was ist Ableiten?

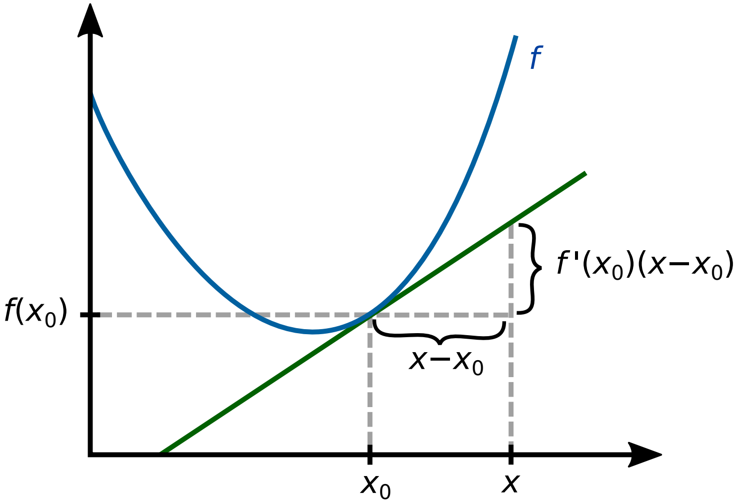
Definition

Ableitung von f an der Stelle x:

$$f'(x) := \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Was bedeutet das?

Beim Ableiten approximiert man den Graph einer Funktion an einer Stelle x_0 durch eine Gerade. Die Definition der Ableitung ist gerade die Berechnungsvorschrift, mit der man den Anstieg $a = f'(x_0)$ dieser Geraden erhält.



f
f'(x0)(x−x0)
f(x0)
x−x0
x0
x

Man kann doch gleich auch eine beliebige Kurve $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ an einem Punkt mit einer Geraden approximieren?

Ja, klar. Alles was wir brauchen ist eine Berechnungsvorschrift, die der Kurve den Richtungsvektor der Geraden zuordnet:

Ableitung einer Kurve

$$\gamma'(t) = \lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h}$$

Eigentlich muss man die gewöhnliche Ableitung nur Komponentenweise berechnen:

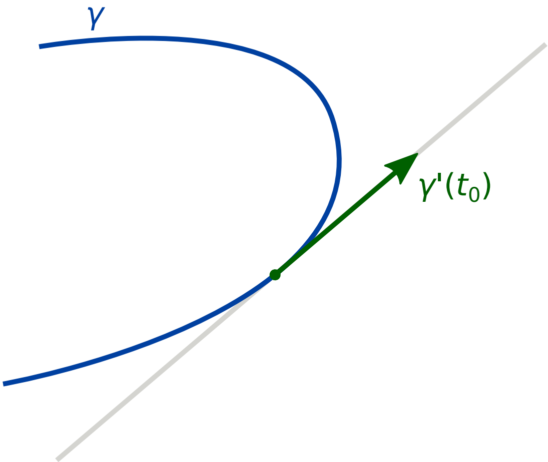
$$\gamma'(t) = \lim_{h \to 0} \frac{\begin{bmatrix} \gamma_x(t+h) \\ \gamma_y(t+h) \end{bmatrix} - \begin{bmatrix} \gamma_x(t) \\ \gamma_y(t) \end{bmatrix}}{h}$$

$$= \lim_{h \to 0} \begin{bmatrix} \frac{\gamma_x(t+h) - \gamma_x(t)}{h} \\ \frac{\gamma_y(t+h) - \gamma_y(t)}{h} \end{bmatrix} = \begin{bmatrix} \gamma_x'(t) \\ \gamma_y'(t) \end{bmatrix}.$$

Oder, um es banaler auszudrücken, einfach mit dem Differentialoperator $D = \mathrm{d}/\mathrm{d}t$ ausmultiplizieren:

$$D\gamma = D(\gamma_x \mathbf{e}_x + \gamma_y \mathbf{e}_y) = \mathbf{e}_x D\gamma_x + \mathbf{e}_y D\gamma_y.$$

Die Basisvektoren $\mathbf{e}_x, \mathbf{e}_y$ dürfen als konstante Faktoren angesehen und vor die Ableitung gezogen werden.


γ
γ'(t₀)

Einwand. Aber eine Oberfläche $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ wird doch auch durch eine Tangentialebene approximiert? Was ist da denn der Anstieg oder die Richtung?

Eine Ebene ist durch

$$E(x,y) = f(x_0,y_0) + a(x - x_0) + b(y - y_0)$$

gegeben. Hält man jetzt $y := y_0$ konstant, dann ergibt sich eine Gerade

$$g(x) = f(x,y_0) + a(x - x_0)$$

mit dem Anstieg a. Es gibt also je Variable einen Anstieg. Diese Anstiege sind die partiellen Ableitungen.

Wir können die Ebenengleichung mittels Skalarprodukt kompakter schreiben:

$$E(x,y) = f(x_0,y_0) + \left\langle \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \right\rangle$$

Der Vektor

$$(\nabla f)(x_0, y_0) := \begin{bmatrix} (\partial_x f)(x_0, y_0) \\ (\partial_y f)(x_0, y_0) \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

heißt *Gradient* von f an der Stelle (x_0, y_0), wobei $\partial_x f$ und $\partial_y f$ die partiellen Ableitungen sind.

Fassen wir doch auch die Koordinaten x, y zu einem Koordinatentupel $\mathbf{x} = (x, y)$ zusammen. Die Gleichung der Ebene lautet nun

$$E(\mathbf{x}) = f(\mathbf{x}_0) + \langle (\nabla f)(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle,$$

was ganz analog zu

$$g(x) = f(x_0) + f'(x_0)(x - x_0)$$

ist.

Einwand. Aber der Definitionsbereich der Fläche $f(x, y)$ kann doch einfach ein affiner Raum sein? Auf einem affinen Raum ist doch nicht unbedingt ein Skalarprodukt definiert? Trotzdem ergibt doch der Gedanke einen Sinn, dass $f(x, y)$ durch eine Ebene approximiert werden kann?

Ja. Das Skalarprodukt ist hier in Wirklichkeit eine duale Paarung aus einem Kovektor mit einem Vektor. Der Gradient

$$(\nabla f)(x_0, y_0) = (\partial_x f)(x_0, y_0)\mathbf{e}_x + (\partial_y f)(x_0, y_0)\mathbf{e}_y$$

muss gegen das Differential

$$(\mathrm{d}f)(x_0, y_0) = (\partial_x f)(x_0, y_0)\mathrm{d}x + (\partial_y f)(x_0, y_0)\mathrm{d}y$$

ausgetauscht werden. Das Differential ist der Kovektor. Die Differentiale $\mathrm{d}x$ und $\mathrm{d}y$ sind die Basisvektoren der Dualbasis zu $(\mathbf{e}_x, \mathbf{e}_y)$.

Duale Paarung?

Das ist fast banal. Bei einer dualen Paarung eines Kovektors $\omega = \omega_x \mathrm{d}x + \omega_y \mathrm{d}y$ mit einem Vektor $\mathbf{v} = v_x \mathbf{e}_x + v_y \mathbf{e}_y$ darf man einfach das Standardskalarprodukt auf die Komponententupel anwenden, auch wenn $B = (\mathbf{e}_1, \mathbf{e}_2)$ keine Orthonormalbasis ist. Duale Paarung:

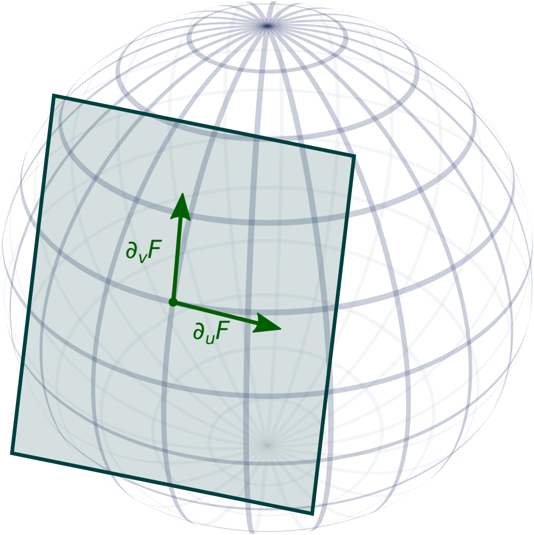
$$\omega(\mathbf{v}) = \omega_x v_x + \omega_y v_y.$$

Bei der Differenz $\mathbf{v} = \mathbf{x} - \mathbf{x}_0$ handelt es sich nicht um einen Punkt, sondern um einen Verschiebungsvektor, der den Punkt $\mathbf{x}_0$ zu $\mathbf{x}$ verschiebt: $\mathbf{x}_0 + \mathbf{v} = \mathbf{x}$. Die Gleichung der Ebene nimmt nun die folgende allgemeine Form an:

$$E(\mathbf{x}) = f(\mathbf{x}_0) + (\mathrm{d}f)(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

Einwand. Aber einer Parameterfläche $F(u, v)$ mit $F\colon \mathbb{R}^2 \to \mathbb{R}^3$ kann man doch auch eine Tangentialebene zuordnen?

Die partielle Ableitung lässt sich Komponentenweise berechnen. Insgesamt ergibt sich ein Tangentialvektor. Die beiden Tangentialvektoren $(\partial_u F)(u_0, v_0)$ und $(\partial_v F)(u_0, v_0)$ bilden dann eine Tangentialbasis des Tangentialraums. Der *Tangentialraum*, das ist wenn man die Tangentialebene als Vektorraum betrachtet.



$\partial_v F$
$\partial_u F$

Die Tangentialvektoren lassen sich nach Wahl eines Koordinatensystems als Komponententupel betrachten und zur Jacobimatrix

$$DF = \begin{bmatrix} \partial_u F_x & \partial_v F_x \\ \partial_u F_y & \partial_v F_y \\ \partial_u F_z & \partial_v F_z \end{bmatrix}$$

zusammenfassen.

Ist $F\colon \mathbb{R}^n \to \mathbb{R}^m$, dann lässt sich F an der Stelle $\mathbf{u}_0$ linear durch einen Tangentialraum approximieren. Die Approximation besitzt die Berechnungsvorschrift

$$T(\mathbf{u}) = F(\mathbf{u}_0) + (DF)(\mathbf{u}_0)(\mathbf{u} - \mathbf{u}_0),$$

was wieder ganz analog zu

$$g(x) = f(x_0) + f'(x_0)(x - x_0)$$

ist. Mit $(DF)(\mathbf{u}_0) \in \mathbb{R}^{m \times n}$ ist die Jacobimatrix gemeint.

Aber eine Matrix ist doch nichts anderes als die Darstellungsmatrix einer linearen Abbildung bezüglich einer Basis. Warum müssen wir eine Basis wählen?

Klar, wir müssen natürlich keine Basis wählen und können die Ableitung $(DF)(\mathbf{u}_0)$ als lineare Abbildung betrachten. Beachte aber dass eine lineare Abbildung immer eine Abbildung zwischen zwei Vektorräumen ist.

Die Differenzen $\mathbf{u}-\mathbf{u}_0$ sind wie gesagt Vektoren. Sie gehören zum Definitionsbereich der linearen Abbildung. Und die Zielmenge der linearen Abbildung? Die ist ebenfalls ein Vektorraum, für den bisher auch immer eine Basis gewählt wurde, denn es war jedes mal der Koordinatenraum.

Einwand. Bei einer Kurve beschreibt der Richtungsvektor den Tangentialraum. Bei einer Funktion muss die Funktion doch erst als Kurve betrachtet werden?

Ja richtig, die Bildmenge der linearen Abbildung ist nur dann der Tangentialraum, wenn eine Parameterdarstellung eines krummen geometrischen Objektes vorliegt.

Einwand. Und wenn bei $F\colon M \to N$ die Definitionsmenge und die Zielmenge selbst so eine krumme Oberfläche ist? Dann ist das doch am Ursprung $\mathbf{u}_0$ kein Vektorraum mehr, und $\mathbf{u} - \mathbf{u}_0$ ergibt keinen sind mehr! Trotzdem wird man irgendwie den Gedanken nicht los, dass F trotzdem irgendwie *smooth* sein könnte, so dass sich die Ableitung bilden lässt.

Huch? Was machen wir jetzt? Nun, die Ableitung soll die Approximation von F durch eine lineare Abbildung sein. Aber die muss immer zwischen zwei Vektorräumen bestehen. Es ist doch irgendwie sinnvoll, dafür an der Stelle $p = \mathbf{u}_0$ den Tangentialraum T_pM und am Bildpunkt $F(p)$ den Tangentialraum $T_{F(p)}N$ zu wählen, so dass F eine lineare Abbildung zwischen diesen beiden Tangentialräumen ist.

Ja. Ernsthaft. Wenn wir dann mit einem Vektor v geringfügig von der Stelle $p = \mathbf{u}_0$ wegschieben, dann lässt sich mit der linearen Abbildung eine Approximation der Funktion berechnen, deren Wert im Tangentialraum zu $F(p)$ liegt.

Ok. Also $F\colon M \to N$ ist eine Funktion zwischen zwei glatten Mannigfaltigkeiten. Weil M glatt ist, gibt es zu jedem Punkt $p \in M$ den Tangetialraum T_pM. Und den Tangentialraum $T_{F(p)}N$ gibt es ebenfalls. Die Funktion F lässt sich nun in der Nähe vom Punkt p durch die lineare Abbildung

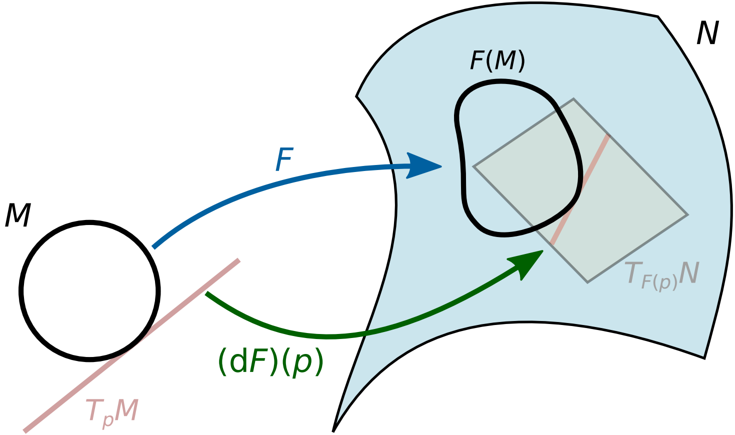
$$(\mathrm{d}F)(p)\colon T_pM \to T_{F(p)}N$$

approximieren.

Man bezeichnet die Ableitung $(\mathrm{d}F)(p)$ auch als das Differential von F an der Stelle p. Die Applikation auf einen Vektor v wird auch *Pushforward* von v genannt. Die Vorstellung dabei ist, dass v von T_pM vorwärts zu $T_{F(p)}N$ gepusht wird.

Äh? Die Ableitung soll doch einen Tangentialraum beschreiben. Warum werden die denn jetzt vorausgesetzt?

Wenn man F tatsächlich als Parameterdarstellung eines krummen geometrischen Objektes betrachtet, dann beschreibt die Bildmenge der linearen Abbildung $(\mathrm{d}F)(p)$ tatsächlich den Tangentialraum T des Objetes am Punkt $F(p)$. Offenbar ist $T \subseteq T_{F(p)}N$. D. h. T darf ein Untervektorraum von $T_{F(p)}N$ sein.



N
F(M)
F
M
$T_{F(p)}N$
(dF)(p)
T_pM

Wenn wir für eine differenzierbare Mannigfaligkeit M eine Karte haben (so wie die Karten der Erdoberfläche), das ist eine diffeomorphe Abbildung

$$\varphi\colon (U \subseteq \mathbb{R}^n) \to V \subseteq M,$$

dann lässt sich φ als gekrümmtes Koordinatensystem auf M betrachten. Die partiellen Ableitungen $(D_k\varphi)(\mathbf{x}_0)$ sind Tangentialvektoren von M am Punkt $p = \varphi(\mathbf{x}_0)$. Die Tangentialvektoren bilden die Tangentialbasis, das ist ein lokales ungekrümmtes Koordinatensystem, welches das gerümmte Koordinatensystem bei p approximiert.

Die disjunkte Zusammenfassung aller Tangentialbasen wird als *Rahmen* bezeichnet. Die Karte φ induziert diesen Rahmen, und der Rahmen induziert das Tangentialbündel TM, weil man für jeden Punkt $p \in M$ den Tangentialraum T_pM aus der Basis aufspannen kann, wobei die Information über die Basis verloren geht.

Aber die Ableitung einer Karte φ ist doch auch so ein komischer Pushforward?

Ja, $(\mathrm{d}\varphi)(\mathbf{x}_0)$ ist eine lineare Abbildung, die jedem $v = \mathbf{x} - \mathbf{x}_0$ einen Tangentialvektor

$$w = (\mathrm{d}\varphi)(\mathbf{x}_0)(v)$$

mit $w \in T_pM$ und $p = \varphi(\mathbf{x}_0)$ zuordnet.

Eine Karte φ induziert als Diffeomorphismus an der Stelle $\mathbf{x} \in \mathbb{R}^n$ eine injektive lineare Abbildung $(\mathrm{d}\varphi)(\mathbf{x})$, deren Koordinatendarstellung die Jacobimatrix

$$(D\varphi)(\mathbf{x})\colon \mathbb{R}^n \to \mathbb{R}^m$$

ist. Die Jacobimatrix muss nicht quadratisch sein, hat aber den Rang n, ordent also einer Basis des $\mathbb{R}^n$ eine Basis des n-dimensonalen Tangentialraums zu.

Eigentlich sollte die lineare Abbildung bijektiv sein und eine quadratische Matrix besitzen, aber dann müsste man einen Rahmen für die Bildmenge von φ wählen. Der Rahmen wird jedoch erst durch $d\varphi$ erzeugt.

Ok, der $\mathbb{R}^n$ ist doch der Koordinatenraum. Dort liegt doch die Standardbasis $\mathrm{E} = (\mathbf{e}_k)_{k=1}^n$ vor. Bsp. $\mathbf{e}_1 = (1, 0)$ und $\mathbf{e}_2 = (0, 1)$. Dann ist

$$\mathbf{w}_k = (\mathrm{d}\varphi)(\mathbf{x}_0)(\mathbf{e}_k)$$

die induzierte Basis von T_pM bei $p = \varphi(\mathbf{x}_0)$.

Warum induziert φ eine injektive lineare Abbildung?

Wenn φ diffeomorph ist, sind φ und φ^{-1} differenzierbar. Dann gilt die Kettenregel:

$$\mathrm{id} = \mathrm{d}(\varphi^{-1} \circ \varphi)(\mathbf{x}) = \mathrm{d}\varphi^{-1}(\varphi(\mathbf{x})) \circ \mathrm{d}\varphi(\mathbf{x}).$$

In Worten: $\mathrm{d}\varphi^{-1}(\varphi(\mathbf{x}))$ ist die Linksinverse von $\mathrm{d}\varphi(\mathbf{x})$. Somit muss $\mathrm{d}\varphi(\mathbf{x})$ injektiv sein.

Ende.

Juli 2020