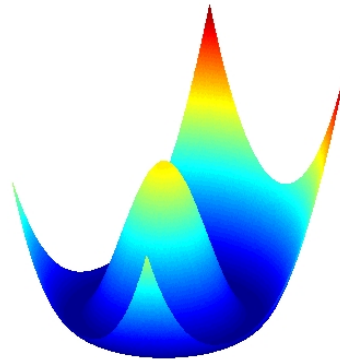


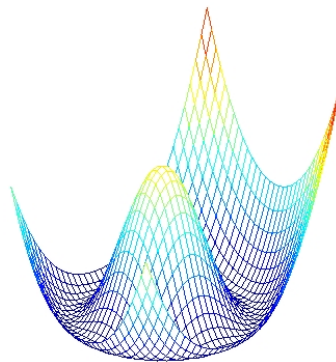
2 Mehrdimensionale Differenzialrechnung

2.1 Geometrische Interpretation

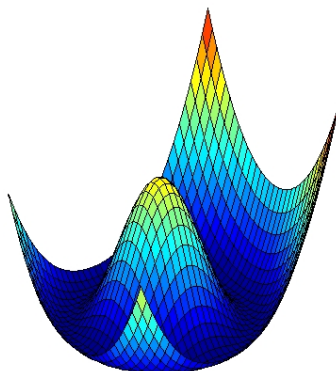
Gegeben sei eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Bekanntlich ist f gerade durch die Werte festgelegt, die f auf dem Definitionsbereich $\mathcal{D}(f) = \mathbb{R}^n$ annimmt, also durch unendlich viele Zahlen. Damit können wir in der Praxis eine Funktion durch Berechnung aller Funktionswerte nur in seltenen Fällen erfassen.



Wählen wir ein beschränktes Gebiet $G \subseteq \mathcal{D}(f)$ und zerlegen G in ein Gitter (beispielsweise äquidistant), so erlangen wir durch Berechnung derjenigen Werte, die f auf den (endlich vielen) Gitterpunkten annimmt, schon mal einen ersten Eindruck von f .



Ist die Funktion hinreichend „glatt“, dann können wir durch ein hinreichend feines Gitter die Funktion so beliebig nahe approximieren, indem wir die Eckpunkte einer Gittermasche aus G verbinden und so statt des interessierenden Graphen Γ_f von G eine Facettenstruktur erhalten, die diesem beliebig nahe kommt.



Was gewinnen wir dabei? Indem wir Γ_f durch endlich viele, affine Ebenen approximieren, kennen wir mit den berechneten Gitterpunkten schon näherungsweise den ganzen Graphen. Im Gegensatz zum Graphen allgemeiner Funktionen sind affine Ebenen nämlich durch drei „Werte“ schon festgelegt: Durch einen Stützvektor und zwei linear unabhängige Vektoren, die in der Ebene liegen. Auf diese Weise lässt sich der gesamte approximative Graph durch endlich viele Werte beschreiben und damit beispielsweise numerisch bestimmen.

2.2 Tangentialebenen

Interessieren wir uns für eine lokale Umgebung U eines einzelnen Punktes $x = (x_1, \dots, x_n)$ auf dem Graphen, dann ist es günstig, eine affine Ebene zu betrachten, in der dieser Punkt liegt und die auf U möglichst wenig von Γ_f abweicht. Geometrisch ist klar, dass dies die (eindeutig bestimmte) „Tangentialebene“ an $(x, f(x))$ sein muss.

Im Eindimensionalen war das die Tangente

$$t^x(t) = f(x) + f'(x)(t - x),$$

also eine affin lineare Abbildung. Den mehrdimensionalen Fall wollen wir auf den Eindimensionalen zurückführen.

Wir wählen dazu die Gerade $G_1 : t \mapsto x + te^{(1)}$ (d.h. G_1 verläuft in $\mathcal{D}(f)$ parallel zur x_1 -Achse durch x). Setzt man

$$g_1(t) := f(x + te^{(1)}),$$

dann liefert g_1 die Funktionswerte von f entlang dieser Geraden. Der Clou: g_1 ist eine Funktion von \mathbb{R} nach \mathbb{R} ! Bestimmen wir nun nach altbewährter [AI]-Methode die Ableitung $g'_1 : \mathbb{R} \rightarrow \mathbb{R}$, dann beschreibt

$$t_1^x(t) := f(x) + g'_1(0)t$$

eine Tangente an $(x, f(x))$. Die Steigung $g'_1(0) \in \mathbb{R}$ erhält einen speziellen Namen: Wir nennen diesen Wert die *erste partielle Ableitung von f an der Stelle x* und schreiben

$$\partial_1 f(x) := g'_1(0).$$

Analog definieren wir $g_i(t) := f(x + te^{(i)})$ und $t_i^x(t) := f(x) + g'_i(0)e^{(i)}$ ($2 \leq i \leq n$) und erhalten so n Tangenten an $(x, f(x))$ mit den Steigungen

$$\partial_i f(x) := g'_i(0).$$

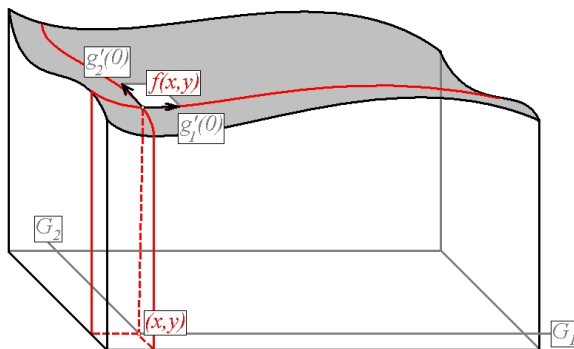
Dann beschreibt

$$(x, f(x)) + \text{span}(e^{(1)} + \partial_1 f(x)e^{(n+1)}, \dots, e^{(n)} + \partial_n f(x)e^{(n+1)})$$

einen affinen, n -dimensionalen Unterraum des \mathbb{R}^{n+1} : Die *Tangentialebene* an f im Punkt $(x, f(x))$. Wir erhalten die Parametrisierung

$$T(y) = f(x) + \sum_{i=1}^n \partial_i f(x)y_i = f(x) + A(x)y \quad \text{mit} \quad A(x) = (\partial_i f(x))_{1 \leq i \leq n}$$

und nennen $A(x) = \mathfrak{J}_f(x)$ die *JACOBI-Matrix von f an der Stelle x* .



Man beachte dabei, dass $A(x)$ als Darstellungsmatrix einer linearen Abbildung keinen Sinn ergibt, wenn nicht bekannt ist, bzgl. welcher Basis wir arbeiten. In unserem Fall ist das wie meistens $\mathfrak{E} := (e^{(1)}, \dots, e^{(n)})$.

2.3 Das totale Differenzial

DEFINITION

Seien U, V endlich dimensionale, normierte Räume, $\Omega \subset U$ offen und $\Omega_x := \{h \in U \mid x + h \in \Omega\}$ für $x \in \Omega$.

Dann heißt $f : \Omega \rightarrow V$ an der Stelle $x \in \Omega$ **differenzierbar**, falls es ein $A(x) \in L(U, V)$ und eine in $0 \in \Omega_x$ stetige Abbildung $r(x, \cdot) : \Omega_x \rightarrow V$ mit $r(x, 0) = 0$ gibt, so dass

$$\forall h \in \Omega_x : f(x + h) = f(x) + A(x)h + \|h\|r(x, h),$$

wobei $A(x)h := A(x)(h)$.

$f : \Omega \rightarrow V$ heißt (in Ω) **differenzierbar**, falls f in allen $x \in \Omega$ differenzierbar ist.

$A : \Omega \rightarrow L(U, V)$ wird auch mit f' (**erste Ableitung**) oder df (**totales Differenzial**) bezeichnet.

BEMERKUNG

(1) Die Definition ist nur sinnvoll, wenn die lineare Abbildung $A(x)$ eindeutig bestimmt ist.

Angenommen, $f(x + h) = f(x) + A_1(x)h + \|h\|r_1(x, h) = f(x + h) = f(x) + A_2(x)h + \|h\|r_2(x, h)$.
Dann gilt

$$(A_1(x) - A_2(x))\frac{h}{\|h\|} = r_2(x, h) - r_1(x, h) \text{ für alle } h \in \Omega_x.$$

Da U offen, gibt es $\epsilon > 0$ mit $B_\epsilon(x) \subseteq U$, also auch $B_\epsilon(0) \subseteq \Omega_x$, d.h. für jedes $y \in \partial V_1(0)$ gilt $h_n := \frac{1}{n}y \in \Omega_x$ für hinreichend großes $n \in \mathbb{N}$.

Da $\frac{h_n}{\|h_n\|} = y$, gilt dann

$$(A_1(x) - A_2(x))y = r_2(x, h_n) - r_1(x, h_n) \xrightarrow{n \rightarrow \infty} 0,$$

also $A_1(x)y = A_2(x)y$ auf $\partial B_1(0)$ und damit $A_1(x) = A_2(x)$.

(2) Ist f differenzierbar, so ist $f' : \Omega \rightarrow L(U, V)$.

(3) Bzgl. der kanonischen Basis \mathfrak{E} hat $f'(x)$ die Darstellungsmatrix

$$\mathfrak{J}_f(x) = \begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix} \quad (\text{JACOBI-Matrix}).$$

Dies folgt sofort aus $f'(x)e_j^{(n)} = \partial_j f(x)$ ($1 \leq j \leq n$).

(4) Sei $f \in \mathfrak{C}^1(U, \mathbb{R})$ mit $U \subseteq \mathbb{R}^n$. Dann ist $\mathfrak{J}_f(x) = (\partial_1 f(x), \dots, \partial_n f(x)) \in \mathbb{R}^{1 \times n}$ und wir setzen

$$\text{grad} f(x) := \nabla f(x) := (\mathfrak{J}_f(x))^T = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \in \mathbb{R}^{n \times 1}.$$

$\nabla f(x)$ („Nabla“) heißt der **Gradient** von f an der Stelle x .

(5) Für die „Richtungsableitungen“ $D_\xi f(x) := f'(x)\xi$ gilt dann:

$$D_\xi f(x) = \sum_{i=1}^n \xi_i \partial_i f(x) = \left\langle \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \right\rangle = \langle \xi, \nabla f(x) \rangle,$$

d.h. $D_\xi f(x)$ lässt sich als Skalarprodukt des Gradienten von f bei x und des Richtungsvektors ξ schreiben.

2.4 Richtungsableitungen

DEFINITION

Seien $f : U \rightarrow \mathbb{R}$, $U \subseteq \mathbb{R}^n$ offen, $x \in U$, $\xi \in \mathbb{R}^n$, $\|\xi\|_2 = 1$ und $\mathcal{D}(g) := \{t \in \mathbb{R} \mid x + t\xi \in U\}$. Setze

$$g : \begin{array}{l} \mathcal{D}(g) \rightarrow \mathbb{R} \\ t \mapsto f(x + t\xi) \end{array} .$$

Existiert $g'(0)$, so nennt man den Wert die **Richtungsableitung** von f im Punkt x in Richtung ξ und schreibt

$$D_\xi f(x) := g'(0).$$

Für die Koordinatenrichtungen $e^{(i)}$ ($1 \leq i \leq n$) nennt man $D_{e^{(i)}} f(x)$ auch **i -te partielle Ableitung** von f an der Stelle x und schreibt

$$\partial_i f(x) := D_{e^{(i)}} f(x).$$

f heißt in x **partiell differenzierbar**, falls $\partial_1 f(x)$, ..., $\partial_n f(x)$ existieren.

BEMERKUNG

Die Berechnung partieller Ableitungen funktioniert wie in 1D:

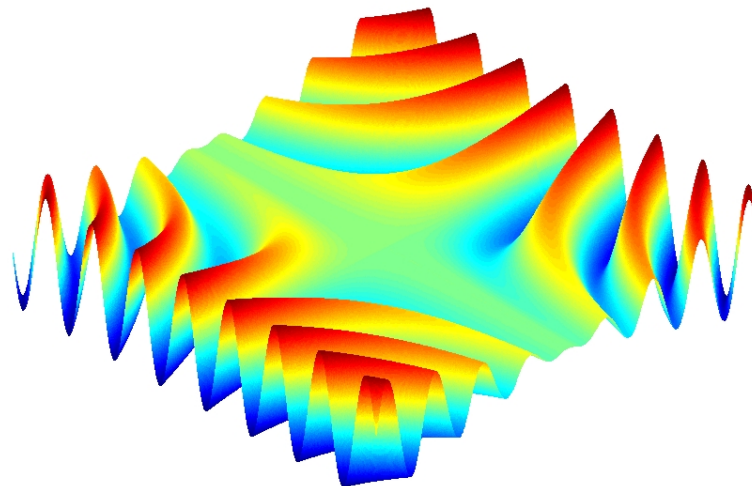
$$\partial_i f(x_1, \dots, x_n) = \frac{d}{dx_i} f(x_i).$$

Die anderen Variablen von f werden also beim partiellen Ableiten wie Konstanten behandelt.

BEISPIEL

Gegeben sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = x \sin(xy).$$



Dann sind die partiellen Ableitungen von f :

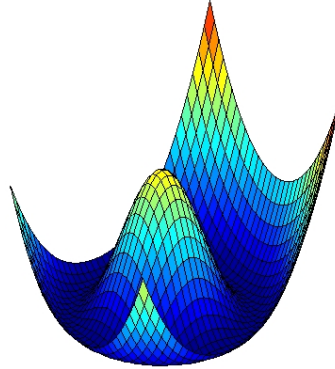
$$\begin{aligned} \partial_x f(x, y) &= \sin(xy) + x \cos(xy)y; \\ \partial_y f(x, y) &= x \cos(xy)x. \end{aligned}$$

BEISPIEL

Wir wenden uns wieder der eingangs betrachteten Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ zu, die durch

$$f(x, y) := 5 \exp(-x^2 - (y - 2)^2) + x^2 + (y - 2)^2$$

gegeben ist.



Die partiellen Ableitungen von f sind

$$\begin{aligned} \partial_x f(x, y) &= 5 \exp(-x^2 - (y - 2)^2)(-2x) + 2x \text{ und} \\ \partial_y f(x, y) &= 5 \exp(-x^2 - (y - 2)^2)(-2(y - 2)) + 2(y - 2). \end{aligned}$$

Damit erhalten wir

$$\mathfrak{J}_f(x, y) = \begin{pmatrix} \partial_x f(x, y) & \partial_y f(x, y) \end{pmatrix} \quad \text{bzw.} \quad \nabla f(x, y) = \begin{pmatrix} \partial_x f(x, y) \\ \partial_y f(x, y) \end{pmatrix}.$$

Die Tangentialebene an einem Punkt (\bar{x}, \bar{y}) wird dann beschrieben durch

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ f(\bar{x}, \bar{y}) \end{pmatrix} + \text{span} \left(\begin{pmatrix} 1 \\ 0 \\ \partial_x f(\bar{x}, \bar{y}) \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \partial_y f(\bar{x}, \bar{y}) \end{pmatrix} \right)$$

bzw. durch die Parametrisierung

$$T(x, y) := f(\bar{x}, \bar{y}) + \partial_x f(\bar{x}, \bar{y})x + \partial_y f(\bar{x}, \bar{y})y.$$

Wir berechnen exemplarisch die Tangentialebenen in $(\bar{x}_1, \bar{y}_1) = (-0, 5; 1)$ und $(\bar{x}_2, \bar{y}_2) = (0; 2)$:

$$\begin{aligned} T_1(x, y) &= [5 \exp(-1, 25) + 1, 25] + [5 \exp(-1, 25) - 1]x + [10 \exp(-1, 25) + 2]y \\ T_2(x, y) &= 5. \end{aligned}$$

