Multivariate Kettenregel

Für die Hintereinanderschaltung

$$h = g \circ f : x \mapsto y = f(x) \mapsto z = g(y) = h(x)\,,$$

stetig differenzierbarer Funktionen $f : \mathbb{R}^n \supseteq U \to \mathbb{R}^\ell$ und $g : \mathbb{R}^\ell \supseteq V \to \mathbb{R}^m$ mit $f(U) \subseteq V$ gilt

$$h'(x) = \underbrace{g'(y)}_{m\times\ell}\underbrace{f'(x)}_{\ell\times n},$$

d.h. die $m \times n$-Jacobi-Matrix von h ist das Produkt der Jacobi-Matrizen von f und g. Die einzelnen Einträge von h' ergeben sich durch Matrixmultiplikation:

$$\frac{\partial h_i}{\partial x_k} = \sum_{j=1}^{\ell} \frac{\partial g_i}{\partial y_j}\frac{\partial f_j}{\partial x_k}, \quad i = 1, \ldots, m,\ k = 1, \ldots, n\,.$$

Diese Identität vereinfacht sich, wenn eine oder zwei der Dimensionen gleich eins sind. Beispielsweise hat eine multivariate Funktion g entlang einer Kurve mit Parametrisierung f,

$$x \mapsto h(x) = g(f_1(x), \ldots, f_\ell(x)),$$

die Ableitung

$$\frac{\mathrm{d}h}{\mathrm{d}x} = \partial_1 g(f(x)) f_1'(x) + \cdots + \partial_\ell g(f(x)) f_\ell'(x) = (\operatorname{grad} g)^{\mathrm{t}}\big|_{f(x)} f'(x),$$

d.h. $h'(x)$ ist das Skalarprodukt aus Gradient von g und Tangentenvektor von f.

Für Funktionen von zwei oder drei Veränderlichen werden oft statt der Index-Schreibweise verschiedene Buchstaben für die Variablen verwendet. Beispielsweise ist die Jacobi-Matrix der Funktion

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} p(u(x,y), v(x,y)) \\ q(u(x,y), v(x,y)) \end{pmatrix}$$

gemäß der Kettenregel das Matrixprodukt

$$\left.\begin{pmatrix} p_u & p_v \\ q_u & q_v \end{pmatrix}\right|_{(u(x,y),v(x,y))} \left.\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}\right|_{(x,y)} .$$

Beweis

Definition der Ableitung und Jacobi-Matrix:

$$\varphi(x+\Delta x) = \varphi(x) + \varphi'(x)\Delta x + o(|\Delta x|)$$

Existenz der Ableitungen $f'(x)$ und $g'(y)$ $\quad\Longrightarrow$

$$g(f(x+\Delta x)) = g(f(x)+\underbrace{f'(x)\Delta x + o(|\Delta x|)}_{\Delta y}) = g(y)+[g'(y)\,\Delta y]+o(|\Delta y|)$$

$\leadsto$ Formel für die Jacobi-Matrix von $h = g \circ f$, da

$$[g'(y)\,\Delta y] = \underbrace{g'(y)f'(x)}_{h'(x)}\,\Delta x + o(|\Delta x|)$$

und $|\Delta y| = O(|\Delta x|)$

Beispiel

Jacobi-Matrix der Hintereinanderschaltung der Funktionen

$$y = f(x) = \begin{pmatrix} x_3 \sin x_1 \\ \mathrm{e}^{x_2}/x_3 \end{pmatrix}, \quad g(y) = \begin{pmatrix} y_1 \\ y_1 + \ln y_2 \\ 0 \\ y_2 \cos y_1 \end{pmatrix}$$

für $x = p = (\pi,\, 0,\, 1)^{\mathrm{t}}$

$$f'(x)_{|x=p} = \begin{pmatrix} x_3 \cos x_1 & 0 & \sin x_1 \\ 0 & \mathrm{e}^{x_2}/x_3 & -\mathrm{e}^{x_2}/x_3^2 \end{pmatrix}_{|x=p} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$g'(y)_{|y=f(p)} = \begin{pmatrix} 1 & 0 \\ 1 & 1/y_2 \\ 0 & 0 \\ -y_2 \sin y_1 & \cos y_1 \end{pmatrix}_{|y=(0,1)} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Kettenregel $\quad\rightsquigarrow\quad h'(\pi,0,1) = g'(0,1)\, f'(\pi,0,1)$ und nach Einsetzen

$$h'(\pi,0,1) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

Beispiel

Kettenregel für eine skalare Funktion $f(x, y)$ entlang einer Kurve $t \mapsto (x(t), y(t))^{\mathrm{t}}$

Spezialisierung der allgemeinen Formel $\quad\rightsquigarrow$

$$\frac{d}{dt} f(x(t), y(t)) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$$

(Skalarprodukt von Gradient und Tangentenvektor)

Anwendung im konkreten Fall

$$f(x, y) = xy^2, \quad x = \cos t,\ y = \sin t \,(\text{Kreis})$$

Gradient: $(f_x,\ f_y)^{\mathrm{t}} = (y^2, 2xy)^{\mathrm{t}}$, Auswertung entlang der Kurve $\quad\rightsquigarrow$
$(\sin^2 t, 2\cos t \sin t)^{\mathrm{t}}$
Tangentenvektor: $(x', y')^{\mathrm{t}} = (-\sin t, \cos t)^{\mathrm{t}}$

Einsetzen in die Kettenregel $\quad\rightsquigarrow$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}f(x(t),y(t)) &= (\sin^2 t, 2\cos t\sin t)\begin{pmatrix}-\sin t\\ \cos t\end{pmatrix}\\ &= -\sin^3 t + 2\cos^2 t\sin t = 2\sin t - 3\sin^3 t\end{aligned}$$

Vergleich mit der direkten Berechnung:
$f(x(t),y(t)) = \cos t\sin^2 t \quad \implies$

$$\frac{\mathrm{d}}{\mathrm{d}t}f(x(t),y(t)) = -\sin t(\sin^2 t) + \cos t(2\sin t\cos t) \quad \checkmark$$

Beispiel

Berechnung des Gradienten der Funktion $h = g \circ f$ für

$$f(x,y) = \begin{pmatrix} x+y \\ x-y \\ x^2+y^2-1 \end{pmatrix}, \quad g(u,v,w) = u^2+v^2+w^2$$

Jacobi-Matrix von f

$$f' = (f_x, f_y) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2x & 2y \end{pmatrix}$$

Gradient von g

$$(\operatorname{grad} g)^{\mathrm{t}} = g' = (2u, 2v, 2w) = 2(x+y, x-y, x^2+y^2-1)$$

Kettenregel, $h'(x,y) = g'(f(x,y))f'(x,y) \quad \Longrightarrow$

$$\underbrace{(\operatorname{grad} h)^{\mathrm{t}}}_{h'} = \underbrace{(\operatorname{grad} g)^{\mathrm{t}}}_{g'} f' = 2 \begin{pmatrix} x+y+x-y+2x(x^2+y^2-1) \\ x+y-x+y+2y(x^2+y^2-1) \end{pmatrix}^{\mathrm{t}}$$

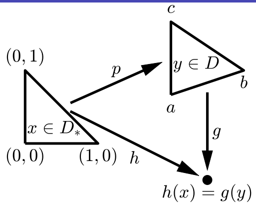
und nach Vereinfachung

$$\operatorname{grad} h = 4(x^2+y^2) \begin{pmatrix} x \\ y \end{pmatrix}$$

Beispiel

Transformation von Gradienten bei affiner Abbildung eines Referenzdreiecks

Parametrisierung eines allgemeinen Dreiecks D mit Eckpunkten $a = (a_1, a_2)^{\mathrm{t}}$, b, c, ausgehend von dem Referenzdreieck $D_* : x_1 + x_2 \leq 1$, $x_k \geq 0$



$$y = p(x) = a + (b - a)x_1 + (c - a)x_2$$

Jacobi-Matrix

$$p'(x) = \frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = (b - a, c - a) = \begin{pmatrix} b_1 - a_1 & c_1 - a_1 \\ b_2 - a_2 & c_2 - a_2 \end{pmatrix}$$

Kettenregel $\quad \Longrightarrow$

$$(\operatorname{grad} h(x))^{\mathrm{t}} = (\operatorname{grad} g(y))^{\mathrm{t}}\, p'(x), \quad h(x) = g(p(x))$$

für skalare Funktionen g und h

Anwendung: Aufstellen von Steifigkeitsmatrizen für Finite Elemente