

Ableitung vektorwertiger Funktionen und die Kettenregel

Wir wollen Funktionen „verschachteln“, also hintereinander ausführen. Und weil das Verschachteln auch wieder eine Funktion ergibt, kann man wiederum fragen, wie man davon die Ableitung berechnen kann. Die Notation kann dabei sehr aufwendig werden, weswegen man sie in der Regel stark vereinfacht, weil ja eh alle wissen, was gemeint ist ... bis eben auf die, die es zum ersten Mal kennenlernen. Deshalb hier einmal ganz ausführlich.

Zuerst die äußere Funktion in der ganz grundsätzlichen Notation für Funktionen [wie in Mathematik 1, $f: D \rightarrow W, z \rightarrow f(z)$]

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m,$$

$$\mathbf{z} = (z_1, \dots, z_n) \rightarrow \mathbf{f}(\mathbf{z}) = \begin{pmatrix} f_1(\mathbf{z}) \\ \vdots \\ f_m(\mathbf{z}) \end{pmatrix} = \begin{pmatrix} f_1(z_1, \dots, z_n) \\ \vdots \\ f_m(z_1, \dots, z_n) \end{pmatrix}$$

In Worten: Der Vektor \mathbf{z} hat n Komponenten z_1, \dots, z_n . Er ist das Argument der Funktion \mathbf{f} , die als Funktionswert einen Vektor mit m Komponenten liefert. Wenn man diese Komponenten mit den ausführlichen Argumenten hinschreibt, wird es eben länglich. Aber noch geht's.

Von dieser Funktion \mathbf{f} können wir nun die erste Ableitung berechnen, wir haben das Gradient genannt. Ein weiterer Name dafür ist „Jacobi-Matrix von \mathbf{f} “.

$$\begin{aligned} \text{grad } \mathbf{f}(\mathbf{z}) &= \begin{pmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1(\mathbf{z})}{\partial z_1} & \cdots & \frac{\partial f_1(\mathbf{z})}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{z})}{\partial z_1} & \cdots & \frac{\partial f_m(\mathbf{z})}{\partial z_n} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f_1(z_1, \dots, z_n)}{\partial z_1} & \cdots & \frac{\partial f_1(z_1, \dots, z_n)}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(z_1, \dots, z_n)}{\partial z_1} & \cdots & \frac{\partial f_m(z_1, \dots, z_n)}{\partial z_n} \end{pmatrix} \end{aligned}$$

Auch hier behalten die einzelnen partiellen Ableitungen ihre Argumente wie die Originalfunktion \mathbf{f} . Man kann die Argumente als Vektor \mathbf{z} hinschreiben oder eben in den einzelnen Komponenten z_1, \dots, z_n . Letzteres wird schon länglicher.

Jetzt die zweite Funktion, die unsere innere Funktion sein wird. Erst einmal ganz analog:

$$\mathbf{g}: \mathbb{R}^k \rightarrow \mathbb{R}^n,$$

$$\mathbf{x} = (x_1, \dots, x_k) \rightarrow \mathbf{g}(\mathbf{x}) = \begin{pmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_n(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} g_1(x_1, \dots, x_k) \\ \vdots \\ g_n(x_1, \dots, x_k) \end{pmatrix}$$

In Worten: Der Vektor \mathbf{x} hat k Komponenten x_1, \dots, x_k . Er ist das Argument der Funktion \mathbf{g} , die als Funktionswert einen Vektor mit n Komponenten liefert. Wenn man die Komponenten mit den ausführlichen Argumenten hinschreibt, wird es auch wieder länglich.

Auch hiervon bilden wir wieder den Gradienten bzw. die Jacobi-Matrix:

$$\begin{aligned} \text{grad } \mathbf{g}(\mathbf{x}) &= \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_k} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial g_1(\mathbf{x})}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial g_n(\mathbf{x})}{\partial x_k} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial g_1(x_1, \dots, x_k)}{\partial x_1} & \dots & \frac{\partial g_1(x_1, \dots, x_k)}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(x_1, \dots, x_k)}{\partial x_1} & \dots & \frac{\partial g_n(x_1, \dots, x_k)}{\partial x_k} \end{pmatrix} \end{aligned}$$

Auch hier behalten die einzelnen partiellen Ableitungen ihre Argumente wie die Originalfunktion \mathbf{g} . Man kann die Argumente als Vektor \mathbf{x} hinschreiben oder eben in den einzelnen Komponenten x_1, \dots, x_k . Letzteres wird eben wiederum länglicher.

Und damit haben wir nun die Notation zusammen, um die Kettenregel in voller Ausführlichkeit hinzuschreiben. Ersteinmal unsere verschachtelte Funktion:

$$\mathbf{f} \circ \mathbf{g}: \mathbb{R}^k \rightarrow \mathbb{R}^m,$$

$$\mathbf{x} = (x_1, \dots, x_k)$$

$$\rightarrow \mathbf{f}(\mathbf{g}(\mathbf{x})) = \begin{pmatrix} f_1(\mathbf{g}(\mathbf{x})) \\ \vdots \\ f_m(\mathbf{g}(\mathbf{x})) \end{pmatrix} = \begin{pmatrix} f_1(g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) \\ \vdots \\ f_m(g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) \end{pmatrix} = \begin{pmatrix} f_1(g_1(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k)) \\ \vdots \\ f_m(g_1(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k)) \end{pmatrix}$$

Auch hier noch einmal in Worten: Die äußere Funktion f hat als Funktionswert einen Vektor mit m Komponenten. Sie hat als Argumente die n Funktionswerte, die ihr die innere Funktion g liefert. Diese innere Funktion g wiederum hat als Argumente die k Komponenten des Vektors x .

So langsam zeichnet sich ab, dass das explizite Aufschreiben der Argumente schlicht Platz braucht und sehr unübersichtlich wird. Aber eine Funktion braucht eh immer Argumente, sonst kann man den Funktionswert ja gar nicht ausrechnen. Deshalb fasst man sich nach etwas Übung gerne kurz. Das wird nun spätestens klar, wenn wir die Kettenregel aufschreiben:

$$\text{grad}(f \circ g)(x) = \text{grad } f(g(x)) \cdot \text{grad } g(x)$$

$$\begin{aligned}
 &= \begin{pmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_n} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_k} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1(g(x))}{\partial z_1} & \cdots & \frac{\partial f_1(g(x))}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(g(x))}{\partial z_1} & \cdots & \frac{\partial f_m(g(x))}{\partial z_n} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(x)}{\partial x_1} & \cdots & \frac{\partial g_n(x)}{\partial x_k} \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\partial f_1(g_1(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k))}{\partial z_1} & \cdots & \frac{\partial f_1(g_1(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k))}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(g_1(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k))}{\partial z_1} & \cdots & \frac{\partial f_m(g_1(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k))}{\partial z_n} \end{pmatrix} \\
 &\quad \cdot \begin{pmatrix} \frac{\partial g_1(x_1, \dots, x_k)}{\partial x_1} & \cdots & \frac{\partial g_1(x_1, \dots, x_k)}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(x_1, \dots, x_k)}{\partial x_1} & \cdots & \frac{\partial g_n(x_1, \dots, x_k)}{\partial x_k} \end{pmatrix}
 \end{aligned}$$

In der letzten Zeile dieser Darstellung steht das, was man eigentlich ausrechnen muss. Deshalb auch wieder in Worten: Die Komponenten der Jacobi-Matrix für f werden durch die Ableitungen von f nach z berechnet, wie ganz oben. Die Stellen, an denen die Zahlenwerte dieser Ableitungen aber nun ausgerechnet werden, sprich die Argumente, die in diese Ableitungen eingesetzt werden, sind die Funktionswerte von g an der Stelle x . *Die Verschachtelung bleibt also erhalten! Daran ändert sich nichts!*

Diese äußere Ableitung wird nun mit der inneren Ableitung multipliziert. Weil beides Matrizen sind, ist das eine Matrixmultiplikation. Diese innere Ableitung ist eben die Jacobi-Matrix der Funktion g und die einzelnen Komponenten dieser Matrix werden an den Stellen x berechnet. Und wenn man nun die Komponenten von g und von x auch noch explizit hinschreibt, passt das nicht mehr in eine Zeile.

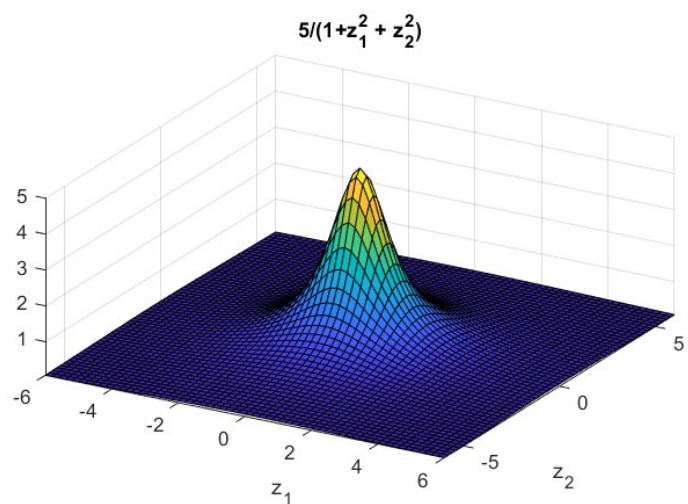
Und jetzt lesen wir die Zeilen der Kettenregel **rückwärts**: Die Argumente der inneren Ableitung müssen wir nicht als Komponenten hinschreiben, weil wir eh die Notation \mathbf{x} für diesen Satz an Variablen (x_1, \dots, x_k) eingeführt haben. Und genauso haben wir \mathbf{g} als Notation für den Vektor (g_1, \dots, g_n) eingeführt. Deshalb müssen wir uns nicht die Mühe machen, die letzte Zeile in voller Länge hinzuschreiben, sondern können Matrizen mit den Argumenten wie im Schritt davor hinschreiben. Aber da stehen jetzt die Argumente, mit denen unsere Funktionen f und \mathbf{g} eh definiert sind. D.h., dass wir hier ein $f(\mathbf{g}(\mathbf{x}))$ und ein $\mathbf{g}(\mathbf{x})$ betrachten, haben wir sowieso schon vorher aufgeschrieben. Dann können wir uns das in der Matrix auch sparen und wir erhalten die erste kurze Matrixnotation, in der wir die Argumente der einzelnen Matrixeinträge nicht mehr explizit hinschreiben. Aber natürlich sind sie nicht weg! Man schreibt sie nur nicht mehr hin, man hat sie „im Sinn“.

Ein erstes Beispiel

Als äußere Funktion f verwenden wir eine „Hut-Funktion“ (die Form erinnert wohl den einen oder die andere an einen Hut ...):

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(\mathbf{z}) = f(z_1, z_2) = \frac{5}{1 + z_1^2 + z_2^2}$$

$$\text{grad } f(\mathbf{z}) = \begin{pmatrix} -\frac{10z_1}{(1 + z_1^2 + z_2^2)^2} & -\frac{10z_2}{(1 + z_1^2 + z_2^2)^2} \end{pmatrix}$$



D.h., $\mathbf{n}=2$ und $\mathbf{m}=1$ und unser $f=f_1$. Die Jacobi-Matrix von f hat also eine Zeile und zwei Spalten.

Jetzt schreiben wir ein erstes $\mathbf{g}(\mathbf{x})$ auf:

$$\mathbf{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, x_2) = \begin{pmatrix} x_1 + 2 \\ x_2 - 2 \end{pmatrix}$$

$$\text{grad } \mathbf{g}(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Die Funktion \mathbf{g} verschiebt also die Punkte der Ebene um einen konstanten Vektor.

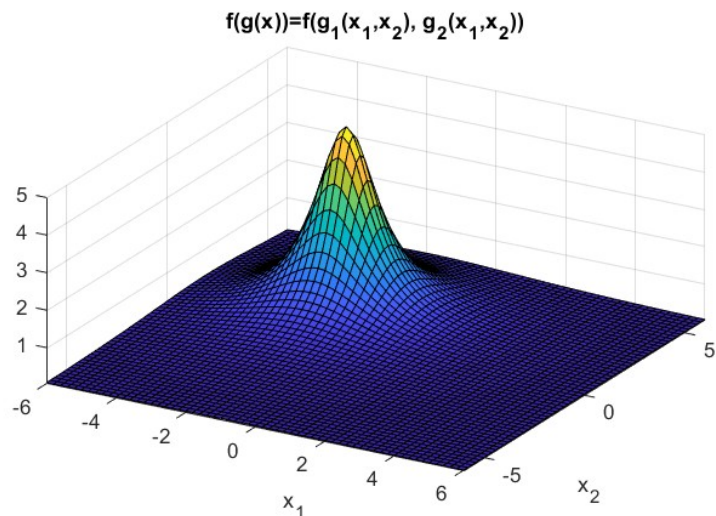
Darüber hinaus ist nun $k=2$ und $n=2$. Die Jacobi-Matrix von \mathbf{g} hat also zwei Zeilen und zwei Spalten.

Die verschachtelte Funktion $f(\mathbf{g}(\mathbf{x}))$ ist dann die folgende.

$$f \circ \mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R} ,$$

$$\begin{aligned} f(\mathbf{g}(\mathbf{x})) &= f(g_1(\mathbf{x}), g_2(\mathbf{x})) \\ &= \frac{5}{1+g_1^2(x_1, x_2)+g_2^2(x_1, x_2)} \\ &= \frac{5}{1+(x_1+2)^2+(x_2-2)^2} \end{aligned}$$

Wir sehen am Bild: Der Peak ist verschoben, hat sich aber in der Form nicht geändert.



Was bedeutet das nun für die Ableitungen? Also erst einmal mit Kettenregel ausrechnen:

$$\text{grad}(f \circ \mathbf{g})(\mathbf{x}) = \text{grad} f(\mathbf{g}(\mathbf{x})) \cdot \text{grad} \mathbf{g}(\mathbf{x})$$

$$\begin{aligned} &= \left(\frac{\partial f_1(g_1(x_1, x_2), g_2(x_1, x_2))}{\partial z_1} \quad \frac{\partial f_1(g_1(x_1, x_2), g_2(x_1, x_2))}{\partial z_n} \right) \cdot \begin{pmatrix} \frac{\partial g_1(x_1, x_2)}{\partial x_1} & \frac{\partial g_1(x_1, x_2)}{\partial x_2} \\ \frac{\partial g_2(x_1, x_2)}{\partial x_1} & \frac{\partial g_2(x_1, x_2)}{\partial x_k} \end{pmatrix} \\ &= \left(-\frac{10g_1(x_1, x_2)}{(1+g_1^2(x_1, x_2)+g_2^2(x_1, x_2))^2} \quad -\frac{10g_2(x_1, x_2)}{(1+g_1^2(x_1, x_2)+g_2^2(x_1, x_2))^2} \right) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \left(-\frac{10g_1(x_1, x_2)}{(1+g_1^2(x_1, x_2)+g_2^2(x_1, x_2))^2} \quad -\frac{10g_2(x_1, x_2)}{(1+g_1^2(x_1, x_2)+g_2^2(x_1, x_2))^2} \right) \\ &= \left(-\frac{10(x_1+2)}{(1+(x_1+2)^2+(x_2-2)^2)^2} \quad -\frac{10(x_2-2)}{(1+(x_1+2)^2+(x_2-2)^2)^2} \right) \end{aligned}$$

In Worten: Erst einmal das Matrixprodukt formal hinschreiben. In der nächsten Zeile haben wir dann g_1 und g_2 in den Gradienten von f (s.o.) eingesetzt und auch den Gradienten von \mathbf{g} sofort komplett aufgeschrieben, denn er ist in diesem Fall die konstante Einheitsmatrix. Das Matrixprodukt können wir nun explizit ausrechnen: Die Multiplikation mit der Einheitsmatrix ändert den Zeilenvektor nicht und deshalb sieht es so aus, als ob die Einheitsmatrix einfach verschwunden wäre. Ist sie aber nicht. Es wurde nur eben das Matrixprodukt ausgerechnet. Und im letzten Schritt setzen wir dann die expliziten Funktionsausdrücke für g_1 und g_2 ein. Und jetzt sieht man, um den Gradienten auszurechnen, braucht man die Argumente. Man sieht aber auch, wenn man das die ganze Zeit mitschleift, hat man viel zu schreiben.

Mit etwas Übung schreibt man das dann so:

$$\begin{aligned} \text{grad}(f \circ \mathbf{g})(\mathbf{x}) &= \text{grad} f(\mathbf{g}(\mathbf{x})) \cdot \text{grad} \mathbf{g}(\mathbf{x}) \\ &= \begin{pmatrix} \frac{\partial f_1(g_1, g_2)}{\partial z_1} & \frac{\partial f_1(g_1, g_2)}{\partial z_2} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -\frac{10g_1}{(1+g_1^2+g_2^2)^2} & -\frac{10g_2}{(1+g_1^2+g_2^2)^2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{10(x_1+2)}{(1+(x_1+2)^2+(x_2-2)^2)^2} & -\frac{10(x_2-2)}{(1+(x_1+2)^2+(x_2-2)^2)^2} \end{pmatrix} \end{aligned}$$

Und weil man ja die Gradienten von f und \mathbf{g} schon vorher aufgeschrieben hat, kann man das dann mit etwas mehr Übung auch noch etwas weiter kürzen:

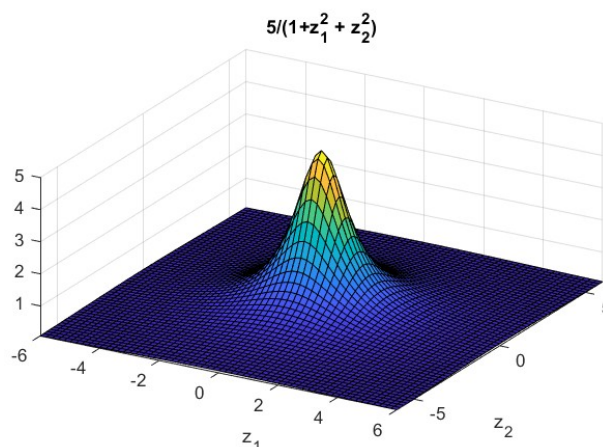
$$\begin{aligned} \text{grad}(f \circ \mathbf{g})(\mathbf{x}) &= \text{grad} f(\mathbf{g}(\mathbf{x})) \cdot \text{grad} \mathbf{g}(\mathbf{x}) = \begin{pmatrix} -\frac{10g_1}{(1+g_1^2+g_2^2)^2} & -\frac{10g_2}{(1+g_1^2+g_2^2)^2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{10(x_1+2)}{(1+(x_1+2)^2+(x_2-2)^2)^2} & -\frac{10(x_2-2)}{(1+(x_1+2)^2+(x_2-2)^2)^2} \end{pmatrix} \end{aligned}$$

Mit diesem Ergebnis können wir nun die ersten Ableitungen dieser verschachtelten Funktion an jeder Stelle \mathbf{x} der Ebene \mathbb{R}^2 berechnen.

Ein zweites Beispiel

Als äußere Funktion verwenden wir das gleiche f :

$$\begin{aligned} f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(\mathbf{z}) = f(z_1, z_2) &= \frac{5}{1+z_1^2+z_2^2} \\ \text{grad} f(\mathbf{z}) &= \begin{pmatrix} -\frac{10z_1}{(1+z_1^2+z_2^2)^2} & -\frac{10z_2}{(1+z_1^2+z_2^2)^2} \end{pmatrix} \end{aligned}$$



D.h., $\mathbf{n}=2$ und $\mathbf{m}=1$ und unser $f=f_1$.

Die Jacobi-Matrix von f hat also wiederum eine Zeile und zwei Spalten.

Nun verwenden wir eine andere Funktion $\mathbf{g}(\mathbf{x})$:

$$\mathbf{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, x_2) = \begin{pmatrix} x_1 + x_2^2 \\ \frac{x_2}{2} \end{pmatrix}$$

$$\text{grad } \mathbf{g}(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 1 & 2x_2 \\ 0 & \frac{1}{2} \end{pmatrix}$$

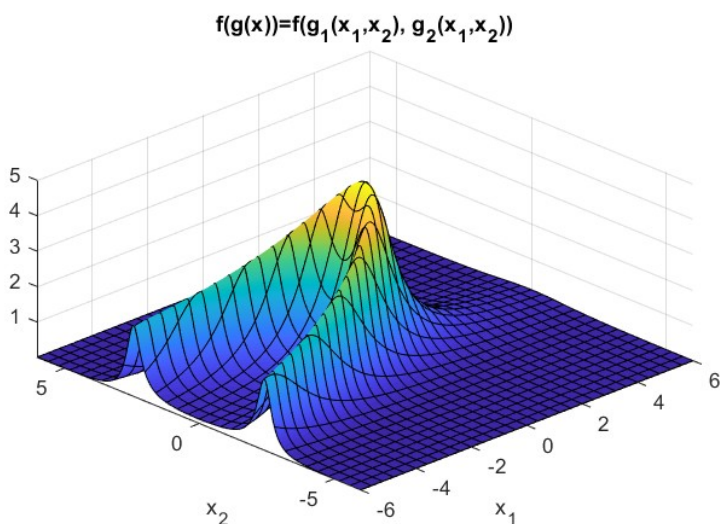
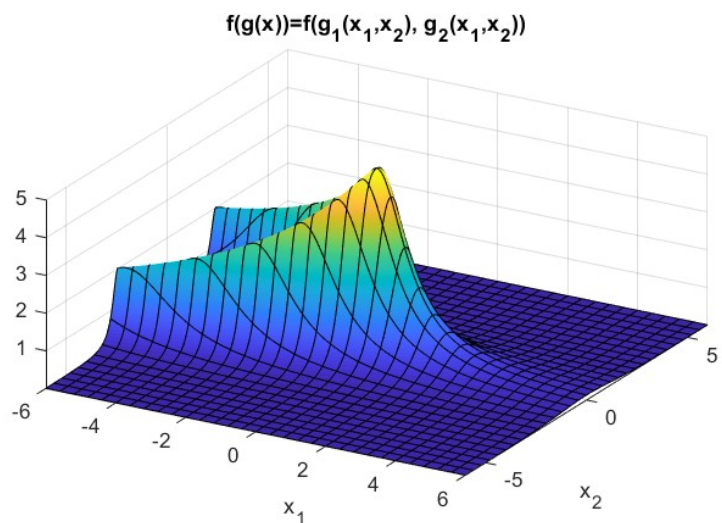
Die Funktion \mathbf{g} verschiebt also Punkte in der Ebene, aber recht kompliziert. Darüber hinaus ist nun auch wieder $\mathbf{k}=2$ und $\mathbf{n}=2$. Die Jacobi-Matrix von \mathbf{g} hat also zwei Zeilen und zwei Spalten.

Wie sieht nun die verschachtelte Funktion aus?

$$f \circ \mathbf{g}: \mathbb{R}^2 \rightarrow \mathbb{R},$$

$$\begin{aligned} f(\mathbf{g}(\mathbf{x})) &= f(g_1(\mathbf{x}), g_2(\mathbf{x})) \\ &= \frac{5}{1 + g_1^2(x_1, x_2) + g_2^2(x_1, x_2)} \\ &= \frac{5}{1 + (x_1 + x_2^2)^2 + \left(\frac{x_2}{2}\right)^2} \end{aligned}$$

Das sieht nun ganz anders aus! Was ist passiert? Nun ja, wenn $x_1 = -x_2^2$ ist, verschwindet ein Term im Nenner! D.h., an diesen Stellen ist der Nenner kleiner als an den Stellen in der Nähe, wo das nicht gilt. Wenn der Nenner kleiner ist, ist die Funktion an diesen Stellen größer. Deshalb ergibt sich dieser parabelförmige Höhenzug für die verschachtelte Funktion. Aber völlig unabhängig davon, rechnen wir den Gradienten wie zuvor aus:



$$\begin{aligned}
 \text{grad}(f \circ \mathbf{g})(\mathbf{x}) &= \text{grad } f(\mathbf{g}(\mathbf{x})) \cdot \text{grad } \mathbf{g}(\mathbf{x}) = \begin{pmatrix} -\frac{10g_1}{(1+g_1^2+g_2^2)^2} & -\frac{10g_2}{(1+g_1^2+g_2^2)^2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 2x_2 \\ 0 & \frac{1}{2} \end{pmatrix} \\
 &= \begin{pmatrix} -\frac{10(x_1+x_2^2)}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} & -\frac{10\frac{x_2}{2}}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 2x_2 \\ 0 & \frac{1}{2} \end{pmatrix} \\
 &= \begin{pmatrix} -\frac{10(x_1+x_2^2)}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} \cdot 1 - \frac{10\frac{x_2}{2}}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} \cdot 2x_2 \\ -\frac{10\frac{x_2}{2}}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} \cdot \frac{1}{2} \end{pmatrix}^T \\
 &= \begin{pmatrix} -\frac{10x_1+20x_2^2}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} \\ \frac{5}{2}x_2 \\ -\frac{5x_2}{\left(1+(x_1+x_2^2)^2+\left(\frac{x_2}{2}\right)^2\right)^2} \end{pmatrix}^T
 \end{aligned}$$

Hier haben wir in den letzten beiden Schritten den Zeilenvektor der besseren Lesbarkeit halber als Spaltenvektor aufgeschrieben, und dann durch das hochgestellte „T“ angezeigt, dass dieser Spaltenvektor noch zu transponieren ist und damit wieder zum Zeilenvektor wird.

Auch mit diesem Ergebnis können wir nun die ersten Ableitungen dieser verschachtelten Funktion an jeder Stelle \mathbf{x} der Ebene \mathbb{R}^2 berechnen.