Musterlösung Serie 6

1. Eine Abbildung $V_1 \times \ldots \times V_n \to W$ heisst multilinear, wenn sie linear in jedem Argument ist. Im Fall $n = 1$ sind dies gerade lineare Abbildungen, im Fall $n = 2$ heissen sie bilinear. (Hier seien die V_j und W endlich dimensionale Vektorräume über $\mathbb{K} = \mathbb{R}$ oder $= \mathbb{C}$.)

   (a) Entscheide, ob die folgenden Abbildungen multilinear sind:

   i. $f\colon \mathbb{R} \to \mathbb{R}, \quad f(x) = x + 1$
   ii. $f\colon \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}, \quad f(X, Y) = AXBYC$, wobei A, B, C (fest gewählte) $(n \times n)$-Matrizen sind.
   iii. $f(x) = ||x||$
   iv. $f(x) = ||x||^2$
   v. $f(x, y, z) = (x, y \times z)$ mit x, y, z in $\mathbb{R}^3$, (,)=Skalarprodukt

   (b) (Polarisationsformel) Jeder Multilinearform $f\colon \overbrace{V \times \ldots \times V}^{n\text{-mal}} \to W$ kann man eine Funktion
   $$F\colon V \to W$$
   zuordnen durch
   $$F(v) = f(v, v, ..., v)$$
   F ist eine polynomiale Funktion, homogen vom Grad n.
   Ist umgekehrt $F\colon V \to W$ polynomial und homogen vom Grad n, so kann man eine symmetrische Multilinearform (die Polarisation von F) definieren durch
   $$\mathrm{Pol}_F\colon V\times\ldots\times V \to W,\ \mathrm{Pol}_F(v_1, \ldots, v_n) = \frac{1}{n!}\frac{\partial}{\partial t_1}\cdots\frac{\partial}{\partial t_n}\bigg|_{t_1=\ldots=t_n=0} F(t_1v_1+\ldots+t_nv_n)$$
   Zeige, dass dann gilt
   $$F(v) = \mathrm{Pol}_F(v, ..., v). \tag{1}$$

   (c) Berechne jeweils den Grad und die Polarisation der folgenden homogenen polynomialen Funktionen

   i. $X \mapsto \mathrm{tr}(X^3)$ für X eine $(n \times n)$-Matrix
   ii. $F\colon \mathbb{R}^2 \to \mathbb{R}, \quad F(x, y) = x^2 + 2xy + y^2$
   iii. $X \mapsto \det(X)$ für X eine $(n \times n)$-Matrix

   (d) Berechne die Ableitungen der Funktionen in 1.(c) i. und ii. unter Benutzung von (1) und der Produktregel aus der Vorlesung.

Lösung:

(a) i. ...ist nicht (multi)linear, unter anderem da $f(0) \neq 0$.
   ii. ...ist bilinear, da für alle $\lambda \in \mathbb{R}$ sowie für alle $M \in \mathbb{R}^{n\times n}$ gilt:
   $f(\lambda X, Y) = \lambda f(X, Y) = f(X, \lambda Y)$
   sowie $f(X + M, Y) = f(X, Y) + f(M, Y)$ und $f(X, Y + M) = f(X, Y) + f(X, M)$.
   iii. ...ist nicht multilinear auf $\mathbb{R}^n$ für $n > 0$, da für einen Vektor x mit $||x|| > 0$ gilt $f(x - x) = f(\overrightarrow{0}) = 0 \neq 2||x|| = f(x) + f(-x)$.

iv. ...ist aus analogen Gründen nicht multilinear auf $\mathbb{R}^n$ für $n > 0$.

v. ...ist als Verknüpfung der multilinearen Abbildungen Skalar- und Vektorprodukt wiederum selbst multilinear.

(b) Berechne:

$$\begin{aligned}
\mathrm{Pol}_F(v,\ldots,v) &= \frac{1}{n!}\frac{\partial}{\partial t_1}\cdots\frac{\partial}{\partial t_n}\Big|_{t_1=\ldots=t_n=0} F(t_1 v + \ldots + t_n v) \\
&= \frac{1}{n!}\frac{\partial}{\partial t_1}\cdots\frac{\partial}{\partial t_n}\Big|_{t_1=\ldots=t_n=0} F((t_1 + \ldots + t_n)v) \\
&= \frac{1}{n!}\frac{\partial}{\partial t_1}\cdots\frac{\partial}{\partial t_n}\Big|_{t_1=\ldots=t_n=0} (t_1 + \ldots + t_n)^n F(v)
\end{aligned}$$

Hier haben wir ausgenutzt, dass $F(v)$ homogen ist vom Grad n, also $F(\lambda v) = \lambda^n F(v)$. Wir rechnen weiter:

$$... = F(v)\,\frac{1}{n!}\frac{\partial}{\partial t_1}\cdots\frac{\partial}{\partial t_n}\Big|_{t_1=\ldots=t_n=0} (t_1 + \ldots + t_n)^n = F(v)\frac{1}{n!}n! = F(v)$$

(c) i. $F(X) = \mathrm{tr}(X^3)$ ist ein Polynom und homogen vom Grad 3. Bei der Berechnung von $F(t_1X_1 + t_2X_2 + t_3X_3)$ interessieren uns nur die Terme, in denen jedes t_i gerade einmal vorkommt. Da die Spur-Abbildung linear ist, erhalten wir

$$\begin{aligned}
\mathrm{Pol}_F(X_1, X_2, X_3) = \frac{1}{6}\,\mathrm{tr}(&X_1X_2X_3 + X_1X_3X_2 + X_2X_1X_3 \\
&+ X_2X_3X_1 + X_3X_1X_2 + X_3X_2X_1).
\end{aligned}$$

ii. $F(x, y) = (x + y)^2$ ist offensichtlich ein Polynom und homogen vom Grad 2. $(F(\lambda x, \lambda y) = \lambda^2 F(x, y))$. Wir berechnen

$$\begin{aligned}
F(t_1x_1 + t_2x_2, t_1y_1 + t_2y_2) &= (t_1(x_1 + y_1) + t_2(x_2 + y_2))^2 \\
&= t_1^2(x_1 + y_1)^2 + 2t_1t_2(x_1 + y_1)(x_2 + y_2) + t_2^2(x_2 + y_2).
\end{aligned}$$

Ableiten nach t_1 und t_2 "wählt" offensichtlich gerade den mittleren Term aus, so dass

$$\mathrm{Pol}_F((x_1, y_1), (x_2, y_2)) = (x_1 + y_1)(x_2 + y_2)$$

iii. Die Determinante ist per Definition ein Polynom in den Matrixeinträgen. Wegen $det(\lambda A) = \lambda^n A$ ist sie homogen vom Grad n.
Seien nun $A_1, \ldots, A_n$ $n \times n$-Matrizen. Wir schreiben $A_i = (a_{i1} \ldots a_{in})$ wobei $a_{ij} \in \mathbb{R}^n$ die j-te Spalte von A_i ist. Wir benutzen aus der linearen Algebra, dass die Determinante linear ist in jeder Spalte separat. Wir erhalten daher

$$\det(t_1A_1 + \ldots + t_nA_n) = \sum_{i_1}\ldots\sum_{i_n} t_{i_1}\cdots t_{i_n} \det(a_{i_1 1}\ldots a_{i_n n})$$

Für die Polarisation suchen wir darin genau die Terme, für die jedes t_i genau einmal vorkommt. Diese Terme sind aber

$$t_1\cdots t_n \sum_\sigma \det(a_{\sigma(1)1}\ldots a_{\sigma(n)n}),$$

wobei die Summe über alle Permutationen σ von $1, \ldots, n$ läuft. Wir erhalten also

$$\mathrm{Pol}_{\det}(A_1, \ldots, A_n) = \frac{1}{n!}\sum_\sigma \det(a_{\sigma(1)1}...a_{\sigma(n)n})$$

(d) i. Wir betrachten zuerst X_2 und X_3 als fix und berechnen das Differential der Abbildung

$$\begin{aligned}\mathrm{Pol}(\cdot, X_2, X_3) &\to \mathbb{R}\\ X_1 &\mapsto \mathrm{Pol}_F(X_1, X_2, X_3)\end{aligned}$$

an einer beliebigen Stelle $X_1 \in \mathbb{R}^{n\times n}$. Wegen der Linearität dieser Abbildung ist dies gegeben durch

$$\begin{aligned}D\,\mathrm{Pol}_F(\cdot, X_2, X_3)_{X_1} : \mathbb{R}^{n\times n} &\to \mathbb{R}\\ A &\mapsto \mathrm{Pol}_F(A, X_2, X_3).\end{aligned}$$

Ebenso sind die Differentiale mit X_2 respektive X_3 als Variable gegeben durch $A \mapsto \mathrm{Pol}_F(X_1, A, X_3)$ und $A \mapsto \mathrm{Pol}_F(X_1, X_2, A)$. Um nun das Differential der Funktion F an einer Stelle X zu berechnen, müssen wir nach der Kettenregel lediglich alle diese Differentiale an der Stelle $X_i = X$ auswerten und addieren:

$$\begin{aligned}DF_X : \mathbb{R}^{n\times n} &\to \mathbb{R}\\ A &\mapsto \mathrm{Pol}_F(A, X, X) + \mathrm{Pol}_F(X, A, X) + \mathrm{Pol}_F(A, X, X)\\ &= 3\,\mathrm{tr}(AX^2 + XAX + X^2A),\end{aligned}$$

wobei wir im letzten Schritt erneut die Linearität der Spur ausgenutzt haben.

ii. Wir schreiben F als Verknüpfung $v \mapsto (v, v) \mapsto \mathrm{Pol}_F(v, v)$. Ableiten liefert die beiden Differentiale

$$\begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\end{pmatrix} \quad \text{und} \quad \begin{pmatrix} x_2 + y_2\\ x_2 + y_2\\ x_1 + y_1\\ x_1 + y_1\end{pmatrix}.$$

Matrixmultiplikation sowie Einsetzen von $x_1 = x_2 = x$, $y_1 = y_2 = y$ liefert wie erwartet $\begin{pmatrix} 2x + 2y\\ 2x + 2y\end{pmatrix}$.

2. (a) Sei $f\colon U \subset \mathbb{R}^m \to \mathbb{C}$ eine C^2-Funktion und $\Psi = (\Psi_1, \dots, \Psi_m)\colon V \subset \mathbb{R}^n \to U$ zweimal stetig differenzierbar. Sei $H_f = f''$ die Hessesche Matrix von f. Zeige, dass für die Hessesche Matrix von $F := f \circ \Psi$ gilt (für $a \in V$):

$$H_F(a) = H_{f\circ\Psi}(a) = (\Psi')^T H_{f(\Psi(a))}(\Psi') + \sum_{j=1}^{m} \partial_j f(\Psi(a)) H_{\Psi_m}(a) \qquad (2)$$

Schreibe ausserdem die obige Formel in Komponenten:

$$\frac{\partial^2}{\partial x_i \partial x_j}(f(\Psi_1(x_1, \dots, x_n), \dots, \Psi_m(x_1, \dots, x_n)) = \dots$$

(b) Die auf $\mathbb{R}^2$ definierte Funktion F sei in Polarkoordinaten gegeben durch $F(r, \phi) = r^3 e^{i\phi}$. Finde einen Ausdruck $f(x, y)$ für diese Funktion in Euklidischen Koordinaten. D. h. finde $f(x, y)$ so, dass $F(r, \phi) = f(r\cos\phi, r\sin\phi)$, also $F = f \circ \Psi$ mit $\Psi(r, \phi) = (r\cos\phi, r\sin\phi)$.

Berechne dann die Hesseschen Matrizen von f und F in $(x, y) = (1, 0)$ (bzw. $(r, \phi) = (1, 0)$). Setzte das Ergebnis in (2) ein und verifiziere so Deine Rechnung.

Lösung:

(a) Wir schreiben die Komponenten in U als $y_1, \dots, y_m$, d.h. $f = f(y_1, \dots, y_m)$. Wir bezeichnen mit $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_m)$ die Ortvektoren, wo benötigt. Zuerst berechnen wir mithilfe der Kettenregel die komponentenweise Darstellung. Es gilt zuerst für die erste Ableitung:

$$\begin{aligned}
\frac{\partial}{\partial x_i}(f \circ \Psi)(x) &= \frac{\partial}{\partial x_i} f(\Psi(x_1, \dots, x_n), \dots, \Psi_m(x_1, \dots, x_n)) \\
&= \frac{\partial f}{\partial y_1}(\Psi(x))\frac{\partial \Psi_1}{\partial x_i} + \dots \frac{\partial f}{\partial y_m}(\Psi(x))\frac{\partial \Psi_m}{\partial x_i} \\
&= \sum_{k=1}^{m} \frac{\partial f}{\partial y_k}(\Psi(x))\frac{\partial \Psi_k}{\partial x_i}
\end{aligned}$$

und damit dann für die zweite Ableitung, mit Produkt- und Kettenregel:

$$\frac{\partial^2}{\partial x_i \partial x_j}(f \circ \Psi)(x) = \sum_{k,l=1}^{m} \frac{\partial^2 f}{\partial y_k \partial y_l}(\Psi(x))\frac{\partial \Psi_k}{\partial x_i}\frac{\partial \Psi_l}{\partial x_j} + \sum_{k=1}^{m} \frac{\partial f}{\partial y_k}(\Psi(x))\frac{\partial^2 \Psi_k}{\partial x_i \partial x_j}.$$

Damit folgt auch der erste Teil der Aufgabe, indem wir die entsprechenden Matrizen komponentenweise schreiben.

(b) Wegen der Eulerschen Formel $e^{i\phi} = \cos\phi + i\sin\phi$ ist hier $F(r, \phi) = r^3 e^{i\phi} = r^2(r\cos\phi + ir\sin\phi)$, also $f(x, y) = (x^2 + y^2)(x + iy)$ Wir berechnen zunächst die Hesseschen von f und F direkt

$$H_f(x, y) = \begin{pmatrix} 6x^2 + 2iy & 2(ix + y) \\ 2(ix + y) & 6iy + 2x \end{pmatrix} \qquad H_F(r, \phi) = e^{i\phi}\begin{pmatrix} 6r & 3ir^2 \\ 3ir^2 & -r^3 \end{pmatrix}$$

Auswerten in $(r, \phi) = (1, 0)$ bzw. $(x, y) = \Psi(1, 0) = (1, 0)$ ergibt:

$$H_f(1, 0) = \begin{pmatrix} 6 & 2i \\ 2i & 2 \end{pmatrix} \qquad H_F(1, 0) = \begin{pmatrix} 6 & 3i \\ 3i & -1 \end{pmatrix}$$

Um Formel (2) zu überprüfen, brauchen wir noch folgende Berechnungen, wobei wir jeweils zuerst ableiten und dann bei den Stellen $(x, y) = (1, 0)$ respektive $(r, \phi) = (1, 0)$ auswerten:

$$\begin{aligned}
\Psi' &= \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
H_{\Psi_1} &= \begin{pmatrix} 0 & -\sin\phi \\ -\sin\phi & -r\cos\phi \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \\
H_{\Psi_2} &= \begin{pmatrix} 0 & \cos\phi \\ \cos\phi & -r\sin\phi \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\partial_x f &= 2x(x + iy) + (x^2 + y^2) = 3 \\
\partial_y f &= 2y(x + iy) + i(x^2 + y^2) = i.
\end{aligned}$$

Einsetzen in die in die Formel (2) ergibt dann

$$\begin{pmatrix} 6 & 3i \\ 3i & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 6 & 2i \\ 2i & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} + i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

und man sieht, dass dies korrekt ist.

3. Sei

$$f(x, y) := \frac{\sin(x^2\tan(y))}{1 + x^2} \quad \text{und} \quad g(x, y) := \begin{cases} \frac{xy^3}{x^2+y^2} & \text{falls } (x, y) \neq (0, 0) \\ 0 & \text{falls } (x, y) = (0, 0) \end{cases}$$

(a) Berechne $\partial_{xx} f$, $\partial_{yy} f$, $\partial_{xy} f$ und $\partial_{yx} f$.

(b) Berechne $\partial_{xx} g$, $\partial_{yy} g$, $\partial_{xy} g$ und $\partial_{yx} g$. Untersuche insbesondere den Punkt $(0,0)$.

(c) Begründe das Resultat aus (b).

Lösung:

Erinnern wir uns an den Satz von Schwarz:

Satz 1 *Die Funktion $f(x,y)$ besitze in $(x_0,y_0) \in \mathbb{R}^2$ die partiellen Ableitungen $\partial_x f(x_0,y_0)$, $\partial_y f(x_0,y_0)$ und $\partial_{xy} f(x_0,y_0)$. Sind $\partial_x f$, $\partial_y f$ und $\partial_{xy} f$ stetig in (x_0,y_0), so existiert auch $\partial_{yx} f(x_0,y_0)$ und es gilt*

$$\partial_{yx} f(x_0,y_0) = \partial_{xy} f(x_0,y_0).$$

(a) Sei $f(x,y) := \frac{\sin(x^2\tan(y))}{1+x^2}$. Wir berechnen für alle $(x,y) \in \mathbb{R}^2$

$$\begin{aligned}
\partial_x f(x,y) &= \frac{2x\tan(y)(1+x^2)\cos(x^2\tan(y)) - 2x\sin(x^2\tan(y))}{(1+x^2)^2};\\
\partial_y f(x,y) &= \frac{\cos(x^2\tan(y))(1+\tan^2(y))x^2}{1+x^2}.
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
\partial_{xx} f(x,y) &= \frac{8x^2\sin\left(x^2\tan(y)\right)}{(1+x^2)^3} - \frac{2\sin\left(x^2\tan(y)\right)}{(1+x^2)^2} - \frac{8x^2\cos\left(x^2\tan(y)\right)\tan(y)}{(1+x^2)^2}\\
&\quad + \frac{2\cos\left(x^2\tan(y)\right)\tan(y)}{1+x^2} - \frac{4x^2\sin\left(x^2\tan(y)\right)\tan^2(y)}{1+x^2},\\
\partial_{yy} f(x,y) &= \frac{-\sin(x^2\tan(y))(1+\tan(y))^2x^4 + 2x^2\cos(x^2\tan(y))\tan(y)(1+\tan^2(y))}{1+x^2}
\end{aligned}$$

und

$$\begin{aligned}
\partial_{yx} f(x,y) = (1+\tan^2(y))\Bigg[&\frac{-2x^3\tan(y)\sin(x^2\tan(y)) + 2x\cos(x^2\tan(y))}{1+x^2}\\
&-\frac{2x^3\cos(x^2\tan(y)}{(1+x^2)^2}\Bigg]
\end{aligned}$$

Mithilfe von Satz 1 folgt

$$\partial_{yx} f(x,y) = \partial_{xy} f(x,y)$$

(b) Sei nun

$$g(x,y) := \begin{cases} \frac{xy^3}{x^2+y^2} & \text{falls } (x,y) \neq (0,0)\\ 0 & \text{falls } (x,y) = (0,0)\end{cases}$$

Für $(x,y) \neq (0,0)$ berechnen wir

$$\begin{aligned}
\partial_x g(x,y) &= \frac{(x^2+y^2)y^3 - 2x^2y^3}{(x^2+y^2)^2} = \frac{y^5 - x^2y^3}{(x^2+y^2)^2};\\
\partial_y g(x,y) &= \frac{3y^2(x^2+y^2) - 2xy^4}{(x^2+y^2)^2}.
\end{aligned}$$

Daraus berechnen wir

$$\begin{aligned}
\partial_{xx} g(x,y) &= \frac{-2xy^3(x^2+y^2)^2 - 4x(x^2+y^2)\left[y^5 - x^2y^3\right]}{(x^2+y^2)^4}\\
\partial_{yy} g(x,y) &= \frac{(x^2+y^2)^2\left[6yx^2 + 12y^3 - 8x - y^3\right] - 4y(x^2+y^2)\left[3y^2(x^2+y^2) - 2xy^4\right]}{(x^2+y^2)^4}
\end{aligned}$$

Für $(x,y) \neq (0,0)$ folgt weiter mit Satz 1

$$\partial_{yx} g(x,y) = \partial_{xy} f(x,y) = \frac{(x^2+y^2)^2\left[5y^4 - 3x^2y^2\right] - 4y(x^2+y^2)\left[y^5 - x^2y^3\right]}{(x^2+y^2)^4}.$$

Es bleibt der Punkt $(0,0)$ zu untersuchen. Für $x = 0$ und $y \neq 0$ folgt

$$\partial_x g(x,y) = \lim_{h\to 0} \frac{g(h,y) - g(0,0)}{h} = \lim_{h\to 0} \frac{y^3}{h^2+y^2} = y.$$

Für $(0,0)$ folgt

$$\partial_x g(0,0) = \lim_{h\to 0} \frac{g(h,0) - g(0,0)}{h} = 0.$$

Zusammenfassend folgt also

$$\partial_x g(0,y) = y \quad \forall\, y \in \mathbb{R}.$$

Sei nun umgekehrt $y = 0$ und x beliebig. Wir berechnen, falls $x \neq 0$

$$\partial_y g(x,0) = \lim_{h\to 0} \frac{g(x,h) - g(x,0)}{h} = \lim_{h\to 0} \frac{xh^2}{x^2+h^2} = 0$$

Im fall $x = 0$ folgt

$$\partial_y g(0,0) = \lim_{h\to 0} \frac{g(0,h) - g(0,0)}{h} = 0.$$

Also folgt

$$\partial_y g(x,0) = 0.$$

Berechnen wir nun die Ableitungen 2. Ordnung in $(0,0)$ folgt sofort

$$\partial_{xy} g(0,0) = 1 \neq 0 = \partial_{yx} g(0,0).$$

(c) Der Satz von Schwarz lässt sich offensichtlich nicht für beliebige Funktionen verwenden. In unserem Fall ist die partielle Ableitung $\partial_y g(x,y)$ offensichtlich unstetig in $(0,0)$, denn

$$\lim_{x\to 0} \partial_y g(x,0) = 0 \neq 3 = \lim_{y\to 0} \partial_y g(0,y).$$

Also müssen alle Vorraussetzungen von Satz 1 erfüllt sein, damit man ihn anwenden kann. Ist eine Vorraussetzung nicht erfüllt hat die Aussage nicht unbedingt Gültigkeit, wie die Funktion in b) beweist.

4. Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f\colon U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Ein kritischer Punkt x_0 von f heisst *nicht degeneriert*, falls $\det(f''(x_0)) \neq 0$ gilt. Zeige, dass alle nicht degenerierten kritischen Punkte isoliert sind. d. h. jeder nicht degenerierte kritische Punkt besitzt eine Umgebung, die keine weiteren kritischen Punkte enthält.

Tipp: *Wende das Inverse Funktionentheorem auf die Funktion* $\nabla f\colon U \to \mathbb{R}^n$ *an.*

Lösung:

Die Ableitung der Funktion $g(x) := \nabla f(x)$ ist gegeben durch

$$\begin{aligned} dg(x) &= \begin{pmatrix} \partial_{x_1} g_1(x) & \partial_{x_2} g_1(x) & \cdots & \partial_{x_n} g_1(x) \\ \partial_{x_1} g_2(x) & \partial_{x_2} g_2(x) & \cdots & \partial_{x_n} g_2(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_n(x) & \partial_{x_2} g_n(x) & \cdots & \partial_{x_n} g_n(x) \end{pmatrix} \\ &= \begin{pmatrix} \partial_{x_1}\partial_{x_1} f(x) & \partial_{x_2}\partial_{x_1} f(x) & \cdots & \partial_{x_n}\partial_{x_1} f(x) \\ \partial_{x_1}\partial_{x_2} f(x) & \partial_{x_2}\partial_{x_2} f(x) & \cdots & \partial_{x_n}\partial_{x_2} f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1}\partial_{x_n} f(x) & \partial_{x_2}\partial_{x_n} f(x) & \cdots & \partial_{x_n}\partial_{x_n} f(x) \end{pmatrix} = d^2 f(x) \end{aligned}$$

Sei nun $x_0 \in U$ ein nicht degenerierter kritischer Punkt von f. Aus $\det(d^2 f(x_0)) \neq 0$ folgt mit obiger Rechnung, dass $dg(x_0)$ invertierbar ist. Wir können also das Inverse Funktionentheorem auf g an der Stelle x_0 anwenden. Dieses besagt, dass offene Umgebungen $x_0 \in V \subset U$ und $0 = g(x_0) \in W \subset \mathbb{R}^n$ exisiteren, sodass die Einschränkung

$$g|_V : V \to W$$

ein C^1-Diffeomorphismus ist. Insbesondere ist die Einschränkung $g|_V$ injektiv und für alle $x \in V$ mit $x \neq x_0$ gilt

$$\nabla f(x) = g(x) \neq g(0) = 0.$$

Folglich besitzt f in V keinen anderen kritischen Punkt ausser x_0 und das zeigt die Behauptung.

5. Sei $A = (a_{ij})_{1 \leq i,j \leq n}$ eine symmetrische, reelle $n \times n$ Matrix. A heisst positiv definit wenn für alle $v \neq 0$ in $\mathbb{R}^n$ gilt: $v^T A v > 0$, und negativ definit wenn für alle $v \neq 0$ in $\mathbb{R}^n$ gilt $v^T A v < 0$. (Dies ist äquivalent dazu, dass alle Eigenwerte von A positiv, bzw. negativ sind.) Definitheit kann man mit dem folgenden Kriterium überprüfen.

   (Sylvesters Kriterium) Sei $A_k := (a_{ij})_{1 \leq i,j \leq k}$ der obere $k \times k$ Block von A, und sei $d_k = \det(A_k)$ für $k = 1, \ldots, n$.
   (1) Dann ist A positiv definit genau dann wenn $d_k > 0$ für $k = 1, \ldots, n$.
   (2) A ist negativ definit genau dann wenn $d_1 < 0$, $d_2 > 0$, $d_3 < 0$, $d_4 > 0$, etc. gilt.

   (a) Benutze das Kriterium um die folgenden Matrizen auf positive bzw. negative Definitheit zu überprüfen.
      i. $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$
      ii. $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$
      iii. $\begin{pmatrix} -1 & x \\ x & -1 \end{pmatrix}$, in Abhängigkeit von x.
      iv. $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$, in Abhängigkeit von a, b, c.

   (b) Zeige (2) in Sylvesters Kriterium unter Benutzung von (1).

   (c) Zeige, dass eine positiv definite Matrix positive Determinante hat ohne Sylvesters Kriterium anzuwenden. (Hinweis: Wie ist die Determinante durch die Eigenwerte ausgedrückt?)

   (d) Zeige, dass falls A positiv definit ist, A_k auch positiv definit ist und folgere damit und mit 5c) die Richtung "$\Rightarrow$" von (1).

   (e) Beweise Sylvesters Kriterium (es fehlt noch Aussage (1), "$\Leftarrow$") durch Induktion über n. Gehe für den Induktionsschritt $n \to n+1$ wie folgt vor:
      i. Wir schreiben die $(n+1) \times (n+1)$-Matrix A als $\begin{pmatrix} A_n & a \\ a^T & b \end{pmatrix}$. Wir setzen $v := -A_n^{-1} a$ und $P = \begin{pmatrix} \mathbb{1}_{n \times n} & v \\ 0 & 1 \end{pmatrix}$. Zeige, dass dann gilt $P^T A P = \begin{pmatrix} A_n & 0 \\ 0 & c \end{pmatrix}$ für ein $c \in \mathbb{R}$.
      ii. Zeige, dass $c > 0$. (Hinweis: $\det(A) = \det(P^T A P)$.)
      iii. Folgere daraus, dass A positiv definit sein muss.

   Schlussbemerkung: Sylvesters Kriterium wie oben ist verwendbar zur Bestimmung der Definitheit, nicht aber für die Semidefinitheit einer Matrix. Mit einer leichten Abwandlung (alle Hauptminoren haben Determinante ≥ 0) kann man auch die positive Semidefinitheit zeigen, dies soll jedoch hier ausser Acht gelassen werden.

   Lösung:

(a) i. $d_1 = 2$, $d_2 = \det A = 1$ und damit ist die Matrix positiv definit.

ii. $d_1 = 2$, $d_2 = 5$, $d_3 = \det A = 4$ und damit ist die Matrix positiv definit.

iii. $d_1 = -1$, $d_2 = 1 - x^2$. Die Matrix ist in keinem Fall positiv definit, sie kann aber negativ definit sein und zwar genau dann wenn $d_2 > 0$, d.h. $1 - x^2 > 0$ respektive $x \in (-1, 1)$.

iv. $d_1 = a$, $d_2 = ac - b^2$. Damit die Matrix positiv definit ist muss gelten, dass $d_1 = a > 0$ und $d_2 = ac - b^2 > 0$, d.h. bei fixiertem $a > 0$ und beliebigem b muss für c gelten: $c > \frac{b^2}{a}$. Damit die Matrix negativ definit ist, muss gelten, dass $a < 0$ und wiederum bei beliebigem b: $c < \frac{b^2}{a}$ (beachte das Vorzeichen, da $a < 0$).

(b) Beachte, dass wenn A die Eigenwerte $\lambda_1, \ldots, \lambda_n$ hat, so hat $-A$ die Eigenwerte $-\lambda_1, \ldots, -\lambda_n$ und damit ist eine Matrix negativ definit genau dann, wenn das Negative der Matrix positiv definit ist. Die Aussage folgt dann, wenn man verwendet, dass für eine k-dimensionale Matrix B und eine reelle Zahl r gilt $\det(rB) = r^k \det B$.

(c) Sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte einer Matrix A, so gilt $\det A = \lambda_1 \cdot \ldots \cdot \lambda_n$ und wenn alle Eigenwerte positiv sind, so ist diese Determinante positiv.

(d) Ist A positiv definit, so bedeutet dies, dass $v^T Av > 0$ für alle $v \neq 0$. Dies gilt insbesondere für alle v, die nur in den ersten k Einträgen nicht verschwindende Einträge haben und sonst 0 sind. Wenn wir einen solchen Vektor $v = (w, 0, \ldots, 0)$ betrachten, wobei $w \in \mathbb{R}^k$, so ist schnell ersichtlich, dass $v^T Av = w^T A_k w$. Diese sind alle positiv, wenn A positiv ist und damit ist A_k positiv definit. Mit Aufgabe 5c) folgt dann, dass $\det A_k > 0$. Da dies für alle $k \leq n$ gilt, haben wir "$\Rightarrow$" von (1) gezeigt.

(e) i. Mit direktem Rechnen ergibt sich:

$$P^T AP = \begin{pmatrix} A_n & 0 \\ v^T A_n + a^T & v^T A_n v + v^T a + a^T v + b \end{pmatrix}.$$

Setzen wir die Definition von v ein und nutzen aus, dass $A^T = A$, dann sehen wir, dass $v^T A_n + a^T = -a^T A_n^{-T} A_n + a^T = -a^T A_n^{-1} A_n + a^T = 0$. Den Wert im letzten Eintrag der Matrix definieren wir als c.

ii. Mit dem Hinweis gilt $0 < \det A = \det P^T AP = c \det A_n = cd_n$. Nach Voraussetzung ist $d_n > 0$ und damit muss auch $c > 0$ gelten, damit $cd_n > 0$.

iii. Um zu zeigen, dass A positiv definit ist, müssen alle Eigenwerte positiv definit sein. Mit der Induktionsvoraussetzung wissen wir, dass A_n positiv definit ist (da die Determinanten der Minoren von A_n genau die $d_1, \ldots, d_n$ sind) und damit positive Eigenwerte $\lambda_1, \ldots, \lambda_n$ hat. Nach der Aufteilung im vorherigen Teilschritt wissen wir damit, dass die Eigenwerte von A genau $\lambda_1, \ldots, \lambda_n, c$ sind und diese sind alle positiv.