**Wirtschaftswissenschaftliches Zentrum**
**Universität Basel**

5

**Mathematik 2**

**Dr. Thomas Zehrt**

Quadratische Formen und Definitheit

# Inhaltsverzeichnis

# 1 Quadratische Formen

Die Koordinatenachsen des $\mathbb{R}^n$ werden wieder mit $x_1, x_2, \ldots, x_n$ bezeichnet.

**Definition 1.1** *Sei $A$ eine symmetrische $(n \times n)$-Matrix. Dann heisst die Funktion*

$$Q_A : \mathbb{R}^n \longrightarrow \mathbb{R}$$

*mit*

$$Q_A(\mathbf{x}) \;=\; \mathbf{x} \bullet A\mathbf{x} \;=\; \mathbf{x}^T A\mathbf{x} \;=\; \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j$$

*die zu $A$ gehörige quadratische Form.*

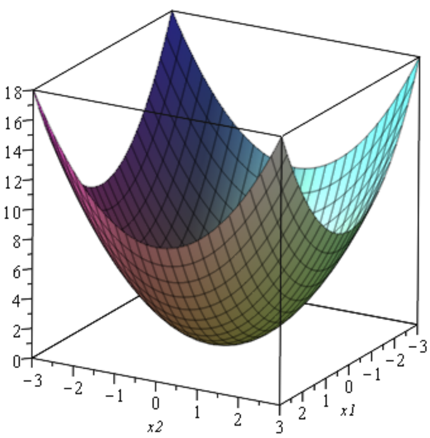
Bemerkungen:

- $Q_A(\mathbf{x})$ ist das Skalarprodukt der Vektoren $\mathbf{x}$ und $A\mathbf{x}$.
- $Q_A$ ist die Funktion, die jedem Vektor $\mathbf{x}$ das Skalarprodukt der Vektoren $\mathbf{x}$ und $A\mathbf{x}$ zuordnet.

**Beispiel 1.1**

$$A \;=\; I \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

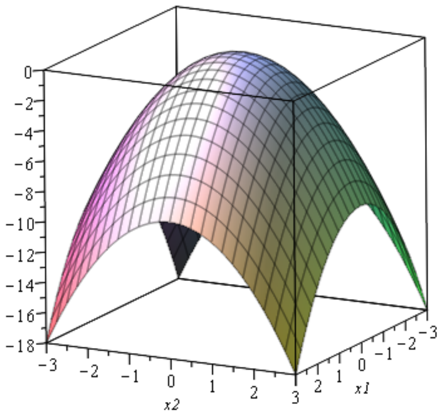
$$Q_A(\mathbf{x}) \;=\; \mathbf{x}^T \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} \;=\; (x_1, x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; (x_1, x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; x_1^2 + x_2^2.$$



**Beispiel 1.2**

$$B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

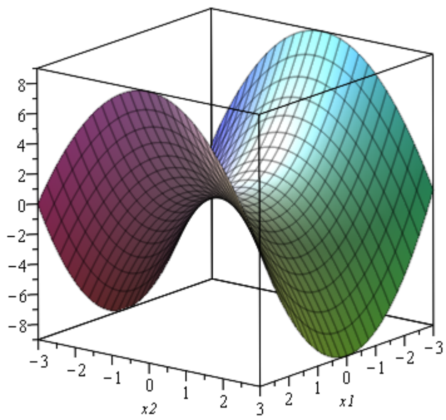
$$Q_B(\mathbf{x}) = (x_1, x_2) \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} = -x_1^2 - x_2^2.$$




**Beispiel 1.3**

$$C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

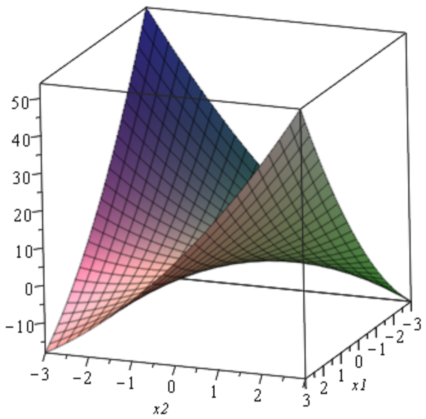
$$Q_C(\mathbf{x}) = (x_1, x_2) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = x_1^2 - x_2^2.$$



**Beispiel 1.4**

$$D \;=\; \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

$$\begin{aligned} Q_D(\mathbf{x}) &= \mathbf{x}^T \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \mathbf{x} \\ &= (x_1, x_2) \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1, x_2) \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 \end{pmatrix} \\ &= x_1(x_1 + 2x_2) + x_2(2x_1 + x_2) \;=\; x_1^2 + 4x_1x_2 + x_2^2. \end{aligned}$$




Einer Diagonalmatrix

$$D \;=\; \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

kann man unmittelbar und ohne grossen Aufwand ihre quadratische Form zuordnen. Insbesondere kommen hier **keine** gemischten Terme $x_i x_j$ vor!

$$Q_D(\mathbf{x}) \;=\; \mathbf{x}^T\, D\, \mathbf{x} \;=\; \sum_{i=1}^{n} d_i\, x_i^2 \;=\; d_1\, x_1^2 + d_2\, x_2^2 + \ldots + d_n\, x_n^2$$

# 2 Definitheit (symmetrischer) Matrizen

**Definition 2.1** *Sei $A$ eine (symmetrische) $(n \times n)$-Matrix. Dann heisst $A$*

- *positiv definit, falls $Q_A(\mathbf{x}) > 0$*
- *positiv semidefinit, falls $Q_A(\mathbf{x}) \geq 0$*
- *negativ definit, falls $Q_A(\mathbf{x}) < 0$ (oder falls $-A$ positiv definit ist)*
- *negativ semidefinit, falls $Q_A(\mathbf{x}) \leq 0$ (oder falls $-A$ positiv semidefinit ist)*

*für alle Vektoren $\mathbf{x} \neq \mathbf{0}$ gilt. Die Matrix $A$ heisst indefinit, wenn es sowohl Vektoren $\mathbf{x}$ mit $Q_A(\mathbf{x}) > 0$ als auch Vektoren $\mathbf{y}$ mit $Q_A(\mathbf{y}) < 0$ gibt.*

Versuchen Sie die folgenden Aussagen zu verstehen:

1. Für alle Matrizen $A$ gilt $Q_A(\mathbf{0}) \;=\; \mathbf{0}^T A\, \mathbf{0} = 0$.
2. Eine positiv (bzw. negativ) definite symmetrische Matrix ist auch positiv (bzw. negativ) semidefinit. Die Umkehrung gilt nicht.
3. Wenden wir die quadratische Form auf einen Standardbasisvektor $\mathbf{e}_i$ $(i = 1, 2, \ldots, n)$ an, erhalten wir den Diagonaleintrag $a_{ii}$ ($i$-te Zeile und $i$-te Spalte) von $A$:

$$\boxed{Q_A(\mathbf{e}_i) \;=\; \mathbf{e}_i^T A\, \mathbf{e}_i \;=\; a_{ii}}$$

   Daraus ergeben sich einige hilfreiche Implikationen (die Umkehrungen sind im Allgemeinen **falsch**!):

   - $A$ positiv definit $\longrightarrow$ alle Diagonaleinträge $> 0$
   - $A$ negativ definit $\longrightarrow$ alle Diagonaleinträge $< 0$
   - $A$ positiv semidefinit $\longrightarrow$ alle Diagonaleinträge $\geq 0$
   - $A$ negativ semidefinit $\longrightarrow$ alle Diagonaleinträge $\leq 0$
   - $A$ hat sowohl positive als auch negative Diagonaleinträge $\longrightarrow$ $A$ indefinit

4. Falls wir die Eigenwerte von $A$ kennen sollten, kennen wir auch die Definitheit.

   **Satz 1 (Sylvester, $n \times n$)** *Eine symmetrische $n \times n$-Matrix $A$ ist*

   - *positiv definit, falls alle Eigenwerte grösser Null sind;*
   - *positiv semidefinit, falls alle Eigenwerte grösser oder gleich Null sind;*
   - *negativ definit, falls alle Eigenwerte kleiner Null sind;*
   - *negativ semidefinit, falls alle Eigenwerte kleiner oder gleich Null sind;*
   - *indefinit, falls positive und negative Eigenwerte existieren.*

## 2.1 Hurwitz für $(n \times n)$-Matrizen

Wir betrachten die symmetrische $(2 \times 2)$-Matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Können wir anhand der Matrixeinträge eine Aussage über die Definitheit von $A$ machen?

**Satz 2 (Hurwitz, $2 \times 2$)** *Eine symmetrische* $(2 \times 2)$*-Matrix*

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

*ist genau dann*

- *positiv definit, wenn sowohl* $a > 0$ *als auch* $\det(A) > 0$*;*
- *negativ definit, wenn sowohl* $a < 0$ *als auch* $\det(A) > 0$*;*
- *indefinit, wenn* $\det(A) < 0$*.*

**Beweisskizze:** Wir beweisen die Behauptung nur für den Fall $a \neq 0$. Dann gilt

$$\begin{aligned}
Q_A(\mathbf{x}) &= (x_1, x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a\, x_1^2 + 2\, b\, x_1\, x_2 + c\, x_2^2 \\
&= a \left( x_1^2 + 2\, \frac{b}{a}\, x_1\, x_2 \right) + c\, x_2^2 \\
&= a \left( x_1^2 + 2 \frac{b}{a}\, x_1\, x_2 + \left( \frac{b}{a}\, x_2 \right)^2 - \left( \frac{b}{a}\, x_2 \right)^2 \right) + c\, x_2^2 \\
&= a \left( \left( x_1 + \frac{b}{a}\, x_2 \right)^2 - \frac{b^2}{a^2}\, x_2^2 \right) + c\, x_2^2 \\
&= a \left( x_1 + \frac{b}{a}\, x_2 \right)^2 - \frac{b^2}{a}\, x_2^2 + c\, x_2^2 \\
&= a \underbrace{\left( x_1 + \frac{b}{a}\, x_2 \right)^2}_{\geq 0} + \left( c - \frac{b^2}{a} \right) \underbrace{x_2^2}_{\geq 0}
\end{aligned}$$

Wann ist nun $Q_A(\mathbf{x})$ für <u>alle</u> $\mathbf{x} \neq \mathbf{0}$ positiv? Genau dann, wenn die folgenden Kriterien erfüllt sind:

- $0 < a$ (setze $x_2 = 0$ und $x_1 \neq 0$)
- $0 < c - \dfrac{b^2}{a} = \dfrac{ac - b^2}{a} = \dfrac{\det(A)}{a}$ (setze $x_1 = -\dfrac{b}{a}\, x_2$)

$\square$

**Verallgemeinerung** Sei A eine symmetrische $(n \times n)$-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Für alle $k = 1, 2, \dots, n$ definieren wir jeweils eine $(k \times k)$-Matrix $A_k$ wie folgt

$$A_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \dots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

Das ist die quadratische Matrix aus der "linken oberen Eck" von $A$.

Wir bezeichnen die Determinante von $A_k$ als den $k$-ten führenden Hauptminor von $A$.

**Satz 3 (Hurwitz, $n \times n$)** *Die symmetrische Matrix $A$ ist genau dann positiv definit, wenn für alle führenden Hauptminoren*

$$\det(A_k) > 0 \qquad (k = 1, 2, \dots, n)$$

*gilt.*

Bemerkung:

- Es müssen also $n$ Determinanten berechnet werden.
- $A$ ist negativ definit, wenn $-A$ positiv definit ist.
  Achtung: Das heisst **nicht**, dass $\det(A_k) < 0$ für alle $k$.

**Beispiel 2.1** *Die Matrix* $A = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ *ist positiv definit, denn für die führenden Hauptminoren gilt:*

$$\det(A_1) = \det(4) = 4 > 0$$

$$\det(A_2) = \det\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} = 12 > 0$$

$$\det(A_3) = \det\begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} = 60 > 0$$

## 2.2 Semi-Hurwitz für $(n \times n)$-Matrizen

Achtung:

Man kann im Satz von Hurwitz nicht einfach $>$ durch $\geq$ ersetzen und dann auf positiv semidefinit schliessen!!!!

**Beispiel 2.2** *Für die Matrix*

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

*gilt offensichtlich* $a = 0 \geq 0$ *und* $\det(A) = 0 \geq 0$*! Ist $A$ positiv semidefinit? Nein, denn offensichtlich ist die zugehörige quadratische Form* $Q_A(\mathbf{x}) = -x_2^2 \leq 0$ *negativ semidefinit.*

Sei A eine symmetrische $(n \times n)$-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Es müssen alle $(k \times k)$-Untermatrizen von $A$ überprüft werden, die aus $A$ entstehen, wenn man jeweils die selben $n - k$ Spalten und Zeilen aus $A$ streicht.

Die Determinanten dieser Matrizen nennt man die k-ten Hauptminoren von $A$.

**Satz 4 (Semi-Hurwitz, $n \times n$)** *Die symmetrische Matrix $A$ ist genau dann positiv semidefinit, wenn alle Hauptminoren (einschliesslich führender Hauptminoren) von $A$ grösser oder gleich Null sind.*

**Aufgabe 2.1** *Untersuchen Sie die Matrix*

$$A = \begin{pmatrix} 3 & 0 & 9 \\ 0 & 2 & 6 \\ 9 & 6 & 45 \end{pmatrix}$$

*auf positive Semidefinitheit. Ist A positiv definit?*

**Lösung:** Die Matrix ist positiv semidefinit. Beweisen Sie das mit dem Kriterium von Hurwitz.

Falls Sie das quadratische Ergänzen beherrschen, können Sie auch einen direkten Beweis führen. Es gilt:

$$\begin{aligned} Q_A(\mathbf{x}) &= \underline{3x_1^2 + 18x_1x_3 + 45x_3^2} + 2x_2^2 + 12x_2x_3 \\ &= \underline{3(x_1 + 3x_3)^2} + 18x_3^2 + 2x_2^2 + 12x_2x_3 \\ &= 3(x_1 + 3x_3)^2 + \underline{18x_3^2 + 2x_2^2 + 12x_2x_3} \\ &= 3(x_1 + 3x_3)^2 + \underline{2(x_2 + 3x_3)^2} \end{aligned}$$

Natürlich ist dieser Term stets $\geq 0$ und er ist genau gleich 0, wenn $x_1 + 3x_3 = 0$ und $x_2 + 3x_3 = 0$ gilt. Lösen wir dieses lineare Gleichungssystem (3 Unbekannte und 2 Gleichungen) erhalten wir für die Nullstellenmenge der quadratischen Form die Gerade:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = t \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix}.$$

# 3 Quadratische Approximation von Funktionen

## 3.1 Taylor-Polynome (Wiederholung)

In der Vorlesung Mathematik 1 hatten wir (genügend oft differenzierbare) reelle Funktionen $y = f(x)$ in der Nähe eines Punktes $a$ durch das Taylor-Polynom $y = P_2(x)$ approximiert:

$$P_2(x) \quad = \quad f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2$$

Es gilt:

$$\begin{aligned} P_2(a) &= f(a) \\ P_2'(a) &= f'(a) \\ P_2''(a) &= f''(a) \end{aligned}$$

Insbesondere glauben wir, dass sich $P_2$ und $f$ nahe bei $a$ sehr ähnlich sind. Das quadratische Polynom $P_2$ sollte also die Krümmung von $f$ nahe bei $a$ (meist) recht gut beschreiben und wir können tatsächlich das wohl bekannte Kriterium zur Unterscheidung zwischen lokalen Maximalstellen und lokalen Minimalstellen ableiten.

Sei also $a$ ein stationärer Punkt von $f$, d.h. $f'(a) = 0$. Dann gilt für alle $x$ nahe bei $a$:

$$f(x) \;\approx\; P_2(x) \quad = \quad f(a) + \frac{f''(a)}{2}\underbrace{(x-a)^2}_{>0}$$

und man erkennt sofort:

- ist $f''(a) > 0$, so ist $\frac{f''(a)}{2}(x-a)^2 > 0$ und

  $$f(x) \quad \approx \quad f(a) + \underbrace{\frac{f''(a)}{2}(x-a)^2}_{>0} \;>\; f(a)$$

  für **alle** $x$ nahe bei $a$. Also muss $a$ eine lokale Minimalstelle sein.

- ist $f''(a) < 0$, so ist $\frac{f''(a)}{2}(x-a)^2 < 0$ und

  $$f(x) \quad \approx \quad f(a) + \underbrace{\frac{f''(a)}{2}(x-a)^2}_{<0} \;<\; f(a)$$

  für **alle** $x$ nahe bei $a$. Also muss $a$ eine lokale Maximalstelle sein.

- ist $f''(a) = 0$ können wir zunächst nichts aussagen! Wir müssten höhere Ableitungen von $f$ untersuchen.

## 3.2 Verallgemeinerte Taylor-Polynome

**Ziel:** Eine genügend oft differenzierbare Funktion $y = f(x_1, x_2) = f(\mathbf{x})$ mit $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ soll in der Nähe eines Punktes $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ durch eine quadratische Funktion $y = P(\mathbf{x})$ approximiert werden.

**Beachte:**

$$\mathbf{x} - \mathbf{a} = \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \end{pmatrix} \qquad \text{und} \qquad (\mathbf{x} - \mathbf{a})^T = (x_1 - a_1 \;\; x_2 - a_2)$$

**Ansatz:** Die allgemeine quadratische Funktion in zwei Variablen lässt sich mit zunächst beliebigen reellen Zahlen $c_1, c_2, c_{11}, c_{12}, c_{21}$ und $c_{22}$ (da wir diese Parameter in einem Vektor bzw. einer Matrix zusammenfassen wollen, wählen wir schon hier die Indexnotation) wie folgt beschreiben. Auch hier ist es praktischer, die Darstellung mit dem Entwicklungspunkt $\mathbf{a}$ zu wählen.

$$\begin{aligned}
& P(\mathbf{x}) \\
= \;& c_0 + c_1(x_1 - a_1) + c_2(x_2 - a_2) + c_{11}(x_1 - a_1)^2 + c_{22}(x_2 - a_2)^2 + 2c_{12}(x_1 - a_1)(x_2 - a_2) \\
= \;& c_0 + \underbrace{(c_1 \;\; c_2)}_{\mathbf{c}^T} \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \end{pmatrix} + (x_1 - a_1 \;\; x_2 - a_2) \underbrace{\begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}}_{C} \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \end{pmatrix} \\
= \;& c_0 + \mathbf{c}^T \, (\mathbf{x} - \mathbf{a}) + (\mathbf{x} - \mathbf{a})^T \, C \, (\mathbf{x} - \mathbf{a})
\end{aligned}$$

**Forderungen:** Im Punkt $\mathbf{a}$ sollen möglichst viele Ableitungen der Funktion $f$ mit den entsprechenden Ableitungen von $P$ übereinstimmen. Alle Ableitungen von $P$ lassen sich leicht berechnen und wir verlangen:

$$\begin{aligned}
f(\mathbf{a}) &= P(\mathbf{a}) &&= c_0 \\
f_{x_1}(\mathbf{a}) &= P_{x_1}(\mathbf{a}) &&= c_1 \\
f_{x_2}(\mathbf{a}) &= P_{x_2}(\mathbf{a}) &&= c_2 \\
f_{x_1x_1}(\mathbf{a}) &= P_{x_1x_1}(\mathbf{a}) &&= 2c_{11} \\
f_{x_2x_2}(\mathbf{a}) &= P_{x_2x_2}(\mathbf{a}) &&= 2c_{22} \\
f_{x_1x_2}(\mathbf{a}) &= P_{x_1x_2}(\mathbf{a}) &&= 2c_{12} \\
f_{x_2x_1}(\mathbf{a}) &= P_{x_2x_1}(\mathbf{a}) &&= 2c_{12}
\end{aligned}$$

Daraus ergibt sich sofort die folgende Beschreibung der quadratischen Approximation einer Funktion $f$. Alle Parameter im allgemeinen Ansatz lassen sich durch die partiellen Ableitungen von $f$ an der Stelle $\mathbf{a}$ ausdrücken.

**Satz 5** *Für die quadratische Approximation der Funktion $f$ an der Stelle $\mathbf{a}$ gilt:*

$$\boxed{P(\mathbf{x}) \;=\; f(\mathbf{a}) + \begin{pmatrix} f_{x_1}(\mathbf{a}) \\ f_{x_2}(\mathbf{a}) \end{pmatrix}^T (\mathbf{x}-\mathbf{a}) + (\mathbf{x}-\mathbf{a})^T \;\frac{1}{2}\begin{pmatrix} f_{x_1x_1}(\mathbf{a}) & f_{x_1x_2}(\mathbf{a}) \\ f_{x_2x_1}(\mathbf{a}) & f_{x_2x_2}(\mathbf{a}) \end{pmatrix}(\mathbf{x}-\mathbf{a})}$$

*Der Vektor*

$$\textbf{grad } f(\mathbf{a}) \;=\; \begin{pmatrix} f_{x_1}(\mathbf{a}) \\ f_{x_2}(\mathbf{a}) \end{pmatrix}$$

*heisst der <u>Gradient</u> der Funktion $f$ im Punkt $a$. Die Matrix*

$$H_f(\mathbf{a}) \;=\; \begin{pmatrix} f_{x_1x_1}(\mathbf{a}) & f_{x_1x_2}(\mathbf{a}) \\ f_{x_2x_1}(\mathbf{a}) & f_{x_2x_2}(\mathbf{a}) \end{pmatrix}$$

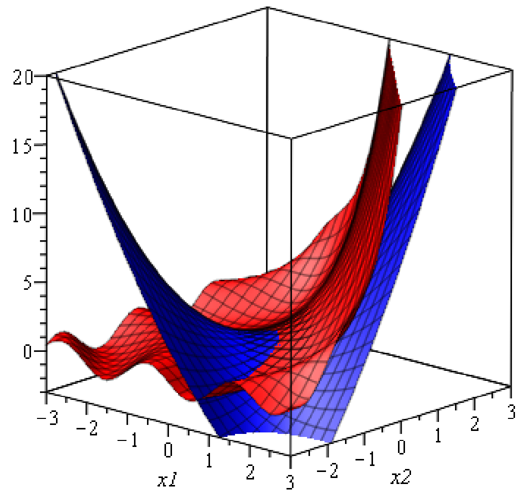
*heisst die <u>Hesse-Matrix</u> der Funktion $f$ im Punkt $a$.*

Beachte:
Für die uns interessierenden Funktionen gilt überall $f_{x_1x_2} = f_{x_2x_1}$, die Hesse-Matrix ist also symmetrisch.

**Beispiel 3.1** *Die quadratische Approximation der Funktion $f(x_1, x_2) = e^{x_1+x_2} + \sin(x_1x_2)$ an der Stelle $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ist*

$$P(\mathbf{x}) \;=\; 1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix}^T \mathbf{x} + \mathbf{x}^T \;\frac{1}{2}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\mathbf{x} \;=\; 1 + x_1 + x_2 + \frac{1}{2}x_1^2 + 2x_1x_2 + \frac{1}{2}x_2^2$$

*Graph von $f$ (rot) und Graph von $P$ (blau)*



**Aufgabe 3.1** *Bestimmen Sie die quadratische Approximation der Funktion* $f(\mathbf{x}) = f(x_1, x_2) = x_1^{1/4} x_2^{3/4}$ *an der Stelle* $\mathbf{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

**Lösung:**

$$P(\mathbf{x}) \quad = \quad 1 + \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix}^T \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} + (x_1 - 1, x_2 - 1) \ \frac{1}{2} \begin{pmatrix} -3/16 & 3/16 \\ 3/16 & -3/16 \end{pmatrix} \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix}$$

## 3.3 Definitheit und lokale Extremalstellen

**Satz 6** *Sei*

- $D \subset \mathbb{R}^n$ *offen,*
- $f : D \to \mathbb{R}$ *eine zweimal stetig differenzierbare Funktion,*
- $\mathbf{a} \in D$ *ein stationärer Punkt von* $f$*,*
- $H_f(\mathbf{a})$ *die Hesse-Matrix von* $f$ *in* $\mathbf{a}$

*Dann gilt:*

$$\boxed{\begin{array}{lcl} H_f(\mathbf{a}) \textit{ positiv definit} & \Longrightarrow & \mathbf{a} \textit{ ist lokale Minimalstelle} \\ H_f(\mathbf{a}) \textit{ negativ definit} & \Longrightarrow & \mathbf{a} \textit{ ist lokale Maximalstelle} \\ H_f(\mathbf{a}) \textit{ indefinit} & \Longrightarrow & \mathbf{a} \textit{ ist Sattelpunkt} \end{array}}$$

**Beweisidee:**
Wir benutzen das 2-te Taylor-Polynom von $f$ in $\mathbf{a}$ und beachten, dass $\mathbf{a}$ ein stationärer Punkt ist, also $\mathbf{grad} f(\mathbf{a}) = \mathbf{0}$ gilt. Der Funktionswertunterschied kann recht gut durch die quadratische Form des Taylor-Poynoms beschrieben werden:

$$\begin{array}{rcl} f(\mathbf{x}) - f(\mathbf{a}) & \approx & P_2(\mathbf{x}) - f(\mathbf{a}) \\ & = & \frac{1}{2}\,(\mathbf{x}-\mathbf{a})^T\, H_f(\mathbf{a})\ \ (\mathbf{x}-\mathbf{a}) \end{array}$$

Genauer genommen gilt in einer kleinen Umgebung $U$ von $\mathbf{a}$ (mit einem nicht besprochenen Zwischenwertsatz):

$$f(\mathbf{x}) - f(\mathbf{a}) \;=\; \frac{1}{2}\,(\mathbf{x}-\mathbf{a})^T\, H_f(\mathbf{a}^\star)\ \ (\mathbf{x}-\mathbf{a})$$

mit einem Punkt $\mathbf{a}^\star$ zwischen $\mathbf{x}$ und $\mathbf{a}$. Da alle zweiten partiellen Ableitungen stetig sind, ändert sich auch $H_f(\mathbf{x})$ stetig in allen Variablen. Insbesondere können wir die Umgebung $U$ so wählen, dass die Matrix $H_f(\mathbf{x})$ für alle $\mathbf{x} \in U$ die selbe Definitheit (positiv definit, negativ definit bzw. indefinit) wie $H_f(\mathbf{a})$ hat. In dieser Umgebung $U$ von $\mathbf{a}$ ist dann die rechte Seite der obigen Gleichung stets positiv, negativ bzw. positiv und negativ. Das bedeutet für alle $\mathbf{x} \in U$ somit

- $f(\mathbf{x}) - f(\mathbf{a}) \geq 0$ oder $f(\mathbf{x}) \geq f(\mathbf{a})$ und $\mathbf{a}$ ist lokale Minimalstelle
- $f(\mathbf{x}) - f(\mathbf{a}) \leq 0$ oder $f(\mathbf{x}) \leq f(\mathbf{a})$ und $\mathbf{a}$ ist lokale Maximalstelle bzw.
- $f(\mathbf{x}) - f(\mathbf{a})$ ist sowohl positiv als auch negativ und $\mathbf{a}$ ist Sattelpunkt.

□

**Spezialfall** $n = 2$ Sei $\mathbf{a} = (a_1, a_2)^T$ ein stationärer Punkt der zweimal stetig differenzierbaren Funktion $f$. Für die Hessematrix gilt dann

$$H_f(\mathbf{a}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{a}) & f_{x_1x_2}(\mathbf{a}) \\ f_{x_2x_1}(\mathbf{a}) & f_{x_2x_2}(\mathbf{a}) \end{pmatrix} \longrightarrow \det H_f(\mathbf{a}) = f_{x_1x_1}(\mathbf{a}) f_{x_2x_2}(\mathbf{a}) - f_{x_1x_2}(\mathbf{a})^2.$$

Zusammen mit den Hurwitz-Kriterien für $2 \times 2$-Matrizen folgt dann

| | | | | |
|---|---|---|---|---|
| $\det H_f(\mathbf{a}) > 0$ und $f_{x_1x_1}(\mathbf{a}) > 0$ | $\Longrightarrow$ | $H_f(\mathbf{a})$ pos. def. | $\Longrightarrow$ | $\mathbf{a}$ ist lokale Minimalstelle |
| $\det H_f(\mathbf{a}) > 0$ und $f_{x_1x_1}(\mathbf{a}) < 0$ | $\Longrightarrow$ | $H_f(\mathbf{a})$ neg. def. | $\Longrightarrow$ | $\mathbf{a}$ ist lokale Maximalstelle |
| $\det H_f(\mathbf{a}) < 0$ | $\Longrightarrow$ | $H_f(\mathbf{a})$ indef. | $\Longrightarrow$ | $\mathbf{a}$ ist Sattelpunkt |

# 4 Übungsaufgaben

1. Berechnen Sie für die Matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

die quadratische Form $Q_A$ und bestimmen sie die Werte $Q_A(\mathbf{x})$ für die Vektoren

$$\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ und } \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$$

2. Untersuchen Sie die folgenden Matrizen auf ihre Definitheit:

a) $\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ c) $\begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix}$ d) $\begin{pmatrix} -1 & 2 \\ 2 & 4 \end{pmatrix}$

e) $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ f) $\begin{pmatrix} 40 & 1 \\ 1 & 2 \end{pmatrix}$ g) $\begin{pmatrix} 40 & -1 \\ -1 & 2 \end{pmatrix}$

h) $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ i) $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ j) $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 6 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

3. Bestimmen Sie die quadratische Approximation der Funktion

$$f(\underbrace{x_1}_{x}, \underbrace{x_2}_{y}) = x\sin(x) + y\cos(y)$$

an den Stellen $(a)$ $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und $(b)$ $\mathbf{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

4. (a) Berechnen Sie mit Hilfe der Lagrange-Methode eine mögliche Extremalstelle $(x^\star, y^\star)$ des folgenden Optimierungsproblems:
   Maximiere die Zielfunktion $f(x,y) = 8\ln(x) + 2\ln(y)$ unter der Nebenbedingung $\phi(x,y) = 2x + y - 5 = 0$.

   (b) Berechnen Sie **grad** $f(x^\star, y^\star)$ und **grad** $\phi(x^\star, y^\star)$ und interpretieren Sie das Resultat an Hand einer Graphik.

5. Sei $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine beliebige (nicht notwendigerweise symmetrische) $(2 \times 2)$-Matrix und $A = \frac{1}{2}(B + B^T)$. Zeigen Sie:

   (a) $A$ ist symmetrisch und

   (b) $Q_A = Q_B$, d.h. beide Matrizen beschreiben die selbe quadratische Form.