Mathematik für Studierende der Erdwissenschaften
Lösungen der Beispiele des 12. Übungsblatts

1. Eingeschränkte Extrema

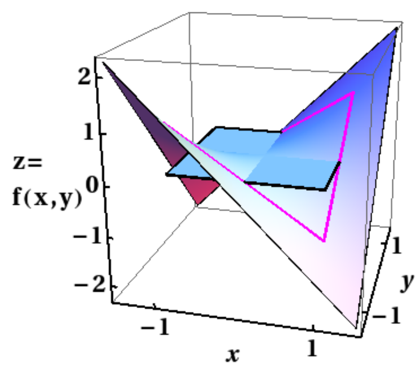
   (a) Die globalen Extrema der Funktion

$$f(x, y) = xy$$

über die Menge

$$Q = \{(x, y) \in \mathbb{R}^2 : \max\{|x|, |y|\} \leq 1\}$$

sollen gefunden werden. Die grafische Darstellung des Problems ist gegeben durch



wobei die sattelförmige (blaue) Fläche der Funktion $f(x, y)$ entspricht, und das flache (zyanfarbige) Quadrat der Menge Q entspricht. Weiters entspricht die (magentafarbige) Kurve den Werten von $f(x, y)$ über den Rand von Q,

$$\partial Q = \{(x, y) \in \mathbb{R}^2 : \max\{|x|, |y|\} = 1\}$$

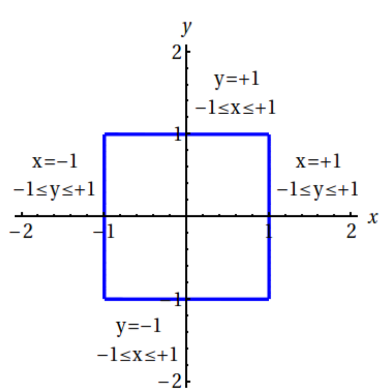
Die partiellen Ableitungen erster Ordnung sind gegeben durch

$$f_x(x, y) = y, \quad f_y(x, y) = x$$

und daher ist $(x, y) = (0, 0)$ der einzige kritische Punkt in der Menge Q, in dem $f_x(0, 0) = 0 = f_y(0, 0)$ gelten. Der Funktionswert an dieser Stelle ist

$$f(0, 0) = 0$$

Die grafische Darstellung des Rands ∂Q ist gegeben durch



Um die Extremwerte von $f(x,y)$ am Rand ∂Q zu zeigen, wird beispielsweise die obere Seite mit $y = +1$ und $-1 \le x \le +1$ im Detail untersucht. Sei $g(x)$ durch die f-Werte auf dieser Seite definiert:

$$g(x) = f(x, +1) = x, \quad x \in [-1, +1]$$

Das Rezept der globalen Extrema zeigt für $g(x)$ mit $D = [-1, +1]$, das globale Minimum für g ist $g(-1) = -1$ und das globale Maximum für g ist $g(+1) = +1$. Analog kann gezeigt werden, die Extremwerte von $f(x,y)$ über die anderen Seiten sind die gleichen,

$$-1 = g(-1) = f(-1, +1) = f(+1, -1), \quad +1 = g(+1) = f(+1, +1) = f(-1, -1)$$

Der interne Wert $f(0,0) = 0$ liegt zwischen den Extremwerten am Rand. Daher ist das globale Minimum von $f(x,y)$ über Q gegeben durch $f(-1,+1) = f(+1,-1) = -1$, und das globale Maximum von $f(x,y)$ über Q ist $f(-1,-1) = f(+1,+1) = +1$.

Das gleiche Argument zeigt, das globale Minimum von $f(x,y)$ über den Rand ∂Q ist gegeben durch $f(-1,+1) = f(+1,-1) = -1$, und das globale Maximum von $f(x,y)$ über ∂Q ist $f(-1,-1) = f(+1,+1) = +1$.

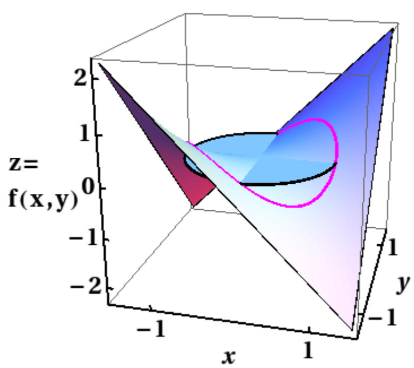
(b) Die globalen Extrema der Funktion

$$f(x,y) = xy$$

über die Menge

$$S = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 1\}$$

sollen gefunden werden. Die grafische Darstellung des Problems ist gegeben durch



wobei die sattelförmige (blaue) Fläche der Funktion $f(x,y)$ entspricht, und die flache (zyanfarbige) Scheibe der Menge S entspricht. Weiters entspricht die (magentafarbige) Kurve den Werten von $f(x,y)$ über den Rand von S,

$$\partial S = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

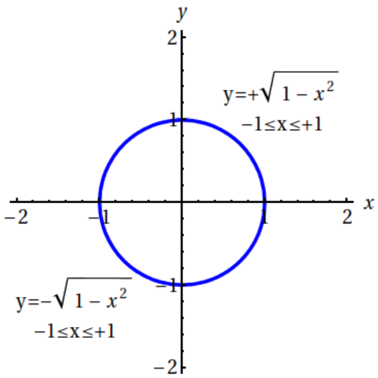
Die partiellen Ableitungen erster Ordnung sind gegeben durch

$$f_x(x,y) = y, \quad f_y(x,y) = x$$

und daher ist $(x,y) = (0,0)$ der einzige kritische Punkt in der Menge S, in dem $f_x(0,0) = 0 = f_y(0,0)$ gelten. Der Funktionswert an dieser Stelle ist

$$f(0,0) = 0$$

Die grafische Darstellung des Rands ∂S ist gegeben durch



Um die Extremwerte von $f(x,y)$ am Rand ∂S zu zeigen, wird beispielsweise die obere Seite mit $y = +\sqrt{1-x^2}$ und $-1 \leq x \leq +1$ im Detail untersucht. Sei $g(x)$ durch die f-Werte auf dieser Seite definiert:

$$g(x) = f(x, +\sqrt{1-x^2}) = x\sqrt{1-x^2}, \quad \text{mit} \quad x \in [-1,+1] \quad \text{und} \quad g'(x) = \frac{1-2x^2}{\sqrt{1-x^2}}$$

Das Rezept der globalen Extrema zeigt für $g(x)$ mit $D = [-1,+1]$, das globale Minimum für g ist $g(-1/\sqrt{2}) = -1/2$ und das globale Maximum für g ist $g(+1/\sqrt{2}) = +1/2$. Analog kann gezeigt werden, die Extremwerte von $f(x,y)$ über die untere Seiten sind die gleichen,

$$-\tfrac{1}{2} = g(-\tfrac{1}{\sqrt{2}}) = f(-\tfrac{1}{\sqrt{2}}, +\tfrac{1}{\sqrt{2}}) = f(+\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}})$$

$$+\tfrac{1}{2} = g(+\tfrac{1}{\sqrt{2}}) = f(+\tfrac{1}{\sqrt{2}}, +\tfrac{1}{\sqrt{2}}) = f(-\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}})$$

Der interne Wert $f(0,0) = 0$ liegt zwischen den Extremwerten am Rand. Daher ist das globale Minimum von $f(x,y)$ über S gegeben durch $f(-1/\sqrt{2}, +1/\sqrt{2}) = -1/2$, und das globale Maximum von $f(x,y)$ über S ist $f(-1/\sqrt{2}, -1/\sqrt{2}) = +1/2$.

Das gleiche Argument zeigt, das globale Minimum von $f(x,y)$ über den Rand ∂S ist gegeben durch $f(-1/\sqrt{2}, +1/\sqrt{2}) = -1/2$, und das globale Maximum von $f(x,y)$ über ∂Q ist $f(-1/\sqrt{2}, -1/\sqrt{2}) = +1/2$.

Die globalen Extrema der gegebenen Funktion über die Ränder der obigen Mengen können – wie im nächsten Beispiel – auch mit Lagrangeschen Multiplikatoren gefunden werden.

2. Geometrische Projektion

   (a) Der Abstand zwischen dem Punkt $(0,1)$ und einer beliebigen Punkt (x,y) in der Parabel

   $$P = \{(x,y) \in \mathbb{R}^2 : y = x^2\}$$

   wird minimiert, wenn die Funktion

   $$f(x,y) = (x-0)^2 + (y-1)^2$$

   minimiert wird unter der Einschränkung

   $$g(x,y) = y - x^2 = 0$$

Die Lösung dieses eingeschränkten Minimierungsproblems entspricht einem kritischen Punkt der Lagrangeschen Funktion,

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

wobei λ ein Lagrangescher Multiplikator für die Einschränkung $g(x, y) = 0$ ist. Die Bedingungen eines kritischen Punkts sind

$$\begin{array}{llll} L_x(x,y,\lambda) & = f_x(x,y) - \lambda g_x(x,y) & = 2x + 2\lambda x & = 0 \\ L_y(x,y,\lambda) & = f_y(x,y) - \lambda g_y(x,y) & = 2(y-1) - \lambda & = 0 \\ L_\lambda(x,y,\lambda) & = -g(x,y) & = -y + x^2 & = 0 \end{array}$$

Eine Lösung der ersten Gleichung dieses Systems ist $x = 0$, und es folgt $y = 0$ aus der dritten Gleichung. Aus der zweiten Gleichung folgt $\lambda = -2$. Also ist die erste kritische Punkte von L gegeben durch

$$x = 0, \quad y = 0, \quad \lambda = -2$$

Die andere Lösung der ersten Gleichung des Systems ist $\lambda = -1$, und es folgt $y = 1/2$ aus der zweiten Gleichung. Aus der dritten Gleichung folgen $x = 1/\sqrt{2}$ oder $x = -1/\sqrt{2}$. Also ist der zweite und der dritte kritische Punkt von L gegeben durch

$$x = -1/\sqrt{2}, \quad y = 1/2, \quad \lambda = -1 \quad \text{und} \quad x = +1/\sqrt{2}, \quad y = 1/2, \quad \lambda = -1$$

Die Werte von $f(x, y)$ für diese drei kritischen Punkte sind

$$f(0,0) = 1, \quad f(+\tfrac{1}{\sqrt{2}}, \tfrac{1}{2}) = \tfrac{3}{4} = f(-\tfrac{1}{\sqrt{2}}, \tfrac{1}{2})$$

Der minimale Wert in dieser Liste ist $\frac{3}{4}$. Daher ist $\frac{3}{4}$ der Abstand zwischen dem Punkt $(0, 1)$ und der Parabel $y = x^2$, und die Punkte auf der Parabel, die am nähesten zum Punkt $(0, 1)$ liegen, sind $(-\frac{1}{\sqrt{2}}, \frac{1}{2})$ und $(+\frac{1}{\sqrt{2}}, \frac{1}{2})$.

(b) Der Abstand zwischen dem Punkt $(2, 0, 0)$ und einer beliebigen Punkt (x, y, z) in der Ebene

$$E = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}$$

wird minimiert, wenn die Funktion

$$f(x, y, z) = (x - 2)^2 + (y - 0)^2 + (z - 0)^2$$

minimiert wird unter der Einschränkung

$$g(x, y, z) = x + y + z - 1 = 0$$

Die Lösung dieses eingeschränkten Minimierungsproblems entspricht einem kritischen Punkt der Lagrangeschen Funktion,

$$L(x, y, z, \lambda) = f(x, y, z) - \lambda g(x, y, z)$$

wobei λ ein Lagrangescher Multiplikator für die Einschränkung $g(x, y, z) = 0$ ist. Die Bedingungen eines kritischen Punkts sind

$$\begin{array}{llll} L_x(x,y,\lambda) & = f_x(x,y,z) - \lambda g_x(x,y,z) & = 2(x-2) - \lambda & = 0 \\ L_y(x,y,\lambda) & = f_y(x,y,z) - \lambda g_y(x,y,z) & = 2y - \lambda & = 0 \\ L_z(x,y,\lambda) & = f_z(x,y,z) - \lambda g_z(x,y,z) & = 2z - \lambda & = 0 \\ L_\lambda(x,y,z,\lambda) & = -g(x,y,z) & = -x - y - z + 1 & = 0 \end{array}$$

Durch die ersten drei Gleichungen dieses Systems folgt

$$x - 2 = y = z = \lambda/2$$

Durch die vierte Gleichung des Systems folgt

$$1 = x + y + z = (2 + \lambda/2) + (\lambda/2) + (\lambda/2) \quad \text{oder} \quad \lambda = -2/3$$

Aus der vorletzten Gleichung folgen

$$x = 2 + \lambda/2 = 5/3, \quad y = \lambda/2 = -1/3, \quad z = \lambda/2 = -1/3$$

Also ist $(5/3, -1/3, -1/3)$ der Punkt in der Ebene $x + y + z = 1$, der am nähesten zum Punkt $(2, 0, 0)$ liegt, und der Abstand zwischen dem Punkt und der Ebene ist gegeben durch den Wert

$$f(5/3, -1/3, -1/3) = 1/3$$

Für die obigen Abstände, d.h. die globalen Minima der obigen Zielfunktionen über die gegebenen Einschränkungsmengen, können – wie im ersten Beispiel – auch ohne Lagrangesche Multiplikatoren gefunden werden.

3. Lineare Regression

   (a) Die partiellen Ableitungen der Funktion

$$E(k, d) = \sum_{i=1}^{n} [(kx_i + d) - y_i]^2$$

nach k und d sind

$$E_k(k, d) = \sum_{i=1}^{n} \partial_k [(kx_i + d) - y_i]^2 = \sum_{i=1}^{n} 2[(kx_i + d) - y_i]x_i$$

$$= 2kn \left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] + 2dn \left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] - 2n \left[\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right] = 2n \left[k\overline{x^2} + d\overline{x} - \overline{xy}\right]$$

und

$$E_d(k, d) = \sum_{i=1}^{n} \partial_d [(kx_i + d) - y_i]^2 = \sum_{i=1}^{n} 2[(kx_i + d) - y_i]$$

$$= 2kn \left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] + 2dn \left[\frac{1}{n}\sum_{i=1}^{n} 1\right] - 2n \left[\frac{1}{n}\sum_{i=1}^{n} y_i\right] = 2n \left[k\overline{x} + d - \overline{y}\right]$$

Die Lösung des lineare Gleichungssystems

$$\left\{\begin{array}{l} E_k(k,d)/(2n) = 0 \\ E_d(k,d)/(2n) = 0 \end{array}\right. \quad \text{oder} \quad \left\{\begin{array}{lrcl} \text{I:} & k\overline{x^2} + d\overline{x} & = & \overline{xy} \\ \text{II:} & k\overline{x} + d & = & \overline{y} \end{array}\right. \quad \text{oder} \quad \begin{bmatrix} \overline{x^2} & \overline{x} \\ \overline{x} & 1 \end{bmatrix} \begin{bmatrix} k \\ d \end{bmatrix} = \begin{bmatrix} \overline{xy} \\ \overline{y} \end{bmatrix}$$

bekommt man durch die gewichtete Summe der Gleichungen $\text{I} - \overline{x} \cdot \text{II}$ oder

$$k(\overline{x^2} - \overline{x}^2) = \underbrace{(k\overline{x^2} + d\overline{x})}_{\text{I, links}} - \overline{x} \cdot \underbrace{(k\overline{x} + d)}_{\text{II, links}} = \underbrace{\overline{xy}}_{\text{I, rechts}} - \overline{x} \cdot \underbrace{\overline{y}}_{\text{II, rechts}}$$

Die Lösung ist

$$k^* = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} \quad \text{und durch II,} \quad d^* = \bar{y} - k^* \bar{x}$$

(b) Die partiellen Ableitung zweiter Ordnung der Funktion $E(k,d)$ sind

$$E_{kk}(k,d) = \sum_{i=1}^{n} \partial_k^2[(kx_i+d)-y_i]^2 \overset{\text{K.R.}}{=} \sum_{i=1}^{n} \partial_k 2[(kx_i+d)-y_i]x_i = 2n\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] = 2n\overline{x^2} > 0$$

$$E_{dd}(k,d) = \sum_{i=1}^{n} \partial_d^2[(kx_i+d)-y_i]^2 \overset{\text{K.R.}}{=} \sum_{i=1}^{n} \partial_d 2[(kx_i+d)-y_i] = 2n > 0$$

und

$$E_{kd}(k,d) = \sum_{i=1}^{n} \partial_d\partial_k[(kx_i+d)-y_i]^2 \overset{\text{K.R.}}{=} \sum_{i=1}^{n} \partial_d 2[(kx_i+d)-y_i]x_i = 2n\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = 2n\overline{x}$$

(c) Wenn $\overline{(x-\bar{x})^2}$ explizit und ausführlich geschrieben wird, ergibt sich

$$\overline{(x-\bar{x})^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}[x_i^2 - 2x_i\overline{x} + \overline{x}^2]$$

$$= \left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - 2\overline{x}\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] + \overline{x}^2\left[\frac{1}{n}\sum_{i=1}^{n} 1\right] = \overline{x^2} - 2\overline{x}\cdot\overline{x} + \overline{x}^2 = \overline{x^2} - \bar{x}^2$$

Da die x-Werte verschieden sind, gilt

$$\overline{x^2} - \bar{x}^2 = \overline{(x-\bar{x})^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2 > 0$$

Es folgt

$$E_{kk}(k,d)E_{dd}(k,d) - E_{kd}(k,d)^2 = 4n^2(\overline{x^2} - \bar{x}^2) > 0$$

(d) Da die Bedingungen auf Seite 180 im Skriptum erfüllt sind, ist $E(k,d)$ global konvex, und (k^*,d^*) ist ein globales Minimum für $E(k,d)$.

4. Polynomiale Regression

(a) In der Zielfunktion

$$E(p_0,p_1,\dots,p_{m-1},p_m) = \sum_{i=1}^{n}[(p_m x_i^m + p_{m-1}x_i^{m-1} + \cdots + p_1 x_i + p_0) - y_i]^2$$

lässt sich der $i$te Term bezüglich eines Skalarprodukts darstellen,

$$(p_0 + p_1 x_i + \cdots p_{m-1}x_i^{m-1} + p_m x_i^m) - y_i$$

$$= \begin{bmatrix} 1 \\ x_i \\ \vdots \\ x_i^{m-1} \\ x_i^m \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{m-1} \\ p_m \end{bmatrix} - y_i = \begin{bmatrix} 1 & x_i & \cdots & x_i^{m-1} & x_i^m \end{bmatrix}^\top \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{m-1} \\ p_m \end{bmatrix} - y_i$$

und der Vektor dieser Terme lässt sich bezüglich aller solchen Skalarprodukte darstellen,

$$
\begin{bmatrix}
(p_0 + p_1 x_1 + \cdots p_{m-1} x_1^{m-1} + p_m x_1^m) - y_1 \\
(p_0 + p_1 x_2 + \cdots p_{m-1} x_2^{m-1} + p_m x_2^m) - y_2 \\
\vdots \\
(p_0 + p_1 x_{n-1} + \cdots p_{m-1} x_{n-1}^{m-1} + p_m x_{n-1}^m) - y_{n-1} \\
(p_0 + p_1 x_n + \cdots p_{m-1} x_n^{m-1} + p_m x_n^m) - y_n
\end{bmatrix}
$$

$$
= \begin{bmatrix}
1 & x_1 & \cdots & x_1^{m-1} & x_1^m \\
1 & x_2 & \cdots & x_2^{m-1} & x_2^m \\
 & & \vdots & & \\
1 & x_{n-1} & \cdots & x_{n-1}^{m-1} & x_{n-1}^m \\
1 & x_n & \cdots & x_n^{m-1} & x_n^m
\end{bmatrix}^\top
\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{m-1} \\ p_m \end{bmatrix}
-
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix}
= M\boldsymbol{p} - \boldsymbol{y}
$$

Die Zielfunktion E ist die Summe der Quadrate dieser Terme, d.h. die quadrierte Norm des obigen Vektors,

$$
E(p_0, p_1, \ldots, p_{m-1}, p_m) = \sum_{i=1}^{n} [(p_m x_i^m + p_{m-1} x_i^{m-1} + \cdots + p_1 x_i + p_0) - y_i]^2
$$

$$
= \left\| \begin{bmatrix}
(p_0 + p_1 x_1 + \cdots p_{m-1} x_1^{m-1} + p_m x_1^m) - y_1 \\
(p_0 + p_1 x_2 + \cdots p_{m-1} x_2^{m-1} + p_m x_2^m) - y_2 \\
\vdots \\
(p_0 + p_1 x_{n-1} + \cdots p_{m-1} x_{n-1}^{m-1} + p_m x_{n-1}^m) - y_{n-1} \\
(p_0 + p_1 x_n + \cdots p_{m-1} x_n^{m-1} + p_m x_n^m) - y_n
\end{bmatrix} \right\|^2 = \|M\boldsymbol{p} - \boldsymbol{y}\|^2
$$

(b) Wie auf dem 4. Ergänzungsblatt gezeigt, lässt sich die Zielfunktion so umschreiben,

$$
E(\boldsymbol{p}) = \|M\boldsymbol{p} - \boldsymbol{y}\|^2 = \boldsymbol{p}^\top M^\top M \boldsymbol{p} - 2(M^\top \boldsymbol{y})^\top \boldsymbol{p} + \boldsymbol{y}^\top \boldsymbol{y}
$$

Wie auf dem 4. Ergänzungsblatt gezeigt, ist der Vektor der partiellen Ableitungen erster Ordnung des ersten Terms gegeben durch

$$
\{\partial_{p_k} \boldsymbol{p}^\top M^\top M \boldsymbol{p}\}_{k=0}^m = 2M^\top M \boldsymbol{p}
$$

Wie auf dem 4. Ergänzungsblatt gezeigt, ist der Vektor der partiellen Ableitungen erster Ordnung des zweiten Terms gegeben durch

$$
\{\partial_{p_k} 2(M^\top \boldsymbol{y})^\top \boldsymbol{p}\}_{k=0}^m = 2M^\top \boldsymbol{y}
$$

Wie auf dem 4. Ergänzungsblatt gezeigt, ist der Vektor der partiellen Ableitungen erster Ordnung des dritten Terms gegeben durch

$$
\{\partial_{p_k} \boldsymbol{y}^\top \boldsymbol{y}\}_{k=0}^m = \mathbf{0}
$$

Zusammengefasst ist der Vektor der partiellen Ableitungen erster Ordnung $\{\partial_{p_k} E(\boldsymbol{p})\}_{k=0}^m$ gegeben durch:

$$
\{\partial_{p_k} E(\boldsymbol{p})\}_{k=0}^m = 2(M^\top M \boldsymbol{p} - M^\top \boldsymbol{y})
$$

(c) Das letzte Ergebnis ist für die partiellen Ableitungen zweiter Ordnung noch abzuleiten. Wie auf dem 4. Ergänzungsblatt gezeigt, ist die Matrix der partiellen Ableitungen zweiter Ordnung des ersten Terms gegeben durch

$$\{\partial_{p_l}\partial_{p_k} 2M^\top M\boldsymbol{p}\}_{k,l=0}^m = 2M^\top M$$

Wie auf dem 4. Ergänzungsblatt gezeigt, ist die Matrix der partiellen Ableitungen zweiter Ordnung des zweiten Terms gegeben durch

$$\{\partial_{p_l}\partial_{p_k} 2M^\top \boldsymbol{y}\}_{k,l=0}^m = 0$$

Zusammengefasst ist die Matrix der partiellen Ableitungen zweiter Ordnung $\{\partial_{p_l}\partial_{p_k} E(\boldsymbol{p})\}_{k,l=0}^m$ gegeben durch:

$$\{\partial_{p_l}\partial_{p_k} E(\boldsymbol{p})\}_{k,l=0}^m = 2M^\top M$$

5. Rechnungen mit Matrizen und Vektoren

(a) Mit

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

wird das lineare Gleichungssystem $A\boldsymbol{x} = \boldsymbol{b}$ folgendermaßen durch Gaußsche Elimination

$$\begin{array}{r} \mathrm{I}: \\ \mathrm{II}: \end{array} \left[\begin{array}{cc|c} 2 & 1 & 4 \\ 1 & 2 & 5 \end{array}\right]$$

$$\begin{array}{r} \mathrm{I}: \\ \mathrm{II} - \frac{1}{2}\mathrm{I} \to \tilde{\mathrm{II}}: \end{array} \left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & \frac{3}{2} & 3 \end{array}\right]$$

und rückwärts Substitution

$$\left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & \frac{3}{2} & 3 \end{array}\right] \quad \begin{array}{r} \Rightarrow x = (4 - 1 \cdot y)/2 = 1 \\ \Rightarrow y = 3/\frac{3}{2} = 2 \quad \uparrow \end{array}$$

mit dem Vektor

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

gelöst.

(b) Durch die Cramersche Regel ist die Lösung des Gleichungssystem $A\boldsymbol{x} = \boldsymbol{b}$ gegeben durch

$$x = \frac{\det \begin{bmatrix} 4 & 1 \\ 5 & 2 \end{bmatrix}}{\det \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}} = \frac{4 \cdot 2 - 1 \cdot 5}{2 \cdot 2 - 1 \cdot 1} = \frac{3}{3} = 1, \quad y = \frac{\det \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix}}{\det \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}} = \frac{2 \cdot 5 - 1 \cdot 4}{2 \cdot 2 - 1 \cdot 1} = \frac{6}{3} = 2$$

(c) Die Inverse

$$A^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

lässt sich durch die folgenden Eliminationsschritte bestimmen

$$\begin{array}{r} \mathrm{I}: \\ \mathrm{II}: \end{array} \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right]$$

$$\begin{array}{r} \mathrm{I}: \\ \mathrm{II}-\frac{1}{2}\mathrm{I} \to \tilde{\mathrm{II}}: \end{array} \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & \frac{3}{2} & -\frac{1}{2} & 1 \end{array}\right]$$

$$\begin{array}{r} \mathrm{I}: \\ \frac{2}{3}\tilde{\mathrm{II}} \to \hat{\mathrm{II}}: \end{array} \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \end{array}\right]$$

$$\begin{array}{r} \mathrm{I}-\hat{\mathrm{II}} \to \tilde{\mathrm{I}}: \\ \hat{\mathrm{II}}: \end{array} \left[\begin{array}{cc|cc} 2 & 0 & \frac{4}{3} & -\frac{2}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \end{array}\right]$$

$$\begin{array}{r} \frac{1}{2}\tilde{\mathrm{I}} \to \hat{\mathrm{I}}: \\ \hat{\mathrm{II}}: \end{array} \left[\begin{array}{cc|cc} 1 & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \end{array}\right]$$

(d) Mit

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, \quad V = \begin{bmatrix} +1 & -1 \\ +1 & +1 \end{bmatrix}$$

bestätigt man durch eine direkte Rechnung, es gilt $AV = V\Lambda$. Die Eigenwerte und die entsprechenden Eigenvektoren

$$\lambda_1 = 3, \quad \boldsymbol{v}_1 = \begin{bmatrix} +1 \\ +1 \end{bmatrix}, \quad \lambda_2 = 1, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$$

der Matrix A werden folgendermaßen bestimmt. Die Nullstellen des Polynoms

$$p(\lambda) = \det(A-\lambda I) = \det\left(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}\right) = (2-\lambda)^2 - 1$$

sind $\lambda_1 = 3$ und $\lambda_2 = 1$. Ein Eigenvektor für λ_1 ist eine nicht triviale Lösung des Systems

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda_1 \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{oder} \quad \begin{bmatrix} 2-\lambda_1 & 1 \\ 1 & 2-\lambda_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\text{oder mit } \lambda_1 = 3, \quad \begin{bmatrix} -1 & +1 \\ +1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{oder z.B.} \quad x = y = 1$$

Ein Eigenvektor für λ_2 ist eine nicht triviale Lösung des Systems

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda_2 \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{oder} \quad \begin{bmatrix} 2-\lambda_2 & 1 \\ 1 & 2-\lambda_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\text{oder mit } \lambda_2 = 1, \quad \begin{bmatrix} +1 & +1 \\ +1 & +1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{oder z.B.} \quad x = -1 = -y$$

(e) Eine Charakterisierung einer SPD-Matrix ist, dass sie symmetrisch ist mit positiven Eigenwerten. Die Matrix A erfüllt $A^\top = A$, und sie ist daher symmetrisch. Da die Eigenwerte $\lambda_1 = 3 > 0$, $\lambda_1 = 1 > 0$, positiv sind, ist A SPD.

6. Eine Charakterisierung einer 2×2 SPD-Matrix ist,

$$\text{für} \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{gelten} \quad a_{11}, a_{22} > 0 \quad \text{und} \quad a_{11}a_{22} > a_{12}a_{21} = a_{12}^2$$

Die Bedingungen für ein lokales Minimum in $f(x_0, y_0)$ einer ausreichend differenzierbaren Funktion $f(x, y)$

$$f_{xx}(x_0, y_0), f_{yy}(x_0, y_0) > 0 \quad \text{und} \quad f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) > (f_{xy}(x_0, y_0))^2$$

ist die Bedingung, dass die Hessematrix ($f_{yx}(x_0, y_0) = f_{xy}(x_0, y_0)$)

$$\nabla^2 f(x_0, y_0) = \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}$$

SPD ist. Die Bedingungen für ein lokales Maximum in $f(x_0, y_0)$ einer ausreichend differenzierbaren Funktion $f(x, y)$

$$f_{xx}(x_0, y_0), f_{yy}(x_0, y_0) < 0 \quad \text{und} \quad f_{xx}(x_0, y_0) f_{yy}(x_0, y_0)) > (f_{xy}(x_0, y_0))^2$$

oder

$$-f_{xx}(x_0, y_0), -f_{yy}(x_0, y_0) > 0 \quad \text{und} \quad (-f_{xx}(x_0, y_0))(-f_{yy}(x_0, y_0)) > (-f_{xy}(x_0, y_0))^2$$

ist die Bedingung, dass die Matrix

$$-\nabla^2 f(x_0, y_0) = \begin{bmatrix} -f_{xx}(x_0, y_0) & -f_{xy}(x_0, y_0) \\ -f_{yx}(x_0, y_0) & -f_{yy}(x_0, y_0) \end{bmatrix}$$

SPD ist.