Bemerkung

1) *Die Bedingung* $\operatorname{grad} f(\mathbf{x}^0) = 0^T$ *definiert gewöhnlich ein nichtlineares Gleichungssystem zur Berechnung von* $\mathbf{x} = \mathbf{x}^0$, *wobei* n *Gleichungen für* n *Unbekannte gegeben sind.*
2) *Die Punkte* $\mathbf{x}^0 \in D^0$ *mit* $\operatorname{grad} f(\mathbf{x}^0) = 0$ *nennt man* **stationäre Punkte** *von* $f(x)$. *Stationäre Punkte sind* **nicht** *immer lokale Extremwerte. Zum Beispiel besitzt die Funktion*

$$f(x, y) := x^2 - y^2$$

den Gradienten

$$\operatorname{grad} f(x, y) = 2(x, -y)^T$$

und hat daher nur einen stationären Punkt $\mathbf{x}^0 = (0, 0)^T$, *dieser ist jedoch ein* **Sattelpunkt** *von* $f(x)$.

Rückblick auf die letzte Vorlesung

1. Mittelwertsatz
2. Mittelwertabschätzung
3. Taylorentwicklungen
4. Extrema

Ausblick auf die heutige Vorlesung

1. Hinreichende und notwendige Bedingungen
2. Eigenwerte der Hesse-Matrix
3. Lokale Auflösbarkeit
4. Satz über implizite Funktionen

Klassifikation stationärer Punkte

Satz

Sei $f(\mathbf{x})$ eine $\mathcal{C}^2$–Funktion auf D^0 und $\mathbf{x}^0 \in D^0$ ein stationärer Punkt von $f(\mathbf{x})$, d.h. $\operatorname{grad} f(\mathbf{x}^0) = 0$.

1) **Notwendige Bedingung II**
   Ist $\mathbf{x}^0$ ein lokales Extremum von $f(\mathbf{x})$, so gilt:
   $\mathbf{x}^0$ *lokales Minimum* $\Rightarrow$ $\mathbf{H}\, f(\mathbf{x}^0)$ *positiv semidefinit*
   $\mathbf{x}^0$ *lokales Maximum* $\Rightarrow$ $\mathbf{H}\, f(\mathbf{x}^0)$ *negativ semidefinit.*

Klassifikation stationärer Punkte II

Satz

2) **Hinreichende Bedingung**
Ist $\mathbf{H}\, f(\mathbf{x}^0)$ *positiv definit (bzw. negativ definit), so ist* $\mathbf{x}^0$ *ein strenges lokales Minimum (bzw. Maximum) von* $f(\mathbf{x})$.
Ist $\mathbf{H}\, f(\mathbf{x}^0)$ *indefinit, so ist* $\mathbf{x}^0$ *ein Sattelpunkt, d.h. es gibt in jeder Umgebung von* $\mathbf{x}^0$ *Punkte* $\mathbf{x}^1$ *und* $\mathbf{x}^2$ *mit*

$$f(\mathbf{x}^1) < f(\mathbf{x}^0) < f(\mathbf{x}^2)$$

Klassifikation Beweis von Teil 1

Beweis

Sei $\mathbf{x}^0$ *ein lokales Minimum. Für* $\mathbf{v} \neq 0$ *und* $\varepsilon > 0$ *hinreichend klein folgt aus der Taylor–Formel*

$$f(\mathbf{x}^0 + \varepsilon\mathbf{v}) - f(\mathbf{x}^0) = \frac{1}{2}(\varepsilon\mathbf{v})^T \mathbf{H}\, f(\mathbf{x}^0 + \theta\varepsilon\mathbf{v})(\varepsilon\mathbf{v}) \geq 0 \qquad (1)$$

mit $\theta = \theta(\varepsilon, \mathbf{v}) \in (0, 1)$.
Hinweis: *Der Gradient in der Taylorentwicklung entfällt, da* $\operatorname{grad} f(\mathbf{x}^0) = 0$ *gilt.*
Wegen (1) gilt auch

$$\mathbf{v}^T \mathbf{H}\, f(\mathbf{x}^0 + \theta\varepsilon\mathbf{v})\mathbf{v} \geq 0. \qquad (2)$$

Klassifikation Beweis von Teil 1 - II

Beweis

Da $f(\mathbf{x})$ *eine* $\mathcal{C}^2$*–Funktion ist, ist die Hesse–Matrix eine* **stetige** *Abbildung. Im Grenzwert* $\varepsilon \to 0$ *folgt daher aus (2)*

$$\mathbf{v}^T \mathbf{H}\, f(\mathbf{x}^0)\mathbf{v} \geq 0,$$

d.h. $\mathbf{H}\, f(\mathbf{x}^0)$ *ist positiv semidefinit.*

Klassifikation: Beweis von Teil 2

Beweis

Ist $\mathbf{H}\, f(\mathbf{x}^0)$ *positiv definit, so* $\mathbf{H}\, f(\mathbf{x})$ *ebenfalls in einer hinreichend kleinen Umgebung* $\mathbf{x} \in K_\varepsilon(\mathbf{x}^0) \subset D$ *positiv definit.*
Dies folgt aus der Stetigkeit der zweiten partiellen Ableitungen.
Für $\mathbf{x} \in K_\varepsilon(\mathbf{x}^0)$, $\mathbf{x} \neq \mathbf{x}^0$ *gilt damit*

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}^0) &= \frac{1}{2}(\mathbf{x} - \mathbf{x}^0)^T \mathbf{H}\, f(\mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0))(\mathbf{x} - \mathbf{x}^0) \\ &> 0 \end{aligned}$$

mit $\theta \in (0,1)$, *d.h.* $f(\mathbf{x})$ *hat in* $\mathbf{x}^0$ *ein strenges lokales Minimum.*

Klassifikation: Beweis von Teil 2 - II

Beweis

Ist $\mathbf{H}\, f(\mathbf{x}^0)$ *indefinit, so existieren zu Eigenwerten von* $\mathbf{H}\, f(\mathbf{x}^0)$ *mit verschiedenen Vorzeichen gewisse Eigenvektoren* $\mathbf{v}, \mathbf{w}$ *mit (zum Beispiel)*

$$\mathbf{v}^T \mathbf{H} f(\mathbf{x}^0)\mathbf{v} > 0 \qquad \mathbf{w}^T \mathbf{H} f(\mathbf{x}^0)\mathbf{w} < 0$$

Klassifikation: Beweis von Teil 2 - III

Beweis.

Analog zu Teil 1) sieht man dann, dass $\varepsilon_1, \varepsilon_2 > 0$ existieren mit

$$f(\mathbf{x}^0 + \varepsilon\mathbf{v}) - f(\mathbf{x}^0) = \frac{1}{2}(\varepsilon\mathbf{v})^T\, \mathbf{H} f(\mathbf{x}^0 + \theta_1\varepsilon\mathbf{v})\,(\varepsilon\mathbf{v}) > 0$$

für alle $\varepsilon \in (0, \varepsilon_1)$ und

$$f(\mathbf{x}^0 + \varepsilon\mathbf{w}) - f(\mathbf{x}^0) = \frac{1}{2}(\varepsilon\mathbf{w})^T\, \mathbf{H} f(\mathbf{x}^0 + \theta_1\varepsilon\mathbf{w})\,(\varepsilon\mathbf{w}) < 0$$

für alle $\varepsilon \in (0, \varepsilon_2)$.

Damit ist gezeigt, dass $\mathbf{x}^0$ ein **Sattelpunkt** von $f(\mathbf{x})$ ist. □

Eigenwerte der Hesse–Matrix

Bemerkung (Geometrische Interpretation)

Die Hesse–Matrix kann positive und negative Eigenwerte besitzen. Die zugehörigen Eigenvektoren geben dabei diejenigen Richtungen an, in denen die Funktion wächst beziehungsweise fällt.

Ausgeartete stationäre Punkte

Bemerkung

1) *Ein stationärer Punkt* $\mathbf{x}^0$ *mit* $\det \mathbf{H}f(\mathbf{x}^0) = 0$ *heißt* **ausgeartet**.
   Die Hesse–Matrix besitzt dann einen Eigenwert $\lambda = 0$.
   Ist $\mathbf{x}^0$ **nicht** *ausgeartet, so existieren genau drei Fälle:*

$$\text{Alle Eigenwerte} > 0 \quad \Rightarrow \quad \mathbf{x}^0 \text{ ist strenges lokales Minimum}$$

$$\text{Alle Eigenwerte} < 0 \quad \Rightarrow \quad \mathbf{x}^0 \text{ ist strenges lokales Maximum}$$

$$\exists \text{ Eigenwerte } \lambda_1 \cdot \lambda_2 < 0 \quad \Rightarrow \quad \mathbf{x}^0 \text{ Sattelpunkt}$$

Definitheit der Hesse–Matrix

Bemerkung (Fortsetzung)

2) *Es gelten die folgenden Implikationen:*

$$\begin{array}{ccc} \mathbf{x}^0 \text{ lokales Minimum} & \Leftarrow & \mathbf{x}^0 \text{ strenges lokales Minimum} \\ \Downarrow & & \Uparrow \\ \mathbf{H}f(\mathbf{x}^0) \text{ positiv semidefinit} & \Leftarrow & \mathbf{H}f(\mathbf{x}^0) \text{ positiv definit} \end{array}$$

Abschätzungen mit der Hesse–Matrix

Bemerkung (Fortsetzung)

3) *Ist $f(\mathbf{x})$ eine $\mathcal{C}^3$–Funktion, $\mathbf{x}^0$ ein stationärer Punkt von $f(\mathbf{x})$ und $\mathbf{H}f(\mathbf{x}^0)$ positiv definit, so gilt die Abschätzung:*

$$(\mathbf{x}-\mathbf{x}^0)^T\,\mathbf{H}f(\mathbf{x}^0)\,(\mathbf{x}-\mathbf{x}^0) \geq \lambda_{min} \cdot \|\mathbf{x}-\mathbf{x}^0\|^2$$

wobei λ_{min} den **kleinsten** *Eigenwert der Hesse–Matrix bezeichnet. Nach dem Satz von Taylor gilt dann:*

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}^0) &\geq \frac{1}{2}\lambda_{min}\|\mathbf{x}-\mathbf{x}^0\|^2 + R_3(\mathbf{x};\mathbf{x}^0) \\ &\geq \|\mathbf{x}-\mathbf{x}^0\|^2\left(\frac{\lambda_{min}}{2} - C\|\mathbf{x}-\mathbf{x}^0\|\right) \end{aligned}$$

Abschätzungen mit der Hesse–Matrix II

Bemerkung (Fortsetzung)

3) *Nach dem Satz von Taylor gilt dann:*

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}^0) &\geq \frac{1}{2}\lambda_{min}\|\mathbf{x} - \mathbf{x}^0\|^2 + R_3(\mathbf{x};\mathbf{x}^0) \\ &\geq \|\mathbf{x} - \mathbf{x}^0\|^2 \left(\frac{\lambda_{min}}{2} - C\|\mathbf{x} - \mathbf{x}^0\|\right) \end{aligned}$$

mit einer geeigneten Konstanten $C > 0$.
In der Nähe von $\mathbf{x}^0$ wächst $f(\mathbf{x})$ also mindestens quadratisch mit dem Abstand von $\mathbf{x}^0$.

Beispiel

Wir betrachten die Funktion

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

und suchen die stationären Punkte:

$$\operatorname{grad} f(x, y) = (y^2 + x(3x + 2), 2y(x - 1))^T.$$

Die Bedingung $\operatorname{grad} f(x, y) = 0$ liefert die beiden stationären Punkte

$$\mathbf{x}^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \mathbf{x}^0 = \begin{pmatrix} -2/3 \\ 0 \end{pmatrix}$$

Die Hesse–Matrizen an den Stellen $\mathbf{x}^0$ und $\mathbf{x}^1$ lauten

$$\mathbf{H}f(\mathbf{x}^0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \qquad \mathbf{H}f(\mathbf{x}^1) = \begin{pmatrix} -2 & 0 \\ 0 & -10/3 \end{pmatrix}$$

Beispiel (Fortsetzung)

$$\mathbf{H}f(\mathbf{x}^0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \qquad \mathbf{H}f(\mathbf{x}^1) = \begin{pmatrix} -2 & 0 \\ 0 & -10/3 \end{pmatrix}$$

Die Matrix $\mathbf{H}f(\mathbf{x}^0)$ ist indefinit, also ist $\mathbf{x}^0$ ein Sattelpunkt, $\mathbf{H}f(\mathbf{x}^1)$ ist negativ definit, also ist $\mathbf{x}^1$ ein strenges lokales Maximum von $f(\mathbf{x})$.

Implizit definierte Funktionen

Untersuche die Lösungsmengen von nichtlinearen Gleichungssystemen der Form

$$\mathbf{g}(\mathbf{x}) = 0$$

mit $\mathbf{g} : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, d.h. wir betrachten m Gleichungen für n Unbekannte.Insbesondere gelte:

$$m < n$$

d.h. wir haben **weniger** Gleichungen als Unbekannte.
Man nennt dann das Gleichungssystem **unterbestimmt** und die Lösungsmenge $G \subset \mathbb{R}^n$ enthält gewöhnlich unendlich viele Punkte.

Lokales Auflösen

Frage: Kann man das System $\mathbf{g}(\mathbf{x}) = 0$ nach bestimmten Unbekannten, zum Beispiel den letzten m Variablen $x_{n-m+1}, \ldots, x_n$, **auflösen**? Existiert eine Funktion $\mathbf{f}(x_1, \ldots, x_{n-m})$ mit

$$\mathbf{g}(\mathbf{x}) = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x_{n-m+1} \\ \vdots \\ x_n \end{pmatrix} = \mathbf{f}(x_1, ..., x_{n-m})$$

Auflösen bedeutet also die letzten m Variablen durch die ersten $n-m$ Variablen zu beschreiben.
Weitere Frage: Nach welchen m Variablen lässt sich das Gleichungssystem auflösen? Ist die Auflösung global auf dem Definitionsbereich D möglich oder nur lokal auf einer Teilmenge $\tilde{D} \subset D$?

Geometrische Interpretation

Die Lösungsmenge G von $\mathbf{g}(\mathbf{x}) = 0$ lässt sich (zumindest lokal) als Graph einer Funktion $\mathbf{f}$ darstellen.

Beispiel

1. $F(x, y) = y - x^2 =$ erlaubt die Auflösung $y = x^2$ auf $\mathbb{R}$.
2. Sei $F(x, y) = x^2 - y^2$. Welche Auflösungen gibt es?

Kreisgleichung als Beispiel für lokales Auflösen

Beispiel

Die Kreisgleichung

$$g(x, y) = x^2 + y^2 - r^2 = 0 \qquad (r > 0)$$

definiert ein **unterbestimmes** nichtlineares Gleichungssystem. Wir haben **zwei** Unbekannte (x, y), aber nur eine Gleichung. Die Kreisgleichung lässt sich **lokal** auflösen und definiert dabei die folgenden vier Funktionen:

Kreisgleichung als Beispiel für lokales Auflösen II

Beispiel

$$\begin{aligned}
y &= \sqrt{r^2 - x^2}, \quad -r \leq x \leq r \\
y &= -\sqrt{r^2 - x^2}, \quad -r \leq x \leq r \\
x &= \sqrt{r^2 - y^2}, \quad -r \leq y \leq r \\
x &= -\sqrt{r^2 - y^2}, \quad -r \leq y \leq r
\end{aligned}$$

Affine Abbildung als Beispiel für lokales Auflösen

Beispiel

Sei $\mathbf{g}(\mathbf{x})$ eine affin–lineare Funktion, d.h.

$$\mathbf{g}(\mathbf{x}) = \mathbf{C}\mathbf{x} + \mathbf{b}, \quad \mathbf{C} \in \mathbb{R}^{(m,n)}, \quad \mathbf{b} \in \mathbb{R}^m$$

Wir spalten die Variablen $\mathbf{x}$ in zwei Vektoren auf

$$\begin{aligned} \mathbf{x}^{(1)} &= (x_1, \ldots, x_{n-m})^T \in \mathbb{R}^{n-m} \\ \mathbf{x}^{(2)} &= (x_{n-m+1}, \ldots, x_n)^T \in \mathbb{R}^n \end{aligned}$$

Affine Abbildung als Beispiel für lokales Auflösen II

Beispiel (Fortsetzung)

Dies führt zur Aufspaltung der Matrix $\mathbf{C}$ mit der Darstellung

$$\mathbf{g}(\mathbf{x}) = \mathbf{B}\mathbf{x}^{(1)} + \mathbf{A}\mathbf{x}^{(2)} + \mathbf{b}$$

mit $\mathbf{B} \in \mathbb{R}^{(m,n-m)}$, $\mathbf{A} \in \mathbb{R}^{(m,m)}$.
Das Gleichungssystem $\mathbf{g}(\mathbf{x}) = 0$ ist genau dann nach den Variablen $\mathbf{x}^{(2)}$ (eindeutig) auflösbar, falls $\mathbf{A}$ regulär ist:

$$\mathbf{g}(\mathbf{x}) = 0 \quad \Leftrightarrow \quad \begin{aligned} \mathbf{x}^{(2)} &= -\mathbf{A}^{-1}(\mathbf{B}\mathbf{x}^{(1)} + \mathbf{b}) \\ &= \mathbf{f}(\mathbf{x}^{(1)}) \end{aligned}$$

Wie kann man die Matrix $\mathbf{A}$ in Abhängigkeit von $\mathbf{g}$ schreiben?

Affine Abbildung als Beispiel für lokales Auflösen III

Beispiel (Fortsetzung)

Aus der Darstellung

$$\mathbf{g}(\mathbf{x}) = \mathbf{B}\mathbf{x}^{(1)} + \mathbf{A}\mathbf{x}^{(2)} + \mathbf{b}$$

erkennt man direkt, dass

$$\mathbf{A} = \frac{\partial \mathbf{g}}{\partial \mathbf{x}^{(2)}}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$$

gilt, d.h. $\mathbf{A}$ ist die Jacobi–Matrix der Abbildung $\mathbf{x}^{(2)} \to \mathbf{g}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ für festes $\mathbf{x}^{(1)}$!

Auflösbarkeitsbedingung

Auflösbarkeit ist also gegeben, falls die Jacobi–Matrix regulär ist. Dies führt auf den

Satz über implizite Funktionen

Satz über implizite Funktionen

Satz

Sei $g : D \to \mathbb{R}^m$ eine $\mathcal{C}^1$–Funktion, $D \subset \mathbb{R}^n$ offen. Die Variablen in D seien $(\mathbf{x}, \mathbf{y})$ mit $\mathbf{x} \in \mathbb{R}^{n-m}$ und $\mathbf{y} \in \mathbb{R}^m$. Der Punkt $(\mathbf{x}^0, \mathbf{y}^0) \in D$ sei eine Lösung von $\mathbf{g}(\mathbf{x}^0, \mathbf{y}^0) = \mathbf{0}$. Falls

$$\frac{\partial \mathbf{g}}{\partial \mathbf{y}}(\mathbf{x}^0, \mathbf{y}^0) := \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\mathbf{x}^0, \mathbf{y}^0) & \dots & \frac{\partial g_1}{\partial y_m}(\mathbf{x}^0, \mathbf{y}^0) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1}(\mathbf{x}^0, \mathbf{y}^0) & \dots & \frac{\partial g_m}{\partial y_m}(\mathbf{x}^0, \mathbf{y}^0) \end{pmatrix}$$

regulär ist, gibt es Umgebungen U von $\mathbf{x}^0$ und V von $\mathbf{y}^0$, $U \times V \subset D$ und eine eindeutig bestimmte stetig differenzierbare Funktion $\mathbf{f} : U \to V$ mit

Satz über implizite Funktionen – Aussage

Satz

$$\mathbf{f}(\mathbf{x}^0) = \mathbf{y}^0 \quad \textit{und} \quad \mathbf{g}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{0} \ (\forall\, \mathbf{x} \in U)$$

und

$$\mathbf{J}\,\mathbf{f}(\mathbf{x}) = -\left(\frac{\partial \mathbf{g}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{f}(\mathbf{x}))\right)^{-1} \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{f}(\mathbf{x}))\right)$$