Gradientenverfahren – Maximierungsprobleme (numerisches Verfahren)

Idee: • An kritischen Punkten verschwindet der Gradient $\nabla f(x) \stackrel{!}{=} 0$

• Der Gradient zeigt in Richtung des steilsten Anstiegs

1. Initialisierung

Wähle einen Startpunkt $x^0 \in \mathbb{R}$, eine Toleranz ε und Parameter $\sigma, \rho \in (0,1)$. Setze $k=0$.

2. Abbruchkriterium

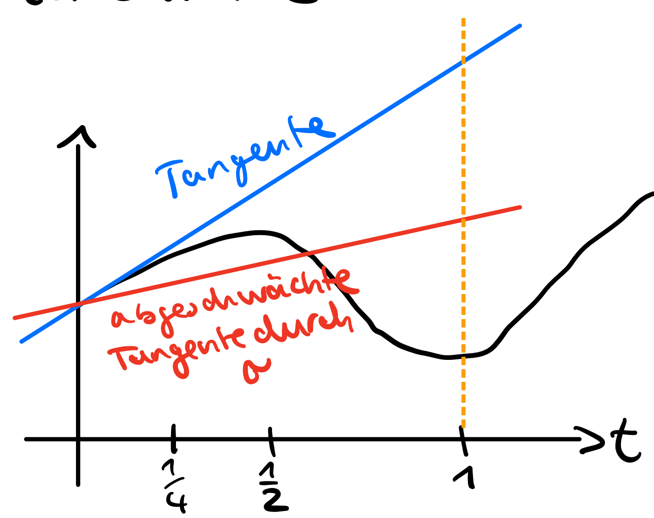
- falls $\|\nabla f(x^k)\|_2 < \varepsilon \Rightarrow$ Abbruch, x^k ist eine geeignete Approximation des lokalen Maxima
- falls $\|\nabla f(x^k)\|_2 \geq \varepsilon \Rightarrow$ weiter zu Schritt 3

Algorithmus 1 Armijo-Regel zur inexakten eindimensionalen Maximierung

Input: Eine stetig differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ sowie $\overline{x}$, $d \in \mathbb{R}^n$ mit $\langle \nabla f(\overline{x}), d\rangle > 0$ und Parameter $\sigma, \rho \in (0,1)$.

1: Setze $t^0 = 1$ und $k = 0$
2: **while** $f(\overline{x} + t^k d) < f(\overline{x}) + t^k \sigma \langle \nabla f(\overline{x}), d\rangle$ **do**
3: Setze $t^{k+1} = \rho t^k$
4: Ersetze k durch $k+1$
5: **end while**

Output: Schrittweite t^k mit Mindestanstieg in f



3. Annäherungsschritt

d = Suchrichtung

- bestimme steilsten Anstieg in Punkt x^k: $d^k = \nabla f(x^k)$
- bestimme Schrittweite $t_0^k = 1$: bestimme $f(x^k + t_0^k d^k)$
  ↳ Armijo-Regel
  bestimme $f(x^k) + t_0^k \sigma \nabla f(x^k)^T d^k$

- falls $f(x^k + t_0^k d^k) \geq f(x^k) + t_0^k \sigma \nabla f(x^k)^T d^k \Rightarrow$ Stopp, Schrittweite t_0^k
- falls $f(x^k + t_0^k d^k) < f(x^k) + t_0^k \sigma \nabla f(x^k)^T d^k \Rightarrow$ Setze $t_1^k = t_0^k \cdot \rho$

↳ bestimme erneut Schrittweite mit t_i^k

→ wiederholen bis $f(x^k + t_i^k d^k) \geq f(x^k) + t_i^k \sigma \nabla f(x^k)^T d^k$ gilt

- Setze $x^{k+1} = x^k + t^k d^k$

4. Setze $k = k+1$ und gehe zu Schritt 2.

Für Minimierungsprobleme ändert sich Schritt 3:

Setze $d^k = -\nabla f(x^k)$

Bei Armijo-Regel ändern sich Ungleichungen

Algorithmus 1 Armijo-Regel zur inexakten eindimensionalen ~~Maximierung~~ Minimierung

---

Input: Eine stetig differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ sowie $\bar{x}, d \in \mathbb{R}^n$ mit $\langle \nabla f(\bar{x}), d \rangle < 0$ und Parameter $\sigma, \rho \in (0,1)$.

Setze $t^0 = 1$ und $k = 0$

while $f(\bar{x} + t^k d) > f(\bar{x}) + t^k \sigma \langle \nabla f(\bar{x}), d \rangle$ **do**

  Setze $t^{k+1} = \rho t^k$

  Ersetze k durch $k+1$

end while

Output: Schrittweite t^k mit Mindest~~an~~stieg in f Mindestabstieg

---

Newtonverfahren (numerisches Verfahren)

1. Initialisierung

- Wähle einen Startpunkt $x^0 \in \mathbb{R}$ und eine Toleranz $\varepsilon > 0$

  Setze $k = 0$.

2. Abbruchkriterium

- falls $\|\nabla f(x^k)\|_2 < \varepsilon \Rightarrow$ Abbruch, x^k ist eine geeignete Approximation des lokalen Maxim

- falls $\|\nabla f(x^k)\|_2 \geq \varepsilon$ $\Rightarrow$ weiter zu Schritt 3

3. Berechne Suchrichtung d^0 als Lösung des Gleichungssystems
   - $D^2 f(x^0)\, d^0 = -\nabla f(x^0)$

   $\Rightarrow d^0 = -\nabla f(x^0) \left[D^2 f(x^0)\right]^{-1}$
   - Setze $x^{k+1} = x^k + t^k d^k$

     $\rightarrow$ ungedämpftes Newtonverfahren : $t = 1$

4. Setze $k = k+1$ und gehe zu Schritt 2.

---

konvexe unrestringierte NLO

- lokale Extrempunkte sind auch globale Extrempunkte
- konvex $\rightarrow$ Minimierung
- konkav $\rightarrow$ Maximierung

C^2-Charakterisierung von Konvexität

Eine zweimal stetig differenzierbare Funktion f ist genau dann konvex, wenn ihre sämtlichen Hessematrizen positiv semidefinit sind, d.h. wenn $D^2 f(x)$ für alle $x \in \mathbb{R}^n$ nur nichtnegative Eigenwerte besitzt.

Aufgabe 1

Gegeben sei die Funktion

$$f(x_1, x_2) = -x_1 - x_1^2 + x_1 x_2^2 - 1.$$

Wenden Sie das Gradientenverfahren zur lokalen Maximierung der Funktion f beginnend mit dem Startpunkt $x^0 = (-1, -1)$ und einer Toleranz $\varepsilon = 10^{-4}$ an. Benutzen Sie zur Schrittweitensteuerung die Armijo-Regel mit $\sigma = 0.2$ und $\rho = 0.5$ gemäß Algorithmus 1:

Algorithmus 1 Armijo-Regel zur inexakten eindimensionalen Maximierung

Input: Eine stetig differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ sowie $\overline{x}, d \in \mathbb{R}^n$ mit $\langle \nabla f(\overline{x}), d \rangle > 0$ und Parameter $\sigma, \rho \in (0, 1)$.

```
Setze t^0 = 1 und k = 0
while f(x̄ + t^k d) < f(x̄) + t^k σ ⟨∇f(x̄), d⟩ do
    Setze t^{k+1} = ρ t^k
    Ersetze k durch k + 1
end while
```

Output: Schrittweite t^k mit Mindestanstieg in f

$$\nabla f(x_1, x_2) = \begin{pmatrix} -1 - 2x_1 + x_2^2 \\ 2x_1 x_2 \end{pmatrix}$$

1. Startpunkt $x^0 = (-1, -1)$, Tolerant $\varepsilon = 10^{-4}$, $\sigma = 0.2$, $\rho = 0.5$, $k = 1$

2. Abbruchkriterium:

$$\|\nabla f(x^0)\|_2 = \|\nabla f(-1, -1)\|_2$$
$$= \left\| \begin{pmatrix} -1 - 2(-1) + (-1)^2 \\ 2(-1)(-1) \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\|_2$$
$$= \sqrt{2^2 + 2^2} = \sqrt{8} \geq \varepsilon$$

3. Steilster Anstieg im Punkt x^0: $d^0 = \nabla f(-1, -1) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$

Bestimme Schrittweite:

Für $t_0^0 = 1$ gilt:

- $f(x^0 + t_0^0 d^0) = f(-1 + 1 \cdot 2, -1 + 1 \cdot 2) = f(1, 1) = -2$
- $f(x^0) + t_0^0 \sigma \nabla f(x^0)^T d^0 = -2 + 1 \cdot 0{,}2 \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix}^T \begin{pmatrix} 2 \\ 2 \end{pmatrix}$

$$= -2 + 1 \cdot 0{,}2 \cdot 8 = -0{,}4$$

$\Rightarrow -2 < -0{,}4$

Für $t_1^0 = \rho \cdot t_0^0 = 0.5\, t_0^0 = 0.5 \cdot 1 = 0.5$ gilt:

- $f(x^0 + t_1^0 d^0) = f(-1 + 0.5 \cdot 2, -1 + 0.5 \cdot 2) = f(0, 0) = -1$
- $f(x^0) + t_1^0 \sigma \nabla f(x^0)^T d^0 = -2 + 0.5 \cdot 0{,}2 \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix}^T \begin{pmatrix} 2 \\ 2 \end{pmatrix} = -1.2$

$\Rightarrow -1 \geq -1{,}2$ $\Rightarrow$ Schrittweite $t^0 = t_1^0 = 0.5$

Setze $x^1 = x^0 + t^0 d^0 = \begin{pmatrix} -1 + 0{,}5 \cdot 2 \\ -1 + 0{,}5 \cdot 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

4. Setze $k = 2$ und gehe zu Schritt 2

2. Abbruchkriterium:

$$\|\nabla f(x^1)\|_2 = \|\nabla f(0,0)\|_2$$

$$= \left\| \begin{pmatrix} -1 & -2 \cdot 0 & + 0^2 \\ & 2 \cdot 0 \cdot 0 & \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\|_2$$

$$= \sqrt{1^2} = 1 \geq \varepsilon$$

3. Steilster Anstieg im Punkt x^1: $d^1 = \nabla f(0,0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$

Bestimme Schrittweite:

- Für $t_0^1 = 1$ gilt:
  - $f(x^1 + t_0^1 d^1) = f(0 + 1 \cdot (-1), 0 + 1 \cdot 0) = f(-1, 0) = -1$
  - $f(x^1) + t_0^1 \sigma \nabla f(x^1)^T d^1 = -1 + 1 \cdot 0{,}2 \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix}^T (-1\ 0)$

$$= -1 + 1 \cdot 0{,}2 \cdot 1 = -0.8$$

$$-1 < -0{,}8$$

Für $t_1^1 = 0.5\, t_0 = 0.5 \cdot 1 = 0.5$ gilt:

- $f(x^1 + t_1^1 d^1) = f(-0.5, 0) = -0{,}75$
- $f(x^1) + t_1^1 \sigma \nabla f(x^1)^T d^1 = -1 + 0.1 = -0.9$

$\Rightarrow -0.75 \geq -0{,}9 \Rightarrow$ Schrittweite $t^1 = t_1^1 = 0.5$

Setze $x^2 = x^1 + t^1 d^1 = \begin{pmatrix} 0 + 0.5\,(-1) \\ 0 + 0.5\,(0) \end{pmatrix} = \begin{pmatrix} -0.5 \\ 0 \end{pmatrix}$

4. Setze $k = 3$ und gehe zu Schritt 2

2. Abbruchkriterium:

$$\|\nabla f(x^2)\|_2 = \|\nabla f(-0.5, 0)\|_2 = \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| = 0 < \varepsilon$$

$\Rightarrow$ Abbruch, $x^{max} = x^2 = \begin{pmatrix} -0{,}5 \\ 0 \end{pmatrix}$ ist eine geeignete Approximation des lokalen Minimalpunktes (hier sogar exakte Lösung)

Aufgabe 2

Gegeben sei die Funktion $f: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1^2 - 2x_1x_2 - x_2^2 x_1 + 3.$

a) Berechnen Sie die kritischen Punkte von f.

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 - 2x_2 - x_2^2 \\ -2x_1 - 2x_1 x_2 \end{pmatrix} \qquad D^2 f(x_1, x_2) = \begin{pmatrix} 2 & -2-2x_2 \\ -2-2x_2 & -2x_1 \end{pmatrix}$$

Zu a)

$$\nabla f(x_1, x_2) \stackrel{!}{=} 0$$

$$2x_1 - 2x_2 - x_2^2 = 0$$

$$-2x_1 - 2x_1 x_2 = 0$$

$$\Leftrightarrow -2x_1(1 + x_2) = 0$$

$\hookrightarrow$ Fall 1: $x_1 = 0$

$\Rightarrow -2x_2 - 2x_2^2 = 0 \Rightarrow x_2 \in \{0, -2\}$

$\Rightarrow z^1 = (0, 0)$, $z^2 = (0, -2)$

$\hookrightarrow$ Fall 2: $x_1 \neq 0$

$\Rightarrow x_2 = -1 \Rightarrow 2x_1 + 2 - 1 = 0$

$\Leftrightarrow x_1 = -\frac{1}{2}$

$\Rightarrow z^3 = (-\frac{1}{2}, -1)$

b) Bestimmen Sie deren Charakteristika (Maximalpunkt, Minimalpunkt, Sattelpunkt). Besitzt die Funktion einen globalen Maximal- oder Minimalpunkt? Begründen Sie Ihre Antwort.

A2b) $D^2 f(z^1) = D^2 f(0,0) = \begin{pmatrix} 2 & -2 \\ -2 & 0 \end{pmatrix}$

Alternative Methode zur Bestimmung der Definitheit: Hurwitz-Kriterium

$\det(D^2 f(z^1) - \lambda I) = \lambda^2 - 2\lambda - 4 \stackrel{!}{=} 0$

$\Rightarrow \lambda_1, \lambda_2 = 1 \pm \sqrt{1+4} = 1 \pm \underbrace{\sqrt{5}}_{>1=\sqrt{1}}$

$\Rightarrow D^2 f(z^1)$ indefinit, z^1 Sattelpunkt

$D^2 f(z^2) = D^2 f(0,-2) = \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix}$

$\det(D^2 f(z^2) - \lambda I) = \lambda^2 - 2\lambda - 4 \stackrel{!}{=} 0$

$\Rightarrow \lambda_1, \lambda_2 = 1 \pm \sqrt{1+4} = 1 \pm \underbrace{\sqrt{5}}_{>1=\sqrt{1}}$

$\Rightarrow D^2 f(z^2)$ indefinit, z^2 Sattelpunkt

$D^2 f(z^3) = D^2 f(-\frac{1}{2}, -1) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

$\Rightarrow$ Eigenwerte ablesen

$\Rightarrow \lambda_1 = 2, \lambda_2 = 1$

$\Rightarrow D^2 f(z^3)$ positiv definit, z^3 lokaler Minimalpunkt

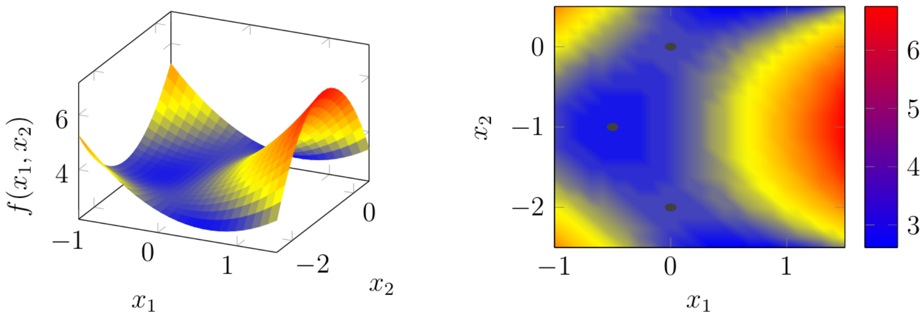
Untersuchung von f auf globalen Maxi-/Minimalpunkt:

Für $x^{(k)} = (k,0)$ $(k \in \mathbb{N})$ gilt: $f(x^{(k)}) = k^2 + 3 \xrightarrow{k\to\infty} \infty$

Für $x^{(k)} = (1,k)$ $(k \in \mathbb{N})$ gilt: $f(x^{(k)}) = 1 - 2k - k^2 + 3 \xrightarrow{k\to\infty} -\infty$

$\longrightarrow$ versuchen für $x_1 = k$ und $x_2 = -1/0/1$

f ist also weder nach oben, noch nach unten beschränkt und hat daher keine globalen Maxi-/Minima



c) Beim ungedämpften Newtonverfahren beträgt die Schrittweite in jedem Schritt $t = 1$. Führen Sie das Verfahren mit Toleranz $\varepsilon = 10^{-2}$ und ausgehend von verschiedenen Startpunkten $x^0 \in \{(-1, 0), (1, 0), (-1, -1)\}$ durch. Welchen kritischen Punkt erhalten Sie jeweils in den einzelnen Fällen?

A2c) Startpunkt $x^0 = (-1, 0)$

1. Initialisierung $x^0 = (-1, 0)$, Toleranz $\varepsilon = 10^{-2}$, $k = 0$

2. Abbruchkriterium

$$\|\nabla f(x^0)\|_2 = \left\| \begin{pmatrix} -2 \\ 2 \end{pmatrix} \right\|_2 = \sqrt{8} \geq \varepsilon$$

3. Berechne Suchrichtung d^0

$$D^2 f(x^0)\, d^0 = -\nabla f(x^0)$$

$$\Leftrightarrow \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \cdot \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = -\begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

Das Verfahren bricht ab, weil $D^2 f(x^0)$ nicht invertierbar ist

→ $D^2 f(x^0)$ hat keinen vollen Rang

→ in Praxis jetzt 1 mal Gradientenverfahren

Startpunkt $x^0 = (1, 0)$:

1. Initialisierung $x^0 = (1, 0)$, Toleranz $\varepsilon = 10^{-2}$, $k = 0$

2. Abbruchkriterium

$$\|\nabla f(x^0)\|_2 = \left\| \begin{pmatrix} 2 \\ -2 \end{pmatrix} \right\|_2 = \sqrt{8} \geq \varepsilon$$

3. Berechne Suchrichtung d^0

$$D^2 f(x^0)\, d^0 = -\nabla f(x^0)$$

$$\Leftrightarrow \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix} \cdot \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = -\begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

2 Möglichkeiten:

- LGS aufstellen
- Matrix invertieren

LGS: $\begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix} \cdot \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = -\begin{pmatrix} 2 \\ -2 \end{pmatrix}$

$$\Leftrightarrow \begin{array}{rrcl} 2d_1^0 & -2d_2^0 & = & -2 \\ -2d_1^0 & -2d_2^0 & = & 2 \end{array} \quad +$$

$$\Leftrightarrow \begin{array}{rrcl} 2d_1^0 & -2d_2^0 & = & 0 \\ & -4d_2^0 & = & 0 \end{array}$$

$$\Rightarrow d_1^0 = -1 \quad , d_2^0 = 0$$

Matrix: invertieren von $\begin{pmatrix} -2 & 2 \\ 2 & 2 \end{pmatrix}$

$$\left(\begin{array}{rr|rr} 2 & -2 & 1 & 0 \\ -2 & -2 & 0 & 1 \end{array}\right) \; + \longrightarrow \left(\begin{array}{rr|rr} 2 & -2 & 1 & 0 \\ 0 & -4 & 1 & 1 \end{array}\right) \; \cdot -\tfrac{1}{2}$$

$$\longrightarrow \left(\begin{array}{rr|rr} 2 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & -4 & 1 & 1 \end{array}\right) \begin{array}{l} \cdot \frac{1}{2} \\ \cdot -\frac{1}{4} \end{array} \longrightarrow \left(\begin{array}{rr|rr} 1 & 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & 1 & -\frac{1}{4} & -\frac{1}{4} \end{array}\right) \Rightarrow \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix}^{-1} = \frac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix} \cdot \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = -\begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix}^{-1} \cdot -\begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} \cdot -\begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow d_1^0 = -1 \quad , d_2^0 = 0$$

Es ergibt sich $d^0 = (-1, 0)$ und damit

$$x^1 = x^0 + t^0 d^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

4. Setze $k=1$ und gehe zu Schritt 2

2. Abbruchkriterium

$$\|\nabla f(x^1)\|_2 = \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\|_2 = 0 < \varepsilon$$

$\Rightarrow$ Abbruch, $x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ist eine geeignete Approximation für einen kritischen Punkt (hier sogar exakte Lösung)

Startpunkt $x^0 = (-1, -1)$

1. Initialisierung $x^0 = (-1,-1)$, Toleranz $\varepsilon = 10^{-2}$, $k=0$

2. Abbruchkriterium

$$\|\nabla f(x^0)\|_2 = \left\| \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\|_2 = \sqrt{1} \geq \varepsilon$$

3. Berechne Suchrichtung d^0

$$D^2 f(x^0) d^0 = -\nabla f(x^0)$$

$$\Leftrightarrow \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} d_1^0 \\ d_2^0 \end{pmatrix} = -\begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Es ergibt sich $d^0 = (\frac{1}{2}, 0)$ und damit

$$x^1 = x^0 + t^0 d^0 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} + 1 \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix}$$

4. Setze $k=1$ und gehe zu Schritt 2

2. Abbruchkriterium

$$\|\nabla f(x^1)\|_2 = \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\|_2 = 0 < \varepsilon$$

$\Rightarrow$ Abbruch, $x^1 = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix}$ ist eine geeignete Approximation für einen kritischen Punkt (hier sogar exakte Lösung)

Aufgabe 3

Gegeben sei die quadratische Funktion $f(x) = \frac{1}{2}x^T Ax + b^T x + c$ mit $x, b \in \mathbb{R}^n,\ n \in \mathbb{N},\ c \in \mathbb{R}$ und $A \in \mathbb{R}^{n\times n}$, wobei A eine symmetrische positiv definite Matrix sei.

Zeigen Sie, dass das ungedämpfte Newtonverfahren die Funktion f für jeden Startpunkt in einem Schritt minimiert. Bestimmen Sie hierzu zunächst Gradient und Hessematrix von f, indem Sie die Funktion ausführlich aufschreiben.

A3) $\nabla f(x) = \nabla(\frac{1}{2}x^T A x) + \nabla(b^T x) + \nabla(c)$

$$\frac{1}{2}x^T A x = \frac{1}{2}(x_1, \dots, x_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \frac{1}{2}(x_1, \dots, x_n)\begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n \end{pmatrix}$$

$$= \frac{1}{2}\big(a_{11}x_1x_1 + \dots + a_{1n}x_1x_n + \dots + a_{n1}x_nx_1 + \dots + a_{nn}x_nx_n\big)$$

$a_{ii} \cdot 2x_i$

$2a_{ij}x_j$

$$= \frac{1}{2}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j\Big)$$

Es taucht x_j für $j \in \{1, \dots, n\}$ also in jeder j-ten Zeile und Spalte auf. Wegen der Symmetrie tritt $a_{ij}x_ix_j$ mit $i \neq j$ 2 mal in der Summe auf.

$$\nabla \tfrac{1}{2} x^T A x = \tfrac{1}{2}\left(2a_{j1}x_1 + \dots + 2a_{jj}x_j + \dots + 2a_{jn}x_n\right)$$
$$= \sum_{i=1}^{n} a_{ji} x_i$$
$$= \begin{pmatrix} a_{j1} & \dots & a_{jn} \end{pmatrix} \begin{pmatrix} x_1 \\ | \\ x_n \end{pmatrix}$$
$$= Ax$$

$$\nabla (b^T x) = b$$
$$\nabla (c) = 0$$
$$\nabla f(x) = Ax + b \qquad D^2 f(x) = A \succ 0 \quad \text{spd}$$

Newtonverfahren mit beliebigem Startpunkt x^0:

Berechne Suchrichtung d^0:

$$d^0 = -D^2 f(x^0)^{-1} \nabla f(x^0) = -A^{-1}(Ax^0 + b) = -x^0 - A^{-1}b$$

Matrix A invertierbar, da spd (symmetrisch und positiv definit)

$$x^1 = x^0 + t^0 d^0 = x^0 + (-x^0 - A^{-1}b) = -A^{-1}b$$

Abbruchkriterium:

$$\|\nabla f(x^1)\|_2 = \|Ax^1 + b\|_2 = \|A(-A^{-1}b) + b\|_2$$
$$= \|-b + b\|_2 = \|0\|_2 = 0 < \varepsilon$$

$\Rightarrow x^1$ ist die gesuchte Lösung

Hinweis: Wenn x^0 bereits der gesuchte Minimalpunkt ist oder zumindest nahe an diesem liegt, bricht das Verfahren sogar bereits vor dem ersten Schritt (d.h. vor den o.g. Schritten zur Berechnung von x^1) wegen $\|\nabla f(x^0)\|_2 \leq \varepsilon$ ab.

Aufgabe 4

a) Welche der folgenden Funktionen sind konvex, welche konkav und welche sind weder konvex noch konkav?

i) $f(x_1, x_2) = (x_1 + 5)^2 + (x_2 - 2)^2$

ii) $f(x_1, x_2) = -x_1 - 2x_1^2 + x_1^2 x_2 - 2x_2^2$

iii) $f(x_1, x_2, x_3) = -x_1^2 + 2x_1x_3 + x_2^2 - 8x_3^2$

A4a) i) $\nabla f(x_1, x_2) = \begin{pmatrix} 2(x_1+5) \\ 2(x_2-2) \end{pmatrix}$ $D^2 f(x_1,x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \succ 0$

$\rightarrow D^2 f(x_1,x_2)$ positiv definit $\rightarrow$ konvex

ii) $\nabla f(x_1,x_2) = \begin{pmatrix} -1 - 4x_1 + 2x_1x_2 \\ x_1^2 - 4x_2 \end{pmatrix}$ $D^2 f(x_1,x_2) = \begin{pmatrix} -4+2x_2 & 2x_1 \\ 2x_1 & -4 \end{pmatrix}$

$\rightarrow$ Man kann x_1 und x_2 so wählen, dass $D^2 f(x_1,x_2)$ indefinit ist

$\rightarrow$ weder konvex, noch konkav

iii) $\nabla f(x_1,x_2,x_3) = \begin{pmatrix} -2x_1 + 2x_3 \\ 2x_2 \\ 2x_1 - 16x_3 \end{pmatrix}$ $D^2 f(x_1,x_2,x_3) = \begin{pmatrix} -2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 16 \end{pmatrix}$

$\rightarrow \lambda_1 = 2$, $\lambda_{2,3} = -9 \pm \underbrace{\sqrt{53}}_{< 9 = \sqrt{81}}$

$\rightarrow D^2 f(x_1,x_2,x_3)$ indefinit $\rightarrow f$ weder konkav, noch konvex

b) Eine Menge $\mathbb{M} \subset \mathbb{R}^n$ heißt konvex, falls gilt

$$x, y \in \mathbb{M} \quad \Rightarrow \quad \alpha x + (1-\alpha)y \in \mathbb{M} \quad \text{für alle } \alpha \in [0,1].$$

Zeigen Sie, dass der Durchschnitt zweier konvexer Mengen wieder konvex ist.

A4b) Seien $\mathbb{M}_1, \mathbb{M}_2 \subset \mathbb{R}^n$ zwei konvexe Mengen.

Fall 1: $\mathbb{M}_1 \cap \mathbb{M}_2 = \emptyset$

- Die leere Menge $\emptyset$ ist konvex.
- Somit ist $\mathbb{M}_1 \cap \mathbb{M}_2 = \emptyset$ konvex.

Fall 2: $\mathbb{M}_1 \cap \mathbb{M}_2 \neq \emptyset$

- Seien $x, y \in \mathbb{M}_1 \cap \mathbb{M}_2$, dann gilt $x, y \in \mathbb{M}_1$ und $x, y \in \mathbb{M}_2$.
- Da $\mathbb{M}_1$ und $\mathbb{M}_2$ konvex sind, gilt

$$\alpha x + (1-\alpha)y \in \mathbb{M}_1 \quad \text{für } \alpha \in [0,1]$$
und
$$\alpha x + (1-\alpha)y \in \mathbb{M}_2 \quad \text{für } \alpha \in [0,1].$$

- Dies bedeutet aber genau, dass $\alpha x + (1-\alpha)y \in \mathbb{M}_1 \cap \mathbb{M}_2$ für $\alpha \in [0,1]$.
- Somit ist $\mathbb{M}_1 \cap \mathbb{M}_2 \neq \emptyset$ konvex.

Aufgabe 5

Gegeben sei die Funktion

$$f(x_1, x_2) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

mit $x \in \mathbb{R}^2$. Begründen Sie, dass die Funktion einen globalen Minimalpunkt besitzt. Berechnen Sie diesen.

hinreichende Bedingung für globale Minimalpunkte

- Hessematrix für alle x positiv semidefinit
- es gibt mindestens einen kritischen Punkt

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 - 2x_2 - 4 \\ 4x_2 - 2x_1 - 6 \end{pmatrix} \qquad D^2 f(x_1, x_2) = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}$$

→ konstant in x

$$p(\lambda) = \det(D^2 f(x_1, x_2) - \lambda I) = (4 - \lambda)^2 - 4$$
$$= (\lambda - 6)(\lambda - 2)$$

$\Rightarrow \lambda_1 = 6,\ \lambda_2 = 2 \Rightarrow D^2 f(x_1, x_2)$ positiv definit für jedes x

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 - 2x_2 - 4 \\ 4x_2 - 2x_1 - 6 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\Rightarrow x_1 = \frac{7}{3},\ x_2 = \frac{8}{3}$$

Da die Funktion f nur einen kritischen Punkt hat, muss das der globale Minimalpunkt sein.

nächstes Tutorium:

- Modellieren
- Beschränktheit, Abgeschlossenheit und Kompaktheit
- Karush-Kuhn-Tucker-Bedingungen