Wissenschaftliches Rechnen - Großübung 6.1

Themen: Nullstellensuche, Mehrdimensionale Analysis

Ugo & Gabriel

31. Januar 2023

Aufgabe 1: Nullstellensuche

1. Vergleichen Sie die Verfahren Bisektion, Regula falsi und das Newton-Verfahren zum Finden von Nullstellen. Gehen Sie dabei auf folgende Aspekte ein:
   a) Startvorraussetzungen
   b) Iterationsschritt
   c) Muss die Ableitung bekannt sein?
   d) Konvergenzgeschwindigkeit
   e) Garantie auf Konvergenz

Lösung

a) Bisektion und Regula falsi: $f(x^+) > 0, f(x^-) < 0$, Newton: $x \in \mathbb{R}$

b) Bisektion: Falls $f(\overline{x}) > 0$ dann $x^+ \leftarrow \overline{x}$, sonst $x^- \leftarrow \overline{x}$ mit $\overline{x} = \frac{x_1+x_2}{2}$
Regula falsi: Falls $f(\tilde{x}) > 0$ dann $x^+ \leftarrow \tilde{x}$, sonst $x^- \leftarrow \tilde{x}$ mit $\tilde{x} = \frac{x^+ f(x^-) - x^- f(x^+)}{f(x^-) - f(x^+)}$
Newton: $x_{t+1} \leftarrow x_t - \frac{f(x_t)}{f'(x_t)}$

c) Bisektion und Regula falsi: nein, Newton: ja

d) Bisektion und Regula falsi: mindestens linear (Regula falsi gewöhnlich schneller), Newton: mindestens quadratisch

e) Bisektion: ja, Regula falsi und Newton: nein

Lösung Ende

2. Gegeben ist die Funktion $f(x) = x^3 - 2x^2 - 3x + 1$. Führen Sie einen Iterationsschritt mit allen drei oben genannten Verfahren durch. Wählen Sie $x^- = 2$ und $x^+ = -1$ für Bisektion und Regula falsi und $x = 2$ für das Newton-Verfahren.

Lösung

- Bisektion: $\overline{x} = \frac{-1+2}{2} = \frac{1}{2}$, $f\left(\frac{1}{2}\right) = -\frac{7}{8}$, $x^+ = -1$, $x^- = \frac{1}{2}$
- Regula falsi: $\tilde{x} = \frac{(-1)\cdot(-5)-2\cdot 1}{-5-1} = -\frac{3}{6} = -\frac{1}{2}$, $f\left(-\frac{1}{2}\right) = \frac{15}{8}$, $x^+ = \frac{15}{8}$, $x^- = 2$
- Newton: $f'(x) = 3x^2 - 4x - 3$, $x = 2 - \frac{-5}{1} = 7$

Lösung Ende

3. Wie kann man Algorithmen zur Nullstellensuche dazu nutzen, Minima bzw. Maxima zu finden?

   Lösung

   Extrema sind Nullstellen der ersten Ableitung.

   Lösung Ende

4. Wie lautet der Iterationsschritt des Newton-Verfahrens zum finden von Optima einer eindimensionalen Funktion $f : \mathbb{R} \to \mathbb{R}$? Was ist der Zusammenhang zur Taylorentwicklung 2. Ordnung von f?

   Lösung

   Der Iterationsschritt sieht wie folgt aus:

   $$x_{t+1} \leftarrow x_t - \frac{f'(x_t)}{f''(x_t)}$$

   Es wird (zumindest implizit) eine Taylorapproximation 2. Ordnung für f durchgeführt und dessen Extremum ist die neue Stelle.

   Lösung Ende

Für eine Funktion $f : \mathbb{C} \to \mathbb{C}$ nennt man $z \in \mathbb{C}$ einen *Fixpunkt*, falls $f(z) = z$. Eine Fixpunktiteration ist ein numerisches Verfahren, bei dem eine Funktion wiederholt auf einen Startwert z_0 angewendet wird, bis man zu einem Fixpunkt z^* konvergiert, d.h. $z^* = f(\ldots f(f(z_0))\ldots)$.

5. Wie kann man ein Fixpunktproblem $g(z) = z$ in ein Nullstellenproblem $f(z) = 0$ umschreiben?

   Lösung

   Man setzt $f(z) = g(z) - z$.

   Lösung Ende

6. Zeigen Sie, dass man ein Nullstellenproblem $f(z) = 0$ in ein Fixpunktproblem einer Funktion $g : \mathbb{C} \to \mathbb{C}$ umschreiben kann mit

   $$g(z) = f(z)\phi(z) + z,$$

   wobei $\phi : \mathbb{C} \to \mathbb{C}$ eine beliebige Funktion ist mit $\phi(z) \neq 0$. Überlegen Sie sich, warum es sinnvoll sein kann ϕ nicht konstant zu wählen.

   Lösung

   Falls $f(z) = 0$, dann gilt $g(z) = 0 \cdot \phi(z) + z = z$. Die Funktion ϕ kann, mit genügend Vorwissen über f, so gewählt werden, dass die Fixpunktiteration schneller oder zu einem anderen Fixpunkt konvergiert.

   Lösung Ende

7. Zeigen Sie, dass das Newton-Verfahren eine Fixpunktiteration ist, indem Sie die Funktion $g : \mathbb{C} \to \mathbb{C}$ angeben, welche iteriert wird.

   Lösung

   Offensichtlich gilt $g(z) = z - \frac{f(z)}{f'(z)}$.

   Lösung Ende

8. Ein numerisches Verfahren hat eine Konvergenzordnung von $p \geq 1$, falls es eine Konstante $L \geq 0$ und ein $t_0 \in \mathbb{N}$ gibt, sodass

$$|z_{t+1} - z^*| \leq L \cdot |z_t - z^*|^p \quad \text{für alle } t \geq t_0.$$

Zeigen Sie die folgenden Aussagen:

a) Die Fixpunktiteration hat mindestens lineare Konvergenzordnung ($p = 1$).

b) Das Newton-Verfahren hat mindestens quadratische Konvergenzordnung ($p = 2$).

Lösung

a) Für f gibt es folgende Taylorapproximation:

$$f(z) = f(z^*) + \underbrace{f'(\xi)(z - z^*)}_{R(z)},$$

wobei der hintere Term $R(z) \in \mathcal{O}(z)$ das Restglied ist. Dann ist:

$$\begin{aligned}
|z_{t+1} - z^*| &= |f(z_t) - z^*| \\
&= |f(z^*) + f'(\xi)(z_t - z^*) - z^*| \\
&= |z^* + f'(\xi)(z_t - z^*) - z^*| \\
&= |f'(\xi)(z_t - z^*)| \\
&\leq |f'(\xi)||z_t - z^*|
\end{aligned}$$

b) Für f gibt es folgende Taylorapproximation:

$$f(z) = f(z^*) + f'(z^*)(z - z^*) + \underbrace{\frac{1}{2} f''(\xi)(z - z^*)^2}_{R(z)},$$

wobei der hintere Term $R(z) \in \mathcal{O}(z^2)$ das Restglied ist. Dann ist:

$$\begin{aligned}
|z_{t+1} - z^*| &= \left| z_t - \frac{f(z_t)}{f'(z_t)} - z^* \right| \\
&= \left| z_t - \frac{f(z^*) + f'(z^*)(z_t - z^*) + \frac{1}{2} f''(\xi)(z - z^*)^2}{f'(z_t)} - z^* \right| \\
&= \left| z_t - \frac{f'(z^*)(z_t - z^*) + \frac{1}{2} f''(\xi)(z - z^*)^2}{f'(z_t)} - z^* \right| \\
&= \left| z_t - (z_t - z^*) + \frac{f''(\xi)}{2 f'(z_t)} (z - z^*)^2 - z^* \right| \\
&= \left| \frac{f''(\xi)}{2 f'(z_t)} (z - z^*)^2 \right| \\
&\leq \left| \frac{f''(\xi)}{2 f'(z_t)} \right| |z - z^*|^2
\end{aligned}$$

Lösung Ende

Aufgabe 2: Gradient und Hessematrix

Gegeben ist eine beliebige quadratische Funktion $f : \mathbb{R}^n \to \mathbb{R}$, die geschrieben werden kann als

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

wobei $\mathbf{A} \in \mathbb{R}^{n\times n}$ symmetrisch ist, $\mathbf{b} \in \mathbb{R}^n$ und $c \in \mathbb{R}$.

1. Geben Sie f für die folgenden Werte an:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \quad c = 5$$

Lösung

Es gilt:

$$f(x, y) = 2x^2 + 2xy + 2y^2 - 2x - y + 5$$

Lösung Ende

2. Berechnen Sie alle kritischen Punkte von f für die vorgegebenen Werte und klassifizieren Sie diese (Maximum, Minimum, Sattelpunkt).

Lösung

Wir berechnen den Gradienten von f:

$$\nabla f(x, y) = \begin{bmatrix} 4x + 2y - 2 \\ 2x + 4y - 1 \end{bmatrix}$$

Der einzige kritische Punkt ist $(x^*, y^*) = (1/2, 0)$. Die Hessematrix ist

$$\mathbf{H}_f(x, y) = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}.$$

Mit etwas Hinsehen wird klar, dass die konstante Hessematrix positiv definit ist. Damit ist der gefundene kritische Punkt ein Minimum.

Lösung Ende

3. Das Taylorpolynom 2. Ordnung einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ am Entwicklungspunkt $\mathbf{x}_0 \in \mathbb{R}^n$ lässt sich durch

$$T_2(f, \mathbf{x}_0, \mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^\mathsf{T}\mathbf{H}_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^\mathsf{T}\nabla f(\mathbf{x}_0) + f(\mathbf{x}_0)$$

berechnen.

Geben Sie das Taylorpolynom 2. Ordnung von f am Punkt $(x_0, y_0) = (-1, 1)$ an.

Lösung

Da f eine quadratische Funktion ist, ist sie ihr eigenes Taylorpolynom.

Lösung Ende

4. Geben Sie den Gradienten und die Hessematrix von f im allgemeinen Fall an.

Lösung

Es gilt:

$$\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$$

sowie

$$\mathbf{H}_f(\mathbf{x}) = \mathbf{A}$$

Lösung Ende

5. In welchen Fällen hat f kritische Punkte? Geben Sie, falls sie existieren, eine Formel für diese an.

Lösung

Es muss gelten, dass $\nabla f(\mathbf{x}) = \mathbf{0}$, also:

$$\mathbf{A}\mathbf{x} + \mathbf{b} = \mathbf{0} \Leftrightarrow \mathbf{A}\mathbf{x} = -\mathbf{b} \Leftrightarrow \mathbf{x} = -\mathbf{A}^{-1}\mathbf{b}$$

Lösung Ende

6. Eine Funktion heißt *konvex*, falls für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ und alle $\alpha \in [0, 1]$ gilt, dass

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y})$$

Zeigen Sie: Die Funktion f ist genau dann konvex, wenn $\mathbf{A}$ positiv semidefinit ist.

Lösung

Zunächst stellen wir fest, dass alle konstanten und linearen Funktionen konvex sind. Da die Summe konvexer Funktionen ebenfalls konvex ist, reicht es zu zeigen, dass die quadratische Form $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ genau dann konvex ist, wenn $\mathbf{A}$ positiv semidefinit ist. Es gilt:

$$\begin{aligned}
f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) &= (\alpha\mathbf{x} + (1-\alpha)\mathbf{y})^\mathsf{T}\mathbf{A}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \\
&= (\alpha\mathbf{x} + (1-\alpha)\mathbf{y})^\mathsf{T}(\alpha\mathbf{A}\mathbf{x} + (1-\alpha)\mathbf{A}\mathbf{y}) \\
&= \alpha^2\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + 2\alpha(1-\alpha)\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{y} + (1-\alpha)^2\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y}
\end{aligned}$$

sowie

$$\alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}) = \alpha\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + (1-\alpha)\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y}$$

Weiter gilt:

$$\begin{gathered}
f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}) \\
\Leftrightarrow \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}) - f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq 0 \\
\Leftrightarrow \alpha\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + (1-\alpha)\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} - \alpha^2\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} - 2\alpha(1-\alpha)\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{y} - (1-\alpha)^2\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} \geq 0 \\
\Leftrightarrow \alpha(1-\alpha)\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \alpha(1-\alpha)\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} - 2\alpha(1-\alpha)\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{y} \geq 0 \\
\Leftrightarrow \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} - 2\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{y} + \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} \geq 0 \\
\Leftrightarrow (\mathbf{x} - \mathbf{y})^\mathsf{T}\mathbf{A}(\mathbf{x} - \mathbf{y}) \geq 0
\end{gathered}$$

Lösung Ende

7. Für ein $\mathbf{A} \in \mathbb{R}^{n\times m}$ und ein $\mathbf{b} \in \mathbb{R}^n$ möchten wir folgende Funktion $f : \mathbb{R}^m \to \mathbb{R}$ minimieren:

$$f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|^2$$

wobei $\mathbf{A}$ überbestimmt und $\|\cdot\|$ die ℓ^2-Norm ist.

a) Um welches Ihnen wohlbekannte Problem handelt es sich hierbei?

b) Finden Sie einen geschlossenen Ausdruck für einen kritischen Punkt $\mathbf{x}^*$, indem Sie den Gradienten von f berechnen und ihn mit dem Nullvektor gleichsetzen.

c) Begründen Sie, warum der gefundene Punkt ein Minimum ist.

d) Begründen Sie anschließend, warum der gefundene Punkt sogar das globale Minimum von f ist.

Lösung

a) Methode der kleinsten Quadrate / Least Squares (Ausgleichsrechnung)

b) Zunächst gilt:

$$\begin{aligned}\|\mathbf{Ax} - \mathbf{b}\|^2 &= (\mathbf{Ax} - \mathbf{b})^\top(\mathbf{Ax} - \mathbf{b}) \\ &= \mathbf{x}^\top\mathbf{A}^\top\mathbf{Ax} - \mathbf{x}^\top\mathbf{A}^\top\mathbf{b} - \mathbf{b}^\top\mathbf{Ax} + \mathbf{b}^\top\mathbf{b} \\ &= \mathbf{x}^\top\mathbf{A}^\top\mathbf{Ax} - 2\mathbf{b}^\top\mathbf{Ax} + \mathbf{b}^\top\mathbf{b}\end{aligned}$$

Dann gilt:

$$\begin{aligned}\nabla(\|\mathbf{Ax} - \mathbf{b}\|^2) = \mathbf{0} &\Leftrightarrow \nabla(\mathbf{x}^\top\mathbf{A}^\top\mathbf{Ax} - 2\mathbf{b}^\top\mathbf{Ax} + \mathbf{b}^\top\mathbf{b}) = \mathbf{0} \\ &\Leftrightarrow \nabla(\mathbf{x}^\top\mathbf{A}^\top\mathbf{Ax}) - \nabla(2\mathbf{b}^\top\mathbf{Ax}) + \nabla(\mathbf{b}^\top\mathbf{b}) = \mathbf{0} \\ &\Leftrightarrow 2\mathbf{A}^\top\mathbf{Ax} - 2\mathbf{A}^\top\mathbf{b} = \mathbf{0} \\ &\Leftrightarrow 2\mathbf{A}^\top\mathbf{Ax} = 2\mathbf{A}^\top\mathbf{b} \\ &\Leftrightarrow \mathbf{A}^\top\mathbf{Ax} = \mathbf{A}^\top\mathbf{b}\end{aligned}$$

Das entspricht der uns wohlbekannten Normalengleichung. Für invertierbares $\mathbf{A}^\top\mathbf{A}$ ergibt sich folgende geschlossene Form:

$$\mathbf{x}^* = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{b}$$

c) Die Hessematrix ist $\mathbf{H}_f(\mathbf{x}) = \mathbf{A}^\top\mathbf{A}$ und für ein überbestimmtes $\mathbf{A}$ positiv definit.

d) Da f konvex ist, ist jedes lokale Minimum auch ein globales Minimum.

Lösung Ende