Aufgabensammlung

Nichtlineare Programmierung – Gradient und Newton

Deutsch

Wolfgang Böhme

wolfgang.boehme@economic-modeling.de

www.economic-modeling.de

25. Oktober 2026

Dokumentationen dieser Webseite wurden mit Unterstützung der KI Claude (Anthropic) entwickelt.

Inhaltsverzeichnis

1 Stufe 1: Leicht **2**
Aufgabe 1.1 2
Aufgabe 1.2 2

2 Stufe 2: Mittel **4**
Aufgabe 2.1 4
Aufgabe 2.2 5

3 Stufe 3: Schwer **7**
Aufgabe 3.1 7
Aufgabe 3.2 8

1 Stufe 1: Leicht

Aufgabe 1.1 – Stationäre Punkte und Optimalitätsbedingungen

Aufgabe 1.1

Gegeben ist die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x_1, x_2) = x_1^2 + 4x_2^2 - 4x_1 - 8x_2 + 10$$

(a) Berechne den Gradienten $\nabla f(x_1, x_2)$ und die Hesse-Matrix $\mathbf{H}(x_1, x_2)$.

(b) Bestimme alle stationären Punkte (setze $\nabla f = \mathbf{0}$).

(c) Prüfe anhand der Hesse-Matrix, ob der stationäre Punkt ein Minimum, Maximum oder Sattelpunkt ist.

(d) Berechne den Funktionswert im Optimum.

Lösung: Aufgabe 1.1

(a) Gradient und Hesse-Matrix:

$$\nabla f = \begin{pmatrix} 2x_1 - 4 \\ 8x_2 - 8 \end{pmatrix}, \qquad \mathbf{H} = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

(b) Stationärer Punkt: $\nabla f = \mathbf{0}$:

$$2x_1 - 4 = 0 \Rightarrow x_1^* = 2, \qquad 8x_2 - 8 = 0 \Rightarrow x_2^* = 1$$

Einziger stationärer Punkt: $\mathbf{x}^* = (2, 1)^\top$.
(c) Art des stationären Punktes:
Eigenwerte von $\mathbf{H}$: $\lambda_1 = 2 > 0$, $\lambda_2 = 8 > 0 \Rightarrow \mathbf{H} \succ 0$ (positiv definit).
$\Rightarrow \mathbf{x}^* = (2, 1)^\top$ ist ein **striktes globales Minimum**.
(d) Funktionswert:

$$f(2, 1) = 4 + 4 - 8 - 8 + 10 = 2$$

Optimum: $\mathbf{x}^* = (2, 1)^\top$, $f^* = 2$

Aufgabe 1.2 – Ein Schritt Gradientenabstieg

Aufgabe 1.2

Gegeben ist $f(x_1, x_2) = (x_1 - 3)^2 + 2(x_2 + 1)^2$ mit Startpunkt $\mathbf{x}^{(0)} = (1, 1)^\top$.

(a) Berechne $\nabla f(\mathbf{x}^{(0)})$ und die Suchrichtung $\mathbf{d}^{(0)} = -\nabla f(\mathbf{x}^{(0)})$.

(b) Führe einen Gradientenabstiegsschritt mit fester Schrittweite $\alpha = 0{,}25$ durch.

(c) Wie viele Iterationen benötigt der Gradientenabstieg mit optimaler Schrittweite, um das Minimum zu erreichen? (Begründung.)

(d) Berechne die Konditionszahl $\kappa = \lambda_{\max}/\lambda_{\min}$ der Hesse-Matrix.

Lösung: Aufgabe 1.2

(a) Gradient und Suchrichtung:

$$\nabla f = \begin{pmatrix} 2(x_1 - 3) \\ 4(x_2 + 1) \end{pmatrix}, \quad \nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 2(1-3) \\ 4(1+1) \end{pmatrix} = \begin{pmatrix} -4 \\ 8 \end{pmatrix}$$

$$\mathbf{d}^{(0)} = -\nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 4 \\ -8 \end{pmatrix}$$

(b) Gradientenabstiegsschritt mit $\alpha = 0{,}25$:

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + 0{,}25 \cdot \mathbf{d}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0{,}25 \begin{pmatrix} 4 \\ -8 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$f(\mathbf{x}^{(1)}) = (2-3)^2 + 2(-1+1)^2 = 1 + 0 = 1$$

(c) Optimale Schrittweite:
Die Hesse-Matrix ist $\mathbf{H} = \mathrm{diag}(2, 4)$. Da $\mathbf{H}$ diagonal und konstant ist, können wir für jede Komponente separat optimieren. Mit optimaler Schrittweite $\alpha_i^* = 1/\lambda_i$:

$$x_1^{(1)} = 1 + \frac{1}{2} \cdot 4 = 3, \quad x_2^{(1)} = 1 + \frac{1}{4} \cdot (-8) = -1$$

$\mathbf{x}^{(1)} = (3, -1)^\top = \mathbf{x}^*$. Der Gradientenabstieg mit komponentenweiser optimaler Schrittweite erreicht das Minimum in **einer Iteration**.

(d) Konditionszahl:

$$\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{4}{2} = 2$$

Eine Konditionszahl von 2 ist sehr gut – der Gradientenabstieg konvergiert schnell.

Optimum: $\mathbf{x}^* = (3, -1)^\top$, $f^* = 0$, $\kappa = 2$

2 Stufe 2: Mittel

Aufgabe 2.1 – Newton-Methode: Zwei Iterationen

Aufgabe 2.1

Gegeben ist die nichtlineare Funktion:

$$f(x_1, x_2) = x_1^4 + x_2^4 - 4x_1x_2 + 1$$

mit Startpunkt $\mathbf{x}^{(0)} = (2, 2)^\top$.

(a) Berechne ∇f und $\mathbf{H}f$ (allgemein).

(b) Führe **zwei** vollständige Newton-Iterationen durch (mit Schrittweite $\alpha = 1$).

(c) Berechne $\|\nabla f(\mathbf{x}^{(k)})\|$ nach jeder Iteration und beurteile die Konvergenz.

(d) Begründe, warum die Newton-Methode hier schnell konvergiert.

Lösung: Aufgabe 2.1

(a) Gradient und Hesse-Matrix:

$$\nabla f = \begin{pmatrix} 4x_1^3 - 4x_2 \\ 4x_2^3 - 4x_1 \end{pmatrix}, \qquad \mathbf{H} = \begin{pmatrix} 12x_1^2 & -4 \\ -4 & 12x_2^2 \end{pmatrix}$$

(b) Newton-Iteration 1 (bei $\mathbf{x}^{(0)} = (2, 2)^\top$):

$$\nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 4 \cdot 8 - 8 \\ 4 \cdot 8 - 8 \end{pmatrix} = \begin{pmatrix} 24 \\ 24 \end{pmatrix}, \qquad \mathbf{H}(\mathbf{x}^{(0)}) = \begin{pmatrix} 48 & -4 \\ -4 & 48 \end{pmatrix}$$

$\det(\mathbf{H}^{(0)}) = 48^2 - 16 = 2304 - 16 = 2288$.
Newton-Gleichungssystem $\mathbf{H}^{(0)}\mathbf{d}^{(0)} = -\nabla f(\mathbf{x}^{(0)})$:

$$\begin{pmatrix} 48 & -4 \\ -4 & 48 \end{pmatrix} \mathbf{d}^{(0)} = \begin{pmatrix} -24 \\ -24 \end{pmatrix}$$

Durch Symmetrie (∇f hat gleiche Komponenten): $\mathbf{d}^{(0)} = (d, d)^\top$ mit $(48-4)d = -24$, also $d = -24/44 = -6/11$.

$$\mathbf{d}^{(0)} = \begin{pmatrix} -6/11 \\ -6/11 \end{pmatrix} \approx \begin{pmatrix} -0{,}5455 \\ -0{,}5455 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \begin{pmatrix} -6/11 \\ -6/11 \end{pmatrix} = \begin{pmatrix} 16/11 \\ 16/11 \end{pmatrix} \approx \begin{pmatrix} 1{,}4545 \\ 1{,}4545 \end{pmatrix}$$

Newton-Iteration 2 (bei $\mathbf{x}^{(1)} \approx (1{,}4545, 1{,}4545)^\top$):

$$\nabla f(\mathbf{x}^{(1)}) \approx \begin{pmatrix} 4(1{,}4545)^3 - 4(1{,}4545) \\ 4(1{,}4545)^3 - 4(1{,}4545) \end{pmatrix} = \begin{pmatrix} 4(3{,}0787) - 5{,}818 \\ 4(3{,}0787) - 5{,}818 \end{pmatrix} \approx \begin{pmatrix} 6{,}497 \\ 6{,}497 \end{pmatrix}$$

$$\mathbf{H}(\mathbf{x}^{(1)}) \approx \begin{pmatrix} 12(2{,}1156) & -4 \\ -4 & 12(2{,}1156) \end{pmatrix} = \begin{pmatrix} 25{,}387 & -4 \\ -4 & 25{,}387 \end{pmatrix}$$

Newton-Schritt: $d = -6{,}497/(25{,}387 - 4) = -6{,}497/21{,}387 \approx -0{,}3038$.

$$\mathbf{x}^{(2)} \approx \begin{pmatrix} 1{,}4545 - 0{,}3038 \\ 1{,}4545 - 0{,}3038 \end{pmatrix} = \begin{pmatrix} 1{,}1507 \\ 1{,}1507 \end{pmatrix}$$

(c) Gradientennorm:

k	$\mathbf{x}^{(k)}$	$f(\mathbf{x}^{(k)})$	$\|\nabla f(\mathbf{x}^{(k)})\|$
0	$(2, 2)$	$2 \cdot 16 - 4 \cdot 4 + 1 = 17$	$\sqrt{2} \cdot 24 \approx 33{,}94$
1	$(1{,}455, 1{,}455)$	$\approx 3{,}82$	$\approx 9{,}19$
2	$(1{,}151, 1{,}151)$	$\approx 1{,}48$	$\approx 3{,}14$

Stationärer Punkt bei $\mathbf{x}^* = (1, 1)^\top$: $4 - 4 = 0$ ✓, $f^* = 2 - 4 + 1 = -1$.
(d) Schnelle Konvergenz: Die Funktion ist glatt ($\mathcal{C}^\infty$) und der Startpunkt liegt relativ nahe am Minimum. Im Bereich des Minimums ist $\mathbf{H} \succ 0$, sodass die Newton-Methode quadratisch konvergiert.
Optimum: $\mathbf{x}^* = (1, 1)^\top$, $f^* = -1$

Aufgabe 2.2 – Vergleich Gradient vs. Newton: Konditionszahl

Aufgabe 2.2

Betrachte die Funktion $f(x_1, x_2) = x_1^2 + 100x_2^2$ (stark elongierte Ellipse).

(a) Berechne Gradient, Hesse-Matrix und Konditionszahl κ.

(b) Führe **drei** Gradientenabstiegsschritte mit optimaler Schrittweite $\alpha^* = 2/(\lambda_{\max} + \lambda_{\min})$ durch, Startpunkt $\mathbf{x}^{(0)} = (1, 1)^\top$.

(c) Führe **einen** Newton-Schritt durch und vergleiche.

(d) Erkläre, warum der Gradientenabstieg bei großer Konditionszahl langsam ist (Zick-Zack-Phänomen).

Lösung: Aufgabe 2.2

(a) Gradient, Hesse-Matrix, Konditionszahl:

$$\nabla f = \begin{pmatrix} 2x_1 \\ 200x_2 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 2 & 0 \\ 0 & 200 \end{pmatrix}, \quad \kappa = \frac{200}{2} = 100$$

(b) Gradientenabstieg, $\alpha^* = 2/(200 + 2) = 1/101 \approx 0{,}0099$:
$k = 0$: $\nabla f(1, 1) = (2, 200)^\top$, $\mathbf{d} = (-2, -200)^\top$.

$$\mathbf{x}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{101} \begin{pmatrix} -2 \\ -200 \end{pmatrix} = \begin{pmatrix} 1 - 2/101 \\ 1 - 200/101 \end{pmatrix} \approx \begin{pmatrix} 0{,}9802 \\ -0{,}9802 \end{pmatrix}$$

$f(\mathbf{x}^{(1)}) = (0{,}9802)^2 + 100(0{,}9802)^2 = 101 \cdot (0{,}9802)^2 \approx 96{,}93$. (Startfunktionswert: $1 + 100 = 101$.)
$k = 1$: $\nabla f(\mathbf{x}^{(1)}) \approx (1{,}960, -196{,}04)^\top$.

$$\mathbf{x}^{(2)} \approx \begin{pmatrix} 0{,}9608 \\ 0{,}9608 \end{pmatrix}, \quad f(\mathbf{x}^{(2)}) \approx 93{,}97$$

$k = 2$: Analog: $\mathbf{x}^{(3)} \approx (0{,}9417, -0{,}9417)^\top$, $f \approx 89{,}06$.
Reduktionsfaktor pro Iteration: $\left(\frac{\kappa-1}{\kappa+1}\right)^2 = \left(\frac{99}{101}\right)^2 \approx 0{,}9610$ – sehr langsame Konvergenz!

(c) Newton-Schritt (ein Schritt):

$$\mathbf{d}^{(0)} = -\mathbf{H}^{-1}\nabla f(\mathbf{x}^{(0)}) = -\begin{pmatrix} 1/2 & 0 \\ 0 & 1/200 \end{pmatrix}\begin{pmatrix} 2 \\ 200 \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}_N^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{x}^* \quad \checkmark$$

Newton erreicht das Minimum in **einem Schritt**, Gradient braucht > 10000 Iterationen!

(d) Zick-Zack-Phänomen:
Bei großem κ sind die Niveaukurven stark elongierte Ellipsen. Der Gradient steht senkrecht auf den Niveaukurven und zeigt in Richtung der kurzen Ellipsenachse (hier x_2-Richtung mit großem $\lambda = 200$). Jeder Gradientenabstiegsschritt überschießt in x_2-Richtung und muss in der nächsten Iteration korrigieren $\Rightarrow$ Zick-Zack. Newton korrigiert durch die Hesse-Matrix die unterschiedlichen Krümmungen.
Newton: 1 Iteration. Gradient: $\approx$ Tausende. Konditionszahl $\kappa = 100$ ist problematisch.

3 Stufe 3: Schwer

Aufgabe 3.1 – Sattelpunkt und Konvergenzanalyse

Aufgabe 3.1

Gegeben ist die Funktion $f(x_1, x_2) = x_1^2 - x_2^2 + 2x_1x_2$.

(a) Berechne alle stationären Punkte.

(b) Klassifiziere den stationären Punkt mithilfe der Hesse-Matrix und erkläre, was das Vorzeichen von $\det(\mathbf{H})$ aussagt.

(c) Welche Probleme entstehen, wenn die Newton-Methode auf dieses Problem angewendet wird? Berechne den Newton-Schritt von $\mathbf{x}^{(0)} = (1, 0)^\top$ und erkläre das Ergebnis.

(d) Wie verhält sich der Gradientenabstieg bei einem Sattelpunkt? Formuliere eine allgemeine Aussage.

Lösung: Aufgabe 3.1

(a) Stationäre Punkte:

$$\nabla f = \begin{pmatrix} 2x_1 + 2x_2 \\ -2x_2 + 2x_1 \end{pmatrix} = \mathbf{0} \Rightarrow \begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 0 \end{cases} \Rightarrow x_1 = x_2 = 0$$

Einziger stationärer Punkt: $\mathbf{x}^* = (0, 0)^\top$.

(b) Klassifikation:

$$\mathbf{H} = \begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix}, \quad \det(\mathbf{H}) = -4 - 4 = -8 < 0$$

Da $\det(\mathbf{H}) < 0$, hat $\mathbf{H}$ einen positiven und einen negativen Eigenwert $\Rightarrow$ $\mathbf{H}$ ist **indefinit** $\Rightarrow$ $(0, 0)$ ist ein **Sattelpunkt** (kein Minimum, kein Maximum).

Eigenwerte: $\lambda_{1,2} = \frac{(2-2) \pm \sqrt{(2+2)^2 + 4 \cdot 4}}{2} = \pm\sqrt{4+4} \cdot \frac{1}{1} = \pm 2\sqrt{2}$.

f ist nach oben *und* unten unbeschränkt.

(c) Newton-Schritt bei Sattelpunkt:

$$\nabla f(1, 0) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \mathbf{H}^{-1} = \frac{1}{-8} \begin{pmatrix} -2 & -2 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & -1/4 \end{pmatrix}$$

$$\mathbf{d} = -\mathbf{H}^{-1} \nabla f = -\begin{pmatrix} 1/4 & 1/4 \\ 1/4 & -1/4 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = -\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = (1, 0)^\top + (-1, 0)^\top = (0, 0)^\top$$

Newton konvergiert in einem Schritt zum Sattelpunkt! Das ist ein Problem: Newton findet **stationäre Punkte**, unterscheidet aber nicht zwischen Minima, Maxima und Sattelpunkten. Eine Prüfung von $\det(\mathbf{H})$ ist unerlässlich.

(d) Gradientenabstieg bei Sattelpunkten:

Der Gradientenabstieg konvergiert theoretisch **nicht** zu Sattelpunkten (da $-\nabla f$ in Abstiegsrichtung zeigt und der Sattelpunkt kein lokales Minimum ist). Allerdings kann er bei bestimmten Startpunkten langsam in Richtung des Sattelpunkts driften, bevor er abweicht. In der Praxis: Starts genau auf dem stabilen Unterraum → Konvergenz zum Sattelpunkt; sonst Divergenz davon.
$(0,0)$ **ist Sattelpunkt. Newton findet ihn in 1 Schritt – aber es ist kein Minimum!**

Aufgabe 3.2 – Rosenbrock: Numerische Analyse

Aufgabe 3.2

Die Rosenbrock-Funktion $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ ist ein klassisches Testproblem.

(a) Zeige, dass $\mathbf{x}^* = (1,1)^\top$ das einzige globale Minimum ist mit $f^* = 0$.

(b) Berechne $\mathbf{H}$ am Minimum $\mathbf{x}^* = (1,1)^\top$ und bestimme die Konditionszahl.

(c) Erkläre, warum der Gradientenabstieg an der Rosenbrock-Funktion extrem langsam ist.

(d) Berechne den ersten Newton-Schritt von $\mathbf{x}^{(0)} = (0,0)^\top$ vollständig.

(e) Vergleiche qualitativ: Wie viele Iterationen brauchen Gradient und Newton jeweils?

Lösung: Aufgabe 3.2

(a) Globales Minimum:
$f(\mathbf{x}) \geq 0$ für alle $\mathbf{x}$ (Summe von Quadraten). $f(\mathbf{x}) = 0$ genau dann, wenn:

$$x_2 - x_1^2 = 0 \quad \text{und} \quad 1 - x_1 = 0 \Rightarrow x_1 = 1,\ x_2 = 1$$

Also ist $(1,1)^\top$ das **einzige globale Minimum** mit $f^* = 0$.
(b) Hesse-Matrix am Minimum:
Allgemein (aus Theorieteil):

$$\mathbf{H} = \begin{pmatrix} -400(x_2 - x_1^2) + 800x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

An $\mathbf{x}^* = (1,1)^\top$: $x_2 - x_1^2 = 0$:

$$\mathbf{H}(\mathbf{x}^*) = \begin{pmatrix} 800 + 2 & -400 \\ -400 & 200 \end{pmatrix} = \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}$$

$\det(\mathbf{H}^*) = 802 \cdot 200 - 400^2 = 160400 - 160000 = 400 > 0$, Spur $= 1002 > 0 \Rightarrow$ $\mathbf{H}^* \succ 0$ (Minimum bestätigt).
Eigenwerte: Charakteristisches Polynom $(\lambda - 802)(\lambda - 200) - 160000 = 0$:

$$\lambda^2 - 1002\lambda + 400 = 0 \Rightarrow \lambda_{1,2} = \frac{1002 \pm \sqrt{1002^2 - 1600}}{2}$$

$$\lambda_1 \approx 1001{,}6, \quad \lambda_2 \approx 0{,}3996$$

$$\kappa(\mathbf{H}^*) = \frac{\lambda_1}{\lambda_2} \approx \frac{1001{,}6}{0{,}3996} \approx \mathbf{2506}$$

Extrem hohe Konditionszahl!

(c) Langsame Konvergenz des Gradientenabstiegs:
Die Rosenbrock-Funktion hat ein schmales, gekrümmtes Tal entlang der Parabel $x_2 = x_1^2$. Der Gradient zeigt quer zum Tal (da die Niveaukurven in Talrichtung sehr dicht liegen). Der Gradientenabstieg:

- Nähert sich schnell dem Talboden (Abstieg in x_2-Richtung).
- Schreitet dann extrem langsam das Tal entlang vor (sehr kleine Gradienten in Talrichtung).
- Konvergenzrate $\approx \left(\frac{2506-1}{2506+1}\right)^2 \approx 0{,}9984$ – nur 0,16% Reduktion pro Schritt!

(d) Newton-Schritt von $\mathbf{x}^{(0)} = (0,0)^\top$:

$$\nabla f(0,0) = \begin{pmatrix} -400 \cdot 0 \cdot (0-0) - 2(1-0) \\ 200(0-0) \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{H}(0,0) = \begin{pmatrix} -400(0) + 0 + 2 & 0 \\ 0 & 200 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 200 \end{pmatrix}$$

Newton-Schritt:

$$\mathbf{d} = -\mathbf{H}^{-1}\nabla f = -\begin{pmatrix} 1/2 & 0 \\ 0 & 1/200 \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = (0,0)^\top + 1 \cdot (1,0)^\top = (1,0)^\top$$

$f(1,0) = 100(0-1)^2 + 0 = 100$. Noch nicht am Minimum, aber guter Fortschritt (f sank von 1 auf 100 – warte!).
Tatsächlich: $f(0,0) = 100 \cdot 0 + 1 = 1$. Nach Newton-Schritt: $f(1,0) = 100$. Hier ist Newton schlechter geworden! Die Hesse-Matrix am Startpunkt ist nicht positiv definit genug für globale Konvergenz.
Mit Liniensuche: $\alpha < 1$ würde f reduzieren. Mit $\alpha \approx 0{,}01$: $\mathbf{x}^{(1)} \approx (0{,}01, 0)^\top$, $f \approx 0{,}98$. Besserer Ausgangspunkt für Newton.

(e) Iterationsvergleich (typisch):

Methode	Typische Iterationen	Erreichbare Genauigkeit
Gradientenabstieg	10 000–100 000	$\|\nabla f\| \approx 10^{-3}$
Newton (mit LS)	20–50	$\|\nabla f\| \approx 10^{-10}$

$\mathbf{x}^* = (1,1)^\top$, $\kappa \approx 2506$, Newton: ≈ 30 Iter., Gradient: $> 10\,000$ Iter.