Fachbereich Mathematik der Universität Hamburg

Dr. Hanna Peywand Kiani

Hörsaalübung 5 Analysis III für Studierende der Ingenieurwissenschaften

Infos/ Lehrmaterial unter

http://www.math.uni-hamburg.de/teaching/export/tuhh/index.html

Extrema unter Nebenbedingungen, Newton Verfahren

Erlaubte Hilfsmittel

für die Klausur Mathe III (Ana III + DGL I): 4 Blätter = 8 Seiten DIN-A4

Nur DGL I (Teile der LUM/BUW Studierenden): 2 Blätter = 4 Seiten DIN-A4

Extrema unter Nebenbedingungen:

Beispiel 1 Gegeben ist das folgende Optimierungsproblem:

$$\begin{aligned} \text{Gesucht sind die Minima/Maxima von} \quad & f(x,y) = x + y \\ \text{unter der Nebenbedingung} \quad & g(x,y) = 4 - x^2 - y^2 \geq 0\,. \end{aligned} \tag{1}$$

$$\operatorname{grad} f(x,y) = (f_x(x,y), f_y(x,y)) = (1,1)$$

$\Longrightarrow$ keine Extrema in ....

Zulässige Menge = Kreisscheibe ist kompakt, f ist stetig.

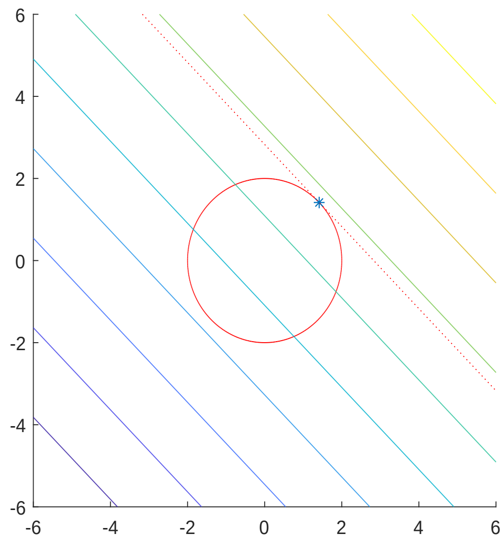
$\Longrightarrow$ es gibt ein globales Minimum und ein globales Maximum.

Durch $g(x,y) = 4 - x^2 - y^2 = 0$ ist eine Höhenlinie von g gegeben.

- **In jedem Punkt steht der Gradient senkrecht auf die Höhenlinie.**

- **Der Gradient zeigt in Richtung des steilsten Anstiegs.**

Analoges gilt für die Höhenlinien/Gradienten von f.



Geratene notwendige Bedingung für (lokale) Optimalität:

$$\operatorname{grad} f(x,y) + \lambda \operatorname{grad} g(x,y) = 0.$$

Damit hätten wir das Gleichungssystem

$$f_x + \lambda g_x = 0, \qquad f_y + \lambda g_y = 0, \qquad g = 0$$

oder mit $F(x,y) := f(x,y) + \lambda\, g(x,y)$

Für zulässige Punkte die Bedingung $\boxed{\operatorname{grad} F(x,y) = \mathbf{0}}$

Mit $f(x,y) = x + y\,, \qquad g(x,y) = 4 - x^2 - y^2 = 0\,.$

$$\begin{array}{lllll} 1 & + & \lambda \cdot (-2x) & = 0 \Longrightarrow \lambda \neq 0 \wedge x = \frac{1}{2\lambda}\,, \\ 1 & + & \lambda \cdot (-2y) & = 0 \Longrightarrow \lambda \neq 0 \wedge y = \frac{1}{2\lambda}\,, \\ 4 & - & x^2 - y^2 & = 0\,. \end{array}$$

Setze die Ergebnisse aus den ersten zwei Zeilen in die letzte Zeile ein:

$$4 - x^2 - x^2 = 0 \Longrightarrow 4 = 2x^2 \Longrightarrow x = y = \pm\sqrt{2}.$$

Wir erhalten also zwei Lösungen: $P_1 = \binom{-\sqrt{2}}{-\sqrt{2}}$, $P_2 = \binom{\sqrt{2}}{\sqrt{2}}$

Wenn die geratene Bedingung tatsächlich notwendig ist, sind die einzigen Kandidaten P_1 und P_2.

Zulässige Menge ist kompakt. Minimum und Maximum werden angenommen.

$f(P_1) = -2\sqrt{2}, \qquad f(P_2) = 2\sqrt{2}.$

Einziges lokale Minimum (und damit das globale Minimum) auf der zulässigen Menge liegt in P_1.

Einzige lokale Maximum (und damit das globale Maximum) auf der zulässigen Menge liegt in P_2.

Optimierung mit Gleichungsnebenbedingungen

hier $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ oder $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$

Problem: $\boxed{f(\boldsymbol{x}) \;=\; \text{min/max} \;\;!}$ unter der(den) Nebenbedigung(en)

$$\boxed{g(\boldsymbol{x}) \;=\; 0} \quad \text{bzw.} \quad \boxed{\begin{aligned} g(\boldsymbol{x}) &= 0 \\ h(\boldsymbol{x}) &= 0 \end{aligned}}$$

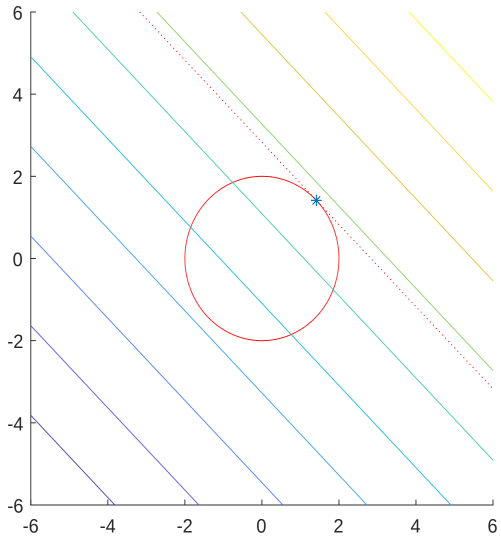
Regularitätsbedingung (RB)

Jacobi-Matrix der Nebenbedingungen hat maximalen Rang:

Bei einer Nebenbedingung $g(\boldsymbol{x}) = 0$ heißt das $\operatorname{grad} g(\boldsymbol{x}) \neq \mathbf{0}$,

bei zwei Nebenbedingungen $g(\boldsymbol{x}) = 0,\; h(\boldsymbol{x}) = 0:$ Rang $\begin{pmatrix} \text{grad}\, g(\boldsymbol{x}) \\ \text{grad}\, h(\boldsymbol{x}) \end{pmatrix} = 2$,

jeweils für die zulässigen Punkte.


6
4
2
0
-2
-4
-6
-6
-4
-2
0
2
4
6

Euler, Lagrange: Definiere mit λ, $\mu \in \mathbb{R}$ die Lagrange Funktion

$$F := f + \lambda g \qquad bzw. \qquad F := f + \lambda g + \mu h$$

Notwendige Bedingung für Min/Max von f unter g=(h=)=0

Wenn die **Regularitätsbdingung** erfüllt ist, ist jedes Extremum von f unter den Nebenbedingungen g=(h=)=0, ein zulässiger stationärer Punkt der **Lagrange-Funktion** $F(x, y, z)$. D.h.

$$\boxed{\text{grad } F(x, y, z) = 0, \quad g = h = 0}$$

Bestimme also stationäre Punkte von F d.h. Punkte mit $\mathrm{grad}\, F = \mathbf{0}$,

für die zusätzlich : $g = 0$ bzw. $g = h = 0$ gilt!

Ausführlicher für $\mathbb{R}^3$: Notwendig für lokale Extrema bei Regularität

(Im Fall $\mathbb{R}^2$: Zeile 3 und 5 streichen und überall sonst z streichen und $\mu = 0$ setzen.)

$F := f + \lambda g + \mu h$

$$\boxed{\begin{aligned} F_x &= f_x + \lambda\, g_x + \mu\, h_x = 0 \\ F_y &= f_y + \lambda\, g_y + \mu\, h_y = 0 \\ F_z &= f_z + \lambda\, g_z + \mu\, h_z = 0 \\ F_\lambda &= g(x,y,z) = 0 \\ F_\mu &= h(x,y,z) = 0 \end{aligned}}$$

$\longrightarrow$ **Kandidaten**

Klassifizierung: (Min, Max oder Sattel)

A) Zulässige Menge kompakt und f stetig : $\Longrightarrow$ Min/Max werden angenommen.
Kandidat mit höchstem Funktionswert = globales Maximum.
Kandidat mit kleinstem Funktionswert = globales Minimum.

B) Bedingungen zweiter Ordnung : (im Fall von 2 Nebenbedingungen)

Sei $\boldsymbol{x}_0$ zulässig (d.h. $g(\boldsymbol{x}_0) = h(\boldsymbol{x}_0) = 0$),
die Regularitätsbedingung erfüllt in $\boldsymbol{x}_0$, und es gelte

$$\exists \lambda, \mu \quad \text{mit} \qquad \mathrm{grad}\ F(\boldsymbol{x}_0) = 0$$

Definiere Tangentialraum:

$$TG(\boldsymbol{x}_0) = \{\boldsymbol{w}\ :< \ \boldsymbol{w}, \mathrm{grad}\, g(\boldsymbol{x}_0)\ > = 0 \ \text{ und } \ < \ \boldsymbol{w}, \mathrm{grad}\, h(\boldsymbol{x}_0)\ > = 0\,\}$$

Dann ist

$$\boxed{\textbf{notwendig für lokales Minimum: } \boldsymbol{w}^T \cdot HF_{\boldsymbol{x}}(\boldsymbol{x}_0) \cdot \boldsymbol{w} \geq 0, \; \forall \boldsymbol{w} \in TG(\boldsymbol{x}_0) \setminus \{\mathbf{0}\}}$$

und

$$\boxed{\textbf{hinreichend für lokales Minimum: } \boldsymbol{w}^T \cdot HF_{\boldsymbol{x}}(\boldsymbol{x}_0) \cdot \boldsymbol{w} > 0, \; \forall \boldsymbol{w} \in TG(\boldsymbol{x}_0) \setminus \{\mathbf{0}\}}$$

Analog

$$\boxed{\textbf{notwendig für lokales Maximum: } \boldsymbol{w}^T \cdot HF_{\boldsymbol{x}}(\boldsymbol{x}_0) \cdot \boldsymbol{w} \leq 0, \; \forall \boldsymbol{w} \in TG(\boldsymbol{x}_0) \setminus \{\mathbf{0}\}}$$

und

$$\boxed{\textbf{hinreichend für lokales Maximum: } \boldsymbol{w}^T \cdot HF_{\boldsymbol{x}}(\boldsymbol{x}_0) \cdot \boldsymbol{w} < 0, \; \forall \boldsymbol{w} \in TG(\boldsymbol{x}_0) \setminus \{\mathbf{0}\}}$$

Das heißt insbesondere : die notwendigen Bedingungen aus dem unrestingierten Fall für Minima (Maxima), nämlich Hesse-Matrix positiv (negativ) semidefinit sind hier keine notwendigen Bedingungen mehr. Die Matrix kann z.B. auch bei einem Minimum negative Eigenwerte haben, sofern die zugehörigen Eigenvektoren keine zulässigen Richtungen sind, d.h. aus der zulässigen Menge raus führen.

Beispiel 2: [Alte Klausuraufgabe]

Bestimmen Sie die globalen Extrema der Funktion

$$f(x,y,z) = x - 8y + z$$

auf dem Schnitt der beiden Kugeloberflächen

$$g(x,y,z) \quad = \quad x^2 + (y+4)^2 + z^2 - 25 \quad = \quad 0$$

und

$$h(x,y,z) \quad = \quad x^2 + y^2 + z^2 - 9 \quad = \quad 0.$$

Lösungsskizze:

Regularitätsbedingung (RB):

$$J(g,h)(x,y,z) = \begin{pmatrix} g_x & g_y & g_z \\ h_x & h_y & h_z \end{pmatrix} = \begin{pmatrix} 2x & 2(y+4) & 2z \\ 2x & 2y & 2z \end{pmatrix}$$

RB verletzt falls

$$\alpha \begin{pmatrix} 2x \\ 2(y+4) \\ 2z \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \implies \begin{cases} \alpha = 1 \lor x = 0 \\ \text{nicht erfüllbar für } \alpha = 1 \\ \alpha = 1 \lor z = 0 \end{cases}$$

RB kann also nur für $x = z = 0$ verletzt sein.

$g(0, y, 0) = 0 + (y+4)^2 + 0 - 25 = 0 \implies y = -4 \pm 5,$

$h(0, y, 0) = 0 + y^2 + 0 - 9 = 0 \implies y = \pm 3.$

Die Regularitätsbedingung ist in allen zulässigen Punkten erfüllt.

Mit $f(x, y, z) = x - 8y + z$ und der Lagrange Funktion $F = f + \lambda g + \mu h$ erhält man als notwendige Bedingungen für Extrema:

$1)\; F_x = 0 :$

$2)\; F_y = 0 :$

$3)\; F_z = 0 :$

$4)\; g = 0 : \qquad x^2 + (y+4)^2 + z^2 - 25 = 0\,,$

$5)\; h = 0 : \qquad x^2 + y^2 + z^2 - 9 = 0\,.$

Aus den letzten beiden Gleichungen folgt

$$(y+4)^2 - y^2 = 16 \iff 8y + 16 = 16 \iff \text{y= 0}\,.$$

Dies eingesetzt in die zweite Gleichung liefert $\lambda = 1$ und damit

I) $1 + 2x + 2\mu x = 0\,,$
II) $1 + 2z + 2\mu z = 0\,,$
III) $x^2 + z^2 - 9 = 0\,,$
$\lambda = 1\,, \qquad y = 0\,.$

I - II :

$(1+\mu)(x-z)=0 \iff \mu=-1$ oder $x=z\,.$

$\mu=-1$: I)

$x=z$: III)

Kandidaten und Funktionswerte für $f(x,y,z) = x - 8y + z$

$$P_1 = \begin{pmatrix} \frac{3}{\sqrt{2}} \\ 0 \\ \frac{3}{\sqrt{2}} \end{pmatrix}, \quad f(P_1) = 3\sqrt{2}\,, \qquad P_2 = \begin{pmatrix} -\frac{3}{\sqrt{2}} \\ 0 \\ -\frac{3}{\sqrt{2}} \end{pmatrix}, \quad f(P_2) = -3\sqrt{2}\,.$$

Da der Schnitt der beiden Kugeloberflächen (leer, Punkt oder Kreisrand) eine kompakte Menge ist werden Minimum und Maximum der stetigen Funktion f angenommen. Vergleich der Funktionswerte zeigt, dass in P_1 das globale Maximum und in P_2 das globale Minimum liegt.

Beispiel 3) Gegeben sei das Extremalproblem

$$f(x, y) \;=\; x^2 \;+\; y^2 \;=\; \text{min!}$$

unter der Nebenbedingung

$$g(x, y) \;=\; e^{x-1} - \arctan(y+1) - 1 \;= 0.$$

a) Überprüfen Sie die Regularitätsbedingung.

Zeigen Sie, dass $\boldsymbol{x}_0 \;=\; (1, -1)^T$ zusammen mit einem geeigneten Multiplikator λ ein zulässiger, stationärer Punkt der zugehörigen Lagrange–Funktion F ist.

b) Untersuchen Sie den stationären Punkt $(1, -1)^T$ auf seinen Typ hin. Stellen Sie dazu die Hesse–Matrix $H_{\boldsymbol{x}}F(\boldsymbol{x}_0)$ auf und überprüfen Sie deren Definitheit auf dem Tangentialraum ker $\mathsf{D}g(\boldsymbol{x}_0) = TG_g(\boldsymbol{x}_0)$.

Lösungsskizze:

Teil a): $g(x,y) = e^{x-1} - \arctan(y+1) - 1$

$\Longrightarrow \boldsymbol{J}g(x,y) = \operatorname{grad} g(x,y) = (e^{x-1}, -\frac{1}{1+(1+y)^2}) \neq (0,0) \qquad \forall (x,y) \in \mathbb{R}^2$

$\Longrightarrow \operatorname{grad} g(x,y)$ hat Rang 1 (Regularitätsbedingung).

Zulässigkeit: $g(1,-1) = e^{1-1} - \arctan(-1+1) - 1$

Lagrange-Funktion: $F = f + \lambda \cdot g = x^2 + y^2 + \lambda \left(e^{x-1} - \arctan(y+1) - 1\right)$.

stationärer Punkt der Lagrange–Funktion bedeutet: $\operatorname{grad} F(1,-1) = (0,0)$.

Zu zeigen: $\exists \lambda$ mit $\operatorname{grad} F(1,-1) = (0,0)$

$\operatorname{grad} F(x,y) =$ $\qquad$ $\operatorname{grad} F(1,-1) =$

Teil b): $\operatorname{grad}_{x,y} F(x,y) = (F_x,\ F_y) = (2x + \lambda e^{x-1}, 2y - \lambda \dfrac{1}{1+(1+y)^2})$

$$\boldsymbol{H}_{x,y}F = \begin{pmatrix} F_{xx} & F_{xy} \\ F_{yx} & F_{yy} \end{pmatrix} =$$

$\Rightarrow \boldsymbol{H}_{x,y}F(1,-1) =$

d.h. $\boldsymbol{H}_{x,y}F(1,-1)$ ist semidefinit (det $\boldsymbol{H}_{x,y}F(1,-1) = 0$).

Untersuchung auf dem Tangentialraum:

$$\text{Ker}Dg(\boldsymbol{x}_0) = TG(\boldsymbol{x}_0) = \{\boldsymbol{w} \in \mathbb{R}^n : \operatorname{grad} g(\boldsymbol{x}_0) \cdot \boldsymbol{w} = 0\}$$

$\operatorname{grad} g(x,y) = (e^{x-1}, -\frac{1}{1+(1+y)^2})$

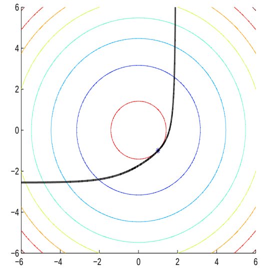
$\boldsymbol{w} = \binom{w_1}{w_2}$ mit $\operatorname{grad} g(1,-1)\binom{w_1}{w_2} = 0 \Rightarrow w_1 = w_2$

$$\text{Ker} Dg(\boldsymbol{x}_0) = TG(\boldsymbol{x}_0) = \left\{ \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2 \;:\; \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix},\; \alpha \in \mathbb{R} \setminus \{0\} \right\}$$

Für alle $\boldsymbol{w} \neq 0$ aus dem Tangentialraum gilt mit einem $\alpha \neq 0$:

$$\begin{aligned} \boldsymbol{w}^T \boldsymbol{H}_{x,y} F(1,-1)\, \boldsymbol{w} &= \alpha \cdot (1,1) \boldsymbol{H}_{x,y} F(1,-1) \alpha \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \alpha^2 \cdot (1,1) \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \alpha^2 \cdot (1,1) \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2\alpha^2 > 0\,. \end{aligned}$$

d.h. die Hesse-Matrix $\boldsymbol{H}_{x,y} F(1,-1)$ ist positiv definit auf dem Tangentialraum. Daher liegt im Punkt $(1,-1)$ ein strenges lokales Minimum vor.



Idee des Newtonverfahrens:

Ziel: $f(x) \stackrel{!}{=} 0$

Idee: $f(x) \approx T_1(x;\, x_k) \;=\; f(x_k) + f'(x_k)(x - x_k)$ (Nahe x_k)

Löse Ersatzproblem: $f(x_k) + f'(x_k)(x - x_k) \;=\; 0 \Longleftrightarrow f'(x_k)\Delta_k \stackrel{!}{=} -f(x_k)$

Definiere $x_{k+1} = x_k + \Delta_k$.

Iteriere!

Geometrisch:

Newton-verfahren:

Gesucht Lösung des System

$$\boldsymbol{f}(\boldsymbol{x}) = 0, \qquad \boldsymbol{x} \in D \subset \mathbb{R}^n,\ f : D \to \mathbb{R}^n$$

Berechne stattdessen Nullstelle der Linearisierung im Punkt $\boldsymbol{x}^{[k]}$

$$T_1(\boldsymbol{x}; \boldsymbol{x}^{[k]}) = \boldsymbol{f}(\boldsymbol{x}^{[k]}) + \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}^{[k]})(\boldsymbol{x} - \boldsymbol{x}^{[k]}) = 0 \Longrightarrow \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}^{[k]})(\boldsymbol{x} - \boldsymbol{x}^{[k]}) = \boldsymbol{f}(\boldsymbol{x}^{[k]})$$

Nehme dieses $\boldsymbol{x}$ als neue Näherung und iteriere. Startwert $x^{[0]}$ vorgeben. Dann:

- Berechne $\boldsymbol{f}(\boldsymbol{x}^{[k]})$,
- Berechne die Jacobi-Matrix $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}^{[k]})$,
- Löse das Gleichungssystem $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}^{[k]}) \cdot \Delta^{[k]} = -\boldsymbol{f}(\boldsymbol{x}^{[k]})$
- Setze $\boldsymbol{x}^{[k+1]} = \boldsymbol{x}^{[k]} + \Delta^{[k]}$.

Beispiel:

Zur Bestimmung des Minimums der Funktion

$$z = F(x,y) := x^2 + 2y^2 - 0.1\cos(x+y) - 3x + 2y$$

soll das Newton–Verfahren auf die Funktion $f(x,y) := \nabla F(x,y)^T$ angewendet werden.

a) Berechnen Sie $f(x,y)$ und die Jacobi-Matrix $Jf(x,y)$.

b) Stellen Sie das Newton–Verfahren auf und führen Sie den ersten Iterationsschritt per Handrechnung durch; Startvektor: $(x^0, y^0) = (0,0)$.

c) Führen Sie die Iteration numerisch durch und berechnen Sie damit die Lösung auf (wenigstens) zehnstellige Genauigkeit.

Lösung:

$$\boldsymbol{f}(x,y) = (2x + 0.1\sin(x+y) - 3,\; 4y + 0.1\sin(x+y) + 2)^T$$

$$f(0,0) = (-3,\; 2)^T$$

$$\boldsymbol{J}\boldsymbol{f}(x,y) = \begin{pmatrix} 2 + 0.1 * \cos(x+y) & 0.1 * \cos(x+y) \\ 0.1 * \cos(x+y) & 4 + 0.1 * \cos(x+y) \end{pmatrix}$$

$$\boldsymbol{J}\boldsymbol{f}(0,0) = \begin{pmatrix} 2.1 & 0.1 \\ 0.1 & 4.1 \end{pmatrix}$$

$$\boldsymbol{J}\boldsymbol{f}(0,0) \cdot \Delta^{[0]} = -\,\boldsymbol{f}(0,0) \iff \begin{pmatrix} 2.1 & 0.1 \\ 0.1 & 4.1 \end{pmatrix} \cdot \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix} = -\begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

$$\boldsymbol{x}^{[1]} = \boldsymbol{x}^{[0]} + \Delta^{[0]}.$$

Matlab für Anfänger (das geht natürlich deutlich eleganter!)

```
format long
k=0
x= 0 %Startwert
y= 0  %Startwert
for j=1:5, %Anzahl Newton-Schritte
c= 0.1 * cos(x+y);
s = 0.1 * sin(x+y);
A=[2+c c; c 4+c]  %Jakobi-Matrix
b=[-2*x-s+3;-4*y-s-2] %Funktion dessen Nullstelle gesucht ist
%Newton-Iteration
dx = A \b ; %Gleichungssystem lösen
k=k+1
x= x+dx(1,1)
y=y+dx(2,1)
end
%Gradient von ursprünglicher Funktion
g= b
```

$$
\begin{array}{llll}
k=0 & x & =0 & y=0 \\
k=1 & x & =1.45348837209302 & y=-0.52325581395349 \\
k=2 & x & =1.45963647269063 & y=-0.52018176365468 \\
k=3 & x & =1.45963810885761 & y=-0.52018094557119 \\
k=4 & x & =1.45963810885773 & y=-0.52018094557114 \\
k=5 & x & =1.45963810885773 & y=-0.52018094557114
\end{array}
$$

Gradient von F im Punkt $x^{[5]}$:

$g = 10^{-15} * (-0.444089209850063, -0.222044604925031)$

Da die Hesse–Matrix von F (also die Jacobi-Matrix von f) positiv definit ist, handelt es sich tatsächlich um eine Näherung für ein Minimum.