Fakultät für Luft- und Raumfahrttechnik
Institut für Angewandte Mathematik
und Wissenschaftliches Rechnen
Lineare und nichtlineare Optimierung
Priv.-Doz. Dr. habil. S.-J. Kimmerle

der Bundeswehr
Universität München

3. Übung

Herbsttrimester 2023

8) Ein numerisches Verfahren erzeuge eine Iterationsfolge $\{x^{[i]}\}_{i\in\mathbb{N}}$, die gegen ein x^* konvergiert. Für die k-te Iterierte gelte $\|x^{[k]} - x^*\| = \varepsilon_0 < 1$. Wie viele Iterationen benötigt das Verfahren um die Genauigkeit $\varepsilon \in (0, \varepsilon_0)$ zu erreichen, wenn die Folge linear mit Konstante $C_1 \in (0, 1)$, bzw. quadratisch mit Konstante $C_2 \in \left(0, \frac{1}{\varepsilon_0}\right)$ konvergiert? Was ergibt sich konkret für $C_1 = C_2 = \frac{1}{2}, \varepsilon_0 = 10^{-3}$ und $\varepsilon = 10^{-10}$?

9) Implementieren Sie das globalisierte Newtonverfahren aus Algorithmus 2.5.9 mit den Parametern $\rho = 10^{-8}, p = 2.1, \beta = 0.5, \sigma = 10^{-4}$ und $\varepsilon = 10^{-12}$. Testen Sie es mit der Rosenbrock-Funktion und dem Startvektor $x^0 = (-1.2, 1)$.
Hinweis: Sie können die Funktionen aus Aufgabe 7 wiederverwenden.

10) Lösen Sie das Problem

$$\min x_1^2 + x_2^2 + x_3^2 \quad \text{u.d.N.} \quad x_1 + x_2 + x_3 = 1.$$

11) Zeigen Sie mit dem Beispiel

$$\max x_2 \quad \text{u.d.N.} \quad x_1 \geq 0, x_1 + x_2 \leq 1, x_1^2 + x_2^2 \leq 1,$$

dass die Lagrange-Multiplikatoren im Allgemeinen nicht eindeutig sind.

12) Zeigen Sie mit dem Beispiel

$$\min (x_1 - x_2)^2 \quad \text{u.d.N.} \quad x_1 \leq 1, x_2 \leq 1,$$

dass es auch KKT-Punkte geben kann, bei denen keine strikte Komplementarität vorliegt, d.h. in den Komplementaritätsbedingungen sind beide Faktoren Null.

13) Betrachten Sie für $-\sqrt{2} \leq \gamma$ das Optimierungsproblem

$$\min x_1 \quad \text{u.d.N.} \quad x_1^2 + x_2^2 \leq 1, x_1 + x_2 \leq \gamma.$$

Bestimmen Sie die KKT-Punkte in Abhängigkeit von γ.

Lösungen

8) Konvergiert die Folge linear mit Konstante $C_1 \in (0,1)$, so gilt für alle $n \in \mathbb{N}$

$$\|x^{[k+n]} - x^*\| \leq C_1 \|x^{[k+n-1]} - x^*\| \leq C_1^n \|x^{[k]} - x^*\| = C_1^n \varepsilon_0.$$

Damit folgt, dass $\|x^{[k+n]} - x^*\| \leq \varepsilon$ erfüllt ist, wenn $C_1^n \varepsilon_0 \leq \varepsilon$ gilt, also wenn

$$n \geq \frac{\ln(\varepsilon/\varepsilon_0)}{\ln(C_1)}.$$

Bei der quadratischen Konvergenz gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} \|x^{[k+n]} - x^*\| &\leq C_2 \|x^{[k+n-1]} - x^*\|^2 \leq C_2 * C_2^2 \|x^{[k+n-2]} - x^*\|^4 \\ &\leq \ldots \leq C_2^{1+2+\ldots+2^{n-1}} \|x^k - x^*\|^{2^n} = C_2^{2^n-1} \varepsilon_0^{2^n}. \end{aligned}$$

Damit folgt, dass $\|x^{[k+n]} - x^*\| \leq \varepsilon$ erfüllt ist, wenn $C_2^{2^n-1} \varepsilon_0^{2^n} \leq \varepsilon$ gilt, also wenn

$$(C_2\varepsilon_0)^{2^n} \leq C_2\varepsilon \quad \Leftrightarrow \quad 2^n \geq \frac{\ln(C_2\varepsilon)}{\ln(C_2\varepsilon_0)} \quad \Leftrightarrow \quad n \geq \frac{\ln\left(\frac{\ln(C_2\varepsilon)}{\ln(C_2\varepsilon_0)}\right)}{\ln(2)}.$$

Speziell für $C_1 = C_2 = \frac{1}{2}, \varepsilon_0 = 10^{-3}$ und $\varepsilon = 10^{-10}$ ergibt sich im linearen Fall $n = 24$ und im quadratischen Fall $n = 2$ Iterationen.

9)
```
function x=Newtonverfahren(x,f,Df,D2f)

sigma=1e-4; %Parameter Armijo-Liniensuche
beta=0.5;
eps=1e-12; %Fehlertoleranz Abbruchkriterium
p=2.1; %Parameter Loesung Newton-LGS sinnvoll
rho=1e-8;
kmax=1e5; %max. Schritte in Newtonrichtung
alphamin=1e-4; %minimale Schrittweite (ergibt Maximalzahl an Armijo-
    Iterationen)
k=0;
fcount=0; %zaehlt Funktionsauswertungen
Dfcount=0; %zaehlt Gradientenauswertungen
D2fcount=0; %zaehlt Hesse-Matrix-Auswertungen

%Ausgabe #Schritte, Schrittweite, Abbruchkriterium, Zwischenwert, #
    Funktions-/Gradienten-/Hesse-Matrix-Auswertungen
%fprintf('%5.0f %6.4f %12.7f %12.7f %12.7f %6.0f %6.0f %6.0f\n',k,alpha,
    term,x(1),x(2),fcount,Dfcount,D2fcount);
fprintf('    k   alpha          term          x(1)          x(2)  fcount  Dfcount
    D2fcount \n');
```

```
fx=feval(f,x);
fcount=fcount+1;
nablafx=feval(Df,x);
Dfcount=Dfcount+1;
term=norm(nablafx);

while term > eps && k<=kmax %Newtonverfahren (glob.)
    Hf=feval(D2f,x);
    D2fcount=D2fcount+1;
    if condest(Hf) > 1e+15
        d=-nablafx;
        condest(Hf)
        pause
    else
        d=-Hf\nablafx;
        if nablafx'*d > -rho*norm(d)^p
            d=-nablafx;
            d
            pause
        end
    end
    alpha=1;
    fxalphad=feval(f,x+d);
    fcount=fcount+1;
    nfd=sigma*nablafx'*d;
    while fxalphad > fx +alpha*nfd && alpha > alphamin %Armijo-Liniensuche
        alpha=alpha*beta;
        fxalphad=feval(f,x+alpha*d);
        fcount=fcount+1;
    end
    x=x+alpha*d; %Update
    fx=fxalphad; %letzter Funktionswerte wird angenommen
    nablafx=feval(Df,x);
    Dfcount=Dfcount+1;
    term=norm(nablafx); %fuer Abbruchkriterium
    k=k+1;
    %Ausgabe #Schritte, Schrittweite, Abbruchkriterium, Zwischenwert, #
        Funktions-/Gradienten-/Hesse-Matrix-Auswertungen
    %if mod(k, 25) == 0 %falls k durch 100 teilbar Pause zur Ansicht der
        Ausgabe beim Testen
    %    pause
    %    fprintf('    k   alpha          term           x(1)           x(2)
        fcount Dfcount D2fcount \n');
    %end
    fprintf('%5.0f %6.4f %12.7f %12.7f %12.7f %6.0f %6.0f %6.0f\n',k,alpha,
        term,x(1),x(2),fcount,Dfcount,D2fcount);
end
```

```
fprintf('\n')
end

function fx=D2Rosenbrock(x)
fx=[ 800*x(1)^2-400*(x(2)-x(1)^2)+2,  -400 *x(1);     -400*x(1), 200];
end
```

Mit diesen Funktionen und den bekannten Funktionen von Blatt 2 erhalten wir:

```
>> x0 = [-1.2; 1];
>> Newtonverfahren(x0, @Rosenbrock, @DRosenbrock, @D2Rosenbrock)
    k  alpha        term          x(1)          x(2)  fcount  Dfcount  D2fcount
    1 1.0000   4.6394262    -1.1752809     1.3806742       2        2         1
    2 0.1250  28.5500805    -0.9329814     0.8112107       6        3         2
    3 1.0000  11.5715209    -0.7825401     0.5897364       7        4         3
    4 1.0000  30.3258946    -0.4599971     0.1075634       8        5         4
    5 1.0000   3.6041023    -0.3930456     0.1500024       9        6         5
    6 0.2500   9.2484184    -0.2094119     0.0067701      12        7         6
    7 1.0000   4.9198007    -0.0657190    -0.0163287      13        8         7
    8 1.0000   8.6643402     0.1420425    -0.0229888      14        9         8
    9 1.0000   1.7788143     0.2311072     0.0454780      15       10         9
   10 0.5000   5.8778039     0.3797428     0.1181458      17       11        10
   11 1.0000   2.1763707     0.4795949     0.2200408      18       12        11
   12 1.0000   9.4012897     0.6534058     0.3967289      19       13        12
   13 1.0000   0.4920626     0.7026236     0.4912576      20       14        13
   14 0.5000   3.9242492     0.8027855     0.6332210      22       15        14
   15 1.0000   1.2421077     0.8634908     0.7419312      23       16        15
   16 1.0000   2.5330674     0.9420787     0.8813362      24       17        16
   17 1.0000   0.2375818     0.9679918     0.9363367      25       18        17
   18 1.0000   0.3482721     0.9962103     0.9916387      26       19        18
   19 1.0000   0.0038742     0.9994794     0.9989483      27       20        19
   20 1.0000   0.0001187     0.9999989     0.9999975      28       21        20
   21 1.0000   0.0000000     1.0000000     1.0000000      29       22        21
   22 1.0000   0.0000000     1.0000000     1.0000000      30       23        22


ans =
      1
      1
```

10) Die Lagrangefunktion lautet $L(x,\mu) = f(x) + \mu h(x) = x_1^2 + x_2^2 + x_3^2 + \mu(x_1 + x_2 + x_3 - 1)$. Da $\nabla h(x) = (1,1,1)^\top$ nie der Nullvektor ist, liegt lineare Unabhängigkeit vor d.h. (LICQ) ist erfüllt, und es folgt aus den KKT-Bedingungen

$$(0,0,0)^\top \overset{!}{=} \nabla_x L(x,\mu) = (2x_1 + \mu, 2x_2 + \mu, 2x_3 + \mu)^\top$$

Somit ist $x_1 = x_2 = x_3 = -\mu/2$ und wegen $x_1 + x_2 + x_3 = 1$ folgt $\mu = -\frac{2}{3}$. Also ist die

Lösung $x = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)^\top$. Mithilfe der pos. Definitheit von $\nabla_{xx}L$ oder durch geometrische Überlegungen sieht man, dass tatsächlich ein Minimum vorliegt.

11) Aus den KKT-Bedingungen ergibt sich für die Lagrangefunktion

$$L(x,\lambda) = -x_2 - \lambda_1 x_1 + \lambda_2(x_1 + x_2 - 1) + \lambda_3(x_1^2 + x_2^2 - 1)$$

die Bedingung

$$(0,0)^\top = \nabla_x L(x,\lambda) = (-\lambda_1 + \lambda_2 + 2\lambda_3 x_1, -1 + \lambda_2 + 2\lambda_3 x_2)^\top .$$

1. Fall $\lambda_1 = 0$: Wegen $x_1, \lambda_2, \lambda_3 \geq 0$ liefert die erste Gleichung $\lambda_2 = 0$ und $\lambda_3 x_1 = 0$. Die zweite liefert dann $2\lambda_3 x_2 = 1$ und somit $\lambda_3 > 0$, was $x_1 = 0$ impliziert. Ferner ist nun die Ungleichung $x_1^2 + x_2^2 = 1$ aktiv, was unter Beachtung von $0 < \lambda_3 = \frac{1}{2x_2}$ zu dem eindeutigen Punkt $x = (0,1)$ führt.

2. Fall $\lambda_1 > 0$: Damit ist die erste Ungleichung aktiv, also $x_1 = 0$. Die erste Stationaritätsgleichung liefert dann $\lambda_2 = \lambda_1 > 0$, und daher ist auch die Ungleichung $x_1 + x_2 = 1$ aktiv, d.h. $x_2 = 1$.

In beiden Fällen ergibt sich also $x = (0,1)$. Für die Lagrange-Multiplikatoren ergibt sich dann aus den Stationaritätsgleichungen $\lambda_2 = \lambda_1$ und $\lambda_3 = \frac{1}{2}(1 - \lambda_2)$. Da alle nicht negativ sein müssen, gilt

$$\lambda \in \left\{ \left(\alpha, \alpha, \frac{1-\alpha}{2} \right) \middle| \alpha \in [0,1] \right\},$$

d.h. die Lagrange-Multiplikatoren sind nicht eindeutig bestimmt.

Man bemerkt, dass (LICQ) im Minimum nicht erfüllt ist (3 Nebenbedingungen bei 2 Raumdimensionen).

12) Aus den KKT-Bedingungen folgt für die Lagrangefunktion

$$L(x,\lambda) = (x_1 - x_2)^2 + \lambda_1(x_1 - 1) + \lambda_2(x_2 - 1)$$

die Bedingung

$$(0,0)^\top = \nabla_x L(x,\lambda) = (2(x_1 - x_2) + \lambda_1, 2(x_2 - x_1) + \lambda_2)^\top .$$

Da sowohl $\lambda_1 \geq 0$, als auch $\lambda_2 \geq 0$ gelten muss, folgt hieraus $x_1 = x_2$. Damit ist aber auch $\lambda_1 = 0 = \lambda_2$. Nun erhält man die Lösungsmenge

$$\{(\alpha,\alpha)^\top \mid \alpha \leq 1\}$$

Insbesondere für das Minimum $(1,1)^\top$ sind dann beide Restriktionen aktiv und gleichzeitig sind die Lagrange-Multiplikatoren Null.

13) Die Lagrangefunktion ist $L(x,\lambda) = x_1 + \lambda_1(x_1^2 + x_2^2 - 1) + \lambda_2(x_1 + x_2 - \gamma)$. Aus den KKT-Bedingungen folgt

$$\begin{aligned} 1 + \lambda_1 2x_1 + \lambda_2 &= 0 \\ \lambda_1 2x_2 + \lambda_2 &= 0 \end{aligned}$$

Ist $\lambda_1 = 0$, so folgt aus der zweiten Gleichung $\lambda_2 = 0$ und damit ist die erste nicht lösbar. Also gilt $\lambda_1 \neq 0$ und wegen der Komplementarität $x_1^2 + x_2^2 = 1$. Durch Differenzbildung der zwei Gleichungen und Auflösung nach λ_1 erhält man $\lambda_1 = \frac{1}{2(x_2 - x_1)}$ und damit $\lambda_2 = \frac{-x_2}{x_2 - x_1}$. Wegen $\lambda_1 \geq 0$ und $\lambda_2 \geq 0$ folgt hieraus $x_2 > x_1$ und $x_2 \leq 0$.

Nun unterscheiden wir zwei Fälle:

$\lambda_2 = 0$: Dann liefert die zweite Gleichung $x_2 = 0$, und damit, da die erste Restriktion aktiv ist, $x_1 = \pm 1$. Da $\lambda_1 > 0$ gelten muss, liefert die erste Gleichung $x_1 < 0$. Insgesamt erhält man den Punkt

$$(x_1, x_2, \lambda_1, \lambda_2) = \left(-1, 0, \frac{1}{2}, 0\right).$$

Dieser ist KKT-Punkt, wenn er zulässig ist, d.h. wenn $-1 \leq \gamma$ gilt. Andernfalls hat man in diesem Fall keinen KKT Punkt.

$\lambda_2 \neq 0$: Dann müssen beide Restriktionen aktiv sein, und somit ist $x_2 = \gamma - x_1$. Dies in die erste Restriktion eingesetzt ergibt:

$$\begin{aligned} & x_1^2 + (\gamma - x_1)^2 = 1 \\ \Leftrightarrow\ & 2x_1^2 - 2\gamma x_1 + \gamma^2 - 1 = 0 \\ \Leftrightarrow\ & x_1 = \frac{2\gamma \pm \sqrt{4\gamma^2 - 8(\gamma^2 - 1)}}{4} \\ \Leftrightarrow\ & x_1 = \frac{\gamma}{2} \pm \frac{1}{2}\sqrt{2 - \gamma^2} \end{aligned}$$

Man sieht also, dass dies nur für $-\sqrt{2} \leq \gamma \leq \sqrt{2}$ möglich ist. Weiter erhält man $x_2 = \gamma - x_1 = \frac{\gamma}{2} \mp \frac{1}{2}\sqrt{2 - \gamma^2}$ und wegen $x_1 < x_2$ erhält man den einen Punkt $(x_1, x_2) = \left(\frac{\gamma}{2} - \frac{1}{2}\sqrt{2 - \gamma^2}, \frac{\gamma}{2} + \frac{1}{2}\sqrt{2 - \gamma^2}\right)$. Da ferner $x_2 \leq 0$ gelten muss, folgt $\gamma \leq -1$. Damit erhält man für $-\sqrt{2} < \gamma \leq -1$ den KKT-Punkt

$$\left(\frac{\gamma}{2} - \frac{1}{2}\sqrt{2 - \gamma^2}, \frac{\gamma}{2} + \frac{1}{2}\sqrt{2 - \gamma^2}, \frac{1}{2\sqrt{2 - \gamma^2}}, \frac{-\frac{\gamma}{2} - \frac{1}{2}\sqrt{2 - \gamma^2}}{\sqrt{2 - \gamma^2}}\right).$$

Für $\gamma = -\sqrt{2}$ sind die Multiplikatoren in der Formel nicht definiert. Man erkennt, dass dann $x_1 = x_2$ ist und somit hat das Gleichungssystem keine Lösung.

Zusammenfassend erhalten wir also für $\gamma = -\sqrt{2}$ keinen KKT-Punkt, für $-\sqrt{2} < \gamma$ jeweils genau einen.