Nadine Kostorz (1972005), Leon Talenberg (1957966), Martin Belica (1775706)

Aufgabe 1

a) Berechnen Sie die Hessematrix $D^2 f$ der Funktion

$$f(x) = x_1^2 x_2 + 2x_1 x_2^2 + x_1 x_2.$$

Beweis: Da f als Polynom in C^2 liegt, folgt:

$$\nabla f(x) = \begin{pmatrix} 2x_1x_2 + 2x_2^2 + x_2 \\ x_1^2 + 4x_1x_2 + x_1 \end{pmatrix} \text{ und damit } D^2 f(x) = \begin{pmatrix} 2x_2 & 2x_1 + 4x_2 + 1 \\ 2x_1 + 4x_2 + 1 & 4x_1 \end{pmatrix}$$

$\square$

b) Implementieren Sie eine Funktion $DsqF(X)$, die zu gegebenem Interval $X \in \mathbb{IR}^2$ die (eintragsweise) natürliche Intervallerweiterung D^2F von D^2f zurück gibt.

$DsqF$ soll ein $(2, 2, 2)$-Array zurück geben, deren Einträge $(i, j, 1)$ die Unter- und $(i, j, 2)$ die Obergrenze von $F_{i,j}(X)$ enthalten, $i, j \in \{1, 2\}$. Die Box X soll als $(2, 2)$-Array übergeben werden, deren Einträge $(i, 1)$ die Untergrenze $\underline{x}_i$ und $(i, 2)$ die Obergrenze $\overline{x}_i$, $i \in \{1, \dots, n\}$ der jeweiligen Intervalle enthalten. Nutzen Sie für die Grundrechenarten der Intervallarithmetik die Funktionen aus Aufgabe S. 3.2.

```
import numpy as np
from math import sqrt

def interval_hull(A):
    result=np.array([min(A), max(A)] )
    return result

def interval_add(x,y):
    result= np.array([x[0]+y[0],x[1]+y[1]])
    return result

def interval_subtract(x,y):
    result= np.array([x[0]-y[1],x[1]-y[0]])
    return result

def interval_multiply(x,y):
    A=np.array([x[0]*y[0],x[1]*y[1],x[0]*y[1],x[1]*y[0]])
    result= interval_hull(A)
    return result

def boxweite(x):
    result = (x[1]-x[0])
    return result

def mittelpunkt(x):
    result =((x[0]+x[1])/2)
    return result
```

```
def DsqF(X):
    result1=np.array(2*X[1,:])
    result2= interval_add(2*X[0,:],4*X[1,:])+1
    result3=interval_add(2*X[0,:],4*X[1,:])+1
    result4=np.array(4*X[0,:])
    res=np.array([[result1,result2],[result3,result4]])
    return res
```

Aufgabe 2

Es seien $A(x) = D^2 f(x)$ die (n,n)-Hessematrix einer Funktion f und alle Einträge a_{ij} faktorisierbar. Für eine Box $X \in \mathbb{IR}^n$ bezeichne $A(X)$ diejenige Matrix, deren Einträge A_{ij} natürliche Intervallerweiterung von a_{ij} sind.

Implementieren Sie eine Funktion *lambda_min* die für eine Box X und die Matrix $A(X)$ eine Unterschranke β des kleinsten Eigenwertes $\lambda_{\min}(D^2 f(x))$ auf X berechnet. Verwenden sie hierfür die Formel aus Skript S. 146.

Die intervallwertige Hessematrix A soll dabei als $(n,n,2)$-Array übergeben werden, deren Einträge $(i,j,1)$ die Unter- und $(i,j,2)$ die Obergrenze von $F_{i,j}(X)$ enthalten, $i,j \in \{1,\ldots,n\}$. Die Box X soll als $(n,2)$-Array übergeben werden, deren Einträge $(i,1)$ die Untergrenze $\underline{x}_i$ und $(i,2)$ die Obergrenze $\overline{x}_i$, $i \in \{1,\ldots,n\}$ der jeweiligen Intervalle enthalten.

```
def lambda_min(A):
    n=np.size(A[:,0,0])
    b=np.zeros(n)
    for i in range (0,n):
        R=np.zeros(n)
        for j in range (0,n):
            if j!=i :
                R[j]=max(|A[i,j,0]|,|A[i,j,1]|)
                print(R[j])
        print(R)
        b[i]=A[i,i,0]-sum(R)
    print(b)
    res=min(b)
    return res
```

Aufgabe 3

Betrachten Sie nun wieder die Funktion

$$f(x) = x_1 2x_2 + 2x_1x_2^2 + x_1x_2$$

auf der Box $X = [0,1] \times [-1,1]$.

a) Nutzen Sie Ihre in Aufgabe S. 4.1. implementierte Funktion, um die Matrix $D^2F(X)$ auszugeben. Verwenden Sie Ihre in Aufgabe 4.2 implementierte Fkt. *lambda_min*, um eine Unterschranke β des kleinsten Eigenwertes der Hessematrix D^2f auf X zu berechnen und geben Sie diese aus.

Beweis:

□

```
def lambda_min(A):
    n=np.size(A[:,0,0])
    b=np.zeros(n)
    for i in range (0,n):
        R=np.zeros(n)
        for j in range (0,n):
            if j!=i :
                R[j]=max(A[i,j,0],A[i,j,1])
                print(R[j])
        print(R)
        b[i]=A[i,i,0]-sum(R)
    print(b)
    res=min(b)
    return res
```

```
X=np.array([[0,1],[-1,1]])
A=DsqF(X)
print(A)
beta=lambda_min(A)
print(beta)
```

```
[[[-2  2]
  [-3  7]]

 [[-3  7]
  [ 0  4]]]
7.0
[ 0.  7.]
7.0
[ 7.  0.]
[-9. -7.]
-9.0
```

b) Bestimmen Sie mit Hilfe von Aufgabenteil a) eine konvexe Relaxierung $\hat{f}_\alpha$ der Funktion f auf X. Geben Sie diese auf der schriftlichen Abgabe an.

Beweis:

$$\hat{f}_\alpha = x_1^2 x_2 + 2x_1 x_2^2 + x_1 x_2 + \frac{9}{2}\left(-x_1 + x_1^2 + x_2^2 - 1\right)$$

□

c) Plotten Sie die beiden Funktionen $\hat{f}_\alpha$ und f auf X in eine Grafik, sodass gut erkennbar ist, dass $\hat{f}_\alpha$ konvexe Relaxierung von f auf X ist.

Beweis:

□

```python
def f(x_1,x_2):
    result = x_1**2*x_2+2*x_1*x_2**2+x_1*x_2
    return result
```

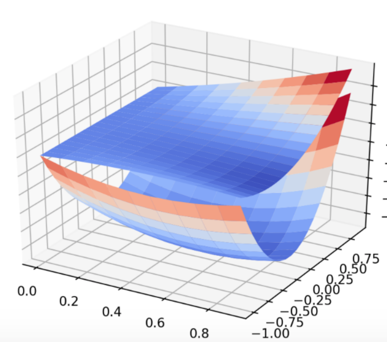
```python
def psi(x_1,x_2):
    result = 1/2*((-x_1)*(1-x_1)+(-1-x_2)*(1-x_2))
    return result
```

```python
def fdach_alpha(x_1,x_2,a):
    result = f(x_1,x_2)+a*psi(x_1,x_2)
    return result
```

```python
import numpy as np
from matplotlib import pyplot as plt
%matplotlib notebook
from mpl_toolkits.mplot3d import axes3d
from matplotlib import cm
```

```python
x = np.arange(0,1,0.1)
y = np.arange(-1,1,0.1)
X,Y = np.meshgrid(x,y)
Z = f(X,Y)
Z1 = fdach_alpha(X,Y,9)
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
surf1 = ax.plot_surface(X, Y, Z, cmap=cm.coolwarm)
surf2 = ax.plot_surface(X, Y, Z1, cmap=cm.coolwarm)
plt.show()
```

Figure 4



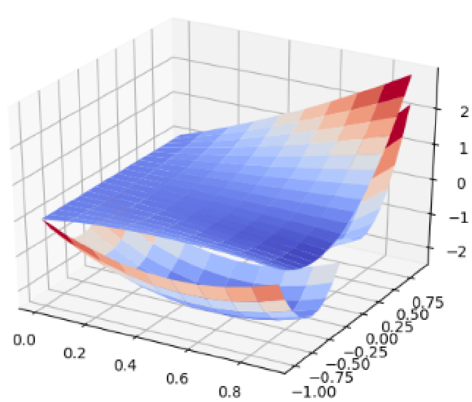
d) Welche Teilboxen X^1 und X^2 ergeben sich, wenn die Box entsprechend Algorithmus 3.3 halbiert wird? Mit der Notation von Algorithmus 3.3: Plotten Sie f auf X, $\hat{f}_\alpha^1$ auf X^1 und $\hat{f}_\alpha^2$ auf X^2 in eine Grafik.

Beweis: Nach Algorithmus 3.3

$$X^1 = [0,1] \times [-1,0], \quad X^2 = [0,1] \times [0,1]$$

```
def psi_1(x_1,x_2):
    result = 1/2*((-x_1)*(1-x_1)+(-1-x_2)*(-x_2))
    return result
def psi_2(x_1,x_2):
    result = 1/2*((-x_1)*(1-x_1)+(-x_2)*(1-x_2))
    return result
def fdach_alpha_1(x_1,x_2,a):
    result = f(x_1,x_2)+a*psi_1(x_1,x_2)
    return result
def fdach_alpha_2(x_1,x_2,a):
    result = f(x_1,x_2)+a*psi_2(x_1,x_2)
    return result
```

```
x = np.arange(0,1,0.1)
y = np.arange(-1,1,0.1)
X,Y = np.meshgrid(x,y)
x1 = x
y1 = np.arange(-1,0,0.1)
X1,Y1 = np.meshgrid(x1,y1)
x2 = x
y2 = np.arange(0,1,0.1)
X2,Y2 = np.meshgrid(x2,y2)
Z = f(X,Y)
Z1 = fdach_alpha_1(X1,Y1,9)
Z2 = fdach_alpha_2(X2,Y2,9)
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
surf1 = ax.plot_surface(X, Y, Z, cmap=cm.coolwarm)
surf2 = ax.plot_surface(X1, Y1, Z1, cmap=cm.coolwarm)
surf3 = ax.plot_surface(X2, Y2, Z2, cmap=cm.coolwarm)
plt.show()
```



□

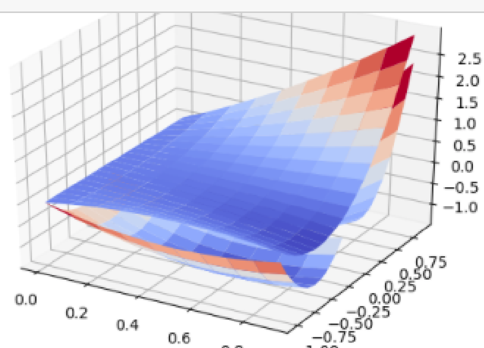
e) Bestimmen Sie nun mit der Funktion *lambda_min* individuelle Unterschranken des kleinsten Eigenwertes der Hessematrix $D^2 f$ auf X^i, $i \in \{1,2\}$ und geben Sie diese aus. Nutzen Sie diese, um (möglicherweise) bessere Relaxierungen $\hat{f}^1_{\alpha_1}$ und $\hat{f}^2_{\alpha_2}$ zu bestimmen. Plotten Sie f auf X, $\hat{f}^1_{\alpha_1}$ auf X^1 und $\hat{f}^2_{\alpha_2}$ auf X^2 in eine Grafik. Vergleichen Sie Ihr Ergebnis mit Aufgabenteil d).

Beweis: Im Vergleich zur d) liegt der Graph der Relaxierung etwas näher an dem der ursprünglichen Funktion. □

```
X_1=np.array([[0,1],[-1,0]])
X_2=np.array([[0,1],[0,1]])
A_1=DsqF(X_1)
A_2=DsqF(X_2)
beta_1=lambda_min(A_1)
beta_2=lambda_min(A_2)
print(beta_1)
print(beta_2)
```

```
3.0
[ 0.  3.]
3.0
[ 3.  0.]
[-5. -3.]
7.0
[ 0.  7.]
7.0
[ 7.  0.]
[-7. -7.]
-5.0
-7.0
```

```
x = np.arange(0,1,0.1)
y = np.arange(-1,1,0.1)
X,Y = np.meshgrid(x,y)
x1 = x
y1 = np.arange(-1,0,0.1)
X1,Y1 = np.meshgrid(x1,y1)
x2 = x
y2 = np.arange(0,1,0.1)
X2,Y2 = np.meshgrid(x2,y2)
Z = f(X,Y)
Z1 = fdach_alpha_1(X1,Y1,5)
Z2 = fdach_alpha_2(X2,Y2,7)
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
surf1 = ax.plot_surface(X, Y, Z, cmap=cm.coolwarm)
surf2 = ax.plot_surface(X1, Y1, Z1, cmap=cm.coolwarm)
surf3 = ax.plot_surface(X2, Y2, Z2, cmap=cm.coolwarm)
plt.show()
```



f) Bestimmen Sie Optimalpunkte von $\hat{f}^1_{\alpha_1}$ auf X^1 und $\hat{f}^2_{\alpha_2}$ auf X^2 mit einem Solver Ihrer Wahl. Runden Sie das Ergebnis auf zwei Nachkommastellen. Bestimmen Sie $\tilde{v}$ und $\hat{v}^*$? Begründen Sie: Würde X^1 oder X^2 als X^* für die nächste Iteration von Algorithmus 3.3 verwendet werden?

Beweis: Es ist $\hat{x}^1 \approx (0.62, -0.47)$, $\hat{x}^2 \approx (0.38, 0.35)$

$$\Rightarrow \tilde{x} = \hat{x}^1$$

$\tilde{v} = f(\hat{x}^1) \approx -0.1982$, $\hat{v}^* = \hat{f}_{\alpha_1}(\hat{x}^1) \approx -1.4099$

$$\Rightarrow X^* = X^1, \text{ da } \hat{v}^1 = \min_{l=1,2} \hat{v}^l$$

$\square$

```
from scipy.optimize import minimize
```

```
def fdach_alpha_1(x):
    result = f(x[0],x[1])+5*psi_1(x[0],x[1])
    return result
def fdach_alpha_2(x):
    result = f(x[0],x[1])+7*psi_2(x[0],x[1])
    return result
```

```
res_1 = minimize(fdach_alpha_1,np.array([0.5,-0.5]))
res_2 = minimize(fdach_alpha_2,np.array([0.5,0.5]))
```

```
res_1.x
```

```
array([ 0.62258489, -0.46863049])
```

```
res_2.x
```

```
array([ 0.37714705,  0.35030619])
```

```
f(0.62,-0.47)
```

```
-0.198152
```

```
f(0.38,0.35)
```

```
0.27664
```

```
fdach_alpha_1(res_1.x)
```

```
-1.409923900695528
```

```
fdach_alpha_2(res_2.x)
```

```
-1.3442389749970767
```

g) Bestimmen Sie, aufbauend auf Aufgabenteil f), ein möglichst kleines Intervall, in dem der optimale Zielfunktionswert v liegt.

Beweis: Es ist

$$-1.4099 \leq v \leq -0.1982 \iff v \in [-1.4099, -0.1982]$$

$\square$

h) Was passiert, wenn Sie versuchen mit dem in Aufgabenteil d) verwendeten Solver das Minimum von f auf X mit Startpunkt $x_0 = (0,0)^T$ auszurechnen? Vergleichen Sie Ihr Ergebnis mit Aufgabenteil f) und g).

Beweis: Mit dem Solver kommt als Minimalpunkt $x^* = (0,0)$ raus, da

$$f((0,0)) = 0 \notin [-1.4099, -0.1982],$$

liegt der Optimalwert im Vergleich zu den vorherigen Aufgaben über der Oberschranke $\tilde{v} = -0.1982 \Rightarrow$ schlechteres Ergebnis. $\square$

Aufgabe 4

Gegeben seien stetige Funktionen $f, g\colon X \to \mathbb{R}$, auf einer konvexen und kompakten Menge $X \subseteq \mathbb{R}^n$, die nicht notwendigerweise eine Box ist. Darüber hinaus sei f nicht konvex und

$$\hat{f} \coloneqq f + g$$

eine konvexe Relaxierung von f auf X.
Sei o.B.d.A X nicht-leer, dann sonst ist in den folgenden Teilaufgaben nichts zu zeigen.

a) Folgt hieraus, dass g eine konvexe Funktion ist? Beweisen Sie Ihre Behauptung.

Beweis: Die gilt im Allgemeinen nicht. Betrachtet man den Fall $n = 2$ und $X = [-1, 1]^2$ (wodurch X konvex ist), so folgt mit

$$f(x) = \frac{1}{2}\left(3x_1^2 - x_2^2\right) \quad \text{und} \quad \tilde{g}(x) = \frac{1}{2}\left(-x_1^2 + 3x_2^2\right),$$

dass f und $\tilde{g}$ als Polynome in $C^1(X)$ liegen. Außerdem sind $f, \tilde{g}$ nicht konvex, was aus Satz 2.2.2 (C^1-Charakterisierung von Konvexität) folgt, da

$$f\left(\begin{pmatrix}1\\1\end{pmatrix}\right) = 1 \leq 1 + 2 = f\left(\begin{pmatrix}1\\-1\end{pmatrix}\right) + \left\langle \begin{pmatrix}3\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}1\\-1\end{pmatrix} \right\rangle$$
$$\tilde{g}\left(\begin{pmatrix}1\\-1\end{pmatrix}\right) = 1 \leq 1 + 2 = \tilde{g}\left(\begin{pmatrix}-1\\-1\end{pmatrix}\right) + \left\langle \begin{pmatrix}1\\-3\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix} - \begin{pmatrix}-1\\-1\end{pmatrix} \right\rangle$$

Es gilt außerdem, dass

$$\frac{3}{2}x_1^2 \geq f(x) \geq -\frac{1}{2}x_2^2$$

für alle $x \in X$, sodass $f(x) \in [-0.5, 1.5]$ und analog $\tilde{g}(x) \in [-0.5, 1.5]$. Somit folgt mit $g \coloneqq g - 1.5$, dass als Verschiebung von $\tilde{g}$ die Funktion g ebenfalls nicht konvex ist, $\hat{f}(x) = f(x) + g(x) = x_1^2 + x_2^2 - 1.5 \leq f(x)$ für alle $x \in X$. Außerdem ist $\hat{f}$ in $C^2(X)$ als Polynom und damit

$$D^2\hat{f} = \begin{pmatrix}2 & 0\\0 & 2\end{pmatrix} \succ 0$$

wodurch $\hat{f}$ nach Satz 2.5.3 (C^2-Charakterisierung von Konvexität) konvex ist. Insgesamt erhalten wir eine konvexe Funktion $\hat{f}$ für die $\hat{f}(x) \leq f(x)$ gilt, d.h. $\hat{f}$ ist eine konvexe Relaxierung von f. $\square$

b) Nun gelte $g(x) \geq -2$ für alle $x \in X$. Zeigen Sie, dass für die Minimalwerte v bzw. $\hat{v}$ von f bzw. $\hat{f}$ auf X gilt:

$$0 \leq v - \hat{v} \leq 2.$$

Beweis: Da f und g nach Aufgabe stetig sind, ist $\hat{f}$ auch stetig. Somit nehmen, $f, \hat{f}$ und g als stetige Funktionen auf der kompakten Menge X nach dem Satz von Weierstraß (Satz 1.2.10) ihr Minimum an. Nach a) bzw. Definition 3.2.2 (Konvex relaxierte Funktion), ist $\hat{f}(x) \leq f(x)$ für alle $x \in X$, und damit auch

$$\min_{x \in X} \hat{f}(x) \leq \min_{x \in X} f(x) \iff 0 \leq \min_{x \in X} f(x) - \min_{x \in X} \hat{f}(x) = v - \hat{v}$$

Nach Voraussetzung ist

$$g(x) \geq -2 \iff -g(x) \leq 2.$$

Es folgt mit Übung 1.3.1 (Skalare Vielfache und Summen):

$$\min_{x \in X} \hat{f}(x) = \min_{x \in X} (f(x) + g(x)) \geq \min_{x \in X} f(x) + \min_{x \in X} g(x)$$

$$\iff \min_{x \in X} \hat{f}(x) - \min_{x \in X} f(x) \geq \min_{x \in X} g(x) \iff v - \hat{v} \leq -\min_{x \in X} g(x) = \max_{x \in X} -g(x) \leq 2,$$

wobei wir im letzten Schritt Ausgenutzt haben, dass 2 eine Oberschranke von $-g$ ist, d.h.

$$0 \leq v - \hat{v} \leq 2$$

$\square$

c) Nun seien $X \in \mathbb{IR}$, $f \in C^2(X, \mathbb{R})$, $D^2 f$ faktorisierbar und $g = \alpha\psi(x)$ per αBB Methode bestimmt. Zeigen Sie, dass für die Minimalwerte v bzw. $\hat{v}$ von f bzw. $\hat{f}$ auf X die folgende Abschätzung gilt

$$v - \hat{v} \leq \frac{\alpha}{8} w(X)^2$$

Beweis: Da die Voraussetzungen von a) und b) gelten und nach Übung 3.4.1

$$\psi(x) \geq \min_{x \in X} \psi(x) = -\frac{1}{8} w(X)^2,$$

folgt die Behauptung aus b), indem man

$$g(x) \coloneqq \alpha\psi(x) \quad \forall x \in X$$

setzt, damit nicht $g(x) \geq -2$ sondern $g(x) \geq \alpha \min_{x\in X} \psi(x) = -\frac{\alpha}{8} w(X)^2$, und da nach Satz 3.4.3 folgt, dass $\hat{f} = f + \psi$ eine konvexe Relaxierung von f auf X ist. $\square$

d) Geben Sie eine Funktion f, eine Box X und ein $\alpha \geq 0$ an, sodass das die Voraussetzungen aus c) erfüllt und die Ungleichung aus c) mit Gleichheit erfüllt ist.

Beweis: Sei für $X = [-1, 1]$ die Funktion $f(x) = -x^4 + x^2$, dann ist

$$f(x) \geq 0, \quad \text{da } -x^4 + x^2 \geq 0 \iff x^2 \leq 1$$

Außerdem ist f nach Definition 2.1.1 nicht konvex, da

$$\frac{1}{2} f(-1) + \frac{1}{2} f(0) = 0 < \frac{3}{16} = f(-0.5)$$

Nach Satz 3.5.2 (αBB-Methode) ist allerdings für $x \in X$ mit geeignetem α

$$\hat{f}(x) = f(x) + \alpha\psi(x)$$

die konvexe Relaxierung von f. Da $f(0) = 0 = \min_{x\in X} f(x)$ und nach Übung 3.4.1 $\psi(0) = \psi(m(X)) = \min_{x\in X} \psi(x)$ gilt, ist somit:

$$\min_{x\in X} \hat{f}(x) = \min_{x\in X} \left(f(x) + \alpha\psi(x)\right) = \min_{x\in X} f(x) + \alpha \min_{x\in X} \psi(x)$$

$$\Rightarrow \min_{x\in X} f(x) - \min_{x\in X} \hat{f}(x) = -\alpha \min_{x\in X} \psi(x) = \frac{\alpha}{8} w(X)^2$$

$\square$