Nonlinear Optimization – Sheet 12

Exercise 1

```
import numpy as np
import matplotlib.pyplot as plt


# solves systems as in the algorithm
def solve_system(M,B,zeta,m,n):
    C = np.block([[M,B],[B.transpose(),np.zeros((m,n))]])
    zeta_new = np.pad(zeta,(0,m),"constant")
    return np.linalg.solve(C,-zeta_new)[:n]

def truncated_cg_for_symm_systems(b,d_part,A,M,B,eps_rel):
    m = np.shape(B)[0]
    n = np.shape(B)[1]

    alpha = 0
    beta = 0
    q = 0
    theta = 0

    d = d_part
    zeta = np.dot(A,d) - b
    p = solve_system(M,B,zeta,m,n)
    p_new = 0
    initial_delta = -np.dot(zeta,p)
    delta = initial_delta
    delta_new = 0
    delta_array = [delta]
    while delta >= eps_rel**2 * initial_delta:
        q = np.dot(A,p)
        theta = np.dot(q,p)
        if theta > 0:
            alpha = delta / theta
            d = d + alpha*p
            zeta = zeta + alpha*q
            p_new = solve_system(M,B,zeta,m,n)
            delta_new = - np.dot(zeta,p_new)
            beta = delta_new / delta
            p_new = p_new + beta* p
            p = p_new
            delta = delta_new
            delta_array.append(delta)
        else:
            break
    return(delta_array)



A = np.array([[11,3],[3,11]])
M = np.array([[-1,2],[2,-1]])
B = np.array([[1,0],[0,1]])
```

```
# c = [1,1], so d_part = [1,1]
d_part = np.array([1,1])
b = np.array([24,44])

plt.plot(truncated_cg_for_symm_systems(b,d_part,A,M,B,0.001))
plt.show()
```

We did only implement the projected CG-method but did not implement any visualization.

Exercise 2

Lemma 0.1. *If a function f is Lipschitz-continuous and continuously differentiable on a neighborhood $U(x)$ of x, then f is semismooth at x.*

Proof. As the function is continuously differentiable, the limit of derivatives at iterates is the derivative of the limit of iterates. Consequently, the Bouligand derivative only consists of one element, the derivative of f at x. Consequently, the Clarke derivative has only one element, too. Therefore, the sequences M^k and d^k both converge and we have

$$\lim_{k\to\infty} M^k d^k = \left(\lim_{k\to\infty} M^k\right)\left(\lim_{k\to\infty} d^k\right) = f'(x) \cdot d.$$

In particular, the limit exists. □

(i) Away from 0, $|x|$ is continuosly differentiable. As can be clearly seen from the proof of Lemma 0.1, the Bouligand derivative for $x > 0$ is simply the singleton containing the derivative (which is always 1). Analogous considerations for $x < 0$ yield $\partial_B f(x) = \{-1\}$.

From these facts, it easily follows that $\partial_B f(x) = \{\pm 1\}$. For the Clarke derivative, we obtain

$$\begin{cases} \{1\} & x > 0 \\ [-1, 1] & x = 0 \\ \{-1\} & x < 0 \end{cases}.$$

(ii) By Rademachers theorem, the set of points in the given neighborhood U where f is Lipschitz-continuous has Lebesgue-measure 0. As a result, there exists sequences $x_k \to x$, $x_k \in D_F$. As the function is Lipschitz-continuous in this neighborhood, we conclude

$$|f'(y)| = \lim_{z\to y} \frac{|f(z) - f(y)|}{|z - y|} \le L \qquad \forall y \in U,$$

where L may denote the Lipschitz constant. Consider a sequence $x_k \to x$ and the induced sequence $f'(x_k)$. This sequence doesn't necessarily converge, but via Bolzano-Weierstrass, we obtain the existence of a convergent subsequence. As a result, the Bouligand generalized differential is nonempty. Also, it is bounded by L.

Assumption: $\partial_B f(x)$ is not closed. Then we have a sequence $M^k \in \partial_B f(x)$ with limit M that is not in $\partial_B f(x)$. For each k we have a sequence $x^{k,\ell}$ s.t. $f'(x^{k,\ell}) \to M^k$. W.l.o.g replace each such sequence with a subsequence that satisfies $|f'(x^{k,\ell}) - M^k| < \frac{1}{n} \forall \ell \ge n$. Now consider the diagonal sequence $x^{k,k}$.

Let $\varepsilon > 0$. Then choose k_1 s.t.

$$|M^k - M| < \frac{\varepsilon}{2} \forall k \ge k_1.$$

Let $k_2 \in \mathbb{N}$ s.t. $\frac{1}{k_2} < \frac{\varepsilon}{2}$. Then $\forall k$

$$|f'(x^{k,\ell}) - M^k| < \frac{\varepsilon}{2} \forall \ell \ge k_2.$$

Let $k_3 = \max(k_1, k_2)$. Then

$$|f'(x^{k_3,k_3}) - M| \le |f'(x^{k_3,k_3}) - M^{k_3}| + |M^{k_3} - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Therefore, the sequence $x^{k,k}$ converges to M and as a result, $M \in \partial_B f(x)$, contradiction.

We conclude that $\partial_B f(x)$ is bounded and closed and thus compact. $\partial f(x)$ as the convex hull of $\partial_B f(x)$ is clearly nonempty and still compact (and obviously convex).

Exercise 3

(i) Let $a, b \geq 0$ and $ab = 0$. $ab = 0 \implies a = 0 \vee b = 0$. W.l.o.g $a = 0$. Then,

$$\min(a, b) = \min(0, b) = 0$$

as $b \geq 0$. Also,

$$\sqrt{a^2 + b^2} - a - b = b - b = 0.$$

Let, on the other hand

$$\min(a, b) = 0.$$

Then, w.l.o.g $a = 0$ and $b \geq 0 \implies ab = 0, a, b \geq 0$. Furthermore,

$$\sqrt{a^2 + b^2} - a - b = 0 \Longleftrightarrow a^2 + b^2 = a^2 + 2ab + b^2 \implies ab = 0.$$

W.l.o.g $a = 0$. If $b < 0$, we have $0 = \sqrt{b^2} - b = -2b \implies b = 0$, contradiction. Therefore $b \geq 0$

(ii) (a) At any point (a, b) we have

$$|\min(a + x, b + y) - \min(a, b)| = |\min(x, y)| \leq 1 \cdot |(a + x, b + y) - (a, b)|.$$

Therefore, $\min(a, b)$ is Lipschitz-continuous on $\mathbb{R}^2$. On H^+ and H^-, every point has a neighborhood as required in Lemma 0.1 where it is Lipschitz-continuous and continuously differentiable.

In order to show that it is semismooth everywhere, we therefore only need to consider points $z \in H$, i.e. of the form (a, a). There, according to example 13.6, the Clarke derivative is given by $G := \{(\alpha, 1 - \alpha) | \alpha \in [0, 1]\}$.

$$\begin{aligned} M^k \cdot d^k &= \begin{cases} d_2^k & | d_2^k < d_1^k \\ \alpha d_1^k + (1 - \alpha) d_2^k = d_2^k & | d_1^k = d_2^k \\ d_1^k & | d_1^k < d_2^k \end{cases} \\ &= \min(d_1, d_2) \end{aligned}$$

Let $\varepsilon > 0$. We know that d^k converges and therefore also d_1^k and d_2^k. Take N s.t. $\forall k \geq N$ $|d_1^k - d_1| < \varepsilon$ and $|d_2^k - d_2| < \varepsilon$. Then $|\min(d_1^k, d_2^k) - \min(d_1, d_2)| < \varepsilon$. We conclude that $M^k \cdot d^k$ converges to $\min(d_1, d_2)$ and in particular, the desired limit exists.

(b) $f := \Phi_{FB}$ is differentiable everywhere apart from 0. Also, it is locally Lipschitz-continuous everywhere apart from 0 as a concatenation of locally Lipschitz-continuous functions. In 0, it is also Lipschitz-continuous:

$$|\sqrt{x^2 + y^2} - 0| \leq 1 \cdot \|(x, y) - (0, 0)\|_2.$$

With Lemma 0.1, we conclude that it is semismooth everywhere apart from 0. In order to show the semismoothness in 0, we compute the generalized derivatives.

We have

$$\|f'(x^k, y^k) + (1, 1)\|_2 = \sqrt{\frac{4(x^k)^2 + 4(y^k)^2}{(x^k)^2 + (y^k)^2}} = 2$$

and in particular

$$\|\lim_{k \to \infty} f'(x^k, y^k) + (1, 1)\|_2 = \lim_{k \to \infty} \|f'(x^k, y^k) + (1, 1)\|_2 = 2.$$

Therefore, the generalized Bouligand derivative forms a circle of radius two around $(-1,-1)$ and the generalized Clarke derivative is the corresponding filled circle. We derive that

$$\|d\| \leq 2+\sqrt{2} \leq 4 \qquad \forall d \in \partial f(x) \forall x \in \mathbb{R}^n.$$

If $d=0$, i.e. $d^k \to 0$, we have

$$\|\lim_{k\to\infty} M^k d^k\|_2 = \lim_{k\to\infty} \|M^k\|_2 \|d^k\|_2 \leq 4 \cdot \lim_{k\to\infty} \|d^k\|_2 = 0.$$

Consequently, $\lim\limits_{k\to\infty} M^k d^k = 0$.

If $d \neq 0$ and $t^k \neq 0 \, \forall k$ (w.l.o.g. if there are only finitely many $t^k = 0$. In any other case, consider the counterexample at the end of the proof.) For big enough k, $d^k \neq 0$. Then, $M^k \in \partial f(t^k d^k) = f'(t^k d^k) = \left(\frac{x^k}{\sqrt{(x^k)^2+(y^k)^2}} - 1, \frac{y^k}{\sqrt{(x^k)^2+(y^k)^2}} - 1 \right)$. We obtain

$$\lim_{k\to\infty} M^k d^k = t^k \cdot \left(\frac{(x^k)^2+(y^k)^2}{\sqrt{(x^k)^2+(y^k)^2}} - x^k - y^k \right)$$

$$= \lim_{k\to\infty} t^k \left(\sqrt{(x^k)^2+(y^k)^2} - x^k - y^k \right) = 0 \cdot (\sqrt{d_1^2+d_2^2} - d_1 - d_2) = 0,$$

the limit exists.

If $d \neq 0$ and $t^k = 0 \forall k$ (w.l.o.g. if there are infinitely many such k because then we can just choose a subsequence), we find a counterexample by choosing $d^k = d = (1,0)$ and

$$M^k = \begin{cases} (1,-1) & k = 2l \\ (-1,1) & k = 2l+1 \end{cases}.$$

The choice for M^k is possible as $\partial f(z + t^k d^k) = \partial f(0) = \{d \in \mathbb{R}^n \colon \|d + (1,1)\| \leq 2\}$. Then

$$\lim_{k\to\infty} M^k d^k = \lim_{k\to\infty} (-1)^k$$

This limit doesn't exist.

Exercise 4

(a) This is obvious from looking at equation (13.15) for $m = 0$.

(b) Consider

$$\begin{aligned} \min\left(\mu^{m+1}, d^{m+1} - \ell\right) &= \begin{cases} \mu_i & i \in I\left(d^{m+1}, \mu^{m+1}\right) \\ d_i^{m+1} - \ell_i & i \in I\left(d^{m+1}, \mu^{m+1}\right) \end{cases} \\ &= \begin{cases} \mu_i & i \in I\left(d^{m}, \mu^{m}\right) \\ d_i^{m+1} - \ell_i & i \in I\left(d^{m}, \mu^{m}\right) \end{cases} \end{aligned}$$

according to the lecture notes below equation 13.15

$$\begin{aligned} &= \begin{cases} 0 \\ 0 \end{cases} \\ &= 0. \end{aligned}$$

The other required equalities are clear from (a).