Modern Optimization Techniques – Group 01

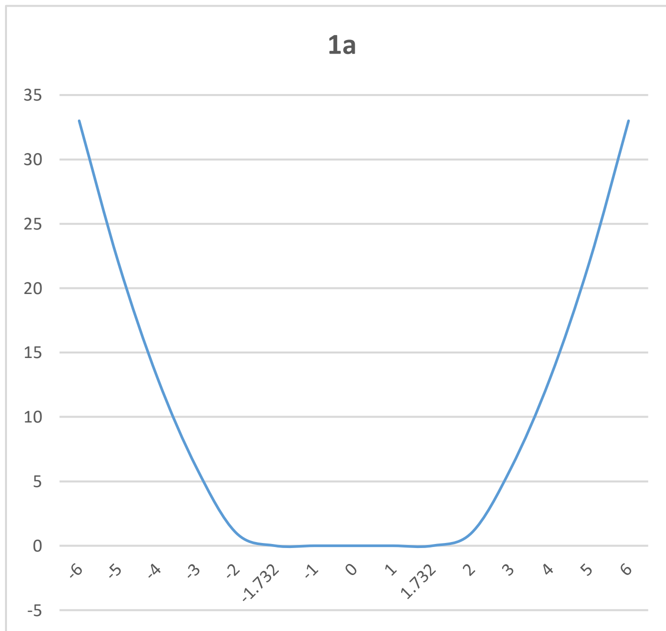
Exercise Sheet 06

Submitted by: Muhammad Inaam Ashraf (Matrikel-Nr: 307524)

Semester 2 MSc. Data Analytics

Question 1: Subgradients and Subdifferentials

(a). $\boldsymbol{f(x) = max\{0, x^2 - 3\}}$



Plotting the function gives the graph on the right side.
We can also re-write the function as

$$f(x) = \begin{cases} x^2 - 3 \; for \;\; x^2 - 3 > 0 \\ 0 \qquad for \;\; x^2 - 3 \leq 0 \end{cases}$$

Now,

$$x^2 - 3 > 0$$

$$x^2 > 3$$

$$-1.732 > x > 1.732$$

So,

$$f(x) = \begin{cases} x^2 - 3 \; for - 1.732 > x > 1.732 \\ 0 \qquad for \;\; - 1.732 \leq x \leq 1.732 \end{cases}$$

Derivative of the function is given by

$$\frac{\partial f(x)}{\partial x} = \begin{cases} 2x \qquad for - 1.732 > x > 1.732 \\ 0 \qquad for \;\; - 1.732 < x < 1.732 \end{cases}$$

As can be seen, at two points, this function is not differentiable and these points are

$$\boldsymbol{x_0 = -1.732 \; and + 1.732}$$

At these points, we can have a range of subgradients between the gradient from one side and the one from the other.

$$\partial f(x_0) = set \; of \; subgradients \; between \; 0 \; and \; 2x_0$$

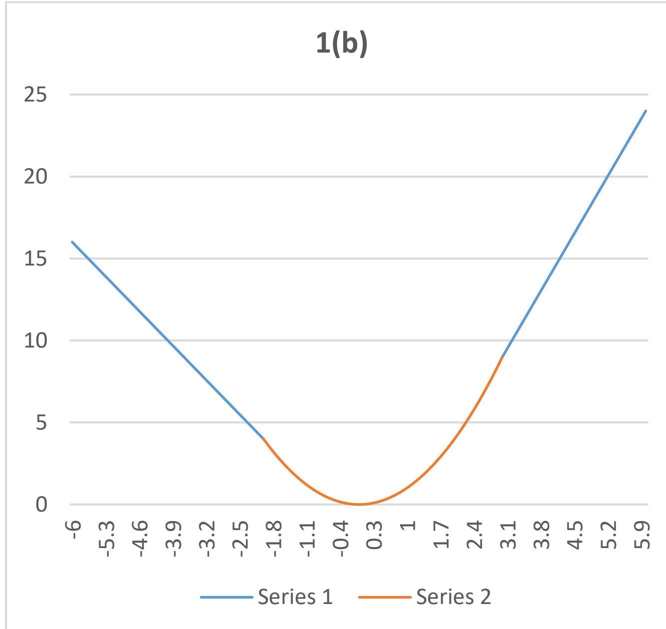
Therefore, for $x_0 = -1.732$

$$\boldsymbol{\partial f(-1.732) = [2(-1.732), 0]}$$

And for $x_0 = 1.732$

$$\boldsymbol{\partial f(-1.732) = [0, 2(1.732)]}$$

(b). $f(x) = \begin{cases} -3x - 2 & x \in (-\infty, -2] \\ x^2 & x \in (-2, 3) \\ 5x - 6 & x \in [3, \infty) \end{cases}$



Plotting the function gives the graph on the right side.

As can be seen from the plot, the given function is **not convex**. This is because if we draw a tangent at points -2 and 3, it will overlap with the function and will not be below the function at all times. Therefore,

$$x_0 = -2 \; and \; 3$$

and

$$\partial f(-2) = [\,]$$

$$\partial f(3) = [\,]$$

For a non-convex function, subgradients make less sense. Generalized subgradients are used that are beyond the scope of this lecture as discussed by Professor Lars in the lecture (Slide 8).

Question 2: Subgradients and Subdifferentials of L1-Norm

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

The L1- norm can also be written as

$$f(x) = ||x||_1 = |x_1| + |x_2| + |x_3| + \cdots + |x_n|$$

In its current form, it is not differentiable. But as given in the question, we can also write it as maximum of 2^n functions as follows:

$$||x||_1 = \max\{s^T x \mid s_i \in \{-1,1\}\}$$

Now, we try to identify an active function $s^T x$ and can write for a single i

$$||x_i||_1 = \begin{cases} +1x_i & for \; x_i > 0 \\ -1x_i & for \; x_i < 0 \\ \pm 1x_i & for \; x_i = 0 \end{cases}$$

As can be seen, we have chosen

$$s_i = +1 \; for \; x_i > 0$$

$$s_i = -1 \; for \; x_i < 0$$

$$s_i = \pm 1 \; for \; x_i = 0$$

Now, we can compute subgradient for a single i as follows:

$$g_i = \begin{cases} +1 & for\ x_i > 0 \\ -1 & for\ x_i < 0 \\ \pm 1 & for\ x_i = 0 \end{cases}$$

Since, this is the ith subgradient, the subdifferential will be a convex hull of all such subgradients and can be written as

$$\partial f(x) = \{g \mid ||g||_\infty \leq 1, g^T x = \left||x|\right|_1\}$$