Authors:

- Mhd Jawad Al Rahwanji - 7038980 - mhal00002@stud.uni-saarland.de
- Christian Singer - 7039059 - chsi00002@stud.uni-saarland.de

Exercise 5.1 - Newton's Method

a) The taylor series of $f(x)$:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \ldots$$

Using the first 3 terms (2nd order) because that way we'd have a parabola:

$$f(x) \approx f(x_0) + f'(x_0)x - f'(x_0)x_0 + \frac{1}{2}f''(x_0)x^2 - f''(x_0)x_0x + \frac{1}{2}f''(x_0)x_0^2$$

We now take the derivative of the 2nd order approximation of the function and set it equal to 0 to find its minimum.

$$\frac{f(x)}{dx} \approx f'(x_0) + f''(x_0)x - f''(x_0)x_0 = 0$$

$$f''(x_0)x = f''(x_0)x_0 - f'(x_0)$$

$$x = x_0 - \frac{f'(x_0)}{f''(x_0)}$$

Similarly, in the case of a multivariable f:

$$x_{t+1} = x_t - f''(x_t)^{-1}f'(x_t)$$

We swap $f''(x_t)^{-1}$ for $H(f(x_t))^{-1}$ and $f'(x_t)$ for $\nabla f(x_t)$:

$$x_{t+1} = x_t - aH(f(x_t))^{-1}\nabla f(x_t)$$

b) We have the following:

$$f(x) = f(x_1, x_2) = x_1^2 - 3x_1 + x_2^2 - x_1x_2,$$

$$\epsilon = 0.5, \text{ and}$$

$$\nabla f(x) = \nabla f(x_1, x_2) = (2x_1 - x_2 - 3,\ 2x_2 - x_1)$$

The gradient descent iterations:

Iteration	x	*f*(x)	∇*f*(x)	*L2-norm(∇f(x))*
0	(1, 1)	-2	(-2, 1)	2.236
1	(2, 0.5)	-2.75	(0.5, -1)	1.118
2	(1.75, 1)	-2.937	(-0.5, 0.25)	0.559
3	(2, 0.875)	-2.984	(0.125, -0.25)	0.279
4	**(1.9375, 1)**	**-2.996**	**(-0.125, 0.0625)**	**0.139**

c) We have the following:

$$f(x) = f(x_1, x_2) = x_1^2 - 3x_1 + x_2^2 - x_1x_2, \quad a = 0.5,$$

$$\nabla f(x) = \nabla f(x_1, x_2) = (2x_1 - x_2 - 3,\ 2x_2 - x_1),$$

$$H(f(x)) = H(f(x_1, x_2)) = [(2,\ -1), (-1,\ 2)], \text{ and}$$

$$H(f(x))^{-1} = H(f(x_1, x_2))^{-1} = [(\tfrac{2}{3},\ \tfrac{1}{3}),\ (\tfrac{1}{3},\ \tfrac{2}{3})]$$

Iteration	x	*f*(x)	∇*f*(x)	*L2-norm*(∇*f*(x))
0	(1, 1)	-2	(-2, 1)	2.236
1	(1.5, 1)	-2.75	(-1, 0.5)	1.118
2	(1.75, 1)	-2.937	(-0.5, 0.25)	0.559
3	(1.875, 1)	-2.984	(-0.25, 0.125)	0.279
4	(1.9375, 1)	-2.996	(-0.125, 0.0625)	0.139

$H(f(x))$ is positive definite therefore f is convex and has a global minimum at $\hat{x} = (2,\ 1)$ and $f(\hat{x}) = -3$. It is what we get from the Newton method.

d) We have the following:

$$f(x) = 2x^3 - 5x,$$

$$a = 0.5,$$

$$f'(x) = 6x^2 - 5,$$

$$f''(x) = 12x, \text{ and}$$

$$f''(x)^{-1} = \frac{1}{12x}$$

When $x_0 = 0$, we get a division by zero problem which makes performing Newton's method impossible. In this case, the convergence, or lack thereof, is entirely dependent on x_0. On another note, the function has a local minimum and no global minimum since it isn't convex.