1 Nelder-Mead method

In this part optimization results are presented for Nelder-Mead and GD methods. We are comparing the performance for functions from previous homework. The results for f_1 are presented in Table 1. Each value for specific setting (Steps or elapsed time) is presented as absolute difference ($|y(x^*) - y(x_n)|$). We can see that Newton method is showing best results, closely followed by BFGS and Nelder-Mead. In this setting we used starting simplex of diameter $D \approx 1$. For gradient descent we used fixed learning rates($\gamma = 0.001$ $\mu = 0.009$). Also it is worth noticing that GD,Polyak and Nesterov are having same performance. Timewise, all of the methods (besides Newton) get stuck close to the theoretical optimum with absolute difference of $\approx e^-17$, while Nelder-Mead is having absolute difference of $\approx e^-16$ magnitude.

N / Time limit	Gradient descent	Polyak GD	Nesterov GD	AdaGrad GD	Newton method	BFGS	NelderMead
N = 2	0.19396	0.19396	0.19396	56.78663	0.07929	0.04138	0.19791
N = 5	0.18830	0.18829	0.18829	2.50917	0.02148	$2.775e^{-17}$	0.19791
N = 10	0.17944	0.17943	0.17943	0.5794	0.00243	$2.775e^{-17}$	0.19791
N = 100	0.08651	0.08645	0.08645	$3.306e^{-12}$	0.0	$2.775e^{-17}$	$1.2706e^{-13}$
T = 0.1 s	(Steps,min): $(14642, 2.775e^{-17})$	(Steps,min): $(13273, 2.775e^{-17})$	(Steps,min): $(10390, 2.775e^{-17})$	(Steps,min): $(12212, 2.775e^-17)$	(Steps,min): $(19527, 0.0)$	(Steps,min): $(3, 2.775e^{-17})$	(Steps,min): $(539, 1.1102e^{-16})$
T = 1 s	(Steps,min): $(172766, 2.775e^{-17})$	(Steps,min): $(138063, 2.775e^{-17})$	(Steps,min): $(111802, 2.775e^{-17})$	(Steps,min): $(120040, 2.775e^{-17})$	(Steps,min): $(196930, 0.0)$	(Steps,min): $(3, 2.775e^{-17})$	(Steps,min): $(5322, 1.1102e^{-16})$
T = 2 s	(Steps,min): $(346840, 2.775e^{-17})$	(Steps,min): $(277356, 2.775e^{-17})$	(Steps,min): $(229919, 2.775e^{-17})$	(Steps,min): $(256961, 2.775e^{-17})$	(Steps,min): $(400150, 0.0.)$	(Steps,min): $(3, 2.775e^{-17})$	(Steps,min): $(10618, 1.1102e^{-16})$

Table 1: Results for $f_1(x, y, z)$ - absolute difference ($|y(x*) - y(x_n)|$)

For $f_2(x, y, z)$, results are presente in Table 2. BFGS method is closest to the optimum after 100 steps, followed by next best results using Nelder-Mead method. When we consider time performance results, we can notice that GD, Polyak,Nesterov are closer to optimum after $1s$, rather then $2s$ run. This trend is also visible in Newton and BFGS method. Nelder-Mead is converging nicely.

N / Time limit	Gradient descent	Polyak GD	Nesterov GD	AdaGrad GD	Newton method	BFGS	NelderMead
N = 2	0.04237	0.04162	0.04218	672.9987	1.87405	0.40476	8.875
N = 5	0.01822	0.01821	0.01822	66.38255	1.70704	0.00393	5.69913
N = 10	0.0181	0.0181	0.0181	6.01104	1.9019	0.0004	1.20708
N = 100	0.01687	0.01687	0.01687	1.14332	2.2119	$8.09e^{-23}$	0.00401
T = 0.1 s	(Steps,min): $(14156, 3.971e^{-8})$	(Steps,min): $(11761, 3.826e^{-7})$	(Steps,min): $(9979, 2.08e^{-6})$	(Steps,min): $(10817, 0.00019)$	(Steps,min): $(13113, 4.456e^{-10})$	(Steps,min): $(18, 8.0909e^{-23})$	(Steps,min): $(524, 9.8607e^{-31})$
T = 1 s	(Steps,min): $(140772, 2.9874e^{-26})$	(Steps,min): $(117289, 2.98e^{-26})$	(Steps,min): $(99356, 2.98e^{-26})$	(Steps,min): $(109609, 4.07181e^{-27})$	(Steps,min): $(130356, 5.3374e^{-27})$	(Steps,min): $(18, 8.09e^{-23})$	(Steps,min): $(5080, 9.8607e^{-31})$
T = 2 s	(Steps,min): $(346840, 2.775e^{-17})$	(Steps,min): $(277356, 2.775e^{-17})$	(Steps,min): $(229919, 2.775e^{-17})$	(Steps,min): $(256961, 2.775e^{-17})$	(Steps,min): $(400150, 0.0.)$	(Steps,min): $(3, 2.775e^{-17})$	(Steps,min): $(10145, 9.8607e^{-31})$

Table 2: Results for $f_2(x, y, z)$ - absolute difference ($|y(x*) - y(x_n)|$)

In Table 2 performance of Nelder-Mead is presented for starting simplex of different diameter. All values are actual best score minimal values of the $f_1(x, y, z)$. We can see that best converge is performed with smallest diameter (≈ 0.1). As we construct simplex with larger diameter convergence rate are lower and overall performance is worse.

Iterations	Diameter 0.1	Diameter 0.5	Diameter 1	Diameter 2	Diameter 4
N = 2	0.0	0.0	0.0	0.0	0.0
N = 5	−0.0711	0.0	0.0	0.0	0.0
N = 10	−0.1681	−0.1266	0.0	−0.1173	0.0
N = 100	−0.1979	−0.1979	−0.1979	−0.1979	−0.1979

Table 3: Results for different different starting diameter simplex in Nelder Mead ($min\ f_1(x, y, z) = -0.1979$)

2 Blackbox function

In Table 4, blackbox procedure results are presented for three options ($i \in \{1, 2, 3\}$). In this case we can see that best overall result for three cases is shown for case of $D \approx 2$. We can see that $D \approx 4$ is giving worst results for $i \in \{1, 3\}$, while $D \approx 0.1$ is producing worst result for $i = 2$.

i	Diameter ≈ 0.1	Diameter ≈ 0.5	Diameter ≈ 1	Diameter ≈ 2	Diameter ≈ 4
1	3.708	2.8857	2.8888	2.8888	4.2353
2	1787.488	1147.389	720.58	70.027	128.86
3	1.4945	1.3414	1.2358	1.2358	1.5497

Table 4: f_{min} for black box procedure with Nelder Mead

All of the results are stored in project .txt files. The source code is in .py files.