Joanna Grace S. Saladino
Math 288: Numerical Optimization
Problem Set 2

1. Consider the following function with the given global minimizer (x^*, y^*).

Function	Bounds	Global Minimizer
$f_1(x,y) = 100(y-x^2)^2 + (x-1)^2$	$-30 \leq x, y \leq 30$	(1,1)
$f_2(x,y) = 100\sqrt{\lvert y - 0.01x^2 \rvert} + 0.01\lvert x+10 \rvert$	$-15 \leq x \leq -5, -3 \leq y \leq 3$	(-10,1)
$f_3(x,y) = x\sin(4x) + 1.1\sin(2y)$	$0 \leq x, y \leq 10$	(9.039, 8.668)

TABLE 1. Average and Best minimizer using Genetic Algorithm (GA) and Nelder Mead Simplex Method (NMSM)

Function	GA		NMSM	
	Average	Best	Average	Best
f_1	(0.59, 1.60)	(1.01, 1.02)	(1.00, 1.00)	(1.00, 1.00)
f_2	(-8.76, 0.91)	(-8.63, 0.75)	(-9.60, 1.04)	(-10.31, 1.06)
f_3	(8.88, 8.35)	(9.03, 8.66)	(6.69, 5.07)	(9.039, 8.668)

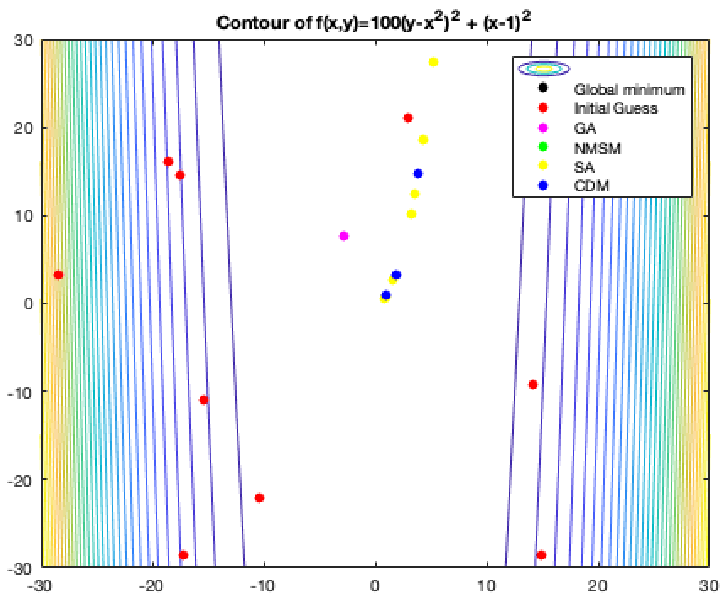
TABLE 2. Average and Best minimizer using Simulated Annealing (SA) and Coordinate Descent Method (CDM)

Function	SA		CDM	
	Average	Best	Average	Best
f_1	(1.78,5.18)	(1.00, 1.01)	(1.36, 2.58)	(1.00, 1.00)
f_2	(-9.28, 0.98)	(-10.59, 1.12)	(-10.71,1.18)	(-10.21, 1.04)
f_3	(8.72, 8.63)	(9.04, 8.67)	(6.846, 4.758)	(9.039,8.668)

1.1 The contour and scatter plot of the obtained minimizers of the first function

$$f_1(x,y) = 100(y-x^2)^2 + (x-1)^2$$

FIGURE 1. Contour of f_1 and the scatter plot of the obtained minimizers



It is evident from Table 1 and Table 2, that the *best* values from all four algorithms are near the global minimizer (1,1) but the *average* values are far from it especially in the case of SA and CDM. This observation is supported by Figure 1 which shows that the obtained minimizers from the 10 runs of SA and CDM didn't accurately estimate the global minimizer. Unlike SA and CDM, all ten runs from NMSM obtained (1,1) as a minimizer which means that despite the initial points, the NMSM successfully get the global minimizer. Although, GA didn't accurately get the global minimizer, all the obtained minimizer from this algorithm are all near the global minimum making it a good estimate for it.
It is worth noting also that among the 10 initial points, the obtained minimizer from the *fifth* run which has na initial point (19.3878, 21.4378) was the farthest from the global minimizer and this is true for NMSM, SA and CDM.

1.2 The contour of the second function

$$f_2(x,y) = 100\sqrt{|y - 0.01x^2|} + 0.01|x + 10|$$

and scatter plot of the obtained minimizers is

FIGURE 2. Contour of f_2 and the scatter plot of the obtained minimizers

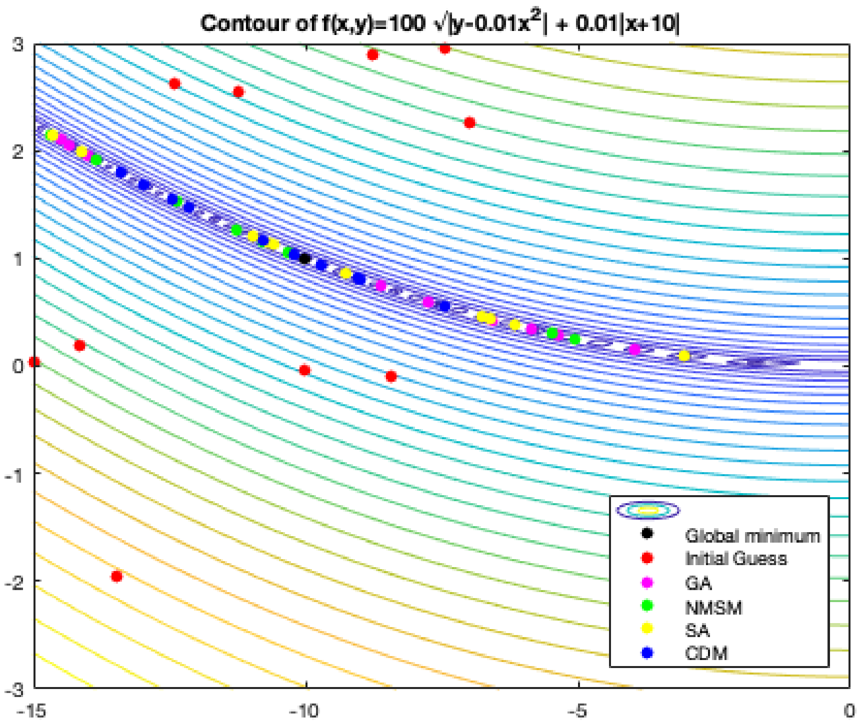


Table 1 and Table 2 shows that the *best* minimizers of NMSM, SA, and CDM had accurately estimate the the global minimizer (-10,1). Although, all four *average* values can be considered as good estimate for the global minimizer of f_2 since they are near it, the obtained average using CDM gave the the least function value.
Observe that f_2 have multiple local minima all throughout the $x \in [-15, 5]$ region and the global minimizer is located somewhere in between (see Figure 2). Hence, it is easy for any algorithm to get stuck at a local minima unless of course if the initial point is sufficiently near the global minima. In our case, however, all the initial points are considerably far from the global minimizer so results from the 10 independent runs of the 4 algorithm were somewhat varied. As a matter of fact, none of the obtained minimizer converge at the global minimizer but there were few *(the best values)* that converges somewhere near it.

The best value of NMSM is (-10.31, 1.06) has an initial point (-6.9934, 2.2560); the best value of SA is (-10.59, 1.12) has an initial point (-13.4679, -1.9568); the best value of CDM is (-10.21, 1.04) has an initial point (-13.4679,-1.9568). This initial points were not the closest at the global minimizer but they served as good initial points because the algorithms were able to estimate the global minimizer. This just shows that eventhough the initial point is near the global minimizer like (-10.0266, -0.0387), this doesn't necessarily mean that the algorithm will converge at the global minimum especially if there were numerous local minima around it.
Therefore, none of the algorithms really outperform the others. All four algorithm didn't get the global minimizer and most of its minimizers from the independent runs were stuck at the local minima. But having said that, their best values could serve as good estimators to the global minimizer.

1.3 Third function

$$f_3(x, y) = x\sin(4x) + 1.1\sin(2y)$$

Observe that f_3 has many local minima and all of the initial guess is located near these local minima.

Clearly, all the *best* values is close at the global minimizer (see Table 1 and 2). Among the four algorithms, the GA and SA have the best *average* minimizer. Apparently, GA performs great at this kind of functions because it doesn't require initial points. Moreover, since mutation is possible in GA it is plausible for the algorithm to go out at the local minima.

Apart from GA, results from the 10 runs of SA were also near the global minimizer. Figure 5 shows that the obtained minimizers using SA flocks near the global minima and at one local minima.

Unlike GA and SA, the average of NMSM and CDM is far from the global minimizer. As I've mentioned above, it is easy for any algorithm to get stuck at the local minima if the initial point is near a local minima. This is clearly what happened in this case. The initial points are located near the local minima so the NMSM and CDM algorithm mostly got stuck at the local minima nearest to it. It is evident from Figure 5 and 6 that the obtained minimizers were located at the minima near its initial points.

Therefore, for functions like f_3 GA and SA outperform the CDM and NMSM in estimating the global minimizer.

FIGURE 3. Contour of f_3 and the scatter plot of the minimizers of GA

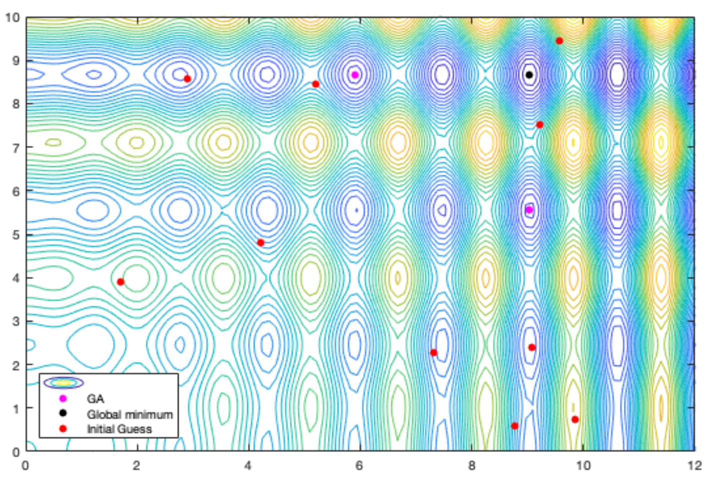


FIGURE 4. Contour of f_3 and the scatter plot of the minimizers of NMSM

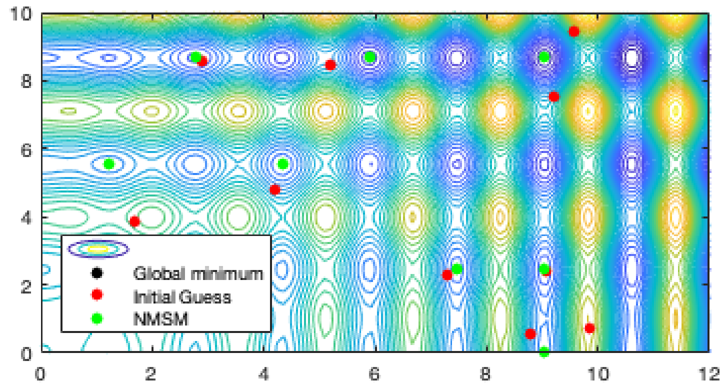


FIGURE 5. Contour of f_3 and the scatter plot of the minimizers of SA

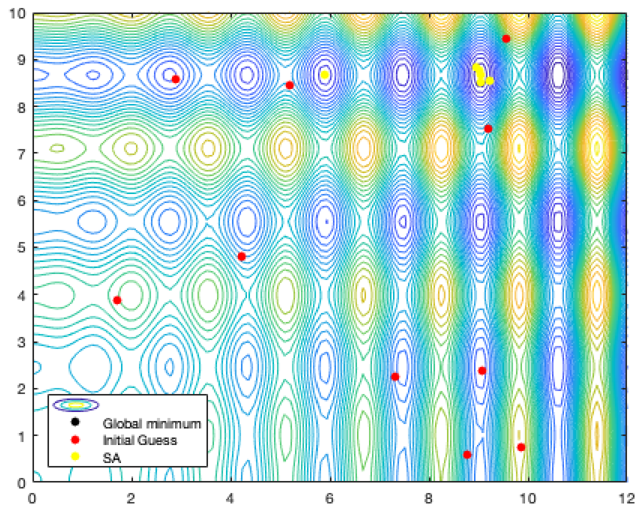
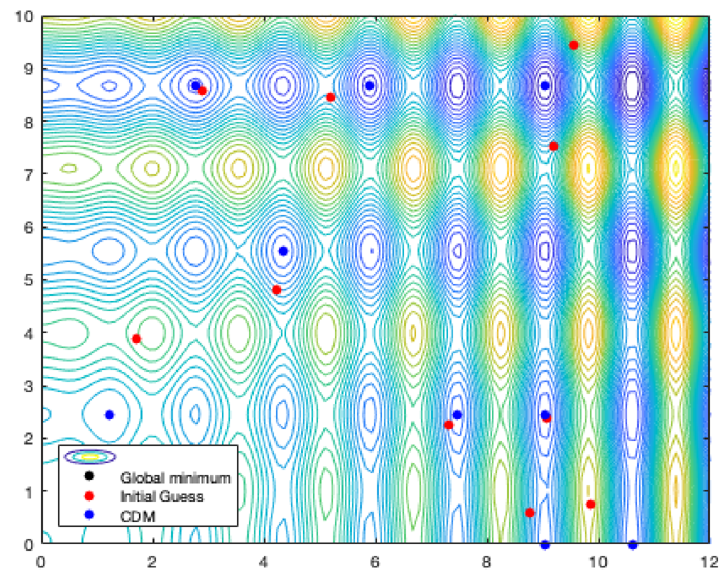


FIGURE 6. Contour of f_3 and the scatter plot of the minimizers of CDM



2. Let

$$A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 3 \end{bmatrix}$$

To find the Cholesky factorization of A we need to minimize

$$f(x) = ||A - LL^T||_F^2$$

where $x \in \mathbb{R}^6$, $x_i \in [-4, 4]$ for all $i = 1, 2, ...6$ and $L = \begin{bmatrix} x(1) & 0 & 0 \\ x(2) & x(3) & 0 \\ x(4) & x(5) & x(6) \end{bmatrix}$.

Using Genetic Algorithm, we'll get

$$L = \begin{bmatrix} -1.4140 & 0 & 0 \\ 1.4147 & -1.4137 & 0 \\ -0.0001 & -1.4150 & -0.9987 \end{bmatrix}, \quad \text{and} \quad L^T = \begin{bmatrix} -1.4140 & 1.4147 & -0.0001 \\ 0 & -1.4137 & -1.4150 \\ 0 & 0 & -0.9987 \end{bmatrix}$$

3. The amount of pressure required to sink a large, heavy object in a soft homogeneous soil that lies above a hard base soil can be predicted by the amount of pressure required to sink smaller objects in the same soil. Specifically, the amount of pressure p required to sink a circular plate of radius r a distance d in the soft soil, where the hard base soil lies a distance $D > d$ below the surface, can be approximated by an equation of the form

$$p = k_1 e^{k2r} + k_3 r$$

where k_1, k_2, k_3 were all constant, with $k_2 > 0$ and $k_3 < 0$ depending on d and the consistency of the soil but not on the radius of the plate.

3.a) To solve for k_1, k_2, k_3 we need to minimize the following minimization problem

$$f(k) = \min \ \frac{1}{2}||p - k_1 e^{k_2 r} - k_3 r||^2 + \mu(\max(0, -k_2) + \max(0, k_3))$$

where $k \in \mathbb{R}^3$ and $\mu = 1000$. Using CDM, the estimated values of k_1, k_2, k_3 is 8.7713, 0.2597, and -1.3723, respectively.

3.b) To solve for the minimum size of circular plate that would be required to sustain a load of 500 (lb/in^2) we need to minimize

$$g(r) = \frac{1}{2}||500 - k_1 e^{k_2 r} - k_3 r||^2$$

where $k_1 = 8.7713$, $k_2 = 0.2597$, and $k_3 = -1.3723$. Again, using CDM, we get $r = 15.7315$.

4) Estimating the values of $s = [k_1,\ k_2,\ k_3]$ using Simulated Annealing we get $k_1 = 0.0020,\ k_2 = 0.0006,\ k_3 = 0.4416$. Performing the Nelder Mead Simplex Method using the value of s as its starting point to further improve the estimation we get

$$k_1 = 0.0020,\ k_2 = 0.0006,\ k_3 = 0.5072$$

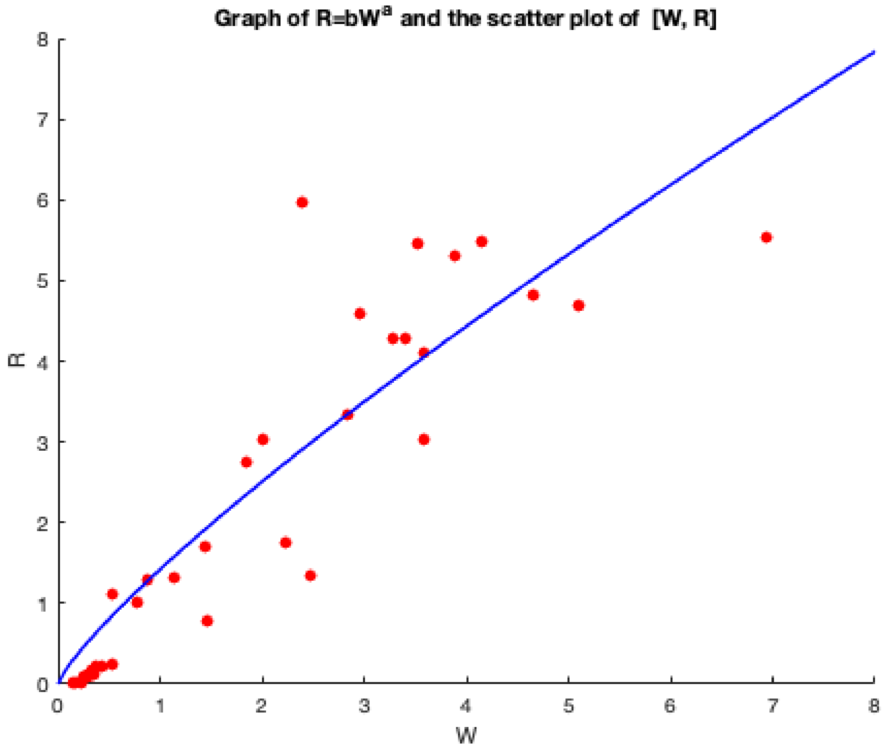
5) Given the sphinx moth data, you are tasked to determine the relationship between W, the live weight of the sphinx moth larvae in grams, and R, the oxygen consumption of the larvae in ml/hour. For biological reasons, it is assumed that a relationship is in the form

$$R = bW^a$$

exists between W and R where a and b are unknown parameters. Calculating the values of a and b by minimizing

$$\sum_{i=1}^{37} (R_i - bW_i^a)^2$$

using Simulated Annealing we get $a = 0.8230$ and $b = 1.4159$. The graph of the $R = bW^a$ using the calculated a and b and the scatter plot of $[WR]$ was shown in **??**


Graph of R=bW^a and the scatter plot of [W, R]
R
W
0
1
2
3
4
5
6
7
8