Numerical Optimization

Assignment 1

Xingrong Zong, tnd179

February 8th, 2022

Introduction

This assignment performs the implementation and testing results of the functions given in the case-studies by using python and online calculator. Also, all the following graphs were plotted under the conditions with 2-dimension and $-1 \leq x \leq 1$

Programming Exercises

The Ellipsoid Function

$f_1(x) = \sum_{i=1}^{d} \alpha^{\frac{i-1}{d-1}} x_i^2, \alpha = 1000$
$\nabla f_1(x) = 2\alpha^{\frac{i-1}{d-1}} x_i$

$$\nabla^2 f_1(x) = \begin{bmatrix} 2\alpha^{\frac{0}{d-1}} & 0 & \cdots & 0 \\ 0 & 2\alpha^{\frac{1}{d-1}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2\alpha^{\frac{d-1}{d-1}} \end{bmatrix}$$

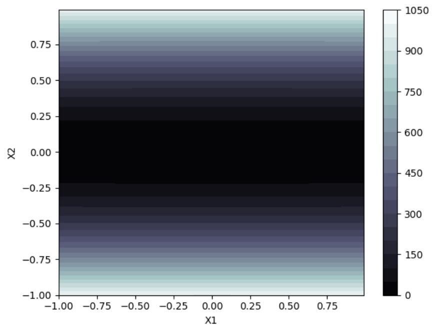

Figure 1: $f_1(x)$ in 2D

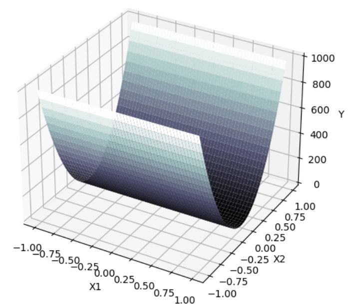

Figure 2: $f_1(x)$ in 3D

The Rosenbrock Banana Function

$f_2(x) = (1 - x_1)^2 + 100 \cdot (x_2 - x_1^2)^2$

$\nabla f_2(x) = (2x_1 - 2 + 400x_1^3 - 400x_1x_2, 200x_2 - 200x_1^2)$

$$\nabla^2 f_2(x) = \begin{bmatrix} 2 + 1200x_1^2 - 400x_2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

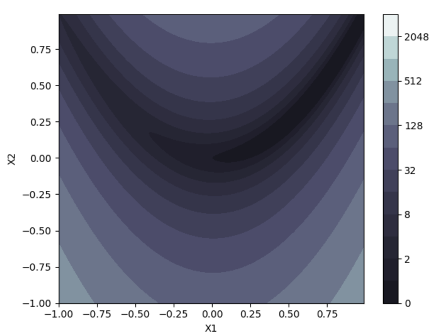

Figure 3: $f_2(x)$ in 2D

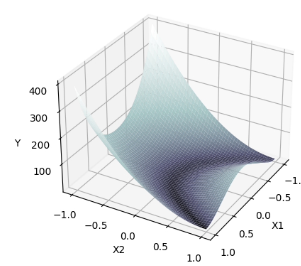

Figure 4: $f_2(x)$ in 3D

The Log-Ellipsoid Function

$f_3(x) = \log(\epsilon + f_1(x)), \epsilon = 10^{-4}$

$\nabla f_3(x) = \frac{2\alpha^{\frac{i-1}{d-1}} x_i}{\epsilon + f_1(x)}$

$$\nabla^2 f_3(x) = \begin{bmatrix} \frac{2\alpha^{\frac{0}{d-1}}(\epsilon+f_1(x))-(2\alpha^{\frac{0}{d-1}}x_1)^2}{(\epsilon+f_1(x))^2} & \frac{-4\alpha^{\frac{1}{d-1}}x_1x_2}{(\epsilon+f_1(x))^2} & \cdots & \frac{-4\alpha^{\frac{d-1}{d-1}}x_1x_d}{(\epsilon+f_1(x))^2} \\ \frac{-4\alpha^{\frac{1}{d-1}}x_2x_1}{(\epsilon+f_1(x))^2} & \frac{2\alpha^{\frac{2}{d-1}}(\epsilon+f_1(x))-(2\alpha^{\frac{2}{d-1}}x_2)^2}{(\epsilon+f_1(x))^2} & \cdots & \frac{-4\alpha^{\frac{d}{d-1}}x_2x_d}{(\epsilon+f_1(x))^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-4\alpha^{\frac{d-1}{d-1}}x_dx_1}{(\epsilon+f_1(x))^2} & \frac{-4\alpha^{\frac{d}{d-1}}x_dx_2}{(\epsilon+f_1(x))^2} & & \frac{2\alpha^{\frac{2(d-1)}{d-1}}(\epsilon+f_1(x))-(2\alpha^{\frac{2(d-1)}{d-1}}x_d)^2}{(\epsilon+f_1(x))^2} \end{bmatrix}$$

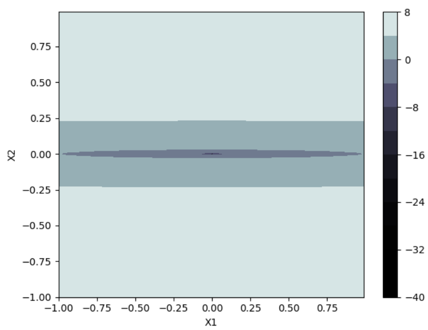

Figure 5: $f_3(x)$ in 2D

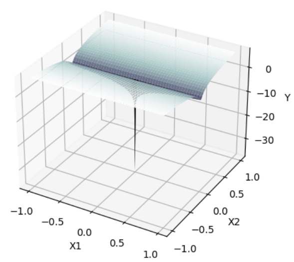

Figure 6: $f_3(x)$ in 3D

The Attractive-Sector Function

$f_4(x) = \sum_{i=1}^{d} h(x_i)^2 + 100 \cdot h(-x_i)^2,$

$h(x) = \frac{\log(1+\exp(q \cdot x))}{q}, q = 10^4$

$\nabla f_4(x) = \frac{\exp(q^2 \cdot x_i)}{1+\exp(q^2 \cdot x_i)} - 100 \cdot \frac{\exp(-q^2 \cdot x_i)}{1+\exp(-q^2 \cdot x_i)}$

$$\nabla^2 f_4(x) = \begin{bmatrix} \frac{101q^2\exp(-q^2 \cdot x_1)}{(1+\exp(-q^2 \cdot x_1))^2} & 0 & \cdots & 0 \\ 0 & \frac{101q^2\exp(-q^2 \cdot x_2)}{(1+\exp(-q^2 \cdot x_2))^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{101q^2\exp(-q^2 \cdot x_d)}{(1+\exp(-q^2 \cdot x_d))^2} \end{bmatrix}$$

The Sum of Different Powers Function

$f_5(x) = \sum_{i=1}^{d} (x_i^2)^{1+\frac{i-1}{d-1}}$

$\nabla f_5(x) = 2(1 + \frac{i-1}{d-1})x_i^{1+2\frac{i-1}{d-1}}$

$\nabla^2 f_5(x) = (2 + 6\frac{i-1}{d-1} + 4(\frac{i-1}{d-1})^2)x_i^{2\frac{i-1}{d-1}}$

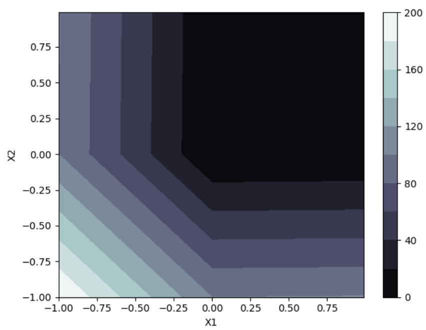

Figure 7: $f_4(x)$ in 2D

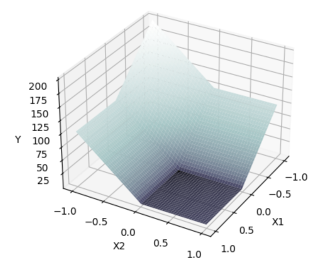

Figure 8: $f_4(x)$ in 3D

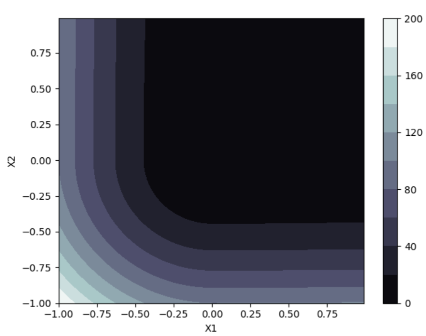

Figure 9: $f_5(x)$ in 2D

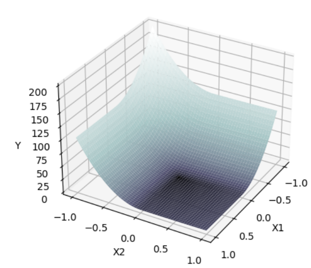

Figure 10: $f_5(x)$ in 3D

Theoretical Exercises

The Hessian matrix H of f is a square $N \times N$ matrix, it is defined and arranged as follows:

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_N} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_N \partial x_1} & \frac{\partial^2 f}{\partial x_N \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_N^2} \end{bmatrix} \text{ or } (H_f)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

$f(x) = \sum_{i=1}^{N} g_i(x_i), x \in^N$

When $i \neq j$:

$\frac{\partial^2 f}{\partial x_i x_j} = \frac{\partial}{\partial x_j}[\frac{\partial}{\partial x_i}[g_1(x_1) + g_2(x_2) + ... + g_N(x_N)]]$

$= \frac{\partial}{\partial x_j}[\frac{\partial}{\partial x_i} g_1(x_1) + \frac{\partial}{\partial x_i} g_2(x_2) + ... + \frac{\partial}{\partial x_i} g_N(x_N)]$

$= \frac{\partial}{\partial x_j}[0 + 0 + \cdots + g_i^{'}(x_i)] = 0$

When $i = j$:

$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i}[\frac{\partial}{\partial x_i}[g_1(x_1) + g_2(x_2) + ... + g_N(x_N)]]$

$= \frac{\partial}{\partial x_i}[\frac{\partial}{\partial x_i} g_1(x_1) + \frac{\partial}{\partial x_i} g_2(x_2) + ... + \frac{\partial}{\partial x_i} g_N(x_N)]$

$= \frac{\partial}{\partial x_i}[0 + 0 + \cdots + g_i^{'}(x_i)] = g_i^{''}(x_i)$

$$H_f = \nabla^2 f(x) = \begin{bmatrix} g_1^{''}(x_1) & 0 & \ldots & 0 \\ 0 & g_2^{''}(x_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & g_N^{''}(x_N) \end{bmatrix}$$

Therefore, $(\nabla^2 f(x))_{ii} = g_i^{''}(x_i)$.

Conclusion

From this assignment, I gained a brief view on the goal of this course based on how to use python to implement different algorithms in order to get a visualized and further understanding behind the algorithms.