THE UNIVERSITY OF HONGKONG
DEPARTMENT OF PHYSICS

Assignment 2

Course: *Machine Learning in Physics (PHYS3151)* – Prof. *Zi Yang Meng*
Tutor: *Mr. Tim Lok Chau*
Due date: *3rd March, 2025*

1. Gradient descent, steepest descent and Conjugate gradient methods

   Given

$$A = \begin{pmatrix} 4 & -3 \\ -3 & 7 \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} -1 \\ 8 \end{pmatrix} \qquad c = 2$$

   (a) Use gradient descent method, steepest descent method, as well as conjugate gradient method with initial guess $\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, find the the optimal $\mathbf{x}$ where the quadratic form $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x} + c$ attains its minimum. Plot the iteration paths of the three methods. Stop the iterations when $\delta < \epsilon^2 \times ||r^{(0)}||^2$, where $\epsilon = 0.001$.

   (b) Check that, for the steepest descent method, $\mathbf{r}^{(k)} \cdot \mathbf{r}^{(k-1)} = 0$, i.e. Any two consecutive search lines are orthogonal.

   (c) Compare their performance by comparing the convergence number of iterations and plotting $||\mathbf{r}^{(i)}||$ against iterations, that is the vector norm of the residual $\mathbf{r}^{(i)} = A\mathbf{x}^{(i)} - \mathbf{b}$.

2. Application to solving system of linear equations

   Given

$$\begin{cases} x \qquad - z = 0 \\ \qquad 3y \qquad = 6 \\ -x \qquad + 4z = 3 \end{cases}$$

   (a) Solve the above system of linear equations analytically. Show your solving procedures step by step.

   (b) Implement methods in question 1 to solve the above system of linear equations. You should choose suitable A (**a positive semi-definite and symmetric matrix**), $\mathbf{b}$ and other parameters (e.g. initial conditions, stopping condition ϵ value, etc.). **Set** $c = 0$. Round the answer to 3 decimal places and compare the answer to that of part a).

(c) Plot the three paths of iterations in part b).
The function we are trying to minimize is $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - b^T\mathbf{x} + c$ and in the plot below, each surfaces is $f(\mathbf{x}) = constant$. At the end of your optimization, you should see $f(\mathbf{x}) = C_m$ and C_m is the minimum constant that you can arrive at. $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and three axes are shown in the graph below.

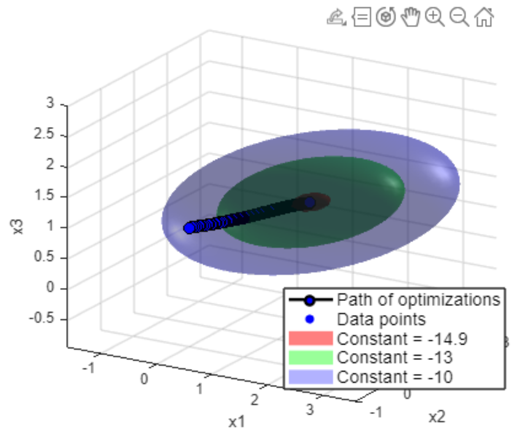

Figure 1: The result is obtained by the gradient descent method.

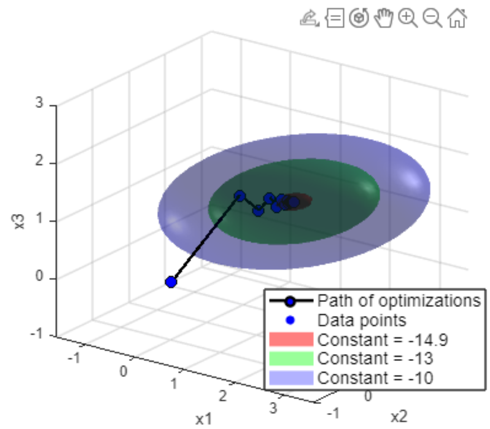

Figure 2: The result is obtained by the steepest descent method.

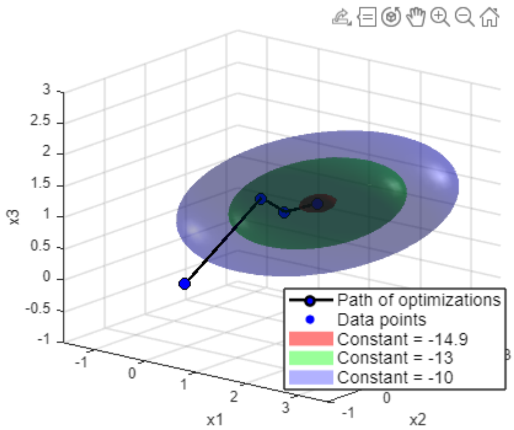

Figure 3: The result is obtained by the conjugate gradient method.