ID2090 ASSIGNMENT-4 REPORT

SHWETHA V ME23B073

Question-1 Matrix Operations

I started off by searching up the terms that were new to me.

- YAML (YAML Ain't Markup Language :P): Every object occupies a particular location in memory. Data serialization is the process of converting an object into a stream of bytes to more easily save or transmit it. YAML, JSON, XML etc are all examples of data serialization languages often used for writing configuration files. Due to YAML's syntax, it's considered easier to read, comprehend, and write.
- Spectral theorem: found a yt video and paper on it as well.
- Learnt about different special types of matrices and the operations on them here.

I started by figuring out which library to use, sage or numpy or sympy, and decided to go with sympy as i found a really easy to understand documentation of various built in functions to perform the matrix operations asked in the assignment questions.

I first executed the code on Google Colab, and after i implemented each of the functionalities and got it working, wrote it as a shell script.

Figure 1: XML vs JSON vs YAML.

Question-2 Optimisation

Objective

The primary objective of this task is to find the optimal parameters (a, b, c) of the plane equation $ax + by + cz = 1$ such that it minimizes the sum of squares of distances from the plane to the given set of points. This optimization problem can be formulated as follows:

$$\text{Minimize } L = \sum_{i=1}^{n}(ax_i + by_i + cz_i - 1)^2 \tag{1}$$

where n represents the number of points, and (x_i, y_i, z_i) denote the coordinates of the i-th point.

Methodology

To address this optimization problem, we employ Newton's steepest descent method, a variant of Newton's method used to find the local minimum of a function. The methodology involves several key steps:

- Objective function: L quantifies discrepancies between the plane and the points.
- Gradient and Hessian: To guide the optimization process, we compute the gradient and Hessian matrix of the objective function with respect to the parameters a, b, and c. The gradient provides information about the direction of steepest ascent, while the Hessian matrix characterizes the curvature of the objective function.

  $$\text{Gradient:} \quad \nabla L = \left(\frac{\partial L}{\partial a}, \frac{\partial L}{\partial b}, \frac{\partial L}{\partial c}\right) \tag{2}$$

  $$\text{Hessian:} \quad H_{ij} = \frac{\partial^2 L}{\partial \theta_i \partial \theta_j} \tag{3}$$

  where θ_i represents the parameters a, b, or c. These derivatives provide crucial information about the direction of steepest ascent and the curvature of the objective function, respectively.

- Newton's Steepest Descent Method: Newton's steepest descent method is employed to iteratively update the parameters of the plane equation. This method utilizes both first and second-order derivatives of the objective function to efficiently converge to the optimal solution. The parameters are updated using the following rule:

  $$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - (H^{-1})\nabla L \tag{4}$$

  where $\boldsymbol{\theta}$ represents the parameters (a, b, c), H is the Hessian matrix, and ∇L is the gradient of the objective function.

- Convergence Criteria: To assess convergence, we track changes in the objective function or gradient norm. The optimization continues until these changes dip below a predefined threshold, signaling proximity to a local minimum.

Implementation details

Bash script

1. Initialize plane equation parameters (a, b, c) randomly.
2. Read point coordinates from the input file.
3. Define a function to compute plane-point distances.
4. Calculate the error function L, gradient, and Hessian matrices.
5. Update parameters using Newton's steepest descent until convergence.
6. Display optimal parameters and iteration count at convergence.

Results

Upon convergence, we obtain optimal (a, b, c) parameters for the plane equation, best-fitting the given 3D points.

Observations

Throughout optimization, I noted:

- Gradual reduction in the objective function indicating algorithmic convergence.
- Sensitivity of convergence(no. of iterations) to initial parameter choice.
- The change in the objective function provided a reliable criterion for assessing convergence. Once the change fell below the predefined threshold, the optimization process was terminated.