name: Sadith Lina Apaza Huayta.

Comparing the Moore-Penrose Pseudoinverse and Gradient Descent for Solving Linear Regression Problems: A Performance Analysis

Introduction:

Linear regression is used to model the relationship between a dependent variable and one or more independent variables. The objetive is to find the coefficient vector β that minimizes the error between observed and predicted values. The general model is expressed as: $y = X\beta + \varepsilon$ where y represents the observed values, X is the feature matrix, β is the vector of unknown parameters, and ε corresponds to random error. The cost funtion to minimize is:

$$S(\beta) = \|X\beta - y\|^2$$

Background and Theoretical Framework

Linear regression models the relationship between a dependent variable y and independent variables X through: $y = X\beta + \varepsilon$ where β are the coefficients estimated by minimizing the sum of squared residuals: $\hat{\beta} = \arg\min_{\beta} \|X\beta - y\|^2$ The Moore-Penrose pseudoinverse gives an exact analytical solution: $\hat{\beta}_{pinv} = (X^T X)^{-1} X^T y$ or, in general, $\hat{\beta}_{pinv} = X^+ y$. It is precise but computationally expensive for large datasets or ill-conditioned matrices. Gradient Descent, on the hand, is an iterative optimization algorithm that updates the parameters as: $\beta^{(t+1)} = \beta^{(t)} - 2\alpha X^T(X\beta^{(t)} - y)$ where α is the learning rate. It is more scalable but may converge slowly or inaccurately if the data is poorly conditioned.

Methodology.

For synthetic data, feature matrices X were generated with controlled condition numbers to simulate well and ill conditioned cases. The target variable y followed the model $y = X\beta^* + \varepsilon$, allowing the true solution to be known. Both methods were appliend: the pseudoinvere was computed using numpy.linalg.pinv while gradient descent used a fixed learning rate $\alpha = 0{,}01$ and stopped when changes in parameters become very small or after 10,000 iterations. Perfoimace was evaluated using three metrics: Runtimee, Mean Squared Error, Iterations to convergence.

Analysis of Results: The results show that the Moore Penrose pseudoinverse is consistently faster and more accurate than Gradient descent, especially for well and moderately conditioned data.