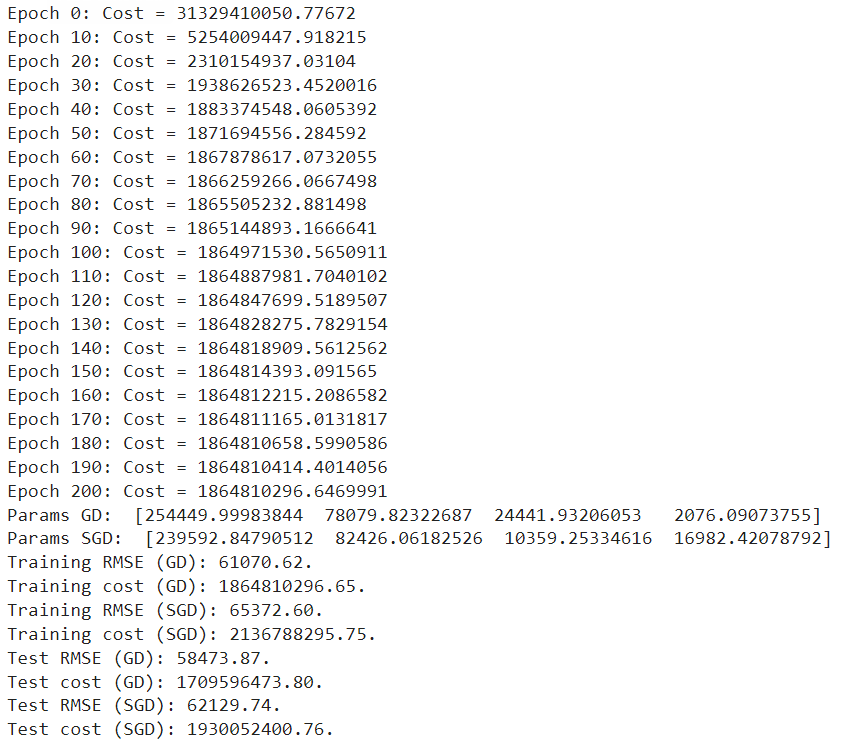
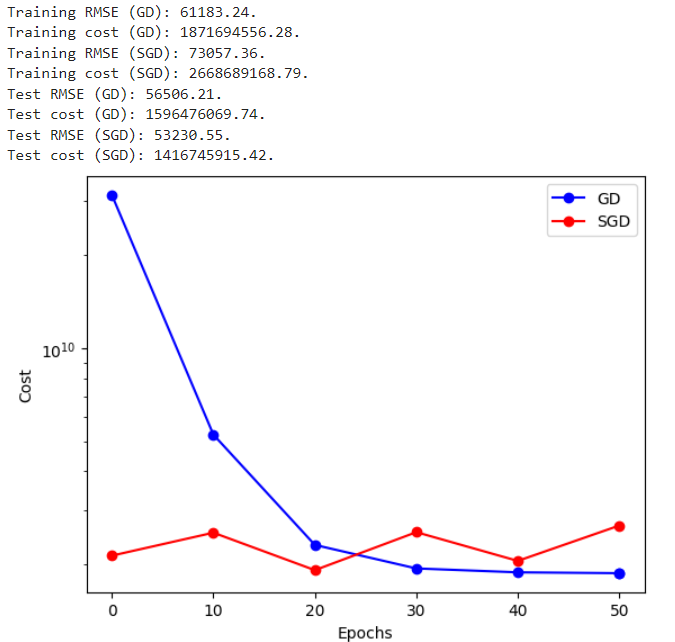
**Results:**



A graph with red and blue lines

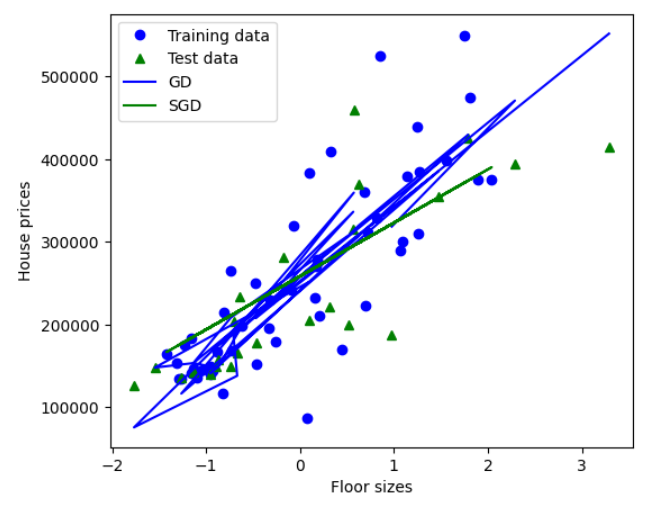
Description automatically generatedA diagram of a training data

Description automatically generated

A diagram of a training data

Description automatically generated

A graph with red and blue lines

Description automatically generated  
  
****

With 200 epochs, both SGD and batch GD have sufficient time to converge to a stable solution for many datasets.

However, for larger datasets or more complex models, 200 epochs might not be enough, especially for SGD, which updates the parameters more frequently but with smaller steps.

Increasing the number of epochs to 300 provides more iterations for both methods to refine their parameters.

This can lead to improved convergence and potentially better performance, especially for datasets with high variance or noise.

With only 50 epochs, the training may not have enough iterations to converge effectively, especially for more complex models or datasets.

SGD might not reach a stable solution within 50 epochs, as it updates parameters more frequently but with smaller steps, requiring more iterations to converge.

Batch GD may show better convergence behavior compared to SGD with a limited number of epochs due to its updates being based on the entire dataset, allowing for more consistent progress towards the optimal solution.

In summary, the number of epochs impacts the convergence and performance of both SGD and batch GD. While fewer epochs may lead to underfitting and insufficient convergence, more epochs can provide better convergence and potentially improve the performance of the trained model. The optimal number of epochs depends on factors such as the dataset size, model complexity, and desired level of convergence.

**Discussions:**

Using data from text files, the given code trains a multiple linear regression model using batch and stochastic gradient descent procedures.

Functionality: The code reads data, calculates mean and standard deviation, trains gradient descent, calculates root mean square error (RMSE), and plots.

It offers flexibility in optimization techniques by covering both batch gradient descent and stochastic gradient descent.

Reading data from text files, standardizing features, training models, assessing performance, and producing plots for display are all done by the main software.

Data handling: The read\_data function is used to read data from text files, which makes it simple to import datasets that are saved in external files.

It is necessary to standardize features in order to guarantee improved convergence during training.

Gradient Descent: The implementation of both stochastic gradient descent (train\_SGD) and batch gradient descent (train) is accurate.

the above formulas are used to compute the gradients with accuracy.

Model Evaluation: To assess how well trained models perform on both training and test data, the method computes RMSE and cost.

This offers information about the predictive accuracy and fit of the models to the data.

graphs: To evaluate the effectiveness of batch gradient descent versus stochastic gradient descent, and to visually represent the training process, the code creates graphs.

Visualizing the learnt model's predictions on both training and test data is done using the floor size vs. house price plot.

Issues with Implementation:

Data File Path: Verify that the train\_file\_path and test\_file\_path file paths correctly point to the Google Drive locations of the dataset files.

Data Formatting: Verify that the data files' layout complies with the requirements of the code (column features, last column labels, etc.).

Hyperparameters: To improve model performance and convergence, modify the hyperparameters (learning rate, number of epochs) as necessary.

Preparing data: Verify that data pretreatment procedures like normalization are suitable for the dataset and do not cause any problems.

Debugging: Use print statements or debugging tools to find and fix problems step-by-step if you run into any errors or unexpected behavior.

With proper implementation and resolution of these possible issues, the code ought to be able to use Google Colab to train a multiple linear regression model on the given dataset.