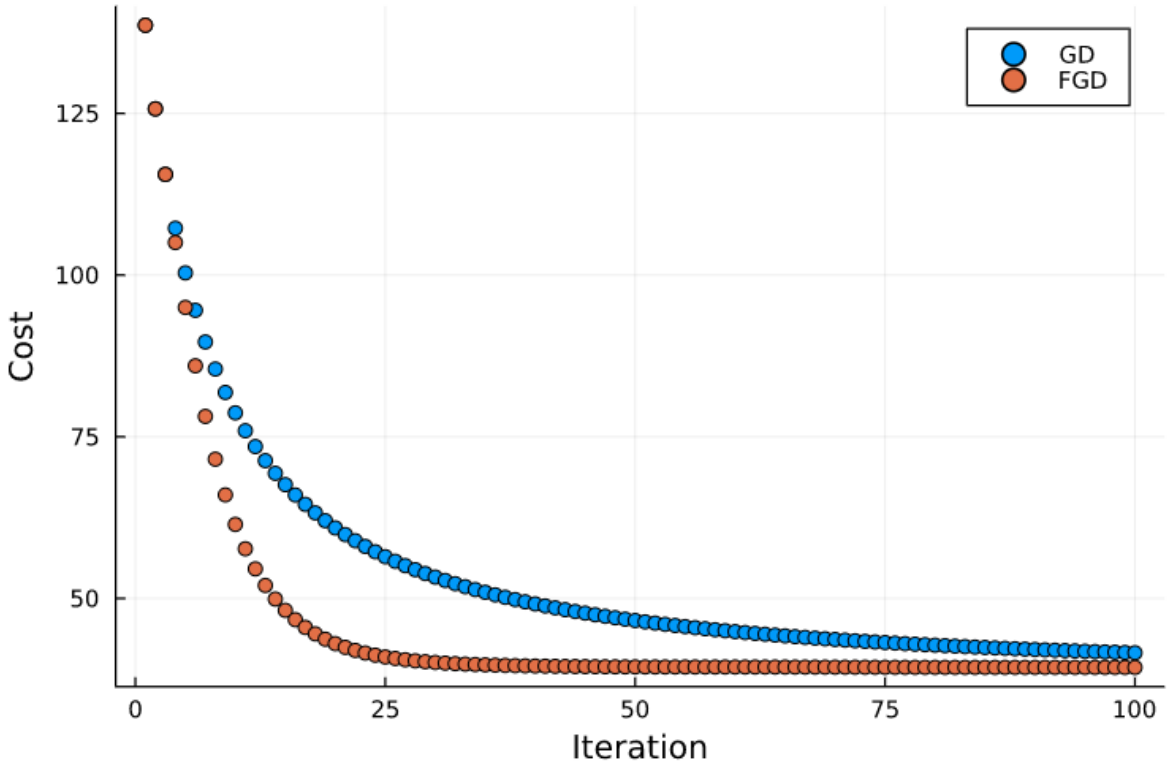
Name:Yuxuan Zhang, Jiayun Ni, Yuang Huang

Part 1: Logistic regression set-up

What is the value of the Lipschitz constant?
23.98

Plot of the cost function versus iteration:



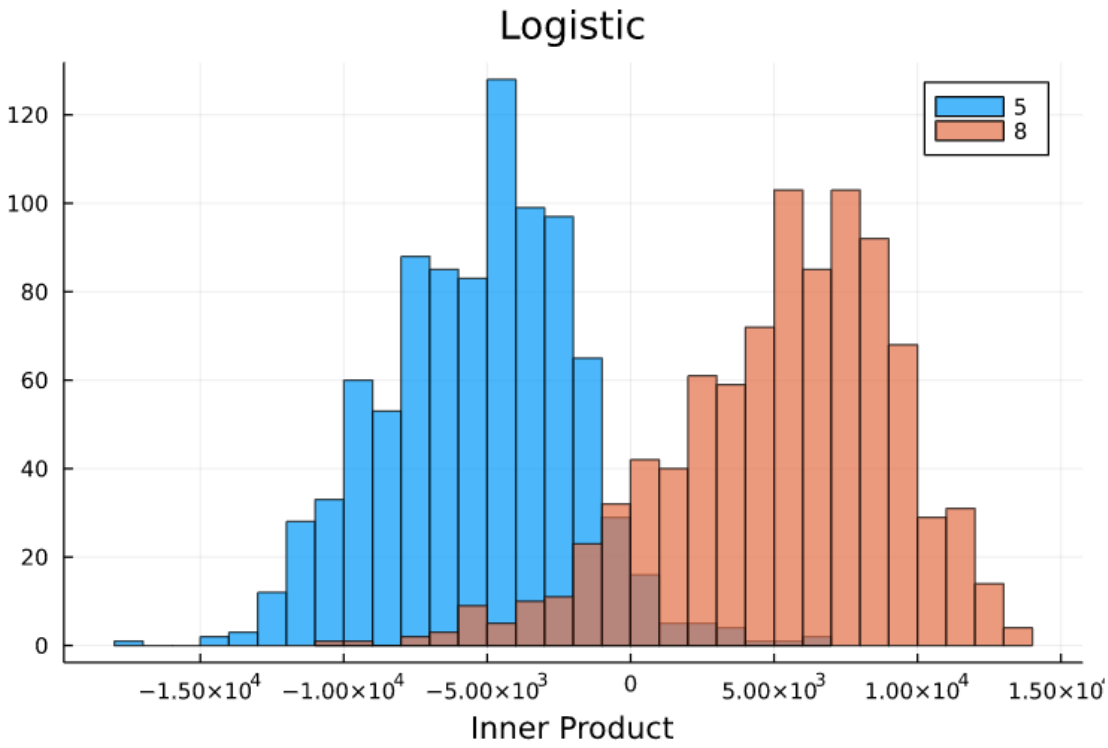
Has the FGD converged in 100 iterations? How do you know? (1 sentence)
FGD converged in 100 iterations, as the value of the cost function remains 39.33 for more than 50 iterations.

Part 2: Test logistic regression

Image of logistic regression weights:

What is the classification accuracy?
Accuracy for 5s: 0.9622
Accuracy for 8s: 0.8922

Histogram of inner products:

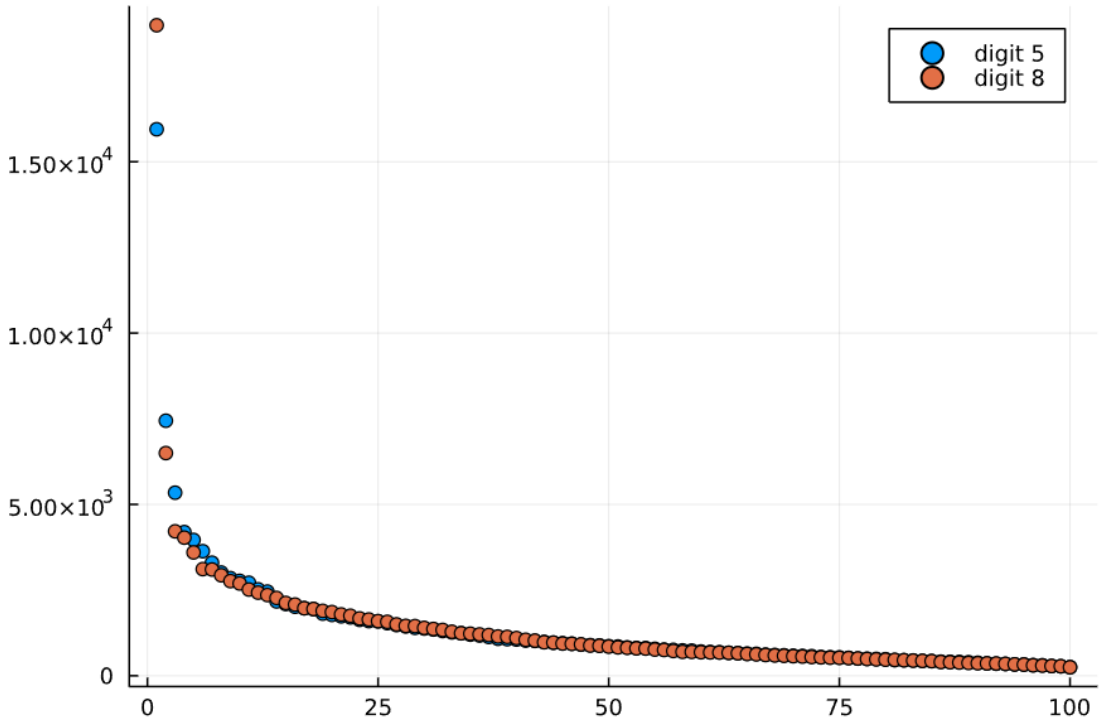


Part 3: Subspace based classification

What is the classification accuracy when using 3 basis vectors?
Accuracy for 5s: 0.9111
Accuracy for 8s: 0.8744

How many basis vectors did you decide to use? Why? (1-2 sentences)
We can see from this figure that when the number of the basic vectors is larger than 20, the two scree plots nearly overlap. So, we tried and found that when r0=20,r1=21, both classification accuracies of digit 5 and 8 reach a relatively high level simultaneously.

```
# Examine the scree plots to choose a rank!
plot(1:ntrain, s0, line=(:dots, :blue), label="digit 5")
plot!(1:ntrain, s1, line=(:dots,:red), label="digit 8")
```

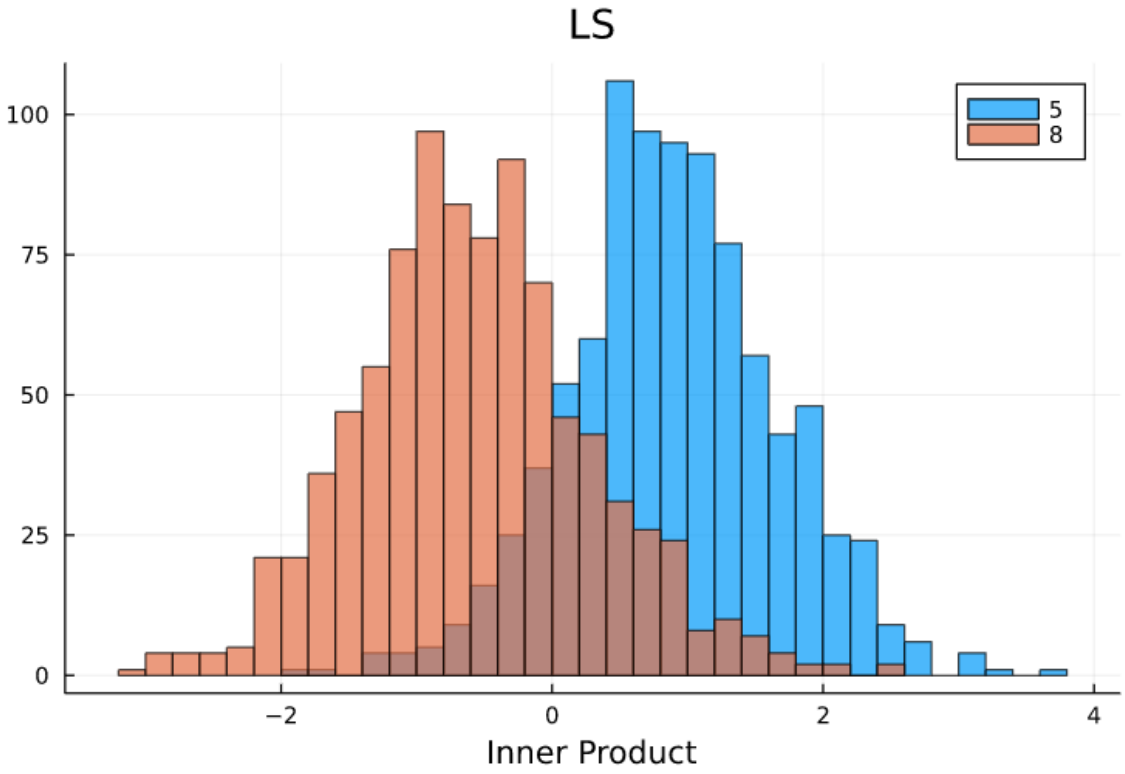


What is the classification accuracy when using your chosen number of basis vectors?
Accuracy for 5s: 0.9533
Accuracy for 8s: 0.9533

Part 4: Least-squares based classification

What is the classification accuracy?
Accuracy for 5s: 0.8867
Accuracy for 8s: 0.7722

Histogram of the inner products:



How does the histogram for the LS-based and logistic-regression based classifiers relate to the classification accuracies? (1-2 sentences)
There is a larger overlap between histograms of digit 5 and 8 when the LS-based classifiers are used, which corresponds to lower accuracies. When it comes to the logistic-regression, the overlap between histograms of digit 5 and 8 is comparatively smaller, corresponding to higher accuracies.

Why does the following block of code give 100% classification accuracy? (1-2 sentences)

```
tmp0 = sum(sign.(reshape(train0, nx*ny, :)' * x_ls) .== 1)
@show tmp0 / ntrain
tmp1 = sum(sign.(reshape(train1, nx*ny, :)' * x_ls) .== -1)
@show tmp1 / ntrain;
```

Because the train0 and train1 are test0 and test1,respectively, and the training converges which leads to 100% accuracy.
Meanwhile, train0 and train1 are almost full rank matrices. The LS-based can return accurate values corresponding to training data.

Part 5: Nearest-neighbor by angle based classification

What is the classification accuracy?
Accuracy for 5s: 0.9289
Accuracy for 8s: 0.9444