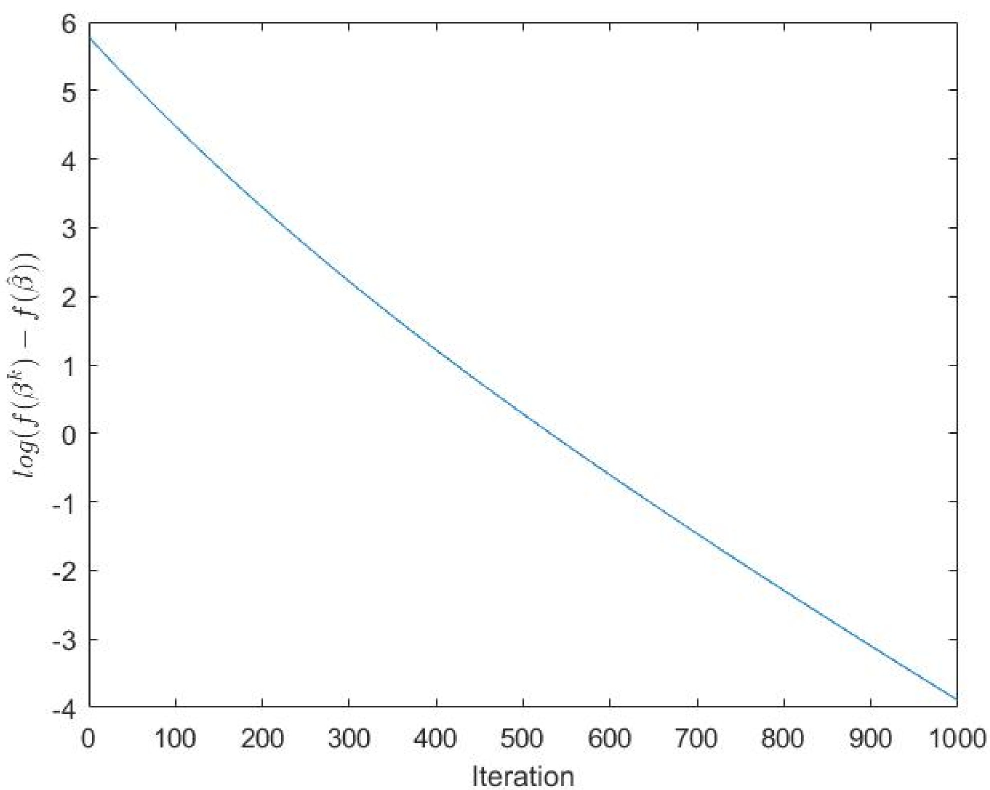
Andrew ElHabr
ISyE 6740
10/17/17

Homework 2

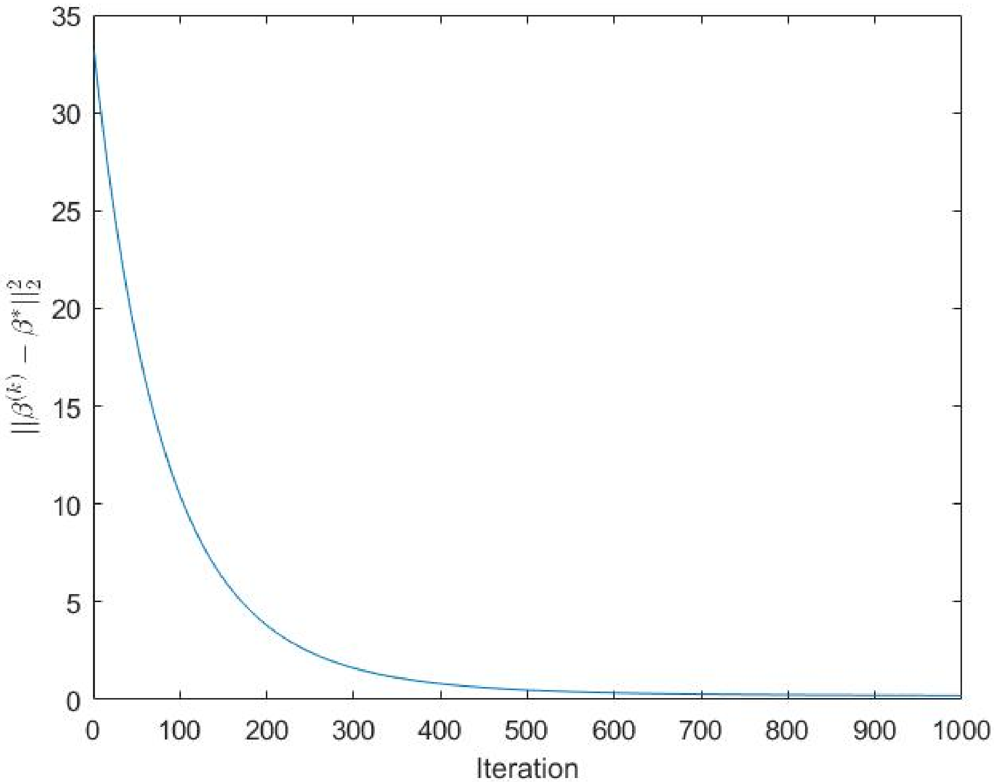
1. (a) The ordinary least squares estimator is implemented in the code as

$$\widehat{\beta} = (X'X)^{-1}X'y.$$

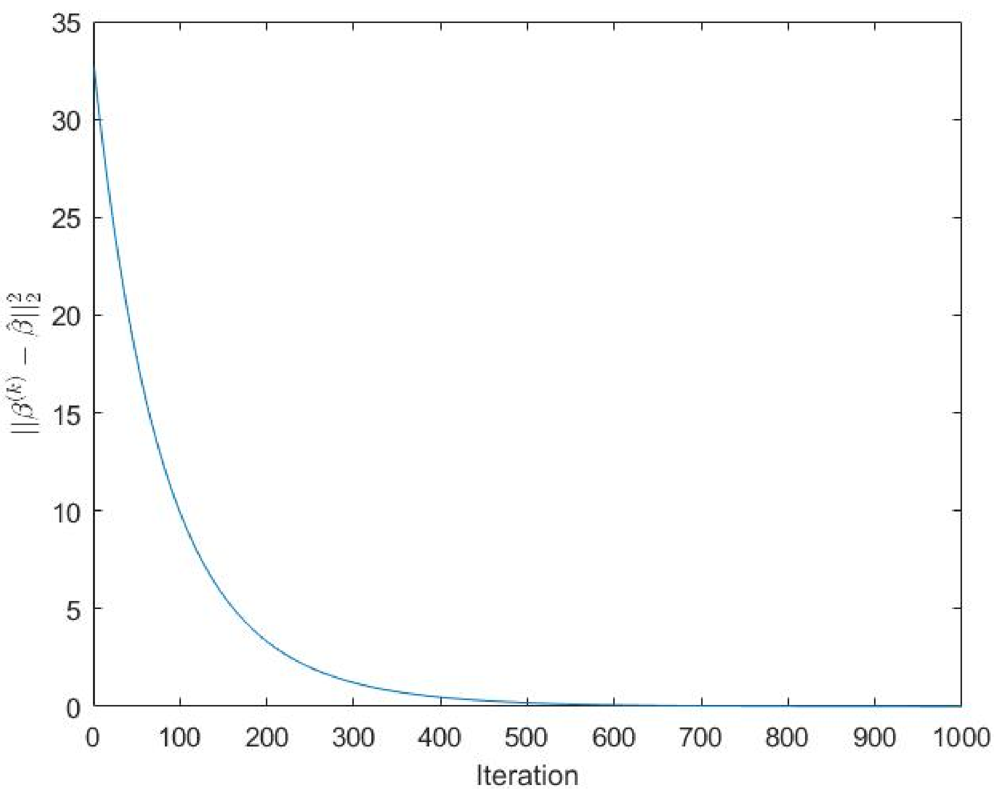
   (b) The squared error is 0.1991.



2. (a)

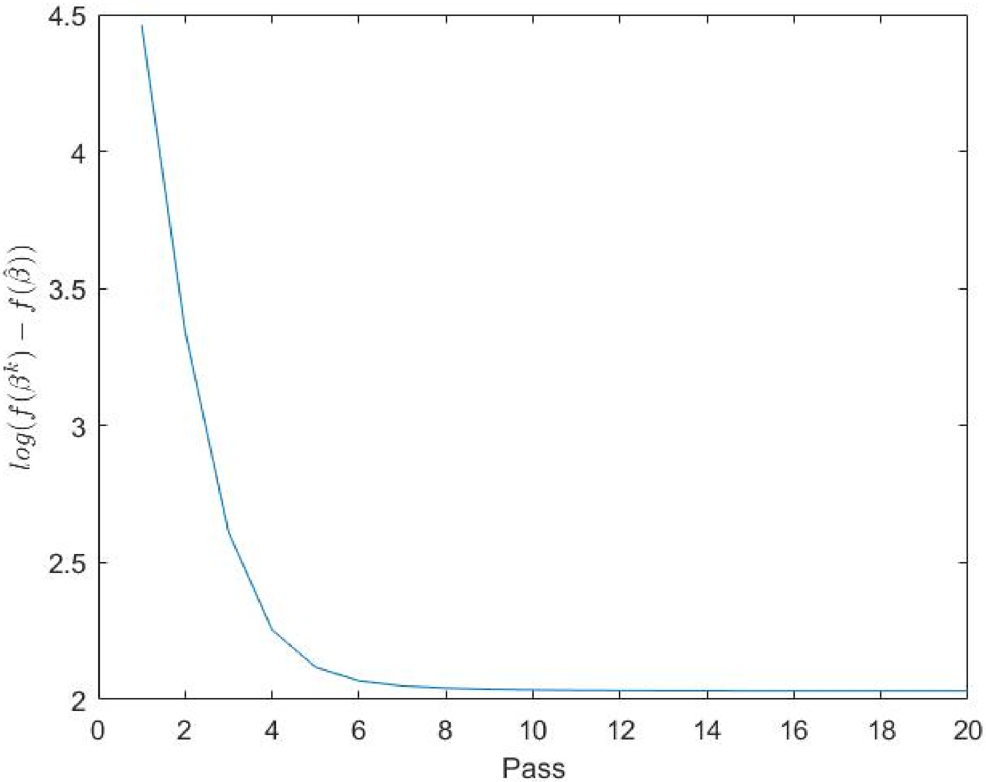

35
30
25
20
15
10
5
0
0
100
200
300
400
500
600
700
800
900
1000
Iteration
$||\beta^{(k)} - \beta^*||_2^2$

(b)

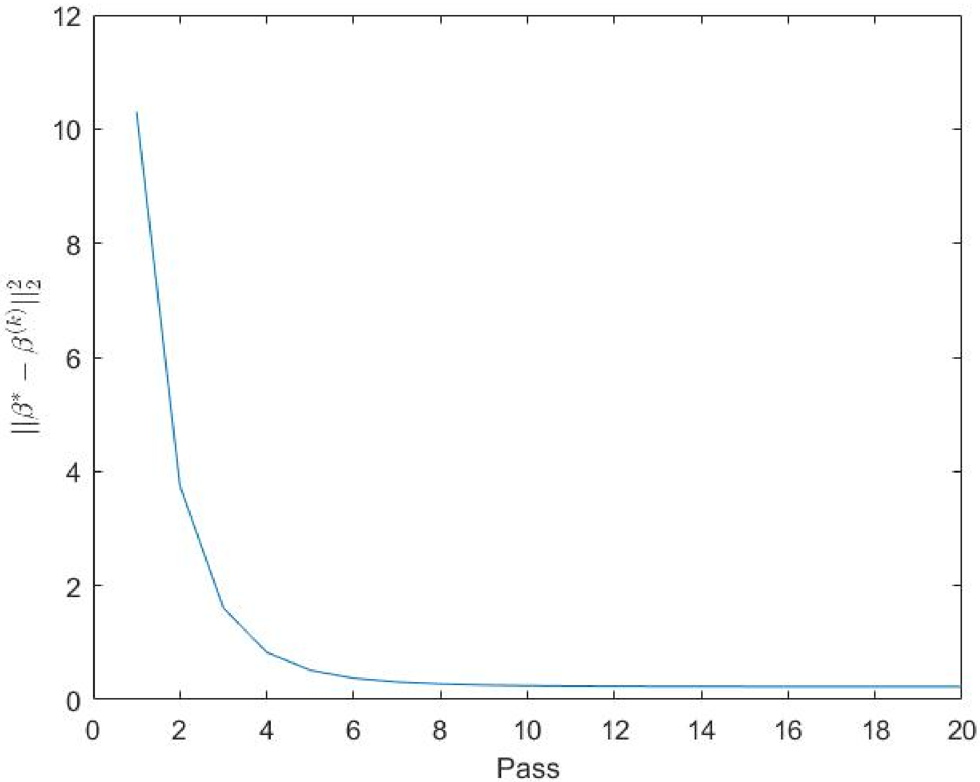

35
30
25
20
15
10
5
0
0
100
200
300
400
500
600
700
800
900
1000
Iteration
$||\beta^{(k)} - \hat{\beta}||_2^2$

(c)

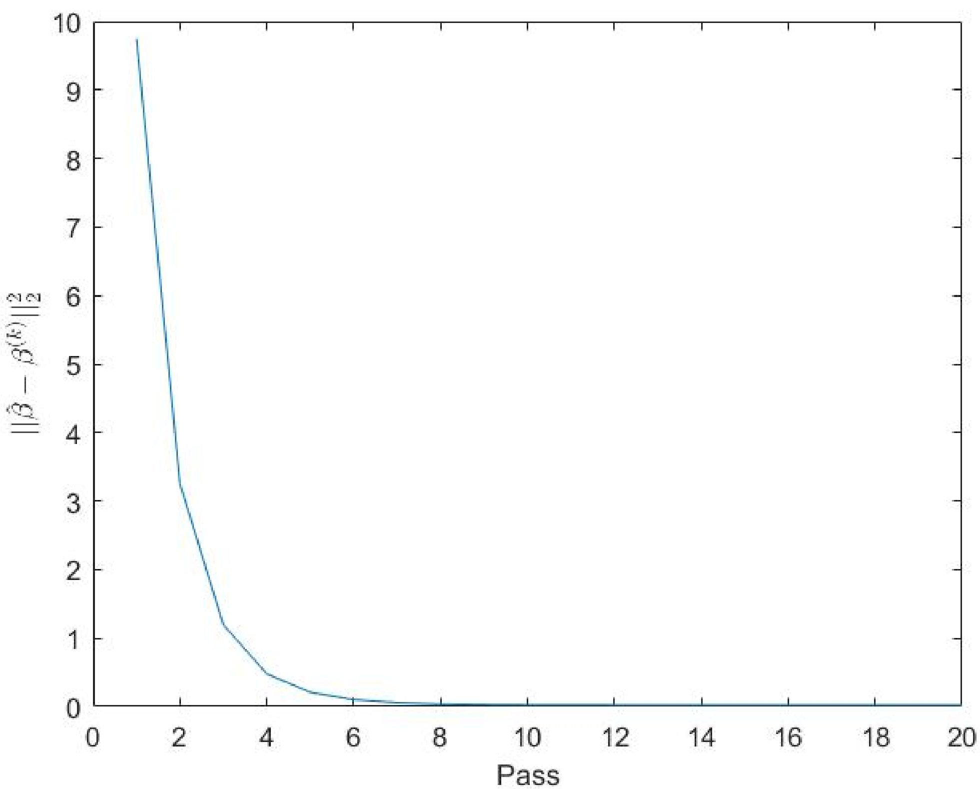
(d) We notice that β^k converges to $\widehat{\beta}$ faster than β^* .



3. (a)

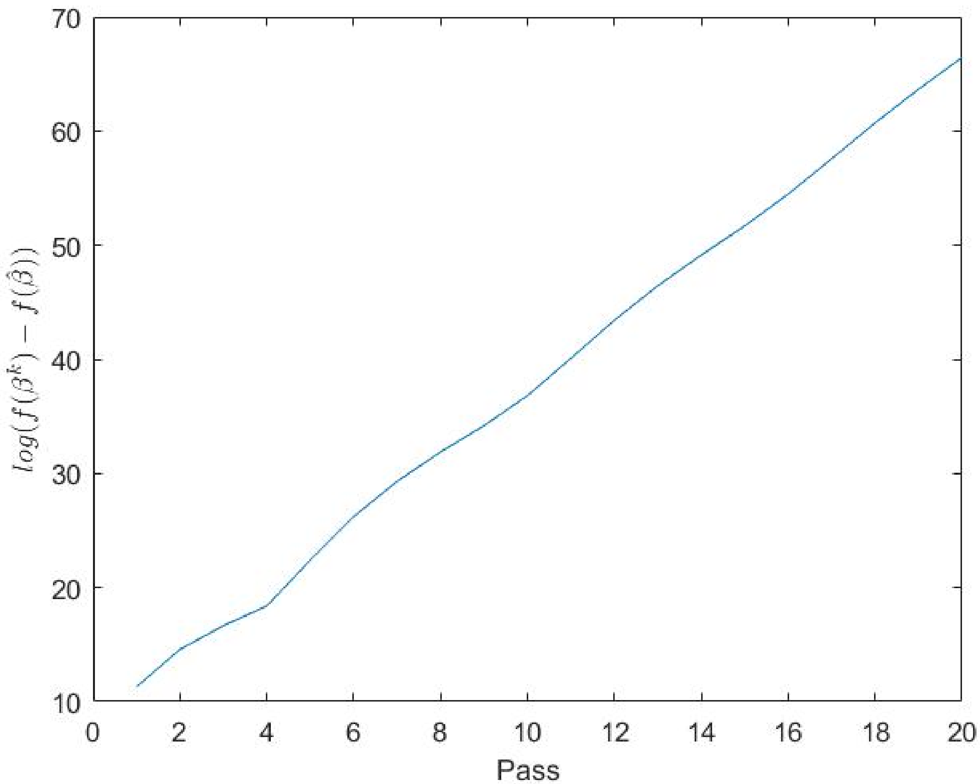

12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16 18 20
Pass
$||\beta^* - \beta^{(k)}||_2^2$

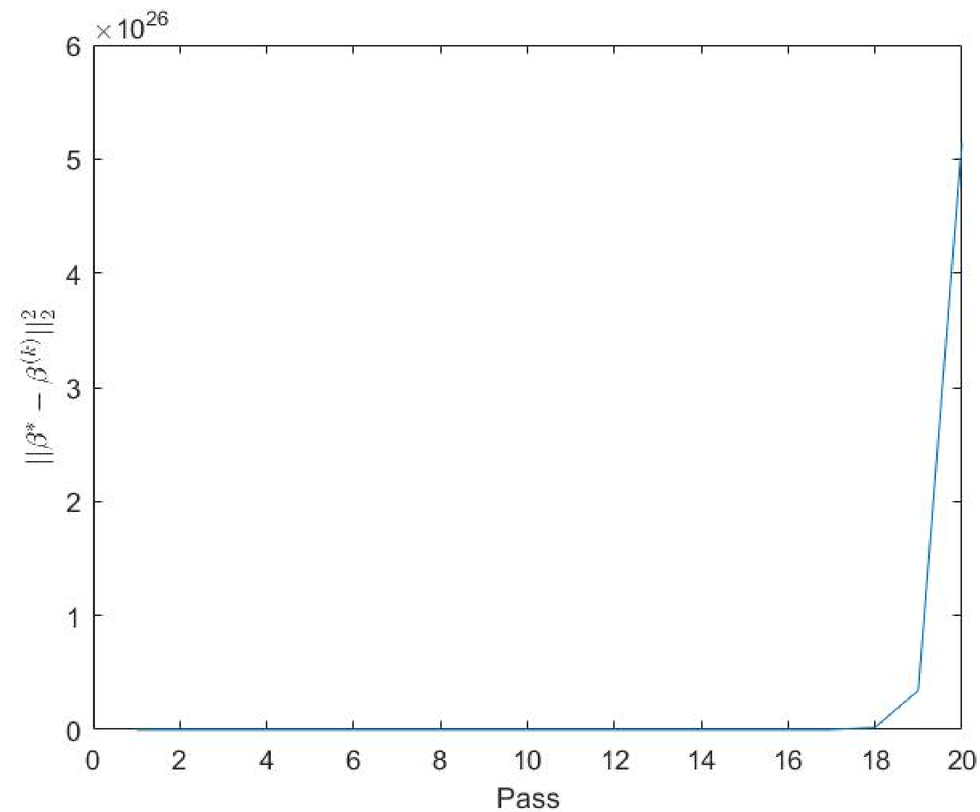
(b)

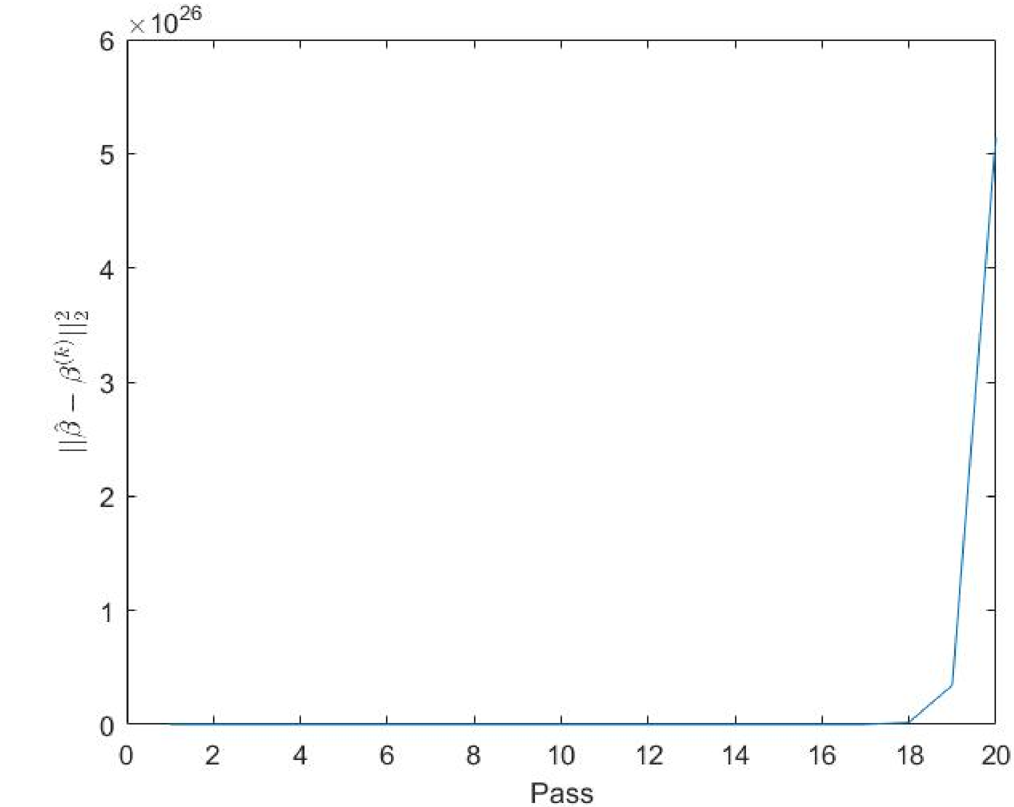

10
9
8
7
6
5
4
3
2
1
0
0 2 4 6 8 10 12 14 16 18 20
Pass
$||\hat{\beta} - \beta^{(k)}||_2^2$

(c)

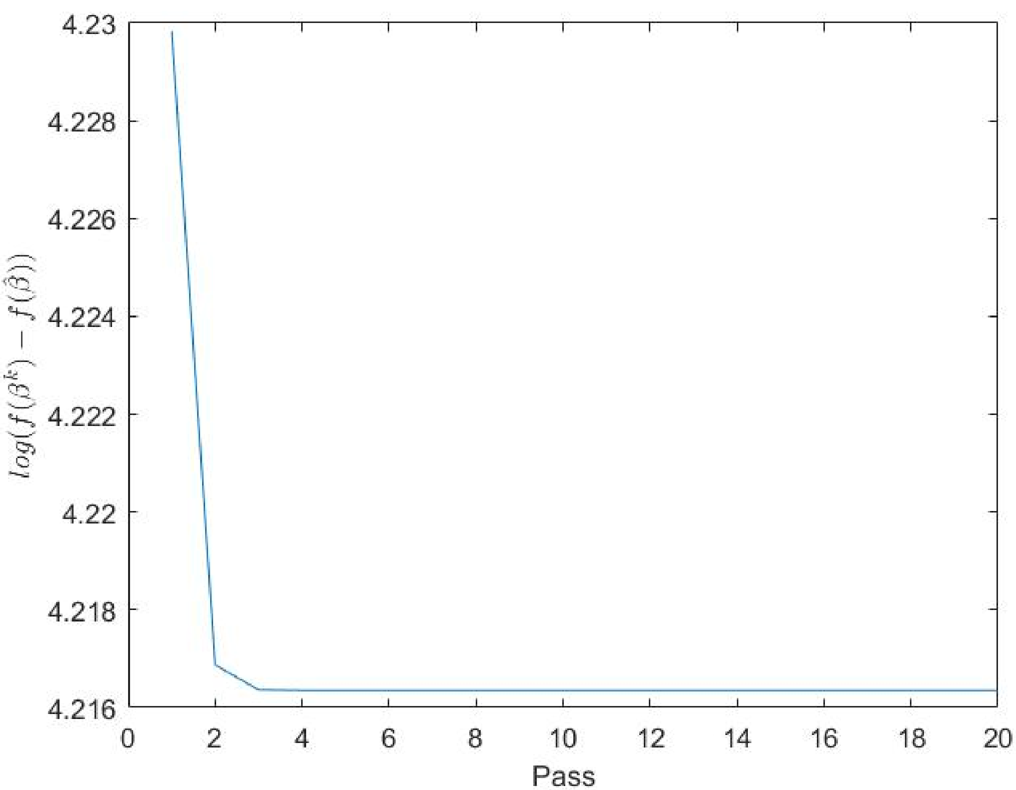
(d) For a step size of $1.7/L$, we have the following figures:

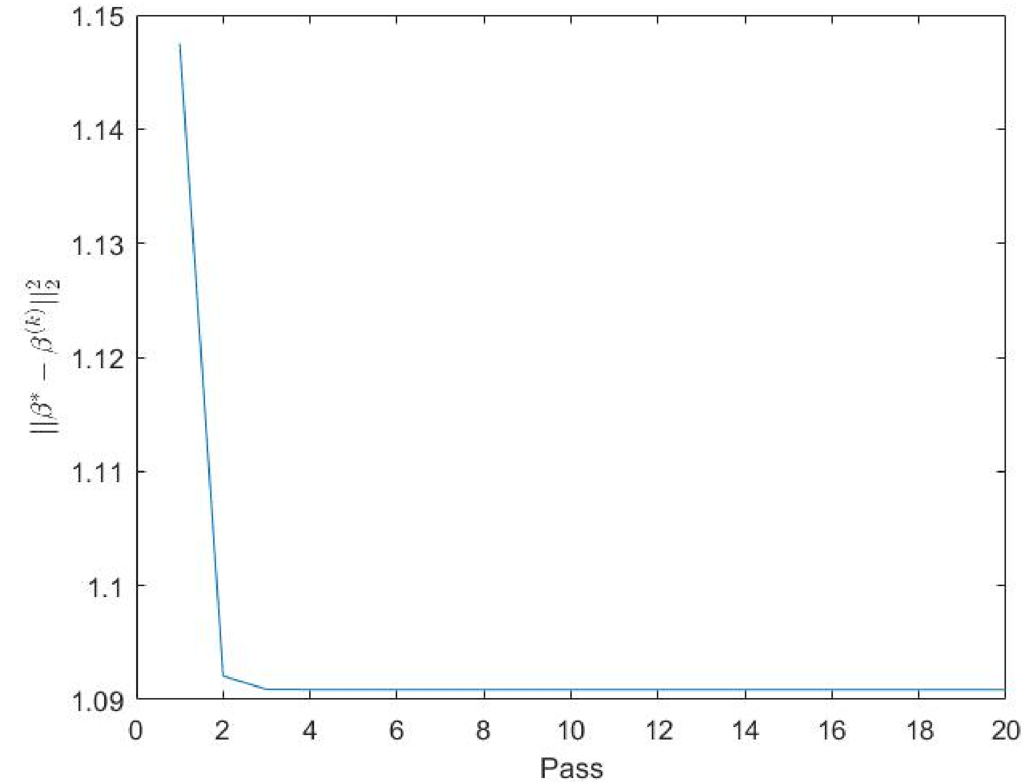


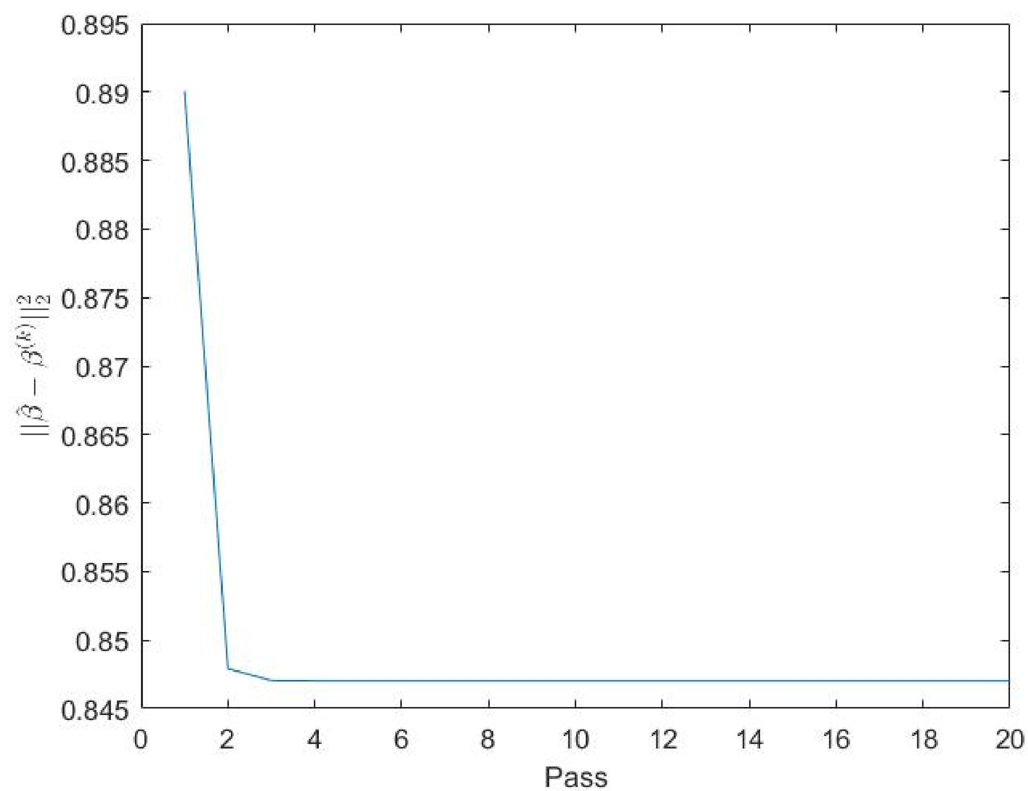




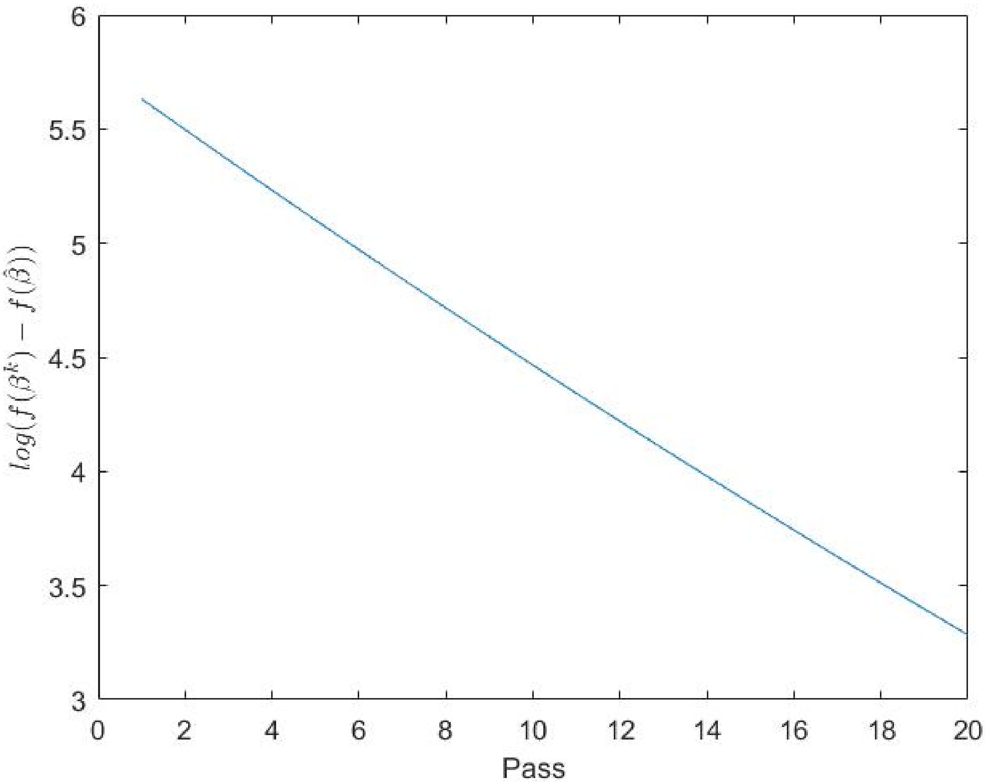
For a step size of $1/L$, we have the following figures:

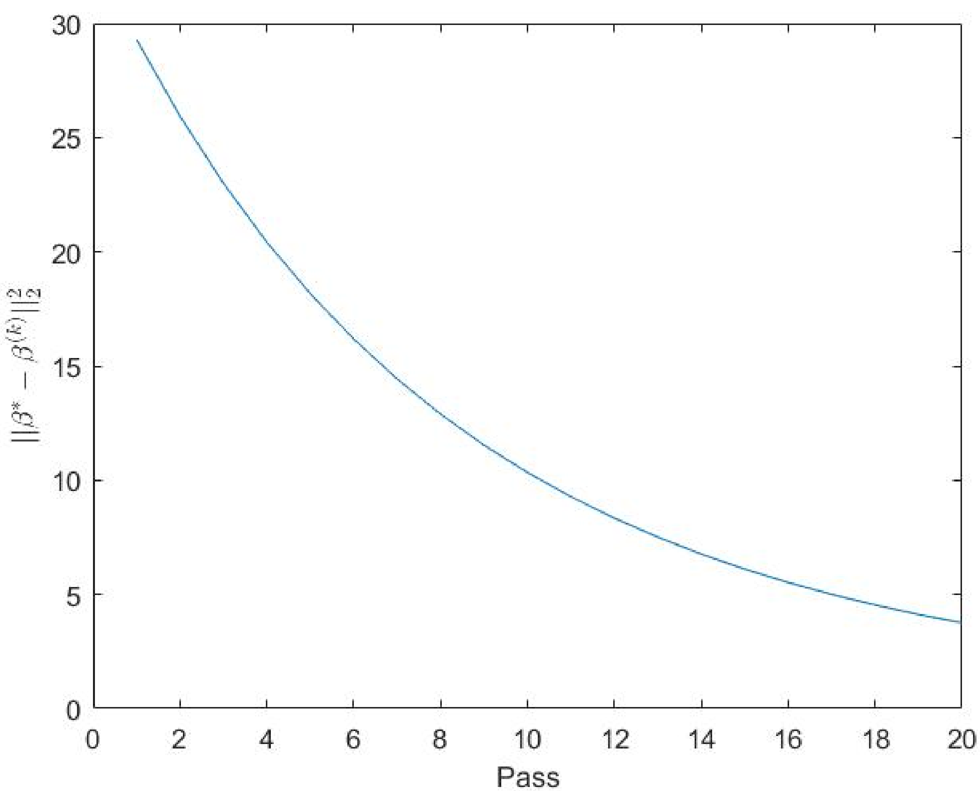


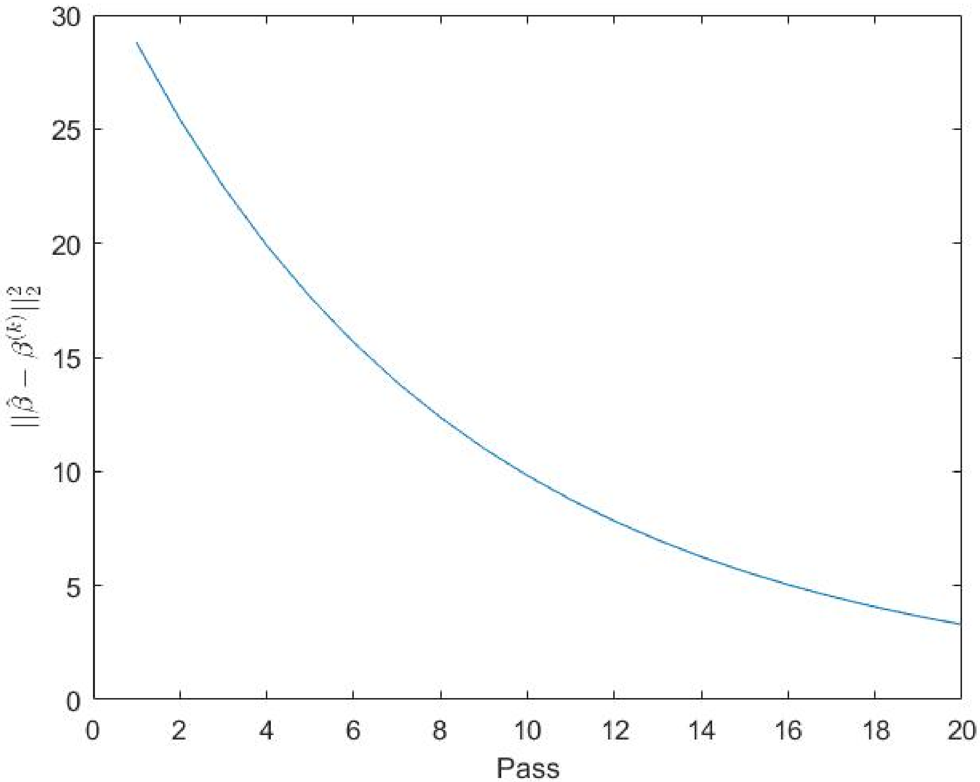




For a step size of $0.01/L$, we have the following figures:


$log(f(\beta^k) - f(\hat{\beta}))$
Pass


$||\beta^* - \beta^{(k)}||_2^2$
Pass



(e) We notice that β^k converges to $\widehat{\beta}$ and β^* only for a step size of $0.1/L$ and appears to diverge for a step size of $1.7/L$. In addition, we notice that β^k converges to some offset values of $\widehat{\beta}$ and β^* for a step size of $1/L$. Finally, it appears that β^k converges for a step size of $0.01/L$, but its convergence is very slow. Ideally, we should run more iterations and passes to confirm that β^k for this step size. These results show the importance of choosing a good step size when using the Stochastic Gradient Descent Algorithm.

4. (a) For GDA, the training error is 0.1233, and the testing error is 0.14.
   (b) For NB-GDA, the training error is 0.12219, and the testing error is 0.135.
   (c) For NB-BDA, the training error is 0.12191, and the testing error is 0.127.
   (d) For QDA, the training error is 0.19217, and the testing error is 0.204.