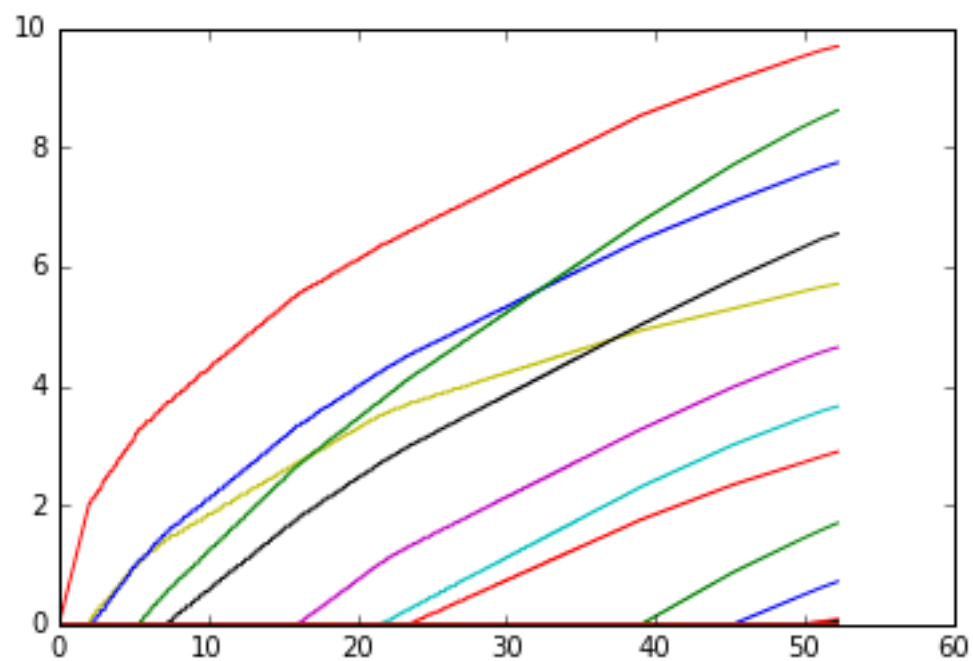


Stagewise Regression at  $\epsilon = 10^{-4}$



Lasso Regression ( $n=50$ ,  $p=200$ )

While Stagewise regression gives a non-continuous, stepped graph for values of epsilon of  $10^{-3}$  and above, comparing Lasso Regression's solution path to Stagewise Regression's solution path, we can see that for an epsilon value of  $10^{-4}$ , stagewise regression gives the same result as lasso regression with  $n=50$  and  $p=200$ . Stagewise regression gives smoother paths than lasso, and takes longer to overfit data, as compared to lasso.

All the following plots were taken with these initial values:

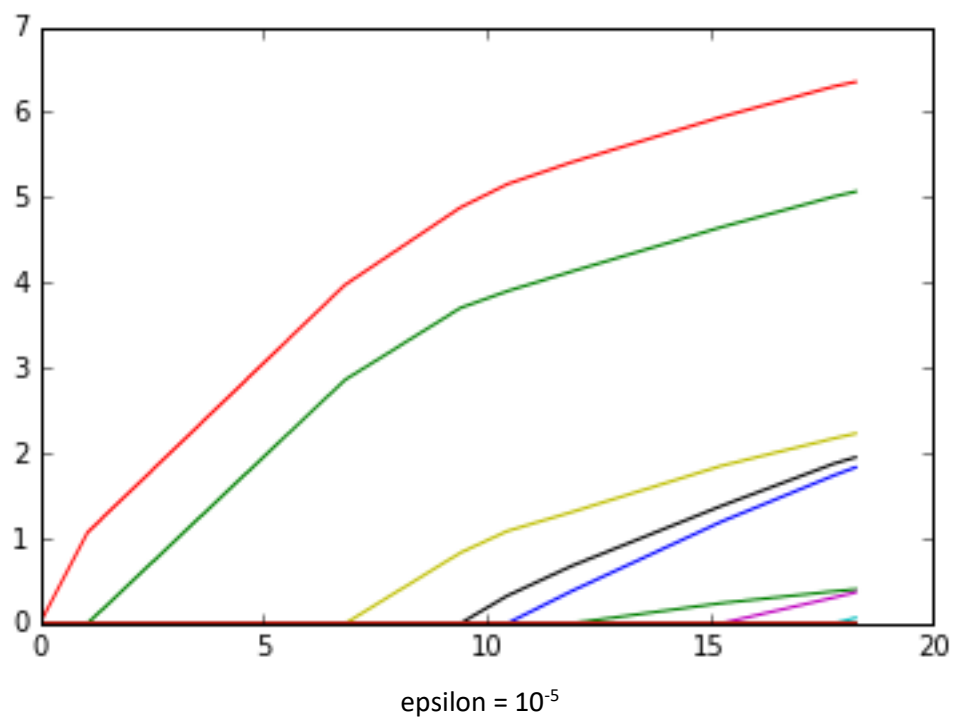
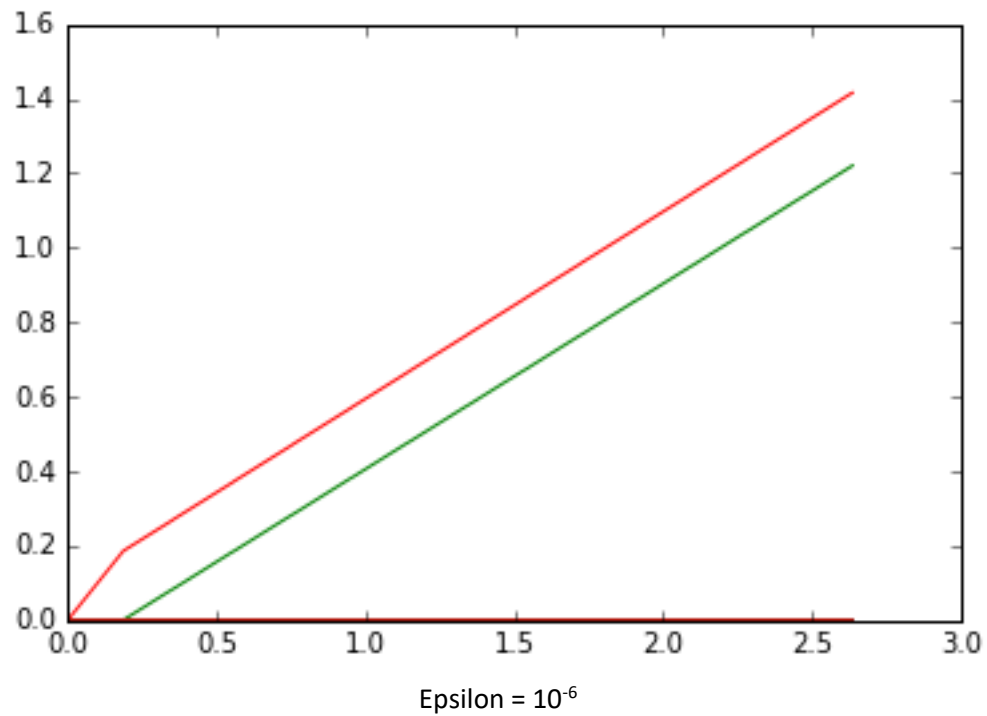
$n=100$

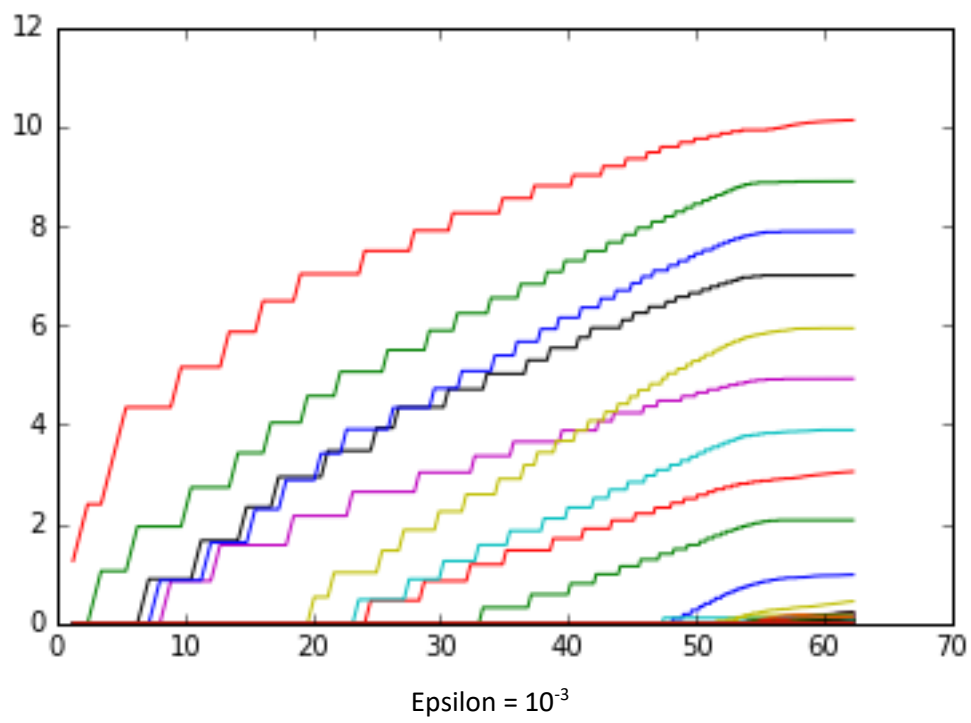
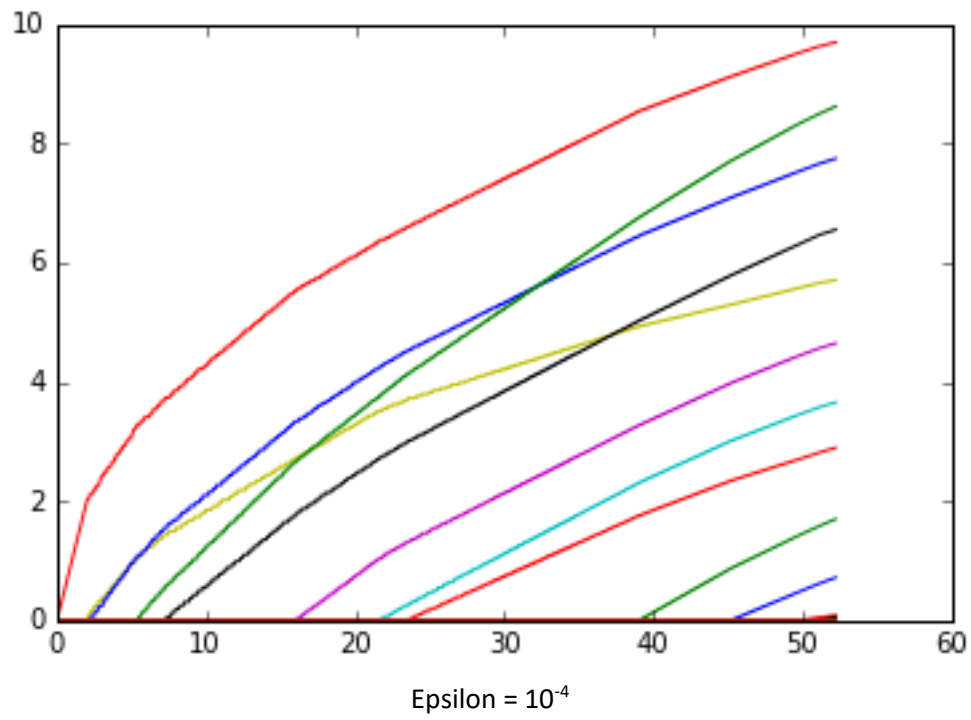
$p=500$

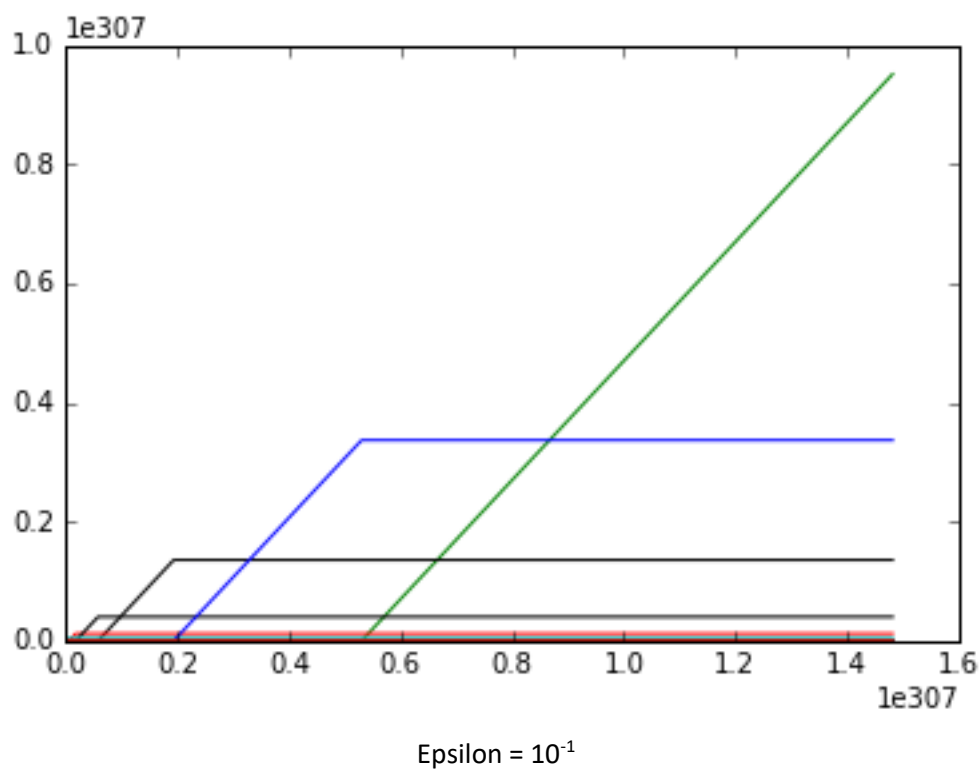
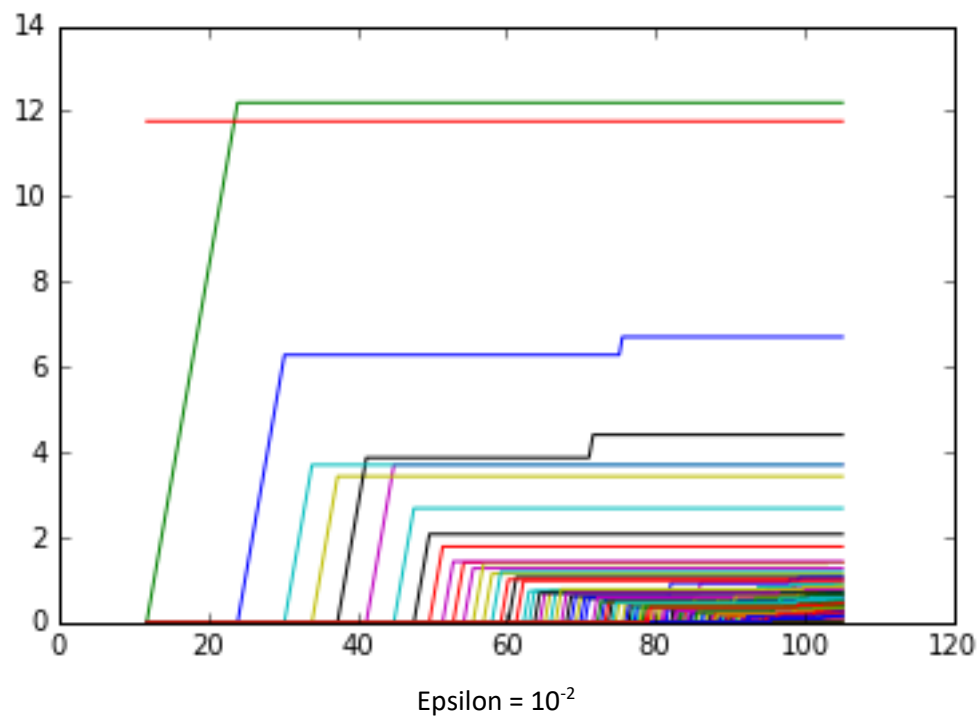
$s=10$

$\text{numIter}=3000$

Epsilon was varied exponentially from 0.000001 to 0.1







## Code to reproduce plot in python:

### Input code:

```
n = 100
p = 500
s = 10
X = np.random.standard_normal((n,p))
beta_true = np.zeros(p)
beta_true[0:s] = range(1, s+1)
Y = np.add(np.dot(X,beta_true),np.random.standard_normal(n,))
swRegression(X,Y)
```

### Plot code:

```
u = np.transpose(np.dot(np.ones((1,p)),abs(beta_all)))
v = np.transpose(beta_all)
plt.figure()
plt.plot(u, v, label='Stagewise Regression')
```

## Code to reproduce plot in R:

### Input code:

```
n = 100
p = 500
s = 10
X = matrix(rnorm(n*p),nrow = n)
beta_true = matrix(rep(0,p), nrow = p)
beta_true[1:s] = 1:s
Y=X%%beta_true + rnorm(n)
swRegression(X,Y)
```

### Plot code:

```
matplot(t(matrix(rep(1, p), nrow = 1)%*%abs(beta_all)), t(beta_all), type = 'l')
```