**Name: Eric Agyemang**

HOMEWORK 3

3.a. iii. Steady decrease

With increase in s we are making the model even more flexible as any restriction on beta is reducing. This will result in decreased in RSS

b. Decrease initially, and then eventually start increasing in a typical U shape

As the model is becoming more and more flexible the test RSS will reduce first and then start increasing when overfitting will start

c. Steadily Increase

Variance steadily increase with the increase in how the model is flexible

d. Steadily decrease

Bias decreases with the increase in the model flexibility

e. Remain constant

Irreducible error is independent of model parameters and thus independent of s

5.a. Given the setting we have x1,1=x1,2=x1 and similarly x2 Thus the ridge regression problem reduces to minimizing

b. On differentiating the expression for ridge regression wrt β1 and β2 & equating it to zero we get expression β1 = β2

c.

9.a. ii. lm (formula = Apps ~., data = train)

Residuals:

Min 1Q Median 3Q Max

-5023.2 -408.6 -46.6 336.7 7242.5

Coefficients:

Estimate Std. Error t value

(Intercept) -140.27742 518.76903 -0.270

PrivateYes -511.17144 176.18600 -2.901

Accept 1.63646 0.05066 32.302

Enroll -1.08353 0.24583 -4.408

Top10perc 58.88924 7.33783 8.025

Top25perc -21.37240 6.10623 -3.500

F.Undergrad 0.07107 0.04330 1.641

P.Undergrad 0.05855 0.03783 1.548

Outstate -0.08871 0.02447 -3.625

Room.Board 0.14479 0.06418 2.256

Books -0.19750 0.29231 -0.676

Personal 0.03265 0.07920 0.412

PhD -8.78197 6.16418 -1.425

Terminal -4.79463 6.69874 -0.716

S.F.Ratio 17.93368 16.68176 1.075

perc.alumni 2.29991 5.51102 0.417

Expend 0.07830 0.01574 4.974

Grad.Rate 9.14788 3.83929 2.383

Pr(>|t|)

(Intercept) 0.786955

PrivateYes 0.003872 \*\*

Accept < 2e-16 \*\*\*

Enroll 1.27e-05 \*\*\*

Top10perc 6.66e-15 \*\*\*

Top25perc 0.000505 \*\*\*

F.Undergrad 0.101333

P.Undergrad 0.122318

Outstate 0.000317 \*\*\*

Room.Board 0.024493 \*

Books 0.499547

Personal 0.680391

PhD 0.154844

Terminal 0.474463

S.F.Ratio 0.282848

perc.alumni 0.676608

Expend 8.92e-07 \*\*\*

Grad.Rate 0.017541 \*

---

Signif. codes:

0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1113 on 525 degrees of freedom

Multiple R-squared: 0.9308, Adjusted R-squared: 0.9286

F-statistic: 415.4 on 17 and 525 DF, p-value: < 2.2e-16

ii. [1] 769127.1

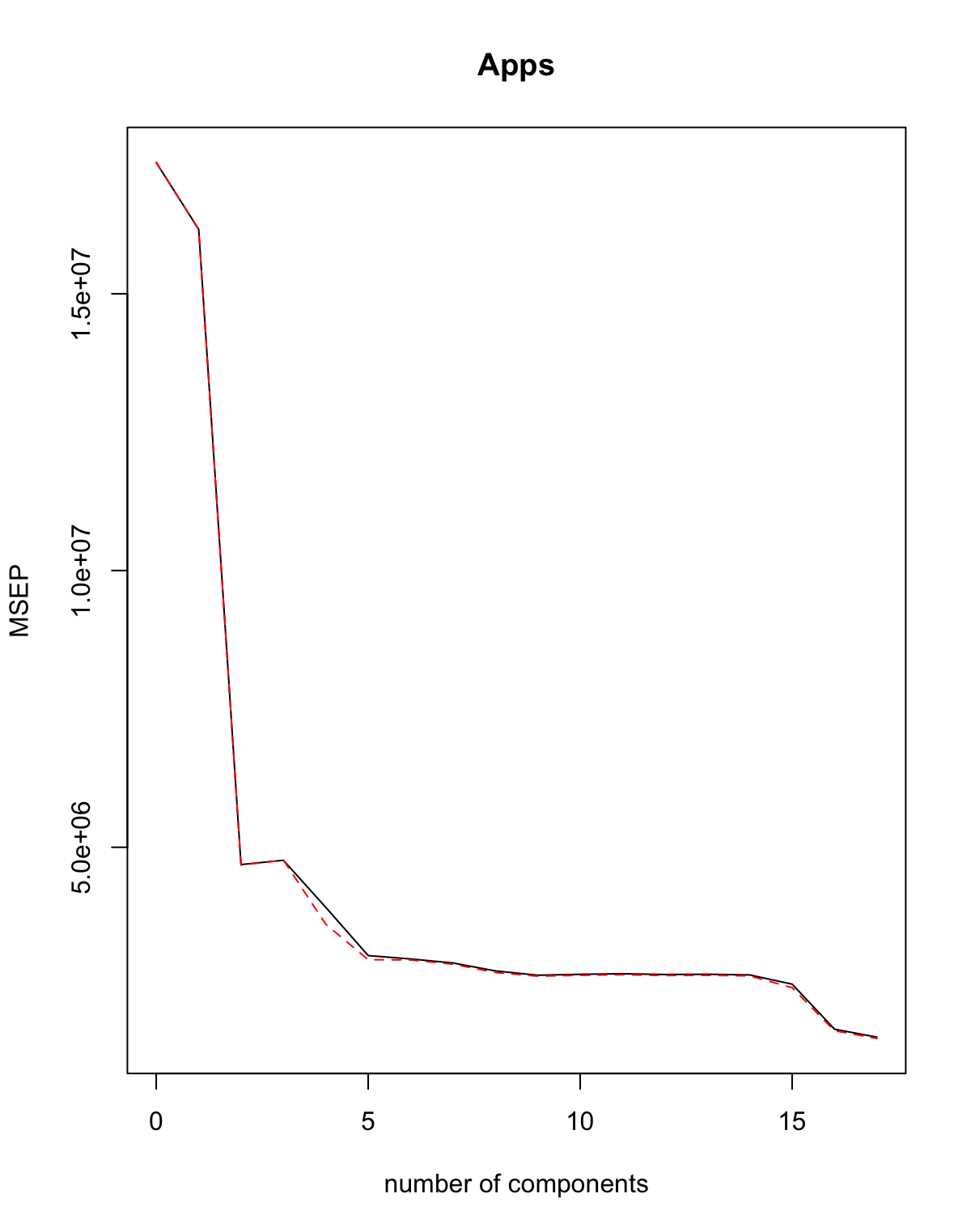
c.iii. [1] 0.01

iii. [1] 769103.1

D.i. [1] 0.01

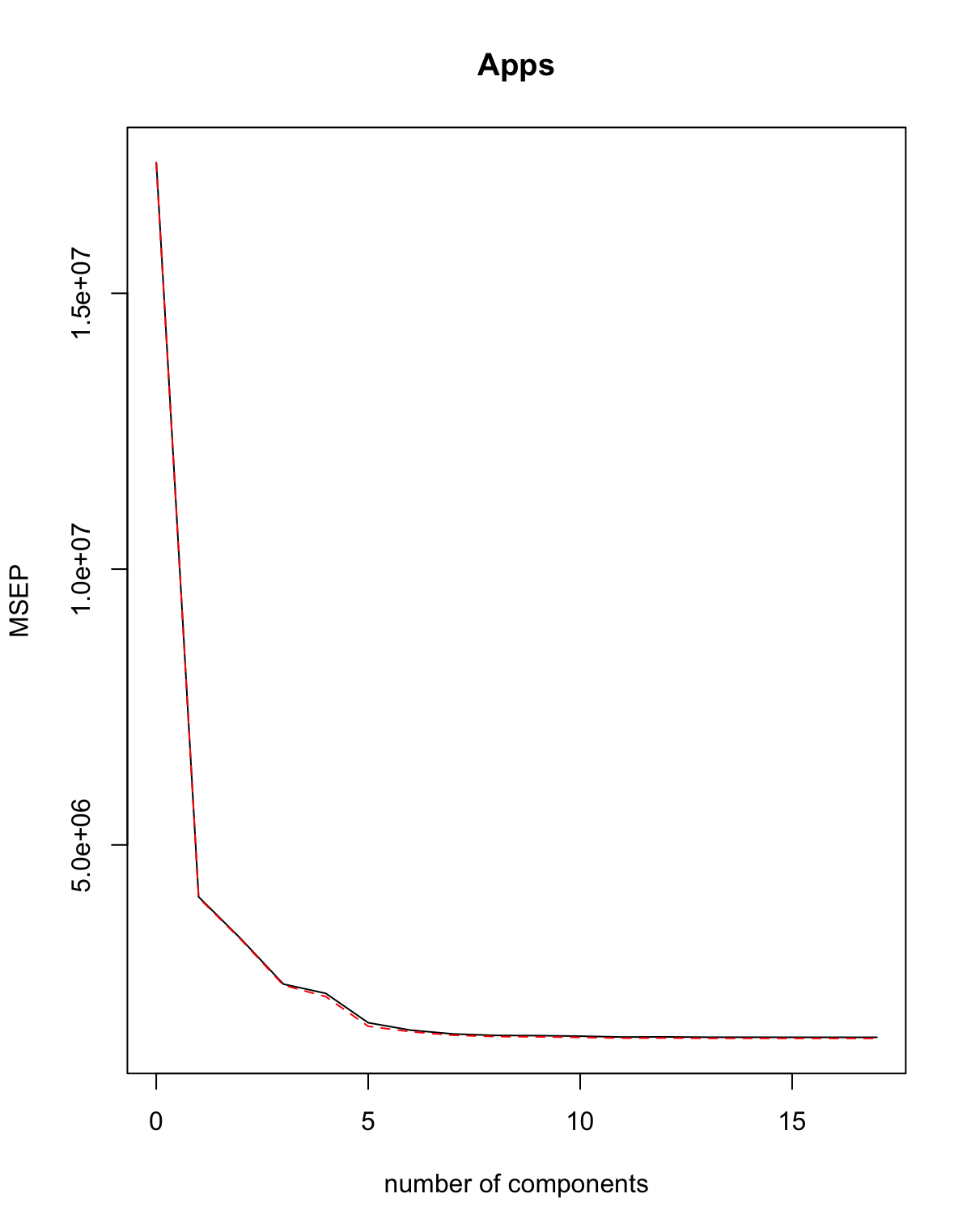
ii. [1] 714910.4

iii.

e.i

ii. [1] 769127.1

We see that the cross validation error is minimum for M=17.If we use 17 components it gives MSE on test set as **7.69127110^{5}**.This is similar to the one obtained for least squares method

F.i. 

ii. [1] 775233.6

We see that the cross validation error is minimum for M=10.If we use 10 components it gives MSE on test set as **7.752336110^{5}**.

g.i. **Least Square** Test R-Square: **0.918807**

**Ridge Model** Test R-Square: **0.9188096**

**Lasso Model** Test R-Square: **0.9188143**

**PCR Model** Test R-Square: **0.918807**

**PLS Model** Test R-Square: **0.9181624**

It can be seen that the Lasso model predicts the highest R-Square. Though all the models have similar performance

This was expected as the minimum MSE was found for Lasso model across all the model