Xu_Steven_hw3

Steven Xu

10/1/2019

1.

For a fixed λ_2, the elastic net problem can be transformed to an equivalent lasso problem by augmenting both the response vector and design matrix by means of

$$\mathbf{X}^* = (1+\lambda_2)^{-1/2}\begin{pmatrix}\mathbf{X}\\ \sqrt{\lambda_2}\mathbf{I}\end{pmatrix},\ \mathbf{y}^* = \begin{pmatrix}y\\ \mathbf{0}\end{pmatrix}$$

Then the loss function will consist only a L_1 penalty on

$$\beta^* = \sqrt{1+\lambda_2}\beta$$

with penalty coefficient $\gamma = \lambda_1/\sqrt{1+\lambda_2}$

The solution to the elastic net can then be retrieved by transforming the solution to the above system by

$$\hat{\beta} = \frac{1}{\sqrt{1+\lambda_2}}\hat{\beta}^*$$

2.

(a)

```
##
## Call:
## lm(formula = lpsa ~ ., data = pro_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.64870 -0.34147 -0.05424  0.44941  1.48675
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.429170   1.553588   0.276  0.78334
## lcavol       0.576543   0.107438   5.366 1.47e-06 ***
## lweight      0.614020   0.223216   2.751  0.00792 **
## age         -0.019001   0.013612  -1.396  0.16806
## lbph         0.144848   0.070457   2.056  0.04431 *
## svi          0.737209   0.298555   2.469  0.01651 *
## lcp         -0.206324   0.110516  -1.867  0.06697 .
## gleason     -0.029503   0.201136  -0.147  0.88389
## pgg45        0.009465   0.005447   1.738  0.08755 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7123 on 58 degrees of freedom
## Multiple R-squared:  0.6944, Adjusted R-squared:  0.6522
## F-statistic: 16.47 on 8 and 58 DF,  p-value: 2.042e-12
```

Based on the ouput we can see that $R^2 = 0.6944$, the set of significant predictors are lcavol, lweight, lbph, svi. TrainErr = 0.4392 and TestErr = 0.5213.

(b)-(c)

```
## The sets of regression coefficients for M1 - M8 along with the estimated TrainErr,BIC,AIC are
```

```
## M1
## (Intercept)      lcavol
##   1.5163048   0.7126351
##
## TrainErr = 0.6646057
## BIC = -18.96422
## AIC = -23.37361
##
## M2
## (Intercept)      lcavol     lweight
##  -1.0494396   0.6276074   0.7383751
##
## TrainErr = 0.5536096
## BIC = -27.00272
## AIC = -33.6168
##
## M3
## (Intercept)      lcavol     lweight         svi
##  -1.0227780   0.5199861   0.7367954   0.5379032
##
## TrainErr = 0.5210112
## BIC = -26.86414
## AIC = -35.68291
##
## M4
## (Intercept)      lcavol     lweight        lbph         svi
##  -0.3259212   0.5055209   0.5388292   0.1400111   0.6718487
##
## TrainErr = 0.489776
## BIC = -26.80161
## AIC = -37.82507
##
## M5
##  (Intercept)       lcavol      lweight         lbph          svi
## -0.465877591  0.472278483  0.563935476  0.137116261  0.578163005
##        pgg45
##  0.004330753
##
## TrainErr = 0.4786485
## BIC = -24.1367
## AIC = -37.36485
##
## M6
##  (Intercept)       lcavol      lweight         lbph          svi
## -0.728972257  0.549778034  0.563105747  0.125978836  0.756354835
##          lcp        pgg45
## -0.190824719  0.007541236
```

```
##
## TrainErr = 0.4558176
## BIC = -23.20654
## AIC = -38.63939
##
## M7
##  (Intercept)       lcavol      lweight          age         lbph
##  0.259061747  0.573930391  0.619208833 -0.019479879  0.144426474
##          svi          lcp        pgg45
##  0.741781258 -0.205416986  0.008944996
##
## TrainErr = 0.4393627
## BIC = -21.46527
## AIC = -39.10281
##
## M8
##  (Intercept)       lcavol      lweight          age         lbph
##  0.429170133  0.576543185  0.614020004 -0.019001022  0.144848082
##          svi          lcp      gleason        pgg45
##  0.737208645 -0.206324227 -0.029502884  0.009465162
##
## TrainErr = 0.4391998
## BIC = -17.28543
## AIC = -37.12766
```

The model that minimizes BIC is M2, with set of important variables being lcavol, lweight. The TestErr of the refitted OLS is 0.4925.

The model that minimizes AIC is M7, with set of important variables being lcavol, lweight, age, lbph, svi, lcp, pgg45. The TestErr of the refitted OLS is 0.5165135.

3.

Part 1

Before running further analysis, the design matrix is centered by column means and scaled by column L_2-norms. By centering the design matrix, the estimated intercept for all below models are directly calculated by taking the sample mean of the response.

Since by default **cv.lars()** and **lars()** output the estimated fraction (s), the corresponding λ is found by method of root searching using **uniroot**.

(a)

```
## The best lambda is 0.1412
```

```
## The selected model and its estimated regression coefficients are
```

```
##     lcavol    lweight        age       lbph        svi        lcp
##  5.4020501  2.2804329 -0.8347302  1.5704452  2.2116353 -1.4482073
##    gleason      pgg45
##  0.0000000  1.6488464
```

```
## TestErr = 0.4873755
```

(b)

```
## The best lambda is 1.2459
```

```
## The selected model and its estimated regression coefficients are
```

```
##    lcavol   lweight       age      lbph       svi       lcp   gleason
## 4.6207321 1.7085083 0.0000000 0.4569907 1.1002017 0.0000000 0.0000000
##     pgg45
## 0.2754862
```

```
## TestErr = 0.4569148
```

Part 2

Convergence is detected if the change of L_1-norm is small ($<10^{-5}$) for more than 5 iterations.

```
## The best lambda is 1.0476 with a BIC of -22.58
```

```
## The selected model and its estimated coefficients are
```

```
##   lcavol  lweight      age     lbph      svi      lcp  gleason    pgg45
## 4.705844 1.983815 0.000000 1.137093 1.605886 0.000000 0.000000 0.713330
```

```
## TestErr = 0.4583168
```

Part 3

```
## The best lambda is 0.4914 with a BIC of -22.4531
```

```
## The selected model and its estimated coefficients are
```

```
##    lcavol   lweight       age      lbph       svi       lcp   gleason
## 5.1962616 2.0264436 0.0000000 0.9734609 1.4426308 0.0000000 0.0000000
##     pgg45
## 0.6029959
```

```
## TestErr = 0.4420785
```