R Notebook

Code

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

Hide

```
data = mtcars
```

Hide

```
model = lm(mpg~., data = data)
```

Hide

```
model
```

```

Call:
lm(formula = mpg ~ ., data = data)

Coefficients:
(Intercept)          cyl         disp           hp         drat           wt         qsec           vs           am
   12.30337     -0.11144      0.01334     -0.02148      0.78711     -3.71530      0.82104      0.31776      2.52023
       gear         carb
    0.65541     -0.19942
```

Hide

```
summary(model)
```

```
Call:
lm(formula = mpg ~ ., data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.4506 -1.6044 -0.1196  1.2193  4.6271

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 12.30337   18.71788   0.657   0.5181
cyl         -0.11144    1.04502  -0.107   0.9161
disp         0.01334    0.01786   0.747   0.4635
hp          -0.02148    0.02177  -0.987   0.3350
drat         0.78711    1.63537   0.481   0.6353
wt          -3.71530    1.89441  -1.961   0.0633 .
qsec         0.82104    0.73084   1.123   0.2739
vs           0.31776    2.10451   0.151   0.8814
am           2.52023    2.05665   1.225   0.2340
gear         0.65541    1.49326   0.439   0.6652
carb        -0.19942    0.82875  -0.241   0.8122
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.65 on 21 degrees of freedom
Multiple R-squared:  0.869, Adjusted R-squared:  0.8066
F-statistic: 13.93 on 10 and 21 DF,  p-value: 3.793e-07
```

Hide

```
 0.2739 > 0.05
```

```
[1] TRUE
```

Hide

```
formula(model)
```

```
mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
```

Hide

```
backward_ = step(model, direction = "backward")
```

```
Start:  AIC=70.9
mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb

       Df Sum of Sq    RSS    AIC
- cyl   1    0.0799 147.57 68.915
- vs    1    0.1601 147.66 68.932
- carb  1    0.4067 147.90 68.986
- gear  1    1.3531 148.85 69.190
- drat  1    1.6270 149.12 69.249
- disp  1    3.9167 151.41 69.736
- hp    1    6.8399 154.33 70.348
- qsec  1    8.8641 156.36 70.765
<none>              147.49 70.898
- am    1   10.5467 158.04 71.108
- wt    1   27.0144 174.51 74.280

Step:  AIC=68.92
mpg ~ disp + hp + drat + wt + qsec + vs + am + gear + carb

       Df Sum of Sq    RSS    AIC
- vs    1    0.2685 147.84 66.973
- carb  1    0.5201 148.09 67.028
- gear  1    1.8211 149.40 67.308
- drat  1    1.9826 149.56 67.342
- disp  1    3.9009 151.47 67.750
- hp    1    7.3632 154.94 68.473
<none>              147.57 68.915
- qsec  1   10.0933 157.67 69.032
- am    1   11.8359 159.41 69.384
- wt    1   27.0280 174.60 72.297

Step:  AIC=66.97
mpg ~ disp + hp + drat + wt + qsec + am + gear + carb

       Df Sum of Sq    RSS    AIC
- carb  1    0.6855 148.53 65.121
- gear  1    2.1437 149.99 65.434
- drat  1    2.2139 150.06 65.449
- disp  1    3.6467 151.49 65.753
- hp    1    7.1060 154.95 66.475
```

```
<none>              147.84 66.973
- am    1   11.5694 159.41 67.384
- qsec  1   15.6830 163.53 68.200
- wt    1   27.3799 175.22 70.410

Step:  AIC=65.12
mpg ~ disp + hp + drat + wt + qsec + am + gear

       Df Sum of Sq    RSS    AIC
- gear  1     1.565 150.09 63.457
- drat  1     1.932 150.46 63.535
<none>              148.53 65.121
- disp  1    10.110 158.64 65.229
- am    1    12.323 160.85 65.672
- hp    1    14.826 163.35 66.166
- qsec  1    26.408 174.94 68.358
- wt    1    69.127 217.66 75.350

Step:  AIC=63.46
mpg ~ disp + hp + drat + wt + qsec + am

       Df Sum of Sq    RSS    AIC
- drat  1     3.345 153.44 62.162
- disp  1     8.545 158.64 63.229
<none>              150.09 63.457
- hp    1    13.285 163.38 64.171
- am    1    20.036 170.13 65.466
- qsec  1    25.574 175.67 66.491
- wt    1    67.572 217.66 73.351

Step:  AIC=62.16
mpg ~ disp + hp + wt + qsec + am

       Df Sum of Sq    RSS    AIC
- disp  1     6.629 160.07 61.515
<none>              153.44 62.162
- hp    1    12.572 166.01 62.682
- qsec  1    26.470 179.91 65.255
- am    1    32.198 185.63 66.258
- wt    1    69.043 222.48 72.051
```

```
Step:  AIC=61.52
mpg ~ hp + wt + qsec + am

       Df Sum of Sq    RSS    AIC
- hp    1     9.219 169.29 61.307
<none>              160.07 61.515
- qsec  1    20.225 180.29 63.323
- am    1    25.993 186.06 64.331
- wt    1    78.494 238.56 72.284

Step:  AIC=61.31
mpg ~ wt + qsec + am

       Df Sum of Sq    RSS    AIC
<none>              169.29 61.307
- am    1    26.178 195.46 63.908
- qsec  1   109.034 278.32 75.217
- wt    1   183.347 352.63 82.790
```

Hide

```
summary(backward_)
```

```
Call:
lm(formula = mpg ~ wt + qsec + am, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.4811 -1.5555 -0.7257  1.4110  4.6610

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.6178     6.9596   1.382 0.177915
wt           -3.9165     0.7112  -5.507 6.95e-06 ***
qsec          1.2259     0.2887   4.247 0.000216 ***
am            2.9358     1.4109   2.081 0.046716 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.459 on 28 degrees of freedom
Multiple R-squared:  0.8497,    Adjusted R-squared:  0.8336
F-statistic: 52.75 on 3 and 28 DF,  p-value: 1.21e-11
```

Setwise forward

Hide

```
forward_ = step( lm(mpg~1, data=data), direction = "forward", scope = formula(model) )
```

```
Start:  AIC=115.94
mpg ~ 1

       Df Sum of Sq     RSS     AIC
+ wt    1    847.73  278.32  73.217
+ cyl   1    817.71  308.33  76.494
+ disp  1    808.89  317.16  77.397
+ hp    1    678.37  447.67  88.427
+ drat  1    522.48  603.57  97.988
+ vs    1    496.53  629.52  99.335
+ am    1    405.15  720.90 103.672
+ carb  1    341.78  784.27 106.369
+ gear  1    259.75  866.30 109.552
+ qsec  1    197.39  928.66 111.776
<none>              1126.05 115.943

Step:  AIC=73.22
mpg ~ wt

       Df Sum of Sq    RSS    AIC
+ cyl   1    87.150 191.17 63.198
+ hp    1    83.274 195.05 63.840
+ qsec  1    82.858 195.46 63.908
+ vs    1    54.228 224.09 68.283
+ carb  1    44.602 233.72 69.628
+ disp  1    31.639 246.68 71.356
<none>              278.32 73.217
+ drat  1     9.081 269.24 74.156
+ gear  1     1.137 277.19 75.086
+ am    1     0.002 278.32 75.217

Step:  AIC=63.2
mpg ~ wt + cyl

       Df Sum of Sq    RSS    AIC
+ hp    1   14.5514 176.62 62.665
+ carb  1   13.7724 177.40 62.805
<none>              191.17 63.198
+ qsec  1   10.5674 180.60 63.378
+ gear  1    3.0281 188.14 64.687
```

```
+ disp  1    2.6796 188.49 64.746
+ vs    1    0.7059 190.47 65.080
+ am    1    0.1249 191.05 65.177
+ drat  1    0.0010 191.17 65.198

Step:  AIC=62.66
mpg ~ wt + cyl + hp

       Df Sum of Sq    RSS    AIC
<none>              176.62 62.665
+ am    1    6.6228 170.00 63.442
+ disp  1    6.1762 170.44 63.526
+ carb  1    2.5187 174.10 64.205
+ drat  1    2.2453 174.38 64.255
+ qsec  1    1.4010 175.22 64.410
+ gear  1    0.8558 175.76 64.509
+ vs    1    0.0599 176.56 64.654
```

Hide

```
forward_
```

```

Call:
lm(formula = mpg ~ wt + cyl + hp, data = data)

Coefficients:
(Intercept)           wt          cyl           hp
   38.75179     -3.16697     -0.94162     -0.01804
```

Hide

```
summary(forward_)
```

```
Call:
lm(formula = mpg ~ wt + cyl + hp, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.9290 -1.5598 -0.5311  1.1850  5.8986

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 38.75179    1.78686  21.687  < 2e-16 ***
wt          -3.16697    0.74058  -4.276 0.000199 ***
cyl         -0.94162    0.55092  -1.709 0.098480 .
hp          -0.01804    0.01188  -1.519 0.140015
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.512 on 28 degrees of freedom
Multiple R-squared:  0.8431,    Adjusted R-squared:  0.8263
F-statistic: 50.17 on 3 and 28 DF,  p-value: 2.184e-11
```

Hide

```
0.0984 > 0.01
```

```
[1] TRUE
```

Para um nivel de significancia 0.001 há evidencia suficiente para aceitar H_0:cyl=0

Hide

```
library(glmnet)
```

```
Aviso: pacote ‘glmnet’ foi compilado no R versão 4.4.2Carregando pacotes exigidos: Matrix
Loaded glmnet 4.1-8
```

Hide

```
x = model.matrix(data=data)
```

```
Error in model.matrix.default(data = data) :
  argumento "object" ausente, sem padrão
```

Hide

```
cv.glmnet(x, y, alpha = 1)
```

```

Call:  cv.glmnet(x = x, y = y, alpha = 1)

Measure: Mean-Squared Error

    Lambda Index Measure    SE Nonzero
min 0.6648    23   8.343 2.241       4
1se 1.5357    14  10.378 3.820       3
```

Hide

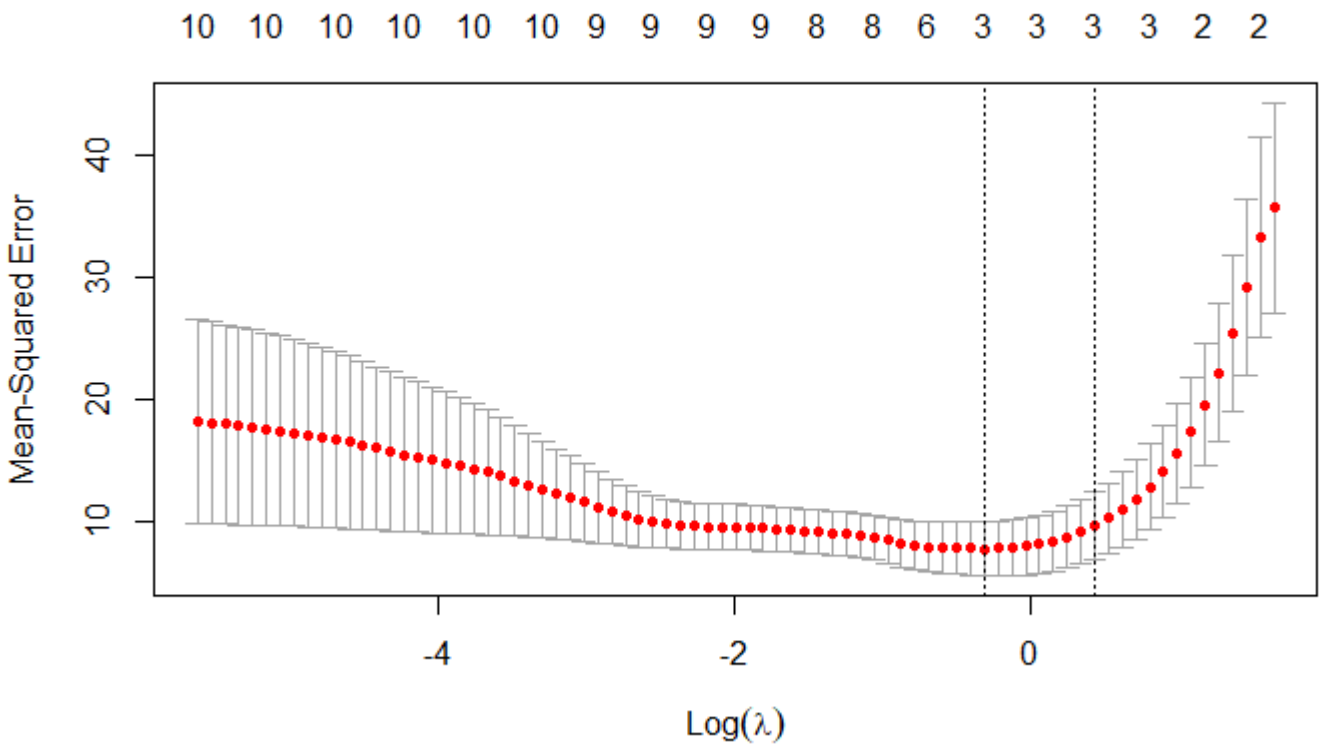
```
lasso
```

```

Call:  cv.glmnet(x = x, y = y, alpha = 1)

Measure: Mean-Squared Error

    Lambda Index Measure    SE Nonzero
min 0.7296    22   7.706 2.202       3
1se 1.5357    14   9.598 2.784       3
```


10 10 10 10 10 10 9 9 9 9 8 8 6 3 3 3 3 2 2
Mean-Squared Error
10
20
30
40
-4
-2
0
Log(λ)