Linear-Regression.R

patriciamaya

2020-12-05

```
library(ISLR)
```

```
## Warning: package 'ISLR' was built under R version 4.0.2
```

```
data("Carseats")
attach(Carseats)

set.seed(123)
indx <- sample(2,nrow(Carseats), replace=T, prob = c(0.8, 0.2))
train <- Carseats[indx ==1, ]
test <- Carseats[indx ==2, ]

#lm - linear model (~)
#lm(num target ~ inputs, data= train)

lmModel <- lm(Sales ~ ., data= train)
summary(lmModel)
```

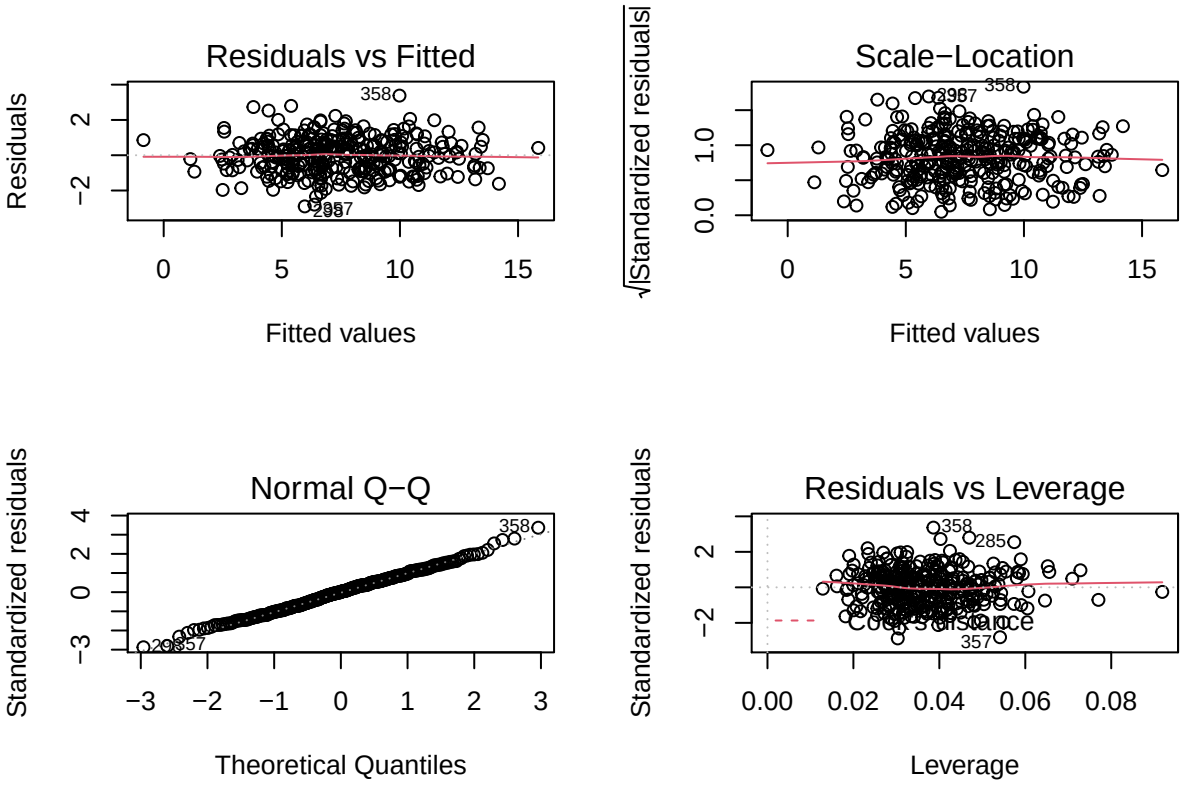
```
##
## Call:
## lm(formula = Sales ~ ., data = train)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.8993 -0.7146  0.0192  0.6676  3.3789
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.843e+00  6.790e-01   8.605 3.75e-16 ***
## CompPrice        9.105e-02  4.495e-03  20.254  < 2e-16 ***
## Income           1.659e-02  2.069e-03   8.017 2.18e-14 ***
## Advertising      1.228e-01  1.222e-02  10.047  < 2e-16 ***
## Population       8.435e-06  4.182e-04   0.020    0.984
## Price           -9.555e-02  2.896e-03 -32.999  < 2e-16 ***
## ShelveLocGood    4.857e+00  1.703e-01  28.511  < 2e-16 ***
## ShelveLocMedium  1.884e+00  1.387e-01  13.583  < 2e-16 ***
## Age             -4.679e-02  3.610e-03 -12.960  < 2e-16 ***
## Education       -1.783e-02  2.249e-02  -0.793    0.428
## UrbanYes         1.947e-01  1.251e-01   1.556    0.121
## USYes           -1.624e-01  1.672e-01  -0.971    0.332
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.024 on 313 degrees of freedom
## Multiple R-squared:  0.876,  Adjusted R-squared:  0.8716
## F-statistic:   201 on 11 and 313 DF,  p-value: < 2.2e-16
```

```
#estimate- coefficient of each variable in the linear model
#if coefficient is larger-> means variable is more important & sign means direction of relationship
#if we increase price, sales decrease.
#Pr(>|t|) -p value, whether there is a significant relationship between each variable & target variable
#we want r-squared to be large-> indicates good fit of model

#Population around the store does NOT affect sales much, also education level, if urban, and if store
#is in USA.

layout(matrix(c(1,2,3,4), 2, 2))
plot(lmModel)
```




```
#Residuals vs Fitted - we don't want to see any pattern, points should be
  #scattered at random, should be a horizontal line.

#Scale-Location - variation of residuals,
  #line should be horizontal-> means variation is the same around all points.
#Normal Q-Q - check that residuals have normal distribution, should be a 45 degree line.

#Residual vs Leverage - helps identify influential points.
  #no influential points in our model.

#mean squared error
predictions <- predict(lmModel, newdata = test)

mean(test$Sales-predictions)^2
```

```
## [1] 0.02526576
```

```
## [1] 0.02526576

fitted(lmModel) #predictions for train values
```

```
##         1         2         3         6         7         9        10
##  7.269306 12.450736  9.240229  9.630682  6.225408  5.988411  5.751118
##        12        13        14        15        17        18        19
## 11.885747  3.642487 11.999492  9.737271  8.480791 11.813086 13.427286
##        22        23        25        26        27        28        29
## 11.057106  5.697161 10.041736 13.329523  8.488569  5.280351  4.445458
##        30        33        34        35        36        37        38
##  6.066833  6.022051  8.485141  4.630409 10.701775 10.386680  6.858415
##        39        40        41        42        43        44        45
##  6.005887  3.533196  2.544935  6.367282 11.123760  5.209829  5.010800
##        46        47        48        49        51        52        53
##  4.587227 13.138225  5.958299  4.551147  3.284849  4.434350  6.366582
##        54        55        56        57        58        60        61
##  5.814153  5.451772  4.847478 10.723155  1.129605  5.467949  7.670400
##        62        63        64        66        69        70        71
##  6.148943  2.664257  8.546695  4.232405 12.058326  7.751798  8.886307
##        72        73        74        75        76        77        78
##  6.502409  6.621525 12.417926  6.743824  7.601618  9.769819  9.000272
##        79        80        81        82        83        84        85
##  3.769988  8.602035  8.863534  6.914214 11.455758  5.461889  2.469757
##        86        90        91        92        93        94        95
##  7.593101  7.087390  5.534545  4.472604  6.661793  8.808447  7.773796
##        96        97        98        99       100       101       102
##  4.470777  9.535457  8.355042 11.767615  4.717640  6.450947  7.932666
##       103       105       108       109       110       112       113
##  5.786150  4.748785  8.557242  4.171312  7.567753  6.521143  8.492899
##       115       116       117       119       120       121       122
##  8.014251  6.636094  5.112505  7.107561  6.699438  6.811407 10.162209
##       123       124       125       127       128       129       130
##  6.860236  7.374707 10.033625 10.281886  6.587000  5.169832  5.456701
##       131       133       134       135       136       138       140
##  9.354223  8.120641  7.383417  4.098863  6.115451  5.062190 12.499747
##       141       142       143       144       146       147       148
##  7.341957  5.389058  6.800197  2.496577  9.584844  3.205531 11.713774
##       149       150       152       153       154       155       156
##  6.670213 11.235010 10.042160  9.082304  6.877537  8.549643  8.119811
##       157       158       159       160       161       162       163
##  7.079880  9.450994 12.378212  9.262807  6.125237  2.910791  3.405649
##       164       165       166       167       168       169       170
##  6.484732  7.134335  1.300615  5.934335  6.860316  7.683714 11.632781
##       171       172       174       175       176       177       178
##  6.761838 10.430247  6.536572 -0.856677  6.828125  5.903369 10.971593
##       180       182       183       184       185       186       187
##  6.646598  8.243220  5.220831  6.660692  8.718632  9.212235  8.827054
##       188       191       192       194       196       197       198
##  4.845222 10.000174  6.987208 13.203263  4.579002  2.553638  3.199190
##       199       200       201       203       204       205       207
```

```
##  4.2647069  5.6602425  5.3017797  3.7333899  2.9051601  9.5343270  4.4268623
##        208        209        210        211        212        213        214
##  5.3937343  6.8942761  3.9532442  3.7886330  8.8614246 12.1089720  8.3309733
##        215        216        217        218        221        223        224
##  6.5457253  2.3791112  5.6559879  3.9812124 10.0910918  7.1298377  4.5342890
##        225        226        227        228        229        231        232
##  5.3413905  6.4286845  7.7511719  7.7351550  5.0695438  4.3891536  7.7745751
##        233        234        235        236        237        239        241
## 12.6058744  9.1639460  8.1574246  5.4903078  9.4029116  5.9549368 10.6296185
##        242        243        244        245        246        247        250
## 10.3772065  4.1998998  7.7054614  9.1667266 10.3775925  7.6177916  4.1136495
##        251        252        253        254        255        256        257
##  7.6539196  5.2429787  8.6788989  5.3312258 10.7566131  7.6513321  4.6186521
##        258        259        263        265        266        267        268
##  7.2154583  4.9503808  6.6255777  6.5730107  5.0693283 10.7895976  6.8026194
##        269        270        272        273        274        275        276
##  8.0131995  5.2299713  5.2645851 13.7184162  9.9185985  8.6822732  5.5392131
##        278        279        280        281        282        283        284
##  8.5010857  6.4700292  4.7297836  3.1309884 11.2277740  8.3790868  6.6363156
##        285        286        287        288        289        290        291
##  4.4387722  8.3166803  7.2759106  6.5059319  5.8611104  7.9092930  8.9143538
##        292        293        295        298        299        300        301
##  7.3112624 13.1880103 13.3738165  5.9692861  9.7799873 10.6090493  7.9778469
##        302        303        304        305        306        307        308
##  7.0522786  4.3623586 11.0311400 10.6939608  7.5186100  5.1730926  5.5940017
##        309        310        311        312        313        314        315
##  8.2445785 10.0934330  9.7805712  6.5237696  6.0504319 10.5363559  8.7214363
##        318        319        322        323        324        325        326
##  7.1242762  8.9349126  9.1637972  9.1818061  9.8106525  2.7837925 11.0344024
##        328        329        331        332        333        335        336
##  6.7117069  4.3700741  5.4849993  8.4604242  7.4115345  7.2647218  7.5453393
##        337        338        339        341        342        343        344
##  4.6252253  6.6341776  6.5666191  8.2739531  6.2091645  8.5548294  5.4276928
##        345        346        348        349        350        351        353
##  8.8829301  4.8935931  7.8735736 14.1888310 10.7123828  9.4902091 11.4564821
##        354        355        357        358        359        361        362
##  9.2532619  4.3656946  6.3784885  9.9811372  5.1952094  9.9755176  8.3115072
##        364        365        366        367        368        369        370
## 10.6671625  9.9862564  3.7980568  5.6405996 13.5035464 10.7964694 10.8238967
##        371        372        374        375        377        378        379
##  7.1825044  8.2551579  6.5400318  7.6387671 15.8549829  6.6835138  6.4169409
##        381        382        383        384        385        387        388
##  7.7114663  2.5632389  6.6772413  9.1624060 12.1634356  6.2885357  8.7721120
##        389        390        392        393        394        395        396
##  9.5610293  9.0200603  6.2595241  5.6717540  6.3903289  5.8986975 13.1526578
##        397        398        399
##  6.7352176  7.0322852  5.5213980
```

```
coefficients(lmModel)   #coefficients of regression
```

```
##     (Intercept)       CompPrice          Income     Advertising      Population
##    5.842839e+00    9.105120e-02    1.658512e-02    1.227960e-01    8.434767e-06
##           Price   ShelveLocGood ShelveLocMedium             Age       Education
##   -9.555257e-02    4.856521e+00    1.884479e+00   -4.679088e-02   -1.783126e-02
##        UrbanYes           USYes
```

```
##    1.946842e-01   -1.624238e-01
```

```
residuals(lmModel) #residuals from actual train values - predicted train values
```

```
##           1           2           3           6           7           9
##  2.230694030 -1.230735704  0.819770826  1.179317660  0.404592547  0.551588688
##          10          12          13          14          15          17
## -1.061118049  0.074253427  0.337512644 -1.039492364  1.432729453 -0.900790488
##          18          19          22          23          25          26
##  0.476914082  0.482714358  1.072893677 -0.617160949  0.098264051  1.570477230
##          27          28          29          30          33          34
## -0.158568562 -0.010350553 -1.455458019  1.743167126  0.177949539  0.284859082
##          35          36          37          38          39          40
## -1.960409363  0.368225301 -1.496680105 -1.908415281  0.584112912 -0.293195755
##          41          42          43          44          45          46
## -0.474935404  1.592718155 -0.693759606 -1.089829057 -0.850800387 -0.027226995
##          47          48          49          51          52          53
## -0.698224835 -1.578298726 -0.641146792 -1.864848716 -0.014350196  1.543418266
##          54          55          56          57          58          60
##  1.105847319 -0.551772363  2.002522027  1.186845148 -0.219605349 -0.257949217
##          61          62          63          64          66          69
##  0.649599794  1.171057359 -0.844256730 -0.076695497  0.667595210  1.331673510
##          70          71          72          73          74          75
##  0.238201538  0.573693350 -0.002409133 -1.101524754  0.192074307 -0.543823736
##          76          77          78          79          80          81
##  0.948382009  0.870180973 -1.300271663  0.660012304  0.537964840 -0.853533793
##          82          83          84          85          86          90
##  0.605786098  0.164241631 -1.041889288 -0.239757190  0.876898810  0.862610434
##          91          92          93          94          95          96
## -0.204544558  0.337396123 -2.131792916  0.051552936  0.616204205  1.109222668
##          97          98          99         100         101         102
## -0.055457085 -0.905042128  0.722384610  0.162359956 -2.340947097 -1.732665897
##         103         105         108         109         110         112
## -0.486149727 -0.128784582 -0.007241515 -0.701311868  1.412247213  0.098856851
##         113         115         116         117         119         120
## -1.822898941  1.295748957  1.903906079 -0.032504917  0.462438692  0.670561808
##         121         122         123         124         125         127
##  0.058593189  1.507790674  0.019764208  0.815292805 -1.163624695  0.988113927
##         128         129         130         131         133         134
## -0.067000358 -0.209832125 -0.986700529 -0.944223132  1.419358770  0.236582738
##         135         136         138         140         141         142
## -0.428863087  0.324549366  1.457809894 -0.199746712 -1.311957432  1.140942340
##         143         144         146         147         148         149
##  0.639803368 -1.966577256 -0.814843590  0.694468902 -1.203773476  0.889786936
##         150         152         153         154         155         156
##  0.244990458  0.727840370 -1.442303972 -0.947537326 -1.659642585 -0.409810576
##         157         158         159         160         161         162
##  0.410119514  0.759005836  0.151787620  0.057192641 -1.455237342  0.019209227
##         163         164         165         166         167         168
##  0.224351142 -0.804731508  1.085664683 -0.930614715  0.775664562 -0.150316057
##         169         170         171         172         174         175
## -0.383713630 -0.152780662  1.248162171  2.059753194 -0.156571838  0.856677157
##         176         177         178         180         182         183
##  0.711875090 -0.293368677 -0.491593310  1.133401565 -0.813220184 -0.480830451
##         184         185         186         187         188         191
```

```
## -1.340692228  1.231367699  0.857764905 -0.147053843  1.184777764 -1.210174011
##          192          194          196          197          198          199
## -0.317207490  0.076736663 -0.389002427  1.546362317 -0.679189629 -0.644706852
##          200          201          203          204          205          207
##  0.759757481  0.258220315  0.366610134 -0.855160144 -0.794326991  0.543137732
##          208          209          210          211          212          213
##  2.796265738  0.885723929 -0.933244249  0.571366991  0.528575444 -0.068971996
##          214          215          216          217          218          221
## -0.100973299 -1.715725255 -0.039111244  0.074012066  0.358787602  0.498908226
##          223          224          225          226          227          228
##  0.360162349 -1.084289032 -1.241390490  0.251315512  0.048828059  0.954845045
##          229          231          232          233          234          235
##  0.330456157  0.770846439  0.315424851  0.534125599 -0.513946015  1.272575437
##          236          237          239          241          242          243
##  0.039692244 -0.082911572  1.405063249 -0.319618495  1.632793502  0.480100221
##          244          245          246          247          250          251
##  0.114538605 -0.386726567 -0.377592535 -0.717791553  0.936350531  1.506080422
##          252          253          254          255          256          257
## -1.522978686 -0.368898920  0.308774204 -1.176613076  0.058667897 -0.418652120
##          258          259          263          265          266          267
##  1.454541697 -1.480380813 -0.255577670  0.376989273  0.240671724 -1.689597634
##          268          269          270          272          273          274
## -0.972619387 -1.483199517 -0.219971333 -0.714585093 -0.738416223  0.121401494
##          275          276          278          279          280          281
## -1.462273184  1.130786877 -0.701085729  0.749970793 -1.309783596 -0.270988415
##          282          283          284          285          286          287
## -0.037773995 -0.639086750 -1.276315555  2.531227797 -0.716680319  0.254089378
##          288          289          290          291          292          293
##  0.374068148  1.118889649  0.840707010  0.575646249 -0.671262426 -1.368010275
##          295          298          299          300          301          302
## -0.713816466 -2.899286090  1.200012750 -1.209049287  0.592153081  0.357721358
##          303          304          305          306          307          308
##  0.917641417 -1.021139992  1.236039216  0.511390016 -0.393092577  0.305998264
##          309          310          311          312          313          314
##  0.995421550  1.086567015 -0.250571201 -0.373769624  0.749568058 -1.206355943
##          315          318          319          322          323          324
## -1.001436329 -0.714276207  1.145087355 -1.643797249 -0.021806078  0.549347475
##          325          326          328          329          331          332
## -0.123792486  0.665597581 -0.481706880 -1.220074099 -0.494999319  1.639575790
##          333          335          336          337          338          339
## -1.671534537  0.365278177 -1.365339314  0.544774705  1.975822420 -0.596619135
##          341          342          343          344          345          346
## -0.773953143  1.170835487 -0.744829401  0.562307168 -0.452930144 -0.083593073
##          348          349          350          351          353          354
## -0.993573632 -1.618830960 -1.392382757 -0.850209132  1.983517946  0.196738102
##          355          357          358          359          361          362
##  0.934305384 -2.798488538  3.378862813 -1.025209406 -1.205517634  0.368492828
##          364          365          366          367          368          369
## -0.407162452  0.513743585  2.731943190  0.339400418  0.866453568 -0.086469417
##          370          371          372          374          375          377
## -0.563896737  0.497495578  0.824842109 -0.960031808  1.801232868  0.415017124
##          378          379          381          382          383          384
##  0.126486208 -0.306940862  1.928533721  1.336761074 -1.727241323  0.187593957
##          385          387          388          389          390          392
```

```
##  0.686564379 -0.968535721 -0.102111983 -1.421029263 -0.580060304 -0.159524076
##          393          394          395          396          397          398
## -1.141754048 -0.820328934 -0.548697514 -0.582657813 -0.595217563  0.377714820
##          399
##  0.418602026
```