Collinearity and Model Selection Application in R

```
library(car)
library(ggplot2)
helpdata <- read.csv("help.csv")
```

Continuing with the regression of number of drinks on substance, age, and their interaction from previous lectures using the HELP data, we will examine multicollinearity.

Collinearity

Although not a statistical assumption, multicollinearity can affect interpretation of regression coefficients and leads to large confidence intervals. Multicollinearity can be detected using a statistic called the variance inflation factor (VIF). For any predictor variable, the square root of the VIF indicates the degree to which the confidence interval for that variable's regression parameter is inflated relative to a model with uncorrelated predictors. VIF values are provided by the `vif()` function in the **car** package. As a general rule, \(\sqrt{vif} > 2\) indicates a multicollinearity problem.

```
lm1 <- lm(i1 ~ substance * age, data=helpdata)
vif(lm1)
```

```
## there are higher-order terms (interactions) in this model
## consider setting type = 'predictor'; see ?vif
```

```
##                      GVIF Df GVIF^(1/(2*Df))
## substance     563.622268  2        4.872446
## age             2.607256  1        1.614700
## substance:age 522.910185  2        4.781969
```

```
sqrt(vif(lm1)) > 2
```

```
## there are higher-order terms (interactions) in this model
## consider setting type = 'predictor'; see ?vif
```

```
##                 GVIF    Df GVIF^(1/(2*Df))
## substance       TRUE FALSE            TRUE
## age            FALSE FALSE           FALSE
## substance:age   TRUE FALSE            TRUE
```

Model Selection

The adjusted \(R^2\) is obtained by using the `summary()` function:

```
summary(lm1)
```

```
##
## Call:
## lm(formula = i1 ~ substance * age, data = helpdata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.653  -9.625  -4.832   5.576 102.891
```

```
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             7.9130     6.7925   1.165   0.2447
## substancecocaine        7.8539    10.1649   0.773   0.4401
## substanceheroin        -2.6009     9.6617  -0.269   0.7879
## age                     0.5571     0.1744   3.195   0.0015 **
## substancecocaine:age   -0.6625     0.2770  -2.391   0.0172 *
## substanceheroin:age    -0.4504     0.2653  -1.698   0.0902 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.7 on 447 degrees of freedom
## Multiple R-squared:  0.2268, Adjusted R-squared:  0.2181
## F-statistic: 26.22 on 5 and 447 DF,  p-value: < 2.2e-16
```

In the case of normally distributed errors and the linear model, Mallows' \(C_p\) is equivalent to the AIC. The AIC is obtained using the `AIC()` function:

```
AIC(lm1)
```

```
## [1] 3897.112
```

The BIC is obtained using the `BIC()` function:

```
BIC(lm1)
```

```
## [1] 3925.924
```

To compute the MSE (mean squared error):

```
mean(residuals(lm1)^2)
```

```
## [1] 309.2391
```

To calculate LOOCV for a linear model:

```
loocv <- mean((residuals(lm1)/(1-hatvalues(lm1)))^2)
loocv
```

```
## [1] 317.771
```

To perform k-fold cross-validation, use the `cv.m()` function from the `DAAG` package.

```
library(DAAG)
```

```
cv.lm(data=helpdata, form.lm=lm1, m=10, plotit = F)
```

```
##
## fold 1
## Observations in test set: 45
##                    14        16        24        37         51        54
## Predicted    9.443399  8.298460 12.499756  8.298460 12.0781898 34.095625
## cvpred       9.125682  8.806993 13.278575  8.806993 12.7057618 30.365685
## i1          13.000000  0.000000  4.000000  3.000000 13.0000000 37.000000
## CV residual  3.874318 -8.806993 -9.278575 -5.806993  0.2942382  6.634315
##                   67        75        80       102       134       139
## Predicted   13.02671  24.62532  24.06824 27.410704  12.81593 10.111566
## cvpred      13.99459  27.03336  26.83734 28.013457  13.70818  9.539215
## i1          12.00000   0.00000   1.00000 36.000000   2.00000  8.000000
## CV residual -1.99459 -27.03336 -25.83734  7.986543 -11.70818 -1.539215
```

```
##                   145        181        195       206       207       208
## Predicted   10.111566  9.0450330 12.3943647  12.07819 12.078190 30.753164
## cvpred       9.539215  9.1084961 13.1353714  12.70576 12.705762 29.189571
## i1           0.000000  9.0000000 13.0000000   0.00000  4.000000 26.000000
## CV residual -9.539215 -0.1084961 -0.1353714 -12.70576 -8.705762 -3.189571
##                   224       230       256       262       267       285
## Predicted    36.32393  12.18358  8.725073  30.19609  8.618420 11.762015
## cvpred       31.14976  12.84897  8.979280  28.99355  8.936209 12.276152
## i1          102.00000   1.00000  0.000000  14.00000  0.000000  3.000000
## CV residual  70.85024 -11.84897 -8.979280 -14.99355 -8.936209 -9.276152
##                   286       288       304       309        314       316
## Predicted    39.10932  9.578300 24.625320 26.853627 11.8674065 11.762015
## cvpred       32.12986  9.323856 27.033363 27.817438 12.4193555 12.276152
## i1          142.00000 64.000000 18.000000 32.000000 13.0000000  3.000000
## CV residual 109.87014 54.676144 -9.033363  4.182562  0.5806445 -9.276152
##                   327       337       344       348       352       366
## Predicted   13.02671  26.29655  9.471646  7.658540  12.18358  9.364993
## cvpred      13.99459  27.62142  9.280784  8.548562  12.84897  9.237712
## i1           6.00000  13.00000  4.000000  6.000000   0.00000  6.000000
## CV residual -7.99459 -14.62142 -5.280784 -2.548562 -12.84897 -3.237712
##                   379       387       397       408       414       415
## Predicted   12.499756 37.438085 29.639011 22.954090  7.978500  30.75316
## cvpred      13.278575 31.541798 28.797533 26.445306  8.677777  29.18957
## i1           6.000000 38.000000 20.000000 20.000000  6.000000  10.00000
## CV residual -7.278575  6.458202 -8.797533 -6.445306 -2.677777 -19.18957
##                   429       437       448
## Predicted   33.53855  8.085153  7.871846
## cvpred      30.16967  8.720849  8.634705
## i1          51.00000  0.000000  0.000000
## CV residual 20.83033 -8.720849 -8.634705
##
## Sum of squares = 24739.61    Mean square = 549.77    n = 45
##
## fold 2
## Observations in test set: 46
##                    18        22        26        33        36       40
## Predicted    12.71054  12.28897 12.183581  12.71054  29.08193 11.23506
## cvpred       13.17674  12.52171 12.357954  13.17674  30.36207 10.88414
## i1            1.00000   0.00000  3.000000   0.00000   2.00000 24.00000
## CV residual -12.17674 -12.52171 -9.357954 -13.17674 -28.36207 13.11586
##                    44          71        73        78        79        81
## Predicted   24.6253201 12.60514800  7.658540  12.07819  37.43808  12.39436
## cvpred      24.8791309 13.01297962  7.559287  12.19420  40.64258  12.68547
## i1          25.0000000 13.00000000  2.000000   1.00000  19.00000   2.00000
## CV residual  0.1208691 -0.01297962 -5.559287 -11.19420 -21.64258 -10.68547
##                   88        97       111       150       156       159
## Predicted   12.28897  27.41070 11.867406  31.31024  39.66639 38.552238
## cvpred      12.52171  28.30597 11.866684  33.10354  43.38405 42.013316
## i1           6.00000   9.00000 19.000000  19.00000  12.00000 41.000000
## CV residual -6.52171 -19.30597  7.133316 -14.10354 -31.38405 -1.013316
##                   161       180       185       186       188       204
## Predicted    8.405113  8.831726  35.20978 31.86732  8.405113 27.967781
## cvpred       8.091373  8.395422  37.90111 33.78891  8.091373 28.991335
## i1           0.000000 58.000000  19.00000 58.00000  0.000000 23.000000
```

```
## CV residual -8.091373 49.604578 -18.90111 24.21109 -8.091373 -5.991335
##                   211       214       238       246       252       253
## Predicted    32.98147  25.73947  12.60515  9.045033  30.75316  8.725073
## cvpred       35.15964  26.24987  13.01298  8.547446  32.41817  8.319410
## i1           10.00000  15.00000   3.00000  5.000000  10.00000  0.000000
## CV residual -25.15964 -11.24987 -10.01298 -3.547446 -22.41817 -8.319410
##                   261       273       274       294       296       307
## Predicted    8.938380 20.725783  27.41070 25.739474  9.684953  9.898260
## cvpred       8.471434 20.081559  28.30597 26.249866  9.003520  9.155544
## i1           3.000000 29.000000   0.00000 35.000000  0.000000 26.000000
## CV residual -5.471434  8.918441 -28.30597  8.750134 -9.003520 16.844456
##                   320       324       338       350       378       393       400
## Predicted   12.07819 12.28897  8.085153  8.938380  29.63901 30.753164  8.618420
## cvpred      12.19420 12.52171  7.863336  8.471434  31.04744 32.418172  8.243397
## i1          58.00000 10.00000  0.000000 36.000000   6.00000 42.000000 26.000000
## CV residual 45.80580 -2.52171 -7.863336 27.528566 -25.04744  9.581828 17.756603
##                   410       413        444
## Predicted    30.75316 30.753164 26.2965503
## cvpred       32.41817 32.418172 26.9352331
## i1            8.00000 28.000000 26.0000000
## CV residual -24.41817 -4.418172 -0.9352331
##
## Sum of squares = 14874.99    Mean square = 323.37    n = 46
##
## fold 3
## Observations in test set: 46
##                     7        9        12        13        53        56
## Predicted   10.602707 35.76685  40.22347  40.22347 11.129665  9.258340
## cvpred      11.199962 36.38217  40.81960  40.81960 11.622524  9.158864
## i1          13.000000 71.00000  13.00000  20.00000  3.000000 13.000000
## CV residual  1.800038 34.61783 -27.81960 -20.81960 -8.622524  3.841136
##                    65        77        82        92        96       123
## Predicted    9.364993  8.085153  21.28286 11.551232 12.921323 12.28897
## cvpred       9.288062  7.737685  21.96053 11.960573 13.059234 12.55216
## i1           0.000000  6.000000   0.00000  2.000000 19.000000  3.00000
## CV residual -9.288062 -1.737685 -21.96053 -9.960573  5.940766 -9.55216
##                   144       148       170        171       172       176
## Predicted    12.49976  7.978500  8.405113 11.7620148 12.28897  25.18240
## cvpred       12.72118  7.608487  8.125279 12.1295982 12.55216  25.84328
## i1            0.00000  0.000000  1.000000 13.0000000 10.00000   6.00000
## CV residual -12.72118 -7.608487 -7.125279  0.8704018 -2.55216 -19.84328
##                   179       182       187       194       201       223
## Predicted    7.658540  9.364993  8.831726 11.55123  29.08193  31.31024
## cvpred       7.220893  9.288062  8.642072 11.96057  29.72603  31.94474
## i1           0.000000  0.000000 32.000000 38.00000  16.00000  16.00000
## CV residual -7.220893 -9.288062 23.357928 26.03943 -13.72603 -15.94474
##                   225       232       241       255       269       301
## Predicted    11.17810  9.684953 29.639011 27.967781 35.209778 27.410704
## cvpred       11.48443  9.675656 30.280708 28.616671 35.827494 28.061993
## i1            1.00000 24.000000 28.000000 33.000000 42.000000 26.000000
## CV residual -10.48443 14.324344 -2.280708  4.383329  6.172506 -2.061993
##                   343       347       351       363       365       367
## Predicted     9.54879  13.02671  8.191806 31.867318 11.551232  7.658540
## cvpred       10.35484  13.14375  7.866883 32.499422 11.960573  7.220893
```

```
## i1            0.00000   3.00000 18.000000 38.000000  6.000000  0.000000
## CV residual -10.35484 -10.14375 10.133117  5.500578 -5.960573 -7.220893
##                   368       371       385       391       404        406
## Predicted    9.471646 28.52486  8.191806  31.86732  7.871846 22.9540900
## cvpred       9.417260 29.17135  7.866883  32.49942  7.479289 23.6245632
## i1           0.000000 24.00000 19.000000  19.00000 13.000000 24.0000000
## CV residual -9.417260 -5.17135 11.133117 -13.49942  5.520711  0.3754368
##                   425       426       436       445
## Predicted   13.026715 12.078190  8.831726  25.18240
## cvpred      13.143746 12.383135  8.642072  25.84328
## i1          12.000000  4.000000 32.000000   4.00000
## CV residual -1.143746 -8.383135 23.357928 -21.84328
##
## Sum of squares = 8130.77    Mean square = 176.76    n = 46
##
## fold 4
## Observations in test set: 46
##                      5        10       23        34        66        76
## Predicted   12.394365  9.471646 26.85363  12.28897 31.310241  7.978500
## cvpred      12.768086  9.423125 26.72823  12.66023 31.284363  7.760946
## i1          10.000000 20.000000 34.00000   0.00000 26.000000 21.000000
## CV residual -2.768086 10.576875  7.27177 -12.66023 -5.284363 13.239054
##                   116        121       127       132       140        143
## Predicted    12.07819 25.7394736 31.86732  8.831726  11.65662 12.3943647
## cvpred       12.44451 25.5891967 31.85388  8.710763  12.01308 12.7680856
## i1            1.00000 26.0000000 53.00000  4.000000   0.00000 12.0000000
## CV residual -11.44451  0.4108033 21.14612 -4.710763 -12.01308 -0.7680856
##                   154       163       196       203       210       215
## Predicted   12.60515  8.298460  25.73947 20.168706 11.235057 12.288973
## cvpred      12.98380  8.117127  25.58920 19.894031 11.581653 12.660228
## i1          29.00000 19.000000   0.00000 26.000000 13.000000 19.000000
## CV residual 16.01620 10.882873 -25.58920  6.105969  1.418347  6.339772
##                   218       219        220       228       229       278
## Predicted    12.07819  8.618420 13.1321062 12.183581 12.60515  31.86732
## cvpred       12.44451  8.473309 13.5230881 12.552371 12.98380  31.85388
## i1            2.00000 13.000000 14.0000000  4.000000 27.00000  19.00000
## CV residual -10.44451  4.526691  0.4769119 -8.552371 14.01620 -12.85388
##                   279       281        282       287       303       308
## Predicted    24.06824 25.73947 12.0781898 27.96778  26.29655 10.324873
## cvpred       23.88065 25.58920 12.4445131 27.86726  26.15871 10.372941
## i1           13.00000 67.00000 13.0000000 53.00000   3.00000  4.000000
## CV residual -10.88065 41.41080  0.5554869 25.13274 -23.15871 -6.372941
##                   323       325       341       356       357       364
## Predicted   12.394365  7.978500 12.499756  33.53855  8.085153  24.62532
## cvpred      12.768086  7.760946 12.875943  33.56243  7.879673  24.45016
## i1           6.000000  0.000000  5.000000   6.00000  6.000000  12.00000
## CV residual -6.768086 -7.760946 -7.875943 -27.56243 -1.879673 -12.45016
##                   369       372       374       388       394        409
## Predicted   7.6585397 32.42439 28.52486 29.639011  12.92132 25.1823969
## cvpred      7.4047654 32.42340 28.43678 29.575813  13.30737 25.0196801
## i1          8.0000000 51.00000 73.00000 26.000000   1.00000 26.0000000
## CV residual 0.5952346 18.57660 44.56322 -3.575813 -12.30737  0.9803199
##                   417       434       442       446
## Predicted   12.710540  27.96778  32.42439  7.978500
```

```
## cvpred      13.091658  27.86726  32.42340  7.760946
## i1           4.000000   7.00000  22.00000  3.000000
## CV residual -9.091658 -20.86726 -10.42340 -4.760946
##
## Sum of squares = 10229.57    Mean square = 222.38    n = 46
##
## fold 5
## Observations in test set: 45
##                      6        11       15        17        27       35        42
## Predicted   10.813490  8.938380 27.96778  9.045033  9.045033 19.05455  11.02427
## cvpred      10.760125  9.668888 27.63973  9.806874  9.806874 17.70685  10.95277
## i1           4.000000  0.000000 51.00000  0.000000  7.000000 32.00000   0.00000
## CV residual -6.760125 -9.668888 23.36027 -9.806874 -2.806874 14.29315 -10.95277
##                     61         68        83       131       138       146
## Predicted   27.967781   9.684953 27.410704 35.20978 12.39436  8.511766
## cvpred      27.639731  10.634787 27.018926 35.71020 12.20495  9.116946
## i1          35.000000   0.000000 26.000000 64.00000 38.00000  1.000000
## CV residual  7.360269 -10.634787 -1.018926 28.28980 25.79505 -8.116946
##                    149       152       158       164       173       178
## Predicted   34.652701  11.97280  37.99516 11.867406 12.499756  11.97280
## cvpred      35.089394  11.81967  38.81423 11.723345 12.301277  11.81967
## i1          26.000000   1.00000  19.00000  8.000000  3.000000   1.00000
## CV residual -9.089394 -10.81967 -19.81423 -3.723345 -9.301277 -10.81967
##                    184       212       216       227       254       260
## Predicted   26.296550 29.639011  11.65662 11.44584  34.09562  8.938380
## cvpred      25.777315 29.502147  11.53070 11.33806  34.46859  9.668888
## i1          33.000000 23.000000   0.00000 27.00000  24.00000  0.000000
## CV residual  7.222685 -6.502147 -11.53070 15.66194 -10.46859 -9.668888
##                    271       280       306       311       312       318
## Predicted    9.898260  11.65662  33.53855 30.196087 12.18358  29.63901
## cvpred      10.910758  11.53070  33.84778 30.122952 12.01231  29.50215
## i1           6.000000   1.00000  12.00000 34.000000 38.00000  18.00000
## CV residual -4.910758 -10.53070 -21.84778  3.877048 25.98769 -11.50215
##                    321       328       334       340       342       345       354
## Predicted    10.96479  8.405113 27.967781 11.44584 12.815931 11.97280  8.831726
## cvpred       12.29061  8.978961 27.639731 11.33806 12.590243 11.81967  9.530903
## i1            2.00000  0.000000 25.000000 26.00000 10.000000 29.00000 13.000000
## CV residual -10.29061 -8.978961 -2.639731 14.66194 -2.590243 17.18033  3.469097
##                    375       386       398       405        424       440
## Predicted    9.898260  9.364993  7.551886 30.75316 12.4997564 29.63901
## cvpred      10.910758 10.220830  7.875077 30.74376 12.3012767 29.50215
## i1           9.000000  3.000000  0.000000 51.00000 12.0000000 41.00000
## CV residual -1.910758 -7.220830 -7.875077 20.25624 -0.3012767 11.49785
##                    452
## Predicted    9.548790
## cvpred       9.604261
## i1          13.000000
## CV residual  3.395739
##
## Sum of squares = 7127.94    Mean square = 158.4    n = 45
##
## fold 6
## Observations in test set: 45
##                      25        29        38        47        48       60
```

```
## Predicted   10.708098  12.71054 26.296550 10.004913 11.445840 12.07819
## cvpred      10.150043  11.20273 24.797516 10.747188 10.537874 10.87030
## i1           6.000000   0.00000 27.000000 15.000000  7.000000 32.00000
## CV residual -4.150043 -11.20273  2.202484  4.252812 -3.537874 21.12970
##                    62        74        84       90       94        95
## Predicted   22.9540900  12.49976  8.725073 26.29655 11.86741 13.026715
## cvpred      20.7322563  11.09192  8.756954 24.79752 10.75949 11.368941
## i1          20.0000000 102.00000  0.000000 64.00000 61.00000  2.000000
## CV residual -0.7322563  90.90808 -8.756954 39.20248 50.24051 -9.368941
##                   104       129      142      162      175      183       189
## Predicted    10.00491 27.410704 25.73947 26.29655 12.28897 12.81593  12.18358
## cvpred       10.74719 26.152603 24.11997 24.79752 10.98111 11.25813  10.92571
## i1            0.00000 25.000000 39.00000 59.00000 19.00000 35.00000   0.00000
## CV residual -10.74719 -1.152603 14.88003 34.20248  8.01889 23.74187 -10.92571
##                   192       217      221      243      249       259       265
## Predicted   21.2828598 10.004913 34.65270 11.86741 12.92132  8.405113  12.39436
## cvpred      18.6996263 10.747188 34.96067 10.75949 11.31354  8.259396  11.03651
## i1          18.0000000 13.000000 51.00000 51.00000 26.00000  0.000000   0.00000
## CV residual -0.6996263  2.252812 16.03933 40.24051 14.68646 -8.259396 -11.03651
##                   270       272       275       290       319       326
## Predicted   10.111566 11.762015  7.765193  9.471646  7.765193  28.52486
## cvpred      10.913041 10.704088  7.264279  9.917924  7.264279  27.50769
## i1           2.000000 19.000000  0.000000  2.000000  0.000000   6.00000
## CV residual -8.913041  8.295912 -7.264279 -7.917924 -7.264279 -21.50769
##                   329       339       353       383       395       396
## Predicted   24.625320 12.605148  7.551886 10.813490 12.815931 29.639011
## cvpred      22.764886 11.147323  6.932573 10.205448 11.258132 28.862776
## i1          32.000000  2.000000 45.000000  1.000000 18.000000 35.000000
## CV residual  9.235114 -9.147323 38.067427 -9.205448  6.741868  6.137224
##                   401       402      411      428      438       439       450
## Predicted   24.06824  32.98147 26.29655 12.28897 29.08193  36.88101  35.20978
## cvpred      22.08734  32.92804 24.79752 10.98111 28.18523  37.67084  35.63821
## i1          43.00000  19.00000 61.00000  3.00000 76.00000  26.00000  13.00000
## CV residual 20.91266 -13.92804 36.20248 -7.98111 47.81477 -11.67084 -22.63821
##
## Sum of squares = 25163.73    Mean square = 559.19    n = 45
##
## fold 7
## Observations in test set: 45
##                     4         8        20        46        49        52
## Predicted    9.471646  23.51117 12.921323  27.96778 8.5117664 28.524857
## cvpred       9.834785  22.90656 13.183413  27.17643 8.5357865 27.710164
## i1           5.000000  12.00000 23.000000  13.00000 9.0000000 34.000000
## CV residual -4.834785 -10.90656  9.816587 -14.17643 0.4642135  6.289836
##                    57        58       64       112       117       118
## Predicted   12.183581  30.75316 30.19609  30.75316 30.753164 11.8674065
## cvpred      12.409944  29.84510 29.31136  29.84510 29.845097 12.0784571
## i1           3.000000   3.00000 59.00000  13.00000 22.000000 13.0000000
## CV residual -9.409944 -26.84510 29.68864 -16.84510 -7.845097  0.9215429
##                   120       125       133       137      147      157       165
## Predicted   29.63901  8.831726  7.871846  11.86741 27.96778 10.32487 12.815931
## cvpred      28.77763  8.968786  7.669787  12.07846 27.17643 10.98945 13.072917
## i1          19.00000  0.000000  3.000000   1.00000 19.00000  4.00000 16.000000
## CV residual -9.77763 -8.968786 -4.669787 -11.07846 -8.17643 -6.98945  2.927083
```

```
##                    166       174       193       198       205       222
## Predicted   12.3943647  9.471646  8.61842  7.338580  25.73947 12.078190
## cvpred      12.6309348  9.834785  8.68012  6.948122  25.04150 12.299448
## i1          12.0000000 20.000000  0.00000 27.000000  13.00000 10.000000
## CV residual -0.6309348 10.165215 -8.68012 20.051878 -12.04150 -2.299448
##                    231       237       242       244       266       268
## Predicted   29.08193  9.151686 11.656623  31.86732  8.085153 29.63901
## cvpred      28.24390  9.401786 11.857466  30.91256  7.958454 28.77763
## i1          54.00000 16.000000 13.000000 134.00000  1.000000 25.00000
## CV residual 25.75610  6.598214  1.142534 103.08744 -6.958454 -3.77763
##                    276       289       291       298       299       335
## Predicted     9.791606  7.658540 34.09562 12.288973  8.511766  26.85363
## cvpred       10.267785  7.381121 33.04750 12.520439  8.535787  26.10896
## i1            0.000000  0.000000 51.00000  6.000000 12.000000  13.00000
## CV residual -10.267785 -7.381121 17.95250 -6.520439  3.464213 -13.10896
##                    380       389       403       407       423       430
## Predicted    8.725073 25.73947  13.13211  7.551886 11.762015 12.288973
## cvpred       8.824453 25.04150  13.40440  7.236788 11.967962 12.520439
## i1           2.000000 83.00000   1.00000 13.000000 13.000000  5.000000
## CV residual -6.824453 57.95850 -12.40440  5.763212  1.032038 -7.520439
##                    433       447
## Predicted   26.2965503 25.18240
## cvpred      25.5752302 24.50776
## i1          26.0000000 56.00000
## CV residual  0.4247698 31.49224
##
## Sum of squares = 20441.41    Mean square = 454.25    n = 45
##
## fold 8
## Observations in test set: 45
##                     19        30        32        39        45        50
## Predicted   8.1918064 27.41070 11.867406  9.578300 12.183581  28.52486
## cvpred      8.2496293 28.26633 11.462405  9.840242 11.707278  29.39748
## i1          9.0000000 20.00000  6.000000  3.000000  6.000000   5.00000
## CV residual 0.7503707 -8.26633 -5.462405 -6.840242 -5.707278 -24.39748
##                    55        59        63        70        85        93
## Predicted   29.081934  8.405113  27.96778 12.815931 10.324873 30.196087
## cvpred      29.963062  8.494339  28.83191 12.197023 10.696726 31.094217
## i1          36.000000  0.000000   7.00000  6.000000  9.000000 33.000000
## CV residual  6.036938 -8.494339 -21.83191 -6.197023 -1.696726  1.905783
##                   101       103       106       113       128       135
## Predicted    8.831726 29.081934 12.288973  8.298460 13.237498  32.98147
## cvpred       8.983758 29.963062 11.788902  8.371984 12.523519  33.92210
## i1           0.000000 31.000000  2.000000  0.000000  7.000000  13.00000
## CV residual -8.983758  1.036938 -9.788902 -8.371984 -5.523519 -20.92210
##                   155       160       169       191       247       250
## Predicted   39.666392  8.831726  29.63901  10.81349  25.73947  24.62532
## cvpred      40.709032  8.983758  30.52864  10.64616  26.56960  25.43844
## i1          38.000000  1.000000  12.00000   0.00000   3.00000  15.00000
## CV residual -2.709032 -7.983758 -18.52864 -10.64616 -23.56960 -10.43844
##                   292       293       300       310       317       330
## Predicted    12.39436 30.753164  25.73947 10.813490 12.92132  34.09562
## cvpred       11.87053 31.659794  26.56960 10.646163 12.27865  35.05326
## i1            1.00000 24.000000   7.00000 13.000000 49.00000   6.00000
```

```
## CV residual -10.87053 -7.659794 -19.56960  2.353837 36.72135 -29.05326
##                   333       355      360       373      376       381      384
## Predicted    13.34289  9.791606 10.43153 29.081934 29.63901 31.310241 13.23750
## cvpred       12.60514 10.084952 10.81908 29.963062 30.52864 32.225372 12.52352
## i1            0.00000  4.000000 13.00000 35.000000 51.00000 26.000000 49.00000
## CV residual -12.60514 -6.084952  2.18092  5.036938 20.47136 -6.225372 36.47648
##                   390       392       412       420       422       427       431
## Predicted    8.191806 12.49976 10.538180 27.967781  13.23750 13.026715 11.65662
## cvpred       8.249629 11.95215 10.941435 28.831907  12.52352 12.360271 11.29916
## i1          32.000000 30.00000 13.000000 26.000000   0.00000 12.000000 68.00000
## CV residual 23.750371 18.04785  2.058565 -2.831907 -12.52352 -0.360271 56.70084
##                   432
## Predicted   22.954090
## cvpred      23.741711
## i1          29.000000
## CV residual  5.258289
##
## Sum of squares = 12385.61    Mean square = 275.24    n = 45
##
## fold 9
## Observations in test set: 45
##                    2        41        43        86        87        89
## Predicted   28.52486 12.815931  9.791606 13.026715 10.111566 32.981471
## cvpred      28.47922 13.357226  8.715039 13.636727  8.837292 33.389348
## i1          56.00000  6.000000 13.000000 10.000000  4.000000 26.000000
## CV residual 27.52078 -7.357226  4.284961 -3.636727 -4.837292 -7.389348
##                   99       105       107       109       114       126
## Predicted   26.296550 23.511167  12.60515 25.739474  8.511766  11.76201
## cvpred      26.024155 22.955325  13.07772 25.410389  8.226030  11.95972
## i1          18.000000 26.000000   0.00000 34.000000  0.000000   0.00000
## CV residual -8.024155  3.044675 -13.07772  8.589611 -8.226030 -11.95972
##                   153       168       197       200       202       213       226
## Predicted    30.75316 27.96778 29.63901  8.191806  12.39436 10.324873 12.394365
## cvpred       30.93428 27.86545 29.70675  8.103777  12.79822  8.918793 12.798223
## i1           12.00000 50.00000 46.00000 12.000000   1.00000 42.000000  6.000000
## CV residual -18.93428 22.13455 16.29325  3.896223 -11.79822 33.081207 -6.798223
##                   233       234       236       239       245       251
## Predicted    25.18240 24.068243 11.024273 11.55123  19.05455 8.1918064
## cvpred       24.79662 23.569091 10.981464 11.68022  18.04520 8.1037775
## i1           10.00000 30.000000  2.000000 34.00000   5.00000 9.0000000
## CV residual -14.79662  6.430909 -8.981464 22.31978 -13.04520 0.8962225
##                   257       263       283       284       295       313
## Predicted    8.8317263 10.111566  7.658540  9.471646 28.52486  34.65270
## cvpred       8.3482822  8.837292  7.900024  8.592787 28.47922  35.23065
## i1           8.0000000 12.000000 20.000000  0.000000 67.00000   0.00000
## CV residual -0.3482822  3.162708 12.099976 -8.592787 38.52078 -35.23065
##                   322       346        349       358       359       382
## Predicted   12.499756  34.09562 11.9727981 27.967781  34.09562  12.28897
## cvpred      12.937974  34.61688 12.2392204 27.865453  34.61688  12.65847
## i1           6.000000  20.00000 13.0000000 25.000000  21.00000   0.00000
## CV residual -6.937974 -14.61688  0.7607796 -2.865453 -13.61688 -12.65847
##                   416       418       419       421       435       441
## Predicted    8.938380 23.511167  8.405113 11.97280  7.765193  9.151686
## cvpred       8.389033 22.955325  8.185279 12.23922  7.940774  8.470535
```

```
## i1           0.000000 25.000000  2.000000 24.00000  5.000000 18.000000
## CV residual -8.389033  2.044675 -6.185279 11.76078 -2.940774  9.529465
##                   443       449
## Predicted    9.258340  8.831726
## cvpred       8.511285  8.348282
## i1          53.000000  0.000000
## CV residual 44.488715 -8.348282
##
## Sum of squares = 10840.45    Mean square = 240.9    n = 45
##
## fold 10
## Observations in test set: 45
##                     1         3        21        28        31        69
## Predicted   11.867406  8.085153  9.684953 26.853627  31.86732 11.235057
## cvpred      11.820977  8.099894  9.556404 26.493057  31.56807 11.465271
## i1          13.000000  0.000000 26.000000 24.000000   3.00000 18.000000
## CV residual  1.179023 -8.099894 16.443596 -2.493057 -28.56807  6.534729
##                    72       91        98      100       108       110       115
## Predicted   11.340448 12.92132  8.405113 29.63901 27.41070 11.34045 12.078190
## cvpred      11.524555 12.41382  8.391196 29.31251 27.05695 11.52456 11.939546
## i1           5.000000 26.00000  0.000000 51.00000 51.00000 39.00000 13.000000
## CV residual -6.524555 13.58618 -8.391196 21.68749 23.94305 27.47544  1.060454
##                   119       122       124       130       136       141       151
## Predicted   12.60515  35.76685 12.92132  35.20978  33.53855 12.183581 12.288973
## cvpred      12.23597  35.51531 12.41382  34.95142  33.25975 11.998831 12.058115
## i1          26.00000  13.00000 24.00000  15.00000  20.00000 13.000000  3.000000
## CV residual 13.76403 -22.51531 11.58618 -19.95142 -13.25975  1.001169 -9.058115
##                   167       177       190       199       209       235       240
## Predicted   25.182397  33.53855 11.44584 10.602707 11.972798 27.96778 12.921323
## cvpred      24.801384  33.25975 11.58384 11.109564 11.880262 27.62084 12.413822
## i1          26.000000 102.00000  6.00000  3.000000 13.000000 43.00000  8.000000
## CV residual  1.198616  68.74025 -5.58384 -8.109564  1.119738 15.37916 -4.413822
##                   248       258       264       277       297       302
## Predicted    11.34045 24.625320  11.86741 12.710540 12.4997564 25.73947
## cvpred       11.52456 24.237493  11.82098 12.295253 12.1766841 25.36527
## i1            0.00000 27.000000   1.00000 22.000000 13.0000000 41.00000
## CV residual -11.52456  2.762507 -10.82098  9.704747  0.8233159 15.63473
##                   305       315       331       332       336       361
## Predicted   32.424394  9.045033  26.29655 11.972798 23.511167  8.725073
## cvpred      32.131964  8.973800  25.92917 11.880262 23.109712  8.682498
## i1          38.000000  0.000000   3.00000  6.000000 18.000000 37.000000
## CV residual  5.868036 -8.973800 -22.92917 -5.880262 -5.109712 28.317502
##                   362       370       377       399       451       453
## Predicted   29.08193 28.524857 10.111566 12.710540  9.471646 32.98147
## cvpred      28.74862 28.184729  9.944807 12.295253  9.362203 32.69585
## i1          25.00000 32.000000  6.000000 11.000000  1.000000 51.00000
## CV residual -3.74862  3.815271 -3.944807 -1.295253 -8.362203 18.30415
##
## Sum of squares = 12419.01    Mean square = 275.98    n = 45
##
## Overall (Sum over all 45 folds)
##       ms
## 323.0752
```

The `step()` function will do stepwise model selection based on AIC, either forward or backward.