Table of Contents

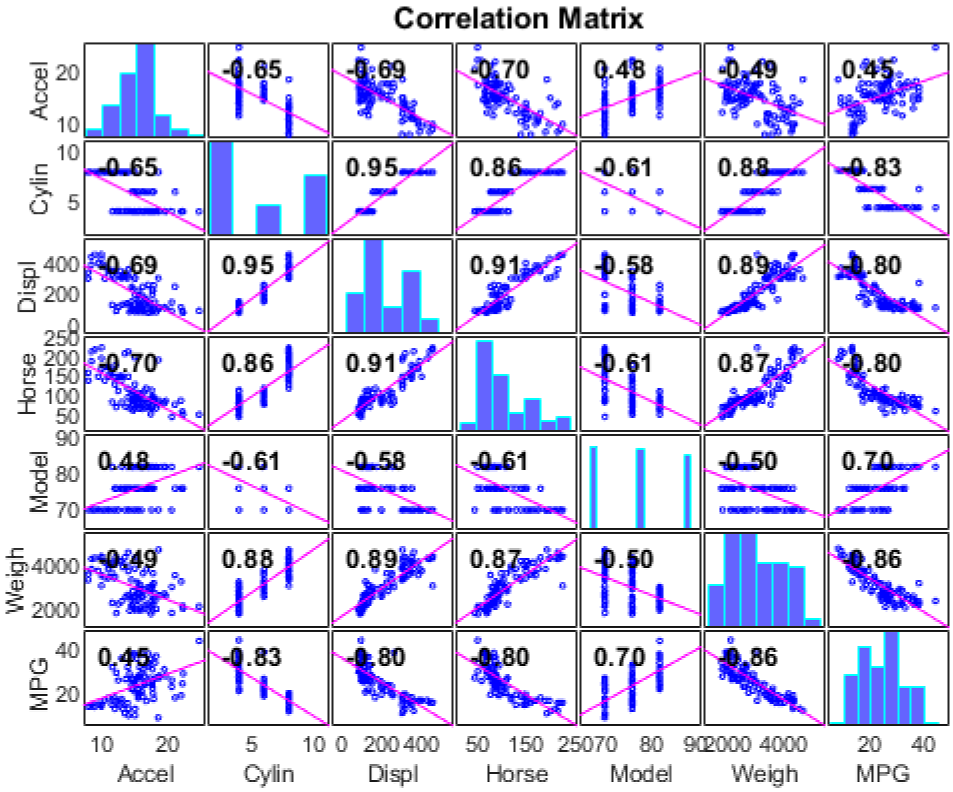
Multiple Linear Regression

Farzad Rezazadeh Pilehdarboni, 25.05.2022

```
close all;clear;clc
```

3a)

```
% Load the data set
load carsmall
ds =
 table(Acceleration,Cylinders,Displacement,Horsepower,Model_Year,Weight,MPG, ...
    'VariableNames',
{'Acceleration','Cylinders','Displacement','Horsepower','Model_Year','Weight','MPG'});

% Scatterplot matrix of data set
figure(1)
corrplot(ds(:,1:7))
```



3b)

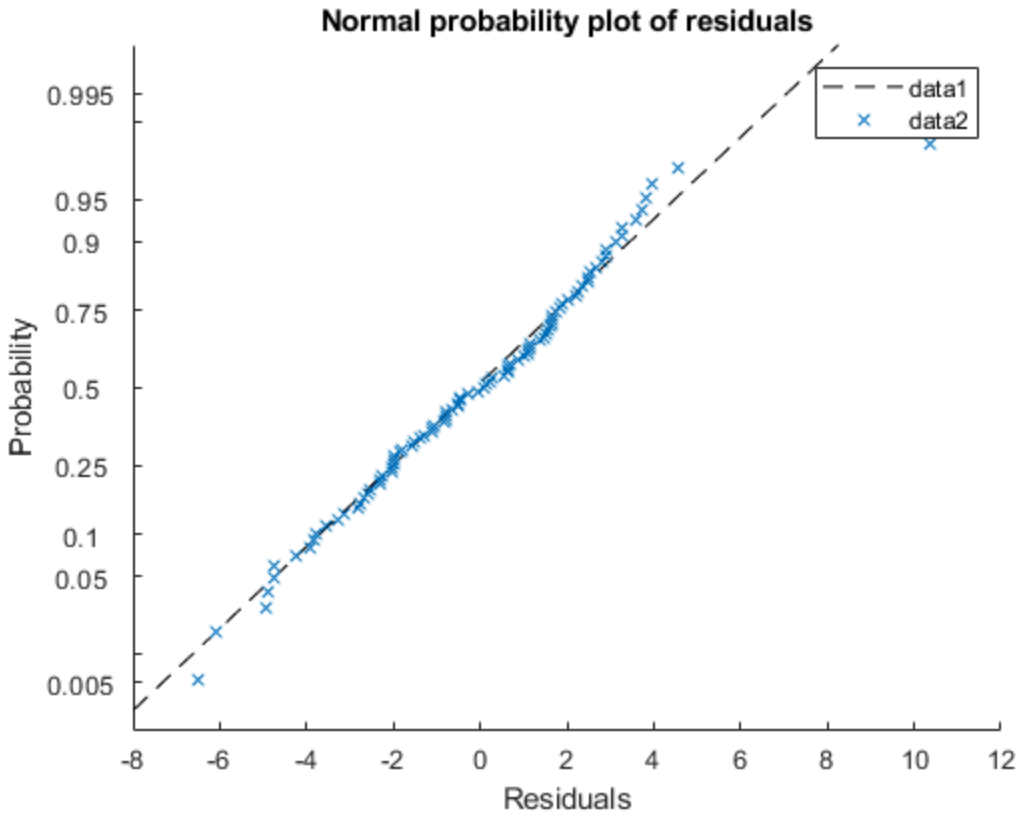
```
% Fit multiple regression model with all predictors
mult_mod = fitlm(ds);
disp(mult_mod)

% Fit multiple regression model with only Model_Year and Weight
mult_mod_2 = fitlm(ds,'MPG ~ 1 + Model_Year + Weight');
disp(mult_mod_2)
```

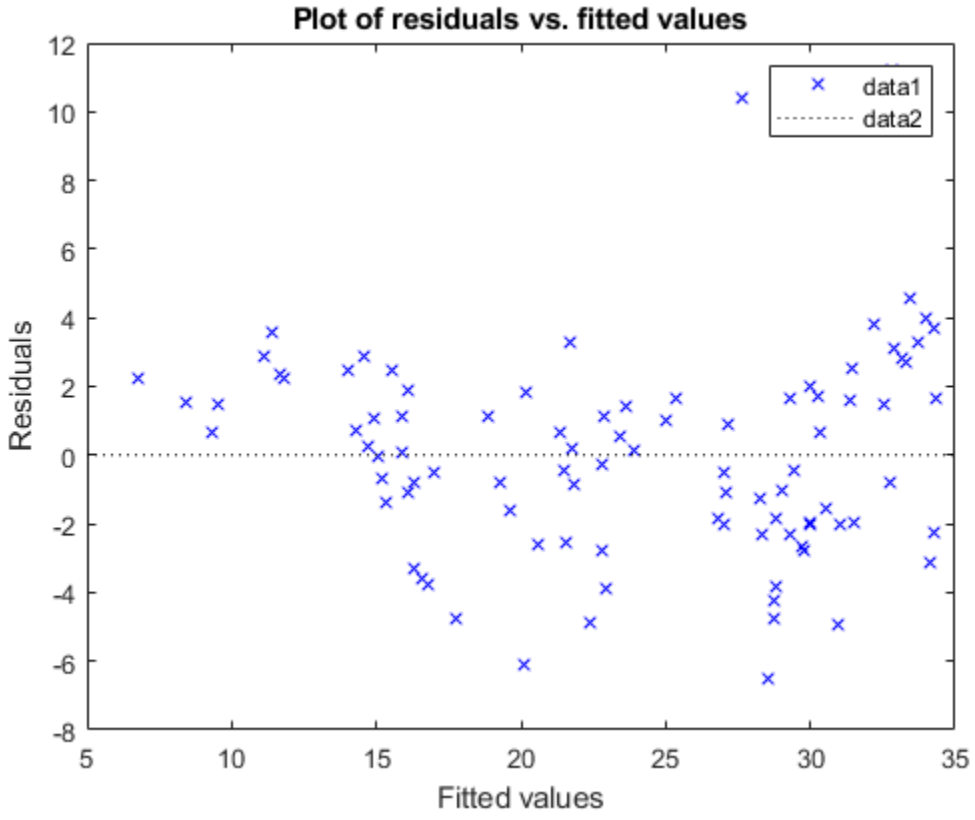
```
Linear regression model:
    MPG ~ [Linear formula with 7 terms in 6 predictors]

Estimated Coefficients:
                     Estimate        SE         tStat       pValue
                    __________    _________    ________    __________

    (Intercept)        -3.4351          8.1    -0.42409       0.67256
    Acceleration     -0.051911       0.1589    -0.32669        0.7447
    Cylinders         -0.83318      0.65105     -1.2797       0.20408
    Displacement      0.012826     0.011866      1.0809       0.28276
    Horsepower      -0.0024624     0.020801    -0.11838       0.90604
    Model_Year         0.65845     0.087631      7.5139    5.0347e-11
    Weight          -0.0066985    0.0011323     -5.9158    6.5402e-08
```

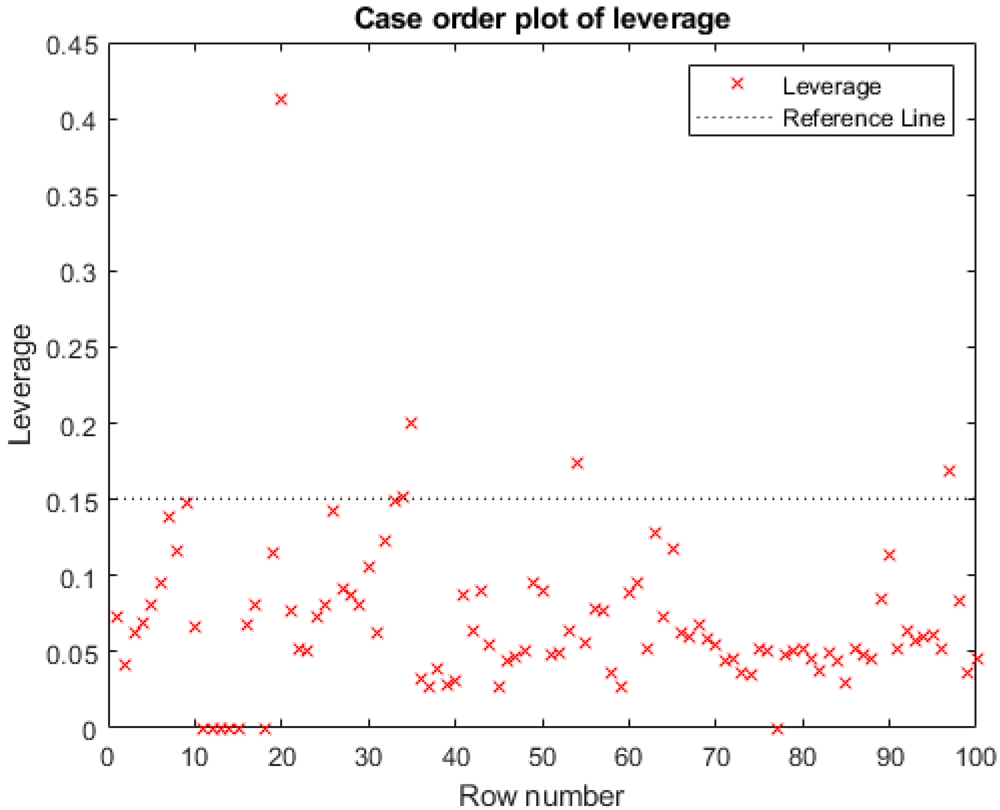
```
Number of observations: 93, Error degrees of freedom: 86
Root Mean Squared Error: 3.08
R-squared: 0.864,  Adjusted R-Squared: 0.855
F-statistic vs. constant model: 91.4, p-value = 3.54e-35

Linear regression model:
    MPG ~ 1 + Model_Year + Weight

Estimated Coefficients:
                   Estimate        SE         tStat        pValue
                   ____________    ____________    ________    __________

    (Intercept)       -5.7045         6.3858    -0.89331       0.37405
    Model_Year        0.65127       0.074313      8.7639    9.8456e-14
    Weight         -0.0068023    0.00044659      -15.232    9.0903e-27


Number of observations: 94, Error degrees of freedom: 91
Root Mean Squared Error: 3.06
R-squared: 0.858,  Adjusted R-Squared: 0.855
F-statistic vs. constant model: 275, p-value = 2.7e-39
```

3c)

```
% Diagnostic plots
figure(2)
plotResiduals(mult_mod,'probability')
legend('show')
figure(3)
plotResiduals(mult_mod,'fitted')
legend('show')
figure(4)
plotDiagnostics(mult_mod)
legend('show')
```


Normal probability plot of residuals
data1
data2
Probability
0.995
0.95
0.9
0.75
0.5
0.25
0.1
0.05
0.005
-8
-6
-4
-2
0
2
4
6
8
10
12
Residuals


Plot of residuals vs. fitted values
data1
data2
Residuals
12
10
8
6
4
2
0
-2
-4
-6
-8
5
10
15
20
25
30
35
Fitted values



3d)

```
% Add interaction terms
mult_mod_inter = fitlm(ds,'interactions');
disp(mult_mod_inter)
```

```
Linear regression model:
    MPG ~ [Linear formula with 22 terms in 6 predictors]

Estimated Coefficients:
                                  Estimate         SE         tStat
 pValue
                                 __________    _________    ________

 ________

    (Intercept)                     -83.816        93.99    -0.89175
 0.37554
    Acceleration                    -4.0957        4.426    -0.92539
 0.3579
    Cylinders                        12.954       17.987     0.72017
 0.47378
    Displacement                   -0.73987      0.42774     -1.7297
 0.088026
    Horsepower                       1.3975      0.65317      2.1396
 0.035823
```

```
    Model_Year                    1.7486       1.0572       1.6541
 0.10252
    Weight                      0.018631     0.039118      0.47628
 0.63534
    Acceleration:Cylinders       0.17396      0.37726      0.46111
 0.64613
    Acceleration:Displacement  -0.00039026    0.0057898   -0.067404
 0.94645
    Acceleration:Horsepower     -0.0035815    0.0080672    -0.44396
 0.65842
    Acceleration:Model_Year      0.054619      0.04961        1.101
 0.27463
    Acceleration:Weight        -0.00031117   0.00039598    -0.78583
 0.43458
    Cylinders:Displacement       0.013537     0.015236     0.88848
 0.37728
    Cylinders:Horsepower        -0.068402      0.04412     -1.5504
 0.1255
    Cylinders:Model_Year         -0.19811      0.23246    -0.85224
 0.39695
    Cylinders:Weight             0.0013104    0.0018601    0.70445
 0.48346
    Displacement:Horsepower     0.00044117   0.00047666    0.92555
 0.35782
    Displacement:Model_Year      0.0095651    0.0051404     1.8608
 0.066918
    Displacement:Weight        -3.3689e-05   3.5083e-05    -0.96025
 0.34019
    Horsepower:Model_Year        -0.01741     0.0083306    -2.0899
 0.040214
    Horsepower:Weight            5.6659e-05   6.3532e-05    0.89181
 0.37551
    Model_Year:Weight          -0.00032913   0.00046387    -0.70953
 0.48032


Number of observations: 93, Error degrees of freedom: 71
Root Mean Squared Error: 2.56
R-squared: 0.922,  Adjusted R-Squared: 0.899
F-statistic vs. constant model: 40.2, p-value = 3.21e-31
```

3e)

```
% Add quadratic terms
transf_mult_mod = fitlm(ds, ...
    'MPG ~ 1 + Acceleration + Horsepower + Horsepower^2 + Weight +
 Displacement + Displacement^2 + Cylinders + Cylinders^2 + Model_Year');
disp(transf_mult_mod)


Linear regression model:
    MPG ~ [Linear formula with 10 terms in 6 predictors]
```

```
Estimated Coefficients:
                       Estimate        SE         tStat        pValue
                      __________    __________    ________    __________

    (Intercept)           9.548        12.542     0.76128       0.44865
    Acceleration        -0.36911        0.1854     -1.9909       0.04978
    Cylinders           -0.45117        3.4131    -0.13219       0.89516
    Displacement        -0.01651      0.037429    -0.44109       0.66029
    Horsepower          -0.22139      0.077023     -2.8743     0.0051413
    Model_Year           0.66384      0.084713      7.8364    1.3761e-11
    Weight              -0.004153     0.0013459     -3.0857      0.002759
    Cylinders^2          0.036757       0.26909      0.1366       0.89168
    Displacement^2     2.5227e-05    6.2183e-05     0.40569       0.68601
    Horsepower^2        0.0006474    0.00024896      2.6005      0.011019


Number of observations: 93, Error degrees of freedom: 83
Root Mean Squared Error: 2.93
R-squared: 0.882,  Adjusted R-Squared: 0.869
F-statistic vs. constant model: 68.6, p-value = 1.1e-34
```

Published with MATLAB® R2022a