Read in the bus cost data, which are described in Problem 5 of Chapter 6 and further examined in Problems 1 to 3 of Chapter 9. The data describe the total operating costs of 246 bus companies, in thousands of dollars, and the total miles driven by each company’s buses, which serves as a measure of output. We’ll read in the data.
and then divide each of these variables by 1000 so that total costs are expressed in millions of dollars and total miles in thousands.

```matlab
M = csvread('buscost.csv',1,0);
Obs = M(:,1);
TotalCost = M(:,2)./1000.0;
RVM = M(:,3)./1000.0;
RVM_SQ = RVM.*RVM;
LogTotalCost = log(TotalCost);
LogRVM = log(RVM);
LogRVM_SQ = LogRVM.*LogRVM;
```

```matlab
plot(RVM,TotalCost,'ko')
xlabel('Miles Driven (in thousands)')
ylabel('Cost (in millions)')
print -depsc BusCost
```

```matlab
plot(LogRVM,LogTotalCost,'ko')
xlabel('Log of Miles Driven')
ylabel('Log of Cost')
print -depsc LogBusCost
```

```matlab
format short g
```

Regression with only the RVM explanatory variable and a constant

```matlab
OutPut = regstats(TotalCost,RVM,'linear')
```

```matlab
OutPut =
```

```matlab
source: 'regstats'
Q: [246x2 double]
R: [2x2 double]
beta: [2x1 double]
covb: [2x2 double]
yhat: [246x1 double]
r: [246x1 double]
mse: 41.442
rsquare: 0.88792
adjrsquare: 0.887
leverage: [246x1 double]
hatmat: [246x246 double]
s2_i: [246x1 double]
beta_i: [2x246 double]
standres: [246x1 double]
studres: [246x1 double]
dfbetas: [2x246 double]
```
Show coefficients, standard errors, and t-stats
[OutPut.tstat.beta OutPut.tstat.se OutPut.tstat.t]

ans =

    -1.5194        0.51296       -2.9621
     5.018         0.11413       43.966

Show s-squared
OutPut.mse

ans =

    41.442

Conduct a Wald test of the hypothesis that the RVM coefficient = 0
% at the 0.01 significance level
beta = OutPut.beta

beta =

    -1.5194
     5.018

Var = OutPut.covb

Var =

    0.26313      -0.035116
   -0.035116       0.013027

R = [0 1];
r = 0;

diff = R*beta - r
diff =

      5.018

Vdiff = R*Var*R'

Vdiff =

      0.013027

% The Wald test statistic is
W = diff'*inv(Vdiff)*diff

W =

      1933

% Show the chi-squared density function for 1 degree of freedom
df = size(R,1);

Sig = 0.01;
Reject = chi2inv(1-Sig,df)

Reject =

      6.6349

% Compare the W statistic to the rejection value
W

W =

      1933

% Regression with RVM and its square RVM_SQ and a constant
% Ask for all type of output from the regstats command
OutPut = regstats(TotalCost,[RVM RVM_SQ],'linear')

OutPut =

      source: 'regstats'
      Q: [246x3 double]
      R: [3x3 double]
beta: [3x1 double]
covb: [3x3 double]
yhat: [246x1 double]
r: [246x1 double]
mse: 41.555
rsquare: 0.88807
adjrsquare: 0.88669
leverage: [246x1 double]
hatmat: [246x246 double]
s2_i: [246x1 double]
beta_i: [3x246 double]
standres: [246x1 double]
studres: [246x1 double]
dfbetas: [3x246 double]
dfit: [246x1 double]
dffits: [246x1 double]
covratio: [246x1 double]
cookd: [246x1 double]
tstat: [1x1 struct]
fstat: [1x1 struct]

% Show coefficients, standard errors, and t-stats
[OutPut.tstat.beta OutPut.tstat.se OutPut.tstat.t]

ans =

-1.3056  0.63224  -2.065
4.8505   0.31046  15.624
0.01176  0.020273 0.58007

% Show s-squared
OutPut.mse

ans =

41.555

% Conduct a Wald test of the hypothesis that both RVM coefficients = 0
beta = OutPut.beta

beta =

-1.3056
4.8505
0.01176

Var = OutPut.covb

Var =

\[
\begin{bmatrix}
0.39973 & -0.14162 & 0.0074732 \\
-0.14162 & 0.096383 & -0.0058518 \\
0.0074732 & -0.0058518 & 0.00041098
\end{bmatrix}
\]

R = [0 1 0; 0 0 1];
r = [0;0] ;

diff = R*beta - r

diff =

\[
\begin{bmatrix}
4.8505 \\
0.01176
\end{bmatrix}
\]

Vdiff = R*Var*R'

Vdiff =

\[
\begin{bmatrix}
0.096383 & -0.0058518 \\
-0.0058518 & 0.00041098
\end{bmatrix}
\]

% The Wald test statistic is
W = diff'*inv(Vdiff)*diff

W =

1928.1

% Use the chi-squared density function for 2 degrees of freedom
% to test this hypothesis at the 0.01 significance level
df = size(R,1)

df =

2
Sig = 0.01;
Reject = chi2inv(1-Sig,df)

Reject =

9.2103

% Compare the W statistic to the rejection value
W

W =

1928.1

% Now repeat the analysis above using the log versions of total
% costs and miles driven
%
% Regression with only the LogRVM explanatory variable and a constant
% Ask for all type of output from the regstats command
OutPut = regstats(LogTotalCost,LogRVM,'linear')

OutPut =

source: 'regstats'
   Q: [246x2 double]
   R: [2x2 double]
   beta: [2x1 double]
   covb: [2x2 double]
   yhat: [246x1 double]
   r: [246x1 double]
   mse: 0.080412
   rsquare: 0.95909
   adjrsquare: 0.95876
   leverage: [246x1 double]
   hatmat: [246x246 double]
   s2_i: [246x1 double]
   beta_i: [2x246 double]
   standres: [246x1 double]
   studres: [246x1 double]
   dfbetas: [2x246 double]
   dffit: [246x1 double]
   dffits: [246x1 double]
   covratio: [246x1 double]
   cookd: [246x1 double]
tstat: [1x1 struct]
fstat: [1x1 struct]

% Show coefficients, standard errors, and t-stats
[OutPut.tstat.beta OutPut.tstat.se OutPut.tstat.t]

ans =
   1.241    0.018566    66.842
   1.1183    0.014785    75.633

% Show s-squared
OutPut.mse

ans =
   0.080412

% Conduct a Wald test of the hypothesis that the RVM coefficient = 0
% at the 0.01 significance level
beta = OutPut.beta

beta =
   1.241
   1.1183

Var = OutPut.covb

Var =
   0.00034471   -6.2425e-005
  -6.2425e-005    0.00021861

R = [0 1];
r = 0;

diff = R*beta - r

diff =
   1.1183
Vdiff = R*Var*R'

Vdiff =

0.00021861

% The Wald test statistic is
W = diff'*inv(Vdiff)*diff

W =

5720.4

% Show the chi-squared density function for 1 degree of freedom
df = size(R,1);

Sig = 0.01;
Reject = chi2inv(1-Sig,df)

Reject =

6.6349

% Compare the W statistic to the rejection value
W =

W =

5720.4

% Regression with LogRVM and its square LogRVM_SQ and a constant
% Ask for all type of output from the regstats command
OutPut = regstats(LogTotalCost,[LogRVM LogRVM_SQ],'linear')

OutPut =

source: 'regstats'
Q: [246x3 double]
R: [3x3 double]
beta: [3x1 double]
covb: [3x3 double]
yhat: [246x1 double]
r: [246x1 double]
mse: 0.079391  
rsquare: 0.95978  
adjrsquare: 0.95928  
leverage: [246x1 double]  
hatmat: [246x246 double]  
s2_i: [246x1 double]  
beta_i: [3x246 double]  
standres: [246x1 double]  
studres: [246x1 double]  
dfbetas: [3x246 double]  
dfit: [246x1 double]  
dffits: [246x1 double]  
covratio: [246x1 double]  
cookd: [246x1 double]  
tstat: [1x1 struct]  
fstat: [1x1 struct]  

% Show coefficients, standard errors, and t-stats
[OutPut.tstat.beta OutPut.tstat.se OutPut.tstat.t]  
ans =

    1.2159    0.022201     54.765
    1.1091    0.015361     72.202
   0.017594   0.0086463    2.0348

% NOTE that now the squared term is statistically significant
% Show s-squared
OutPut.mse  
ans =

    0.079391

% Conduct a Wald test of the hypothesis that both RVM coefficients = 0
beta = OutPut.beta  
beta =

    1.2159
    1.1091
   0.017594

Var = OutPut.covb
Var =

0.0004929 -6.1938e-006 -0.0001068
-6.1938e-006 0.00023597 -3.8806e-005
-0.0001068 -3.8806e-005 7.4758e-005

R = [0 1 0; 0 0 1];
r = [0;0];

diff = R*beta - r

diff =

1.1091
0.017594

Vdiff = R*Var*R'

Vdiff =

0.00023597 -3.8806e-005
-3.8806e-005 7.4758e-005

% The Wald test statistic is
W = diff'*inv(Vdiff)*diff

W =

5798.1

% Use the chi-squared density function for 2 degrees of freedom
% to test this hypothesis at the 0.01 significance level
df = size(R,1);

Sig = 0.01;
Reject = chi2inv(1-Sig,df)

Reject =

9.2103

% Compare the W statistic to the rejection value
\[ W = 5798.1 \]