

Diagnose Outliers Using Residuals

Load the examgrades data set.

```
load examgrades
```

Use the last exam scores as response data and the first two exam scores as predictor data.

```
y = grades(:,5);  
X = [ones(size(grades(:,1))) grades(:,1:2)];
```

Perform multiple linear regression with $\alpha = 0.01$.

```
[~,~,r,rint] = regress(y,X,0.01);
```

Diagnose outliers by finding the residual intervals rint that do not contain 0.

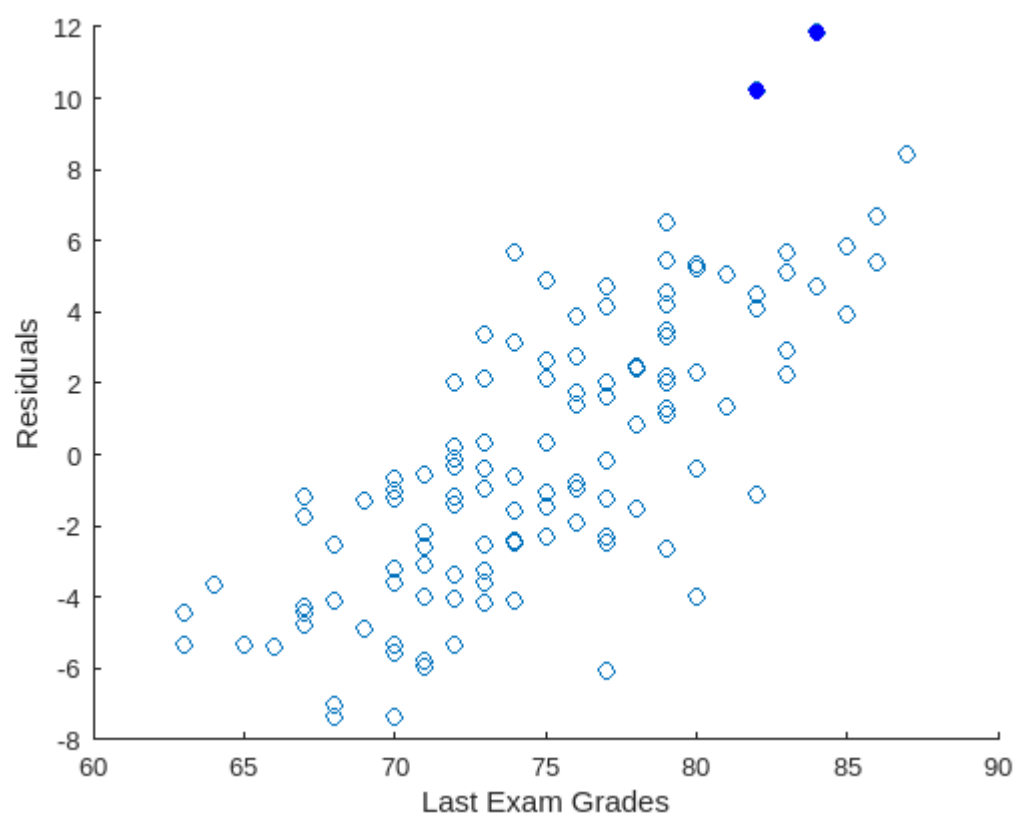
```
contain0 = (rint(:,1)<0 & rint(:,2)>0);  
idx = find(contain0==false)
```

```
idx = 2x1  
    53  
    54
```

Observations 53 and 54 are possible outliers.

Create a scatter plot of the residuals. Fill in the points corresponding to the outliers.

```
hold on  
scatter(y,r)  
scatter(y(idx),r(idx),'b','filled')  
xlabel("Last Exam Grades")  
ylabel("Residuals")  
hold off
```



Copyright 2012 The MathWorks, Inc.