4 Model Fitting and Diagnostics

To date we have looked at the theory underpinning multiple linear regression analysis, where parameter estimates are found using *ordinary least squares*, and we have relied on the assumption of normality to make inferences from the fitted model. In particular, we have seen how to test between subsets of variables and assess the fit of the chosen model.

In this chapter we will consider a more systematic approach to model fitting to help choose the 'best' subset of variables, and methods to assess whether the modelling assumptions (that the error terms are independently and normally distributed with zero mean and constant variance) have been met.

4.1 Choosing a subset of variables

A systematic way of investigating which variables could be omitted from the regression is to carry out and compare the regression analyses for all possible subsets of the regressor variables. However, this can prove to be prohibitively time consuming if k is large, since there are $2^k - 1$ subsets of regressor variables.

4.1.1 All possible subsets

We have seen in Section 3.5 how the measures of $s = \sqrt{MS_R}$ and R^2 (or adjusted R^2, $\bar{R}^2$) from each model can be used to assess the fit, and can form a basis for selecting between models. The table of the following page shows these values for each of the possible regression models fitted to the `oil` dataset, using subsets of the four explanatory variables; `gravity` (x_1), `pressure` (x_2), `distil` (x_3) and `endpoint` (x_4).

Recall that the tentative result of the analysis in Section 3.5 selected model 11 (with explanatory variables `distil` and `endpoint`).

The plots in the Figure at the top of page 3 show both s and R^2 versus the number of parameters for the 'best' regression model for subsets of parameters of size p, $p = 1, \ldots, 5$. These plots indicate support for model 11, which achieves a value of MS_R close to that of the *full* model on all explanatory variables, and explains almost the same proportion of the total sum of squares (R^2). This model would appear to be a *parsimonious* choice.

Table: All possible regressions from the `oil` dataset

	Regressors	No. parms						
Model	Included	p	$SS_R(p)$	df	$s = \sqrt{MS_R(p)}$	R^2	$\bar{R}^2$	p-value[†]
1	None (β_0)	1	3564.08	31	10.72	0.00	0.00	—
2	x_1	2	3347.82	30	10.56	0.06	0.03	
3	x_2	2	3038.34	30	10.06	0.15	0.12	*
4	x_3	2	3210.38	30	10.34	0.10	0.17	
5	x_4	2	1759.69	30	7.66	0.51	0.50	**
6	x_1, x_2	3	3037.97	29	10.24	0.15	0.09	
7	x_1, x_3	3	3205.74	29	10.51	0.10	0.04	
8	x_1, x_4	3	861.95	29	5.45	0.76	0.74	**
9	x_2, x_3	3	3016.59	29	10.20	0.15	0.10	
10	x_2, x_4	3	369.87	29	3.57	0.90	0.89	**
11	x_3, x_4	3	170.61	29	2.43	0.95	0.95	**
12	x_1, x_2, x_3	4	3008.76	28	10.37	0.16	0.07	
13	x_1, x_3, x_4	4	146.00	28	2.28	0.96	0.95	**
14	x_1, x_2, x_4	4	265.48	28	3.08	0.93	0.92	**
15	x_2, x_3, x_4	4	160.62	28	2.40	0.95	0.95	**
16	x_1, x_2, x_3, x_4	5	134.80	27	2.23	0.96	0.96	**

[†] The p-value column indicates whether the ANOVA F-statistic for the test $H_0 : \beta_i = 0, \forall i = 1, \ldots, k$ (for the terms included in the model), is significant at the 5% level (*) or 1% level (**) of significance.

4.1.2 Sequential methods for variable selection

It is appealling, where the number of possible subsets of regressors is large, to consider automative methods for the selection of the best subset of variables. Sequential methods are popular, where models are systematically built up, adding variables one by one ('forward selection') to a null model comprising just an intercept (β_0), or systematically reduced from a full model, by deleting variables one by one ('backwards selection').

Assume that there are p terms currently in the model (β_0 plus a subset of $p-1$ regressors from a total of k), and $SS_R(p)$ represents the residual sum of squares for this model.

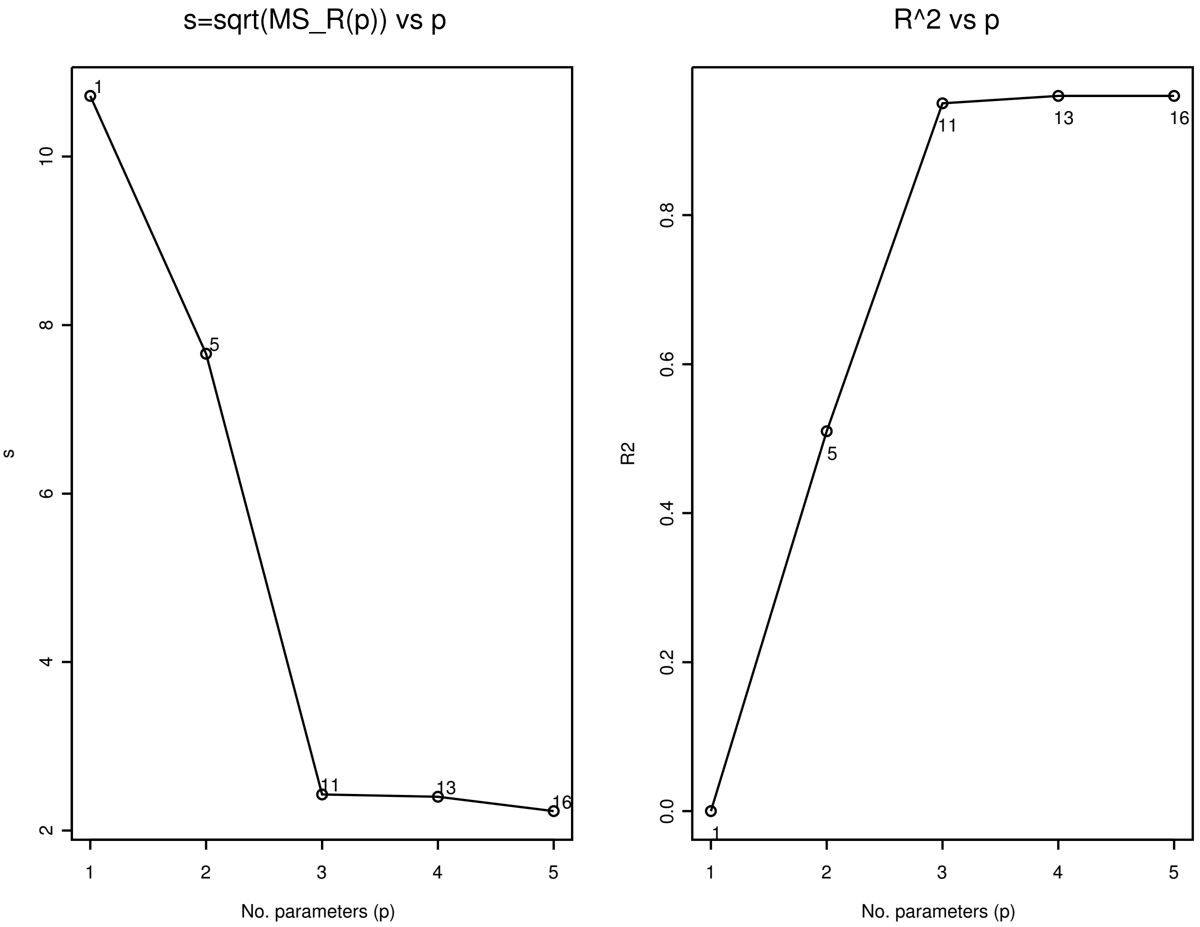
Forward Selection

Starting with the p parameters in the current model:

(1) Fit regressions adding each of the remaining $k - p + 1$ terms individually in turn to the model with the existing p parameters, noting the value of $SS_R(p+1)$ in each case. Note the variable x_{add} whose inclusion gives the smallest value of $SS_R(p+1)$, say $SS'_R(p+1)$.

(2) Calculate

$s = \sqrt{MS_R(p)}$ and R^2 versus no. of parameters (p)
for regression models from the `oil` dataset



s=sqrt(MS_R(p)) vs p

R^2 vs p

$$F_0 = \frac{SS_R(p) - SS'_R(p+1)}{MS'_R(p+1)}$$

(3) If F_0 is greater than a pre-selected '*F to enter*' value, add x_{add} to the model, and return to step (1) with an increased model containing ($p' = p + 1$) terms; otherwise stop.

Backward Selection

Again, starting with the p parameters in the current model:

(1) Fit regressions dropping each of the remaining $p - 1$ terms individually in turn to the model with the existing p parameters, noting the value of $SS_R(p-1)$ in each case. Note the variable x_{drop} whose exclusion gives the smallest value of $SS_R(p-1)$, say $SS'_R(p-1)$.

(2) Calculate

$$F_0 = \frac{SS'_R(p-1) - SS_R(p)}{MS_R(p)}$$

(3) If F_0 is less than an arbitrary pre-chosen '*F to remove*' value, remove x_{drop} from the model, and return to step (1) with a reduced model containing ($p' = p - 1$) terms; otherwise stop.

An alternative procedure is *stepwise selection*, which starts as the forward selection method, but checks at each stage to see whether any variable in the current model can be removed (in the presence of the new term). The process stops once F_0 is less than *F to enter* for all possible additions and F_0 is greater than *F to remove* for all possible exclusions.

Note that with both the forward and backward selection methods, the maximum number of regressions to be performed is $k + (k-1) + (k-2) + \ldots + 1 = \frac{1}{2}k(k+1)$, much fewer than the 2^k required in the 'all subsets' approach. (The maximum number of regressions for the stepwise selection approach is slightly greater).

4.2 The AIC and a stepwise approach in R

The basis for stepwise selection described above is the basis of *stepwise* procedures in a number of statistical packages. An alternative approach is used in R, based on a (likelihood based) goodness-of-fit criterion known as AIC, *Akaike information criterion*, which is defined by

$$\begin{aligned} \text{AIC} &= -2 \text{ maximized log-likelihood} + 2 \text{ number of parameters} \\ &= n \ln(SS_R/n) + 2p + const. \end{aligned} \tag{4.1}$$

When comparing models with different numbers of parameters, the model with the smallest value of the AIC is judged to be the most appropriate.

- As we are only interested in comparing the AICs of different models, it is sufficient to evaluate the AIC up to an arbitrary additive constant.
- Note that the AIC weights the desirability of having a small value of the residual sum of squares SS_R against that of having a small number of parameters p.

The `stepAIC` from `MASS` package or `step` from `stats` are used to carry out a stepwise regression procedure which, by sequentially deleting or adding regressor variables, attempts to find a "best" set of regressor variables. Note that `step` is a minimal implementation and `stepAIC` could be used for a wider range of object classes.
As with the stepwise methods, this procedure is not guaranteed to find in any sense the "best" set of regressor variables, although it may be a useful guide. It is not necessarily the case that there even is a clearly "best" set of regressor variables. If the procedure terminates with, say, m regressor variables, there may be another set of m regressor variables which give a better fit. We also have to specify the starting point for the iteration. In the following output for our example, the object `oil0.lm` is the result of fitting no regressor variables to the response variable `spirit`. It is used to specify the starting point for the iteration of no regressor variables present. Thus the `stepAIC` function has `oil0.lm` as its first argument. The second argument is the *scope* argument, which specifies the set of regressor variables to be considered for inclusion.

At each stage, for each candidate model, the output lists the change in the regression sum of squares due to the deleted or added variable, the resulting residual sum of squares (RSS) and the criterion statistic (AIC).
Finally, a brief summary is provided for the chosen model, in the present case the model with the four regressor variables endpoint, distil, gravity and pressure.

```
> library(MASS)
> oil0.lm <- lm(spirit ~ 1, data = oil)
> stepAIC(oil0.lm, ~ gravity + pressure + distil + endpoint, data = oil)
Start:  AIC=152.81
spirit ~ 1

           Df Sum of Sq    RSS    AIC
+ endpoint  1   1804.38 1759.7 132.23
+ pressure  1    525.74 3038.3 149.71
+ distil    1    353.70 3210.4 151.47
+ gravity   1    216.26 3347.8 152.81
<none>                  3564.1 152.81

Step:  AIC=132.23
spirit ~ endpoint

           Df Sum of Sq    RSS     AIC
+ distil    1   1589.08  170.6  59.557
+ pressure  1   1389.83  369.9  84.317
+ gravity   1    897.75  861.9 111.391
<none>                  1759.7 132.229
- endpoint  1   1804.38 3564.1 152.814

Step:  AIC=59.56
spirit ~ endpoint + distil

           Df Sum of Sq    RSS     AIC
+ gravity   1     24.61  146.0  56.572
<none>                   170.6  59.557
+ pressure  1      9.99  160.6  59.626
- distil    1   1589.08 1759.7 132.229
- endpoint  1   3039.77 3210.4 151.469

Step:  AIC=56.57
spirit ~ endpoint + distil + gravity

           Df Sum of Sq    RSS     AIC
+ pressure  1     11.20  134.8  56.019
<none>                   146.0  56.572
- gravity   1     24.61  170.6  59.557
- distil    1    715.95  861.9 111.391
- endpoint  1   3059.74 3205.7 153.423

Step:  AIC=56.02
spirit ~ endpoint + distil + gravity + pressure

           Df Sum of Sq     RSS     AIC
<none>                   134.80  56.019
- pressure  1     11.20  146.00  56.572
- gravity   1     25.82  160.62  59.626
- distil    1    130.68  265.48  75.706
- endpoint  1   2873.95 3008.76 153.393

Call:
lm(formula = spirit ~ endpoint + distil + gravity + pressure,
    data = oil)

Coefficients:
(Intercept)     endpoint       distil      gravity     pressure
    -6.8208       0.1547      -0.1495       0.2272       0.5537
```

We carry out the stepwise procedure with all regressor variables present, which corresponds to the object `oil.lm`. In this case we find that the procedure deletes no variables, so that the suggested model is again the one with all four regressor variables.

```
> stepAIC(oil.lm, ~ gravity + pressure + distil + endpoint, data = oil)
Start:  AIC=56.02
spirit ~ gravity + pressure + distil + endpoint

           Df Sum of Sq     RSS     AIC
<none>                   134.80  56.019
- pressure  1     11.20  146.00  56.572
- gravity   1     25.82  160.62  59.626
- distil    1    130.68  265.48  75.706
- endpoint  1   2873.95 3008.76 153.393

Call:
lm(formula = spirit ~ gravity + pressure + distil + endpoint,
    data = oil)

Coefficients:
(Intercept)      gravity     pressure       distil     endpoint
    -6.8208       0.2272       0.5537      -0.1495       0.1547
```

4.3 Diagnostics: Fitted values, residuals and leverages

After fitting a multiple regression model, the residuals should be examined for evidence of systematic deviation from the model. We can look at the residuals and standardized residuals from the fitted models, and also consider leverage (the potential for influence), and influence (does omitting a point change the fitted regression).

4.3.1 Residuals

We will look first at residuals. Recall that the observed values of the residuals are given by

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{H})\mathbf{y} \tag{4.2}$$

The resulting vector of residuals $\mathbf{e} = (e_1, e_2, \ldots, e_n)^T$ is used to investigate the validity of the assumptions about the vector of errors $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)^T$, which is not itself observable. Recall, in terms of the multivariate normal distribution, it was assumed that

$$\boldsymbol{\epsilon} \sim \text{MVN}(\mathbf{0}, \sigma^2\mathbf{I}).$$

From(4.2), we have

$$\begin{aligned} \text{E}(\mathbf{e}) &= \text{E}((\mathbf{I} - \mathbf{H})\mathbf{y}) = (\mathbf{I} - \mathbf{H})\text{E}(\mathbf{y}) \\ &= (\mathbf{I} - \mathbf{H})\mathbf{X}\boldsymbol{\beta} = \mathbf{0}, \end{aligned}$$

Furthermore,

$$\begin{aligned} \text{Cov}(\mathbf{e}, \mathbf{e}) &= \text{Cov}((\mathbf{I} - \mathbf{H})\mathbf{y}, (\mathbf{I} - \mathbf{H})\mathbf{y}) = (\mathbf{I} - \mathbf{H})\text{Cov}(\mathbf{y}, \mathbf{y})(\mathbf{I} - \mathbf{H})^T \\ &= \sigma^2(\mathbf{I} - \mathbf{H}), \end{aligned} \tag{4.3}$$

since $\mathbf{I} - \mathbf{H}$ is symmetric and idempotent. Hence, using the property of the normal distribution that a linear combination of normal variates is also normally distributed, under our assumptions about the error distribution,

$$e_i \sim \text{N}(0, (1 - h_{ii})\sigma^2) \qquad i = 1, \ldots, n,$$

where h_{ii} is the ith diagonal element of $\mathbf{H}$. Note that the e_i are <u>not</u> comparable. That is, they have different standard errors and they are, in general, correlated, since from (4.3)

$$\text{Cov}(e_i, e_j) = -h_{ij}\sigma^2.$$

The *standardized (Pearson) residuals* e_i' are defined by

$$e_i' = \frac{e_i}{s\sqrt{1 - h_{ii}}} \qquad i = 1, \ldots, n.$$

If the assumptions of the regression model are correct, the standardized residuals are *approximately* NID$(0, 1)$. Through various descriptive statistics and plots of the residuals, standardized residuals and fitted values, the data can be examined for *outliers* and for indications of departures from the assumptions about the error distribution.

Residual Plots

The following plots may be useful to check the model assumptions concerning the error terms.

Normality

- A normal probability plot (`qqnorm`) of the standardized residuals. Departures from normality are indicated by deviations from a straight line.

(Zero Mean and) Constant Variance

- Plot the standardized residuals e_i' against the fitted values $\hat{y}_i$. The points should be scattered evenly around zero, with no systematic pattern.

Independence

- Plot the standardized residuals against the *serial order* in which the observations were taken. Again, if the modelling assumptions are correct a random scattering of points with no visible trend is expected.

Outliers

Possible *outliers* may be indicated by points with large standardized residuals. Note that (under normality), approximately 95% of observations should have standardized residuals in the range (-2.0, 2.0). A 'rule of thumb' often adopted is to consider an observation with an absolute value of standardized residual > 2.5 say as an outlier, and the accuracy of such an observed value should be investigated.

Model Adequacy

The standardized residuals can be plotted against the individual regressor (explanatory) variables in turn. All such plots should indicate random scatter of equal width about zero. There are two possible situations to consider.

1. For regressors that are in the model: non-linearity suggests that higher order terms involving those regressors should be added to the model.

2. For regressors that are not included in the model: any systematic pattern with the residuals suggests that those regressors should be added to the model.

4.3.2 Fitted Values and Leverages

We now turn our attention to the fitted values $\hat{y}_i$, elements of $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$.
From the multiple regression model

$$\mathbf{y} \sim \mathrm{MVN}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}).$$

it follows that

$$\mathrm{E}(\hat{\mathbf{y}}) = \mathrm{E}(\mathbf{H}\mathbf{y}) = \mathbf{H}\mathrm{E}(\mathbf{y}) = \mathbf{H}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta},$$

Furthermore,

$$\mathrm{Cov}(\hat{\mathbf{y}}, \hat{\mathbf{y}}) = \mathbf{H}\mathrm{Cov}(\mathbf{y}, \mathbf{y})\mathbf{H}^T = \mathbf{H}\sigma^2\mathbf{I}\mathbf{H} = \sigma^2\mathbf{H},$$

using the properties that $\mathbf{H}$ is symmetric and idempotent. In particular, the variance of the i-th fitted value is given by

$$\mathrm{Var}(\hat{y}_i) = h_{ii}\sigma^2 \qquad i = 1, \ldots, n,$$

where h_{ii} is the ith diagonal element of the hat matrix $\mathbf{H}$ and is known as the *leverage* of the ith observation. In normal linear models, points that have values for the explanatory variables that are very different from those of other points have the potential to dominate the regression analysis.

For the simple linear regression (on one explanatory variable x), it is easily seen that

$$h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{(n-1)s_x^2} \tag{4.4}$$

(c/f solution to Exercises 2).

Hence in this (simple) setting, points which have a high leverage are those which are far from the mean of the explanatory variable x. In the multiple regression context h_{ii} measures the difference from the centroid of the x's , taking account of the covariance structure of the x's. This is seen by considering the (reparameterized) regression model

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}_1 + \mathbf{e}$$

where $\mathbf{y}_i^* = (y_i - \bar{y})$, $i = 1, \ldots, n$, and $\mathbf{X}_{ij}^* = (x_{ij} - \bar{x}_j)$, $i = 1, \ldots, n$, $j = 1, \ldots, k$. In this formulation $\mathbf{b}_1$ is the ordinary least squares estimate of the regression slopes *not* including an intercept.

Now, the hat value for the ith observation is

$$h_{ii}^* = \mathbf{x}_i^{*T}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{x}_i^* = h_{ii} - \frac{1}{n}$$

i.e.

$$h_{ii} = \frac{1}{n} + \mathbf{x}_i^{*T}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{x}_i^* \tag{4.5}$$

where $\mathbf{x}_i^{*T} = (x_{i1} - \bar{x}_1, \ldots, x_{ik} - \bar{x}_k)$ is the ith row of $\mathbf{X}^*$. It is in fact the case that

$$h_{ii} = \frac{1}{n} + \frac{D^2(\mathbf{x}_i, \bar{\mathbf{x}})}{n-1}$$

where $D^2(\mathbf{x}_i, \bar{\mathbf{x}})$ is the *Mahalanobis* distance squared of $\mathbf{x}_i$ from $\bar{\mathbf{x}}$ in the k-dimensional space - a measure of *statistical distance* that accounts for the correlation between variables. Note that the Mahalanobis distances and hat values are invariant with respect to any non-singular linear transformation of $\mathbf{X}$.

The leverages h_{ii} satisfy

$$\sum_{i=1}^{n} h_{ii} = \mathrm{tr}(\mathbf{H}) = p,$$

as seen in Section 2.4. Thus the average value of the h_{ii}, $i = 1, \ldots, n$, is equal to p/n. More detailed analysis shows that

$$\frac{1}{n} \leq h_{ii} \leq 1, \qquad i = 1, \ldots, n.$$

The leverage h_{ii}, which depends only on the values of the regressor variables and not on the values of the response variable, may be viewed as a measure of the distance of the ith observation from the sample mean vector of all n observations in the space of the regressor variables. If an individual leverage h_{ii} is large then it may be that the corresponding observation is *influential* in determining the estimated regression coefficients, that is, the removal of the observation from the data set results in non-trivial changes in the estimates of the regression coefficients. So observations with large leverage should be treated with caution. A '*rule of thumb*' suggests that leverages greater than $2p/n$ (twice the average) should be treated as large.

There are functions in R that allow the examination of residuals, leverages and other diagnostics from a linear regression.

4.3.3 Influence: Cook's distance

In general, an observation may be influential either when it has a large standardized residual (the response variable is an outlier) or when it has a large leverage (the vector of regressor variables is an outlier). When such observations occur, they should be investigated to see if there is any obvious explanation as to why they have occurred. The regression analysis may be repeated after the removal of some or all of these observations to check what the effect will be.

Another measure of influence is *Cook's distance*, D_i, which is a measure of the squared distance between the least squares estimate $\mathbf{b}$ based on all n points in the sample and the least squares estimate $\mathbf{b}_{(i)}$ obtained when the ith point is deleted.

It can be shown that the *influence vector* $\mathbf{b} - \mathbf{b}_{(i)}$ may be written efficiently as

$$\mathbf{b} - \mathbf{b}_{(i)} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_i \frac{e_i}{1 - h_{ii}} \tag{4.6}$$

Then, the (Mahalanobis) distance between the fitted regression coefficients calculated with and without the ith observation (scaled by the number of parameters $(k+1)$) may be written

$$\frac{(\mathbf{b} - \mathbf{b}_{(i)})^T\mathbf{X}^T\mathbf{X}(\mathbf{b} - \mathbf{b}_{(i)})}{(k+1)s^2} = \frac{(\hat{\mathbf{y}} - \hat{\mathbf{y}}_{(i)})^T(\hat{\mathbf{y}} - \hat{\mathbf{y}}_{(i)})}{(k+1)s^2} \tag{4.7}$$

This is Cook's distance, and is equivalent to the F-statistic for testing the hypothesis $\boldsymbol{\beta} = \mathbf{b}_{(i)}$.

Using (4.6), Cook's distance may be written

$$D_i = \frac{e_i'^2}{(k+1)} \frac{h_{ii}}{1 - h_{ii}} = \frac{e_i'^2}{(k+1)} \frac{\mathrm{Var}(\hat{y}_i)}{\mathrm{Var}(e_i)} \qquad i = 1, \ldots, n.$$

Hence, Cook's distance can be calculated following a regression analysis without having to perform any *dropped case* regressions. The above formula also shows that Cook's distance D_i may take a large value either if the standardized residual e_i' is large or the leverage h_{ii} is large. That is the actual *influence* exerted by a point is a combination of its leverage (potential to influence) and residual (lack of fit to the model).

4.4 Diagnostic plots in R

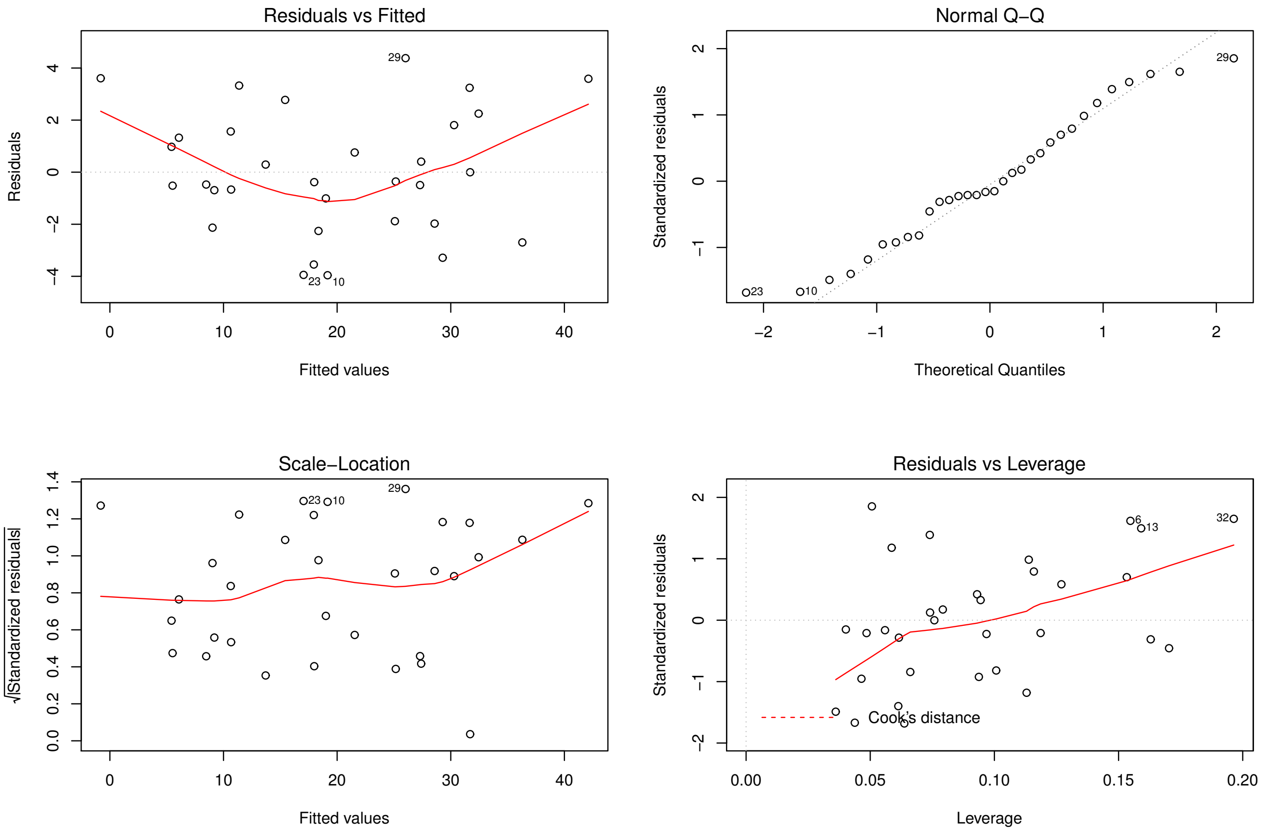
We end this Chapter by looking at some residual plots for the model `oil2.lm`, the *chosen* model using the regressors `distil` and `endpoint`. Vectors of fitted values and (raw) residuals may be obtained from the model object and plotted as follows (plot not reported):

```
> fits <- fitted(oil2.lm)     # alternatively fits <- oil2.lm$fitted
> resids <- resid(oil2.lm)    # alternatively resids <- oil2.lm$residuals

> plot(fits, resids, main = "Fitted Values vs Residuals")
```

Alternatively, the generic `plot` function in R will produce *default* residual plots, using say

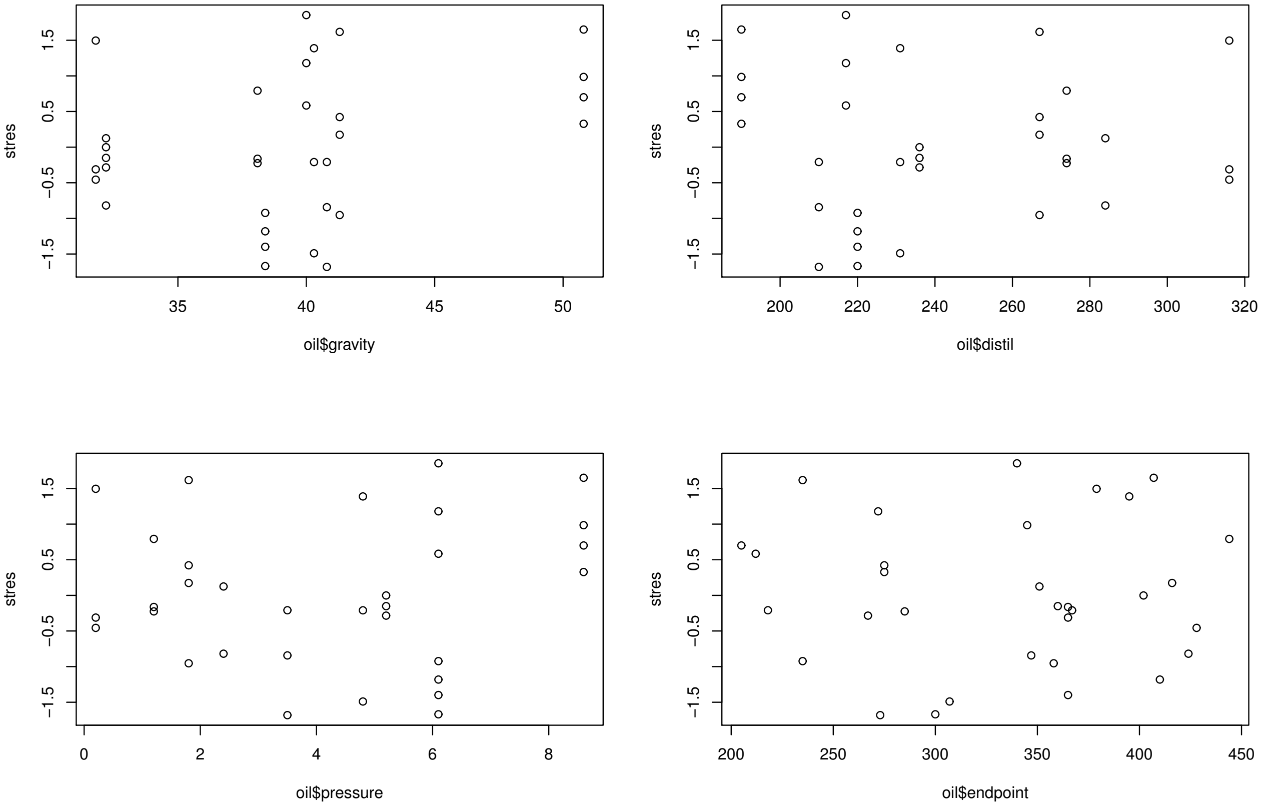
```
> par(mfrow = c(2, 2))
> plot(oil2.lm)
```



The top left plot of Residuals vs Fitted, does not show any indication of increasing variance with mean, which means that the constant variance assumption holds here. The other feature to look for is a pattern in the average value of the residuals as the fitted values change. The solid curve shows a running average of the residuals to help judging this: there is a slight pattern here, which is not extreme however. The lower left plot shows the square root of the absolute value of the standardized residuals against fitted value (again with a running average curve). If all is well the points should be evenly spread with respect the vertical axis here, with no trend in their average value. A trend in average value is indicative of a problem with the constant variance assumption, and is not a concern in this case. The top right plots the ordered standardized residuals against quantiles of a standard normal: a systematic deviation from a straight line is not present here indicating no departure from normality in the residuals. The lower right plot is looking at leverage and influence of residuals, by plotting standardized residuals against a measure of leverage. A combination of high residuals and high leverage indicates a point with substantial influence on the fit. A standard way of measuring this is via Cook's distance (which measures the change in all model fitted values on omission of the data point in question). It turns out that Cook's distance is a function of leverage and the standardized residuals, so contours of Cook's distance values are shown on the plot. Cook's distances over 0.5 are considered borderline problematic, while over 1 is usually considered highly influential, so points to the right of these contours warrant investigation. Although a couple of points have rather high leverage, their actual influence on the fit is not unduly large.

Finally we show plots of the standardized residuals from `oil2.lm` against each of the four possible explanatory variables. There are no obvious patterns against any, whether included in the model or not, so that the adequacy of the model is not questioned.

```
> stres <- stdres(oil2.lm)    # standardized residuals
> par(mfrow = c(2, 2))
> plot(oil$gravity, stres)
> plot(oil$distil, stres)
> plot(oil$pressure, stres)
> plot(oil$endpoint, stres)
```


stres
oil$gravity
stres
oil$distil
stres
oil$pressure
stres
oil$endpoint