Regression Analysis lab 6

1 Influence diagnostics

1.1 Import data

```
delivery<-read.csv(file="D:/chilo/Regression 6/delivery.csv", header=T)
delivery
```

```
   observation  time cases distance
1            1 16.68     7      560
2            2 11.50     3      220
3            3 12.03     3      340
4            4 14.88     4       80
5            5 13.75     6      150
6            6 18.11     7      330
7            7  8.00     2      110
8            8 17.83     7      210
9            9 79.24    30     1460
10          10 21.50     5      605
11          11 40.33    16      688
12          12 21.00    10      215
13          13 13.50     4      255
14          14 19.75     6      462
15          15 24.00     9      448
16          16 29.00    10      776
17          17 15.35     6      200
18          18 19.00     7      132
19          19  9.50     3       36
20          20 35.10    17      770
21          21 17.90    10      140
22          22 52.32    26      810
23          23 18.75     9      450
24          24 19.83     8      635
25          25 10.75     4      150
```

1.2 Fit a multiple linear regression

```
attach(delivery)
dfit <- lm(time ~ cases + distance, data=delivery)
summary(dfit)
```

```
Call:
lm(formula = time ~ cases + distance, data = delivery)

Residuals:
   Min     1Q Median     3Q    Max
-5.788 -0.663  0.436  1.157  7.420

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.34123    1.09673    2.13  0.04417 *
cases        1.61591    0.17073    9.46  3.3e-09 ***
distance     0.01438    0.00361    3.98  0.00063 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.26 on 22 degrees of freedom
Multiple R-squared:  0.96,Adjusted R-squared:  0.956
F-statistic:  261 on 2 and 22 DF,  p-value: 4.69e-16

names(dfit)

 [1] "coefficients"  "residuals"     "effects"       "rank"
 [5] "fitted.values" "assign"        "qr"            "df.residual"
 [9] "xlevels"       "call"          "terms"         "model"

dfit$fit

     1      2      3      4      5      6      7      8      9     10
21.708 10.354 12.080  9.956 14.194 18.400  7.155 16.673 71.820 19.124
    11     12     13     14     15     16     17     18     19     20
38.093 21.593 12.473 18.682 23.329 29.663 14.914 15.551  7.707 40.888
    21     22     23     24     25
20.514 56.007 23.358 24.403 10.963

dfit$res  # residuals

       1        2        3        4        5        6        7        8
-5.02808  1.14639 -0.04979  4.92435 -0.44440 -0.28957  0.84462  1.15660
       9       10       11       12       13       14       15       16
 7.41971  2.37641  2.23749 -0.59304  1.02701  1.06754  0.67120 -0.66293
      17       18       19       20       21       22       23       24
 0.43636  3.44862  1.79319 -5.78797 -2.61418 -3.68653 -4.60757 -4.57285
      25
-0.21258

names(summary(dfit))
```

```
 [1] "call"          "terms"         "residuals"     "coefficients"
 [5] "aliased"       "sigma"         "df"            "r.squared"
 [9] "adj.r.squared" "fstatistic"    "cov.unscaled"

sigmahat<-summary(dfit)$sigma
sigmahat

[1] 3.259

sigmahat2<-sigmahat^2
sigmahat2

[1] 10.62
```

1.3 Compute residuals

```
e<-dfit$res
e

##        1        2        3        4        5        6        7        8
## -5.02808  1.14639 -0.04979  4.92435 -0.44440 -0.28957  0.84462  1.15660
##        9       10       11       12       13       14       15       16
##  7.41971  2.37641  2.23749 -0.59304  1.02701  1.06754  0.67120 -0.66293
##       17       18       19       20       21       22       23       24
##  0.43636  3.44862  1.79319 -5.78797 -2.61418 -3.68653 -4.60757 -4.57285
##       25
## -0.21258
```

1.4 Compute standardized residuals

```
MSE<-sigmahat2
d<-e/sqrt(MSE)
d

##        1        2        3        4        5        6        7        8
## -1.54261  0.35171 -0.01528  1.51078 -0.13634 -0.08884  0.25913  0.35484
##        9       10       11       12       13       14       15       16
##  2.27635  0.72908  0.68646 -0.18194  0.31508  0.32752  0.20592 -0.20339
##       17       18       19       20       21       22       23       24
##  0.13387  1.05803  0.55015 -1.77574 -0.80202 -1.13102 -1.41359 -1.40294
##       25
## -0.06522
```

1.5 Compute hat matrix

```
n<-length(delivery$time)
n

[1] 25

delivery[,-c(1,2)]

   cases distance
1      7      560
2      3      220
3      3      340
4      4       80
5      6      150
6      7      330
7      2      110
8      7      210
9     30     1460
10     5      605
11    16      688
12    10      215
13     4      255
14     6      462
15     9      448
16    10      776
17     6      200
18     7      132
19     3       36
20    17      770
21    10      140
22    26      810
23     9      450
24     8      635
25     4      150

X<-cbind(1,delivery[,-c(1,2)])
X

   1 cases distance
1  1     7      560
2  1     3      220
3  1     3      340
4  1     4       80
5  1     6      150
6  1     7      330
```

```
7  1     2      110
8  1     7      210
9  1    30     1460
10 1     5      605
11 1    16      688
12 1    10      215
13 1     4      255
14 1     6      462
15 1     9      448
16 1    10      776
17 1     6      200
18 1     7      132
19 1     3       36
20 1    17      770
21 1    10      140
22 1    26      810
23 1     9      450
24 1     8      635
25 1     4      150
```

```
X <- as.matrix(X)
t(X) %*% X
```

```
             1  cases distance
1           25    219    10232
cases      219   3055   133899
distance 10232 133899  6725688
```

```
XtXi <- solve(t(X) %*% X)
XtXi
```

```
                  1      cases   distance
1         1.132e-01 -4.449e-03 -8.367e-05
cases    -4.449e-03  2.744e-03 -4.786e-05
distance -8.367e-05 -4.786e-05  1.229e-06
```

```
H<-X %*% XtXi %*% t(X)
H
```

```
           [,1]       [,2]      [,3]      [,4]      [,5]     [,6]
 [1,]  0.101802  5.837e-02  0.090697  0.008604  0.003380 0.039834
 [2,]  0.058366  7.070e-02  0.075871  0.057925  0.047449 0.048457
 [3,]  0.090697  7.587e-02  0.098735  0.036708  0.025068 0.046875
 [4,]  0.008604  5.792e-02  0.036708  0.085375  0.078395 0.049268
 [5,]  0.003380  4.745e-02  0.025068  0.078395  0.075010 0.046275
 [6,]  0.039834  4.846e-02  0.046875  0.049268  0.046275 0.042867
 [7,]  0.040771  7.271e-02  0.067402  0.074675  0.063128 0.050943
```

```
 [8,]  0.007503  4.329e-02  0.024011  0.070485  0.068656 0.044449
 [9,]  0.067316 -5.802e-02 -0.025066 -0.088464 -0.053256 0.004162
[10,]  0.138010  7.380e-02  0.124248 -0.004748 -0.014684 0.041311
[11,]  0.027909  3.168e-03  0.002687  0.010255  0.023027 0.028832
[12,] -0.027276  2.327e-02 -0.012502  0.077694  0.082230 0.041279
[13,]  0.055753  6.546e-02  0.070051  0.054435  0.045757 0.046961
[14,]  0.087440  6.089e-02  0.084513  0.023233  0.016821 0.042161
[15,]  0.047542  4.005e-02  0.044380  0.033802  0.033939 0.039241
[16,]  0.123871  4.743e-02  0.094385 -0.021491 -0.022398 0.033881
[17,]  0.016851  4.960e-02  0.034595  0.069555  0.065685 0.045616
[18,] -0.013512  3.993e-02  0.009150  0.084275  0.083203 0.045478
[19,]  0.008791  6.277e-02  0.040814  0.090456  0.081765 0.050884
[20,]  0.037960 -4.559e-05  0.005821 -0.001544  0.012569 0.026715
[21,] -0.047483  2.003e-02 -0.026792  0.090955  0.096218 0.042268
[22,] -0.059641 -5.903e-02 -0.098955  0.015665  0.048629 0.016873
[23,]  0.048081  4.014e-02  0.044761  0.033448  0.033566 0.039215
[24,]  0.109967  5.485e-02  0.092498 -0.001958 -0.005773 0.037810
[25,]  0.027464  6.094e-02  0.050046  0.072999  0.065339 0.048345
          [,7]       [,8]      [,9]     [,10]     [,11]     [,12]
 [1,]  0.040771  0.0075031  0.067316  0.138010  0.027909 -0.027276
 [2,]  0.072709  0.0432876 -0.058016  0.073796  0.003168  0.023266
 [3,]  0.067402  0.0240113 -0.025066  0.124248  0.002687 -0.012502
 [4,]  0.074675  0.0704845 -0.088464 -0.004748  0.010255  0.077694
 [5,]  0.063128  0.0686558 -0.053256 -0.014684  0.023027  0.082230
 [6,]  0.050943  0.0444493  0.004162  0.041311  0.028832  0.041279
 [7,]  0.081799  0.0562497 -0.096213  0.047231 -0.002918  0.043353
 [8,]  0.056250  0.0637256 -0.028787 -0.009141  0.029313  0.077046
 [9,] -0.096213 -0.0287874  0.498292  0.063685  0.174405 -0.003434
[10,]  0.047231 -0.0091413  0.063685  0.196296  0.014676 -0.066089
[11,] -0.002918  0.0293128  0.174405  0.014676  0.086133  0.048872
[12,]  0.043353  0.0770461 -0.003434 -0.066089  0.048872  0.113656
[13,]  0.066935  0.0423733 -0.040412  0.068828  0.009554  0.025533
[14,]  0.049330  0.0185375  0.032414  0.116491  0.021776 -0.010766
[15,]  0.037274  0.0349101  0.052550  0.051555  0.041411  0.031507
[16,]  0.018542 -0.0130706  0.150607  0.169774  0.046622 -0.053558
[17,]  0.060917  0.0606240 -0.039527  0.006338  0.022827  0.067327
[18,]  0.059699  0.0762552 -0.050205 -0.041935  0.029626  0.100295
[19,]  0.080846  0.0728446 -0.108539 -0.003564  0.003905  0.078109
[20,] -0.010769  0.0208486  0.204914  0.029469  0.092330  0.037130
[21,]  0.046669  0.0890938 -0.024027 -0.097621  0.049172  0.136010
[22,] -0.050566  0.0567943  0.287840 -0.130863  0.150906  0.139507
[23,]  0.037185  0.0345888  0.053099  0.052396  0.041403  0.030911
[24,]  0.033229  0.0001633  0.095904  0.149859  0.034135 -0.036931
[25,]  0.071579  0.0592400 -0.069243  0.024682  0.009975  0.056830
          [,13]     [,14]   [,15]     [,16]     [,17]     [,18]     [,19]
```

```
 [1,]  0.055753  0.087440 0.04754  0.123871  0.016851 -0.013512  0.008791
 [2,]  0.065464  0.060890 0.04005  0.047434  0.049603  0.039927  0.062775
 [3,]  0.070051  0.084513 0.04438  0.094385  0.034595  0.009150  0.040814
 [4,]  0.054435  0.023233 0.03380 -0.021491  0.069555  0.084275  0.090456
 [5,]  0.045757  0.016821 0.03394 -0.022398  0.065685  0.083203  0.081765
 [6,]  0.046961  0.042161 0.03924  0.033881  0.045616  0.045478  0.050884
 [7,]  0.066935  0.049330 0.03727  0.018542  0.060917  0.059699  0.080846
 [8,]  0.042373  0.018537 0.03491 -0.013071  0.060624  0.076255  0.072845
 [9,] -0.040412  0.032414 0.05255  0.150607 -0.039527 -0.050205 -0.108539
[10,]  0.068828  0.116491 0.05156  0.169774  0.006338 -0.041935 -0.003564
[11,]  0.009554  0.021776 0.04141  0.046622  0.022827  0.029626  0.003905
[12,]  0.025533 -0.010766 0.03151 -0.053558  0.067327  0.100295  0.078109
[13,]  0.061125  0.057684 0.04012  0.046980  0.047668  0.039391  0.058430
[14,]  0.057684  0.078243 0.04520  0.099675  0.026665  0.003182  0.024666
[15,]  0.040118  0.045200 0.04111  0.051754  0.035744  0.032095  0.033408
[16,]  0.046980  0.099675 0.05175  0.165940 -0.002835 -0.043589 -0.024559
[17,]  0.047668  0.026665 0.03574 -0.002835  0.059432  0.070379  0.072615
[18,]  0.039391  0.003182 0.03209 -0.043589  0.070379  0.096260  0.087119
[19,]  0.058430  0.024666 0.03341 -0.024559  0.072615  0.087119  0.096449
[20,]  0.007011  0.027823 0.04318  0.064558  0.015014  0.017035 -0.009041
[21,]  0.022666 -0.025531 0.02880 -0.082903  0.076707  0.119531  0.091835
[22,] -0.042553 -0.055165 0.03387 -0.047122  0.031995  0.082743  0.002177
[23,]  0.040194  0.045593 0.04118  0.052537  0.035494  0.031582  0.033042
[24,]  0.052944  0.092109 0.04905  0.139068  0.009914 -0.024307 -0.002874
[25,]  0.057111  0.037013 0.03633  0.005897  0.060800  0.066322  0.077645
          [,20]    [,21]     [,22]   [,23]      [,24]     [,25]
 [1,]  3.796e-02 -0.04748 -0.059641 0.04808  0.1099666  0.027464
 [2,] -4.559e-05  0.02003 -0.059035 0.04014  0.0548508  0.060940
 [3,]  5.821e-03 -0.02679 -0.098955 0.04476  0.0924977  0.050046
 [4,] -1.544e-03  0.09095  0.015665 0.03345 -0.0019581  0.072999
 [5,]  1.257e-02  0.09622  0.048629 0.03357 -0.0057726  0.065339
 [6,]  2.672e-02  0.04227  0.016873 0.03921  0.0378101  0.048345
 [7,] -1.077e-02  0.04667 -0.050566 0.03719  0.0332288  0.071579
 [8,]  2.085e-02  0.08909  0.056794 0.03459  0.0001633  0.059240
 [9,]  2.049e-01 -0.02403  0.287840 0.05310  0.0959036 -0.069243
[10,]  2.947e-02 -0.09762 -0.130863 0.05240  0.1498593  0.024682
[11,]  9.233e-02  0.04917  0.150906 0.04140  0.0341350  0.009975
[12,]  3.713e-02  0.13601  0.139507 0.03091 -0.0369308  0.056830
[13,]  7.011e-03  0.02267 -0.042553 0.04019  0.0529436  0.057111
[14,]  2.782e-02 -0.02553 -0.055165 0.04559  0.0921092  0.037013
[15,]  4.318e-02  0.02880  0.033869 0.04118  0.0490544  0.036328
[16,]  6.456e-02 -0.08290 -0.047122 0.05254  0.1390681  0.005897
[17,]  1.501e-02  0.07671  0.031995 0.03549  0.0099136  0.060800
[18,]  1.704e-02  0.11953  0.082743 0.03158 -0.0243072  0.066322
[19,] -9.041e-03  0.09183  0.002177 0.03304 -0.0028743  0.077645
```

```
[20,]  1.017e-01  0.03346  0.151752 0.04327  0.0469728  0.001878
[21,]  3.346e-02  0.16528  0.164458 0.02802 -0.0604601  0.063639
[22,]  1.518e-01  0.16446  0.391575 0.03320 -0.0564665 -0.007622
[23,]  4.327e-02  0.02802  0.033204 0.04126  0.0496818  0.036147
[24,]  4.697e-02 -0.06046 -0.056466 0.04968  0.1206083  0.020003
[25,]  1.878e-03  0.06364 -0.007622 0.03615  0.0200026  0.066643
```

1.6 Compute internally studentized residuals

```
h<-diag(H)
h
```

```
##  [1] 0.10180 0.07070 0.09873 0.08537 0.07501 0.04287 0.08180 0.06373
##  [9] 0.49829 0.19630 0.08613 0.11366 0.06112 0.07824 0.04111 0.16594
## [17] 0.05943 0.09626 0.09645 0.10168 0.16528 0.39158 0.04126 0.12061
## [25] 0.06664
```

```
r<-e/sqrt(MSE*(1-h))
r
```

```
##        1        2        3        4        5        6        7        8
## -1.62768  0.36484 -0.01609  1.57972 -0.14176 -0.09081  0.27042  0.36672
##        9       10       11       12       13       14       15       16
##  3.21376  0.81325  0.71808 -0.19326  0.32518  0.34114  0.21029 -0.22270
##       17       18       19       20       21       22       23       24
##  0.13804  1.11295  0.57877 -1.87355 -0.87784 -1.45000 -1.44369 -1.49606
##       25
## -0.06751
```

1.7 Compute externally studentized residuals

```
dMSE<-((n-3)*MSE-e^2/(1-h))/(n-3-1)
t<-e/sqrt(dMSE*(1-h))
t
```

```
       1        2        3        4        5        6        7        8
-1.69563  0.35754 -0.01572  1.63916 -0.13856 -0.08874  0.26465  0.35939
       9       10       11       12       13       14       15       16
 4.31078  0.80678  0.70994 -0.18897  0.31847  0.33418  0.20566 -0.21783
      17       18       19       20       21       22       23       24
 0.13492  1.11933  0.56981 -1.99668 -0.87309 -1.48962 -1.48247 -1.54222
      25
-0.06596
```

```
t1<-rstudent(dfit)
t1
```

```
       1        2        3        4        5        6        7        8
-1.69563  0.35754 -0.01572  1.63916 -0.13856 -0.08874  0.26465  0.35939
       9       10       11       12       13       14       15       16
 4.31078  0.80678  0.70994 -0.18897  0.31847  0.33418  0.20566 -0.21783
      17       18       19       20       21       22       23       24
 0.13492  1.11933  0.56981 -1.99668 -0.87309 -1.48962 -1.48247 -1.54222
      25
-0.06596
```

1.8 Compute Cook's distance

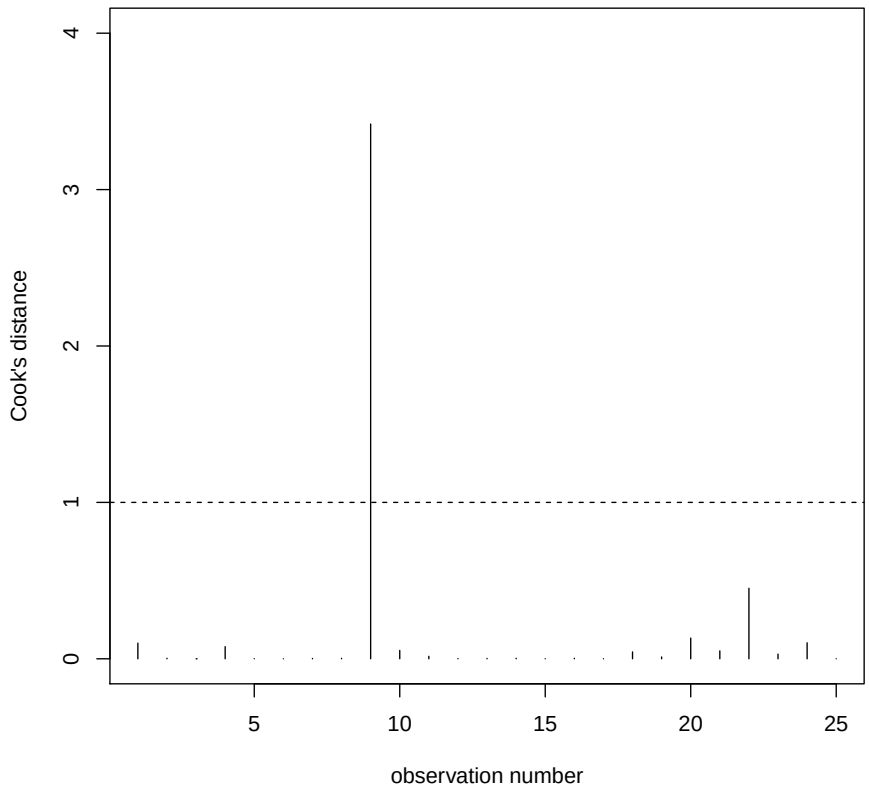
```
dfit <- lm(time ~ cases + distance, data=delivery)
n<-length(delivery$time)
n
```

```
[1] 25
```

```
p<-3
COOK <- cooks.distance(dfit)
COOK
```

```
        1         2         3         4         5         6         7
1.001e-01 3.376e-03 9.456e-06 7.765e-02 5.432e-04 1.231e-04 2.172e-03
        8         9        10        11        12        13        14
3.051e-03 3.419e+00 5.385e-02 1.620e-02 1.596e-03 2.295e-03 3.293e-03
       15        16        17        18        19        20        21
6.320e-04 3.289e-03 4.013e-04 4.398e-02 1.192e-02 1.324e-01 5.086e-02
       22        23        24        25
4.510e-01 2.990e-02 1.023e-01 1.085e-04
```

```
plot(COOK, xlab = 'observation number',
ylab ="Cook's distance", type = 'h', ylim = c(0, 4))
abline(h = 1, lty = 2)
identify(COOK)
```




```
integer(0)
```

1.9 DFBETAS residuals

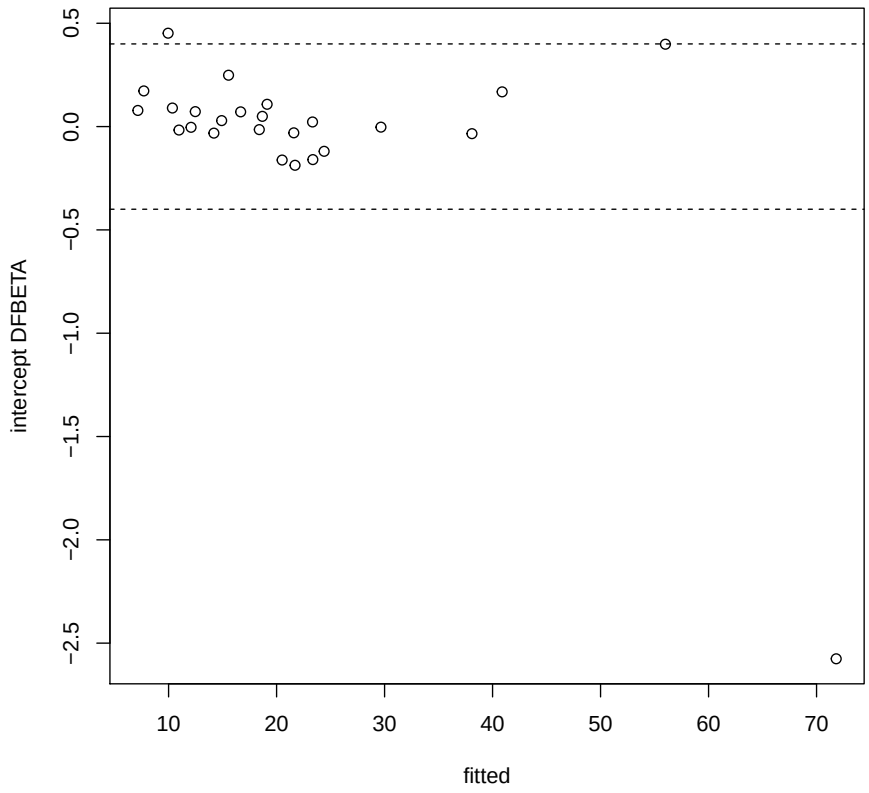
```
dfit <- lm(time ~ cases + distance, data=delivery)
n<-length(delivery$time)
n

[1] 25

p<-3
DFBETAS <- dfbetas(dfit)
DFBETAS
```

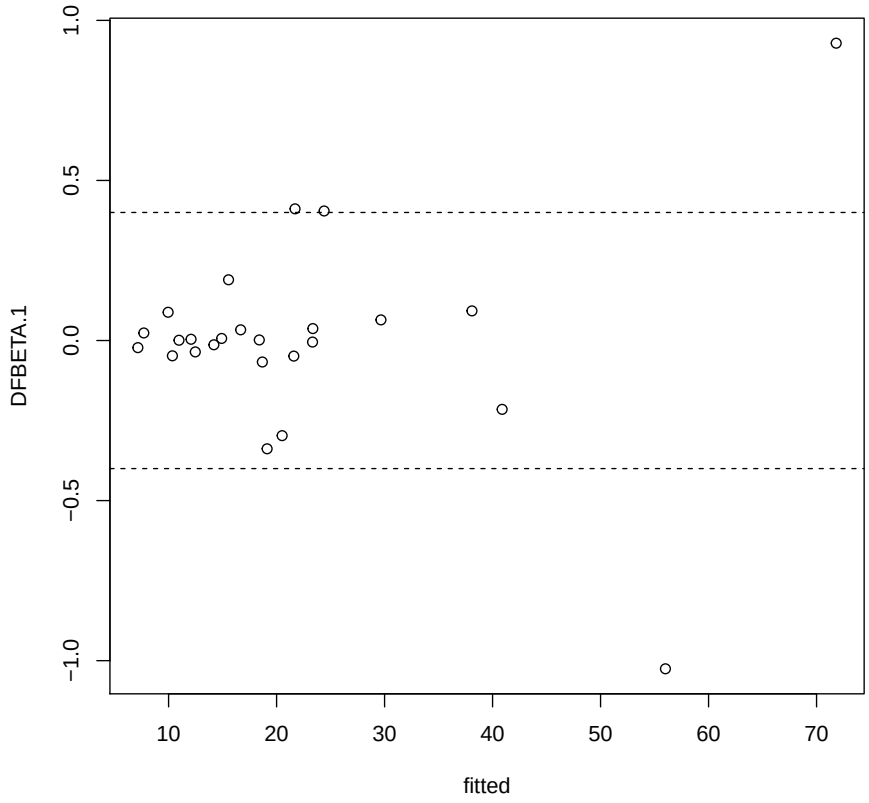
```
   (Intercept)      cases  distance
1    -0.187267  0.4113119 -0.434862
2     0.089793 -0.0477642  0.014414
3    -0.003515  0.0039484 -0.002846
4     0.451965  0.0882803 -0.273373
5    -0.031674 -0.0133001  0.024240
6    -0.014681  0.0017921  0.001079
7     0.078071 -0.0222783 -0.011019
8     0.071203  0.0333823 -0.053824
9    -2.575740  0.9287433  1.507551
10    0.107919 -0.3381629  0.341327
11   -0.034275  0.0925272 -0.002686
12   -0.030269 -0.0486664  0.053973
13    0.072366 -0.0356212  0.011335
14    0.049517 -0.0670869  0.061817
15    0.022279 -0.0047895  0.006838
16   -0.002693  0.0644208 -0.084188
17    0.028856  0.0064876 -0.015697
18    0.248558  0.1897331 -0.272431
19    0.172559  0.0235737 -0.098969
20    0.168037 -0.2149950 -0.092915
21   -0.161929 -0.2971751  0.336406
22    0.398566 -1.0254141  0.573140
23   -0.159852  0.0372930 -0.052652
24   -0.119720  0.4046226 -0.465447
25   -0.016816  0.0008499  0.005592
```

```
DFBETAS.0 <- DFBETAS[,1]
DFBETAS.1 <- DFBETAS[,2]
DFBETAS.2 <- DFBETAS[,3]
plot(dfit$fit, DFBETAS.0,
 xlab = 'fitted', ylab = 'intercept DFBETA')
abline(h = 2/sqrt(n), lty = 2)
abline(h = -2/sqrt(n), lty = 2)
identify(dfit$fit, DFBETAS.0)
```



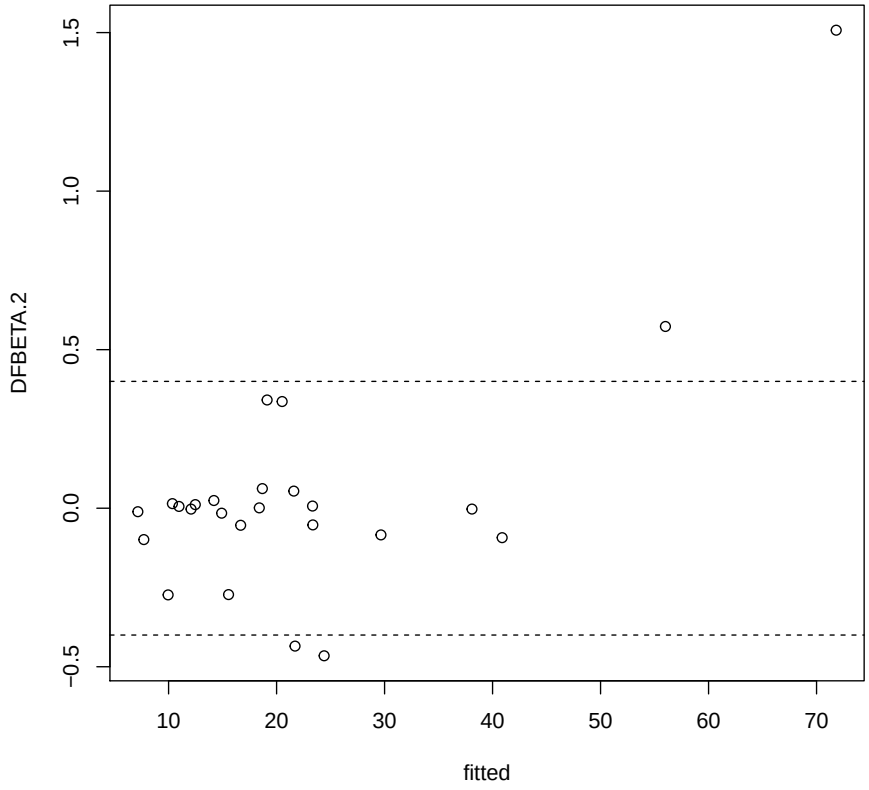

```
integer(0)

plot(dfit$fit, DFBETAS.1,
 xlab = 'fitted', ylab = 'DFBETA.1')
abline(h = 2/sqrt(n), lty = 2)
abline(h = -2/sqrt(n), lty = 2)
identify(dfit$fit, DFBETAS.1)
```




```
integer(0)

plot(dfit$fit, DFBETAS.2,
 xlab = 'fitted', ylab = 'DFBETA.2')
abline(h = 2/sqrt(n), lty = 2)
abline(h = -2/sqrt(n), lty = 2)
identify(dfit$fit, DFBETAS.2)
```




```
integer(0)
```

1.10 DFFITS residuals

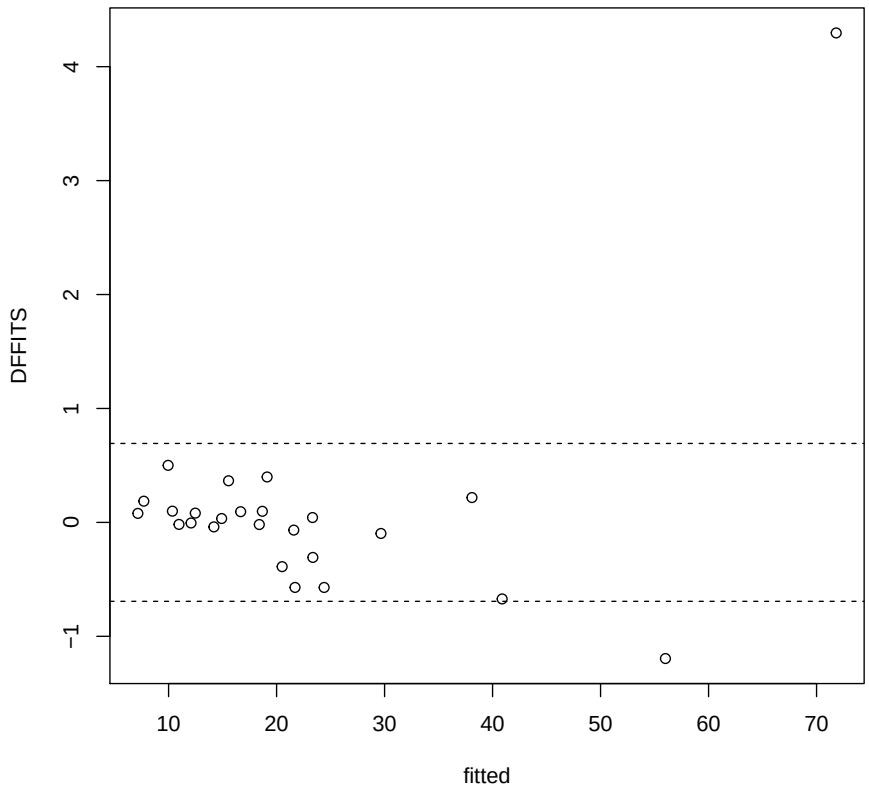
```
dfit <- lm(time ~ cases + distance, data=delivery)
n<-length(delivery$time)
n

[1] 25

p<-3
DFFITS <- dffits(dfit)
DFFITS
```

```
        1         2         3         4         5         6         7
-0.570850  0.098619 -0.005204  0.500802 -0.039459 -0.018779  0.078990
        8         9        10        11        12        13        14
 0.093761  4.296081  0.398713  0.217953 -0.067670  0.081259  0.097363
       15        16        17        18        19        20        21
 0.042584 -0.097160  0.033916  0.365309  0.186168 -0.671771 -0.388501
       22        23        24        25
-1.195036 -0.307539 -0.571140 -0.017626
```

```r
plot(dfit$fit, DFFITS,
 xlab = 'fitted', ylab = 'DFFITS')
abline(h = 2 * sqrt(p/n), lty = 2)
abline(h = -2 * sqrt(p/n), lty = 2)
identify(dfit$fit, DFFITS)
```



```
integer(0)
```