Plot Diagnostics for an lm Object

**Description**

Six plots (selectable by which) are currently available: a plot of residuals against fitted values, a Scale-Location plot of *sqrt(| residuals |)* against fitted values, a Normal Q-Q plot, a plot of Cook's distances versus row labels, a plot of residuals against leverages, and a plot of Cook's distances against leverage/(1-leverage). By default, the first three and 5 are provided.

**Usage**

## S3 method for class 'lm'

plot(x, which = c(1:3, 5),

caption = list("Residuals vs Fitted", "Normal Q-Q",

"Scale-Location", "Cook's distance",

"Residuals vs Leverage",

expression("Cook's dist vs Leverage " \* h[ii] / (1 - h[ii]))),

panel = if(add.smooth) panel.smooth else points,

sub.caption = NULL, main = "",

ask = prod(par("mfcol")) < length(which) && dev.interactive(),

...,

id.n = 3, labels.id = names(residuals(x)), cex.id = 0.75,

qqline = TRUE, cook.levels = c(0.5, 1.0),

add.smooth = getOption("add.smooth"), label.pos = c(4,2),

cex.caption = 1)

**Arguments**

|  |  |
| --- | --- |
| x | lm object, typically result of [lm](http://127.0.0.1:32674/help/library/stats/help/lm) or [glm](http://127.0.0.1:32674/help/library/stats/help/glm). |
| which | if a subset of the plots is required, specify a subset of the numbers 1:6. |
| caption | captions to appear above the plots; [character](http://127.0.0.1:32674/help/library/stats/help/character) vector or [list](http://127.0.0.1:32674/help/library/stats/help/list) of valid graphics annotations, see [as.graphicsAnnot](http://127.0.0.1:32674/help/library/stats/help/as.graphicsAnnot). Can be set to "" or NA to suppress all captions. |
| panel | panel function. The useful alternative to [points](http://127.0.0.1:32674/help/library/stats/help/points), [panel.smooth](http://127.0.0.1:32674/help/library/stats/help/panel.smooth) can be chosen by add.smooth = TRUE. |
| sub.caption | common title—above the figures if there are more than one; used as sub (s.[title](http://127.0.0.1:32674/help/library/stats/help/title)) otherwise. If NULL, as by default, a possible abbreviated version of deparse(x$call) is used. |
| main | title to each plot—in addition to caption. |
| ask | logical; if TRUE, the user is *ask*ed before each plot, see [par](http://127.0.0.1:32674/help/library/stats/help/par)(ask=.). |
| ... | other parameters to be passed through to plotting functions. |
| id.n | number of points to be labelled in each plot, starting with the most extreme. |
| labels.id | vector of labels, from which the labels for extreme points will be chosen. NULL uses observation numbers. |
| cex.id | magnification of point labels. |
| qqline | logical indicating if a [qqline](http://127.0.0.1:32674/help/library/stats/help/qqline)() should be added to the normal Q-Q plot. |
| cook.levels | levels of Cook's distance at which to draw contours. |
| add.smooth | logical indicating if a smoother should be added to most plots; see also panel above. |
| label.pos | positioning of labels, for the left half and right half of the graph respectively, for plots 1-3. |
| cex.caption | controls the size of caption. |

**Details**

sub.caption—by default the function call—is shown as a subtitle (under the x-axis title) on each plot when plots are on separate pages, or as a subtitle in the outer margin (if any) when there are multiple plots per page.

The ‘Scale-Location’ plot, also called ‘Spread-Location’ or ‘S-L’ plot, takes the square root of the absolute residuals in order to diminish skewness (*sqrt(|E|)*) is much less skewed than *| E |* for Gaussian zero-mean *E*).

The ‘S-L’, the Q-Q, and the Residual-Leverage plot, use *standardized* residuals which have identical variance (under the hypothesis). They are given as *R[i] / (s \* sqrt(1 - h.ii))* where *h.ii* are the diagonal entries of the hat matrix, [influence](http://127.0.0.1:32674/help/library/stats/help/influence)()$hat (see also [hat](http://127.0.0.1:32674/help/library/stats/help/hat)), and where the Residual-Leverage plot uses standardized Pearson residuals ([residuals.glm](http://127.0.0.1:32674/help/library/stats/help/residuals.glm)(type = "pearson")) for *R[i]*.

The Residual-Leverage plot shows contours of equal Cook's distance, for values of cook.levels (by default 0.5 and 1) and omits cases with leverage one with a warning. If the leverages are constant (as is typically the case in a balanced [aov](http://127.0.0.1:32674/help/library/stats/help/aov) situation) the plot uses factor level combinations instead of the leverages for the x-axis. (The factor levels are ordered by mean fitted value.)

In the Cook's distance vs leverage/(1-leverage) plot, contours of standardized residuals that are equal in magnitude are lines through the origin. The contour lines are labelled with the magnitudes.

**Author(s)**

John Maindonald and Martin Maechler.

**References**

Belsley, D. A., Kuh, E. and Welsch, R. E. (1980) *Regression Diagnostics.* New York: Wiley.

Cook, R. D. and Weisberg, S. (1982) *Residuals and Influence in Regression.* London: Chapman and Hall.

Firth, D. (1991) Generalized Linear Models. In Hinkley, D. V. and Reid, N. and Snell, E. J., eds: Pp. 55-82 in Statistical Theory and Modelling. In Honour of Sir David Cox, FRS. London: Chapman and Hall.

Hinkley, D. V. (1975) On power transformations to symmetry. *Biometrika* **62**, 101–111.

McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models.* London: Chapman and Hall.

**See Also**

[termplot](http://127.0.0.1:32674/help/library/stats/help/termplot), [lm.influence](http://127.0.0.1:32674/help/library/stats/help/lm.influence), [cooks.distance](http://127.0.0.1:32674/help/library/stats/help/cooks.distance), [hatvalues](http://127.0.0.1:32674/help/library/stats/help/hatvalues).

**Examples**

require(graphics)

## Analysis of the life-cycle savings data

## given in Belsley, Kuh and Welsch.

lm.SR <- lm(sr ~ pop15 + pop75 + dpi + ddpi, data = LifeCycleSavings)

plot(lm.SR)

## 4 plots on 1 page;

## allow room for printing model formula in outer margin:

par(mfrow = c(2, 2), oma = c(0, 0, 2, 0))

plot(lm.SR)

plot(lm.SR, id.n = NULL) # no id's

plot(lm.SR, id.n = 5, labels.id = NULL) # 5 id numbers

## Was default in R <= 2.1.x:

## Cook's distances instead of Residual-Leverage plot

plot(lm.SR, which = 1:4)

## Fit a smooth curve, where applicable:

plot(lm.SR, panel = panel.smooth)

## Gives a smoother curve

plot(lm.SR, panel = function(x, y) panel.smooth(x, y, span = 1))

par(mfrow = c(2,1)) # same oma as above

plot(lm.SR, which = 1:2, sub.caption = "Saving Rates, n=50, p=5")