Improved Spread-Location Visualization

September 11, 1997

Abstract: Cleveland (1979, 1993) introduced the spread-location plot as a diagnostic plot suitable for many types of fitted statistical models. The spread-location plot which plots the absolute residual or square-root absolute residual vs. fitted value along with a robust loess smooth is a useful replacement for the customary practice of plotting residuals vs. fitted values. In this note, we show that neither absolute residual or square-root absolute residual is always appropriate for error distributions likely to be encountered in actual applications. Instead we recommend the simple practice of examining the distribution of the absolute residuals with a boxplot and choosing whatever transformation is necessary to obtain symmetry before constructing the spread-location plot. We conclude with an illustrative example.

Key Words: Loess; Model diagnostic check; Monotone spread; Skewness coefficient; Variance stabilization; Visualizing data

1. Introduction

It is standard practice in regression analysis to plot the residuals against the fitted values. If an increase in variability is detected in this plot then a type of heteroscedasticity exists which can usually be removed by a power transformation (Bartlett, 1947). Cleveland (1979) pointed out that a drawback of this approach is that simply more data at high fitted values can give the impression of an increase in variability. So Cleveland (1979) suggested plotting the absolute residuals. In this plot an increase in variance corresponds to a monotonic trend. Cleveland (1979) introduced the robust loess nonparametric regression smooth to visually assess the trend in this plot. Since the distribution of the absolute residuals are often skewed, Cleveland (1993) recommends plotting the square root absolute residual against fitted value. Lack of skewness is important not only in improving our ability to visually assess the spread-location plot but it is also assumed when fitting a robust loess curve. In this note we show that the square root transformation may not always be the best choice and we recommend the simple expedient of using a boxplot to choose a suitable transformation.

2. Transformation to Symmetry of Absolute Residuals

Let E, the residual, be a random variable with mean zero and let $A = |E|$. In many cases the distribution of A will be positively skewed. An obvious exception would be the case where E has a uniform distribution. In general, we seek a power transformation of A, $A^{(p)} = A^p,\ p \neq 0$ and $A^{(p)} = \log(A),\ p = 0$, for which the lack of symmetry is reduced. It is assumed that with probability 1, $A > 0$.

Consider the case where E has probability density function $f(e)$ and we will assume that the third moment exists. Then the density function of A is $g(a) = f(a) + f(-a)$ which reduces to $g(a) = 2f(a)$ when f is symmetric. The density function for $B = A^{(p)}$ is given by $h(b) = qb^{q-1}g(b)$, where $q = 1/p$ and $p \neq 0$. When $p = 0$, $h(b) = e^b g(e^b)$. Using a symbolic program such as *Mathematica*, we can easily evaluate the skewness coefficient of transformed variable B,

$$\sqrt{\beta}_1 = \frac{\mathrm{E}\left\{(B - E\{B\})^3\right\}}{\mathrm{Var}\,(B)^{3/2}},$$

Table 1. Skewness Coefficient of $|E|^p$

Distribution of E	$p = 1$	$p = 0.5$	$p = 0.4$	$p = 0.3\dot{3}$	$p = 0.25$	$p = 0$
Gaussian	0.99527	0.084073	−0.134382	−0.294799	−0.519216	−1.53514
t_5	2.54964	0.582627	0.276566	0.068672	−0.204957	−1.31309
5% 3 σ contamination	2.67413	0.587239	0.251966	0.028675	−0.260041	−1.13955
Laplace	2	0.631111	0.358632	0.168103	−0.087237	−1.39539

for any particular distribution of E. While only good numerical integration techniques are required to get the results shown in Table 1, the computations are more easily performed in a symbolic computation environment where sophisticated symbolic integration routines are available. Most of the results in Table 1 were obtained symbolically. For convenience the decimal approximations are given. A *Mathematica* notebook containing the detailed derivations of all results in Table 1 is available from my homepage.

As shown in Table 1, for the normal distribution, the square root transformation suggested by Cleveland (1993) is appropriate. But for other distributions, there are better choices. For example, fat-tailed distributions often occur in practice (Tukey, 1960) and Table 1 shows that if the errors are approximated either by a t-distribution on 5 df or by a contaminated normal distribution with a 5% probability of a scale inflation by a factor of 3, then a cube root transformation does better. For Laplace errors, $p = 0.25$ does better.

3. Illustrative Example

The ethanol data of Cleveland (1993, p.217, Figure 4.33) provides as a nice illustration. The boxplots in Figure 1 below show that a log transformation, $p = 0$, makes the absolute error distribution approximately symmetric. Figure 2 compares the spread-location plots using a square-root and log transformation. Notice that the plot with the log transformation is more symmetrically distributed about the fitted curve. Since the amount of variation exhibited in the data is still quite large, compared to the curve, it is safe to assume that monotone spread is not present.

In practice, if a large degree of monotone spread is detected, it may be desirable to use a slightly different power transformation than that initially chosen on the absolute residuals in order to get symmetric deviations from the loess trend on the spread-location visualization.

Figures 1 and 2 here

REFERENCES

Bartlett, M.S. (1947), "The Use of Transformations", *Biometrics*, 3, 37–52.

Cleveland, W. S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots", *Journal of the American Statistical Association*, 74, 829–836.

Cleveland, W. S. (1993), *Visualizing Data*. Summit, New Jersey: Hobart Press.

Tukey, J. W. (1960), "A survey of sampling from contaminated distributions" in Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, Edited by I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow and H. B. Mann, Standford University Press, Standford.

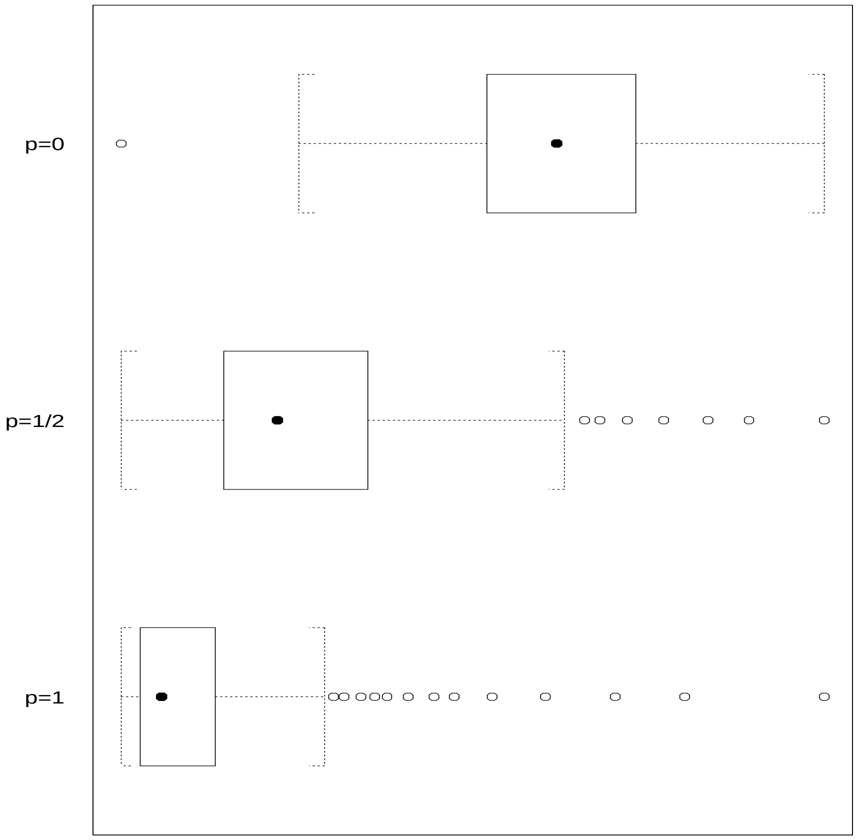

Figure 1: Boxplots absolute residuals and transformed absolute residuals. Data rescaled to fit along common axis.

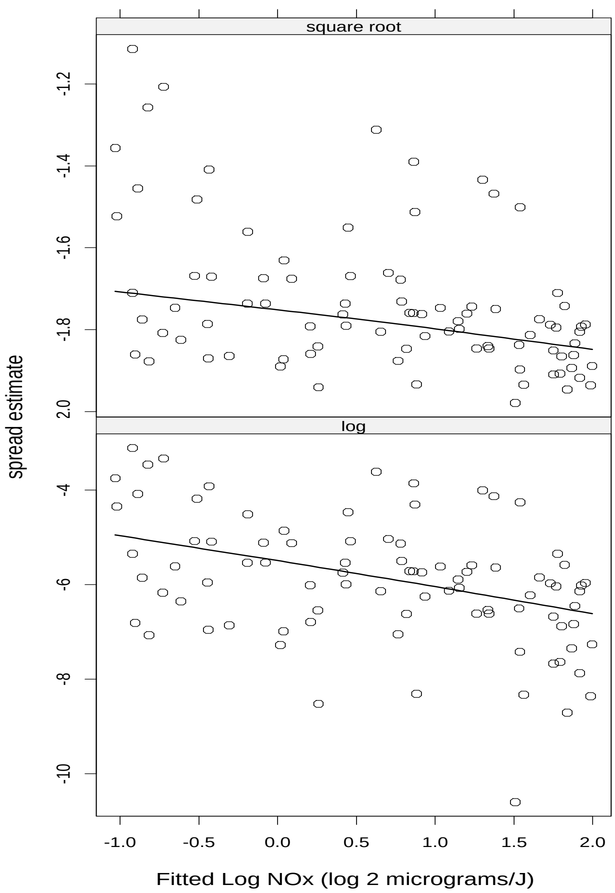

Figure 2: Comparison of spread-location diagnostic plots for the ethanol data