Diagnostic Check For Monotone Spread

A. Ian McLeod
Department of Statistical and Actuarial Sciences
University of Western Ontario
London, Ontario N6A 5B7
Canada

July 10, 1995

Citation: *Computational Statistics & Data Analysis* Vol. 26, pp. 437–443.

Abstract: The spread-location (s-l) plot of Cleveland (1993) is a useful in visually detecting the possible presence of monotone spread. In this article, a diagnostic test statistic based on a signed-power-of-the-mean model is developed which can be used to check if the presence of monotone spread is statistically significant and can suggest possible transformations to remove heteroscedasticity caused by montone spread.

Previously, Carrol and Rupert (1988, p.40, Figure 2.6) have used a similar diagnostic check for regression models. However we show in this article by simulation that robust fitting yields a better estimate. The theorem of Li and Duran (1989) implies that in many situations the diagnostic check will produce a useful estimate of the required power transformation. The usefulness of this estimate of the power transform is also illustrated with the poison data analyzed by Box and Cox (1964). Further examples are given which demonstrate additional insight gained over the original author's analysis from regression and ARIMA modelling.

Documented Splus and Fortran code which implements this test is available from Statlib and my own home page, http://www.stats.uwo.ca/HOMEPAGES/Faculty/mcleod .

Key words: ARIMA models; Diagnostic check; Generalized linear and additive models; Loess; MARS; Monotone spread; Variance stabilizing transformation

1. Introduction

Many statistical models such as in linear and nonlinear regression and linear and nonlinear time series analysis as well as generalized linear/additive models, loess, MARS and smoothing splines produce fitted values for which it is assumed that the variance of the residuals or deviance residuals do not depend on the fitted value. An important type of dependence, monotone spread, occurs when the variance in the residuals tends to increase or decrease with the fitted value (Cleveland, 1993, p.50). Monotone spread is often indicative of other lack of fit problems including skewness of residuals and lack of additivity in the model. Most textbooks on regression mention the importance of plotting the residuals vs. fitted values to look for fan-shaped patterns or other indications of non-constant variance. The s-l plot introduced by Cleveland (1993) is better suited to detecting monotone spread. The purpose of this article is to suggest a modification of Cleveland's method which makes s-l plot more useful in estimating a suitable power transformation and also to provide a significance test of the null hypothesis of model adequacy. In the terminology of Box (1980), a specific diagnostic check for monotone spread is provided.

2. S-L Diagnostic Check

The fitted values and residuals, denoted by, $\hat{y}_i$ and $\hat{e}_i$, $i = 1, \ldots, n$ provide estimates, $\hat{y}_i$ and $|\hat{e}_i|$, of location and scale. In Cleveland's version of the s-l plot, $\sqrt{|\hat{e}_i|}$ is plotted against $\hat{y}_i$ and a robust loess curve with a span of 1 or 2 is fit to aid in discerning any trend. The square root of the absolute residual is used because it makes the distribution of the absolute residual more symmetric. Cleveland recommends that if monotone trend seems apparent, a power transformation can be used to improve the fit in many situations. Our diagnostic check shows what range of power transformations may be suitable.

This diagnostic check is based on the power-of-the-mean model (Carroll and Ruppert, 1988, equation 2.5) which is derived from Bartlett's (1947) theory of variance stabilizing transformations. The power-of-the-mean model states that if a positive random variable, Y, has the property that

$$\mathrm{sd}\,(Y) \propto E(Y)^{1-\lambda} \tag{1}$$

then $Y^{(\lambda)}$ has approximately constant variance, where $Y^{(\lambda)}$ denotes the power transformation,

$$Y^{(\lambda)} = \begin{cases} \frac{Y^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0, \\ \log(Y), & \text{if } \lambda = 0. \end{cases}$$

On the s-l plot, $\lambda = 1 - \beta$, where β is the slope of the regression line of $\log|\hat{e}_i|$ on $\log \hat{y}_i$, estimates λ.

The power-of-the-mean model only applies to positive valued random variables and so one difficulty that occassionally occurs in practice is that the expected value of the fit is assumed to be positive but its estimated value is actually slightly negative. To deal with this situation we introduce the family of signed power transformations,

$$Y^{(\lambda)} = \begin{cases} \mathrm{sign}\,(Y)\frac{|Y|^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0, \\ \mathrm{sign}\,(Y)\log|Y|, & \text{if } \lambda = 0. \end{cases}$$

For signed power transformations, the signed power-of-the-mean model may be shown to be

$$\mathrm{sd}\,(Y) \propto |E(Y)|^{1-\lambda}. \tag{2}$$

Simply using a signed power transformation accomplishes what is needed without the trouble of fiddling with the data by adding some arbitrary constant to the data.

Note that the residuals, $\hat{e}_i$, can not be assumed to be approximately normal since in many situations we know that if there is monotone spread, the residuals will often exhibit non-normal features. Often $\log|\hat{e}_i|$ will also not be approximately normal either. Thus the use of a least-squares line suggested by Carroll and Rupert (1988, p.40 and Figure 2.6) can be improved by using a resistant/robust line fit. Cleveland (1993) has advocated bisquare fitting for resistance/robustness. As pointed out by Mosteller and Tukey (1978, p.378), bisquare fitting provides robustness of both validity and efficiency. Simulation experiments with various regression and anova models have confirmed that the improvement using bisquare over least-squares line fitting in the s-l plot is substantial. The estimated power transformation, $\hat{\lambda}$, is given by one minus the slope of the fitted line. The bootstrap BC_a method (Efron and Tibshirani, 1993) can be used to obtain confidence intervals for λ. At least 1,000 bootstrap iterations are needed. The ABC method of Efron and Tibshirani (1993) could be used to reduce the number of iterations.

This diagnostic test can be viewed and justified as a pure significance test of the null hypothesis that there is no monotone spread in the residuals. As a pure significance test it needs no further justification. However, it is hoped that the estimate of λ provided by this diagnostic will be a useful estimate of the needed power transformation should monotone spread occur. The result of Li and Duran (1989) shows that this will be the case in many situations. Basically, Li and Duran's result implies that in many situations where the model has been mis-specified the resulting fitted values, $\hat{y}$, are proportional to $E(Y)$. Li and Duran (1989) showed that this is exactly so for linear regression with a mis-specified link function in which the explanatory variables have an elliptical distribution. Using an incorrect power transformation often approximates this model so that even in the mis-specified model eq. (2) is approximately true. The next example illustrates that this method does in fact often give a useful estimate of λ. The last two examples show how additional insight over the original analysis could be obtained by using s-l plot diagnostics.

3. SIMULATION EXPERIMENTS

To compare the least-squares and bisquare methods of estimating λ in the s-l plot consider the twoway model $y_{ij} = \exp(\mu + \alpha_i + \beta_j + e_{ij})$ where $\mu = 10$, $\alpha_i = (i-3)$, $i = 1, \ldots, 5$, $\beta_j = (j-3)$, $j = 1, \ldots, 5$ and e_{ij} are independent normal random variables with mean 0 and variance 1. This model was simulated 10^4 times using the Splus random number generator *rnorm* and an additive twoway anova model was fit to the y_{ij}. The s-l plot was used to estimate $\lambda = 0$ using a least squares line fit and a bisquare fit. The boxplot of the estimate, $\hat{\lambda}$, shown in Figure 1, shows the superiority of bisquare fitting although there is some undesirable bias apparent with both estimators. However, this bias would decrease as the sample size increases.

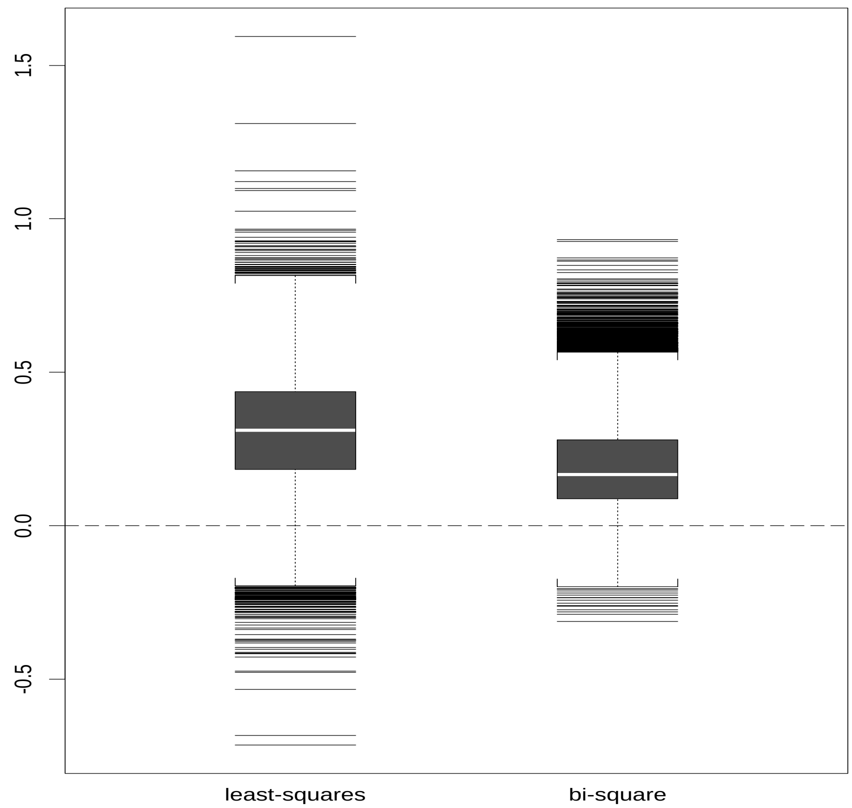

Figure 1: Boxplot of estimate, $\hat{\lambda}$, of $\lambda = 0$ in 10,000 simulations.

The estimated relative efficiency of the bisquare method vs. the least squares method as judged by the mean-square error criterion is the ratio of the observed mean squared error of the least squares method divided by the bisquare method. For this simulation, the estimated relative efficiency is 2.19 with a standard error 0.026. As is evident from Figure 1, the bisquare method produces a better estimate. See the appendix for a general formula for calculating the standard error of an estimated efficiency.

For this particular model, λ could be even more efficiently estimated using the Box-Cox analysis (Box and Cox, 1964). The comparison-value method (Emerson and Hoaglin, pp.200-201) could also be used. However, the purpose of this experiment was just to compare the two methods of estimating λ on a s-l plot.

Similar simulation experiments with other models, including ARIMA and regression models, and with other distributions for the error terms, including t and contaminated normal, and with other values of the transformation parameter, λ, all show that, the bisquare method is significantly more efficient than least squares for estimating λ in all cases.

4. Illustrative Examples

For efficiency the bootstrapping of the bi-square fit is performed using Fortran code which is interfaced to the Splus function for s-l plotting.

4.1 Rubber Abrasion Loss Data

The rubber abrasion loss data consists of a response, abrasion loss, and two explanatory variables, tensile strength and hardness. After deleting 3 outlying data points, Cleveland (1993, p.209) fit a piecewise multiple linear regression model using bisquare fitting. Based on an s-l plot, (Cleveland, 1993, Figure 4.26) which is also reproduced in Figure 2, Cleveland comments that the magnitude of the downward monotone spread is too small to be significant. However, it is not clear cut whether or not there is a statistically significant slope and such diagnostic information is certainly of value.

In this case $\hat{\lambda} = 1.37$ and based on 25,000 bootstrap iterations the 95% BC_a confidence interval for λ is $(0.92, 2.08)$. So as Cleveland suggested there is no strong indication of monotone spread.

4.2 Poison Data

Box and Cox (1964) found that the optimal Box-Cox transformation for this data was $\lambda = -0.75$ with a 95% confidence interval, $(-1.13, -0.37)$. After fitting a model with second order interactions included, our diagnostic statistic is $\hat{\lambda} = -0.99$ with 95% BC_a confidence interval, $(-1.53, -0.47)$ using 1000 bootstrap iterations.

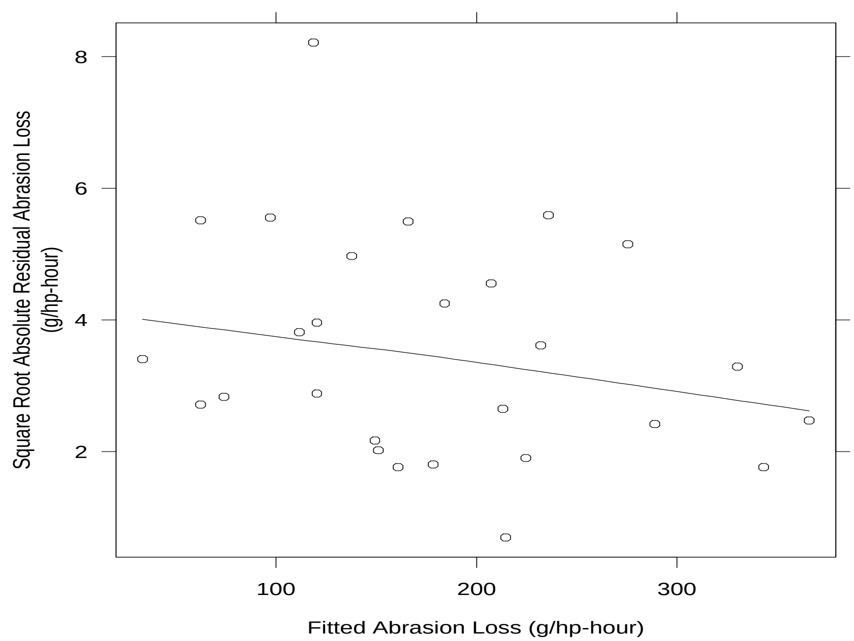

Figure 2: S-L Plot of Cleveland's Model for the Rubber Abrasion-Loss Data.

4.3 Ganglion Data

Cleveland (1993, p.105) fits a quadratic polynomial model to a response variable, CP-ratio, and an explanatory variable, area. Cleveland's s-l plot analysis suggested that the fit exhibits monotone spread and so he used a log transformation. Our diagnostic check statistic for this model is $\hat{\lambda} = -0.29$ with 95% BC_a confidence interval, $(-1.64, 0.57)$ using 1000 bootstrap iterations. This suggests that a reciprocal transformation could also be used. Since the response is a ratio, the reciprocal also has a natural interpretation. The reciprocal transformation is also suggested as a possible transformation by the Box-Cox likelihood analysis.

4.4 Sales of Company X

In a case study of the monthly sales of a company Chatfield and Prothero (1973) first took a log transformation of the data and then fit several multiplicative seasonal ARIMA time series models (Box and Jenkins, 1976). Their model A (Chatfield and Prothero, 1973, p.307), was a seasonal ARIMA $(1,1,0)(0,1,1)_{12}$ which passed all standard residual diagnostic checks for ARIMA models but which failed to produce acceptable forecasts. Our diagnostic check of this model yields, $\hat{\lambda} = 2.57$ with 90% and 95% BC_a confidence intervals $(1.15, 3.79)$ and $(0.83, 4.08)$, respectively. 1000 bootstrap iterations. This suggests that perhaps there is a problem with using the log transformation. In fact, Tunnicliffe-Wilson (1973) pointed out that the problem was that the log transformation had overtransformed the data and that a Box-Cox analysis suggested a different transformation.

5. CONCLUSION

The s-l diagnostic test statistic which is based on an estimate of λ in the signed power-of-the-mean model enhances the usefulness of the s-l diagnostic plot. The s-l diagnostic plot shows graphically whether or not monotone spread exists and also if the power-of-the-mean model is appropriate. The diagnostic test checks if this relationship is statistically significant. The statistical significance is difficult and not always possible to estimate based solely on the plot, so the diagnostic test is a useful adjunct to the plot.

There are various other methods of estimating power transformations from data. The most widely used when a normal distribution is assumed for the transformed data is the likelihood method of Box and Cox (1964) and in many situations, if it is available, it is the best method. Under more general distributional assumptions there are methods of transforming to remove interaction (Mosteller and Tukey, 1978, p.201; Emerson and Hoaglin, 1983, p.200) or to enhance the symmetry of the data (Tukey, 1977; Velleman and Hoaglin, 1981, p.50; Hinkley, 1975). Box (1980) and Cook and Weisberg (1994) have suggested other graphical methods for diagnostic checking and visualizing transformations. The s-l method has the advantages of simplicity and wide applicability. The s-l diagnostic test and plot is based on a direct visualization of monotone spread which seems to occur often in practice and in many modelling situations. S-l plots are also useful in exploratory data analysis (Mosteller and Tukey, 1978; Emerson and Strenio,1983; Emerson, 1983).

References

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

Box, G.E.P. (1980), Sampling and Bayes' Inference in Scientific Modelling and Robustness *Journal of the Royal Statistical Society*, A, 143, 383–430.

Box, G.E.P. and Cox, D.R. (1964), Analysis of Transformations, *Journal of the Royal Statistical Society*, B 26, 211-252.

Box, G.E.P. and Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

Carroll, R.J. and Ruppert, D. (1988), *Transformation and Weighting in Regression*, London: Chapman and Hall.

Chatfield, C. and Prothero, D.L. (1973), Box-Jenkins Seasonal Forecasting: Problems in a Case Study, *Journal of the Royal Statistical Society*, A 136, 295–308.

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

Cook, R.D. and Weisberg, S. (1994), Transforming a response variable for linearity, *Biometrika* 81, 731–737.

Emerson, J.D. and Strenio, J. (1983), "Boxplots and Batch Comparisons" in *Understanding Robust and Exploratory Data Analysis*, Edited by D. C. Hoaglin, F. Mosteller and J. W. Tukey, New York, N.Y.: Wiley.

Emerson, J.D. (1983), "Mathematical Aspects of Transformation" in *Understanding Robust and Exploratory Data Analysis*, Edited by D. C. Hoaglin, F. Mosteller and J. W. Tukey, New York, N.Y.: Wiley.

Emerson, J.D. and Hoaglin, D.C. (1983), "Analysis of Two-way Tables by Medians" in *Understanding Robust and Exploratory Data Analysis*, Edited by D. C. Hoaglin, F. Mosteller and J. W. Tukey, New York, N.Y.: Wiley.

Efron, B. and Tibshirani, R.J. (1993), *An Introduction to the Bootstrap*, New York: Chapman and Hall.

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

Li, K.C. and Duan, N. (1989), Regression analysis under link violation, *Annals of Statistics* 17, 1009–1052.

Mosteller, F. and Tukey, J.W. (1977), *Data Analysis and Regression*, Reading: Addison-Wesley.

Tukey, J.W. (1977), *Exploratory Data Analysis*, Reading: Addison-Wesley.

Tunnicliffe-Wilson, G. (1973), Discussion of "Box-Jenkins Seasonal Forecasting: Problems in a Case Study", by Chatfield and Prothero, *Journal of the Royal Statistical Society*, A 136, 315–319.

Velleman, P.F. and Hoaglin, D.C. (1981), *Applications, Basics and Computing of Exploratory Data Analysis*, Boston: Duxbury.