ARTICLE TEMPLATE

Automated reading of residual plots with computer vision models

Weihao Li[a]

[a]Department of Econometrics and Business Statistics, Monash University, Clayton, VIC, Australia

ARTICLE HISTORY
Compiled November 14, 2023

ABSTRACT
TBD.

KEYWORDS
TBD

1. Introduction

Residuals, within regression analysis, represent the differences between fitted values and observed data points, capturing the unexplained elements in the regression model. The practice of plotting residuals, advocated by influential regression literature (Cook and Weisberg 1982; Draper and Smith 1998; Belsley, Kuh, and Welsch 1980; Montgomery, Peck, and Vining 1982), serves as a standard procedure in regression diagnostics. This visual examination is crucial for identifying deviations from the model assumptions like linearity, homoscedasticity, and normality.

Generating a residual plot in most statistical software is often as straightforward as executing a line of code or clicking a button. However, accurately interpreting a residual plot can be challenging. Consider Figure 1 as an example, the residuals display a triangular shape pointing to the left. While this might suggest heteroskedasticity, it's important to avoid over-interpreting the visual pattern. In this case, the fitted model is correctly specified, and the triangular shape is actually a result of the skewed distribution of the predictors, rather than indicating a flaw in the model.

A residual plot can exhibit various visual features, but it's crucial to recognize that some may arise from the characteristics of predictors and the inherent randomness of the error, rather than indicating a violation of model assumptions (Li et al. 2023). The concept of visual inference, as proposed by Buja et al. (2009), provides an inferential framework to assess whether residual plots indeed contain visual patterns inconsistent with the model assumptions. The fundamental idea involves testing whether the actual residual plot significantly differs visually from null plots, which are created using residuals generated from the null distribution. Typically, this is accomplished through the lineup protocol. In this approach, the real residual plot is embedded within a lineup alongside several null plots. If the real residual plot can be distinguished from the lineup, it provides evidence for rejecting the null hypothesis.
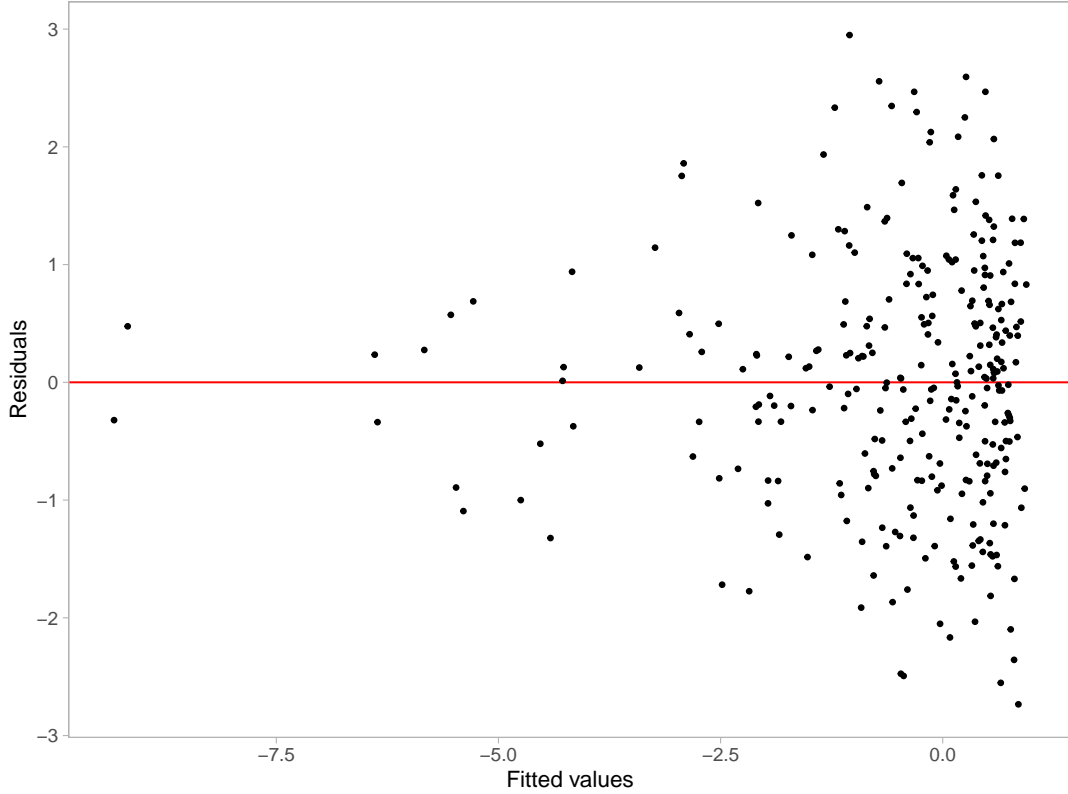

CONTACT Weihao Li. Email: `weihao.li@monash.edu`

Figure 1. An example residual vs fitted values plot (red line indicates 0). The vertical spread of the data points varies with the fitted values. This often indicates the existence of heteroskedasticity.

Delivering a residual plot as a lineup is generally considered a good practice. However, as pointed out by Li et al. (2023), a primary limitation of the lineup protocol is its reliance on human judgments. Unlike conventional statistical tests that can be performed numerically and automatically in statistical software, the lineup protocol requires human evaluation of images. This characteristic makes it less suitable for large-scale applications, given the associated high labour costs and time requirements.

In this paper, we attempt to address this issue by integrating computer vision models into residual plot diagnostics. The paper is structured as follows: ...

2. Methodogology

2.1. *Generation of simulated training data*

2.2. *Architecture of the computer vision model*

2.3. *Training process and hyperparameter tuning*

3. Results

3.1. *Model evaluation*

- Metrics for model performance
- Shap values

- Heatmap

3.2. *Comparison with human visual inference*

3.2.1. *Overview of the human subject experiment*

3.2.2. *Comparison*
- power comparison
- decisions

3.3. *Case study 1: . . .*

3.4. *Case study 2: . . .*

3.5. *Case study 3: datasaurus*

4. Dicussion

5. Conclusion

- Summary of findings
- Contributions to the field
- Future directions for research

References

Belsley, David A, Edwin Kuh, and Roy E Welsch. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity.* John Wiley & Sons.

Buja, Andreas, Dianne Cook, Heike Hofmann, Michael Lawrence, Eun-Kyung Lee, Deborah F Swayne, and Hadley Wickham. 2009. "Statistical inference for exploratory data analysis and model diagnostics." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367 (1906): 4361–4383.

Cook, R Dennis, and Sanford Weisberg. 1982. *Residuals and influence in regression.* New York: Chapman and Hall.

Draper, Norman R, and Harry Smith. 1998. *Applied regression analysis.* Vol. 326. John Wiley & Sons.

Li, Weihao, Dianne Cook, Emi Tanaka, and Susan VanderPlas. 2023. "A Plot is Worth a Thousand Tests: Assessing Residual Diagnostics with the Lineup Protocol." *arXiv preprint arXiv:2308.05964* .

Montgomery, Douglas C, Elizabeth A Peck, and G Geoffrey Vining. 1982. *Introduction to linear regression analysis.* John Wiley & Sons.