Assumptions in Linear Regression

1. There exists a **linear relation** between the response and predictor variable(s).
2. The error terms have **mean 0**.
3. The error terms have the **constant variance**, σ^2.
4. The error terms are uncorrelated.
5. The errors follow a **Normal distribution**.

Assumptions in Linear Regression

We can use a scatterplot of y against x and/or a residual plot of residuals against $\hat{y}$ to assess assumptions 1 to 3.

Examination of Scatterplot

We can examine the scatterplot of y against x to check for assumptions 1 to 3.

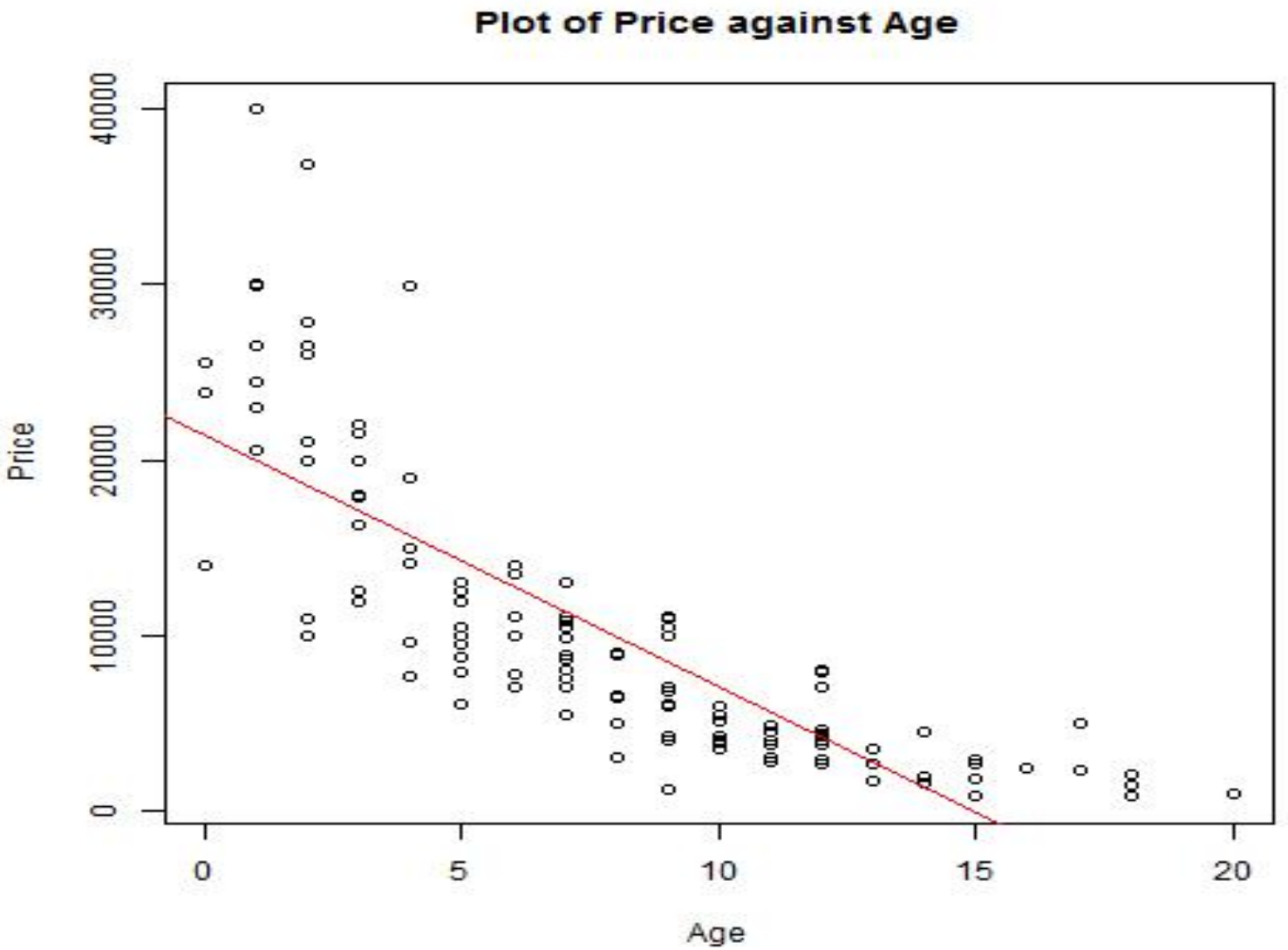
- General pattern is linear (no curved pattern) (assumption 1).
- Data points evenly scattered around fitted line (assumption 2).
- Vertical variation of data points constant (assumption 3).

Examination of Residual Plot

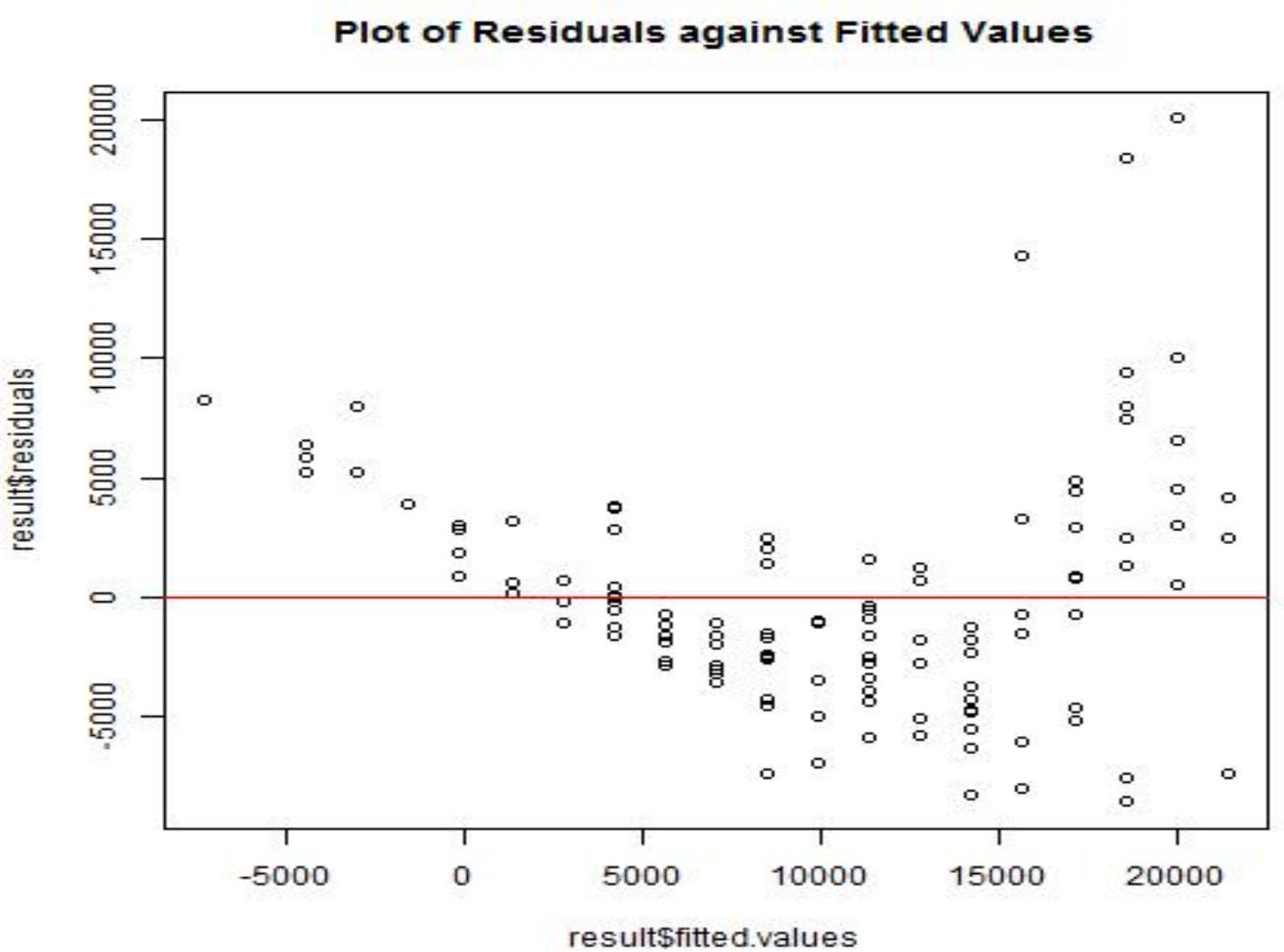
We can examine the plot of residuals against $\hat{y}$ to check for assumptions 1 to 3.

- Residuals evenly scattered around the horizontal axis with no apparent pattern (assumption 1, 2).
- The residuals should have similar vertical variation across the horizontal axis (assumption 3).

Example: Price of Used Mazdas



Example: Price of Used Mazdas

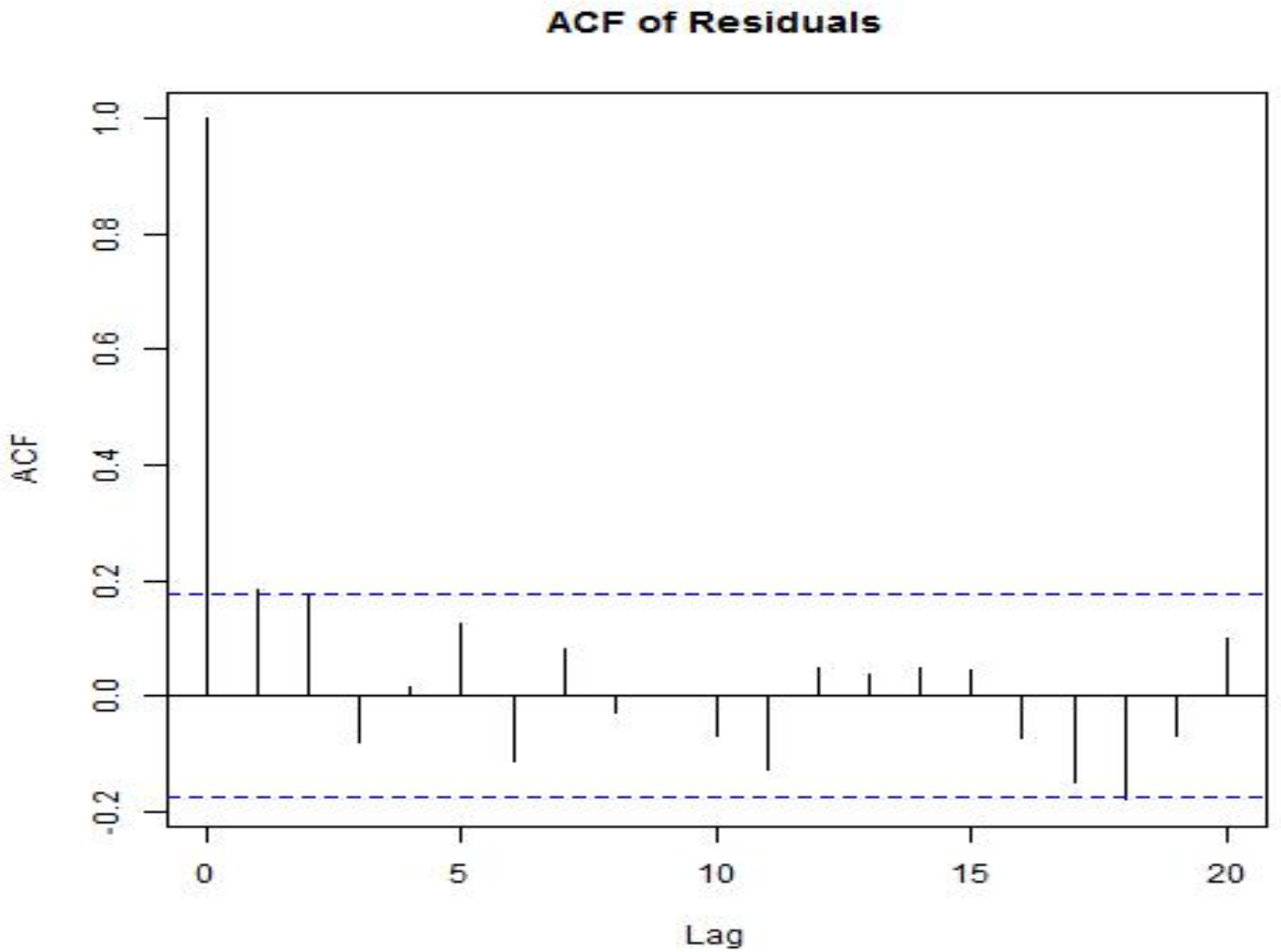


Examination of Autocorrelation Plot

We can examine an autocorrelation plot of the residuals to check for assumption 4.

- Autocorrelation function (ACF) should be insignificant for nonzero lags.

Example: Price of Used Mazdas



Examination of Normal Probability Plot

We can examine a normal probability plot of the residuals to check for assumption 5.

- Normal probability plot for the residuals should fall close to the line representing the expected value under normality (assumption 5).

Example: Price of Used Mazdas