Manual Linear Regression for CBI-CME Velocity Analysis

Kelly French

July 17, 2025

Original Data, No Zeros

The slope m and intercept b of the best-fit line are given by:

$$m = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{-0.04}{1.09 \times 10^{-10}} = -361{,}169{,}714.05$$

$$b = \bar{y} - m\bar{x} = 727.764 - (-361{,}169{,}714.05)(6.08 \times 10^{-7}) = 947.50$$

The correlation coefficient r is:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}} = \frac{-0.04}{\sqrt{(1.09 \times 10^{-10})(1.54 \times 10^{8})}} = -0.31$$

The t-statistic is:

$$t = \frac{m}{\text{SE}_m}$$

$$t = \frac{-361{,}169{,}714.05}{46{,}394{,}577.11} = -7.78$$

The corresponding p-value is:

$$p = 2 \cdot P(T > |t|), \quad T \sim t(n-2)$$

$$p = 2 \cdot P(T > 7.78) = 3.15 \times 10^{-14}$$

Results

- Slope: $m = -361{,}169{,}714.05$
- Intercept: $b = 947.50$
- Correlation coefficient: $r = -0.31$
- p-value: $p = 3.15 \times 10^{-14}$

Comparison with Python Regression

- Slope: manual = −361,169,714.05, Python = −361,169,714.05
- Intercept: manual = 947.50, Python = 947.50
- Correlation coefficient r: manual = −0.305, Python = −0.305
- R^2: manual = 0.0931, Python = 0.0931
- p-value: manual = 3.15×10^{-14}, Python = 3.15×10^{-14}

180° Position Angle Shift From Original Data, No Zeros

The slope m and intercept b are calculated as:

$$m = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{-0.005}{5.40 \times 10^{-11}} = -105{,}969{,}191.42$$

$$b = \bar{y} - m\bar{x} = 727.76 - (-105{,}969{,}191.42)(4.55 \times 10^{-7}) = 775.98$$

The correlation coefficient is:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}} = \frac{-0.005}{\sqrt{(5.40 \times 10^{-11})(1.54 \times 10^{8})}} = -0.06$$

The t-statistic is:

$$t = \frac{m}{\mathrm{SE}_m}$$

$$t = \frac{-105{,}969{,}191.42}{72{,}479{,}871.48} = -1.53$$

The corresponding p-value is:

$$p = 2 \cdot P(T > |t|), \quad T \sim t(n-2)$$

$$p = 2 \cdot P(T > 1.53) = 0.13$$

Results

- Slope: $m = -105{,}969{,}191.42$
- Intercept: $b = 775.98$
- Correlation coefficient: $r = -0.06$
- p-value: $p = 0.13$

Comparison with Python Regression

- Slope: manual = $-105{,}969{,}191.42$, Python = $-105{,}969{,}191.42$
- Intercept: manual = 775.98, Python = 775.98
- Correlation coefficient r: manual = -0.063, Python = -0.063
- R^2: manual = 0.00395, Python = 0.00395
- p-value: manual = 0.127, Python = 0.127