Mathematical Formulas Used

Best Fit Line using Normal Equation

To find the optimal weights $\mathbf{w}$ in linear regression, we use:

$$\boxed{\mathbf{w} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}}$$

- $\mathbf{X}$ is the design matrix.
- $\mathbf{y}$ is the target vector.
- $\mathbf{w}$ is the weight vector.

Residuals

Residuals represent the error between actual and predicted values:

$$\boxed{\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\mathbf{w}}$$

Coefficient of Determination (R^2)

$$\boxed{R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

- $\bar{y}$ is the mean of true values.
- A higher R^2 (closer to 1) indicates better model fit.

Adjusted R^2

$$\boxed{\text{Adjusted } R^2 = 1 - \left(\frac{(1 - R^2)(n - 1)}{n - p - 1} \right)}$$

Where:

- n = number of data points.
- p = number of predictors.

Mean Absolute Percentage Error (MAPE)

$$\boxed{\text{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad \text{for all } y_i \neq 0}$$

Note: Values where $y_i = 0$ are excluded to prevent division by zero.

Accuracy (Derived from MAPE)

$$\boxed{\text{Accuracy} = 100 - \text{MAPE}}$$

Mean Squared Error (MSE)

$$\boxed{\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Root Mean Squared Error (RMSE)

$$\boxed{\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}$$

Mean Absolute Error (MAE)

$$\boxed{\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|}$$