Enhanced Guide to Linear, Huber, and Lasso Regression

1. Introduction to Regression Models

This guide focuses on three core regression methods used in machine learning and statistics: Linear Regression, Huber Regression, and Lasso Regression. Each has different strengths and weaknesses depending on your data quality, dimensionality, and presence of outliers.

2. Linear Regression

Linear Regression models the relationship between a scalar dependent variable and one or more independent variables by fitting a linear equation. It minimizes the sum of squared errors:

$$\min \sum (y_i - \hat{y}_i)^2$$

Pros: Interpretable, fast, works well on clean data.
Cons: Extremely sensitive to outliers and irrelevant variables. Assumes linearity, homoscedasticity, and normality of residuals.
Use when: You have clean, low-noise data and all features are believed to be important.

3. Huber Regression

Huber Regression uses a loss function that is quadratic for small errors and linear for large errors, offering a balance between Linear and Lasso regression:

$$L_\delta(r) = 0.5r^2 \text{ if } |r| \leq \delta, \text{ else } \delta(|r| - 0.5\delta)$$

Delta (δ): The threshold where loss switches from L2 to L1. Common default is $\delta = 1.35$.
Pros: Robust to outliers while maintaining sensitivity to smaller residuals.
Cons: No built-in feature selection, delta needs tuning.
Use when: Data has outliers or heavy-tailed errors, but still follows a mostly linear pattern.

4. Lasso Regression

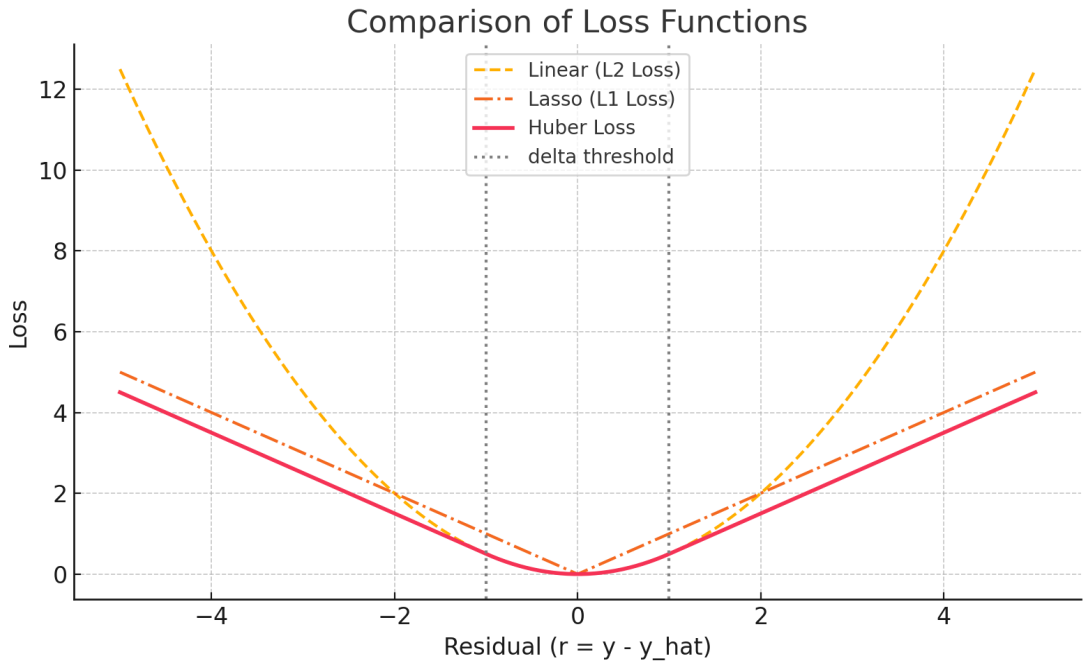
Lasso (Least Absolute Shrinkage and Selection Operator) adds an L1 penalty to the standard least squares cost function:

$$\min \sum (y_i - \hat{y}_i)^2 + \alpha \sum |\beta_j|$$

Alpha (α): Controls the strength of the penalty. Higher values force more coefficients to zero.
Pros: Performs feature selection, simplifies the model, reduces multicollinearity.
Cons: May discard important variables if they're highly correlated; underestimates large coefficients.
Use when: Many features exist and you suspect some are irrelevant or noisy.

5. Visual Comparison of Loss Functions

How each model's loss behaves across residual size:



6. Summary Table

Model	Strengths	Weaknesses	When to Use
Linear	Simple, interpretable	Sensitive to outliers	Clean data
Huber	Robust to outliers	Requires delta tuning	Noisy/robust cases
Lasso	Sparse model, feature selection	May drop correlated features	High-dimensional data

7. Tips for Choosing a Model

- Always scale your features for Lasso and Huber Regression.
- Use cross-validation to select α (for Lasso) and δ (for Huber).
- Start with Linear Regression. Move to Huber if robustness is needed. Use Lasso when you suspect irrelevant features.