

Comparativo de Métodos de Regressão Linear

Método	Quando Usar	Vantagem Principal	Desvantagem Principal
OLS (Ordinary Least Squares)	Quando os dados seguem os pressupostos clássicos (linearidade, homocedasticidade, normalidade dos resíduos)	Simples, interpretável, base de muitos modelos	Sensível a outliers, multicolinearidade e heterocedasticidade
Ridge Regression (L2)	Quando há multicolinearidade entre as variáveis explicativas	Reduz overfitting com penalização suave	Não realiza seleção de variáveis (mantém todas no modelo)
Lasso Regression (L1)	Quando se deseja seleção automática de variáveis	Zera coeficientes irrelevantes	Pode eliminar variáveis úteis com alta correlação
Elastic Net	Quando há muitas variáveis correlacionadas e se deseja regularização + seleção	Combina os pontos fortes de Lasso e Ridge	Mais complexo de ajustar (dois hiperparâmetros a calibrar)
Robust Regression	Quando há outliers nos dados	Não é influenciado por valores extremos	Menos eficiente com dados limpos
Quantile Regression	Quando se quer modelar medianas ou outros quantis , não a média	Permite entender diferentes partes da distribuição	Mais complexo e difícil de interpretar
Weighted Least Squares (WLS)	Quando há heterocedasticidade (variância dos erros não constante)	Corrige problemas de variância desigual	Requer conhecimento prévio dos pesos
Generalized Least Squares (GLS)	Quando há correlação nos resíduos (ex: séries temporais)	Lida com erros correlacionados	Requer modelar corretamente a estrutura dos erros

Bayesian Linear Regression	Quando se deseja incorporar incerteza ou conhecimento prévio	Produce distribuições de probabilidade dos coeficientes	Computacionalmente mais caro e menos intuitivo
Principal Component Regression (PCR)	Quando há muitas variáveis explicativas correlacionadas	Reduz a dimensionalidade de antes da regressão	Perde interpretabilidade dos coeficientes