OLS linear regression assumptions

Assumption	If broken …
Linear relationship between inputs and targets	Inappropriate application/unreliable results; results in non-normality of errors; use a nonlinear modeling technique
N > **p**	Underspecified/unreliable results; use LASSO(L1)/elastic net penalized regression
No strong multicollinearity	Ill-conditioned/unreliable results; Use ridge(L2/Tikhonov)/elastic net penalized regression
No influential outliers	Biased predictions, parameters, and statistical tests; use robust methods, i.e. IRLS, Huber loss, investigate/remove outliers
No strong Heteroskedasticity	Lessened predictive accuracy, invalidates statistical tests
Limited correlation between input rows (no autocorrelation)	Invalidates statistical tests; use time-series methods

Modern approaches – elastic net

$$\tilde{\beta} = \min_{\beta} \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} * \beta_j \right)^2 + \lambda \sum_{j=1}^{p} (\alpha * \beta_j^2 + (1-\alpha) * |\beta_j|) \right\}$$

Modern approaches – elastic net

λ - Controls magnitude of penalties. Variable selection conducted by refitting model many times while varying λ. Decreasing λ allows more variables in the model.

L1/LASSO penalty – for variable selection.

L2/Ridge/Tinkhonov Penalty – helps address multicollinearity.

$$\tilde{\beta} = \min_{\beta} \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} * \beta_j \right)^2 + \lambda \sum_{j=1}^{p} (\alpha * \beta_j^2 + (1 - \alpha) * |\beta_j|) \right\}$$

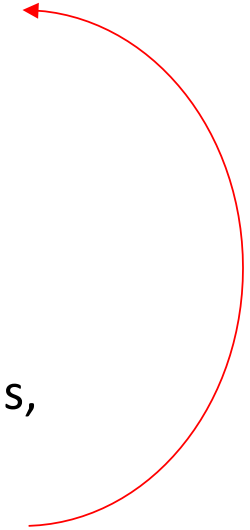
Least squares minimization – finds β's for linear relationship.

α - tunes balance between L1 and L2 penalties.

Modern approaches – iteratively reweighted least squares

Iteratively reweighted least squares complements fitting methods in the presence of outliers by:

- Initially giving all observations equal weight then …
  - Training the model to estimate the β's and find a linear relationship/linear equation
  - Calculate the residuals given these β's/linear equation
  - ***Reweighting*** observations that cause high residuals to have a lower impact in the trained model
  - Re-train to find new β's/linear equation
  - Continue calculating residuals, reweighting observations, and retraining until β's become stable and weighted residuals are small …



iterate