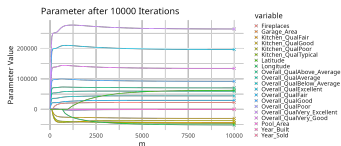


Introduction to Machine Learning

Boosting

Gradient Boosting: CWB and GLMs



Learning goals

- Understand relationship of CWB and GLM

RELATION TO GLM / 2

But: We do not *require* an exponential family distribution and we can - in principle - apply it to any differentiable loss

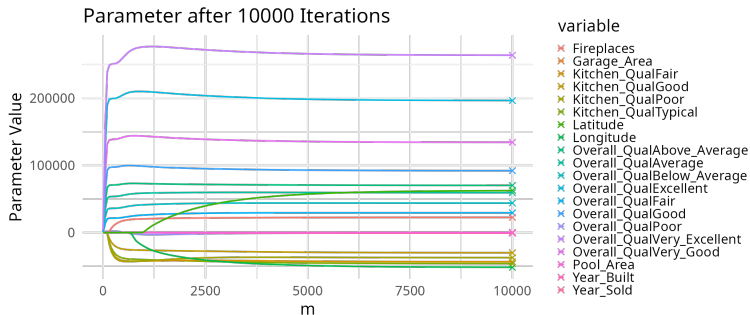
Usually we do not let the boosting model converge fully, but use **early stopping** for the sake of regularization and feature selection.

Even though resulting model looks like a GLM, we do not have valid standard errors for our coefficients, so cannot provide confidence or prediction intervals or perform tests etc. → post-selection inference.



EXAMPLE: CWB PARAMETER CONVERGENCE

The following figure shows the parameter values for $m \leq 10000$ iterations as well as the estimates from a linear model as crosses (GLM with normally distributed errors):



Throughout the fitting of CWB, the parameters estimated converge to the GLM solution. The used data set is Ames Housing.