Introduction to Machine Learning
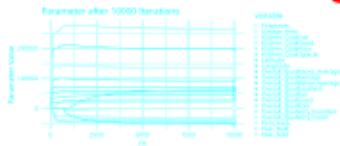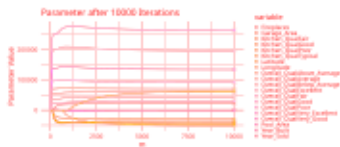
Gradient Boosting: CWB and GLMs

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. $\rightarrow$ post-selection inference.