

# Using Statistics to Determine the Learning Rate for Gradient Descent

*Wednesday, 3 August 2022 17:30 (20 minutes)*

While gradient descent is ubiquitous in Machine Learning, there is no adaptive way to select a learning rate yet. This forces practitioners to do “hyperparameter tuning”. We review how optimization schemes can be motivated using Taylor approximations and develop intuition why this results in unknown hyperparameters. We then replace the Taylor approximation with a statistical Best Linear Unbiased Estimator (BLUE) and derive gradient descent again. But this time with calculable learning rates.

**Primary author:** BENNING, Felix (Universität Mannheim)

**Session Classification:** Session B2 Statistics

**Track Classification:** Statistics