Implicit Regularization: Gradient Flow and Friends
Advanced Topics in Statistical Learning, Spring 2023
Ryan Tibshirani

(This is just an outline of what we will cover. Fully fleshed out notes may be coming at a later point.)

types of implicit regularization:

- early stopping
- weight decay
- dropout
- batch normalization
- bagging / subsampling

why do any of this? the basic idea is to endow regularity in a convenient (implementation-wise or computation-wise) manner

gradient flow & ridge:

- introduce gradient flow
- grad flow for least squares
- matrix exponential facts
- bias and variance
- risk coupling
- optimal tuning
- asymptotics

forward stagewise & lasso:

- explain origin story
- least angle regression
- infinitesimal stagewise
- monotone lasso

generic stagewise (steepest descent):

- where do these connections come from?
- steepest descent is the unifying view
- can also look at infinitesimal dynamics