Although we can’t write down explicit formulas for the bias and variance of the lasso estimate (e.g., when the true model is linear), we know the general trend.

Recall that ˆlasso = argmin 2Rp ky Xk2 2 + kk1

Generally speaking: I The bias increases as (amount of shrinkage) increases I The variance decreases as (amount of shrinkage) increases

What is the bias at = 0? Standard least squares RSS fit of high-order polynomial, lamida=0

The variance at = 1? W=0 for lamida=wuqing

In terms of prediction error (or mean squared error), the lasso performs comparably to ridge regression

The lasso1 estimate is defined as

βˆlasso = argmin β∈Rp ky − Xβk 2 2 + λ X p j=1 |βj |

= argmin β∈Rp ky − Xβ| 2 2 | {z } Loss + λ kβk1 | {z } Penalty