Regularization

Chelsea Parlett-Pelleriti

Review: Loss Functions

MINIMIZE:

$$\sum \left(x_i - \hat{x}_i \right)^2$$

how off we were

true value

model's guess

Overfitting

Ridge

MINIMIZE:

$$\sum \left(x_i - \hat{x}_i \right)^2 + \lambda \sum \beta_j^2$$

how off we were

true value

model's guess

how HARSHLY we penalize

how big the coefs are

LASSO

MINIMIZE:

$$\sum \left(x_i - \hat{x}_i \right)^2 + \lambda \sum |\beta_j|$$

how off we were

true value

model's guess

how HARSHLY we penalize

how big the coefs are

Ridge vs. LASSO