

Regularization

Dr. Chelsea Parlett-Pelleriti

Review: Loss Functions

MINIMIZE:

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

how off
we were

true
value

model's
guess

Overfitting

Ridge

MINIMIZE:

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

Annotations:

- x_i : true value
- \hat{x}_i : model's guess
- λ : how HARSHLY we penalize
- β_j^2 : how big the coefs are
- Red bracket over $(x_i - \hat{x}_i)^2$: how off we were

LASSO

MINIMIZE:

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

Annotations:

- x_i : true value
- \hat{x}_i : model's guess
- λ : how HARSHLY we penalize
- $|B_j|$: how big the coefs are
- Red bracket over $(x_i - \hat{x}_i)^2$: how off we were

Ridge vs. LASSO

