



Regularization

Chelsea Parlett-Pelleriti



Review: Loss Functions

MINIMIZE:

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

Diagram illustrating the components of the loss function:

- A red bracket above the term $(x_i - \hat{x}_i)$ is labeled "how off we were".
- A green arrow points from the text "true value" to x_i .
- An orange arrow points from the text "model's guess" to \hat{x}_i .

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