Linear regression

Swiss Institute of Bioinformatics

- **Regularization**
  - regularization methods seek to both minimize the sum of the squared error of the model on the training data (using ordinary least squares) but also to reduce the complexity of the model (like the number or absolute size of the sum of all coefficients in the model).
  - Two popular examples of regularization procedures for linear regression are:
    - Lasso Regression: where Ordinary Least Squares is modified to also minimize the absolute sum of the coefficients (called L1 regularization).
    - Ridge Regression: where Ordinary Least Squares is modified to also minimize the squared absolute sum of the coefficients (called L2 regularization).
  - These methods are effective to use when there is collinearity in your input values and ordinary least squares would overfit the training data.

Linear regression

Swiss Institute of Bioinformatics

- Prevent *overfitting: regularization:*
  - shrinks the coefficient estimates towards zero
  - this technique discourages learning a more complex model, to avoid the risk of overfitting.

$$\sum_{i}^{n}(y_i - \hat{y}_i)^2 + \lambda \sum_{j}^{p} \|\beta_j\|$$

l1 regularization (Lasso)

$$\sum_{i}^{n}(y_i - \hat{y}_i)^2 + \lambda \sum_{j}^{p} \beta_j^2$$

l2 regularization (Ridge)

λ corresponds to the weight that you put on that regularization

- Hyperparameters
  - Type of regularization
  - λ: regularization strength