Lecture 12:

1. *causation*: what is the effect of a particular variable on an outcome?
2. *prediction*: find some function that provides a good prediction of y as a function of x

Causation: Interested in

Prediction: Interested in

OLS estimates both parameters and their standard deviations

Minimize SSE:

Updating weights:

Penalty term:

: hyperparameter

L1 regularization (Lasso):

Before regularization we rescale our features; convert to zero mean and standardize to unit std. We use StandardScaler in sklearn for this.

Fit to the distribution in the training data first, then rescale train and test!:

<https://stats.stackexchange.com/questions/174823/how-to-apply-standardization-normalization-to-train-and-testset-if-prediction-i>

Otherwise the interacted varaibles are not gaussian distributed.