Get data: (x,y)

Choose loss function L(y,y^)

Choose hypothesis class (linear classifier, decision tree)

Minimize the sum of the losses over all data points

Solve

h is in the set of hypotheses (set:H)

normalization: x – x-bar(mean of x)

covariance matrix เกิดมาจาก matrix multiplication of Vector X Vector Y

A screenshot of a math test

Description automatically generated

What is the joint probability distribution part? (Do we assume uniform distribution?)

If not uniform, How do we derive this formula?

Optimization is the step to minimize empirical risk minimization.