

Supervised Learning: Linear Regression

April 9, 2021

1 Motivation

A college's acceptances are a yes/no decision. The factors the college considers are:

$x_1 := \text{test score} / 10$

$x_2 := \text{grades} / 10$

Can we predict whether a new (x_1, x_2) gets accepted or rejected?

2 Classification Problem

This is a classification problem, so our goal is not to fit a function to the data, but instead fit a function such that it splits the data into discrete, already defined categories. In the case of the college problem, we plot the points on an x_1 vs x_2 plot, and come up with the equation of a line $f(x_1, x_2) = \beta_2 x_2 + \beta_1 x_1 + \beta_0$ that splits the data in half. Then, for any given (x_1, x_2) , our prediction equation is:

$$\hat{y} = \begin{cases} 0 & f(x_1, x_2) \leq 0 \\ 1 & f(x_1, x_2) > 0 \end{cases}$$

2.1 Generalizing to Higher Dimensions

Generally, if each point on our plot has n coordinates, our dividers are going to be n -planes, or an $n-1$ subspace of \mathbb{R}^n