Modeling with Proportionality

1 Introduction

One of the first steps in modeling is to make simplifying assumptions, and one of the simplest assumptions is that one variable is simply a constant multiple of the other. This type of relationship is called **proportionality**.

Definition

Two variables y and x are **proportional** (to each other) if one is always a constant multiple of the other—that is, if

$$y = kx$$

for some nonzero constant k. We write $y \propto x$.

The definition means that the graph of y versus x lies along a straight line through the origin. This graphical observation is useful in testing whether a given data collection reasonably assumes a proportionality relationship. If proportionality is reasonable, a plot of one variable against the other should approximate a straight line through the origin. The constant (k) of proportionality is the slope of this line.

2 Fitting Straight Lines Analytically

2.1 Passes through Origin

Suppose we need to fit a linear function to a given data set that passes through the origin. Then the function is in the form

$$f(x_i) = mx_i$$

where, m is the slope of the fitted function. Then the sum of the squares of the distance between actual and predicted data is given by,

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

where, y_i represents actual data and $f(x_i)$ gives the predicted data. To identify the minimum of the function, we need to calculate the first derivative of the function and set it to zero. But take the derivative concerning what? Obviously, we need to find m. Observe that S is a function of m. So we consider,

derivativeof inner term, then derivative of outer term and then = it to zero to find the min point

$$\frac{dS}{dm} = \sum_{i=1}^{n} 2(y_i - mx_i)(-x_i) = 0$$

refer document to find how the derivation comes

Thus

$$-2\sum_{i=1}^{n}(y_i - mx_i)x_i = 0$$

Hence, we get

$$\sum_{i=1}^{n}(x_iy_i - mx_i^2) = 0$$

By separating summation into two terms,

$$\sum_{i=1}^{n} x_iy_i - m\sum_{i=1}^{n} x_i^2 = 0$$

Thus the slop is,

$$m = \frac{\sum_{i=1}^{n} x_iy_i}{\sum_{i=1}^{n} x_i^2}$$

take second derivative to find whether a min or max function

2.2 Intersect Y-axis at any Point

Now suppose we need to fit a linear function to given data set which intersect the Y-axis at any point. Then the function is in the form

$$f(x_i) = mx_i + b$$

where, m is the slope of the fitted function and b is the Y-intersection. Then the sum of the squares of the distance between actual and predicted data S_1 is given by,

$$S_1 = \sum_{i=1}^{n}(y_i - f(x_i))^2 = \sum_{i=1}^{n}(y_i - (mx_i + b))^2$$

Thus to find m and b, we need to consider the first derivative of S_1 with respect to both m and b.