Linear Regression

1

Regression

2

We frequently measure two or more variables on the same individual (case, object, etc). We do this to explore the nature of the relationship among these variables.

Examples

The tensile strength of wrapping paper versus the percent of hardwood in the pulp batch

Maintenance cost of tractors versus the age of the tractor

The repair time for a computer versus the number of components which have to be changed

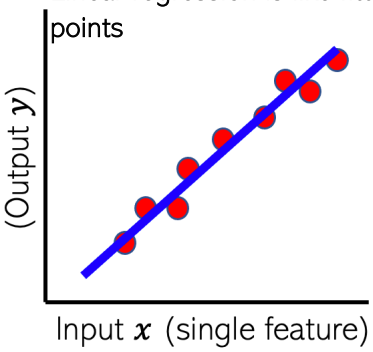
Cholesterol level versus age

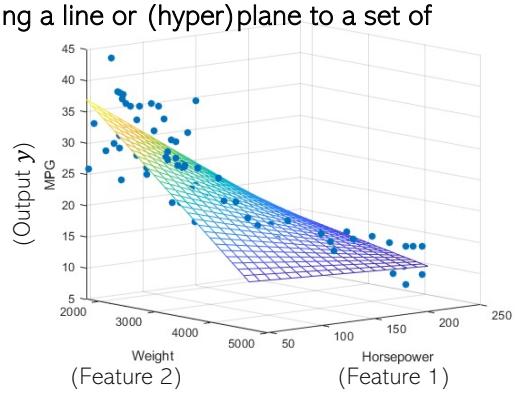
2

Linear Regression

3

Linear regression is like fitting a line or (hyper)plane to a set of points



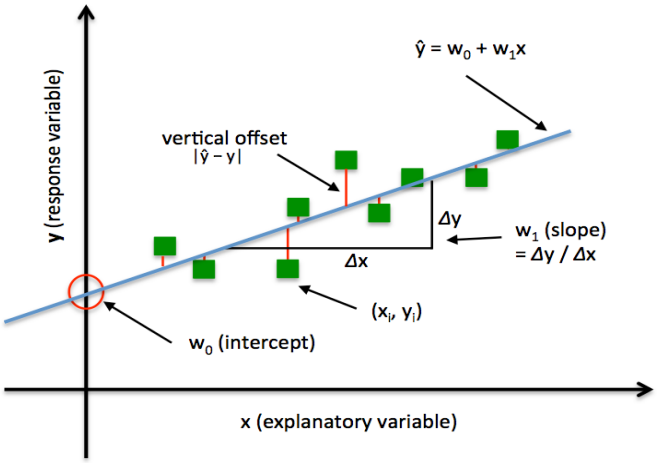


The line/plane must also predict outputs for the test inputs well

3

Simple Linear Regression

4



4

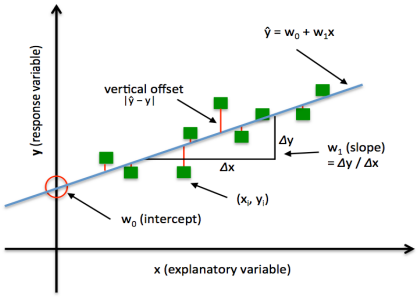
Solving the Simple Linear Regression Problem

5

- We basically want to find {w0, w1} that minimize deviations from the predictor line (model)

$$\arg\min_{w_0,w_1} \sum_{i}^{n} (y_i - w_0 - w_1 x_i)^2$$

- How do we do it?
  - Iterate over all possible w values along the two dimensions?
  - No, we can do this by solving the optimization problem



5

Simple Linear Regression: Example

6

Problem: A company that repairs small computers needs to develop a better way of providing customers typical repair time estimates. To begin this process, they compiled data on repair times (in minutes) and the number of components needing repair or replacement from the previous week. The data, sorted by number of components are as follows:

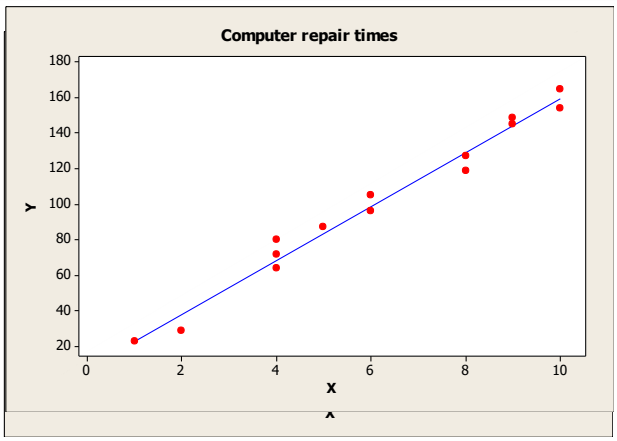
i	Number of components x_i	Repair time y_i
1	1	23
2	2	29
3	4	64
4	4	72
5	4	80
6	5	87
7	6	96
8	6	105
9	8	127
10	8	119
11	9	145
12	9	149
13	10	165
14	10	154

Paired Observations (x_i, y_i)

6

Simple Linear Regression: Example

7



$$\hat{\beta}_1 = \frac{S_{XY}}{S_{XX}}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

7

Thank You!

8