FROM LINEAR ALGEBRA TO MACHINE LEARNING

Omar Gutiérrez

September 3, 2018

@trinogz

Overview

Motivation

- **Linear algebra** is important to understand machine learning.
- As well as **calculus**, **probability theory**, and **statistics**.
- It is rewarding to take the **hard path** to learn machine learning (IMHO).

Learning from errors

```
30  # compute distances using self-created function
31  distances2 <- matrix(nrow=size, ncol=size)
32  for (p in 1:size) {
33    for (q in 1:size) {
34      row_p = iris2[p,]
35      row_q = iris2[q,]
36      distances2[p, q] <- euclidean_distance(row_p, row_q)
37    }
38  }
```

Learning from errors

```
14  # function for calculating Euclidean Distance
15  euclidean_distance <- function(p, q) {
16      ed = 0
17      for (i in 1:4) {
18          ed <- ed + (p[,i] - q[,i]) ^ 2
19      }
20      ed <- sqrt(ed)
21      return(ed)
22  }
```

A vector is a collection of numbers

$$\vec{a} = \boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Length of a vector

$$|\boldsymbol{a}| = \sqrt{\sum_{i=1}^{n} a_i^2}$$

Distance beetwen vectors

$$d(a, b) = ||a - b||$$
$$= \sqrt{\sum_{i=0}^{n} (a_i - b_i)^2}$$

Dot product

$$
\begin{aligned}
a \cdot b &= \sum_{i=0}^{n} a_i b_i \\
&= a_0 b_0 + a_1 b_1 + \ldots + a_n b_n
\end{aligned}
$$

So,

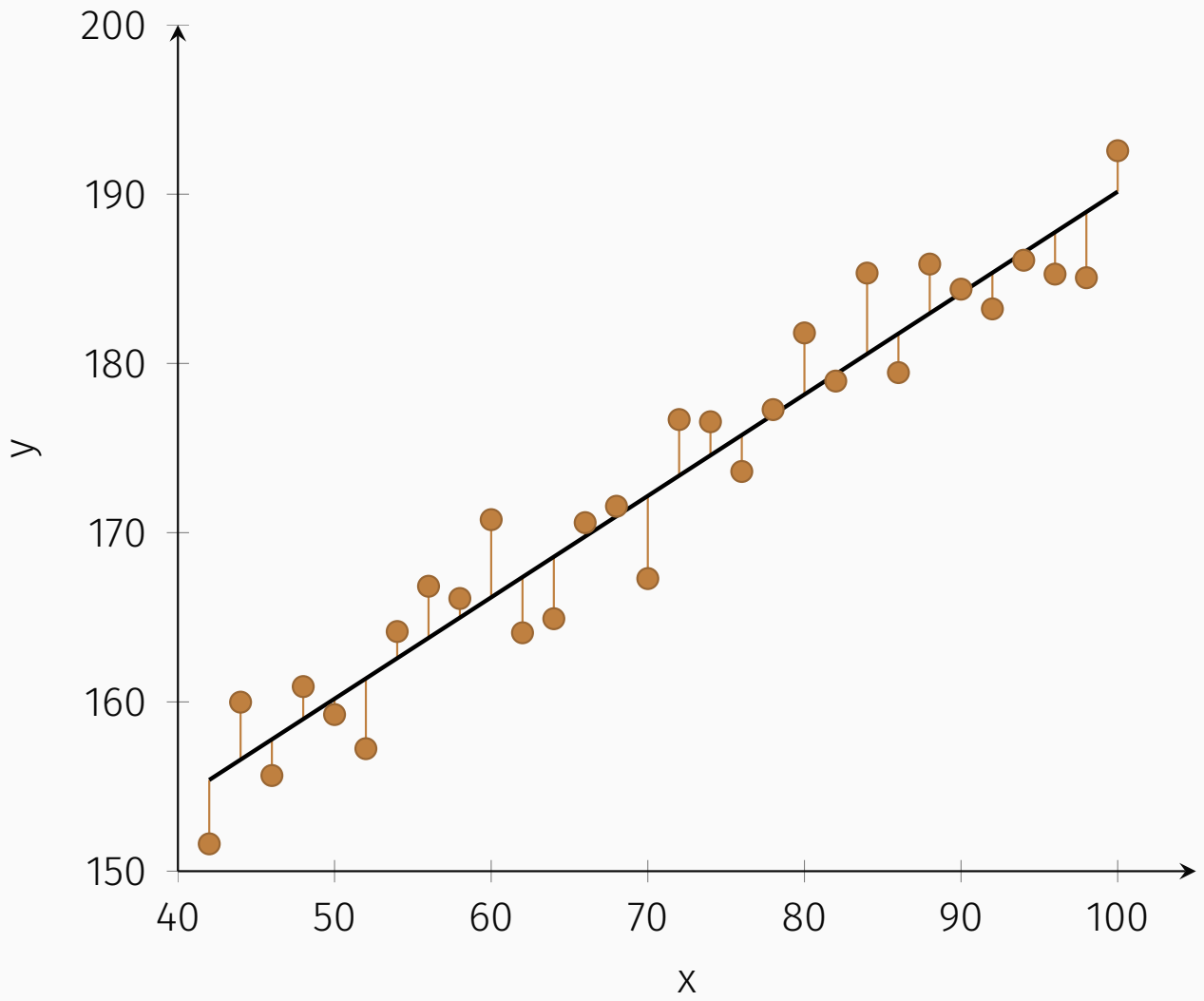
$$
\begin{aligned}
a \cdot a &= a_0 a_0 + a_1 a_1 + \ldots + a_n a_n \\
&= a_0^2 + a_1^2 + \ldots + a_n^2 \\
&= |\boldsymbol{a}|^2
\end{aligned}
$$

WINTER IS COMING

The AI winter is coming

- Is really coming? No.
- However, we already had an AI winter.
- The research on neural nets was stopped for many years, after Minsky and Papert proved that a single layer perceptron was not able to deal with the exclusive-or problem.

Linear Regression



Linear Regression

- We want to calculate the intercept a and the slope b.

$$\arg\min_{a,b} \sum_i (y_i - (ax_i + b))^2 = \arg\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||^2$$

- The solution to this optimization problem is:

$$\boldsymbol{w}^* = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}\,.$$

References

- **Mathematics for Machine Learning: Linear Algebra** offered by Coursera.
- **The Math of Intelligence** offered by Siraj Raval.

Thank you.
Questions?
Comments?