linear-regression.pdf

quinta-feira, 6 de julho de 2023 16:40

linear-regressio...

Deep Learning

Linear Regression

Tiago Vieira

Institute of Computing
Universidade Federal de Alagoas

February 9, 2023

Summary

Introduction

Let's start with a very simple example: Linear Regression.

- ML algorithm: Performance (P) improvement with respect to a given Task (T) via Experience (E).

Introduction

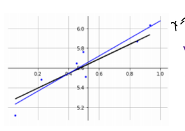


Figure: Example of linear regression. The black solid line represents $f(x)$ corresponding to the real phenomenon generating the data. Blue dots represent the samples obtained after the insertion of the noise. The blue solid line represents the model obtained by linear regression.

$y = f(x) + \varepsilon$ → (erro) função que descreve

Introduction

Linear regression:
Input:

$$\mathbf{x} \in R^n. \quad (1)$$

Output is a linear function of the input:

$$y \in R. \quad (2)$$

That is:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathcal{R}^n \mapsto y \in \mathcal{R} \quad (3)$$

- Parameters are values that control the behavior of the system.
- w_i are coefficients that are multiplied by feature x_i before adding the contributions of all attributes.
- $[w_i]$ is a set of *weights* determining how much each attribute contributes to the prediction $\hat{y}$.
- $w_i > 0 : \uparrow x_i \quad \uparrow \hat{y}$
- $w_i < 0 : \uparrow x_i \quad \downarrow \hat{y}$
- $w_i = 0 : \quad x_i$ has no influence on $\hat{y}$
- $w_i \gg 0 : \quad x_i$ has a large influence on $\hat{y}$

Problem statement:

- Task T: Predict y from a given input $\mathbf{x}$ by computing $\hat{y} = \mathbf{w}^T\mathbf{x}$. → Transposta dos pesos

Suppose:

- Design matrix containing m examples:

$$X^{test} = \left[- - \mathbf{x}^{(i)} - -\right]_{i=1}^{m} \quad (4)$$

- Vector of regression targets:

$$\mathbf{y}^{(test)}. \quad (5)$$

Performance (P) is given by the Mean Squared Error (MSE) computed on the model's predictions on test examples $\mathbf{y}^{(test)}$:

$$MSE_{test} = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{y}^{(test)} - y^{(test)}\right)_i^2 \quad (6)$$

- One can see that the error is zero for $y = \hat{y}$.

Also:

$$MSE_{test} = \frac{1}{m}\|\hat{y}^{(test)} - y\|_i^2 \quad (7)$$

so that the error increases whenever the euclidean distance between predictions and ground-truth targets increases.

ML algorithm's goal:

- Improve weights $\mathbf{w}$ in order to reduce the error $MSE_{(test)}$.
- But the algorithm can "learn" (gain experience) through observations in training dataset $(\mathbf{x}^{(i)}, y^{(i)})_{i=1}^{m}$.

How do we do that?

- Minimizing the mean squared error using the gradient and making it equal to zero:

$$\nabla_{\mathbf{w}} MSE_{(train)} = 0 \quad (8)$$

$$\nabla_{\mathbf{w}} \frac{1}{m}\|\hat{y}^{(train)} - y^{(train)}\|_2^2 = 0$$

$$\frac{1}{m}\nabla_{\mathbf{w}}\|X^{(train)}_{(m,n)}\mathbf{w}_{(n,1)} - \mathbf{y}^{(train)}_{(m,1)}\|_2^2$$

Solving for $\mathbf{w}$ gives:

$$\boxed{\mathbf{w} = \left(X^TX\right)^{-1}X^T\mathbf{y}} \quad (9)$$

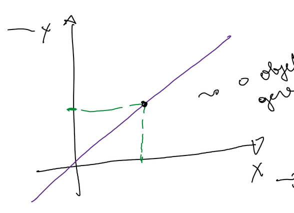
Inversa

- Known as the *normal equations*.
- Simple learning algorithm.

Thank you!
tvieira@ic.ufal.br

* Vídeos

* Regressão Linear



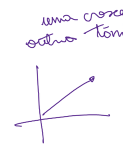
Variável dependente → Resposta

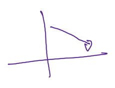
→ o objetivo é encontrar a função que gere uma linha de modo que conseguimos predizer outros pontos

X → variável independente → Explanatória

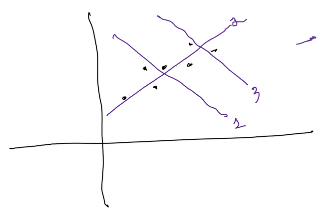
→ Buscar entender como as variáveis dependentes e independentes se relacionam. Se de forma positiva (igual) ou negativa (dispar)

uma cresce e a outra também





→ A base são diversas observações, ou seja, pontos no gráfico e buscar a linha que encaixe



→ Qual linha melhor representa?

A/2

Depois você precisa minimizar o erro, ou seja, distância entre o estimado e o verdadeiro

$\hat{Y}$ = estimated value

$$\boxed{\hat{y} = b_0 + b_1 x}$$

valor de x (independente)

curvatura da reta

valor de y

$$\boxed{y = mx + b}$$

→ Usa-se metodos dos minimos quadrados para ajustar a linha aos dados

→ Depois calcula R^2 e um p-value para R^2

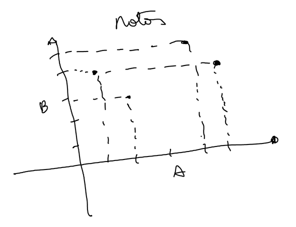
→ A ideia é plotar uma reta e ir rotacionando ela de acordo com a distância do estimado para o real até que chegue num valor bom

* Correlação Linear

Correlação entre duas variáveis diferentes

- Diagrama de dispersão

Grafico de duas variáveis para analisar se ambas tem uma correlação



Se não estiverem muito dispersos, então provavelmente há relação e tem como traçar uma regressão linear

→ Usa-se o coeficiente de correlação

$rxy

- $R = 0$ → Sem correlação
- $R < 50\%$ → Fraca correlação
- $R > 50\%$ → Forte correlação
- $R = 100\%$ → Perfeita correlação

→ Reta de Regressão

$$\hat{Y} = aX + b$$

$$\boxed{b = \frac{\sum_1 Y}{n} - a \cdot \frac{\sum_1 X}{n}}$$

$$\boxed{a = \frac{n \cdot \sum_1 X \cdot Y - \sum_1 X \cdot \sum_1 Y}{n \cdot \sum_1 X^2 - \left(\sum_1 X\right)^2}}$$

Example: $\hat{Y} = 8{,}34\,X + 19{.}37$

previsão

o input que damos

$X = 18$

$\hat{Y} = 169$