Matrix

Tuan Nguyen

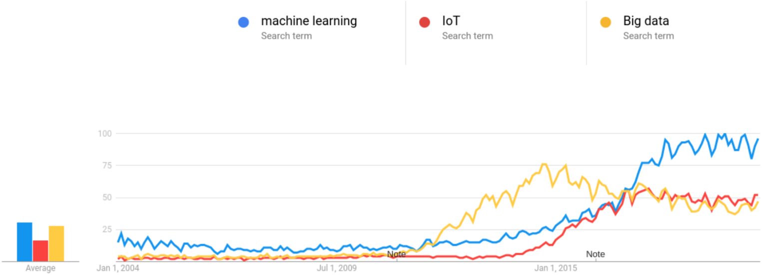
Ngày 17 tháng 12 năm 2020

Overview

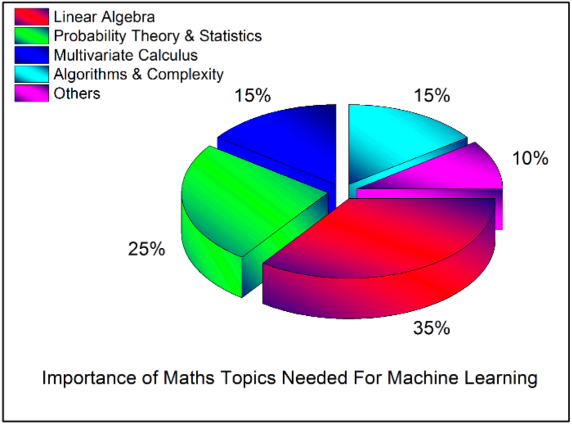
Introduction

- Lecturer at Faculty of Mathematical Economics, National Economics University.
- Founder AI For Everyone (AI4E).
- Publish Deep Learning cơ bản ebook.
- Deep Learning, Python teacher at VIASM.

Machine Learning trends

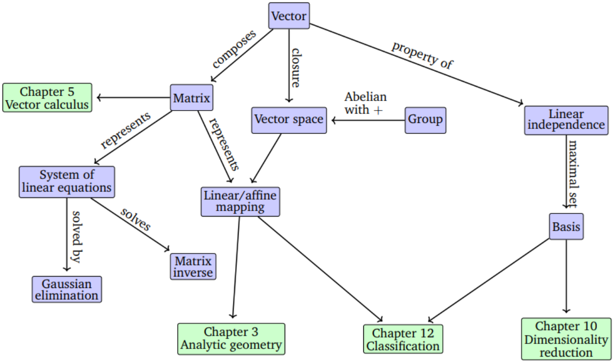


Hình 1: Machine Learning trend



Hình 2: Maths for Machine Learning

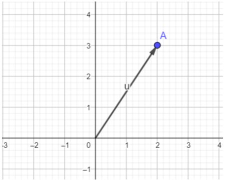
Linear algebra is the study of vectors and certain algebra rules to manipulate vectors



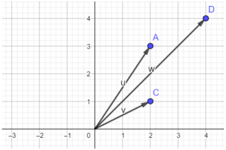
Hình 3: Linear algebra

Vector

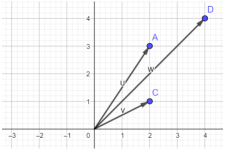
Vectors are special objects that can be added together and multiplied by scalars to produce another object of the same kind.



(a) Vector



(b) Add vector



(c) Scale vector

Hình 4: Vector properties

Vector representation

Geometric vectors are only for 2D and 3D. How if the vector has more dimensions.

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} \in \mathbb{R}^n, v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3$$

Add

$$\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \dots \\ a_n + b_n \end{bmatrix}$$

Multiply

$$\alpha * \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \alpha * a_1 \\ \alpha * a_2 \\ \dots \\ \alpha * a_n \end{bmatrix}$$

Matrix

Matrix is a grid of numbers with size: number_of_row * number_of_column

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots. & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

$$= concat(\begin{bmatrix} a_{11} \\ a_{12} \\ \dots \\ a_{1n} \end{bmatrix}, \begin{bmatrix} a_{21} \\ a_{22} \\ \dots \\ a_{2n} \end{bmatrix}, \dots, \begin{bmatrix} a_{m1} \\ a_{m2} \\ \dots \\ a_{mn} \end{bmatrix})$$

Matrix addition

$A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times n}$

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \dots & \dots. & \dots & \dots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

Hadamard product

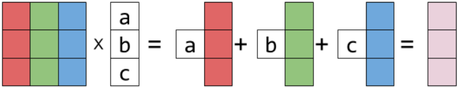
$A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times n}$

$$A \otimes B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2n}b_{2n} \\ \dots & \dots. & \dots & \dots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \dots & a_{mn}b_{mn} \end{bmatrix}$$

Matrix multiplication

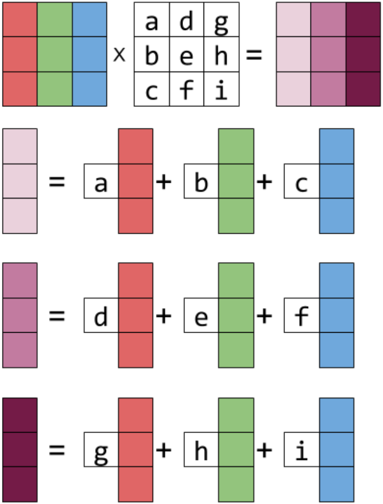
$$A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n*k}, C = AB, C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \times 2 + 2 \times 3 \\ 2 \times 2 + 0 \times 3 \\ 1 \times 2 + 1 \times 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix}$$



Hình 5: Different view

$$\begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix}$$



Hình 6: Matrix multiplication

Identity matrix

Identity matrix is a square matrix, with the value: $I_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Matrix properties

- $\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times q} : (AB)C = A(BC)$
- $\forall A, B \in \mathbb{R}^{m \times n}, C, D \in \mathbb{R}^{n \times p}$
  - $(A + B)C = AC + BC$
  - $A(C + D) = AC + AD$
- $\forall A \in \mathbb{R}^{m \times n} : I_m A = A I_n = A$, note that: $I_m \neq I_n$

Transpose

$A \in \mathbb{R}^{m \times n} => B = A^T \in \mathbb{R}^{n \times m}, a_{ij} = b_{ji}$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots. & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} => B = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots. & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

A matrix $A \in \mathbb{R}^{n \times n}$ is **symmetric** if $A^T = A$

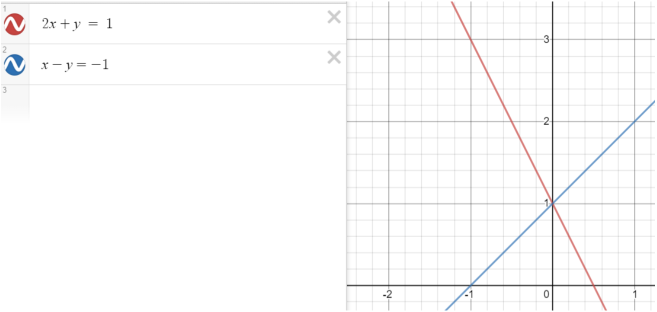
Inverse

Square matrix $A \in \mathbb{R}^{n \times n}$, matrix $B \in \mathbb{R}^{n \times n}$ have the property that $AB = BA = I_n$, B is called the inverse of A and donated A^{-1}.

Not every matrix A possesses an inverse A^{-1}, if the inverse does exist, A is called regular/invertible/nonsingular matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} => A^{-1} =?$$

$$\begin{cases} 2x + y = 1 \\ x - y = -1 \end{cases} \Leftrightarrow \begin{cases} 3x = 0 \\ x - y = -1 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 1 \end{cases}$$



Hình 7: Intersection of lines

Linear equation (cont.)

$$\begin{cases} 2x + y = 1 \\ x - y = -1 \end{cases} \Leftrightarrow \begin{cases} 2x_1 + x_2 = 1 \\ x_1 - x_2 = -1 \end{cases}$$

$$\Leftrightarrow \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Leftrightarrow x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

General formula:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \Leftrightarrow Ax = b$$

$$\Leftrightarrow \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots. & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{bmatrix}$$

How to solve:

$$\begin{bmatrix}1 & 0 & 8 & -4\\ 0 & 1 & 2 & 12\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} = \begin{bmatrix}42\\ 8\end{bmatrix}$$

$$\Leftrightarrow x_1\begin{bmatrix}1\\ 0\end{bmatrix} + x_2\begin{bmatrix}0\\ 1\end{bmatrix} + x_3\begin{bmatrix}8\\ 2\end{bmatrix} + x_4\begin{bmatrix}-4\\ 12\end{bmatrix} = \begin{bmatrix}42\\ 8\end{bmatrix} \Leftrightarrow \sum_{i=1}^{4} x_i c_i = b$$

We can see that $x = [42, 8, 0, 0]^T$ is a solution, it is called particular solution or special solution.

$$\begin{bmatrix}1 & 0 & 8 & -4\\ 0 & 1 & 2 & 12\end{bmatrix}\left(\begin{bmatrix}42\\ 8\\ 0\\ 0\end{bmatrix}\right) = \begin{bmatrix}42\\ 8\end{bmatrix}$$

Also note that:

$$\begin{bmatrix}8\\2\end{bmatrix} = 8\begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\end{bmatrix} \text{ or } c_3 = 8c_1 + 2c_2$$

$$\Leftrightarrow \begin{bmatrix}1 & 0 & 8 & -4\\0 & 1 & 2 & 12\end{bmatrix}\left(\alpha_1\begin{bmatrix}8\\2\\-1\\0\end{bmatrix}\right) = \alpha_1(8c_1 + 2c_2 - c_3) = 0$$

So that $\begin{bmatrix}1 & 0 & 8 & -4\\0 & 1 & 2 & 12\end{bmatrix}\left(\begin{bmatrix}42\\8\\0\\0\end{bmatrix} + \alpha_1\begin{bmatrix}8\\2\\-1\\0\end{bmatrix}\right) = \begin{bmatrix}42\\8\\0\\0\end{bmatrix}$

Finally: $x = \begin{bmatrix}42\\8\\0\\0\end{bmatrix} + \alpha_1\begin{bmatrix}8\\2\\-1\\0\end{bmatrix} + \alpha_2\begin{bmatrix}-4\\12\\0\\-1\end{bmatrix}, \alpha_1, \alpha_2 \in \mathbb{R}$ is a general solution.

Linear equation (cont.)

How to solve system of equations:

- Find a particular solution to $Ax = b$.
- Find all solutions to $Ax = 0$.
- Combine the solutions from steps 1. and 2. to the general solution.

Elementary transformations

Elementary transformations keep the solution set the same, but that transform the equation system into a simpler form:

- Exchange of two equations (rows in the matrix representing the system of equations)
- Multiplication of an equation (row) with a constant
- Addition of two equations (rows)

$$\begin{cases} -2x_1 + 4x_2 - 2x_3 - x_4 + 4x_5 = -3 \\ 4x_1 - 8x_2 + 3x_3 - 3x_4 + x_5 = 2 \\ x_1 - 2x_2 + x_3 - x_4 + x_5 = 0 \\ x_1 - 2x_2 - 3x_4 + 4x_5 = a \end{cases} \quad \text{=> augmented matrix}$$

$$\left(\begin{array}{ccccc|c} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right)$$

The leading coefficient of a row pivot (first nonzero number from the left) is called the pivot

Elementary transformations (cont.)

It means
$$\begin{cases} x_1 - 2x_2 + x_3 - x_4 + x_5 = 0 \\ x_3 - x_4 + 3x_5 = -2 \\ x_4 - 2x_5 = 1 \\ 0 = a + 1 \end{cases}$$

Only a = -1 this system can be solved.

=> What is general solution?

Row-Echelon Form

A matrix is in row-echelon form if:

- All rows that contain only zeros are at the bottom of the matrix
- Looking at nonzero rows only, the first nonzero number from the left pivot (also called the pivot or the leading coefficient) is always strictly to the leading coefficient right of the pivot of the row above it.

Remark: The variables corresponding to the pivots in the row-echelon form are called basic variables and the other basic variable variables are free variables.

$$\left(\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right)$$

x_1, x_3, x_4 are basic variables, x_2, x_5 are free variables

Row-Echelon Form (cont.)

An equation system is in reduced reduced row-echelon form if:

- It is in row-echelon form.
- Every pivot is 1.
- The pivot is the only nonzero entry in its column.

Gaussian elimination is an algorithm that elimination performs elementary transformations to bring a system of linear equations into reduced row-echelon form.

$$A = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix} \Rightarrow \text{What is the solution of } Ax = 0?$$

We can use Gaussian elimination to calculate the inverse of matrix.