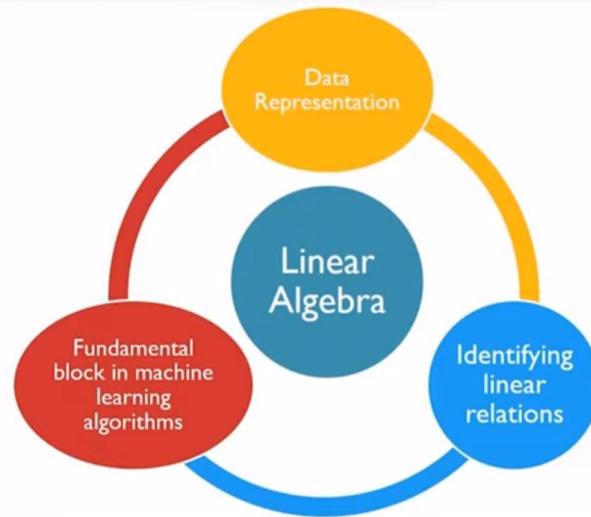


Overview



Linear Algebra

2

Matrix theory and linear algebra

Matrix Theory and Linear Algebra

- Matrices can be used to represent samples with multiple attributes in a compact form
- Matrices can also be used to represent linear equations in a compact and simple fashion
- Linear algebra provides tools to understand and manipulate matrices to derive useful knowledge from data



Linear Algebra

LINEAR ALGEBRA AND MATRICES



Data science for Engineers

Matrices for data science: Data representation

- Usually matrices are used to store and represent the data on machines
- Matrix is a very natural approach for organizing data
- In general, data is organized in the following fashion
 - Rows represent samples
 - Columns represent the values of the variables (or attributes)
 - It is also possible to use rows for variables and columns for samples
 - However, we will stick to rows as samples and columns as variables in all of the material that will be presented



Linear Algebra

Data representation: Examples

- A real life example
 - Consider a reactor which needs to be controlled using multiple attributes from various sensors like Pressure (Pa), Temperature (K), Density (gm/m^3) etc.
 - Independently, the sensors have generated 1,000 data points
 - This complete set of information is contained in

$$\begin{array}{c}
 P \quad T \quad \rho \\
 \begin{matrix} 1 \\ \vdots \\ 1000 \end{matrix} \begin{bmatrix} 300 & 300 & 1000 \\ \vdots & \vdots & \vdots \\ 500 & 1000 & 5000 \end{bmatrix}
 \end{array}$$



Data representation: Examples

- Example 2:

$$X = [1, 2, 3]^T$$

$$Y = [2, 4, 6]^T$$

- X and Y are vectors pertaining to some attributes
- We define the A matrix using a column bind of X and Y thus representing data in a matrix format (the code for the same is attached)

R Code
`x=c(1,2,3)`
`y=c(2,4,6)`
`A=cbind(x,y)`
`print(A)`

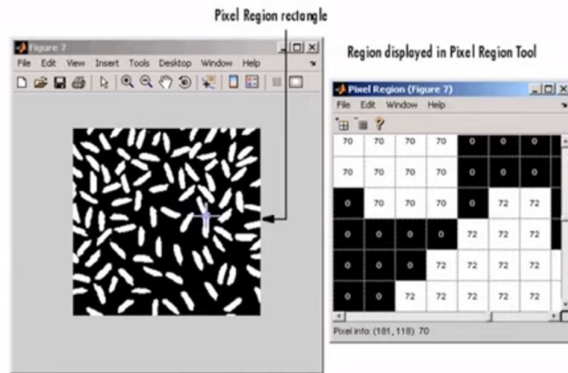
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$$

Output
`> print(A)`
 x y
`[1,] 1 2`
`[2,] 2 4`
`[3,] 3 6`



Data representation: Examples

- The simplicity in representation will become apparent when the image below is considered



Data representation: Examples

- Storing
 - The image is stored in the machine as a large matrix of pixel values across the image.
 - Thus, storing the pixel value matrix is equivalent to storing the image for the machine
- Identification
 - Several machine learning algorithms are deployed in order to “teach” the machine how to identify a particular image.
 - Linear algebra and matrix operations are at the heart of these machine learning algorithms.

Data as matrix: Summary



IDENTIFICATION OF INDEPENDENT ATTRIBUTES



Further analysis

- Now that we can represent the data into a matrix format, we ask the following questions
 - Are all the attributes in the data matrix relevant/ important?
 - Is there any method which can identify if some attributes are related to the other attributes?
 - If yes, how do we identify the linear relationship?
 - Can we use this to reduce the size of the data matrix?



Identification of independent attributes: Example

- Consider the ideal reactor example with multiple (say, 4) attributes like Pressure, Temperature, Density, Viscosity, etc. with 500 samples.
- Thus we have a 500×4 matrix such that
$$A = [P \ T \ D \ \eta]$$
- P, T, D and η are vectors of 500 samples from the pressure, temperature, density and viscosity sensors.
- How does one identify the number of independent attributes?

Identification of independent attributes: Example

- Domain knowledge
$$D \sim f(P, T)$$
- Thus, in some sense **D** is a function of **P** and **T**
- Implying that at least one attribute is dependent on the others
- This variable can be calculated as a linear combination of the other variables
- The physics of the problem helps us identify the relationship in the data matrix
- We now ask if the data itself will help us identify these relationships

Number of independent attributes: Rank of a matrix

- Let us assume that we have many more samples than attributes for now
- Is there any approach which can be used to identify the number of linear relationships between the attributes purely using data?
- This is addressed by the concept of the **rank** of the matrix.
- **Rank** of a matrix refers to the number of linearly independent rows or columns of the matrix
- The rank of a matrix can be found using the rank command: rank(A)

Rank of a matrix: Example 2

- Consider another example

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ 3 & 6 & 0 \end{bmatrix}$$

- We observe that
 - (Col. 2) = 2 x (Col. 1)
 - (Col. 3) is independent
- Thus, the rank of this matrix is 2

R Code

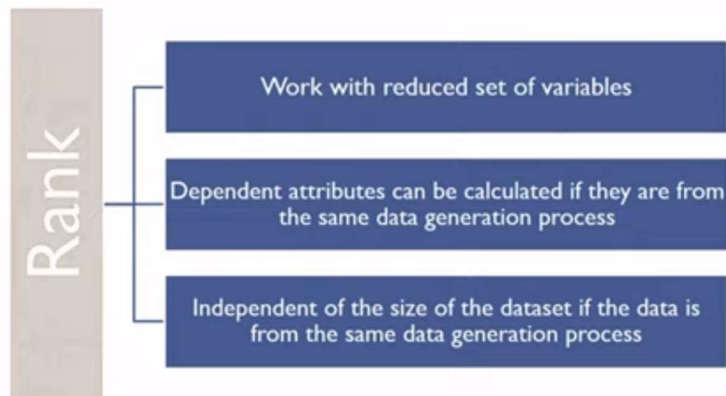
```
A=matrix(c(1,2,3,2,4,6,1,0,0),ncol=3,byrow=F)
library(pracma)
Rank(A)
```

Output

```
> Rank(A)
[1] 2
```



Rank: Advantages and summary





IDENTIFICATION OF LINEAR RELATIONSHIPS AMONG ATTRIBUTES



16:48 / 28:39



YouTube



Data science for Engineers

Linear relationships among attributes

- Now that we have identified the number of linearly independent attributes:
 - How does one identify those linear relations among the attributes?
- Such questions are addressed by the linear algebraic concepts of null space and nullity



Null space for data science

- The null space of a matrix \mathbf{A} consists of all vectors $\boldsymbol{\beta}$ such that $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ and $\boldsymbol{\beta} \neq \mathbf{0}$
- Nullity of a matrix is the number of vectors in the null space of the given matrix
- The size of the null space of a matrix provides us with the number of linear relations among the attributes
- And the null space vectors $\boldsymbol{\beta}$ are useful to identify these linear relationships



Null space : general description

- Let us suppose
- $\mathbf{A} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$ is a data matrix and there is one vector in the null space of \mathbf{A} , i.e, $\boldsymbol{\beta} = [\beta_1 \dots \beta_n]^T$, then as per the definition, $\boldsymbol{\beta}$ satisfies all the equations given below
- $x_{11}\beta_1 + x_{12}\beta_2 + \cdots x_{1n}\beta_n = 0$
- \vdots
- $x_{m1}\beta_1 + x_{m2}\beta_2 + \cdots x_{mn}\beta_n = 0$

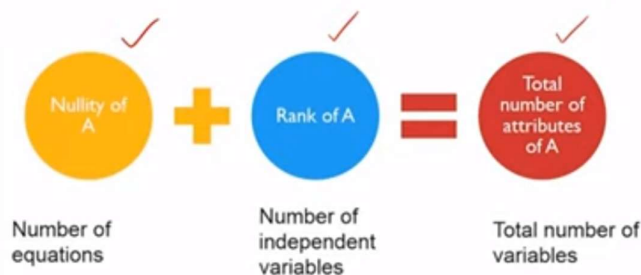


Null space: The idea

- Notice that if $A\beta = \mathbf{0}$, every row of A when multiplied by β goes to zero
- This implies that variable values in each sample (represented by a row) behave the same
- This helps in identifying the linear relationships in the attributes
- Every null space vector corresponds to one linear relationship
- This idea is demonstrated further using examples

Rank nullity theorem

- Consider the data matrix A with the null space and nullity as defined before
- The rank- nullity theorem helps us to relate the nullity of the data matrix to the rank and the number of attributes in the data
- According to the rank-nullity theorem



Summary till now

Data Matrix

- The available data is expressed in the form of a data matrix
- This data matrix is further used to do the necessary operations

Null Space

- Defined as a collection of vectors satisfying $A\beta = 0$
- Helps in identifying the linear relationships between the attributes directly

Nullity

- Nullity is the size of the null space of the data matrix
- Useful to identify the number of linear relationships in the attributes
- Rank- Nullity theorem

Null space: An Example

- Consider the matrix A with attributes $\{x_1, x_2\}$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

Number of columns in A = 2

Rank of A = 2

Thus, nullity = 0

- This implies that the null space of the matrix A does not contain any vectors
- Thus we can claim that all the attributes are linearly independent

Console output

```
> A
     [,1] [,2]
[1,]  1   2
[2,]  3   4
[3,]  5   6
```

Console output

```
> print(ncol(A))
[1] 2
> print(rank(A))
[1] 2
> print(nullity(A))
[1] 0
```

R Code

```
A=matrix(c(1,3,5,2,4,6),ncol=2, byrow=F)
columns=ncol(A)
library(pracma)
rank=Rank(A)
nullity=columns-rank
```

Null space: Further Example

$$A\beta = 0$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- Thus we obtain,

$$b_1 + 2b_2 = 0$$

$$b_3 = 0$$

- The null vector is $B = [b_1 \ b_2 \ b_3]^T = [-2b_2 \ b_2 \ 0]^T = k[-2 \ 1 \ 0]^T$
- We see that we obtain a direct linear relationship between the attributes of A using null space and rank-nullity theorem
- The same concept can be extended for bigger data set



Overall summary

