

LINEAR ALGEBRA IN MACHINE LEARNING

CONTENTS OF REPORT

1. ABSTRACT
2. INTRODUCTION
3. WHY YOU SHOULD LEARN LA BEFORE MACHINE LEARNING
 - 3.1. Linear Algebra is the elementary unit for Machine learning
 - 3.2. The Machine Learning intuition
 - 3.3. Linear Algebra Helps Build Better ML Algorithms from Scratch
 - 3.4. Linear Algebra Will Improve One's Grasp on Statistics
4. WHY DO WE USE LINEAR ALGEBRA IN MACHINE LEARNING?
5. VECTORS
 - 5.1. DEFINITION
 - 5.2. PURPOSE
6. MATRICES
 - 6.1. DEFINITION
 - 6.2. PURPOSE
7. TENSORS
 - 7.1. DEFINITION
 - 7.2. PURPOSE
8. SOME MACHINE LEARNING ALGORITHMS
 - 8.1. Linear regression algorithm
 - 8.2. What is a Support Vector and what is SVM?
 - 8.3. Support Vector Machine Algorithm
 - 8.4. How to create a machine based on support vectors?
9. APPLICATIONS OF LINEAR ALGEBRA IN ML
 - 9.1. Linear Algebra notation
 - 9.2. Linear Algebra arithmetic
 - 9.3. Linear Algebra for statistics
 - 9.4. Matrix factorisation

9.5. Linear least squares

10. EXAMPLES OF LINEAR ALGEBRA IN MACHINE LEARNING

- 10.1. Datasets and information files
- 10.2. Images and Photos
- 10.3. One Hot Encoding
- 10.4. Linear Regression
- 10.5. Regularization
- 10.6. Principle element evaluation
 - 10.6.1. Single cost decomposition
 - 10.6.2. Cryptographic Analysis

11. CONCLUSION

12. REFERENCES

ABSTRACT

Linear Algebra is a part of mathematics that includes vectors, matrices, and linear transformation which is applicable in different fields such as physics, engineering, and computer science and in which one of them is machine learning and linear algebra plays a key role in machine learning algorithms, data analysis and classifications. And machine learning is used to enable computer programs to decide without being explicitly programmed and its application is mostly close to linear algebra as graphing, predicting, and regression.

In this report, we will look at what linear algebra and machine learning are and as the category of machine learning is closely related to statistics, first we will investigate a brief explanation of the statistical concepts. Then we will see the application of linear algebra, especially matrices and vectors in machine learning algorithms such as linear regression and support vector machines.

INTRODUCTION

Linear Algebra is a continuous form of mathematics and is applied throughout science and engineering because it allows you to model natural phenomena and compute them effectively. Linear algebra is quite effective in regard to understanding Machine learning technologies better. It allows us to represent all the conditions based on the flexibility of vector space while allowing individuals to map data to specific dimensions precisely within the vector space.

Machine learning is about discovering structures and patterns that exist in a set of things. This is done using the language of maths, and we need to translate each thing into numbers. For example, a single number can't sum up all the relevant facts about a thing very well; normally 'interesting' things are more complex. Instead, we take a number of different measurements on each thing and collect them into a vector of numbers that stands in for the thing itself.

We can use linear algebra to build better supervised as well as unsupervised machine learning algorithms. Logistic regression, linear regression, decision trees, and support vector machines (SVM) are a few supervised learning algorithms that we can create from scratch with the help of linear algebra.

On the other hand, you can also use it for unsupervised Machine Learning algorithms, including the single value decomposition (SVD), clustering, and components analysis. Linear algebra will help us develop a more in-depth understanding of the machine learning project we are working on, and thus will give us the flexibility to customize different parameters.

Like every cake is based in flour, flour is an important ingredient for a bakery, Linear algebra is that flour for machine learning. Similarly every machine learning model is dependent upon linear algebra. It is not the only ingredient required, for sure. Apart from flour, cakes need sugar, eggs, and butter to taste better, similarly Machine learning models need vector

calculus, probability, and optimization. Applied machine learning, like a bakery, is mainly focused on combining these mathematical ingredients in smarter ways to create useful models.

We don't need to understand Linear Algebra before getting started with Machine Learning, but at some point, you may want to achieve a better understanding of how the different Machine Learning algorithms actually work at the backend. This will help you in making better decisions at the time of developing a Machine Learning system. So, if you really have a desire to be a professional in the field of machine learning, you must master the parts of Linear Algebra as it is important. In Linear Algebra, data is represented in form of linear equations, which are presented in the form of vectors and matrices, so in this report we will have a look at the structure of vectors and matrices and how they are useful in ML.

WHY YOU SHOULD LEARN LA BEFORE MACHINE LEARNING:

1.Linear Algebra is the elementary unit for Machine learning:

Calculus set the scene for LA when it comes to learning advanced math. Differential Calculus and Integral Calculus also set a base to apply limits, differentiation and integration techniques to vectors and multi-dimensional configurations such as tensors. This is called Matrix Calculus. Knowing this will help you to understand the areas such as linear functions and systems of linear equations. This will further help in advanced topics such as Linear Programming, Vectors in space and the Simplex method. In fact, all these concepts will be a cakewalk for you if some time is dedicated in learning LA and ML.

2.The Machine Learning intuition:

Learning Linear Algebra will uplift your intuition towards ML by providing more viewpoints. The vectors and matrices while come across while learning Linear Algebra will make your thinking broader and more idealistic. You may feel motivated while utilising more parameters to a ML component, using more matrix operations, visualising, and coming up with different plotting graphs, or even applying a unique and better code. It has numerous possibilities. For example, consider a ML project in healthcare. The cases used here can be diagnostics, preventive care, insurance, and patient health history. Even though the datasets are already available, the approach to build a ML model depends on perspectives such as data exploration, regression, training and test data and graphical depictions, along with the other features.

3. Linear Algebra Helps Build Better ML Algorithms from Scratch:

Linear Algebra will certainly help in developing popular ML algorithms essentially categorised as Supervised Learning and Unsupervised Learning from scratch. Logistic regression, decision trees, support vector machines and ensemble methods fall under supervised learning algorithms. On the other hand, single value decomposition (SVD), clustering and component analysis fall under unsupervised machine learning algorithms. LA will make a deeper understanding possible for the ML project which gives the flexibility to customise any parameters involved. This is helpful as it will lead to excellent usage of resources.

4. Linear Algebra Will Improve One's Grasp on Statistics:

For Machine Learning, statistics comprises a vital part to organise and comprehend data. LA acts as a necessity to have a deep understanding of statistical concepts. Notations, methods, and operations in Linear Algebra will assist in grasping advanced topics in statistics such as multivariate analysis. For example, let us consider a doctor who has patient data such as heart rate, blood pressure, height, weight, along with some other data. Now these can be used to form the multiple variables in the data set. Let us assume that increased weight leads to higher blood pressure. Thus, forming a linear relation between them — the increase in one variable leads to the increase in other.

WHY DO WE USE LINEAR ALGEBRA IN MACHINE LEARNING?

Linear Algebra mainly focuses on linear equations such as:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

Linear maps such as

$$(x_1, \dots, x_n) \rightarrow a_1x_1 + a_2x_2 + \dots + a_nx_n,$$

and their representation in vector spaces and matrices. It is the branch of mathematics that includes the study of linear transformations between vector spaces, such as scaling up or down a shape and even rotating it or moving it etc.

Machine Learning is all about creating an algorithm that can learn from data to make a prediction and Linear Algebra makes running of these algorithms possible on Large Data Sets that often comprise hundreds to millions of individual data items. Since Machine Learning majorly relies on Linear Algebra, so it is essential to understand what vectors and matrices are, what are the operations that you can perform with them and how they can be useful.

VECTORS

DEFINITION

A vector is a quantity that have a magnitude and a direction. For example, the velocity of a rocket approaching towards the space, is a 3-dimentional vector; its magnitude is the speed of the rocket and its direction is upward. A vector can be represented by an array of numbers called scalars. Each scalar corresponds to the magnitude of that vector in regards of its dimensions.

PURPOSE

Vectors have a lot of purposes in Machine Learning, most significantly to represent the observations and predictions. For an example let us assume we built a Machine Learning system to classify the videos into 3 different categories as – Good video, Spam video and Clickbait based on the we have stored about them. For each individual video, we have a vector representing what we know about that video, such as:

$$\mathbf{viddata} = \begin{pmatrix} 12.2 \\ 5.3 \\ 4.33 \\ 8.0 \end{pmatrix}$$

This vector could represent a video that is 12.2 minutes long, and only 5.3% viewers watch it for more than a minute, on average it gets 4.33 views daily, and it was reported 8 times as spam. As you can see, each axis may have its own different meaning.

Based on this vector our Machine Learning system may predict that there is an 82% probability that it is a spam video, 15% that it is clickbait, and 3% that it is a good video. This could be represented as the following vector:

$$\mathbf{probabilities} = \begin{pmatrix} 0.82 \\ 0.15 \\ 0.03 \end{pmatrix}$$

MATRIX

DEFINITION

A matrix is a two-dimensional array of scalars with one or multiple columns and one or multiple rows. The notation used for representing a matrix is often an uppercase letter, such as A, B, C etc., and the entries are referred by their two-dimensional subscript of row 'i' and column 'j', such as a_{ij} . For example:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

PURPOSE

The first place where you may encounter a matrix in machine learning is in model training data that is comprised of many rows and columns and is often represented using the capital letter “X”.

The geometric analogy that is used to help understand vectors and some of their operations does not hold upon matrices. Further, a vector itself can also be considered a matrix with a single column and multiple rows.

Often the dimensions of the matrix are denoted as m and n for the number of rows and the number of columns, respectively.

TENSORS

DEFINITION

A tensor is an apparent form of vectors and matrices and can be easily understood as a multidimensional array.

A vector can be considered as a one-dimensional or first order tensor and a matrix can be considered as a two-dimensional or second order tensor.

Tensor notation is somewhat similar to that of matrix notation with a capital letter representing a name of the tensor and lowercase letters with subscript integers representing their scalar values within the tensor.

$t_{111}, t_{121}, t_{131}$	$t_{112}, t_{122}, t_{132}$	$t_{113}, t_{123}, t_{133}$
$T = (t_{211}, t_{221}, t_{231}),$	$(t_{212}, t_{222}, t_{232}),$	$(t_{213}, t_{223}, t_{233})$
$t_{311}, t_{321}, t_{331}$	$t_{312}, t_{322}, t_{332}$	$t_{313}, t_{323}, t_{333}$

Many of the operations that can be performed with scalars, vectors, and matrices can easily be reformulated and thus can be used to be performed with tensors.

PURPOSE

As a tool, tensor algebra and tensors are widely used in the fields of physics and engineering. It is basically a term or a set of techniques known in machine learning in the training and operation of deep learning models can be described using the tensors.

In deep learning algorithms it is common to see a lot of discussion around tensors as the foundation or mainspring of data structure.

Even Google’s flagship machine learning library: “TensorFlow” has Tensor in its name.

Tensors are usually a type of data structure used in linear algebra, and like vectors and matrices, you can calculate arithmetic operations with them.

SOME MACHINE LEARNING ALGORITHMS:

1.LINEAR REGRESSION ALGORITHM:

Linear regression is one of the most used methods of data modelling that has a very simple mathematical basis. Whenever we can identify a linear relationship between two variables, we can use this kind of regression to predict the values of these variables based on the value of the other variable.

The purpose of the linear relationship is to see that by increasing one variable, the other variable increases (or decreases) And by decreasing it, the second variable mutually decreases (or increases), and this increase or decrease has a direct relation (simple coefficient) to the value of the first variable, which we call the independent variable. An easy way to figure out this relationship is to plot the value of one variable on the other in a graph, and if the resulting figure was like a straight line, we can conclude that the relationship between these two variables is a linear relationship. This graph is also called the distribution graph.

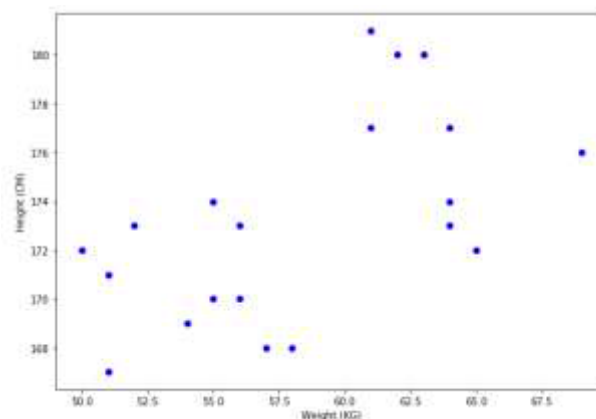


Chart:- height and weight values of the twenty people.

After finding the linear relationship between the two variables, we simply calculate the linear relationship formula between them and use it to predict new values of one variable based on the value of the other variable. That's exactly what linear regression does, finding the linear relationship between two variables by calculating the coefficients of a linear relationship and using it to predict values. The formula for a line is: $Y = w_1 * X + w_0$ where w_1 is the slope of the line and w_0 is the width of the source. In this relation, X is called the independent variable and Y is the dependent variable, although it is easy to obtain X in Y and replace the dependent and independent variable as desired.

What is a Support Vector and what is SVM?

In simple language, support vectors are a set of points in the n-dimensional space of the data that defines the boundaries of the categories, and the bounding and categorization of the data is done by them, and by moving one of them, the output of the categorization may change. For example, in the above figure, the vector (45,150) is a member of a support vector belonging to a woman. In 2D space, the support vector will form a line, in a 3D space form a plane and in a dimensional space form a cloud.

An SVM is a cluster or boundary that measures the best categorization and separation of data by benchmarking support vectors.

In SVM, only the data contained in the support vectors are based on machine learning and model building, and this algorithm is not sensitive to other data points and aims to find the best boundary between the data so that it is as far as possible from all categories (their support vectors).

2.Support Vector Machine Algorithm

SVM (support vector machine) is one of the most common algorithms and methods in the field of data classification. Suppose we have a dataset where 50% of the population is male and 50% is female. This dataset can be customers of an online store. By having a subset of this data that identifies the gender of individuals, we want to create rules that will help us determine the gender of the rest of the population with great accuracy. Detecting the gender of store visitors allows us to display separate ads for men and women and increase store profitability. We call this process classification in data science. To fully explain the problem, suppose two of the parameters we are going to determine are gender, height, and hair length. The graph of the height and length distribution of individuals is shown below, where the sexes of individuals with two square (male) and circle (female) symbols are shown separately.

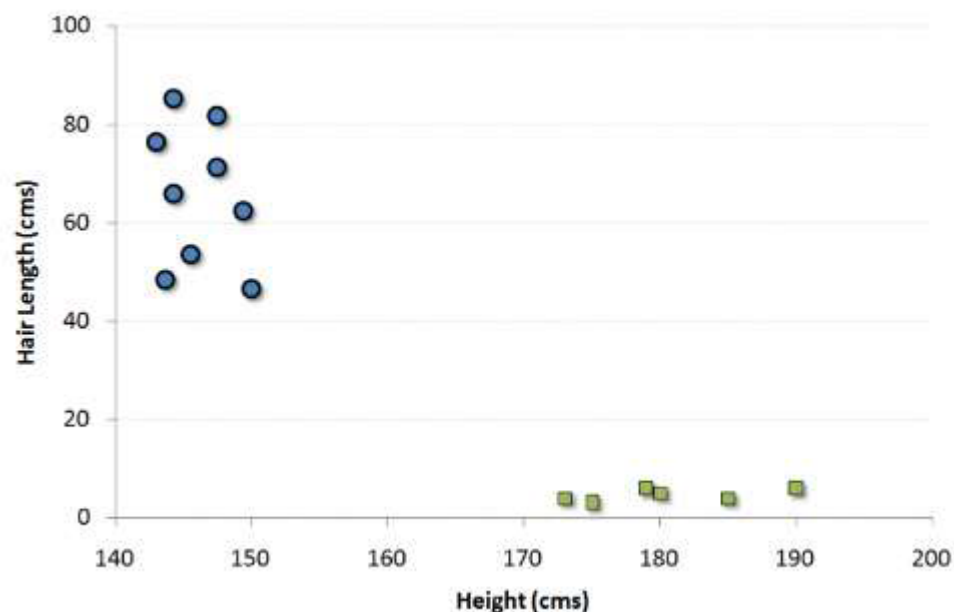


Chart -: Height and hair length graph

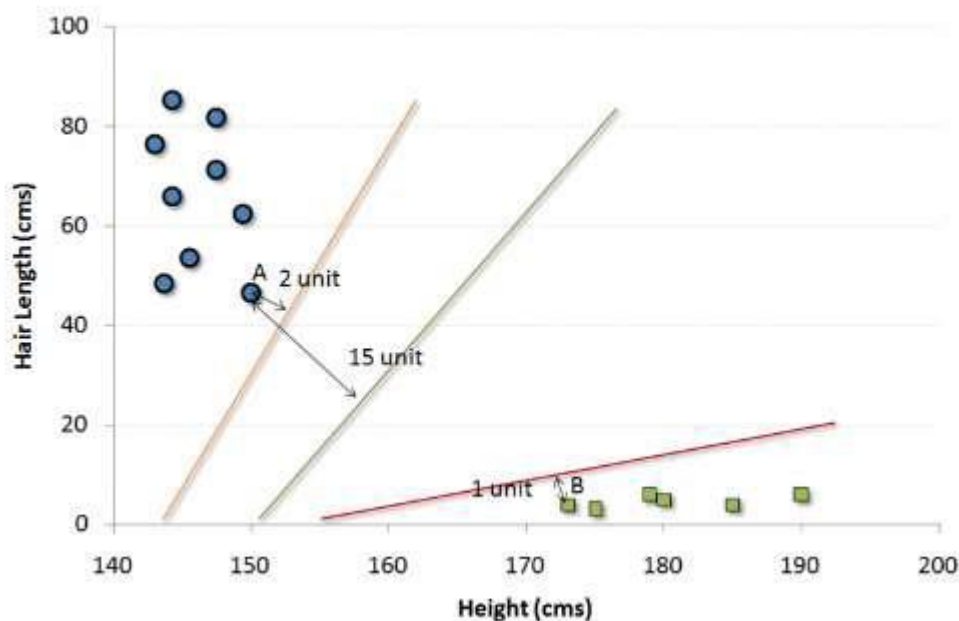
Looking at the graph above, the following facts are easily visible:

1. Men in this collection have a higher average height.
2. Women have longer average hair lengths.

If we were given a new data with 180cm height and 4cm hair length, our best guess for this person would be the men's category.

How to create a machine based on support vectors?

In addition to the data in the example above, we can have many boundaries, three of which are shown below.



The question is, what is the best line in this issue?

An easy way to do this is to create an optimal cluster, calculate the distance of the boundaries obtained with the support vectors of each cluster (the most boundary points of each cluster or class), and finally select the boundary that is the largest of all available clusters. Which in the form of the midline above, a good approximation of the border that is far from both.

APPLICATIONS OF LINEAR ALGEBRA IN ML

Linear Algebra Notation

You can examine and write vector and matrix notation. The set of rules is defined in books, papers, and web sites the usage of vector and matrix notation. The arithmetic of information and notation lets in you describe information control and particular operators. You can examine and write this notation. This talent lets in you to-

1. Read the outline of the algorithms withinside the textbooks
2. Description of recent techniques in studies papers
3. Briefly describe your techniques for different doctors

Programming languages together with Python offer powerful methods to enforce simplicity and algebraic notation in a faster way. Awareness of marking and the way it's far perceived in itself. The language or library lets in the gadget to run small and green mastering algorithms.

Linear Algebra Arithmetic

Linear algebra consists of mathematics operations with notation sharing. Knowing the way to add, subtract, and multiply scalars, vectors, and matrices are essential. An assignment for you is coming to the linear algebra subject in which operations together with matrix multiplication and Tensor Multiplication no longer practice as an instantaneous multiplication of the factors of the systems are, and at the start glance, incomprehensible.

Then, maximum if now no longer all of those features are efficiently carried out and supplied via means of the API Call in Modern Linear Algebra Libraries. Besides, knowledge of vector and matrix operations Matrix notation is carried out and required as a part of powerful analyzing and writing.

Linear Algebra for Statistics

To study facts you want to study linear algebra. Especially the multivariate facts. Mathematics is every other pillar that helps with facts and Machine mastering. They are often involved with describing and knowledge information. As the arithmetic of information, linear algebra has left its fingerprint on many associated fields of arithmetic, consisting of facts.

To examine and interpret information, you want to study the notation and operation of Linear algebra. Modern facts use each notation and gear of linear algebra to explain gear and strategies of statistical techniques. Vectors for the methods and styles of information, for covariate matrices describing the relationships amongst numerous Gaussian variables.

Staples Machine Learning is likewise the end result of a few collaborations among the 2 fields. Methods together with Principal Component Analysis or PCA for brief may be used for information limitations.

Matrix Factorization

The shape of notation and mathematics is the concept of matrix factorization, additionally referred to as Matrix Decomposition. You want to realize the way to validate the matrix and what it means. The Matrix Rationalization is a vital device in linear algebra and is broadly used as a complicated concept.

Linear algebra (Eg, matrix inversion), and gadget mastering.

Also, there's a selection of various matrix factors, each with various strengths and capabilities, a number of which you could apprehend as “gadget mastering” strategies, together with SVD for unmarried-cost decomposition or loss of information. To Read and give an explanation for high-order matrix operations, you want to apprehend the matrix issue.

Linear least squares

You want to realize the way to use the matrix issue to resolve linear least squares. Linear algebra was initially advanced to resolve structures of linear equations. These are the equations in which there are extra equations than unknown variables. As a result, they're hard to resolve arithmetically due to the fact there's no unmarried solution. After all, there's no line or the aircraft isn't enough to provide the specified information without a few errors. These styles of troubles may be formulated as a minimization of squares.

EXAMPLES OF LINEAR ALGEBRA IN MACHINE LEARNING

1. Datasets and information files

In gadget mastering, you healthy the version withinside the dataset. It's a desk like a hard and fast table of numbers in which every row represents the statement and every column represents the feature of the statement.

Below is a fraction of the Iris Flower Dataset 1

1	5.1,3.5,1.4,0.2,Iris-setosa
2	4.9,3.0,1.4,0.2,Iris-setosa
3	4.7,3.2,1.3,0.2,Iris-setosa
4	4.6,3.1,1.5,0.2,Iris-setosa
5	5.0,3.6,1.4,0.2,Iris-setosa

This information is the principle information shape in a matrix, linear algebra. Yet, whilst you are partitioning information into inputs and outputs to in shape supervised gadget mastering fashions for measuring and flowering species, you've got a matrix (X) and a vector (Y).

Vector is every other vital information shape in linear algebra. Each row has the identical duration, i.e., the identical quantity of columns, consequently we will say that the information is vectorized in which rows may be supplied to a version one after the other or in batch and the version may be pre-configured to assume rows of a hard and fast width.

2. Images and photos

Perhaps, they're familiar with running with photographs or photos in laptop imaginative and prescient packages.

Each picture graph you figure with is a desk shape with width and peak and pixel cost in every Mobil Euler for black and white photographs or three-pixel values consistent with Mobil Euler consistent with colour picture graph. An image is every other instance of a matrix from linear algebra. Operations at the picture graph, together with cropping, scaling, shearing, and so forth are all defined by the usage of the notation and operations of linear algebra.

3. One Hot Encoding

Sometimes you figure with categorised information in gadget mastering. Perhaps the magnificence labels for class troubles, or possibly express enter variables. It is not an unusual place to encode express variables to make it less difficult to paint with and study via the means of a few strategies. A famous encoding for express variables is the one-warm encoding. A one-warm encoding is in which a desk is created to symbolize the variable with one column for every class and a row for every instance withinside the dataset. A take a look at or one-cost is introduced withinside the column for the specific cost for a given row, and a 0-cost is introduced to all different columns. For instance, the variable colour variable with the three rows:

For example, the colour variable with the 3 rows:

1	red
2	green
3	blue
4	...

Might be encoded as:

1	red, green, blue
2	1, 0, 0
3	0, 1, 0
4	0, 0, 1
5	...

Each row is encoded as a binary vector, a vector with 0 or one value and that is an instance of sparse illustration, an entire sub-subject of linear algebra.

4. Linear Regression

Linear regression is an old school manner of deciphering facts among relationships. It is frequently utilized in gadget mastering to simplify numerical values in less difficult regression troubles. There are numerous methods to explain and resolve the linear regression problem, which is, to discover a hard and fast of multiples whilst every enter variable is extended and introduced to the output variable is the great reference together. If you've got got used a gadget mastering device or library, the maximum not unusual place manner of fixing linear regression is through a least-squares optimization this is solved the usage of matrix factorization techniques from linear regression, together with an LU decomposition or a singular-cost decomposition or SVD. Even the not unusual place manner of summarizing the linear regression equation makes use of linear algebra notation: $y = A \cdot b$ Where y is the output variable A is the dataset and b are the version coefficients.

Regularization

In implementing gadget mastering, we frequently search for the handiest viable fashions that acquire the great talent on our problem. Simpler fashions are frequently higher at generalizing from particular examples to unseen information. In many techniques that contain coefficients, together with regression techniques and synthetic neural networks, less difficult fashions are frequently characterised by way of means of fashions which have smaller coefficient values. An approach that is frequently used to inspire a version to reduce the dimensions of coefficients even as it's far being healthy on information is referred to as regularization. Common implementations consist of the L 2 and L 1 styles of regularization. Both of those styles of regularization are a degree of the importance or duration of the coefficients as a vector and are techniques lifted immediately from linear algebra referred to as the vector norm.

Principal element evaluation

Often a dataset has many columns, possibly tens, hundreds, thousands, or extra. Modelling information with many capabilities is hard, and fashions constructed from information that consist of inappropriate capabilities are frequently much less skilful than fashions skilled from the maximum applicable information. It is tough to realize which capabilities of the information are applicable and which might be now no longer. Methods for mechanically decreasing the quantity of columns of a dataset are referred to as dimensionality discount, and possibly the maximum famous technique is referred to as the primary element evaluation or PCA for brief. This technique is utilized in gadget mastering to create projections of high-dimensional information for each visualization and education fashions. The middle of the PCA technique is a matrix factorization technique from linear algebra. The eigen-decomposition may be used, and extra strong implementations can also additionally use the singular-cost decomposition or SVD.

Example:

- **Single cost decomposition**

Another famous dimensional discount technique is the unmarried-cost decomposition technique or SVD in brief. According to the matrix and technique name, it distinguishes the factorization technique from linear algebra. It is broadly utilized in linear algebra and may be used immediately in packages together with characteristic selection, visualization, noise discount, and extra.

- **Cryptographic Analysis**

In the sub-subject of gadget mastering to paintings with textual content information, referred to as herbal language processing, it is far not unusual to consult files as a big matrix of phrases. For instance, columns of the matrix can be phrases and rows of sentences recognised withinside the vocabulary, paragraphs, pages, or files of textual content with cells withinside the matrix marked as matter or frequency. It is a small matrix illustration of the textual content. Matrix multiplication strategies together with unmarried-cost decomposition may be implemented. A small matrix this is very applicable for distillation of illustration. It is abstract. Processed files on this manner are clean to compare, question and use. The foundation for a supervised gadget mastering version. This shape of information practise is referred to as latent. LSA for Semantic Analysis or abbreviated as it is also referred to as Lenten Semantic Indexing or LSI.

CONCLUSION

In this report we introduced to Linear Algebra, Machine Learning and Machine Learning algorithms such as linear regression and support vector machine that explicitly have linear relationships to independent and non-independent variables, we learnt the basics of vectors, matrices and tensors and how they are useful in machine learning. We got to know how we can use matrices for datasets and how the vectors are used for classification and boundaries of datasets in n-dimensional space.

Linear algebra is something Machine Learning experts can't live without. You will never evolve from amateur to good specialist without mastering this field of applied mathematics. Usually, math is not the key to success, but in many cases, it helps, and in a few (like developing deep learning models), it is essential. Linear algebra ideas and methods are easiest to comprehend while solving interesting problems, it will help to understand abstract concepts the best. You can see that the deeper you dive into machine learning, the more linear algebra you see there.

Linear regression is an old method from statistics for describing the relationships between variables. It is often used in machine learning for predicting numerical values in simpler regression problems. And Support vector machine are very powerful algorithms in classification and data separation.

REFERENCES

- <https://pabloinsente.github.io/intro-linear-algebra>
- <https://machinelearningmastery.com/introduction-to-tensors-for-machine-learning/>
- <https://machinelearningmastery.com/examples-of-linear-algebra-in-machine-learning/>
- <https://www.educba.com/linear-algebra-in-machine-learning/>
- <https://blog.usejournal.com/easy-as-pie-linear-algebra-for-machine-learning-98acfec084f4>