Lesson One: Introduction to Linear Algebra and Matrices

- Linear algebra is a sub-field of mathematics concerned with vectors, matrices, and operations on these data structures. It is absolutely key to machine learning.
- Linear Algebra is an important part of Mathematical background required not only for Mathematicians but also for other Scientists
- As a machine learning practitioner, you must have an understanding of linear algebra.
- There is need to emphasize the importance of linear algebra to machine learning, vector, and matrix operations, matrix factorization, principal component analysis, and much more.

Why Is Linear Algebra Important to Machine Learning?

- So, why is linear algebra used so much to describe machine learning algorithms?
- Linear algebra is about vectors and matrices and in machine learning we are always working with vectors and matrices (arrays) of data.
- Linear algebra is essentially the mathematics of data.
- It provides useful shortcuts for describing data as well as operations on data that we need to perform in machine learning methods.

FIVE Areas of Linear Algebra to focus on

You don’t need to know all of linear algebra but five key areas of linear algebra are recommended:

1. Learn Linear Algebra Notation

- ✓ Learn how to read and write vector and matrix notation.
- ✓ Algorithms are described in books, papers, and on websites using vector and matrix notation.
- ✓ Linear algebra is the mathematics of data and the notation allows you to describe operations on data precisely with specific operators. You need to be able to read and write this notation.

2. Learn Linear Algebra Arithmetic

- ✓In partnership with the notation of linear algebra are the arithmetic operations performed. You need to know how to add, subtract, and multiply scalars, vectors, and matrices.
- ✓A challenge for newcomers to the field of linear algebra are operations such as matrix multiplication and tensor multiplication that are not implemented as the direct multiplication of the elements of these structures, and at first glance appear non-intuitive.

3. Learn Linear Algebra for Statistics

- Learn linear algebra to be able to learn statistics especially multivariate statistics.
- Statistics is concerned with describing and understanding data. In order to be able to read and interpret statistics, you must learn the notation and operations of linear algebra.

4. Learn Matrix Factorization

- Building on notation and arithmetic is the idea of matrix factorization, also called matrix decomposition. You need to know how to factorize a matrix and what it means.
- Matrix factorization is a key tool in linear algebra and used widely as an element of many more complex operations in both linear algebra (such as the matrix inverse) and machine learning (least squares, PCA and more).

5. Learn Linear Least Squares

- You need to know how to use matrix factorization to solve linear least squares.
- Linear algebra was originally developed to solve systems of linear equations. These are equations where there are more equations than there are unknown variables. As a result, they are challenging to solve arithmetically because there is no single solution as there is no line or plane that can fit the data without some error.
- Such Problems can be framed as the minimization of squared error, called **least squares**, and can be recast in the language of linear algebra, called **linear least squares**.

Introduction to Matrices

- Here, we introduce the concept of matrices.
- Basic definitions are given including the order of a matrix, equal matrices, row matrix, column matrix etc.
- The second subtopic deals with operations on matrices. These include addition, subtraction, scalar multiplication, dot product and matrix multiplication.
- Each concept is illustrated by several examples.

Matrix Defined

- Matrices are used as a shorthand for keeping essential data arranged in rows and columns i.e. matrices are used to summarize data in tabular form.
- **Definition:** A matrix is an ordered rectangular array of numbers, usually enclosed in parenthesis or square brackets. Capital (Upper – case) letters are used to denote matrices.
- An array of numbers organized in form of rows and columns

Order of a Matrix

- The size of a Matrix is specified by

(i) The number of rows (horizontal) and

(ii) The number of columns (vertical).

Types of Matrices

- Row Matrix
- Column Matrix
- Square Matrix
- Diagonal Matrix
- Scalar Matrix
- Identity matrix
- Null Matrix
- Upper Triangular Matrix
- Lower Triangular Matrix
- Symmetric Matrix
- Skew Symmetric Matrix

A general matrix of order m $\times$ n is

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{22} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{mn} \end{bmatrix} \quad \begin{array}{l} = \lfloor a_{ij} \rfloor \\ i \rightarrow ith \ \ row \\ j \rightarrow jth \ column \end{array}$$

A Square Matrix is a one with the same number of rows and columns i.e m $\times$ m matrix. Two matrices are of the same size if they have the same order.
A vector is a matrix with one row (1 $\times$ n) or one column (n $\times$ 1). A row vector is of the form 1 $\times$ n, and a column vector is of the form m $\times$ 1.
A zero matrix of order m $\times$ n is the matrix with $a_{ij} = 0 \quad i = 1, \cdots m, \quad j = 1 \ldots n$.
Similarly we talk of zero rows and column vector.

$$0 = (0, 0, \cdots 0) \text{ or } \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Equal Matrices:

Two matrices A & B are said to be equal if they have the same order (size) m × n and $a_{ij} = b_{ij} \ \forall i \quad \& \forall j$

Operation on Matrices

Addition and Subtraction of matrices

- This is performed on matrices of the **same order** (size).
- Let A and B be m x n matrices.

$$A + B = \left[a_{ij}\right] \pm \left[b_{ij}\right] = \left[a_{ij} \pm b_{ij}\right] = \left[c_{ij}\right] \; (m \times n)$$

Scalar Multiplication

- Performed on any matrix and the resulting matrix is the same size

- Each entry is multiplied by the same number (scalar).

$$cA = c\left[a_{ij}\right] = \left[ca_{ij}\right].$$

Dot product: let $\vec{a}$ and $\vec{b}$ be any two vector of size n (matrices with a single column or row).

$\vec{a} = [a_1, a_2, \cdots a_n]$, $\vec{b} = (b_1, b_2, \cdots b_n)$. The dot product of $\vec{a}$ & $\vec{b}$ denoted $\vec{a} \cdot \vec{b}$, is given by

$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i$. The dot product is also scalar product.

Dot product of a row & column vector of order n.

$$(a_1, a_2 \cdots a_n) \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i = \vec{a} \cdot \vec{b}$$

Dot product examples

first row of matrix A, [1 3], and first column of matrix B, [3 2]. The dot product [1 3] • [3 2] = 1(3) + 2(3) = 9.

$$A = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \quad B = \begin{bmatrix} 7 \\ 9 \\ 11 \end{bmatrix} \quad A^T = [1\ 3\ 5]$$

$$A^T B = [1\ 3\ 5] \begin{bmatrix} 7 \\ 9 \\ 11 \end{bmatrix} = 1 \cdot 7 + 3 \cdot 9 + 5 \cdot 11 = 89$$

Matrix Multiplication:

Let $A=[a_{ik}]$ be an m × n matrix, and $B=[b_{kj}]$ an n × s matrix. The matrix product AB is the m × s matrix $C=[c_{ij}]$ where c_{ij} the dot product of the i^{th} row of A and the j^{th} column of B.

i.e. $AB=C$, $[a_{ik}][b_{kj}]=[c_{ij}]$; $C_{ij}=A_i\cdot B_j=\sum_{k=1}^{n}a_{ik}b_{kj}$

Remark:

1.Let $A(m\times n)$, $B(s\times r)$ be two matrices .

$C=AB$ exists iff $n=s$ & C is $m\times n$

$C=BA$ exists iff $r=m$ & C is $s\times n$

2. It's possible for AB to be defined while BA is not defined. i.e. matrix multiplication is not commutative.

Examples

1. Let $A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{pmatrix}$. Then

$$A + B = \begin{pmatrix} 1+3 & -2+0 & 3+2 \\ 4-7 & 5+1 & -6+8 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 5 \\ -3 & 6 & 2 \end{pmatrix}$$

2. $3A = \begin{pmatrix} 3\cdot 1 & 3\cdot(-2) & 3\cdot 3 \\ 3\cdot 4 & 3\cdot 5 & 3\cdot(-6) \end{pmatrix} = \begin{pmatrix} 3 & -6 & 9 \\ 12 & 15 & -18 \end{pmatrix}$

3. $2A-3B=\begin{pmatrix}2 & -4 & 6\\ 8 & 10 & -12\end{pmatrix}+\begin{pmatrix}-9 & 0 & -6\\ 21 & -3 & -24\end{pmatrix}=\begin{pmatrix}-7 & -4 & 0\\ 29 & 7 & -36\end{pmatrix}$

4. $\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}=\begin{pmatrix}1\cdot 1+2\cdot 0 & 1\cdot 1+2\cdot 2\\ 3\cdot 1+4\cdot 0 & 3\cdot 1+4\cdot 2\end{pmatrix}=\begin{pmatrix}1 & 5\\ 3 & 11\end{pmatrix}$

$$\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}=\begin{pmatrix}1\cdot 1+1\cdot 3 & 1\cdot 2+1\cdot 4\\ 0\cdot 1+2\cdot 3 & 0\cdot 2+2\cdot 4\end{pmatrix}=\begin{pmatrix}4 & 6\\ 6 & 8\end{pmatrix}$$

The above example shows that matrix multiplication is not commutative, i.e.the products AB and BA of matrices need not be equal.

1.3 Assessment Questions

1. Let $A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix}$. Find (a) AB, (b) BA

2. Given $A = (2, -1)$ and $B = \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & -3 \end{pmatrix}$, find (a) AB, (b) BA

3. Given $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}$, find (a) AB, (b) BA

Questions? Comments?