EC3357:Machine Learning

Lecture 2: Linear Algebra

What is Linear Algebra

- Linear algebra is central to almost all areas of mathematics. For instance, linear algebra is fundamental in modern presentations of geometry, including for defining basic objects such as lines, planes and rotations.
- Also, functional analysis, a branch of mathematical analysis, may be viewed as basically the application of linear algebra to spaces of functions.
- Linear algebra is also used in most sciences and fields of engineering, because it allows modeling many natural phenomena, and computing efficiently with such models.
- For nonlinear systems, which cannot be modeled with linear algebra, it is often used for dealing with first-order approximations, using the fact that the differential of a multivariate function at a point is the linear map that best approximates the function near that point.

What is Linear Algebra

- Linear algebra is the branch of mathematics concerning linear equations such as:

$$a_1x_1 + \cdots + a_nx_n = b,$$

- linear maps such as:

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \ldots + a_nx_n,$$

- and their representations in vector spaces and through matrices.

Numerical Linear Algebra

- The application of linear algebra in computers is often called numerical linear algebra.
- It is more than just the implementation of linear algebra operations in code libraries; it also includes the careful handling of the problems of applied mathematics, such as working with the limited floating point precision of digital computers.
- Computers are good at performing linear algebra calculations, and much of the dependence on Graphical Processing Units (GPUs) by modern machine learning methods such as deep learning is because of their ability to compute linear algebra operations fast.

Linear Algebra and Statistics

- Some clear fingerprints of linear algebra on statistics and statistical methods include:
  - Use of vector and matrix notation, especially with multivariate statistics.
  - Solutions to least squares and weighted least squares, such as for linear regression.
  - Estimates of mean and variance of data matrices.
  - The covariance matrix that plays a key role in multinomial Gaussian distributions.
  - Principal component analysis for data reduction that draws many of these elements together.

Applications of Linear Algebra

- Matrices in Engineering, such as a line of springs.
  - Graphs and Networks, such as analyzing networks.
  - Markov Matrices, Population, and Economics, such as population growth.
  - Linear Programming, the simplex optimization method.
  - Fourier Series: Linear Algebra for functions, used widely in signal processing.
  - Linear Algebra for statistics and probability, such as least squares for regression.
  - Computer Graphics, such as the various translation, rescaling and rotation of images.

Reference: Introduction to Linear Algebra, Gilbert Strang

Linear Algebra in Machine Learning

- 1. Dataset and Data Files
- 2. Images and Photographs
- 3. One Hot Encoding
- 4. Linear Regression
- 5. Regularization
- 6. Principal Component Analysis
- 7. Singular-Value Decomposition
- 8. Latent Semantic Analysis
- 9. Recommender Systems
- 10. Deep Learning

Vector

- A vector is a tuple of one or more values called scalars.
- Vectors are built from components, which are ordinary numbers.
- You can think of a vector as a list of numbers, and vector algebra as operations performed on the numbers in the list.

Reference: No Bullshit Guide To Linear Algebra, 2017

Vector

- Vectors are often represented using a lowercase character such as v; for example:

$$v = (v1 , v2 , v3)$$

- Where v 1 , v 2 , v 3 are scalar values, often real values.
- Vectors are also shown using a vertical representation or a column; for example:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Vector

- It is common to represent the target variable as a vector with the lowercase y when describing the training of a machine learning algorithm.
- It is common to introduce vectors using a geometric analogy, where a vector represents a point or coordinate in an n-dimensional space, where n is the number of dimensions, such as 2.
- The vector can also be thought of as a line from the origin of the vector space with a direction and a magnitude.

Vector Arithmetic

- **Addition**
- Two vectors of equal length can be added together to create a new third vector.

$$c = a + b$$

- The new vector has the same length as the other two vectors.
- Each element of the new vector is calculated as the addition of the elements of the other vectors at the same index; for example:

$$c = (a1 + b1, a2 + b2, a3 + b3)$$

Vector Arithmetic

- **Addition**
- Putting another way,

$$c[0] = a[0] + b[0]$$

$$c[1] = a[1] + b[1]$$

$$c[2] = a[2] + b[2]$$

Vector Arithmetic

- **Subtraction**
- One vector can be subtracted from another vector of equal length to create a new third vector.

$$c = a - b$$

- As with addition, the new vector has the same length as the parent vectors and each element of the new vector is calculated as the subtraction of the elements at the same indices.

$$c = (a_1 - b_1 , a_2 - b_2 , a_3 - b_3)$$

Vector Arithmetic

- **Subtraction**
- Put another way:

$$c[0] = a[0] - b[0]$$

$$c[1] = a[1] - b[1]$$

$$c[2] = a[2] - b[2]$$

Vector Arithmetic

- **Multiplication**
- Two vectors of equal length can be multiplied together.

$$c = a \times b$$

- As with addition and subtraction, this operation is performed element-wise to result in a new vector of the same length.

$$c = (a_1 \times b_1 , a_2 \times b_2 , a_3 \times b_3) \quad or$$

$$c = (a_1b_1 , a_2b_2 , a_3b_3)$$

Vector Arithmetic

- **Multiplication**
- Put another way:

$$c[0] = a[0] \times b[0]$$

$$c[1] = a[1] \times b[1]$$

$$c[2] = a[2] \times b[2]$$

Vector Arithmetic

- **Division**
- Two vectors of equal length can be divided.

$$c = \frac{a}{b}$$

- As with other arithmetic operations, this operation is performed element-wise to result in a new vector of the same length.

$$c = (\frac{a_1}{b_1}, \frac{a_2}{b_2}, \frac{a_3}{b_3})$$

Vector Arithmetic

- **Division**
- Put another way:.

    c[0] = a[0]/b[0]

    c[1] = a[1]/b[1]

    c[2] = a[2]/b[2]

Vector Arithmetic

- **Dot Product**
- We can calculate the sum of the multiplied elements of two vectors of the same length to give a scalar.
- This is called the dot product, named because of the dot operator used when describing the operation.
- The dot product is the key tool for calculating vector projections, vector decompositions, and determining orthogonality.
- The name dot product comes from the symbol used to denote it.

$$c = a \cdot b$$

Vector Arithmetic

- **Dot Product**
- The operation can be used in machine learning to calculate the weighted sum of a vector.
- The dot product is calculated as follows:

$$c = (a_1 \times b_1 + a_2 \times b_2 + a_3 \times b_3)$$

or

$$c = (a_1b_1 + a_2b_2 + a_3b_3)$$

Vector Arithmetic

- **Scalar Multiplication**
- A vector can be multiplied by a scalar, in effect scaling the magnitude of the vector. To keep notation simple, we will use lowercase s to represent the scalar value.

$$c = s \times v \quad or$$

$$c = sv$$

- The multiplication is performed on each element of the vector to result in a new scaled vector of the same length.

$$c = (s \times v_1, s \times v_2, s \times v_3)$$

Matrix

- Matrices are a foundational element of linear algebra.
- Matrices are used throughout the field of machine learning in the description of algorithms and processes such as the input data variable (X) when training an algorithm.
- A matrix is a two-dimensional array of scalars with one or more columns and one or more rows.
- A matrix is a two-dimensional array (a table) of numbers.

Matrix

- The notation for a matrix is often an uppercase letter, such as A, and entries are referred to by their two-dimensional subscript of row ($_i$) and column ($_j$), such as $a_{i,j}$.
- For example, we can define a 3-row, 2-column matrix:

  $A = ((a_{1,1}, a_{1,2}), (a_{2,1}, a_{2,2}), (a_{3,1}, a_{3,2}))$
- It is more common to see matrices defined using a horizontal notation.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix}$$

Matrix Arithmetic

Matrix Addition

- Two matrices with the same dimensions can be added together to create a new third matrix.
- C = A + B
- The scalar elements in the resulting matrix are each of the matrices being added.

$$C = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} \\ a_{3,1} + b_{3,1} & a_{3,2} + b_{3,2} \end{pmatrix}$$

Matrix Arithmetic

Matrix Subtraction

- Two matrices with the same dimensions can be subtracted together to create a new third matrix.
- C= A – B
- The scalar elements in the resulting matrix are each of the matrices being subtracted.

$$C = \begin{pmatrix} a_{1,1} - b_{1,1} & a_{1,2} - b_{1,2} \\ a_{2,1} - b_{2,1} & a_{2,2} - b_{2,2} \\ a_{3,1} - b_{3,1} & a_{3,2} - b_{3,2} \end{pmatrix}$$

Matrix Arithmetic

- **Matrix Multiplication (Hadamard Product)**
- Two matrices with the same size can be multiplied together, and this is often called element-wise matrix multiplication or the Hadamard product.
- It is not the typical operation meant when referring to matrix multiplication, therefore a different operator is often used, such as a circle $\circ$.

$$C = A \circ B$$

$$C = \begin{pmatrix} a_{1,1} \times b_{1,1} & a_{1,2} \times b_{1,2} \\ a_{2,1} \times b_{2,1} & a_{2,2} \times b_{2,2} \\ a_{3,1} \times b_{3,1} & a_{3,2} \times b_{3,2} \end{pmatrix}$$

Matrix Arithmetic

- **Matrix Division**
- One matrix can be divided by another matrix with the same dimensions.

$$C = \frac{A}{B}$$

- The scalar elements in the resulting matrix are calculated as the division of the elements in each of the matrices.

$$C = \begin{pmatrix} \frac{a_{1,1}}{b_{1,1}} & \frac{a_{1,2}}{b_{1,2}} \\ \frac{a_{2,1}}{b_{2,1}} & \frac{a_{2,2}}{b_{2,2}} \\ \frac{a_{3,1}}{b_{3,1}} & \frac{a_{3,2}}{b_{3,2}} \end{pmatrix}$$

Matrix Arithmetic

- **Matrix-Matrix Multiplication**
- Matrix multiplication, also called the matrix dot product is more complicated than the previous operations and involves a rule as not all matrices can be multiplied together.

$$C = A \cdot B \text{ or}$$

$$C = AB$$

- The rule for matrix multiplication is as follows:
- The number of columns (n) in the first matrix (A) must equal the number of rows (m) in the second matrix (B).

Example

Given the following two matrices,

$$[A] = \begin{bmatrix} 5 & 2 & 3 \\ 1 & 2 & 7 \end{bmatrix} \qquad [B] = \begin{bmatrix} 3 & -2 \\ 5 & -8 \\ 9 & -10 \end{bmatrix}$$

Find their product,

$$[C] = [A][B]$$

Example (in Phyton)

```
# add matrices
from numpy import array
A = array([[1, 2, 3], [4, 5, 6]])
print(A)
B = array([[1, 2, 3], [4, 5, 6]])
print(B)
C = A + B
print(C)
```

Finding the Inverse of a 2x2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Step-1 First find what is called the Determinant

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

This is calculated as ad-bc

Step-2 Then swap the elements in the leading diagonal

$$\begin{bmatrix} d & b \\ c & a \end{bmatrix}$$

Step-3 Then negate the other elements

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Step-4 Then multiply the Matrix by 1/determinant

$$\frac{1}{ad-cb}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Inverse of a Matrix—Definition

- Let A be a square n x n matrix.

- If there exists an n x n matrix A^{-1} with the property that

$$AA^{-1} = A^{-1}A = I_n$$

  then we say that A^{-1} is the inverse of A.

E.g. Verifying that a Matrix Is an Inverse

- Verify that *B* is the inverse of *A*, where:

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$$

- We perform the matrix multiplications to show that $AB = I$ and $BA = I$.

Example Find Inverse of A

Step 1 – Calc Determinant

$$A = \begin{pmatrix} 4 & 8 \\ 1 & 3 \end{pmatrix}$$

Determinant (ad-cb) = 4x3-8x1 = 4

Step 2 – Swap Elements on leading diagonal

$$\textbf{step2} \quad \begin{pmatrix} 3 & 8 \\ 1 & 4 \end{pmatrix}$$

Step 3 – negate the other elements

$$\textbf{step3} \quad \begin{pmatrix} 3 & -8 \\ -1 & 4 \end{pmatrix}$$

Step 4 – multiply by 1/determinant

$$\textbf{step4} \quad \frac{1}{4}\begin{pmatrix} 3 & -8 \\ -1 & 4 \end{pmatrix}$$

$$A^{-1} = \begin{bmatrix} 0.75 & -2 \\ -0.25 & 1 \end{bmatrix}$$

check

$$AA^{-1} = \begin{bmatrix} 4 & 8 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0.75 & -2 \\ -0.25 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3-2 & -8+8 \\ 0.75-0.75 & -2+3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

E.g. Finding the Inverse of a 2 x 2 Matrix

- Let A be the matrix

$$A = \begin{bmatrix} 4 & 5 \\ 2 & 3 \end{bmatrix}$$

- Find A^{-1} and verify that

$$AA^{-1} = A^{-1}A = I_2$$

Matrix transpose

- **Definition:**

The ***transpose*** of an *m* x *n* matrix *A* is the *n* x *m* matrix *AT* obtained by interchanging rows and columns of *A*,

$$\text{i.e.,}\quad (A^T)_{ij} = A_{ji} \quad \forall\, i, j.$$

Example

$$A = \begin{bmatrix} 1 & 3 & 5 & -2 \\ 5 & 3 & 2 & 1 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 5 \\ 3 & 3 \\ 5 & 2 \\ -2 & 1 \end{bmatrix}$$

Transpose operation can be viewed as flipping entries about the diagonal.

Properties of transpose

apply twice -- get back
to where you started

(1) $(A^T)^T = A$

(2) $(A + B)^T = A^T + B^T$

(3) For a scalar c, $(cA)^T = cA^T$

(4) $(AB)^T = B^T A^T$

To prove this, we show that
$[(AB)^T]_{ij} =$

Inverse Matrix using NumPy

```
# Import required package
import numpy as np

# Taking a 4 * 4 matrix
A = np.array([[6, 1, 1, 3],
              [4, -2, 5, 1],
              [2, 8, 7, 6],
              [3, 1, 9, 7]])

# Calculating the inverse of the matrix
print(np.linalg.inv(A))
```