Linear Algebra

Lecture-03

By Noorullah ibrahimi

Learning Objectives

- **Diagonal of a Matrix:** Identifies special elements (main diagonal) important in identity and triangular matrices.
- **Trace of a Matrix:** Sum of diagonal elements, used as an invariant in linear algebra and optimization.
- **Inner Product of Matrices:** Measures similarity between matrices, defines norms and orthogonality in data analysis.

Learning outcomes

- **Diagonal of a Matrix:** Be able to identify and use main diagonal elements in forming identity, triangular, and diagonal matrices.
- **Trace of a Matrix:** Be able to compute the trace.
- **Inner Product of Matrices**: Be able to calculate inner products.

The diagonal of a Matrix

Defination (Matrix Diagonal): given a square Matrix ***X*** the diagonal of a matrix is a vector form from the diagonal elements of a matrix

Example. Given a matrix ***X*** , ***Z*** defined as

$$X = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 4 & 0 \\ 7 & 3 & 9 \end{bmatrix}, \quad D\mathbf{i}ag = (X) = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} \quad Z = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}, \quad D\mathbf{i}ag(Z) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

Diagonal of a non-square matrix

The diagonal of a non-square matrix is the set of elements

$$a_{11}, a_{22}, a_{33}, \ldots, a_{kk}$$

where

k=min(m,n) **for an m×n**

Example

Take this 2 × 3 matrix (not square):

$$A = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 7 \end{bmatrix}$$

$$\mathbf{Diag(A)=}\begin{bmatrix} \mathbf{1} \\ \mathbf{6} \end{bmatrix}$$

The Trace of a Matrix

Definition (Matrix Trace): given a square matrix$X, Z$the Trace of a matrix is the sum of all elements in the diagonal of a matrix.

Example. Given a matrix $X,$ Zdefined as

$$X = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 4 & 0 \\ 7 & 3 & 9 \end{bmatrix}$$

Tr(X) = 1 + 4 + 9

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

Tr(Z)= 1+4 = 5

Matrix inner product

The matrix inner product is different from the vector inner product with two matrices X and Y their inner product is defined as

$$\langle X, Y\rangle = \mathrm{Tr}(XY^T) \tag{1}$$

For example, given

$$X = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \tag{2}$$

Then

$$\langle X, Y\rangle = Tr\left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}^T\right) = Tr\left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}\right) \tag{3}$$

$$= Tr\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) = 1 + 4 = 5 \tag{4}$$

Take a minute and see if you can find $\langle \mathrm{Y}, \mathrm{X}\rangle$

$$\langle X, Y \rangle = \mathrm{Tr}(XY^T) \tag{1}$$

For example, given

$$X = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \tag{2}$$

Then

$$\langle Y, X \rangle = Tr\left(\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^T \right) = Tr\left(\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \right) \tag{3}$$

$$= \ Tr\left(\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \right) = 1 + 4 = 5 \tag{4}$$

We previously knew that $XY^T = \begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$, we now need to know YX^T , we can find it easily where since

$$YX^T = (XY^T)^T = \begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}^T = \begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix}$$

Matrix/Vector Multiplication Practice

Given

$$\boldsymbol{x} = \begin{bmatrix}1\\ 2\end{bmatrix}, X = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, y = \begin{bmatrix}3\\ 1\\ 3\end{bmatrix}, Y = \begin{bmatrix}1 & 3 & 0\\ 2 & 3 & 1\end{bmatrix}, z = [1\ \ 3\ \ 0], Z = \begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 1\\ 1 & 1 & 1\end{bmatrix}$$

1. $Y^T x$
2. $\mathrm{Z} \odot \mathrm{Z}$
3. $2Z\mathrm{y}$
4. $\langle X, X\rangle$
5. $x \otimes y$
6. $D\mathbf{i}ag(X)$
7. $\mathrm{Tr}(Z)$
7. $\mathrm{Tr}(Z)$

In addition, given

$$\boldsymbol{u} = \begin{bmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\beta}\\ \boldsymbol{\gamma}\end{bmatrix}, \boldsymbol{v} = \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}$$

Solve for

$\mathbf{1} \cdot \boldsymbol{v} \otimes \boldsymbol{v}$ $\quad$ $\mathbf{2}.\ \mathbf{Tr}(\boldsymbol{v} \otimes \boldsymbol{v})$ $\quad$ $\mathbf{3}.\,\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{u}$ $\quad$ $\mathbf{4}.\,\boldsymbol{Z}\boldsymbol{u}$

Solution to exercise

1. $Y^T x = \begin{bmatrix} 1 & 2 \\ 3 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 * 1 & + & 2 * 2 \\ 3 * 1 & + & 3 * 2 \\ 0 * 1 & + & 1 * 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \\ 2 \end{bmatrix}$

2. $\mathrm{Z} \odot \mathrm{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \odot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

3. $2Zy + 1 = 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} + 1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 2 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \\ 14 \end{bmatrix} + 1 = \begin{bmatrix} 7 \\ 11 \\ 15 \end{bmatrix}$

4. $\langle X, X \rangle = \mathrm{Tr}(XX^T) = Tr\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = \mathrm{Tr} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1 + 1 = 2$

5. $x \otimes y = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 3 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 3 \\ 6 & 2 & 6 \end{bmatrix}$

6. $Diag(X) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$7. \operatorname{Tr}(Z) = Tr\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 1\\ 1 & 1 & 1\end{bmatrix} = 1 + 2 + 1 = 4$$

$$8. x^T Y z^T = \begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}1 & 3 & 0\\ 2 & 3 & 1\end{bmatrix}\begin{bmatrix}1\\ 3\\ 0\end{bmatrix} = \begin{bmatrix}5 & 9 & 2\end{bmatrix}\begin{bmatrix}1\\ 3\\ 0\end{bmatrix} = 32$$

Solution to symbolic Representation

$$3. Z^T u = \begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 1\\ 1 & 1 & 1\end{bmatrix}^T \begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} = \begin{bmatrix}1 & 0 & 1\\ 0 & 2 & 1\\ 0 & 1 & 1\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} = \begin{bmatrix}\alpha + \gamma\\ 2\beta + \gamma\\ \beta + \gamma\end{bmatrix}$$

$$4. Yu = \begin{bmatrix}1 & 3 & 0\\ 2 & 3 & 1\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} = \begin{bmatrix}\alpha + 3\beta\\ 2\alpha + 3\beta + \gamma\end{bmatrix}$$

Notice Large equations can be compressed into matrix format

Notice how we can rewrite an equation with vector matrix format

$$f(x) = 1x_1 + 3x_2 = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax$$

Imagine if we have 1000 variables, it would be annoying to write it each time

but with matrix notation we can write it compactly

$$f(x) = 1x_1 + 2x_2 + 3x_3 + \cdots + 1000x_{1000} = \begin{bmatrix} 1 & 2 & 3 \cdots 1000 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_{1000} \end{bmatrix}$$

$$= ax$$

This is a function where $f\colon \mathbb{R}^{1000} \to \mathbb{R}$ but the same idea apply if we have output of multiple dimensions. $f\colon \mathbb{R}^{1000} \to \mathbb{R}^{20}$

Compressing Large Equations with multiple inputs/outputs

This is an example of function with multiple inputs and multiple outputs

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

$$f(x) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 + 2x_2 \\ 0x_1 + 1x_2 \\ 3x_1 + 2x_2 \end{bmatrix}$$

If you put an input of $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ of 2 dimensions, you will get a three dimensions output

$$f(x) = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$$

In this case we can represent this systems of linear equations simply as

$$f(x) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Ax$$

Compact way of representing thousands of equations simultaneously

Here is a problem you can see this type of problem in the SAT

A concert charges $20 for adult tickets and $10 for student's tickets. If the concert sold 100 tickets and made $2000. how many of each type of tickets were sold?

The problem gives us two pieces of information

- Let a be the number of adult tickets sold
- Let s be the number of students tickets sold

Then, we know that

1. adult + student is 100 tickets or a + s = 100
2. $20a + 10s = 2000$

With two equations and two unknowns, you should be able to find a and s.

this is called **linear system of equations** or just **linear system**

$$\begin{aligned} a + s &= 100 \\ 20a + 10s &= 2000 \end{aligned}$$

The standard way we learn to this is through substitution

Given the linear system of

$a + s = 100$

$20a + 10s = 2000$

we let a be

$a = 100 - s$

And we substitute a into the 2nd equations

$20(100 - s) + 10s = 2000$

this allow us to solve for s

$2000 - 20s + 10s = 2000$

$-10s = 0$

$s = 0$

This approach is useful with only a few variables to solve, but what if we have a bigger problem?

It would take a long time to solve this. What is the alternative?

Gaussian Elimination.

$$2x_1 - 3x_2 + 4x_3 + 5x_4 - 6x_5 + 7x_6 - 8x_7 + 9x_8 - 10x_9 + 11\mathrm{x}10 = 12$$
$$-3x_1 + 4x_2 - 5x_3 + 6 \times 4 - 7x_5 + 8x_6 - 9x_7 + 10x_8 - 11x_9 + 12x_{10} = 13$$
$$4x_1 - 5x_2 + 6x_3 - 7x_4 + 8x_5 - 9x_6 + 10x_7 - 11x_8 + 12x_9 - 13x_{10} = 14$$
$$5x_1 - 6x_2 + 7x_3 - 8x_4 + 9x5 - 10x_6 + 11x_7 - 12x_8 + 13x_9 - 14x_{10} = 15$$
$$-6x_1 + 7x_2 - 8x_3 + 9x_4 - 10x_5 + 11x_6 - 12x_7 + 13x_8 - 14x_9 + 15x_{10} = 16$$
$$7x - 8x_2 + 9x_3 - 10x_4 + 11x_5 - -12x_6 + 13x_7 - 14x_8 + 15x_9 - 16x_{10} = 17$$
$$-8x_1 + 9\chi_2 - 10x_3 + 11x_4 - 12x_5 + 13x_6 - 14\chi_7 + 15x_8 - 16x_9 + 17x_{10}18$$
$$9x_1 - 10x_2 + 11 \times 3 - 12x_4 + 13x_5 - 14x_6 + 15x_7 - 16x_8 + 17x_9 - 18x_{10} = 19$$
$$-10x_1 + 11x_2 - 12x_3 + 13x_4 - 14x_5 + 15x_6 - 16x_7 + 17x_8 - 18x_9 + 19x_{10} = 20$$
$$11x_1 - 12x_2 + 13x_3 - 14x_4 + 15x_5 - 16x_8 + 17x_7 - 18x_8 + 19x_9 + 20x_{10} = 21$$

Manipulating the truth

Let’s say have a vector x and two function defined as

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \, and \, \begin{cases} 2x_1 + x_2 = 1 \\ -x_2 = 1 \end{cases}$$

In this case we are telling you that these 2 equations are true given some x

We can also see that these statements both are true if we put $x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ into the equations

$$\begin{cases} 2(1) + (-1) = 1 \rightarrow 1 = 1 \\ -(-1) = 1 \rightarrow 1 = 1 \end{cases}$$

An interesting fact about true statements is that they can be manipulated and remain true. For example we can multiply both sides of the equation by a number……

$$2x_1 + x_2 = 1$$
$$2(2x_1 + x_2) = 2(1)$$

Let us verify the last equation if it is still true for $x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$2\big(2(1) + (-1)\big) = 2(1) \Rightarrow 4 - 2 = 2 \Rightarrow 2 = 2$$

Even though we have two statements that look different actually they are actually the same statements presented in 2 different ways

$$2x_1 + x_2 = 1 \quad \text{and} \quad 4x_1 + 2x_2 = 2$$

We are going to symbolically represent what we did to equation 1 as

$$2e_1$$

These are a couple of fact to note

1. A true statement requires an equal sign like 5 = 5 we can very easily see that 3(5) = 3(5)
2. These mathematical “True” statements are only true given a specific x. note if we use a different x the statements are no longer true. For example given

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} then \begin{cases} 2(1) + (0) \neq 1 \\ \quad -(0) \neq 1 \end{cases}$$

However under correct condition the truth can be manipulated and remain true. This allow us to conclude the firs observation

Conclusion1: the result of multiplication and division of a true statement on both side of the equation is still true

The 2nd conclusion

let’s we have a vector x and two functions defined as follow

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{cases} 2x_1 + x_2 = 1 \\ \quad -x_2 = 1 \end{cases}$$

Notice the order of the true statements does not matter

$$\begin{cases} 2x_1 + x_2 = 1 \\ \quad -x_2 = 1 \end{cases} \quad This\ is\ the\ original\ true\ statement$$

$$\begin{cases} \quad -x_2 = 1 \\ 2x_1 + x_2 = 1 \end{cases} \quad this\ is\ the\ fillped\ \ statement\ but\ they\ still\ true$$

Here we see that equation 1 e1 become e2. we are going to mathematically represent this as

$$e_1 \leftrightarrow \ e_2$$

The 3rd conclusion

let’s we have a vector x and two functions defined as follow

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{cases} 2x_1 + x_2 = 1 \\ \quad -x_2 = 1 \end{cases}$$

Notice if we combine two true statements the result is still true

$$\begin{array}{r} 2x_1 + x_2 = 1 \\ -x_2 = 1 \\ \hline 2x_1 + 0x_2 = 2 \end{array}$$

We are going to symbolically represent what we just did as

$$e_1 + e_2$$

The combination of equation 1 (e1) and equation 2 (e2) yields a brand new equation

$$2x_1 = 2$$

and this equation is still true for x1 = 1 let's check

$$2(1) = 2 \Rightarrow 2 = 2$$

Final conclusion: The combination of two true statements yields another true statement

Combining the three conclusions

By combining the three conclusion it allow us to solve really complex problems. Let's say that we first replace equation 1 (e1) with e1+ e2 to get two new statements

$$\begin{cases} 2x_1 + 0x_2 = 2 \\ 0x_1 - x_2 = 1 \end{cases}$$

Now, let’s perform $0.5e_1$ and $-1e_2$ to get yet another statement as

$$\begin{cases} 2x_1 + 0x_2 = 2 \\ 0x_1 - x_2 = 1 \end{cases} \xrightarrow[-1e_2]{0.5e_1} \begin{cases} x_1 + 0x_2 = 1 \\ 0x_1 + x_2 = -1 \end{cases} = \begin{cases} x_1 = 1 \\ x_2 = -1 \end{cases}$$

Take a note and study what just happened

In this special example, we already know x1 and x2. However if we don’t know this solution, we can manipulate the existing equations to tell us the solutions for x1 and x2
This work for any size of the problem, the goal is to manipulate the equations to get them into a special form that look like this

$$\begin{aligned} x_1 + 0x_2 + 0x_3 + 0x_4 &= 3 \\ 0x_1 + x_2 + 0x_3 + 0x_4 &= 2 \\ 0x_1 + 0x_2 + x_3 + 0x_4 &= 1 \\ 0x_1 + 0x_2 + 0x_3 + x_4 &= 1 \end{aligned}$$

Once we get into this format the solution for x1, x2, x3, x4 are automatically known. This process is called Gaussian Elimination

Let’s look at a simple example

Given the following system of linear equations we first multiply e2 by 2

$$\begin{cases} e_1: 2x_1 + 2x_2 = 2 \\ e_2: x_1 - x_2 = 1 \end{cases} \xrightarrow[\text{2e2}]{} \begin{cases} \text{e1}: 2x_1 + 2x_2 = 2 \\ \text{e2}: 2x_1 - 2x_2 = 2 \end{cases}$$

now we can combine the two equations

$$\begin{cases} \text{e1}: 2x_1 + 2x_2 = 2 \\ \text{e2}: 2x_1 - 2x_2 = 2 \end{cases} \xrightarrow[e_1 + e_2 \to e_1]{} \begin{cases} \text{e1}:\ 4x_1 + 0x_2 = 4 \\ e2:\ 2x_1 - 2x_2 = 2 \end{cases}$$

Notice that we eliminated the blue term. We cam also further simplify the e1 by dividing it by 4

$$\begin{cases} \text{e1}:\ 4x_1 + 0x_2 = 4 \\ \text{e2}:\ 2x_1 - 2x_2 = 2 \end{cases} \xrightarrow[e_1/4]{} \begin{cases} \text{e1}: x_1 + 0x_2 = 1 \\ \text{e2}:\ 2x_1 - 2x_2 = 2 \end{cases}$$

To remove x1 from e2 we multiply e1 by -2 and add to e2

$$\begin{cases} \text{e1}: x_1 + 0x_2 = 1 \\ \text{e2}:\ 2x_1 - 2x_2 = 2 \end{cases} \xrightarrow[-2e_1 + e_2 \to e_2]{} \begin{cases} \text{e1}: x_1 + 0x_2 = 1 \\ \text{e2}:\ 0x_1 - 2x_2 = 0 \end{cases}$$

After dividing e2 by 2 we have essentially solved the problem

$$\begin{cases} \text{e1: } x_1 + 0x_2 = 1 \\ \text{e2: } 0x_1 - 2x_2 = 0 \end{cases} \xrightarrow[e_2/-2]{} \begin{cases} e1\text{: } x_1 + 0x_2 = 1 \\ \text{e2: } 0x_1 + x_2 = 0 \end{cases}$$

We have x1 = 1 and x2 = 0

To double check our solution, we just plug it back in

$$\begin{cases} \text{e1: } 2(1) + 2(0) = 2 \\ e2\text{: } (1) - (0) = 1 \end{cases}$$

Notation Simplification

Notice that when we use Gaussian Elimination so far, we have to copy the x 1 and x 2 over and over again. This is really not necessary. Instead, we can simplify the copying by writing the system in matrix format.

Instead of writing

$$\begin{aligned} &\text{e1:} \quad 2x_1 + x_2 = 1 \\ &\text{e2:} \quad -x_2 = 1 \end{aligned}$$

The matrix notation becomes

$$\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We now can ignore x1, x2 and rewrite it as

$$\left[\begin{array}{cc|c} 2 & 1 & 1 \\ 0 & -1 & 1 \end{array}\right]$$

We call this notation **augmented matrix**

All of the row operations and combination stay the same