Introduction to Linear Algebra

1 Why Linear Algebra is Important

Linear algebra is one of the fundamental branches of mathematics, with applications spanning many different fields. It provides the tools to work with vectors and matrices, which are central to the representation and manipulation of linear systems. Linear algebra is vital for:

- **Computer Science**: Algorithms, machine learning (e.g., neural networks), computer graphics, and optimization rely on matrix operations and vector spaces.
- **Physics and Engineering**: Linear systems of equations, differential equations, and the study of physical systems are often described using linear algebra.
- **Economics and Finance**: Modeling multi-variable systems, like portfolio optimization or resource allocation, requires the use of linear algebra.
- **Data Science**: Principal component analysis (PCA), singular value decomposition (SVD), and other techniques for dimensionality reduction are all rooted in linear algebra.

In this set of lecture notes, we will cover fundamental concepts such as vectors, dot products, matrices, matrix operations, and more.
3Blue1Brown video series: `https://youtu.be/kjBOesZCoqc?si=UUkI12_ND45JzhCQ`

2 Vectors

A **vector** is an element of a vector space, often represented as an ordered list of numbers. Vectors are used to represent quantities that have both a magnitude and direction. A vector in $\mathbb{R}^n$ is written as:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \quad \text{where } v_i \in \mathbb{R}$$

For example, a vector in $\mathbb{R}^3$ could be:

$$\mathbf{v} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$$

Vectors are used in many fields to represent quantities such as force, velocity, or position.

3 Dot Product

The **dot product** (also known as scalar product) of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is defined as the sum of the products of their corresponding components:

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \sum_{i=1}^{n} u_iv_i$$

For example, for $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, the dot product is:

$$\mathbf{u} \cdot \mathbf{v} = 1 \times 4 + 2 \times 5 + 3 \times 6 = 4 + 10 + 18 = 32$$

The dot product has geometric significance: it can be used to find the angle between two vectors.

4 Inner Product

The **inner product** generalizes the dot product to higher dimensions and different vector spaces. In $\mathbb{R}^n$, the inner product is the same as the dot product, but in more abstract vector spaces, the inner product can involve integrals or other operations.

5 Outer Product

The **outer product** of two vectors $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$ is a matrix obtained by multiplying each component of $\mathbf{u}$ by each component of $\mathbf{v}$:

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^T = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix} = \begin{pmatrix} u_1v_1 & u_1v_2 & \dots & u_1v_m \\ u_2v_1 & u_2v_2 & \dots & u_2v_m \\ \vdots & \vdots & \ddots & \vdots \\ u_nv_1 & u_nv_2 & \dots & u_nv_m \end{pmatrix}$$

For example, the outer product of $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 3 & 4 \end{pmatrix}$ is:

$$\mathbf{u} \otimes \mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \times 3 & 1 \times 4 \\ 2 \times 3 & 2 \times 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$$

6 Matrices

A **matrix** is a rectangular array of numbers arranged in rows and columns. A matrix with m rows and n columns is called an $m \times n$ matrix and is written as:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Matrices can represent systems of linear equations, transformations in space, and many other concepts.

7 Vector-Matrix Multiplication

The product of a matrix A of size $m \times n$ and a vector $\mathbf{v}$ of size $n \times 1$ is a new vector $\mathbf{w}$ of size $m \times 1$. The operation is defined as follows:

$$A\mathbf{v} = \mathbf{w}, \quad \text{where} \quad w_i = \sum_{j=1}^{n} a_{ij} v_j$$

For example, for the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and vector $\mathbf{v} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$, we have:

$$A\mathbf{v} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \times 5 + 2 \times 6 \\ 3 \times 5 + 4 \times 6 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}$$

8 Matrix-Matrix Multiplication

The product of two matrices A (size $m \times n$) and B (size $n \times p$) is a new matrix C (size $m \times p$), where the element c_{ij} of matrix C is given by:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

For example, multiplying $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$ results in:

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

9 Matrix Transpose

The **transpose** of a matrix A is a new matrix A^T obtained by swapping rows and columns of A. If A is an $m \times n$ matrix, then A^T will be an $n \times m$ matrix. Formally:

$$(A^T)_{ij} = A_{ji}$$

For example, for $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, the transpose is:

$$A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

10 Matrix Determinant

The **determinant** of a square matrix is a scalar value that provides important information about the matrix, such as whether it is invertible. The determinant of a 2×2 matrix is given by:

$$\det(A) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

For larger matrices, the determinant can be computed using cofactor expansion.

11 Matrix Inverse

The **inverse** of a matrix A, denoted A^{-1}, is the matrix such that:

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

where I is the identity matrix, which has 1s on the diagonal and 0s elsewhere. For a matrix to be invertible, it must be square (i.e., it has the same number of rows and columns), and its determinant must be non-zero.

11.1 Inverse of a 2×2 Matrix

For a 2×2 matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

the inverse of A is given by the formula:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

where $\det(A)$ is the determinant of the matrix A, computed as:

$$\det(A) = ad - bc$$

The matrix A is invertible if and only if $\det(A) \neq 0$. If $\det(A) = 0$, then A does not have an inverse.

Example

Let's compute the inverse of the following matrix:

$$A = \begin{pmatrix} 4 & 7 \\ 2 & 6 \end{pmatrix}$$

First, we calculate the determinant of A:

$$\det(A) = (4)(6) - (7)(2) = 24 - 14 = 10$$

Since the determinant is non-zero, the matrix is invertible. Now we can apply the formula for the inverse:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 6 & -7 \\ -2 & 4 \end{pmatrix}$$

Thus, the inverse of matrix A is:

$$A^{-1} = \begin{pmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{pmatrix}$$

Finally, we can verify that $AA^{-1} = I$ by performing matrix multiplication:

$$AA^{-1} = \begin{pmatrix} 4 & 7 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

This confirms that A^{-1} is indeed the inverse of A.

12 Eigenvalues and Eigenvectors

For a square matrix A, an **eigenvector** is a non-zero vector $\mathbf{v}$ such that:

$$A\mathbf{v} = \lambda\mathbf{v}$$

where λ is a scalar known as the **eigenvalue** associated with the eigenvector $\mathbf{v}$. Eigenvalues and eigenvectors play a central role in many areas of mathematics, physics, and engineering, particularly in the study of linear transformations, stability analysis, and diagonalization of matrices.

12.1 Finding Eigenvalues and Eigenvectors of a 2×2 Matrix

Given a 2×2 matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

we can find its eigenvalues and eigenvectors using the following steps.

Step 1: Find the Eigenvalues

The eigenvalues are found by solving the characteristic equation:

$$\det(A - \lambda I) = 0$$

where I is the identity matrix and λ is the eigenvalue. For a 2×2 matrix, this results in a quadratic equation.
The matrix $A - \lambda I$ is:

$$A - \lambda I = \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix}$$

The determinant of this matrix is:

$$\det(A - \lambda I) = (a - \lambda)(d - \lambda) - bc = 0$$

This is a quadratic equation in λ, which we can solve to find the eigenvalues.

Step 2: Find the Eigenvectors

For each eigenvalue λ, we find the corresponding eigenvector $\mathbf{v}$ by solving the system of linear equations:

$$(A - \lambda I)\mathbf{v} = 0$$

This gives us the eigenvectors associated with each eigenvalue.

Example

Consider the matrix:

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

Step 1: Find the Eigenvalues
The matrix $A - \lambda I$ is:

$$A - \lambda I = \begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix}$$

The determinant is:

$$\det(A - \lambda I) = (4-\lambda)(3-\lambda) - (1)(2) = (12 - 7\lambda + \lambda^2) - 2 = \lambda^2 - 7\lambda + 10$$

We solve the quadratic equation:

$$\lambda^2 - 7\lambda + 10 = 0$$

Using the quadratic formula:

$$\lambda = \frac{-(-7) \pm \sqrt{(-7)^2 - 4(1)(10)}}{2(1)} = \frac{7 \pm \sqrt{49-40}}{2} = \frac{7 \pm \sqrt{9}}{2} = \frac{7 \pm 3}{2}$$

Thus, the eigenvalues are:

$$\lambda_1 = \frac{7+3}{2} = 5, \quad \lambda_2 = \frac{7-3}{2} = 2$$

Step 2: Find the Eigenvectors
Now, we find the eigenvector corresponding to each eigenvalue.
Eigenvector for $\lambda_1 = 5$:
We solve the system:

$$(A - 5I)\mathbf{v} = 0$$

The matrix $A - 5I$ is:

$$A - 5I = \begin{pmatrix} 4-5 & 1 \\ 2 & 3-5 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}$$

We solve the equation:

$$\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives the system of equations:

$$-1v_1 + v_2 = 0 \quad \text{or} \quad v_1 = v_2$$

Thus, the eigenvector corresponding to $\lambda_1 = 5$ is any scalar multiple of:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Eigenvector for $\lambda_2 = 2$:
We solve the system:

$$(A - 2I)\mathbf{v} = 0$$

The matrix $A - 2I$ is:

$$A - 2I = \begin{pmatrix} 4-2 & 1 \\ 2 & 3-2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$$

We solve the equation:

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives the system of equations:

$$2v_1 + v_2 = 0 \quad \text{or} \quad v_2 = -2v_1$$

Thus, the eigenvector corresponding to $\lambda_2 = 2$ is any scalar multiple of:

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Conclusion

The eigenvalues of $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$ are $\lambda_1 = 5$ and $\lambda_2 = 2$, and the corresponding eigenvectors are:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

13 Diagonal Matrix

A **diagonal matrix** is a square matrix in which all the off-diagonal elements are zero, meaning $a_{ij} = 0$ for $i \neq j$. For example, the matrix:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

is a diagonal matrix, where $\lambda_1, \lambda_2, \lambda_3$ are the diagonal elements. Diagonal matrices are of great importance due to their simplicity and the ease with which many operations can be performed.

13.1 Why Diagonal Matrices Are Important

Diagonal matrices have several properties that make them particularly useful in both theoretical and practical applications:

- **Simplicity of operations**: Matrix operations such as matrix multiplication, matrix inversion, and finding powers of a matrix are extremely simple for diagonal matrices. For example, the product of two diagonal matrices is just the product of their corresponding diagonal elements.
- **Diagonalization**: Many matrices can be transformed into a diagonal matrix through a process called diagonalization. If a matrix can be diagonalized, it simplifies the understanding of the linear transformation represented by the matrix, as diagonal matrices are easier to interpret.
- **Linear independence**: The columns of a diagonal matrix corresponding to non-zero diagonal elements represent linearly independent vectors.
- **Eigenvalue decomposition**: Diagonal matrices are closely connected to eigenvalue decomposition, which is a fundamental concept in fields like machine learning, quantum mechanics, and numerical analysis.

13.2 Eigenvalues and Eigenvectors of Diagonal Matrices

Diagonal matrices are particularly simple when it comes to finding eigenvalues and eigenvectors. If a matrix D is diagonal, the eigenvalues are simply the diagonal entries of the matrix, and the eigenvectors are the standard basis vectors.

Eigenvalues of a Diagonal Matrix

Consider a diagonal matrix:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

The eigenvalues of this matrix are directly the diagonal elements: λ_1, λ_2, and λ_3. This is because, for a diagonal matrix D, the characteristic equation $\det(D - \lambda I) = 0$ reduces to solving:

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) = 0$$

Thus, the eigenvalues are simply λ_1, λ_2, and λ_3.

Eigenvectors of a Diagonal Matrix

For each eigenvalue λ_i, the corresponding eigenvector is a standard basis vector. The standard basis vectors in $\mathbb{R}^3$ are:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

For example, for the eigenvalue λ_1, the corresponding eigenvector is $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, for λ_2, the eigenvector is $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and for λ_3, the eigenvector is $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Example

Consider the diagonal matrix:

$$D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

The eigenvalues of D are simply the diagonal elements:

$$\lambda_1 = 3, \quad \lambda_2 = 5, \quad \lambda_3 = 7$$

The corresponding eigenvectors are the standard basis vectors:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

13.3 Diagonalization of a Matrix

A matrix A is said to be **diagonalizable** if there exists an invertible matrix P and a diagonal matrix D such that:

$$A = PDP^{-1}$$

Diagonalizing a matrix allows us to express the matrix as a product of simpler matrices, and it greatly simplifies many matrix operations such as taking powers of a matrix. In particular, many matrices that represent transformations in physics, statistics, and engineering can be diagonalized, making the analysis of these transformations much easier.

14 Higher-Order Matrices and Tensors (optional extra material)

In this final section, we explore the concept of higher-order matrices, introduce tensors, and explain important matrix operations like computing the dimension, rank, and how to use Gauss-Jordan elimination to transform a matrix into row echelon form. This form allows for efficient computation of matrix rank, determinant, and inverses.

14.1 Dimension of a Matrix

The **dimension** of a matrix refers to the number of rows and columns in the matrix. A matrix with m rows and n columns is said to have the dimension $m \times n$. For example, a 3×2 matrix looks like this:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

If a matrix is square (i.e., it has the same number of rows and columns, $m = n$), then it is called a **square matrix**. The dimension of a matrix is important in determining whether certain operations, like multiplication or inversion, are possible.

14.2 Rank of a Matrix

The **rank** of a matrix is the maximum number of linearly independent rows or columns in the matrix. Rank reveals the amount of independent information contained in a matrix, and it plays a central role in determining the solutions to linear systems.
A matrix can have:

- **Full rank**, if the rank is equal to the smaller of the number of rows or columns.
- **Deficient rank**, if the rank is less than the smallest dimension.

The rank of a matrix provides important information about its invertibility:

- A square matrix is invertible if and only if it has full rank.
- The rank also determines how many solutions a system of linear equations has. If the rank is less than the number of variables, the system may have infinitely many solutions or no solutions.

14.3 Gauss-Jordan Elimination and Row Echelon Form

Gauss-Jordan elimination is an algorithm used to transform a matrix into a simpler form known as the **row echelon form** or the **reduced row echelon form** (RREF). This process consists of using elementary row operations to achieve an upper triangular matrix, where all elements below the main diagonal are zeros.
`https://en.wikipedia.org/wiki/Gaussian_elimination`
Steps in Gauss-Jordan Elimination:

1. Identify a pivot (non-zero element) in a column and use it to eliminate all entries below it by subtracting suitable multiples of the row containing the pivot.
2. Repeat this process for each column, moving from left to right, transforming the matrix into an upper triangular form.
3. Optionally, transform the upper triangular matrix into reduced row echelon form by making all the diagonal elements 1 and eliminating entries above each pivot.

Computing the Rank, Determinant, and Inverse Using Row Echelon Form

- **Rank**: The rank of a matrix is the number of non-zero rows in its row echelon form. After applying Gauss-Jordan elimination, simply count the non-zero rows to determine the rank.

- **Determinant**: For a square matrix, the determinant can be calculated as the product of the diagonal elements after transforming the matrix into an upper triangular form. If any diagonal element is zero, the determinant is zero, meaning the matrix is singular and not invertible.

- **Inverse**: Gauss-Jordan elimination can also be used to compute the inverse of a square matrix. By applying the algorithm to the augmented matrix $[A \mid I]$, where A is the matrix to be inverted and I is the identity matrix, the transformation will yield $[I \mid A^{-1}]$ if A is invertible.

Example: Gauss-Jordan Elimination

Consider the following matrix:

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

We perform elementary row operations to transform A into row echelon form:
1. Multiply the first row by $-\frac{1}{2}$ to get a leading 1 in the first column. 2. Use this pivot to eliminate the entries below it by subtracting suitable multiples of the first row from the other rows. 3. Continue this process for the other rows and columns until we get an upper triangular matrix.
After applying Gauss-Jordan elimination, we obtain the row echelon form of A:

$$\text{Row echelon form of } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus, the rank of A is 3, indicating that A has full rank and is invertible.

14.4 Tensors: Generalization of Matrices

Tensors are a generalization of vectors and matrices to higher dimensions. While vectors are 1-dimensional arrays and matrices are 2-dimensional arrays, tensors can have any number of dimensions. For example, a 3-dimensional tensor might represent data in a 3D space, such as a cube of numbers. Tensors are widely used in fields such as physics, machine learning, and data science because they can represent complex, multi-dimensional data.

- **Rank-1 tensor** is a vector (1-dimensional array).
- **Rank-2 tensor** is a matrix (2-dimensional array).
- **Rank-3 and higher tensors** are multi-dimensional arrays.

Tensors are crucial in many modern applications such as deep learning, where data and transformations are often represented as tensors. Tensor operations generalize matrix operations, and tensor decompositions extend the idea of matrix factorizations like the SVD (singular value decomposition) to higher dimensions.

14.5 Conclusion

In this section, we covered higher-order matrices, including the dimension, rank, and how Gauss-Jordan elimination can be used to compute the row echelon form of a matrix. This form is instrumental in determining the rank, determinant, and inverse of a matrix. We also introduced tensors, which generalize the concepts of vectors and matrices to higher dimensions, playing an essential role in modern applications like machine learning and physics.