

Notes for MA 26500 - Linear Algebra I

Shubham Saluja Kumar Agarwal

March 6, 2024

These are lecture notes for spring 2024 MA 26500 at Purdue as taught by professor Siamak Yassemi. Modify, use, and distribute as you please.

Contents

Course Introduction	1
Systems of Equations in Linear Algebra	2
Row Reduction and Echelon Forms	5
Vector Equations	6
Linear Combinations	7
Matrix Equation $\mathbf{Ax} = \mathbf{b}$	8
Solution Sets of Linear Systems	8
Linear Independence	10
Introduction to Linear Transformations	10
Matrix Operations	12
Inverse of a Matrix	14
Subspaces of \mathbb{R}^n	15
Dimension and Rank	16
Determinant	16
Properties of the Determinant	17
Cramer's Rule	17
Vector Space	18
Dimensions of Linear Transformations	20
Eigenvectors and Eigenvalues	20
Diagonalization	21

Course Introduction

This course serves as an introduction to the fundamental concepts and applications of linear algebra, a branch of mathematics that explores vector spaces, linear transformations, and systems of linear equations.

Systems of Equations in Linear Algebra

A linear equation is an equation of the form:

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

Where a_1, a_2, \dots and b are given constants. Note that the exponents of all x terms is 1. Some examples are:

$$2x_1 + 3x_2 = 4$$

$$5x_1 + 6x_2 = 10$$

as opposed to

$$2x_1x_2 + x_3 = 9$$

$$2x_1 + \sqrt{x_2} = 8$$

A system of equations is a collection of one or more linear equations involving the same variable set. An example of this would be:

$$\begin{cases} 3x_1 + 5x_2 + x_3 = 3 \\ 7x_1 - 2x_2 + 4x_3 = 4 \\ -6x_1 + 2x_3 = 2 \end{cases}$$

where the constant coefficient of x_2 in the third equation is 0.

A solution is a list of numbers (s_1, s_2, \dots, s_n) that makes each equation a true statement when we replace $x_1 = s_1, x_2 = s_2, \dots, x_n = s_n$. If we have a system of two linear equations, the solution will be the intersection of the two lines that define the equations on the cartesian plane. The system of equations

$$3x_1 - x_2 = 5$$

$$16x_1 - 2x_2 = 6$$

is mapped to the following graph:



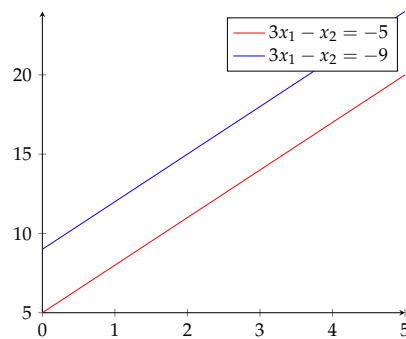
There are three kinds of systems:

1. One solution
2. Infinitely many solutions
3. No solutions

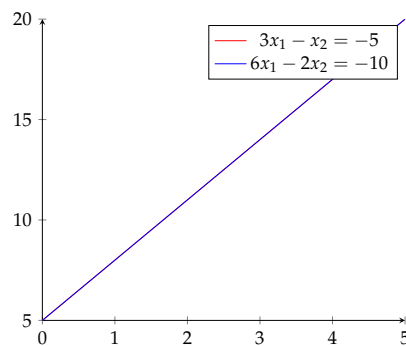
The above example has one solution. This is the intersection of the two lines. This could be further generalized to have far more dimensions as well as far more variables, but such systems are no longer representable in the number of dimensions we can perceive.

However, if the lines were parallel, the other two possibilities surge. If the equations are linearly dependent, there would be infinitely many solutions. However, if the value of b were different for the two, while all the variable coefficients were linearly dependent, it would have none.

The following is an example of no solution:



While this is infinitely many solutions:



Any system of equations has a matrix notation. For example, the above system can be represented as the following.

$$\begin{bmatrix} 3 & -1 & -5 \\ 16 & -2 & 6 \end{bmatrix}$$

Which has the coefficients of the x values as the first n values, and the value of b as the last one. This is called the augmented matrix of the

system of equations. As opposed to:

$$\begin{bmatrix} 3 & -1 \\ 16 & -2 \end{bmatrix}$$

which is the coefficient matrix.

These systems can be solved using the three elementary row operations. This is called row reduction.

The three rules are:

1. Interchange: Exchange the positions of any two rows
2. Multiply: Multiply any row by a constant
3. Addition: Replace the value of a row, with its sum and that of one of the other rows multiplied by a scalar.

By using these three rules, one can reduce it to the following form:

$$\begin{bmatrix} 1 & a_2 & a_3 & a_4 \\ 0 & 1 & b_3 & b_4 \\ 0 & 0 & 1 & c_4 \end{bmatrix}$$

This will result in a trivially solvable unique-solution system of equations. This is also known as a **consistent** system of equations. We can substitute the value of x_3 , which is c_4 , into the equation in the row above, and solve it, as it will now be a single equation with a single variable. This can be done sequentially until all values are found.

Note: If the resulting form is instead

$$\begin{bmatrix} 1 & a_2 & a_3 & a_4 \\ 0 & 1 & b_3 & b_4 \\ 0 & 0 & 0 & c_4 \end{bmatrix}$$

with $c_4 \neq 0$, there would be no solutions, also known as an **inconsistent** system. On the other hand, if $c_4 = 0$, there will be infinitely many solutions, but it will still be considered a **consistent** system. Another point to note is that if the augmented matrices of two systems of equations are equivalent, that is, if you can transform one matrix into the other through elementary row operations, the systems of equations will have the same solutions.

Row Reduction and Echelon Forms

The leading entry of a row refers to the left-most non-zero element in a row. Thus, the leading entries in the following matrix are:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 6 & 9 & 6 \\ 0 & 0 & 8 & 7 \end{bmatrix}$$

A matrix is in reduced echelon form (REF) if it has the following three properties:

1. All non-zero rows are above any rows of all zeros.
2. Each leading entry in a row is in a column to the right of the leading entry in the column above it.
3. All entries below a leading entry are zeros.

If it has the following two properties, it will be a row reduced echelon form (RREF):

4. The leading entry in each non-zero row is 1.
5. Each leading 1 is the only non-zero element in its column.

Example of echelon form:

$$\begin{bmatrix} 1 & 2 & 5 & 8 \\ 0 & 3 & 9 & 6 \\ 0 & 0 & 9 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Example of REF:

$$\begin{bmatrix} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Theorem: Each matrix is equivalent to one and only one RREF.

A pivot position is a location in a matrix that corresponds to a leading 1 in the RREF of a matrix.

A pivot column is a column that contains a pivot position.

Tip: if there is a 1 in the first positions of a row, when the matrix is first "observed", it should be moved to the top of the matrix, as it will most probably make all subsequent computations much easier.

Row reduction algorithm:

1. Begin with the leftmost non-zero position, and make it the first pivot position.
2. Make all positions under the selected pivot position 0.
3. Interchange columns to make sure that the rows with the most leading zeros are closest to the bottom
4. Cover the first column and row, and repeat steps 1 to 3 on the resulting matrix.
5. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot.

The solution of the RREF is the solution of the system of equations that has an augmented matrix with the aforementioned RREF.

Another thing we can note is that all pivot positions refer to basic variables. If there are any non-pivot columns, the variable referring to these columns will be free.

For example a 4×7 matrix is underconstrained and with free variables.

Vector Equations

A matrix with only one column is called a column vector or vector. For example:

$$u = \begin{bmatrix} 3 \\ 8 \\ 12 \end{bmatrix} \text{ or } w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

Two vectors are equal if and only if their corresponding terms are equal. There are some basic operations that can be conducted with them.

Addition between two vectors of the same size:

$$u = \begin{bmatrix} 3 \\ 8 \\ 12 \end{bmatrix} + \begin{bmatrix} 9 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+9 \\ 8+4 \\ 12+5 \end{bmatrix} = \begin{bmatrix} 12 \\ 12 \\ 17 \end{bmatrix}$$

Scalar Multiplication:

$$u = 3 * \begin{bmatrix} 3 \\ 8 \\ 12 \end{bmatrix} = \begin{bmatrix} 3*3 \\ 8*3 \\ 12*3 \end{bmatrix} = \begin{bmatrix} 9 \\ 24 \\ 36 \end{bmatrix}$$

Note: these operations both have graphical representations. Addition is addition of two vectors, as is done by the parallelogram method,

and scalar multiplication is equivalent to proportionally scaling a vector.

A line can be defined as the set of all scalar multiples of a vector.

Algebraic properties of a vector in \mathbb{R}^n :

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
3. $\mathbf{u} + \mathbf{0} = \mathbf{u}$
4. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
7. $c(d\mathbf{u}) = cd(\mathbf{u})$
8. $1\mathbf{u} = \mathbf{u}$

Linear Combinations

A vector is a linear combination of other vectors if:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 \dots c_n\mathbf{u}_n = \mathbf{v}$$

For any set of vectors, there is an infinite amount of linear combinations, as the values of the constants they are multiplied with are not restrained by anything, and they could be a part of any domain, be it \mathbb{R} , \mathbb{I} , or \mathbb{C} .

To find if a vector is a linear combination of other matrices, we can create an augmented matrix with the desired vector as the constant column, and all relevant vectors, as the elements of the coefficient matrix part of the augmented matrix.

Should there be a free variable in the resolution of the augmented matrix, there will be infinitely many linear combinations that fulfill the condition we desire.

The $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n)$ is the set of all linear combinations of these vectors. So, if asked whether a vector is in the span of other vectors, it is equivalent to asking if there exists a linear combination of vectors that results in the desired vector.

Matrix Equation $\mathbf{Ax} = \mathbf{b}$

Let \mathbf{A} be an $m \times n$ matrix. The product of \mathbf{A} and \mathbf{x} will be the linear combination of the columns of \mathbf{A} using the entries of \mathbf{x} as weights.

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n$$

Any linear combination can be written in the form \mathbf{Ax} , and, if it is a system of linear equations, it can be written as $\mathbf{Ax} = \mathbf{b}$.

The matrix equation $\mathbf{Ax} = \mathbf{b}$ has the same solution as $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n = \mathbf{b}$ which in turn has the same solution set as the augmented matrix $\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}$.

In essence, what we are searching for is the value of \mathbf{x} .

Existence of solutions:

The equation $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if \mathbf{b} is a linear combination of \mathbf{A} .

The following statements are equivalent:

- For each \mathbf{b} in \mathbb{R}^n the equation $\mathbf{Ax} = \mathbf{b}$ has a solution
- The equation $\mathbf{Ax} = \mathbf{b}$ has a solution
- Each \mathbf{b} in \mathbb{R}^n is a linear combination of the columns of \mathbf{A}
- \mathbf{A} has a pivot position in every row.

Row-Vector Rule for Computing \mathbf{Ax} :

The valid product of two matrices is the sum of the linear combination of each row with their respective terms in the \mathbf{x} matrix.

$$\begin{bmatrix} 1 & 3 & 4 & 8 \\ 5 & 1 & 4 & 6 \\ 6 & 9 & 1 & 7 \\ 0 & 7 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1x_1 + 3x_2 + 4x_3 + 8x_4 \\ 5x_1 + 1x_2 + 4x_3 + 6x_4 \\ 6x_1 + 9x_2 + 1x_3 + 7x_4 \\ 0x_1 + 7x_2 + 4x_3 + 3x_4 \end{bmatrix}$$

If \mathbf{A} is an $m \times n$ matrix:

- $\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{Au} + \mathbf{Av}$
- $\mathbf{A}(c\mathbf{u}) = c(\mathbf{Au})$

Solution Sets of Linear Systems

If a system can be written as $\mathbf{Ax} = \mathbf{0}$, it is homogenous. This will always have the trivial solution of $\mathbf{x} = \mathbf{0}$.

Thus, there is either a unique solution or infinitely many solutions. That is, if there is a non-zero solution, there are infinitely many solutions.

This happens because if \mathbf{x} is a solution, then $c\mathbf{x}$ will also be a solution. The trivial solution will also need to exist, as anything multiplied by 0 is 0.

So, the homogenous equation $\mathbf{Ax} = 0$, has a non-trivial solution if and only if there is a free variable.

Suppose $\mathbf{Ax} = \mathbf{b}$ is consistent with solution \mathbf{p} , the solution set of the system will be the set of all the vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}$ where \mathbf{v} is the solution set of the homogenous equation.

3 planes containing the origin could have three kinds of intersections:

1. Point: origin — 0 free variables
2. Line: through origin — 1 free variable
3. Plane: same plane through origin — 2 free variables

This can allow us to express the solution set in terms of vectors scalarly multiplied by the free variables, should there be any.

Let us say there are two free variables. All the other variables can be expressed in terms of these free variables.

Let us take the following example:

$$\begin{bmatrix} 1 & 2 & 0 & -3 \\ 3 & 6 & 0 & -9 \end{bmatrix} = 0$$

Has three free variables, and is equivalent to $\begin{bmatrix} 1 & 2 & 0 & -3 \end{bmatrix} = 0$ with only x_1 being non-free.

Thus, the solution set can be represented as:

$$\mathbf{x} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

This can be done for any \mathbb{R}^n homogenous matrix.

1. Make each free variable 1 for its corresponding x_n
2. Go to each row, and solve the 1 variable equation, with free variable we are solving for being 1.
3. Write the aforementioned solution in the corresponding position of the multiplied matrix to x_n

Linear Independence

A set of vectors is linearly independent if the vector equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution. It is linearly dependent if there exist solutions distinct to the trivial solution that fulfill the aforementioned property.

The columns of matrix \mathbf{A} are linearly independent if and only if $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

If every column of \mathbf{A} has a pivot position for $\mathbf{Ax} = \mathbf{b}$, there is only one solution, and thus it is linearly independent.

So,

- One vector is always linearly independent as long as it is non-zero.
- Two vectors are linearly dependent if and only if one vector is a scalar multiple of the other.
- For three or more vectors, if a vector is a linear combination of the others, it is linearly dependent.

If a set has more vectors than entries in each vector, then the set is linearly dependent. For example:

$$\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right\}$$

Note: If \mathbf{A} is an $n \times p$ matrix, it has n rows and p unknowns.

If a set contains the $\mathbf{0}$ vector, it is linearly dependent.

Introduction to Linear Transformations

A transformation or function T from \mathbb{R}^n to \mathbb{R}^m is a rule that takes each vector in \mathbb{R}^n and assigns to it a vector in \mathbb{R}^m .

The set \mathbb{R}^n is called the domain, and the set \mathbb{R}^m is called the codomain of T .

The notation is:

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

so $\mathbf{x} \in \mathbb{R}^n$ and $T(\mathbf{x}) \in \mathbb{R}^m$.

A Matrix Transformation consists of a matrix $m \times n$ can define a transformation from \mathbb{R}^n to \mathbb{R}^m . That is:

$$\mathbf{x} \mapsto \mathbf{A}_{n \rightarrow m} \mathbf{x}$$

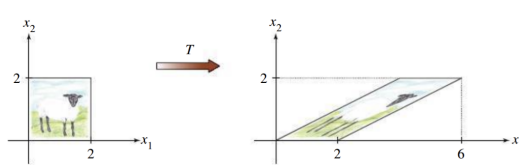
or

$$T(\mathbf{x}) = \mathbf{Ax}$$

In essence, a matrix transformation is the multiplication of the matrix \mathbf{x} with a matrix \mathbf{A} that results in the final transformed matrix that we desire.

So, when asked about the existence of $T(\mathbf{x}) = \mathbf{b}$, all that is necessary is the creation and resolution of the augmented matrix of \mathbf{A} and \mathbf{b} .

This allows for a multitude of geometric transformations such as:



A linear transformation has the following properties:

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
- $T(c\mathbf{u}) = cT(\mathbf{u})$
- $T(\mathbf{0}) = \mathbf{0}$ — which is quite useful for an initial verification of linear transformations
- $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$

Note: all matrix transformations are linear transformations, but not all linear transformations are matrix transformations.

If a matrix is in the range of T , it is a linear transformation of the columns of the matrix \mathbf{A} that defines $T(\mathbf{x})$.

Let T be a linear transformation. There exists a matrix such that $T(\mathbf{x}) = \mathbf{Ax}$. So, we will now work on finding \mathbf{A} .

We know that:

$$T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = T \left(x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$

For any linear transformation there exists a unique matrix \mathbf{A} , called the standard matrix of T , that, in most cases, can be found using the above property.

One just needs to find the values of the transformation of all the vectors of the form:

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

and each transformation of each vector will be its respective column of the matrix \mathbf{A} .

A transformation is one-to-one if for each point in \mathbb{R}^m , that is, the range, every point has at most one vector in \mathbb{R}^n that transforms into it.

A transformation is onto if for each point in \mathbb{R}^n , that is, the domain, every point has a transformation.

- A linear transformation T is one-to-one if and only if $T(x) = 0$ has just the trivial solution.
- T is one-to-one \iff the columns are linearly independent.
- T maps \mathbb{R}^n onto $\mathbb{R}^m \iff$ the columns of \mathbf{A} span \mathbb{R}^m .
- T is onto \iff there is a pivot position in every row.
- If $n > m$, the transformation cannot be one-to-one.
- If $m > n$, the transformation cannot be onto.
- If $m = n$, T is onto $\iff T$ is one-to-one.

Matrix Operations

- Sum:

$$u = \begin{bmatrix} 3 \\ 8 \\ 12 \end{bmatrix} + \begin{bmatrix} 9 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+9 \\ 8+4 \\ 12+5 \end{bmatrix} = \begin{bmatrix} 12 \\ 12 \\ 17 \end{bmatrix}$$

Requires the matrices to be the same size.

- Scalar multiplication:

$$u = 3 * \begin{bmatrix} 3 \\ 8 \\ 12 \end{bmatrix} = \begin{bmatrix} 3*3 \\ 8*3 \\ 12*3 \end{bmatrix} = \begin{bmatrix} 9 \\ 24 \\ 36 \end{bmatrix}$$

If $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are matrices of the same size:

- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- $\mathbf{A} + \mathbf{0} = \mathbf{A}$
- $r(\mathbf{A} + \mathbf{B}) = r\mathbf{A} + r\mathbf{B}$
- $(r + s)\mathbf{A} = r\mathbf{A} + s\mathbf{A}$
- $r(s\mathbf{A}) = s(r\mathbf{A})$

- Matrix Multiplication:

Let \mathbf{A} be an $m \times n$ matrix and \mathbf{B} be an $n \times p$ matrix, and \mathbf{B} has the columns $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$. \mathbf{AB} is defined as the matrix with columns $\mathbf{Ab}_1, \mathbf{Ab}_2, \dots, \mathbf{Ab}_n$.

Note: the number of columns of A , is the same as the number of rows of B .

$$\begin{bmatrix} 3 & 4 \\ 8 & 5 \\ 12 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3(2) + 4(1) & 3(1) + 4(3) \\ 8(2) + 5(1) & 8(1) + 5(3) \\ 12(2) + 1(1) & 12(1) + 1(3) \end{bmatrix} = \begin{bmatrix} 10 & 15 \\ 21 & 23 \\ 25 & 15 \end{bmatrix}$$

Let A_{ij} be the entry in row i and column j of matrix A .

$$(AB)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + A_{i3}B_{3j} + \cdots + A_{in}B_{nj}$$

Matrix multiplication has the following properties:

- $A(BC) = (AB)C$
- $A(B + C) = AB + AC$
- $(B + C)A = BA + CA$
- $r(AB) = (rA)B = A(rB)$
- $I_m A = A = A I_m$

with I_m being the identity matrix.

Note: $AB \neq BA$ unless $B = I_n \vee B = 0$

Note: $AB = 0 \not\Rightarrow A = 0 \vee B = 0$

Powers of a matrix:

If A is an $n \times n$ matrix, then A^k is the matrix that results from:

$$A^k = \underbrace{A \times A \times \cdots \times A}_{k \text{ times}}$$

- Transpose of a Matrix:

Given an $m \times n$ matrix, the transpose of this matrix is the $n \times m$ matrix, denoted by A^T , whose columns are formed by the corresponding rows of A .

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}^T = \begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix}$$

It has the following properties:

- $(A^T)^T = A$
- $A + B^T = A^T + B^T$
- $(rA)^T = rA^T$
- $(AB)^T = A^T B^T$

Let A be an $n \times n$ matrix. This means A is a square matrix.

If this A is of the form:

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$$

it is an upper triangular matrix.

If this \mathbf{A} is of the form:

$$\begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$$

it is a lower triangular matrix.

If this \mathbf{A} is of the form:

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

it is a diagonal matrix.

If this \mathbf{A} is of the form:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

it is a zero matrix.

Inverse of a Matrix

An $n \times n$ matrix \mathbf{A} is said to be invertible, if there is an $n \times n$ matrix \mathbf{C} such that $\mathbf{CA} = \mathbf{AC} = \mathbf{I}_n$. If it is not invertible, it is singular. Thus, if it is invertible, it is also non-singular.

The inverse of \mathbf{A} is denoted by \mathbf{A}^{-1} .

Let $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, \mathbf{A} is invertible and $\mathbf{A}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Note: $ad - bc$ is called the determinant of the matrix.

If \mathbf{A} is invertible, for any \mathbf{b} in \mathbb{R}^n , the equation $\mathbf{Ax} = \mathbf{b}$ has the unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

This has a few additional properties:

- If \mathbf{A} is invertible, then \mathbf{A}^{-1} is also invertible, and $\mathbf{A} = (\mathbf{A}^{-1})^{-1}$
- If \mathbf{A} and \mathbf{B} are both invertible, then \mathbf{AB} is invertible and $(\mathbf{AB})^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}$.
- If \mathbf{A} is invertible, then \mathbf{A}^T is invertible and $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

An elementary matrix is defined as a matrix that can be obtained from performing a single elementary operation on the identity matrix.

If an elementary operation is performed on an $n \times n$ matrix \mathbf{A} , the resulting matrix can be represented by \mathbf{EA} , where \mathbf{E} is the result of the same operation applied to the identity matrix.

This will allow us to define any matrix that results from elementary operations of another as a multiplication of elementary matrices with the original matrix.

A matrix \mathbf{A} is invertible $\iff \mathbf{A}$ is row equivalent to \mathbf{I}_n .

The algorithm to find \mathbf{A}^{-1} for any matrix is the following:

1. Write down the matrix $[\mathbf{A}|\mathbf{I}_n]$.
2. Row reduce \mathbf{A} until it is equal to \mathbf{I}_n .
3. Perform the same operations to \mathbf{I}_n .
4. The result of what was originally \mathbf{I}_n will be \mathbf{A}^{-1} .

If we have a square matrix \mathbf{A} of size $n \times n$:

- \mathbf{A} is invertible
- \mathbf{A} is row equivalent to \mathbf{I}_n
- \mathbf{A} has n pivot positions
- $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution
- Columns of \mathbf{A} are linearly independent
- The linear transformation $\mathbf{x} \mapsto \mathbf{Ax}$ is one-to-one
- The equation $\mathbf{Ax} = \mathbf{b}$ has at least one solution for each \mathbf{b} in \mathbb{R}^n
- The columns of \mathbf{A} span \mathbb{R}^n
- The linear transformation $\mathbf{x} \mapsto \mathbf{Ax}$ maps \mathbb{R}^n to \mathbb{R}^n
- There is an $n \times n$ matrix \mathbf{C} such that $\mathbf{CA} = \mathbf{I}$
- There is an $n \times n$ matrix \mathbf{D} such that $\mathbf{AD} = \mathbf{I}$
- \mathbf{A}^T is an invertible matrix

Subspaces of \mathbb{R}^n

A subspace of \mathbb{R}^n is any set H in \mathbb{R}^n that has three properties:

1. The zero vector is in H
2. For each \mathbf{u} and \mathbf{v} in H , $\mathbf{u} + \mathbf{v}$ is in H .
3. For each \mathbf{u} in H , $c\mathbf{u}$ is in H , for all c .

Let us now look at two important subspaces of \mathbb{R}^n :

- Column Space: the span of three vectors $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$, is called $\text{Col}(\mathbf{A})$, and is a subspace of \mathbb{R}^n .
- Null Space: the set $\text{Nul}(\mathbf{A})$ of all solutions $\mathbf{Ax} = \mathbf{0}$ is called the null space. It is also a subspace of \mathbb{R}^n .

A basis for a subspace H of \mathbb{R}^n is the following:

$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for H

$\iff (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \text{ are linearly independent}) \wedge (\text{span}(\mathbf{V}) = \mathbb{R}^n)$

So, the basis of $\text{Col}(\mathbf{A})$ is the (**pivot columns of \mathbf{A}**).

The basis of $\text{Nul}(\mathbf{A})$ is the span of the vectors resulting from the free variable representation of the homogenous solution. For example:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

and the basis would be:

$$\text{Nul}(\mathbf{A}) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Dimension and Rank

Coordinate vectors: Let the set $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ be a basis of a subspace H . For each \mathbf{x} in H , the coordinates of \mathbf{x} relative to the basis B , are c_1, c_2, \dots, c_n such that the vector is $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \dots + c_n\mathbf{b}_n$.

Dimension of a subspace: The number of vectors in any subspace H . The dimension of the $\mathbf{0}$ subspace is defined as zero.

Rank of a matrix: the dimension of the column space of said matrix. It is, in essence the number of pivot columns of the matrix.

Nullity of a matrix: the dimension of the null space of said matrix. It is, in essence the number of free variables in the matrix.

If a matrix \mathbf{A} has n columns: $\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n$.

Any set of elements of H that spans H is a basis for H .

If a matrix \mathbf{A} is of dimensions $m \times n$, then $\text{rank}(\mathbf{A}) \leq \min(m, n)$

Determinant

Determinant: The determinant of an $n \times n$ matrix \mathbf{A} , $\det(\mathbf{A})$, is based on the elements of the matrix.

If $\mathbf{A} = \begin{bmatrix} a \end{bmatrix}$, the determinant will be a .

If $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant will be $ad - bc$.

The next can also be defined, but it is going to be very long.

So the definition is actually:

Let $\mathbf{A} = [a_{ij}]_{n \times n}$.

The cofactor C_{ij} of \mathbf{A} is $(-1)^{i+j} \det(A_{ij})$, where A_{ij} is the submatrix formed by deleting the i th row and j th column. So,

$$\det(\mathbf{A}) = \sum_{k=1}^n a_{kj} C_{kj} = \sum_{k=1}^n a_{ik} C_{ik}, \text{ for any } i, j \leq n$$

Note: the determinant will be the same regardless of what column or row you choose to cofactor through.

Properties of the Determinant

There are three Elementary Row Operations: swap, add, multiply.

- If you swap, the determinant of the new matrix will be $-1 \times$ (old determinant).
- If you add, the determinant will not change.
- If you multiply a row by k , the determinant will also be multiplied by k .

$$\text{Let } \det(\mathbf{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = D_A.$$

$$\begin{aligned} \Rightarrow \begin{vmatrix} c & d \\ b & a \end{vmatrix} &= -D_A \\ \begin{vmatrix} a & b \\ c+ka & d+kb \end{vmatrix} &= D_A \\ \begin{vmatrix} ka & kb \\ ca & db \end{vmatrix} &= kD_A \end{aligned}$$

Another property is $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

These properties lead to the following theorems:

If \mathbf{A} is invertible, $\det(\mathbf{A}) \neq 0$.

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$$

$$\det \mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}$$

$$\det k\mathbf{A} = k^n \det(\mathbf{A})$$

Cramer's Rule

For an $n \times n$ matrix \mathbf{A} , and any \mathbf{b} in \mathbb{R}^n .

Let $\mathbf{A}_i(\mathbf{b})$ be the matrix resulting from replacing column i with \mathbf{b} .

Cramer's Rule states that for an invertible matrix \mathbf{A} , $\mathbf{Ax} = \mathbf{b}$ has entries given by:

$$x_i = \frac{\det(\mathbf{A}_i(\mathbf{b}))}{\det(\mathbf{A})}$$

Adjoint of a matrix: The adjoint of an invertible matrix is the transpose of the matrix formed by each of the cofactors.

This leads to the following property:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A})$$

This method allows us to easily find singular terms of the inverted matrix. For example, if one was asked to find the term b_{ij} from the inverted matrix, we can calculate it through $\frac{c_{ji}}{\det(A)}$.

Another clever theorem is if we have an $n \times n$ matrix, the hyper-volume of this hyper-parallelogram defined by the n vectors is $|\det(A)|$.

And another:

If we have an area S

$$\text{area}(T(S)) = \text{area}(S) * \det(\mathbf{A})$$

with \mathbf{A} being the transform matrix of the transformation.

Vector Space

Vector Space: a non-empty set V of vectors on which two operations called addition and multiplication by scalar are defined. This is subject to the ten axioms:

If $\mathbf{u}, \mathbf{v} \in V$, and $c, d \in \mathbb{R}$

1. $\mathbf{u} + \mathbf{v} \in V$
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{u} = \mathbf{u} + (\mathbf{v} + \mathbf{u})$
4. $\mathbf{0} + \mathbf{u} = \mathbf{u}$
5. $\forall \mathbf{u} \in V, -\mathbf{u} \in V$
6. $c\mathbf{u} \in V$
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
9. $c(d\mathbf{u}) = cd(\mathbf{u})$
10. $1\mathbf{u} = \mathbf{u}$

Let $\mathbf{M}_{m \times n}$ be a matrix with real entries.

Addition is the sum of each term with their respective term in another matrix of the same dimensions.

Scalar multiplication is the multiplication of each term with the given scalar.

Subspace: A subspace H of a vector space V is a subset that complies with three properties:

1. $\mathbf{0} \in H$
2. $\mathbf{u}, \mathbf{v} \in H \implies \mathbf{u} + \mathbf{v} \in H$
3. $c \in \mathbb{R}, \mathbf{u} \in H \implies c\mathbf{u} \in H$

This is, in essence, the generalized form of "Subspaces of \mathbb{R}^n ".

Theorem. If V is a vector space, and $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$ are vectors within V , then $H = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n)$ is a subspace of V .

Theorem. If \mathbf{A} is an $m \times n$ matrix, the null space of \mathbf{A} will be a subspace of \mathbb{R}^n .

So, if were to be given a system of m equations and n variables (with no constant terms), by passing all variables to the left side of their respective equations, and turning the result into a matrix. The Null of the resulting matrix will be a subspace.

Theorem. If \mathbf{A} is an $m \times n$ matrix, $\text{Col}(\mathbf{A})$ will be a subspace of \mathbb{R}^m .

Linear Transformation from vector space V to W : a rule that assigns to each vector $\mathbf{v} \in V$ a unique vector $\mathbf{w} = T(\mathbf{v}) \in W$. It must also have the properties:

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
- $T(c\mathbf{u}) = cT(\mathbf{u})$

An example of a linear transformation is $\mathbb{P}_2 \mapsto \mathbb{R}^2$. This means that we need to find a standard basis, like we did for the matrix transformations. We will find that the standard matrix is in fact $\{1, x, x^2\}$. So, to find the transformation matrix, we need to find the values of the basis at the given condition.

Kernel: set of all \mathbf{v} in V , such that $T(\mathbf{v}) = \mathbf{0}$

Range: set of all $\{\mathbf{w} \in W \mid \exists \mathbf{v} \in V \ \& \ T(\mathbf{v}) = \mathbf{w}\}$

Theorem. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$ be vectors in a vector space V . They are linearly independent if and only if the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n = \mathbf{0}$$

has only the trivial solution.

Basis of V : set of vectors such that they are linearly independent and span V . Must have the exact same amount of elements as unknowns. That is \mathbb{P}_n will have exactly $n + 1$ elements in its basis. On the other hand $\mathbf{M}_{n \times m}$ will have mn .

The standard basis, on the other hand, is the value that is used to find the transformation matrix. For polynomials, it is:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

with $n + 1$ terms in each vector.

For the matrix, it will be the set of all unique matrices with a 1 in a position within it, and 0 in every other position.

Theorem. The columns of an invertible matrix of side n will be a basis for \mathbb{R}^n .

Theorem. If there is a set of vectors that spans a vector space, all vectors that are linear combinations of others can be removed (sequentially), and what remains will be a basis for said vector space.

Dimensions of Linear Transformations

We already know the definitions of

Eigenvectors and Eigenvalues

Eigenvector: for an $n \times n$ matrix \mathbf{A} , is a non-zero vector such that $\mathbf{Ax} = \lambda\mathbf{x}$ for some scalar λ .

Eigenvalue: scalar value mentioned in the above definition: λ .

The steps to finding these are:

1. Find all values λ such that $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. These are your eigenvalues.
2. For each λ , find the Null Space of $\mathbf{A} - \lambda\mathbf{I}$, that is, find the non-zero vector such that $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$. Each of these will be your eigenvectors.

Eigenspace: for an $n \times n$ matrix \mathbf{A} , is the subspace of \mathbb{R}^n that contains all the eigenvectors of \mathbf{A} and $\mathbf{0}$. It is denoted by E_λ .

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

Theorem. The eigenvalues of a triangular matrix are its main diagonal.

Theorem. 0 is an eigenvalue if and only if the determinant of the matrix is already 0

Theorem. If all eigenvalues are distinct, the corresponding eigenvectors will be linearly independent.

Characteristic equation: $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$

The characteristic equation is an n degree polynomial with n roots, taking into account the multiplicities of the roots themselves.

Algebraic Multiplicity: Number of times a root appears in the characteristic equation.

Geometric Multiplicity: Dimension of the eigenspace corresponding to an eigenvalue.

Similarity: Two matrices \mathbf{A}, \mathbf{B} are similar if $\exists \mathbf{P} | (\mathbf{P}^{-1} \mathbf{A} \mathbf{P} = \mathbf{B}) \vee (\mathbf{P} \mathbf{B} \mathbf{P}^{-1} = \mathbf{A})$

Theorem. Similar matrices have the same characteristic equation and the same eigenvalues with the same multiplicities.

Note: Row equivalence is different from similarity. Matrix operations modify eigenvectors and eigenvalues.

Diagonalization

Theorem. $\mathbf{A} = \mathbf{Q} \mathbf{B} \mathbf{Q}^{-1} \implies \mathbf{A}^n = \mathbf{Q} \mathbf{B}^n \mathbf{Q}^{-1}$

A triangular matrix will have eigenvalues equal to the values on the diagonal, and any matrix similar to it will have the same eigenvalues.

Theorem. For a diagonal matrix \mathbf{D} , \mathbf{D}^k will be the same as the original matrix with each of the terms raised to the k^{th} power.

Theorem. $\mathbf{A} = \mathbf{P} \mathbf{D} \mathbf{P}^{-1} \implies \mathbf{A} \mathbf{P} = \mathbf{P} \mathbf{D}$

Theorem. \mathbf{P} is formed by all eigenvectors, while \mathbf{D} is formed by all eigenvalues.