Properties of Matrices

Rose-Hulman Institute of Technology

David Sachmanyan

March 9, 2026

This will not teach you everything you need to know about matrices or even teach you how to do some of the topics mentioned. This of this more as a reference sheet to familiarize yourself with the operations and theorems of matrices. You probably won't be able to solve many problems after only reading this, you will need practice.

Contents

1. The Matrix Object 2

2. The Determinant, Transpose, and Trace 3

3. Row Reduction and Echelon Form 4

4. Analyzing Systems of Linear Equations 5

5. The Inverse of a Matrix 7

6. Eigenvalues and Eigenvectors 9

7. Fundamental Subspaces 11

1. The Matrix Object

Similar to other mathematical objects such as numbers (for instance, real numbers) and n-th dimensional vectors denoted as $\mathbb{R}$ and $\mathbb{R}^n$ respectfully. For example, if we wish to a is a real number, we would write $a \in \mathbb{R}$ and if $\boldsymbol{v}$ is some n-dimensional real vector, we write $\boldsymbol{v} \in \mathbb{R}^n$. If we have some real $m \times n$ real matrix where m is the number of rows and n is the number of columns, we can say $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. In textbooks or other sources, people often use $A \in \mathbb{R}^{m\times n}$.

Definition 1.1 (Matrix)**.** Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ where A is defined via

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{mn} \end{pmatrix}. \tag{1.1}$$

In a sense, matrices are just a collection of numbers which we can index at row i and column j via $(A)_{ij} = a_{ij}$. Similar to real numbers and vectors, matrices are equipped with standard operations such as addition and multiplication. If we have the matrices $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{R})$, we define $C = A + B$ as

$$c_{ij} = a_{ij} + b_{ij} \tag{1.2}$$

which is "element-wise" addition. For example,

$$A + B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}. \tag{1.3}$$

As you might be able to notice, for addition between two matrices to be defined, they must be of the same dimension, otherwise, they are impossible to add.

As aforementioned, we can also multiply two matrices together. However, there are more restrictions when it comes to multiplication. We define the multiplication between two matrices AB as

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{jk}. \tag{1.4}$$

You may also be able to notice that this operation is similar to taking the dot product (or more abstractly, inner product) between each row of A against each column of B. So, for multiplication to be possible, the number of rows of A must be equal to the columns of B.

Example 1.1 (Matrix multiplication)**.** Consider the following basic example

$$\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1(4) + 3(2) & 1(5) + 3(1) \\ 0(4) + 2(4) & 0(5) + 2(1) \end{pmatrix} = \begin{pmatrix} 10 & 8 \\ 8 & 2 \end{pmatrix}. \tag{1.5}$$

To properly continue our study of matrices, we need to discuss a certain type of matrix. In all vector spaces (you do not need to know what this means), there is always an identity element of scalar multiplication. This is some element which you can multiply another element to get the same element. For instance, the identity element in $\mathbb{R}$ is 1, for $\mathbb{R}^n$ we have a vector of 1s. For $\mathcal{M}_{m\times n}(\mathbb{R})$, we have the *identity matrix*.

Definition 1.2 (Identity Matrix)**.** We define the identity matrix $I_n \in \mathcal{M}_{n\times n}(\mathbb{R})$ via

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix};\ (I)_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}. \tag{1.6}$$

So for any $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $AI = A$. We can simply describe the identity matrix as a matrix with 1 along the diagonal and 0 everywhere else. This type of matrix is extremely important.

2. The Determinant, Transpose, and Trace

In the same way that real numbers come with special operations like absolute value and vectors are defined with magnitude and cross product, matrices also have some special functions and operations we can perform.

Definition 2.1 (Determinant)**.** Right now, the determinant will not have any meaningful use or interpretation other than a computational tool. For a 2×2 real matrix, we define its determinant to be

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc. \tag{2.1}$$

However, as the dimension of our matrix increases, taking the determinant becomes significantly more work. Feel free to memorize other methods of determinants, but a method that works for taking determinants for matrices of all dimensions if *cofactor expansion*. For some $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, we pick some $(A)_{ij}$ and "remove" the row and column containing that element. The determinant of this new matrix. $M \in \mathcal{M}_{n-1\times n-1}(\mathbb{R})$, is called the *minor determinant*. We also define the *cofactor*

$$(C)_{ij} = (-1)^{i+j}(M)_{ij} \tag{2.2}$$

which we multiply my the minor determinant. So basically,

$$\det(A) = \sum (C)_{ij} \det(M)_{ij}. \tag{2.3}$$

This can be complicated without a concrete example so don't worry if you do not understand this just yet. You will usually never have to manually compute large determinants in any course. If you need to, hopefully you can use computer algebra systems like *Maple*.

Proposition 2.1. *If there are any columns or rows of full zeroes, the determinant of the matrix is zero via cofactor expansion. You simply select all the zeroes as the cofactors.*

Remark 2.1. Note that the determinant is only defined for square matrices, meaning $m = n$.

Proposition 2.2. *Performing row operations on a matrix does not change its determinant.*

The next operation is quite silly in the sense here we quite literally flip the rows and columns of a matrix, this is called the *transpose* of a matrix.

Definition 2.2 (Transpose)**.** Given some matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, we write the transpose of A as A^T. If the columns of a matrix as spanned via the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^m$, we can write

$$A = \begin{pmatrix} | & | & & | \\ \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \\ | & | & & | \end{pmatrix},\ A^\mathsf{T} = \begin{pmatrix} — & \boldsymbol{v}_1 & — \\ — & \boldsymbol{v}_2 & — \\ & \vdots & \\ — & \boldsymbol{v}_n & — \end{pmatrix}. \tag{2.4}$$

Simply put, each row becomes a column. The transpose is a critical operation and will come into play later when we talk about the fundamental matrix subspaces.

Proposition 2.3. *For some $n \times n$ matrix A,* $\det(A) = \det(A^\mathsf{T})$.

Example 2.1 (Transpose of a matrix)**.** Consider the small example below:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},\ A^\mathsf{T} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}. \tag{2.5}$$

It is as simple as just swapping the rows and columns.

Definition 2.3 (Symmetric Matrix)**.** An $n \times n$ matrix is said to be symmetric if $A^T = A$.

Definition 2.4 (Trace)**.** We define the trace of a matrix as the sum of its diagonal elements. So for some square matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$,

$$\operatorname{tr}(A) = \sum_{i=1}^{n} (A)_{ii}. \tag{2.6}$$

Rather than having some deep geometric or even spiritual meaning, the trace of a matrix is simply a computational tool, as far as I know, it does not have any deeper mathematical meaning.

3. Row Reduction and Echelon Form

In the same way expressions or logical statements can be equivalent but not equal, so can matrices. Two matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$ can be *row equivalent*, $A \sim B$. If you are reading this you have most likely already done this at some point, but we can perform three basic elementary row operations on a matrix while maintaining row equivalence:

1. Swapping two rows.
2. Multiplying a row by a non-zero scalar.
3. Adding a scalar multiple of one row to another.

These row operations are critical in performing what is called *Gaussian elimination* to achieve *row echelon form*.

Definition 3.1 (Row Echelon Form)**.** The row echelon form of some matrix A is written as $\mathrm{REF}(A)$. A matrix is considered to be in this form if all rows having zero entries are at the bottom and the leading entry in a row, called the *pivot*, is to the right of the leading entry in the row above. I have boxed the pivots in the following examples.

Remark 3.1. The word echelon comes from the French word chelon which basically means rung or ladder which is why we call matrices of that appearance row echelon form.

Example 3.1. For example, consider the following matrix which we will put into row echelon form

$$\begin{pmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 0 & 2 & 4 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 0 & 2 & 4 \end{pmatrix} \xrightarrow{-2R_1 + R_2 \to R_2} \begin{pmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \\ 0 & 2 & 4 \end{pmatrix} \xrightarrow{R_2 + R_3 \to R_3} \begin{pmatrix} \boxed{1} & 3 & 5 \\ 0 & \boxed{-2} & -4 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.1}$$

Now, we could go further and perform row operations on our matrix until we get as close to possible to the identity matrix. We introduce what is called *row reduced echelon form*.

Definition 3.2 (Row Reduced Echelon Form)**.** Any matrix A is considered to be in row reduced echelon form, $\mathrm{RREF}(A)$, if it is in row echelon form, the leading entry in each non-zero row is 1, and each column containing a pivot has zeroes in all entries above it.

Example 3.2. Continuing with our row reduced matrix in (3.1), we have

$$\begin{pmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{-\frac{1}{2}R_2 \to R_2} \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{-3R_2 + R_1 \to R_1} \begin{pmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.2}$$

If you can notice, part of our matrix seems to contain the identity matrix. While this is not always the case, we can perform row operations on some matrices to make them row equivalent to the identity matrix. We will discuss this idea later when we touch on invertability of a matrix.

This notion of echelon form seems silly, but this idea naturally leads us into the next topic of actually applying matrices to problems such as solving systems of linear equations.

4. Analyzing Systems of Linear Equations

As you might already know, matrices are a very important aspect of linear algebra and a good portion of linear algebra is devoted to the study of linear systems. We will formalize the idea of pivots in a matrix after row reduction as the *rank* of a matrix. However, there are multiple types of "rank". For instance, the rank of the set of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ is the number of vectors left in that set after removing all *linearly dependent* vectors, leaving only the independent vectors.

Definition 4.1 (Linear Independence)**.** For some vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^m$ and scalars $c_1, c_2, \ldots, c_n \in \mathbb{R}$, the sum

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n \tag{4.1}$$

is a linear combination of vectors. The set of vectors is called linearly independent if

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = 0 \tag{4.2}$$

only if $c_1 = c_2 = \cdots = c_n = 0$. If the set is linearly dependent, that means there is a vector in the set which can be re-written as a linear combination of another vector(s). So basically, a set of vectors is linearly independent if every vector is completely unique and cannot be formed by adding scalar multiples of other vectors in that set.

Since we can treat the columns or rows of matrices as vectors, this same idea can be applied to the rows or columns of a matrix. This naturally leads us to the idea of row or column rank (which we will simply just refer to as rank).

Definition 4.2 (Rank)**.** The rank of a matrix is simply the number of pivots in row echelon form, written as $\text{rank}(A)$. The rank of a matrix also tells us the number of linearly independent columns of a matrix. That being, if there is a pivot in a column, that column is linearly independent, meaning columns without a pivot can simply be re-written as a linear combination of the columns with a pivot.

Remark 4.1. Given some $A \in \mathcal{M}_{n\times m}(\mathbb{R})$, if $\text{rank}(A) = n$, we say the matrix A is "full rank".

Proposition 4.1. *Performing row operations on a matrix does not change its rank. That is, all matrices that are row equivalent have the same rank.*

The notion of rank becomes extremely important when we discuss systems of linear equations. For example, consider the set of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{4.3}$$

what can we say about it? Do we know if it has one solution, none? Maybe even infinitely many? We can record our system of linear equations using matrix-vector notation, $A\boldsymbol{x} = \boldsymbol{b}$.

Example 4.1. Consider the system of equations

$$\begin{cases} x + y + z = 4 \\ x + y + 2z = 6 \\ x + 3y + 2z = 8 \end{cases}. \tag{4.4}$$

We can record the coefficients of our variables x, y, z into the matrix A, our variables into the vector $\boldsymbol{x}$ and the right side of the equations into the vector $\boldsymbol{b}$. We can then re-write the equation as

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{=\boldsymbol{x}} = \underbrace{\begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix}}_{=\boldsymbol{b}}. \tag{4.5}$$

If you compute the product $A\boldsymbol{x}$, you will see that we will have the left side of our original system in (4.4). We can then perform row operations on the matrix A to help us solve the system. We can stop performing row operations whenever but most people tend to stop when the matrix is in row echelon form or perhaps even row reduced echelon form depending on how easy it is to get it there.

Remark 4.2. If you can't already notice, I will be denoting matrices with capital letters and vectors with bold lowercase letters. This is pretty standard notation for these objects.

We use some interesting syntax when we solve these systems. If we want to solve the system $A\boldsymbol{x} = \boldsymbol{b}$, we row reduce $A|\boldsymbol{b}$ ("A augmented by $\boldsymbol{b}$") which ends up looking like

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right). \tag{4.6}$$

The vertical line is just there for syntax reasons, if we perform row operations on this matrix, we include the vector $\boldsymbol{b}$. This leads us to an interesting idea.

Theorem 4.1. *For some* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ *and* $\boldsymbol{b} \in \mathbb{R}^m$*, if and only if* $\operatorname{rank}(A) = \operatorname{rank}(A|\boldsymbol{b})$*, the system is consistent. If* $\operatorname{rank}(A) \neq \operatorname{rank}(A|\boldsymbol{b})$*, we say the system is inconsistent.*

Example 4.2 (Inconsistent System)**.** What does an inconsistent system look like? Well, for example, consider the following system expressed in matrix-vector notation:

$$\left(\begin{array}{cc|c} * & * & * \\ 0 & 0 & 1 \end{array}\right). \tag{4.7}$$

If we look at the second row of the augmented matrix, we see that $0x_1 + 0x_2 = 1$ which is clearly not possible. Therefore, we say the system is inconsistent.

We can now discuss the idea of existence and uniqueness, a critical component in studying systems of equations. Take two lines (linear equations) and consider the following possibilities:

1. If they intersect, there is one solution.
2. If they are the same line, there are infinitely many solutions.
3. If they are parallel, there is no solution.

This notion can be expanded to any number of linear equations in a system. Similarly, we have three cases. First, we are going to talk about when there is no solution. For example, if we have the matrix-vector system

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \tag{4.8}$$

we can see that we cannot possible have a solution since $x_1 + 2x_2 = 1$ but $2x_1 + 4x_2 = 3$. Therefore, this system admits no solutions. The next case is when we have one solution. Take a look at the elementary system

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \tag{4.9}$$

where there is clearly one solution where $x_1 = 3$ and $x_2 = 5$. Lastly, we can discuss the case where there is no solution. This is a little more complicated because we need to introduce the idea of *free variables*. For example, if we have

$$\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}, \tag{4.10}$$

you may be able to notice that we cannot write our solution without being in terms of x_2 which we call a free variable. For any value of x_2, our solution is true. It is also possible to have multiple free variables which will result in higher degrees of freedom and higher dimensional solution spaces.

5. The Inverse of a Matrix

If we have an equation $5x = 2$, we can simply solve for x by multiplying by the inverse of 5, 5^{-1}, on both sides. Though the inverse is more commonly written though as $\frac{1}{5}$. In the same way real numbers have inverses, that we can multiply to obtain the multiplicative identity, 1, matrices can also have inverses. However, the keyword is "can" and we are not always guaranteed an inverse.

Definition 5.1 (Inverse of a Matrix)**.** A matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ is considered to be invertible if it meets the following requirements:

1. If the matrix A is square, in this case, $m = n$.
2. If $A \sim I$, meaning $\text{RREF}(A) = I$.

This may be hard to notice at first, but we can encode row operations in a matrix. For instance, consider the following four matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{5.1}$$

If we multiple the first matrix by another 2×2 real matrix, nothing changes, that is the identity matrix. However, the second matrix will scale the first row by a factor of 2. The third matrix will swap the rows. Finally, the third matrix will add the second row to the first row. Knowing this, we can store the k row operations we perform to reduce A to the identity matrix in the matrices R_i where

$$R_k \dots R_3 R_2 R_1 A = I. \tag{5.2}$$

Therefore, it should be clear that we can define the inverse as

$$A^{-1} = \prod_{i=1}^{k} R_i = R_k \dots R_3 R_2 R_1. \tag{5.3}$$

A trick to doing this is to reduce the augmented matrix $A|I$. If you can reduce A to the identity matrix, then you will have $I|A^{-1}$. Basically, the matrix A you want to find the inverse of will become the identity matrix and the identity matrix will become the inverse of A.

Proposition 5.1. *If A^{-1} exists, then $A^{-1}A = I = AA^{-1}$.*

Remark 5.1. Simply because there exist two matrices A, B such that $AB = I$, that does not make these matrices are always invertible. For example, consider the two matrices

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \end{pmatrix}, \; B = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \\ 3/2 & -1/2 \end{pmatrix}. \tag{5.4}$$

We see that $AB = I$ if we multiply them out, but since they are not square matrices, they cannot be *true* inverses.

In the following remark, I emphasize that A, B are not "true" inverses. However, this is because there is actually an idea called the *pseudo-inverse.*

Definition 5.2 (Pseudo-Inverse of a Matrix)**.** For some $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, let $A^{+} = \left(A^{\intercal}A\right)A^{\intercal}$ so that

$$A^{+}A = \left(A^{\intercal}A\right)A^{\intercal}A = I. \tag{5.5}$$

While this will most definitely not be useful for the applications most people will be doing, this does have many uses for when an inverse is required but not always possible.

Definition 5.3 (Non-singular Matrix)**.** If $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and A^{-1} exists, then the matrix A is non-singular. If we say a matrix is singular, it does not have an inverse.

Often times taking the inverse of matrices is boring and time consuming. However, if we have a 2×2 matrix, in the form of

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5.6}$$

we can simplify taking the inverse by applying

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{5.7}$$

This formula does indeed apply to higher dimensional matrices, but it becomes much more complicated. There exists a specific type of matrix called the *adjugate matrix*, $\text{adj}(A)$ where the inverse of any square $m \times n$ matrix can be expressed as

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A) \tag{5.8}$$

While we do not need to know that an adjugate matrix by definition, we can extract an interesting theorem from (5.8).

Theorem 5.1. *The inverse A^{-1} exists iff.* $\det(A) \neq 0$ *for some* $A \in \mathcal{M}_{n \times n}(*)$.

Proof. It should be obvious, if $\det(A) = 0$, we will have a division by 0 which is not defined. □

Proposition 5.2. *If A is an $n \times n$ invertible matrix, then* $\left(A^{\mathsf{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathsf{T}}$.

Proof. The inverse A^{-1} satisfies $A^{-1}A = I$. Therefore, taking the transpose of both matrices gives us

$$A^{\mathsf{T}} \left(A^{-1}\right)^{\mathsf{T}} = I, \tag{5.9}$$

showing us that $\left(A^{-1}\right)^{\mathsf{T}}$ is the inverse of A^{T}, or that $\left(A^{\mathsf{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathsf{T}}$. □

6. Eigenvalues and Eigenvectors

We are used to solving systems in the form $A\boldsymbol{x} = \boldsymbol{b}$ where $A \in \mathcal{M}_{m\times n}(*)$. However, there may be times when we need to solve

$$A\boldsymbol{x} = \lambda\boldsymbol{x}. \tag{6.1}$$

Definition 6.1 (Eigenvalue)**.** An eigenvalue is a scalar λ that solves the equation (6.1) where $A \in \mathcal{M}_{n\times n}(*)$ (note this is a square matrix) and $\boldsymbol{x} \in \mathbb{R}^n$.

If we think of the matrix A as a function or *transformation*, the geometric idea of (6.1) is that what scalar(s) can we multiply by $\boldsymbol{x}$ where A only scales the matrix rather than also changing its direction. This scalar λ is called an eigenvalue. We can modify and rearrange the equation to get

$$A\boldsymbol{x} - \lambda\boldsymbol{x} = 0 \iff A\boldsymbol{x} - \lambda I\boldsymbol{x} = (A - \lambda I)\boldsymbol{x} = 0. \tag{6.2}$$

Since this equation allows for $\boldsymbol{x} = \mathbf{0}$ and we want the *non-trivial solution*, we need the existence of M^{-1} where $M = A - \lambda I$. To guarantee this existence, $\det(M) \neq 0$. So, $A\boldsymbol{x} = \lambda\boldsymbol{x}$ will have nontrivial solutions if and only if

$$\det(A - \lambda I) = 0. \tag{6.3}$$

So, to find the eigenvalues of a given matrix, we need to solve the equation (6.3). Expanded out, this looks like

$$\det(A - \lambda I) = \det\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} - \lambda \end{pmatrix} = 0. \tag{6.4}$$

Some references also call this the *characteristic polynomial* since when the determinant is taken, you will find that it is an n-th term polynomial of λ for which we can simply solve. So, we can write out final equation as

$$\rho_A(\lambda) := \det(A - \lambda I) = 0. \tag{6.5}$$

using the notation of the characteristic polynomial. To better understand this, we can perform some basic examples.

Example 6.1 (Eigenvalues of a 2×2 matrix)**.** Consider the matrix

$$A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}. \tag{6.6}$$

To solve the system $A\boldsymbol{x} = \lambda\boldsymbol{x}$, we of course set its characteristic polynomial to 0 via

$$\begin{aligned} \rho_A(\lambda) = \det(A - \lambda I) &= \det\begin{pmatrix} 1-\lambda & 3 \\ 0 & 2-\lambda \end{pmatrix} && (6.7) \\ &= (1-\lambda)(2-\lambda) - 3(0) && (6.8) \\ &= \lambda^2 - 3\lambda + 2 && (6.9) \\ &= (\lambda - 2)(\lambda - 1) = 0 \implies \lambda = 1, 2. && (6.10) \end{aligned}$$

Therefore, the eigenvalues of A are 1 and 2.

Remark 6.1. If you haven't noticed already, if A is an $n \times n$ square matrix, the number of eigenvalues we have is also n. It is indeed possible to have repeated eigenvalues. If a_{22} was a 1 instead of a 2, we would have the repeated eigenvalues of $1, 1$. On second thought the examples I picked just looks like the eigenvalues are the diagonal entries, but this is not always the case.

Definition 6.2 (Eigenvector)**.** We can continue this further. After finding the eigenvalues of a matrix, we can then plug them back into the equation

$$(A - \lambda I)\boldsymbol{x} = \mathbf{0} \tag{6.11}$$

to solve for the vector $\boldsymbol{x}$, the eigenvector(s). Note that the eigenvectors will be written in terms of some free variable usually. Therefore, there will be infinitely many of them.

Continuing with the previous example, we will use the eigenvalues obtained to find the eigenvectors.

Example 6.2. With the eigenvalues $\lambda = 1, 2$, we can solve for the eigenvectors by solving (6.11). First, let us consider $\lambda = 1$. We solve

$$(A - 1I)\boldsymbol{x} = \begin{pmatrix} 1-1 & 3 \\ 0 & 2-1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6.12}$$

which we can do by augmenting $A - I$ by $\boldsymbol{x}$ and performing row operations. However, the matrix I selected is ******* and we do not need to perform row operations. We just open our eyes and see that

$$\begin{cases} 0x_1 + 3x_2 = 0 \\ 0x_1 + 1x_2 = 0 \end{cases} \tag{6.13}$$

which means that since $3x_2 = x_2 = 0$, it must be that $x_2 = 0$. As aforementioned, we will write our eigenvector in terms of our free variable, x_1. Therefore, the eigenvectors for the matrix A using $\lambda = 1$ are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} x_1. \tag{6.14}$$

Now, we continue with $\lambda = 2$, the other eigenvalue. We do the exact same thing as in (6.12):

$$(A - 2I)\boldsymbol{x} = \begin{pmatrix} 1-2 & 3 \\ 0 & 2-2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{6.15}$$

Once again, if we open our eyes, we see that

$$-1x_1 + 3x_2 = 0 \implies x_1 = 3x_2 \tag{6.16}$$

which means that we can write our eigenvector for $\lambda = 2$ as

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} x_2. \tag{6.17}$$

Finally, we have found both eigenvectors for the matrix A. As you can probably tell, this is not very fun for larger matrices as the systems become harder to solve and there are more eigenvalues and vectors.

Now for fun (despite this being quite common), if we have the system $A\boldsymbol{x} = \lambda\boldsymbol{x}$ where $\lambda \in \mathbb{C}$, each value of λ also admits a *complex conjugate*, $\overline{\lambda}$. Instead of doing extra work to find the eigenvector of the conjugate, we can simply take $\overline{A\boldsymbol{x}} = \overline{\lambda\boldsymbol{x}}$ which is simply $A\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}$. This notion leads up to the following proposition.

Proposition 6.1. *Following from the previous paragraph, for every $\boldsymbol{x} \in \mathbb{C}^n$ that is an eigenvector of $\lambda \in \mathbb{C}$, the vector $\overline{\boldsymbol{x}}$ is an eigenvector of $\overline{\lambda}$.*

We also have the following propositions which are quite interesting and useful.

Proposition 6.2. *A and A^T share the same eigenvalues but not always the same eigenvectors. The eigenvectors are the same if the matrix is diagonal or symmetric.*

Proposition 6.3. *If A^{-1} exists, the eigenvalues are reciprocals of the eigenvalues of A with the same eigenvectors. So, if we know the eigenvalues and eigenvectors of A and there exists some A^{-1}, we know the eigenvalues and eigenvectors of A^{-1}.*

You are basically done, any section past this point you will most likely not need.

7. Fundamental Subspaces

This topic will probably not be needed by most people but we will discuss it anyway because it is interesting. Therefore, I will approach it with a more theoretical stance. Matrices have what is called *subspaces*. The primary two are the *column space* (or row space) and the *null space*.

Definition 7.1 (Column Space)**.** The column space of a matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $R(A)$, is the set of all linear combinations of the columns of A which we will denote via $\mathrm{R}(A)$. More mathematically speaking,

$$\mathrm{R}(A) := \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{y} = A\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^n\}\,. \tag{7.1}$$

An interesting property is that the column space is a subspace of $\mathbb{R}^m$, the codomain. Another interesting fact is if and only if the vector $\boldsymbol{b} \in \mathrm{R}(A)$, then $A\boldsymbol{x} = \boldsymbol{b}$ has a solution. This is critical in understanding the column space.

Proposition 7.1. *For some* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$*,* $\dim \mathrm{R}(A) = \operatorname{rank}(A)$.

When it comes to actually finding the column space of a matrix, we can simply say that

$$\mathrm{R}(A) = \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \tag{7.2}$$

where $\boldsymbol{v}_i$ is column (written as a vector) for which a pivot appears after putting the matrix in REF.

Example 7.1 (Column Space)**.** Consider the following matrix and its REF

$$A = \begin{pmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 0 & 2 & 4 \end{pmatrix},\ \mathrm{REF}(A) = \begin{pmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}. \tag{7.3}$$

We can then write the column space of A as

$$\mathrm{R}(A) = \operatorname{span}\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} \right\}, \tag{7.4}$$

simply the span of the set of the original columns containing a pivot.

Definition 7.2 (Row Space)**.** I mentioned the existence of a row space, this is simply $\mathrm{R}(A^\mathsf{T})$. This is one of the four fundamental subspaces.

Definition 7.3 (Null Space)**.** The null space (sometimes called kernel) of a matrix which we will denote as $\mathrm{N}(A)$, is the set of all vectors that a matrix A maps to the zero vector. So,

$$\mathrm{N}(A) := \{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \boldsymbol{0}\} \tag{7.5}$$

which is a subspace of the domain, $\mathbb{R}^n$.

Definition 7.4 (Left Null Space)**.** Similar to the row space, the left null space is simply $\mathrm{N}\left(A^\mathsf{T}\right)$.

Theorem 7.1 (Rank-Nullity Theorem)**.** *For some* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$*,* $\dim \mathrm{N}(A) = n - \operatorname{rank}(A)$.

An interesting result from the previous theorem is that $\operatorname{rank}(A) + \dim \mathrm{N}(A) = n$. Since the number of vectors that are spanned in the column space is the rank of A, it should be clear that $\operatorname{rank}(A) = \dim \mathrm{R}(A)$. This means that

$$\dim \mathrm{N}(A) + \dim \mathrm{R}(A) = n. \tag{7.6}$$

Remark 7.1. From a systems of equations standpoint, the column space is all vectors that solve the particular solution and the null space is all vectors that solve the homogeneous solution.

What we have discussed so far leads us to one of the most important theorems in linear algebra, you can probably guess this from its name. To better understand it, we need to discuss the idea of an *orthogonal complement*.

Definition 7.5 (Orthogonal Complement)**.** For a subspace $V \subseteq \mathbb{R}^n$, the orthogonal complement is

$$V^\perp := \{\boldsymbol{w} \in \mathbb{R}^n \mid \boldsymbol{v}\cdot\boldsymbol{w} = 0\ \forall \boldsymbol{v} \in V\}. \tag{7.7}$$

Theorem 7.2 (Fundamental Theorem of Linear Algebra)**.** *For some* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$*, the following statements are true*

$$\mathrm{R}\left(A^\intercal\right) = (\mathrm{N}(A))^\perp,\ \mathrm{N}\left(A^\intercal\right) = (\mathrm{R}(A))^\perp. \tag{7.8}$$

Now I will introduce some critical corollaries that stem from the fundamental theorem.

Corollary 7.1. *For* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ *with* $\mathrm{rank}(A) = r$*,* $\dim \mathrm{N}(A) = n - r$ *and* $\dim \mathrm{N}\left(A^\intercal\right) = m - r$*.*

Corollary 7.2. *For some* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$*,* $\mathbb{R}^n = \mathrm{R}\left(A^\intercal\right) \oplus \mathrm{N}(A)$ *and* $\mathbb{R}^m = \mathrm{R}(A) \oplus \mathrm{N}\left(A^\intercal\right)$*.*

In fact, the idea that $\mathrm{rank}(A) = \mathrm{rank}(A^\intercal)$ also stems from this. This tells us that row rank and column rank are equal to each other, this is why we are able to generalize row and column rank to just one "rank".

The following diagram can help visualize the idea of how the subspaces relate to each other.

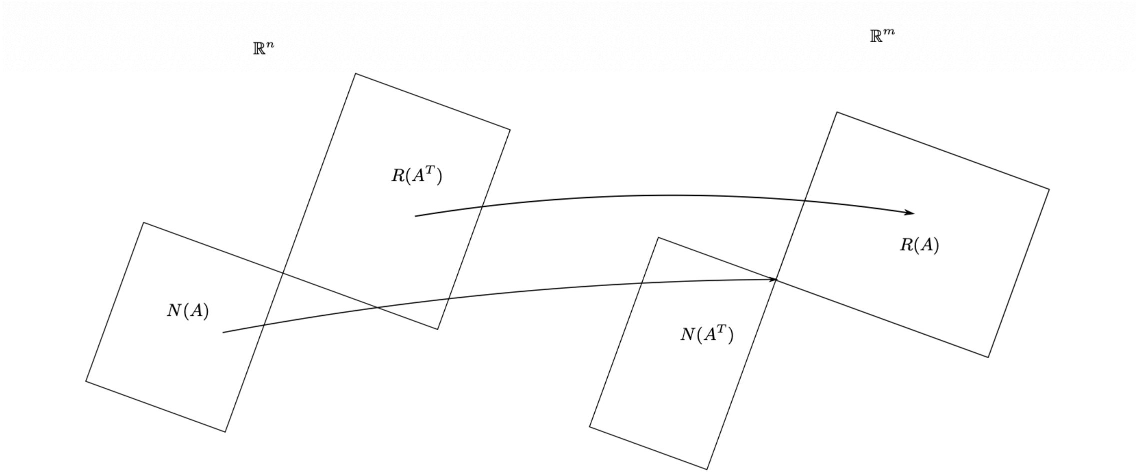

FIGURE 7.1. A geometric representation of the four fundamental subspaces.

This section seems pretty useless, and for most people, it is. These ideas simply build off of pre-existing ideas and help reach the conclusions we have already learned. We will need to dive much deeper for this information to be truly useful.

Remark 7.2. You will often see in other sources that the column space is denoted as $\mathrm{Col}(A)$ and the null space as $\mathrm{Nul}(A)$ or even $\mathrm{Ker}(A)$. In fact, that tends to be the more common way of writing them, but this is simply a matter of notation.