Lecture 3: Linear Algebra

Naz Koont

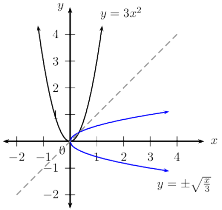
23 August 2021

Today: Linear Algebra

1. Vectors
    - Vector Operations
    - Vector Length, Distance
2. Matrices
    - Matrix Algebra
    - Examples of Matrices
    - Properties of Linear Systems
    - Determinants, Inverse
    - Eigenvalues, Eigenvectors

Follow up to last class

1. How do we know if a function may not have a limit at a due to oscillation, without looking at a graph?
    - Comment from last class: composition of periodic function such as $sin(x)$ with a function that grows without bound at a, such as $1/x$ at $a = 0$.
2. Inverse of non-injective function such as a parabola does not exist because each x is not assigned to a unique $f^{-1}(x)$
3. Minor typo on slide 41 from last class, integral evaluates to $\frac{25}{3}$
4. Lecture videos are currently not viewable, I have let ITG know to fix this.



Motivation

In many financial or economic problems, a single linear equation identifies or characterizes a relationship between two variables, x and y.

- Often, however, there are two or more equations that must be satisfied *simultaneously*.
- **Linear Algebra** provides powerful methods for finding solutions to two or more linear equations.

A linear equation is any equation of the form

$$c_1x_1 + c_2x_2 + ... + c_nx_n = b \text{ for } c_1, ...c_n, b \in \mathbb{R}$$

where c_i are coefficients of the linear equation and b is the constant term.

Geometry of Linear Algebra

To understand the intuition and theory behind the tools that Linear Algebra provides for solving systems of linear equations, it is useful to think about the geometry underlying the results.

- Previously, we have considered points in Euclidean space as ordered n-tuples of numbers.
- We can define these n-tuples as *vectors*.
- In fact, $\mathbb{R}^n$ admits a vector space structure, i.e. our standard rules of addition, subtraction, and scalar multiplication will continue to behave nicely when considering vectors rather than numbers, which will lead to useful calculation methods.

Definition (Vector)

A **vector** x is an array of real numbers, where the corresponding row vector is x' ("x prime").

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad x' = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$$

Vector Addition

Definition (Vector Addition)
To add or subtract two vectors, add or subtract the corresponding components.

$$x + y = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix} = y + x$$

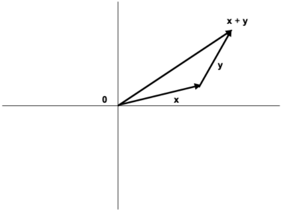


Figure 6.1: Vector Addition

Theorem (Properties of Vector Addition)

1. *Vector addition is commutative, i.e.* $x + y = y + x$
2. *For* $n \neq m$*, cannot add* $x^n + y^m$

Vectors and Scalar Multiplication

Definition (Scalar Multiplication)
For $k \in \mathbb{N}$,

$$k * x = k * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} k * x_1 \\ k * x_2 \\ k * x_3 \end{bmatrix} = x * k$$

Theorem (Properties of Scalar Multiplication)

1. *Scalar multiplication is commutative, i.e.* $k * x = x * k$
2. *Scalar multiplication is distributive over vector addition, i.e.* $k(x + y) = k * x + k * y$ *for all* $x, y \in \mathbb{R}^n; k \in \mathbb{N}$
3. *Scalar multiplication is distributive over scalar addition, i.e.* $(h + k)x = h * x + k * x$ *for all* $x \in \mathbb{R}^n; h, k \in \mathbb{N}$

Vector Length and Distance

Given our understanding of vectors, we want to be able to think about concepts of length and distance.

- We will define these concepts largely as a generalization of the familiar concept of Euclidean distance as defined in the Fundamentals section.
- As a starting point, the concept of an *inner product* will be useful for this purpose.

Definition (Inner Product)
An Inner Product is a linear function that assigns a number given two vectors, i.e. $f : \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$ (and satisfies certain nice properties). The standard inner product, or "Dot Product" of two vectors $x, y \in \mathbb{R}^n$ is:

$$x \cdot y = x'y = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} * \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x_1 * y_1 + x_2 * y_2 + x_3 * y_3 = \sum_{i=1}^{n} x_i * y_i$$

Inner Product

Definition (Inner Product)
An Inner Product is a linear function that assigns a number given two vectors, i.e. $f : \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$ (and satisfies certain nice properties). The standard inner product, or "Dot Product" of two vectors $x, y \in \mathbb{R}^n$ is:

$$x \cdot y = x' y = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} * \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x_1 * y_1 + x_2 * y_2 + x_3 * y_3 = \sum_{i=1}^{n} x_i * y_i$$

Theorem (Properties of Inner Product)

1. *The Inner Product is commutative, i.e.* $x \cdot y = y \cdot x$
2. *For* $x, y \in \mathbb{R}^n$ *and* $k \in \mathbb{R}$*:* $k(x'y) = (k * x)'y = x'(k * y)$

Vector Norm

Given this, we can define the concept of the length, or *norm* of a vector.

Definition (Norm)
Given an Inner Product $x \cdot y$, we can define the norm or length of any vector as $||x|| = \sqrt{x \cdot x}$.

The norm of a vector satisfies several important properties:

Theorem (Properties of a Norm)

1. $||c * x|| = |c| \cdot ||x||$ *for* $c \in \mathbb{R}$
2. $||x|| = 0 \iff x = 0$ *and for all* x, $||x|| \geq 0$
3. *Cauchy-Shwarz Inequality:* $|x \cdot y| \leq ||x|| \cdot ||y||$
4. *Triangle Inequality:* $||x + y|| \leq ||x|| + ||y||$

Vector Distance

Finally, we can define a notion of distance, or a *metric*, between two vectors.

Definition (Distance Between Vectors)
Given an Inner Product $x \cdot y$, we can define a metric, or distance function, between any two vectors as $d(x, y) = ||x - y|| = \sqrt{(x - y) \cdot (x - y)}$

Notice that the definitions of length and distance coincide with their Euclidean formulations with the standard inner product.

Definition (Matrix)
A matrix A is a rectangular array of numbers. A number appearing in a matrix is called an entry of A. If the array has n rows and m columns, we say that A has size n by m, or that A is an n by m matrix. We denote the entry of A appearing in the ith row and jth column as a_{ij}, where i is the row index and j is the column index of this entry.

$$A \in \mathbb{R}^3 \times \mathbb{R}^3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Matrix Algebra

Definition (Matrix Addition and Subtraction)
Addition and subtraction of matrices is well defined only if the matrices involved are of the same size. The sum of two matrices is a matrix, the elements of which are the sums of the corresponding elements of matrices (and analogously for matrix subtraction).

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}$$

Definition (Scalar Multiplication)
In matrix algebra, real numbers are called scalars. Multiplying a matrix by a scalar is called scalar multiplication, and it is carried out by multiplying each element of the matrix by the scalar.

$$3A = \begin{bmatrix} 3a_{11} & 3a_{12} \\ 3a_{21} & 3a_{22} \end{bmatrix}$$

Matrix Algebra

Definition (Matrix Multiplication)

Multiplication of matrices is well defined only if the matrices involved are conformable.

This is the case if the number of columns of the first matrix are the same as the number of rows of the second matrix.

The product of two matrices $A \in \mathbb{R}^{m \times n} * B \in \mathbb{R}^{n \times q}$ is a matrix $C \in \mathbb{R}^{m \times q}$, and its ijth element c_{ij} is obtained by multiplying the elements of the ith row of A by the elements of the jth column of B and adding the resulting products.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} * \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11} b_{11} + a_{12} b_{21} + a_{13} b_{31} & a_{11} b_{12} + a_{12} b_{22} + a_{13} b_{33} \\ a_{21} b_{11} + a_{22} b_{21} + a_{23} b_{31} & a_{21} b_{12} + a_{22} b_{22} + a_{23} b_{32} \end{bmatrix}$$

In general, the product matrix can be expressed as $c_{ij} = \sum_{k=1}^{n} a_{jk} b_{kj}$.

Notice the similarity between the expression for a given element in the resulting matrix c_{ij}, and the notion of taking a dot product between two vectors.

Theorem (Properties of Matrix Multiplication)

1. *Matrix multiplication is not commutative. In fact, even if $A * B$ exists, $B * A$ may not exist depending on the dimensions of the matrices.*
2. *Matrix multiplication is distributive with respect to addition:* $A(B + C) = AB + AC$ *and* $(B + C)A = BA + CA$
3. *Matrix multiplication is associative:* $A(BC) = (AB)C$
4. *Matrix multiplication is commutative with respect to scalars:* $A(\lambda B) = \lambda(AB)$
5. *The unit matrix is the neutral element of matrix multiplication:* $IA = AI = A$
6. *The zero matrix is absorbent:* $0A = A0 = 0$

Matrix Algebra

Matrix multiplication is not commutative. In fact, even if $A * B$ exists, $B * A$ may not exist depending on the dimensions of the matrices.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} * \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix}$$

Matrix Algebra

Matrix multiplication is distributive with respect to addition:
$A(B + C) = AB + AC$ and $(B + C)A = BA + CA$

$A(B + C) =$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \right) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11} + c_{11} & b_{12} + c_{12} \\ b_{21} + c_{21} & b_{22} + c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}(b_{11} + c_{11}) + a_{12}(b_{21} + c_{21}) & a_{11}(b_{12} + c_{12}) + a_{12}(b_{22} + c_{22}) \\ a_{21}(b_{11} + c_{11}) + a_{22}(b_{21} + c_{21}) & a_{21}(b_{12} + c_{12}) + a_{22}(b_{22} + c_{22}) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}(b_{11} + c_{11}) & a_{11}(b_{12} + c_{12}) \\ a_{21}(b_{11} + c_{11}) & a_{21}(b_{12} + c_{12}) \end{bmatrix} + \begin{bmatrix} a_{12}(b_{21} + c_{21}) & a_{12}(b_{22} + c_{22}) \\ a_{22}(b_{21} + c_{21}) & a_{22}(b_{22} + c_{22}) \end{bmatrix}$$

$$= AB + AC$$

Matrix Algebra

Matrix multiplication is associative: $A(BC) = (AB)C$

$A(BC) =$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} * \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \right)$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11}c_{11} + b_{12}c_{21} & b_{11}c_{12} + b_{12}c_{22} \\ b_{21}c_{11} + b_{22}c_{21} & b_{21}c_{12} + b_{22}c_{22} \end{bmatrix}$$

$$\left[\begin{matrix} a_{11}(b_{11}c_{11} + b_{12}c_{21}) + a_{12}(b_{21}c_{11} + b_{22}c_{21}) & a_{11}(b_{11}c_{12} + b_{12}c_{22}) + a_{12}(b_{21}c_{12} + b_{22}c_{22} \\ a_{21}(b_{11}c_{11} + b_{12}c_{21}) + a_{22}(b_{21}c_{11} + b_{22}c_{21}) & a_{21}(b_{11}c_{12} + b_{12}c_{22}) + a_{22}(b_{21}c_{12} + b_{22}c_{22} \end{matrix}\right.$$

$$\left[\begin{matrix} c_{11}(a_{11}b_{11} + a_{12}b_{21}) + c_{21}(a_{11}b_{12} + a_{12}b_{22}) & c_{12}(a_{11}b_{11} + a_{12}b_{21}) + c_{22}(a_{11}b_{12} + a_{12}b_{2} \\ c_{11}(a_{21}b_{11} + a_{22}b_{21}) + c_{21}(a_{21}b_{12} + a_{22}b_{22}) & c_{12}(a_{21}b_{11} + a_{22}b_{21}) + c_{22}(a_{21}b_{12} + a_{22}b_{2} \end{matrix}\right.$$

$= (AB)C$

Examples of Matrices

Definition (Square Matrix)
A matrix that has the same number of rows and columns is called a square matrix

$$A \in \mathbb{R}^2 \times \mathbb{R}^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Definition (Triangular Matrix)
A matrix A that has 0s in every element above or below its diagonal is a triangular matrix. If all elements below the diagonal are 0, then A is an upper-triangular matrix. If all elements above the diagonal are 0, then A is a lower-triangular matrix.

$$U = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \quad L = \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$$

Definition (Diagonal Matrix)
A square matrix that has only nonzero entries on the main diagonal and zeroes everywhere else is known as a diagonal matrix. A diagonal matrix is also triangular.

$$D \in \mathbb{R}^3 \times \mathbb{R}^3 = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Examples of Matrices

Definition (Identity Matrix I)
A special case of a diagonal matrix is the identity matrix, where the diagonal entries are all 1s. We will see that the identity matrix plays the same role in matrix algebra as the number 1 does in the algebra of real numbers.

$$\boldsymbol{I} \in \mathbb{R}^3 \times \mathbb{R}^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition (Null Matrix)
A square matrix with all its entries being 0 is known as the null matrix. The null matrix plays a similar role in matrix algebra as does 0 in the algebra of real numbers.

$$0 \in \mathbb{R}^2 \times \mathbb{R}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Examples of Matrices

Definition (Transpose Matrix)
The transpose matrix A^T is the original matrix A with its rows and columns interchanged

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

1. A matrix that is equal to its transpose is called a symmetric matrix.
2. The transpose of a transpose matrix is the original matrix: $(A^T)^T = A$.
3. The transpose of a sum of matrices is the sum of transposes $(A + B)^T = A^T + B^T$
4. The transpose of a product of matrices: $(AB)^T = B^T * A^T$

Definition (Orthogonal Matrix)

Two vectors $x, y \in \mathbb{R}^n$ are **orthogonal** if $x^T y = 0$ (i.e., if their dot product is 0). A vector $x \in \mathbb{R}^n$ is **normalized** if its norm is 1, $|x| = 1$. A square matrix $U \in \mathbb{R}^{n \times n}$ is orthogonal if all of its columns are orthogonal to each other and are normalized. The columns are then referred to as being orthonormal. Then,

$$U^T U = U U^T = I_n$$

Solving Systems of Linear Equations

Solving a system of linear equations is probably the most important application of linear algebra for our purposes.

- For systems of two linear equation, finding solutions by simple substitution of unknowns works well, but for systems with a larger number of unknowns we need a new tractable solution method.
- One method of doing so involves reformulating a given system of linear equations into a form that has the same solution but is easier to solve.

There are three types of **row operations** that are used to simplify the original system:

1. Interchange the order of any two equations in the system.
2. Multiply any equation in the system by a nonzero constant.
3. Add a constant multiple of any equation to another equation in the system.

Solving Systems of Linear Equations

There are three types of **row operations** that are used to simplify the original system:

1. Interchange the order of any two equations in the system.
2. Multiply any equation in the system by a nonzero constant.
3. Add a constant multiple of any equation to another equation in the system.

And we would like to arrive at a system that has the following properties:

1. The first nonzero coefficient of each equation is one.
2. If an unknown is in the first unknown with a nonzero coefficient in some equation, then that unknown occurs with a 0 coefficient in each of the other equations.
3. The first unknown with a nonzero coefficient in any equation has a larger subscript than the first unknown with a nonzero coefficient in any preceeding equation.

Solving Systems of Linear Equations

Lets see an example of this method. We are given the following system of equations:

$$a_1 - 2a_2 + 2a_4 - 3a_5 = 2 \tag{1}$$
$$2a_1 - 4a_2 + 2a_3 + 8a_5 = 6 \tag{2}$$
$$a_1 - 2a_2 + 3a_3 - 3a_4 + 16a_5 = 8 \tag{3}$$

As a first step, let us eliminate a_1 from every equation except the first by using rule 3. We add $-2 * (1)$ to equation (2), and $-1 * (1)$ to equation (3). While doing this, it will so happen that we also eliminate a_2 from every equation except the first.

$$a_1 - 2a_2 + 2a_4 - 3a_5 = 2 \tag{1}$$
$$2a_3 - 4a_4 + 14a_5 = 2 \tag{2}$$
$$3a_3 - 5a_4 + 19a_5 = 6 \tag{3}$$

Solving Systems of Linear Equations

$$a_1 - 2a_2 + 2a_4 - 3a_5 = 2 \quad (1)$$
$$2a_3 - 4a_4 + 14a_5 = 2 \quad (2)$$
$$3a_3 - 5a_4 + 19a_5 = 6 \quad (3)$$

We now want to make the coefficient of a_3 in equation (2) equal to 1, and then eliminate a_3 from equation (3). We use rule 2 to multiply equation (2) by $\frac{1}{2}$, and then add $-3 * (2)$ to equation (3).

$$a_1 - 2a_2 + 2a_4 - 3a_5 = 2 \quad (1)$$
$$a_3 - 2a_4 + 7a_5 = 1 \quad (2)$$
$$a_4 - 2a_5 = 3 \quad (3)$$

Solving Systems of Linear Equations

$$a_1 - 2a_2 + 2a_4 - 3a_5 = 2 \quad (1)$$
$$a_3 - 2a_4 + 7a_5 = 1 \quad (2)$$
$$a_4 - 2a_5 = 3 \quad (3)$$

We continue by eliminating a_4 by every equation except the (3). We do so by adding $-2 * (3)$ to equation (1) and adding $2 * (3)$ to equation (2). This yields:

$$a_1 - 2a_2 + a_5 = -4 \quad (1)$$
$$a_3 + 3a_5 = 7 \quad (2)$$
$$a_4 - 2a_5 = 3 \quad (3)$$

Solving Systems of Linear Equations

$$a_1 - 2a_2 + a_5 = -4 \qquad (1)$$
$$a_3 + 3a_5 = 7 \qquad (2)$$
$$a_4 - 2a_5 = 3 \qquad (3)$$

We have arrived in a system of linear equations that is easy to solve. We can solve for the first unknown present in each equation in terms of the other unknowns (a_2, a_5). This results in:

$$a_1 = 2a_2 - a_5 - 4 \qquad (1)$$
$$a_3 = -3a_5 + 7 \qquad (2)$$
$$a_4 = 2a_5 + 3 \qquad (3)$$

Solving Systems of Linear Equations

$$a_1 = 2a_2 - a_5 - 4 \tag{1}$$
$$a_3 = -3a_5 + 7 \tag{2}$$
$$a_4 = 2a_5 + 3 \tag{3}$$

Thus for any choice of $a_2, a_5 \in \mathbb{R}$, a vector of the following form is a solution to this system.

$$(a_1, a_2, a_3, a_4, a_5) = (2a_2 - a_5 - 4, a_2, -3a_5 + 7, 2a_5 + 3, a_5)$$

- Notice that this system has infinitely many solutions. In contrast, a system can also have a unique solution, or no solution at all (if these operations ever result in an equation of the form $0 = c$ for $c \neq 0$).
- A system of equations which yields no solution is said to be **inconsistent**.
- A system with infinitely many solutions is said to be **underdetermined**, i.e. there are *free variables* such as a_2, a_5 in the example above.

Elementary Matrix Operations

We can express a system of linear equations as a single matrix equation by writing the constants of the system into a matrix.

- In this representation of the system, the three operations detailed above for solving systems of linear equations are the "elementary row operations" for matrices.
- These operations provide a convenient computational method for determining all solutions to a system of linear equations.

Definition (Elementary Row (Column) Operations)
Let A be an $n \times n$ matrix. Any one of the following three operations on the rows or columns of A is called an elementary row (column) operation:

1. Interchanging any two rows (columns) of A
2. Multiplying any row (column) of A by a nonzero constant
3. Adding any constant multiple of a row (column) of A to another row (column).

Elementary Matrix Operations

Just as before, it is the case that these elementary row operations preserve the solution set of the system of linear equations. Using these operations, we can convert the matrix into a form that is useful for solving the system.

Definition (Reduced Row Echelon Form)
A matrix is said to be in reduced row echelon form if the following three conditions are satisfied:

1. Any row containing a nonzero entry precedes any row in which all the entries are 0 (if any)
2. The first nonzero entry in each row is the only nonzero entry in its column.
3. The first nonzero entry in each row is 1 and it occurs in a column to the right of the first nonzero entry in the preceding row.

Theorem (Unique Reduced Row Echelon Form)

Every matrix admits a unique reduced row echelon form.

Elementary Matrix Operations

Let us see an example of this. Consider the following system:

$$3x_1 + 2x_2 + 3x_3 - 2x_4 = 1 \tag{1}$$

$$x_1 + x_2 + x_3 = 3 \tag{2}$$

$$x_1 + 2x_2 + x_3 - x_4 = 2 \tag{3}$$

Let us rewrite this system into a matrix and use the elementary matrix operations.

$$\left(\begin{array}{cccc|c} 3 & 2 & 3 & -2 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & -1 & 2 \end{array}\right)$$

Elementary Matrix Operations

$$\left(\begin{array}{cccc|c} 3 & 2 & 3 & -2 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & -1 & 2 \end{array}\right)$$

Let us create a 1 in a_{11} by interchanging the first and third rows.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 1 & 1 & 1 & 0 & 3 \\ 3 & 2 & 3 & -2 & 1 \end{array}\right)$$

Let us obtain 0s in a_{21} and a_{31} by adding -1 times the first row to the second row, and -3 times the first row to the third row.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right)$$

Elementary Matrix Operations

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right)$$

Let us obtain a 1 in a_{22} by multiplying the second row by -1.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right)$$

Let us obtain a 0 in a_{32} by adding 4 times the second row to the third row.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{array}\right)$$

Elementary Matrix Operations

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{array}\right)$$

Let us obtain a 1 in a_{33} by multiplying the third row by $-\frac{1}{3}$.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

We have an *upper triangular matrix* i.e., the lower left triangle of the matrix below the diagonal is populated with 0s. Now we want to make the first nonzero entry in each row the only nonzero entry in its column. To do this, lets work upwards (i.e. via *backwards substitution*) and first add 1 times the third row to the second and first row.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

Elementary Matrix Operations

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

Now we add -2 times the second row to the first row, and have created a matrix in row echelon form from applying elementary matrix operations to our original system.

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

This matrix corresponds to a system of linear equations that can be easily solved.

$$x_1 + x_3 = 1 \tag{1}$$

$$x_2 = 2 \tag{2}$$

$$x_4 = 3 \tag{3}$$

A vector of the following form is a solution to this system:

$$(1 - t, 2, t, 3) \quad \text{for } t \in \mathbb{R}$$

Linear Dependence

Definition (Linear Dependence)
A set of vectors $\{x_1, ...x_m\} \subset \mathbb{R}^n$ is said to be linearly independent if no vector can be represented as a linear combination of the remaining vectors.
Conversely, if one vector belonging to the set **can** be represented as a linear combination of the remaining vectors, then the vectors are said to be linearly dependent. That is, if for example

$$x_m = \sum_{i=1}^{m-1} \alpha_i x_i \text{ for some scalar values } \alpha_1, ..., \alpha_{n-1} \in \mathbb{R}$$

Then we say that the vectors $\{x_1, ...x_m\}$ are linearly dependent.

Rank of a Matrix

Definition (Rank of a Matrix)
Given a matrix $A \in \mathbb{R}^{m \times n}$, the size of the largest subset of columns of A that constitute a linearly independent set is the **column rank** of A. With some abuse of terminology, this is often referred to as the number of linearly independent columns of A. Similarly, the **row rank** is the largest number of rows of A that constitute a linearly independent set. For any matrix $A \in \mathbb{R}^{m \times n}$, the column rank **equals** the row rank, and so both quantities are referred to collectively as the **rank** of A, denoted as $rank(A)$.

Theorem (Properties of Matrix Rank)

1. *For $A \in \mathbb{R}^{m \times n}$, $rank(A) \leq min(m, n)$. If $rank(A) = min(m, n)$, then A is said to be full rank.*
2. *For $A \in \mathbb{R}^{m \times n}$, $rank(A) = rank(A^T)$.*
3. *For $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, $rank(AB) \leq min(rank(A), rank(B))$*
4. *For $A, B \in \mathbb{R}^{m \times n}$, $rank(A + B) \leq rank(A) + rank(B)$*

Properties of Linear Systems

Definition (Span)
The span of a set of vectors $\{x_1, ...x_n\}$ is the set of all vectors that can be expressed as a linear combination of $\{x_1, ...x_n\}$. That is,

$$span(\{x_1, ...x_n\}) = \left\{ v | v = \sum_{i=1}^{n} \alpha_i x_i, \alpha_i \in \mathbb{R} \right\}$$

Definition (Range)
The range (or columnspace) of a matrix $A \in \mathbb{R}^{m \times n}$, denoted $R(A)$, is the span of the columns of A.

Definition (Nullspace)
The nullspace of a matrix $A \in \mathbb{R}^{m \times n}$, denoted $N(A)$, is the set of all vectors that equal 0 when multiplied by A, i.e.

$$N(A) = \{x \in \mathbb{R}^n | Ax = 0\}$$

Note that vectors in $R(A)$ are of size m, while vectors in $N(A)$ are of size n, so vectors in $R(A^T)$ and $N(A)$ are both in $\mathbb{R}^n$. In fact, $R(A^T)$ and $N(A)$ are **orthogonal complements**, i.e. disjoint subsets that together span the entire space of $\mathbb{R}^n$. We denote this as $R(A^T) = N(A)^\perp$.

Definition (Basis)
A set of vectors $\{x_1, ...x_n\}$ is a basis for some subset $V \subseteq \mathbb{R}^n$ if and only if each vector $v \in V$ can be uniquely expressed as a linear combination of the vectors $\{x_1, ...x_n\}$.

Theorem (Solutions of a System of Linear Equations)

A system of linear equations has either no solution, exactly one solution, or infinitely many solutions.

Theorem (Solution to a System of Equations)

The system of equations $Ax = b$ for $A \in \mathbb{R}^{n \times n}$ has a solution iff $Rank([A|b] = Rank(A)$. When the system has a solution,

1. *The solution is unique iff the columns of A are linearly independent, i.e. $Rank(A) = n$.*
2. *The system has infinitely many solutions iff $Rank(A) < n$.*

Matrix Division? Matrix Inverse

We have already defined the operations of addition, subtraction, and multiplication on matrices. Can we define rules for diving matrices?

- The answer is yes, but only under certain restrictions. Division is restricted to square matrices that satisfy a condition known as **nonsingularity**, which is equivalent to a square matrix having **full rank**.
- The reason for all of this can again be traced to the relation between matrix algebra and the problem of solving a system of simultaneous linear equations.
- To define the matrix inverse, it is useful to think about the division of real numbers. We can write division by a number b as multiplication by the inverse of b, $\frac{1}{b} = b^{-1}$.
- This inverse further satisfies that $b * b^{-1} = b^{-1} * b = 1$.

Definition (Inverse Matrix)

The inverse matrix A^{-1} of a square matrix A of order n is the matrix that satisfies the condition that $A * A^{-1} = A^{-1}A = I_n$. Note that the inverse is defined only for square matrices, but that not every square matrix has an inverse.

Matrix Inverse

Theorem (Properties of Matrix Inverse)

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$
3. $(A^{-1})^T = (A^T)^{-1}$

When we deal with real numbers, we know that for any nonzero b, the inverse exists. However for A^{-1} to exist, it is not sufficient simply to assume that A is a square matrix that is different from the null matrix.

Definition (Singular Matrix)
Any matrix A for which A^{-1} does not exist is known as a singular matrix. The matrix A for which A^{-1} exists is known as a nonsingular matrix.

Matrix Inverse

Consider the following example. The matrix equation $Ax = b$, where $A \in \mathbb{R}^{n \times n}; x, b \in \mathbb{R}^{n \times 1}$ defines a system of n simultaneous linear equations in n unknowns. Let us solve for x.

$$
\begin{aligned}
Ax &= b \\
A^{-1}Ax &= A^{-1}b \\
I_n x &= A^{-1}b \\
x &= A^{-1}b
\end{aligned}
$$

If A and b are known, this solves for the unknown vector x, provided that A^{-1} exists.

We can see from this example that the existence of the inverse matrix is equivalent to being able to solve a linear system.

Inverse of 2×2 matrix

Definition (Inverse of a 2×2 Matrix)

$$A \in \mathbb{R}^2 \times \mathbb{R}^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies \quad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \equiv = \frac{1}{|A|} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

We can verify that $AA^{-1} = A^{-1}A = I_2$. Clearly in order for the matrix inverse to exist, it must be the case that $|A| \neq 0$.

Notice in the example of a 2×2 matrix, the inverse exists if and only iff $|A| = ab - cd \neq 0$. In fact, this value is referred to as the **determinant** of the matrix A and can be defined generally for a square matrix of size n. The existence of a nonzero determinant gives an equivalence for the invertibility of a given matrix.

Determinant

Definition (Determinant)
For a square matrix A, its determinant, denoted as $\det(A)$, is a value defined inductively in the following way:

1. For a 1×1 matrix $A = a_{11}$, define its determinant as $\det(A) = a_{11}$
2. For an $n \times n$ matrix where $n \geq 2$, define its determinant as:

$$\det(A) \equiv \sum_{j=1}^{n} (-1)^{1+j} a_{1j} det(A_{-i,j})$$

   where $A_{-i,j}$ is the matrix A with the ith row and jth column eliminated (referred to as the (i,j)th cofactor of A).

Theorem (Invertibility and Determinant)

1. *A square matrix A is invertible iff* $\det(A) \neq 0$*.*
2. *For two $n \times n$ matrices, we have that* $\det(AB) = \det(A)\det(B)$*.*
3. *A matrix and its transpose have the same determinant:* $\det(A) = \det(A^T)$*.*

Theorem (Properties of the Determinant)

1. $\det(I_n) = 1$
2. *The determinant of a triangular matrix A is the product of its diagonal elements:* $\det(A) = \Pi_{i=1}^{n} a_{ii}$.
3. *If any two columns of A are equal, then* $\det(A) = 0$.
4. *Antisymmetry: If two columns of A are interchanged, then the determinant changes sign.*
5. *If one adds a scalar multiple of one column to another, the determinant does not change.*

Eigenvalues, Eigenvectors

Definition (Eigenvalues, Eigenvectors)
Let A be an $n \times n$ matrix. A scalar λ is said to be an **eigenvalue** of A if and only if there exists a nonzero vector x in $\mathbb{R}^n$ such that $Ax = \lambda x$. A nonzero vector x in $\mathbb{R}^n$ is said to be an **eigenvector** of A if and only if there exists a scalar λ such that $Ax = \lambda x$.

Theorem (Equivalent Condition for Eigenvalues)

$\lambda \in \mathbb{R}$ *is an eigenvalue of A if and only if* $\det(\lambda I_n - A) = 0$.

Theorem (Eigenvalues and Linear Independence)

Let $A \in \mathbb{R}^{n \times n}$ *with n distinct eigenvalues* $\lambda_1, \lambda_2, ...\lambda_n \in \mathbb{R}$. *Let* $x_1, ..., x_n \in \mathbb{R}^n$ *the corresponding (nonzero) eigenvectors. Then* $x_1, ..., x_n$ *are linearly independent.*

Theorem (Eigenvalues and Invertibility)

A matrix A is invertible if and only if 0 is not one of its eigenvalues.

One of the most important applications of linear algebra, since there is a sense in which a matrix is effectively determined by these values.

Eigenvalues, Eigenvectors

Lets find the Eigenvalues of the following matrix:

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

$$(\lambda I_n - A) = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} \lambda & -1 \\ 2 & \lambda + 3 \end{bmatrix}$$

$$det(\lambda I_n - A) = \lambda^2 + 3\lambda + 2 = 0 \implies \lambda_1 = -1, \lambda_2 = -2$$

The eigenvector corresponding to λ_1:

$$A * v_1 = \lambda_1 v_1 \implies (A - \lambda_1)v_1 = 0$$

$$\begin{bmatrix} -\lambda_1 & 1 \\ -2 & -3 - \lambda_1 \end{bmatrix} v_1 = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} v_1 1 \\ v_1 2 \end{bmatrix} = 0$$

$$v_{11} + v_{12} = 0 \implies v_{11} = -v_{12}$$

$$-2v_{11} - 2v_{12} = 0 \implies v_{11} = -v_{12}$$

$$\begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = k \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ for } k \in \mathbb{R}$$

Today: Linear Algebra

1. Vectors
    - Vector Operations
    - Vector Length, Distance
2. Matrices
    - Matrix Algebra
    - Examples of Matrices
    - Properties of Linear Systems
    - Determinants, Inverse
    - Eigenvalues, Eigenvectors