

# Lecture 26

- System of Linear Equations
- Row Echelon Form
- Gauss-Jordan Elimination

## Recap: Matrices and Linear Equations

We can use matrices and vectors to concisely express systems of linear equations **and to efficiently solve them**.

We begin by defining **vector-vector functions**.

The notation

$$\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

means that  $\mathbf{f}$  is a function that takes a  $n$ -vector as input and produces an  $m$ -vector as output.

One special case is the matrix-vector multiplication function:

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x},$$

where  $\mathbf{A}$  is a  $m \times n$  matrix.

As usual, the vector-vector function  $\mathbf{f}(\mathbf{x})$  is **linear** if it satisfies the **superposition property**:

$$\mathbf{f}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{f}(\mathbf{x}) + \beta\mathbf{f}(\mathbf{y})$$

In fact, if  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is linear, then  $\mathbf{f}$  can be written as  $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ , where  $\mathbf{A}$  is unique.

## Affine Functions

### Affine Function

A vector-function is **affine** if  $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ , where  $\mathbf{A}$  is a  $m \times n$  matrix and  $\mathbf{b}$  is an  $m$ -vector.

We define this linear matrix-vector multiplication function as

$$\begin{aligned}\mathbf{f} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \mathbf{x} &\longmapsto \mathbf{A}\mathbf{x} + \mathbf{b}\end{aligned}$$

Affine functions are often casually referred to as linear, but they are not linear unless  $\mathbf{b} = \mathbf{0}$  because superposition does not hold in general.

---

## Systems of Linear Equations

Define a linear matrix-vector multiplication function as

$$\begin{aligned}\mathbf{f} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \mathbf{x} &\longmapsto \mathbf{Ax}\end{aligned}$$

That is  $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$  where  $\mathbf{A}$  is a  $m \times n$  matrix.

One of the most important applications of linear vector functions is to solve linear systems of equations.

Consider a system of  $m$  equations in  $n$  variables  $x_0, x_1, \dots, x_{n-1}$ :

$$\begin{aligned}A_{00}x_0 + A_{01}x_1 + \cdots + A_{0,n-1}x_{n-1} &= b_0 \\ A_{10}x_0 + A_{11}x_1 + \cdots + A_{1,n-1}x_{n-1} &= b_1 \\ &\vdots \\ A_{m-1,0}x_0 + A_{m-1,1}x_1 + \cdots + A_{m-1,n-1}x_{n-1} &= b_{m-1},\end{aligned}$$

where

- $A_{ij}$  are numbers and are called *coefficients*, and
- $b_j$  are numbers called *right-hand sides*

Then the system of linear equations can be written concisely in matrix notation as

$$\mathbf{Ax} = \mathbf{b},$$

where  $\mathbf{x} = [x_0, x_1, x_{n-1}]^T$  is the vector of variables (unknowns).

An  $n$ -vector of values  $\mathbf{x}$  is a solution to the system of linear equations if  $\mathbf{Ax} = \mathbf{b}$  holds.

A system of linear equations may have:

- no solutions
  - a unique solution
  - many solutions
- 
- The set of linear equations is called **over-determined** if  $m > n$ . In this case, the coefficient matrix  $\mathbf{A}$  is *tall*, that is, we have more equations than variables or unknowns.
  - The set of linear equations is called **under-determined** if  $m < n$ . In this case, the coefficient matrix  $\mathbf{A}$  is *wide*, that is, there are more unknowns than equations.
  - The set of linear equations is called **square** if  $m = n$ . In this case, the coefficient matrix  $\mathbf{A}$  is *square*, that is, the numbers of unknowns and equations is the same.

A set of equations with zero right-hand side,  $\mathbf{Ax} = \mathbf{0}$ , is called a **homogeneous** set of equations. **Any homogeneous set of equations has  $\mathbf{x} = \mathbf{0}$  as a solution.**

Consider the set of equations

$$\begin{cases} 3x_1 &= 30 \\ x_1 + 2x_2 &= 18 \\ x_2 - x_3 &= 2 \end{cases} \iff \begin{cases} 3x_1 + 0x_2 + 0x_3 &= 30 \\ 1x_1 + 2x_2 + 0x_3 &= 18 \\ 0x_1 + 1x_2 - 1x_3 &= 2 \end{cases}$$

In matrix form, we would write

$$\begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 30 \\ 18 \\ 2 \end{bmatrix}$$

---

## Augmented Matrix

Each system of linear equations can be represented by an **augmented matrix**. From the example above, we have the following augmented matrix:

$$\left[ \begin{array}{ccc|c} 3 & 0 & 0 & 30 \\ 1 & 2 & 0 & 18 \\ 0 & 1 & -1 & 2 \end{array} \right]$$

## Operations on Augmented Matrices

There are three elementary row operations we can do on augmented matrices:

### Augmented Matrices Operations

1. **Swap**: exchange any two rows of the augmented matrix (example: swap row 2 with row 3,  $R2 \leftrightarrow R3$ )
2. **Scale**: multiply any row by a non-zero constant (example: scale row 1 by  $\frac{1}{3}$ ,  $R1 \leftarrow \frac{1}{3}R1$ )
3. **Pivot**: add a multiple of a row to another row (example:  $R1 \leftarrow R1 - 3R2$ )

---

## Row Echelon Form

A matrix is said to be in **row echelon form** if:

1. Each leading entry is equal to 1

2. Each leading entry lies to the right of the leading entry above it (upper-diagonal matrices for square matrices)
3. Rows containing zeros are at the bottom of the matrix

For example,

$$\begin{bmatrix} 1 & -4 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$$

is a matrix in **row echelon form**.

The linear transformation **A** for our affine function does not have to be a square matrix. The following examples represent matrices in row echelon form:

over-determined

$$\begin{bmatrix} 1 & -4 & 0 & 2 \\ 0 & 1 & 5 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

under-determined

$$\begin{bmatrix} 1 & -4 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

We will use the **gaussian elimination** algorithm to reduce the augmented matrix into a row echelon form matrix.

Once a matrix is in row echelon form, in order to solve the system of linear equations, we can:

1. Use back substitution
2. Gauss-Jordan elimination

---

## Gauss-Jordan Elimination

In the Gauss-Jordan Elimination algorithm, the augmented matrix is put in **reduced row echelon form (rref)**.

A matrix is said to be in reduced row echelon form if leading row elements equal to 1 and the elements above and below it are equal to 0. For example:

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right]$$

$$\left[ \begin{array}{cc|c} 1 & 0 & 10 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{array} \right]$$

$$\left[ \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right]$$

$$\left[ \begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & -1 \\ 0 & 1 & -\frac{1}{4} & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Transforming a row echelon form to a reduced row echelon form can be accomplished with only *pivot* operations.

Observations:

- Row echelon form is typically not unique
- Reduced echelon form is unique
- Both echelon and reduced echelon forms describe the original system

## Geometric Interpretation

Example:

$$\begin{cases} x + y = 6 \\ x - y = -2 \end{cases}$$

$$\begin{cases} x - y = -2 \\ x - y = -6 \end{cases}$$

A system of equations has a *unique* solution if all the lines (or hyperplanes, for higher-dimensions) intersect at a common point.

A system of equations has an *infinite* number of solutions if the lines are: (1) parallel or (2) identical.

## Consistent vs Dependent Systems

A system of linear equations can have one solution, an infinite number of solutions, or no solution. Systems of equations can be classified by the number of solutions.

- If a system has *at least one solution*, it is said to be **consistent**.
- If a consistent system has *exactly one solution*, it is **independent**.

- If a consistent system has an *infinite number of solutions*, it is **dependent**. When you graph the equations, both equations represent the same line.
- If a system has *no solution*, it is said to be **inconsistent**. The graphs of the lines do not intersect, so the graphs are parallel and there is no solution.

A system is said to be **independent** if the correspondent augmented matrix has a *unique* solution. This means that the matrix **A** in our affine equation,  $\mathbf{y} = \mathbf{Ax}$  has **independent columns** (or rows) which correspond to independent or **informative features**.

- Example: basis vectors

A system is said to be **dependent** if the correspondent augmented matrix has *infinite* solutions. This means that the matrix **A** in our affine equation,  $\mathbf{y} = \mathbf{Ax}$  has **dependent columns** (or rows). One or more columns are a **linear combination** of other columns. We can think of features to be **correlated**.

The Gaussian-Jordan elimination algorithm will allow us to determine whether our system is **independent** or **dependent**.

## Types of Systems

A system can have:

- Infinite number of solutions
  - Example:

$$\begin{cases} x - y = -2 \\ 2x - 2y = -4 \end{cases}$$

- Single unique solution
  - Example:

$$\begin{cases} x - y = -2 \\ 2x - 2y = -4 \\ x + y = 6 \end{cases}$$

- No solutions
  - Example:

$$\begin{cases} x - y = -2 \\ x - y = 0 \end{cases}$$

### Example 1

Solve the following system of linear equations using Gauss-Jordan elimination:

$$\begin{cases} x + 2y = 1 \\ 2x - y = 7 \end{cases}$$

$$\rightarrow (x, y) = (3, -1)$$

### Example 2

Solve the following system of linear equations using Gauss-Jordan elimination:

$$\begin{cases} 3x + 4y = 4 \\ 6x - 2y = 3 \end{cases}$$

$$\rightarrow (x, y) = (2/3, 1/2)$$

### Example 3

Solve the following system of linear equations using Gauss-Jordan elimination:

$$\begin{cases} -x + y = -1 \\ y - z = 6 \\ x + z = -1 \end{cases}$$

$$\rightarrow (x, y, z) = (3, 2, -4)$$

### Example 4

Solve the following system of linear equations using Gauss-Jordan elimination:

$$\begin{cases} x + 2y - z = 9 \\ 2x - y + 3z = -2 \\ 3x - 3y - 4z = 1 \end{cases}$$

$$\rightarrow (x, y, z) = (2, 3, -1)$$

### Example 5

Solve the following system of linear equations using Gauss-Jordan elimination:

$$\begin{cases} 3x + y + 3z = 1 \\ x + 2y - z = 2 \\ 2x - y + 4z = 4 \end{cases}$$

$\rightarrow$  no solution

### Example 6

The following matrices represent systems of 3 equations with 3 variables. Gauss-Jordan elimination was used to arrive at the given matrices. Express the solution indicated by each matrix in the form (x, y, z) or state that an infinite number or no solution exists.

a.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array} \right]$$

b.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

c.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

d.

$$\left[ \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

e.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

f.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

-->

a. (4, -1, 0)

b. no solution

c. infinite solutions (3-2t, t+5, t)

d. infinite solutions (4-2s-3t, s, t)

e. no solution

f. infinite solutions (2, 3, t)



### Example 7

In a physics lab experiment, students are told that a particle moving in a straight line (but not a constant speed!) moves according to the formula  $s = at^2 + bt + c$  where  $s$  is the particle's distance in feet from a fixed point at time  $t$  seconds after launch. Students gather the following data:

t	s
0	5
1	23
2	37

Find the values for  $a$ ,  $b$ , and  $c$  in the above equation for  $s$ . Then use your answer to find how far the particle is from the fixed point 8 seconds after launch.

-->  $(a,b,c) = (-2, 20, 5)$ , so  $s = -2t^2 + 20t + 5$ . When  $t = 8$ ,  $s = 37$  feet.

### Example 8

Consider the following system of linear equations:

$$\begin{cases} 3x = 30 \\ x + 2y = 18 \\ y - z = 2 \end{cases}$$

```
In [ ]: import numpy as np
import numpy.linalg as la
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

**Observation:** All columns (and rows) of the reduced row echelon form (rref) matrix are independent (the inner product is 0, which means they are orthogonal). We have a unique solution.

We are interested in finding out whether the linear transformation  $\mathbf{A}$  will *preserve* all dimensions  $\iff$  linearly independent columns  $\iff$  the system has a unique solution.

In fact, we know that intersections between two lines (or hyperplanes) can happen in any of three different ways:

1. the lines intersect at a unique point (i.e., solution exists and is unique),

2. the lines are coincident (that is, the equations represent the same line and there are infinitely many points of intersection; in this case a solution exists, but is not unique), or
3. the lines are parallel but not coincident (so that no solution exists).