

# Summary of Day 3

William Gunther

May 21, 2014

## 1 Objectives

- Use Gauss-Jordan elimination to solve a system of linear equations.
- Restate problem of system of equations into a problem of vector equations.
- Explore the geometry and algebraic properties of vectors in  $R^n$  (particularly  $R^2$  and  $R^3$  since those are easy to visualize).

## 2 Summary

- The reduced row echelon form of a matrix is a specification of the row echelon form a matrix. It has the following advantage:

**Theorem** Every matrix is row equivalent to a unique matrix in reduced row echelon form.

- With regards to solving systems of linear equations, reduced row echelon form offers another advantage in that a matrix in rref can be solved without the final step of back substitution. This algorithm is called **Gauss-Jordan Elimination** and is very similar to Gaussian Elimination. These are the steps:
  1. Express the system of equations as an augmented matrix.
  2. Use elementary row operations to find a row equivalent matrix in reduced row echelon form.
  3. Solve the variables in the columns with leading entries in terms of free variables.

**Example** Consider the following system of equations:

$$-x_1 + 3x_2 - 2x_3 + 4x_4 = 2$$

$$2x_1 - 6x_2 + x_3 - 2x_4 = -1$$

$$x_1 - 3x_2 + 4x_3 - 8x_4 = -4$$

$$\begin{aligned} \left( \begin{array}{cccc|c} -1 & 3 & 2 & 4 & 2 \\ 2 & 6 & 1 & -2 & -1 \\ 1 & -3 & 4 & 8 & -4 \end{array} \right) & \xrightarrow{\substack{R_3+R_1 \\ R_2+2R_2}} & \left( \begin{array}{cccc|c} -1 & 3 & 2 & 4 & 2 \\ 0 & 12 & 5 & 6 & 3 \\ 0 & 0 & 6 & 12 & -2 \end{array} \right) \\ & \xrightarrow{\substack{R_1-1/4R_2 \\ -R_1 \quad 1/12R_2}} & \left( \begin{array}{cccc|c} 1 & 0 & -3/4 & -5/2 & -5/4 \\ 0 & 1 & 5/12 & 1/2 & 1/4 \\ 0 & 0 & 6 & 12 & -2 \end{array} \right) \\ & \xrightarrow{\substack{R_1+1/8R_3 \\ R_2-5/72R_3 \quad 1/6R_3}} & \left( \begin{array}{cccc|c} 1 & 0 & 0 & -1 & -3/2 \\ 0 & 1 & 0 & -1/3 & 7/18 \\ 0 & 0 & 1 & 2 & -1/3 \end{array} \right) \end{aligned}$$

This is in reduced row echelon form. Now every column without a leading entry is still a free variable (after all, it is in row echelon form as well), therefore we set  $t := x_4$ . But, there is no back substitution as if we solve for the leading entries it is already in terms of the free variables:

$$\begin{aligned}x_1 &= -3/2 + t \\x_2 &= 7/18 + 1/3t \\x_3 &= -1/3 - 2t \\x_4 &= t\end{aligned}$$

or, in vector notation:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3/2 + t \\ 7/18 + 1/3t \\ -1/3 - 2t \\ t \end{pmatrix} = \begin{pmatrix} -3/2 \\ 7/18 \\ -1/3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1/3 \\ 2 \\ 1 \end{pmatrix}$$

**Remark** When you're asked to solve a system, it doesn't really matter which algorithm you use. Personally, I prefer Gauss-Jordan if I want to find the solution. If you just want to analyze the solutions (whether there are any, and if so are there infinitely many, and if so how many free variables are there) then doing Gaussian Elimination to get the matrix in row echelon form will be quicker.

- We will now rephrase the task of solving a linear equation into vector problem. But first we need to know a bit about vectors and vector operations.
- In  $\mathbb{R}^2$  vectors can be represented on the plane as directed line segments. And in  $\mathbb{R}^3$  they can be represented in 3 space similarly. Usually, by convention, we draw vectors with one end starting at the origin and pointing out to the point represented by the  $n$ -tuple, but we view all directed line segments of the same length and direction as the same vector regardless of starting and ending points. When we draw it starting at the origin we call that **standard position**.
- There is a special vector which we denote  $\mathbf{0}$  (or on the chalkboard  $\bar{0}$ ) which is called the **zero vector**. It points in no direction, has no length, and has other algebraic properties which we shall soon discuss.
- Thinking of applications to vectors in other fields, in physics vectors are used to represent various physical quantities which are not easily expressed just as real numbers (such as force, velocity, and displacement).
- Algebraically, we like to think of vectors just as some  $n$ -tuple of real number of data. This makes the following definitions pretty easy to swallow:
- We define an operation called **vector addition** between  $\mathbf{v}$  and  $\mathbf{w}$  of  $\mathbb{R}^n$ . This operation is defined by *componentwise addition*; ie, if  $\mathbf{v} = [v_1, \dots, v_n]$  and  $\mathbf{w} = [w_1, \dots, w_n]$  then:

$$\mathbf{v} + \mathbf{w} = [v_1 + w_1, \dots, v_n + w_n]$$

Geometrically, this operation corresponds to the **parallelogram rule**; put the tail of one of the vectors to the tip of the other, and vice versa, and draw the diagonal of the parallelogram created; this is the sum of the two vectors.

**Remark** For a hint on why this is useful, think of vectors as displacement.

- We define an operation between real numbers  $\mathbb{R}$  (which we call in this context **scalars**) and  $\mathbb{R}^n$  called **scalar multiplication**. This operation is defined by *componentwise multiplication*; ie, if  $\mathbf{v} = [v_1, \dots, v_n]$  and  $c \in \mathbb{R}$  then:

$$c\mathbf{v} = [cv_1, \dots, cv_n]$$

**Remark**  $\mathbb{R}^n$  is what is called a **vector space**; we'll talk more about this in the future.  $\mathbb{R}$  is an algebraic structure called a **field**. Every vector space comes equipped with a field, which is where the scalars live. In this context, you might hear  $\mathbb{R}^n$  called a vector space **over** the field  $\mathbb{R}$ .

Geometrically, scalar multiplication is stretching and shrinking the vector.

- Let's now investigate some properties of these two operations on vectors and how they interact with each other. For these statements, let  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  be vectors, and  $c, d \in \mathbb{R}$  are scalars; note that 0 and  $\mathbf{0}$  are different things.

**Theorem**

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.  $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$
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.  $1\mathbf{u} = \mathbf{u}$

*Proof.* We'll just prove the 6th one.  $\mathbf{u} = [u_1, \dots, u_n]$ . Then:

$$(c+d)\mathbf{u} = [(c+d)u_1, \dots, (c+d)u_n] = [cu_1 + du_1, \dots, cu_n + du_n] = [cu_1, \dots, cu_n] + [du_1, \dots, du_n] = c\mathbf{u} + d\mathbf{u}$$

- A vector  $\mathbf{v}$  is a **linear combination** of vectors  $\mathbf{v}_1, \dots, \mathbf{v}_m$  if there are scalars  $c_1, \dots, c_m$  such that:

$$c_1\mathbf{v}_1 + \dots + c_m\mathbf{v}_m = \mathbf{v}$$

**Example**  $[1, 2]$  is a linear combination of  $[1, 3]$  and  $[1, 1]$ :

$$1/2[1, 3] + 1/2[1, 1] = [1, 2]$$

- We can now rephrase the question of finding solutions to a system of equations to finding out if a vector is a linear combination of some other vectors.

Consider the system:

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= c_1 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= c_m \end{aligned}$$

Any solution to this equation will also satisfy:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix} + \dots + x_m \begin{pmatrix} a_{1m} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Therefore, the question to determine if a system has a solution (and to find solutions) is the same as the question to determine if a vector is a linear combination of other vectors (and to write it as such).