DAY 6: SEPTEMBER 3rd.

4 The Matrix Equation $A\vec{x} = \vec{b}$

From the previous sections, we have seen three different ways of viewing a linear system. First, it can be viewed as a system of linear equations:

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 4 \\ x_1 + 2x_2 - x_3 &= -1 \\ 2x_2 + 4x_3 &= 2 \end{aligned}$$

Second, it can be viewed as an augmented matrix:

$$\begin{bmatrix} 2 & -1 & 3 & 4 \\ 1 & 2 & -1 & -1 \\ 0 & 2 & 4 & 2 \end{bmatrix}$$

And third, we can express it as vector equation:

$$x_1 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$$

In this section, we will see another way to express a linear system.

Definition 4.1. If a matrix A is an $m \times n$ matrix with columns $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ and if $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, then the product of A and $\vec{x}$, denoted by $A\vec{x}$, is defined by:

$$A\vec{x} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

In particular, $A\vec{x}$ is the linear combination of the columns of A using the entries of $\vec{x}$ as weights.

Example 4.2.

(a) Let $A = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 2 & 1 \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}$. Then we have:

$$A\vec{x} = -3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \end{bmatrix} + \begin{bmatrix} -4 \\ 8 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 9 \end{bmatrix}.$$

(b) Let $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^n$. To write the linear combination $\vec{v}_1 - 3\vec{v}_2 + 2\vec{v}_3$ as a matrix times a vector, we place $\vec{v}_1, \vec{v}_2, \vec{v}_3$ as columns of A and the weights $1, -3, 2$ as the entry of $\vec{x}$. We obtain:

$$\vec{v}_1 - 3\vec{v}_2 + 2\vec{v}_3 = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = A\vec{x}.$$

Note:

- If A is $m \times n$ matrix and if $\vec{x} \in \mathbb{R}^n$, then $A\vec{x} \in \mathbb{R}^m$.
- The product $A\vec{x}$ is defined only if the number of columns of A equals to number of entries of $\vec{x}$.

Matrix Equation and Linear System

Let A be an $m \times n$ matrix, $\vec{b} \in \mathbb{R}^m$, and $x_1, x_2, \ldots, x_n$ are variables. Denote $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. The equation $A\vec{x} = \vec{b}$ is called a matrix equation.

By the definition of the product $A\vec{x}$, we can write a matrix equation as a vector equation, an then as a linear system. Conversely, we can write a linear system as a matrix equation $A\vec{x} = \vec{b}$ by defining the matrix A as the coefficient matrix of the linear system and the vector $\vec{b}$ as the constant terms.

Example 4.3. Consider the following linear system:

$$\begin{aligned} x_1 - 3x_2 + x_3 &= 5 \\ -x_2 + 4x_3 &= -1. \end{aligned}$$

The corresponding matrix equation is

$$\begin{bmatrix} 1 & -3 & 1 \\ 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

Theorem 4.4. If A is an $m \times n$ with columns $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ and $\vec{b} \in \mathbb{R}^m$, the matrix equation $A\vec{x} = \vec{b}$ has the same set of solutions as the vector equation:

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b},$$

which is in turn has the same solution set as the linear system with augmented matrix

$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n & \vec{b} \end{bmatrix}$$

Example 4.5. Let $A = \begin{bmatrix} 1 & 3 & -4 \\ 1 & 5 & 2 \\ -3 & -7 & 18 \end{bmatrix}$. For which vector $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ is the equation $A\vec{x} = \vec{b}$ consistent?

To determine the vectors $\vec{b}$ for which the system is consistent, we reduce the corresponding augmented matrix to echelon form. We have

$$\begin{bmatrix} 1 & 3 & -4 & b_1 \\ 1 & 5 & 2 & b_2 \\ -3 & -7 & 18 & b_3 \end{bmatrix} \xrightarrow[R_3+3R_3 \to R_3]{R_2 - R_1 \to R_2} \begin{bmatrix} 1 & 3 & -4 & b_1 \\ 1 & 2 & 6 & b_2 - b_1 \\ 0 & 2 & 6 & b_3 + 3b_1 \end{bmatrix} \xrightarrow{R_3 - R_2 \to R_3} \begin{bmatrix} 1 & 3 & -4 & b_1 \\ 0 & 2 & 6 & b_2 - b_1 \\ 0 & 0 & 0 & b_3 - b_2 + 4b_1 \end{bmatrix}$$

We see that the matrix equation $A\vec{x} = \vec{b}$ is consistent if $b_3 - b_2 + 4b_1 = 0$. Therefore, the equation $A\vec{x} = \vec{b}$ is not consistent for every vector $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. The entries b_1, b_2, b_3 must satisfy the relation $b_3 - b_2 + 4b_1 = 0$. For example, $A\vec{x} = \vec{b}$ is inconsistent for $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ sicne $2 - 1 + 4 \times 1 \neq 0$. And, the equation $A\vec{x} = \vec{b}$ has solution for $\vec{b} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ since $1 - 1 + 4 \times 0 = 0$.

Recall that a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ spans (or generates) $\mathbb{R}^n$ (that is $\text{Span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p) = \mathbb{R}^n$) if every vector in $\mathbb{R}^n$ is a linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$.

Theorem 4.6. Let A be an $m \times n$ matrix. The following statement are equivalent (i.e. they are all true or they are all false):

(a) For each $\vec{b} \in \mathbb{R}^m$, the equation $A\vec{x} = \vec{b}$ has a solution.

(b) Each $\vec{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.

(c) The columns of A span $\mathbb{R}^m$.

(d) In the row reduction process, A has a pivot position in every row (note that here we refer to the matrix A which is the coefficients matrix of the corresponding linear system, not to the augmented matrix).

Computation of $A\vec{x}$

The following is a simple and efficient method for calculating the entries of $A\vec{x}$.

Row-vector Rule:

If the product $A\vec{x}$ is defined, then the ith entry in $A\vec{x}$ is the sum of the products of the corresponding entries from ith row of A and from the vector $\vec{x}$. That is

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{bmatrix}$$

and $b_i = \sum_{j=1}^{m} a_{ij}x_j = a_{i1}x_1 + a_{i1}x_2 + \cdots + a_{in}x_n$.

Example 4.7.

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

Theorem 4.8. If A is an $m \times n$ matrix, $\vec{u}$ and $\vec{v} \in \mathbb{R}^n$, and $c \in \mathbb{R}$, then:

(a) $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$.

(b) $A(c\vec{u}) = c(A\vec{u})$.

5 Solution Sets of Linear Systems

In this section, we use vector notation to give explicit and geometric description of the solution set of a given linear system.

Definition 5.1. A system of linear equation is called homogeneous if it can be written as $A\vec{x} = \vec{0}$.

Example 5.2. The following linear system is homogeneous:

$$x_1 - 3x_2 + 3x_3 = 0$$

$$2x_1 + x_2 - 4x_3 = 0$$

Note: Clearly, the zero vector $\vec{x} = \vec{0}$ is always a solution of $A\vec{x} = \vec{0}$. It is called the **trivial solution.**. The most important question is whether there exists a nontrivial solution, that is a vector $\vec{x} \neq \vec{0}$ that satisfies $A\vec{x} = \vec{0}$.

Fact: The homogeneous system $A\vec{x} = \vec{0}$ has a nontrivial solution if and only if it has a free variable.

Example 5.3. Consider the following homogeneous linear system:

$$x_1 - 2x_2 + 3x_3 = 0$$

$$-2x_1 - 3x_2 - 4x_3 = 0$$

$$2x_1 - 11x_2 + 8x_3 = 0$$

The corresponding matrix equation is:

$$\begin{bmatrix} 1 & -2 & 3 \\ -2 & -3 & -4 \\ 2 & 11 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

To determine if the system has a nontrivial solution, we do row reduction on the augmented matrix to see if there is a free variable. The augmented matrix is given by:

$$\begin{bmatrix} 1 & -2 & 3 & 0 \\ -2 & -3 & -4 & 0 \\ 2 & 11 & 8 & 0 \end{bmatrix}.$$

Applying $R_2 + 2R_1 \to R_2$ and $R_3 - 2R_1 \to R_3$, we have:

$$\begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & -7 & 2 & 0 \\ 0 & -7 & 2 & 0 \end{bmatrix}.$$

Now applying $R_3 + R_2 \to R_3$, we have

$$\begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & -7 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is now in row echelon form and we can see that x_3 does not correspond to any pivot so it is a free variable. It follows that the homogeneous system has a nontrivial solution.1

Parametric Vector Form

Suppose that after row reduction, the augmented matrix of an homogeneous system is given by:

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It follows that x_2 and x_3 are free variables. The corresponding equations are given by:

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ x_2 & \text{ free} \\ x_3 & \text{ free} \end{aligned}$$

A a vector, our general solution $\vec{x}$ is given by

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2x_3 \\ 0 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

In particular, since x_2 and x_3 can be chosen to any number, a solution is a linear combination of the vectors: $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ Therefore, the solution set of the linear system in this particular example is $\text{Span}(\vec{v}_1, \vec{v}_2)$. Since neither of $\vec{v}_1$ nor $\vec{v}_2$ is a scalar multiple of of the other, $\text{Span}(\vec{v}_1, \vec{v}_2)$ is a plan through the origin. We can set x_2 and x_3 as parameters r and t respectively, and we obtain the parametric vector form of the general solution.

Note: The solution set of $A\vec{x} = \vec{0}$ can always be written as $\text{Span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p)$ for some vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$. To give an explicit description of the solutions, we use parametric vector form.

Definition 5.4. An equation of the form

$$\vec{x} = t_1\vec{v}_1 + t_2\vec{v}_2 + \cdots + t_p\vec{v}_p$$

is called **parametric vector equation**.

Example 5.5. For the above example, we set $x_2 = r$ and $x_3 = t$ and then we have the parametric vector equation of the solution as

$$\vec{x} = r \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = r\vec{v}_1 + t\vec{v}_2$$

Solutions of Nonhomogeneous Systems

Suppose now that the system in the above example is not homogeneous (it has some nonzero constants, i.e. we have the form $A\vec{x} = \vec{b}$). In addition, suppose that after row reduction to an echelon form, the augmented matrix is given by:

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The corresponding equations are given by

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ x_2 &\quad \text{free} \\ x_3 &\quad \text{free} \end{aligned}$$

In particular, the system has infinitely many solutions. Hence, as a vector, the general solution of $A\vec{x} = \vec{b}$ in this case is given by

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 + x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2x_3 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

We have seen that

$$x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = r\vec{v}_1 + t\vec{v}_2$$

is the parametric expression of the homogeneous equation. Therefore, we may write $\vec{x} = \vec{p} + r\vec{v}_1 + t\vec{v}_2$ where $\vec{p}$ is a specific solution to the system and $r\vec{v}_1 + t\vec{v}_2$ is the general form of a solution of the homogeneous system.

Theorem 5.6. Suppose that the equation $A\vec{x} = \vec{b}$ has a solution $\vec{p}$. Then all solutions to the equation have the form:

$$\vec{w} = \vec{p} + \vec{v}_h$$

where $\vec{v}_h$ is a solution of the corresponding homogeneous equation $A\vec{x} = \vec{0}$.