LECTURE 6

Systems of Linear Equations

You may recall that in Math 3013, matrices were first introduced as a means of encapsulating the essential data underlying a system of linear equations; that is to say, given a set of n linear equations in m variables

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

one could arrange coefficients a_{ij} of the variables x_j into a rectangular array with n rows and m columns

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

arrange the numbers on the right hand side as an array with one column

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

In fact, once matrix multiplication was been defined it was possible to replace the original set of linear with a single matrix equation

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

where $\mathbf{x}$ is $m \times 1$ matrix holding the variables $x_1, \ldots, x_m$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

But you'll note, in this course, I have thus far avoided any connections to systems of linear equations My reason for this is to constantly stress the abstract vector space point of view. However, now having beat that horse to death, I think we can safely look at a principal application of linear algebra without abandoning our abstract perspective.

TERMINOLOGY 6.1. Let $\mathbb{F}$ be a field. An $\mathbf{n \times m}$ **linear system over** $\mathbb{F}$ is a collection of n linear equations in m unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

where each coefficient $a_{ij} \in \mathbb{F}$ and each $b_i \in \mathbb{F}$. Such a system will be called solvable if there exists a choice of field elements $x_1, \ldots, x_n \in \mathbb{F}$ such that each equation is satisfied. The field elements a_{ij} ,

$1 \leq i \leq n$, $1 \leq j \leq m$ can be arranged in an $n \times m$ matrix with entries in $\mathbb{F}$:

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

This matrix is called the **coefficient matrix** of the linear system. The field elements $b_1, \ldots, b_n$ can be arranged in a $n \times 1$ column vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

This column vector (for lack of a better standard terminology) will be referred to as the **inhomogeneous part** of the system. If it so happens that each $b_i = 0_{\mathbb{F}}$, $1 \leq i \leq n$, then we say that the linear system is **homogeneous**.

THEOREM 6.2. *Consider a $n \times m$ linear system with coefficient matrix $\mathbf{A}$ and inhomogenous part $\mathbf{b} \in \mathbb{F}^n$. For each i between 1 and n, let $\mathbf{c}_i$ denote the element of $\mathbb{F}^n$ formed by writing the entries in the i^{th} column of $\mathbf{A}$ in order (from top to bottom). Then the linear system has a solution if and only if either of the following two conditions is satisfied.*

(i) $\mathbf{b} \in span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m\right)$
(ii) $\dim span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m\right) = \dim span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m, \mathbf{b}\right)$

Proof. Suppose $\mathbf{b}$ lies in the span of $\mathbf{c}_1, \ldots, \mathbf{c}_m$ then there exists elements $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ such that

$$\mathbf{b} = \lambda_1 \mathbf{c}_1 + \cdots + \lambda_m \mathbf{c}_m$$

Both sides of this equation are vectors in $\mathbb{F}^n$. Writing this vector equation component by component

$$\begin{aligned} b_1 &= (\mathbf{b})_1 = \lambda_1 (\mathbf{c}_1)_1 + \cdots + \lambda_m (\mathbf{c}_m)_1 = \lambda_1 a_{11} + \cdots \lambda_m a_{1m} \\ &\vdots \\ b_n &= (\mathbf{b})_n = \lambda_1 (\mathbf{c}_1)_n + \cdots + \lambda_m (\mathbf{c}_m)_n = \lambda_1 a_{1n} + \cdots + \lambda_m a_{nm} \end{aligned}$$

we see $x_1 = \lambda_1, \ldots, x_m = \lambda_m$ furnishes us with a solution of the linear system. On the other hand, if $x_1, \ldots, x_m$ is a solution to the linear system, we can reverse the above argument and regard the original linear system as stipulating a relationship between the entries in the columns of $\mathbf{A}$, the numbers $x_1, \ldots, x_m$ and a vector $\mathbf{b} \in \mathbb{F}^n$. Interpreted this way, the relationship simply express the vector $\mathbf{b}$ as a linear combination of the column vectors of $\mathbf{A}$.

To prove the equivalence of (i) and (ii) we note first that (i) immediately imples (ii). To see that (ii) also implies (i), Suppose that $\dim span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m\right) = \dim span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m, \mathbf{b}\right)$. By Theorem 5.3, pick a subset $\{\mathbf{c}'_1, \ldots, \mathbf{c}'_k\}$ of the column vectors $\{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$ that forms a basis for $span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m\right)$. Then $\mathbf{b} \in span\left(\mathbf{c}'_1, \ldots, \mathbf{c}'_k\right)$ if and only if $\mathbf{b} \in span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m\right)$. Now if $\mathbf{b} \notin span\left(\mathbf{c}'_1, \ldots, \mathbf{c}'_k\right)$, then the can be no dependence relation amongst the vectors $\{\mathbf{c}'_1, \ldots, \mathbf{c}'_k, \mathbf{b}\}$, hence $\{\mathbf{c}'_1, \ldots, \mathbf{c}'_k, \mathbf{b}\}$ would be a linearly independent set, hence

$$k = \dim span\left\{\mathbf{c}'_1, \ldots, \mathbf{c}'_k\right\} = \dim span\left\{\mathbf{c}_1, \ldots, \mathbf{c}_m\right\}$$

but

$$k + 1 = \dim span\left(\mathbf{c}'_1, \ldots, \mathbf{c}'_k, \mathbf{b}\right) = \dim span\left(\mathbf{c}_1, \ldots, \mathbf{c}_m, \mathbf{b}\right)$$

which contradicts our hypothesis. $\square$

EXAMPLE 6.3. Determine if the linear system

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

has a solution.

- We have

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$$

By the preceding theorem, we just to determine if $\mathbf{b}$ lies in the span of

$$\mathbf{c}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \quad , \quad \mathbf{c}_2 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{c}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

First we use row reduction to obtain a basis for $span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$. The coefficient matrix of the vectors $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ with respect to the standard basis of $\mathbb{R}^3$ is

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

This matrix can be row reduced to

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

and so $[1, 2, 1]$ and $[0, 1, 1]$ provide a basis for $span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$. So $\dim span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3) = 2$.

Let us similarly compute the dimension of $span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{b})$

$$\begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & 2 \\ -1 & 0 & 1 \\ 1 & 0 & 3 \end{bmatrix} \xrightarrow{\text{row reduction}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and so $\dim span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{b}) = 3$.

Since $\dim span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3) = 2 \neq 3 = \dim span\,(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{b})$, we conclude that the given linear system has no solution.

Let me now introduce some standard terminology that allows a more succinct expression of the result of Theorem 7.2

DEFINITION 6.4. *The **rank** of a linear system (or of its associated coefficient matrix) is the dimension of the column space of its coefficient matrix.*[1]

DEFINITION 6.5. *Let $\mathbf{A}$ is the coefficient matrix for a $n \times m$ linear system*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

*and let $\mathbf{b}$ is the inhomogeneous part of the same system. The **augmented matrix** for this system is the $n \times (m+1)$ matrix*

$$[\mathbf{A} \mid \mathbf{b}] \equiv \begin{bmatrix} a_{11} & \cdots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} & b_n \end{bmatrix}$$

THEOREM 6.6. *A linear system has a solution if and only if rank of its coefficient matrix equals the rank of its augmented matrix.*

[1]The column space of a matrix is, of course, just the span of its column vectors.

Thus far, we have only developed methods for checking as to whether or not a given linear system has solutions. Before actually working out some methods of solution, let me first give one nice structural theorem about the solutions of such a system.

Recall that a linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

is called *homogeneous* if each b_i, $1 \leq i \leq n$ is equal to $0_{\mathbb{F}}$; otherwise, it is called *non-homogeneous*.

Homogeneous linear systems have a very useful property:

THEOREM 6.7. *Let $S(\mathbf{A}, \mathbf{0})$ be a homogeneous $n \times m$ linear system. Then the solution set of $S(\mathbf{A}, \mathbf{0})$ is a subspace of $\mathbb{F}^m$.*

Let $[\mathbf{c}_1, \ldots, \mathbf{c}_m]$ be the columns of $\mathbf{A}$. By the proof of Theorem 6.2, If $\mathbf{x}, \mathbf{y} \in \mathbb{F}^m$ are solutions of $S(\mathbf{A}, \mathbf{0})$, then

$$\begin{aligned} x_1\mathbf{c}_1 + \cdots + x_m\mathbf{c}_m &= \mathbf{0} \\ y_1\mathbf{c}_1 + \cdots + y_m\mathbf{c}_m &= \mathbf{0} \end{aligned}$$

Now consider a linear combination of $\mathbf{x}$ and $\mathbf{y}$, $\alpha\mathbf{x} + \beta\mathbf{y}$. We have

$$\begin{aligned} (\alpha\mathbf{x} + \beta\mathbf{y})_1 \mathbf{c}_1 + \cdots + (\alpha\mathbf{x} + \beta\mathbf{y})_m \mathbf{c}_m &= \alpha(x_1\mathbf{c}_1 + \cdots + x_m\mathbf{c}_m) + \beta(y_1\mathbf{c}_1 + \cdots + y_m\mathbf{c}_m) \\ &= \alpha \cdot \mathbf{0} + \beta \cdot \mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

And so every linear combination of $\mathbf{x}$ and $\mathbf{y}$ is also a solution of $S(\mathbf{A}, \mathbf{0})$. Hence, the solution set of $S(\mathbf{A}, \mathbf{0})$ is a subspace of $\mathbb{F}^m$. $\square$

Before discussing the solutions of non-homogeneous linear systems, let me introduce a simple geometric construction.

DEFINITION 6.8. *Let $\mathbf{p}_0$ be an element of a vector space V and let S be subspace of V. The **hyperplane through** $\mathbf{p}_0$ **generated by** S is the set*

$$H_{\mathbf{p}_0,S} = \{\mathbf{v} \in V \mid v = \mathbf{p}_0 + \mathbf{s} \quad ; \quad \mathbf{s} \in S\}$$

REMARK 6.9. A special case is when $S = span(\mathbf{d})$. In this case, $H_{\mathbf{p}_0,S}$ is called the **line** through $\mathbf{p}_0$ in the direction of $\mathbf{d}$.

REMARK 6.10. Hyperplanes are subsets of V, but they usually are **not** subspaces. In fact, later we shall prove

$$H_{\mathbf{p}_0,S} \text{ is a subspace} \quad \Longleftrightarrow \quad \mathbf{p}_0 \in S \quad .$$

Back to linear systems.

If we are given a non-homogenous linear system, we can always attach to it a corresponding homogeneous system by considering the homogeneous linear system with exactly the same coefficient matrix as the original linear system.

THEOREM 6.11. *Suppose $S(\mathbf{A}, \mathbf{b})$ is a $n \times m$ linear system with coefficient matrix $\mathbf{A}$ and inhomogeneous part $\mathbf{b}$ and $S(\mathbf{A}, \mathbf{0})$ is the corresponding homogeneous linear system. Then if $\mathbf{x} = [x_1, \ldots, x_m] \in \mathbb{F}^m$ is a solution to $S(\mathbf{A}, \mathbf{b})$ then any other solution $\mathbf{x}'$ of $S(\mathbf{A}, \mathbf{b})$ can be expressed as*

$$\mathbf{x}' = \mathbf{x} + \mathbf{x}_0$$

where $\mathbf{x}_0$ *is a solution to* $S(\mathbf{A},\mathbf{0})$.

Proof. Let $\mathbf{c}_1,\ldots,\mathbf{c}_m$ be the column vectors of $\mathbf{A}$. Suppose $\mathbf{x}$ and $\mathbf{x}'$ are two solutions of $S(\mathbf{A},\mathbf{b})$. From the proof of Theorem 7.2, this means

$$\begin{aligned} x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_m\mathbf{c}_m &= \mathbf{b} \\ x_1'\mathbf{c}_1 + x_2'\mathbf{c}_2 + \cdots + x_m'\mathbf{c}_m &= \mathbf{b} \end{aligned}$$

If we subtract the first equation from the second we get

$$(x_1' - x_1)\,\mathbf{c}_1 + (x_2' - x_2)\,\mathbf{c}_2 + \cdots + (x_m' - x_m)\,\mathbf{c}_m = \mathbf{0}_{\mathbb{F}^n}$$

which says that $\mathbf{x}' - \mathbf{x}$ is a solution of $S(\mathbf{A},\mathbf{0})$. If we denote this solution by $\mathbf{x}_0$ we have

$$\mathbf{x}' - \mathbf{x} = \mathbf{x}_0 \quad \Rightarrow \quad \mathbf{x}' = \mathbf{x} + \mathbf{x}_0 \quad .$$

I'll now rephrase this result in a manner similar to the way the solutions of a linear differential equation are typically described.

Corollary 6.12. *Let* $\mathbf{x}$ *be a solution of an* $n \times m$ *linear system* $S(\mathbf{A},\mathbf{b})$, *and let* S *be the solution set of the corresponding homogeneous linear system* $S(\mathbf{A},\mathbf{0})$. *Then the solution set of* $S(\mathbf{A},\mathbf{b})$ *coincides with the hyperplane through* $\mathbf{x}$ *generated by* S.

Proof. Let $\mathbf{x}$ be given as above and let $\mathbf{y}$ be an arbitrary solution of $S(\mathbf{A},\mathbf{b})$. By the preceding theorem,

$$\begin{aligned} \mathbf{y} - \mathbf{x} &= \mathbf{s} \text{ for some } \mathbf{s} \in S(\mathbf{A},\mathbf{0}) \\ &\Rightarrow \quad \mathbf{y} = \mathbf{x} + \mathbf{s} \text{ for some } \mathbf{s} \in S \\ &\Rightarrow \quad \mathbf{y} \in H_{\mathbf{x},S} \end{aligned}$$

On the other hand, it is easy to see that if $\mathbf{y} \in H_{\mathbf{x},S}$ then $\mathbf{y}$ is a solution of $S(\mathbf{A},\mathbf{b})$. For

$$\begin{aligned} y_1\mathbf{c}_1 + \cdots + y_m\mathbf{c}_m &= (\mathbf{x}+\mathbf{s})_1\,\mathbf{c}_1 + \cdots + (\mathbf{x}+\mathbf{s})_m\,\mathbf{c}_m \\ &= (x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_m\mathbf{c}_m) + (s_1\mathbf{c}_1 + s_2\mathbf{c}_2 + \cdots + s_m\mathbf{c}_m) \\ &= \mathbf{b} + \mathbf{0} \\ &= \mathbf{b} \end{aligned}$$

□