Math 313 Lecture #4
§1.4: Matrix Equations

Recall the linear combination of p vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ in $\mathbb{R}^n$ with weights $c_1, c_2, \ldots, c_p$ in $\mathbb{R}$:

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p.$$

We are going to use this linear combination to define a product of an $m \times n$ matrix A with a vector $\vec{x}$ in $\mathbb{R}^n$.

We view the $m \times n$ matrix A as m columns where the j^{th} column $\vec{a}_j$ is a vector in $\mathbb{R}^m$.

The entries of the vector $\vec{x}$ in $\mathbb{R}^n$ we denote by $x_1, x_2, \ldots, x_n$.

We define the **product** of the $m \times n$ matrix A with the vector $\vec{x}$ to be the vector in $\mathbb{R}^m$ determined by

$$A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots x_n\vec{a}_n.$$

This product only makes sense when the number of columns of A equals the number of entries in $\vec{x}$.

We also write the product of A with $\vec{x}$ as

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

Examples. (a) We can compute a product via the linear combination,

$$\begin{bmatrix} 3 & 2 & -4 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = 5\begin{bmatrix} 3 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 2 \end{bmatrix} + 2\begin{bmatrix} -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 15 \\ 5 \end{bmatrix} + \begin{bmatrix} -2 \\ -2 \end{bmatrix} + \begin{bmatrix} -8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 13 \end{bmatrix}.$$

(b) We can rewrite a linear combination as a product,

$$2\vec{a}_1 - 3\vec{a}_2 + 5\vec{a}_3 = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix} = A\vec{x} \text{ for } \vec{x} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}.$$

We should become familiar and adroit at using the linear combination and the product interchangeably. / / / /

Why the product of a matrix with a vector? Because it enables us to view a linear system as a *matrix equation*. Here is how.

The j^{th} equation in a linear system m equations in n variables $x_1, x_2, \ldots, x_n$ has the form

$$a_{i1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j.$$

By forming the vectors

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix},\ \vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \vec{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix},\ \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

in $\mathbb{R}^m$, we recognize that the linear system can be written as the linear combination

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

which becomes the **matrix equation**

$$A\vec{x} = \vec{b}.$$

A solution of $A\vec{x} = \vec{b}$ is a vector $\vec{s}$ in $\mathbb{R}^n$ which satisfies $A\vec{s} = \vec{b}$.

Theorem 3. The matrix equation $A\vec{x} = \vec{b}$ has the same solution set as that of the vector equation $x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}$, which has the same solution set of the linear system whose augmented matrix is

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

We know from our discussion of span, that $A\vec{x} = \vec{b}$ has a solution when $\vec{b}$ is a linear combination of the columns of A.

But it is possible for $A\vec{x} = \vec{b}$ to be consistent for every choice of $\vec{b}$?

Example. For

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix},\ \text{and } \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

is $A\vec{x} = \vec{b}$ consistent?

We answer this question by row reducing the augmented matrix associated to the linear system:

$$\begin{bmatrix} 1 & 3 & 4 & | & b_1 \\ -4 & 2 & -6 & | & b_2 \\ -3 & -2 & -7 & | & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & | & b_1 \\ 0 & 14 & 10 & | & b_2 + 4b_1 \\ 0 & 7 & 5 & | & b_3 + 3b_1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & | & b_1 \\ 0 & 14 & 10 & | & b_2 + 4b_1 \\ 0 & 0 & 0 & | & * \end{bmatrix}$$

where the bottom right entry (listed as $*$) is

$$b_3 + 3b_1 - (1/2)(b_2 + 4b_1) = b_1 - (1/2)b_2 + b_3.$$

For the matrix equation to be consistent requires that this right bottom entry be zero, and so $A\vec{x} = \vec{b}$ is not consistent for every choice of $\vec{b}$.

This says that the columns of A do not span $\mathbb{R}^3$ because not every vector $\vec{b}$ in $\mathbb{R}^3$ can be written as a linear combination of the columns of A. / / / /

This example illustrates important connections between the columns of A (the coefficient matrix, not the augmented matrix) and the vector $\vec{b}$.

Theorem 4. Let A be an $m \times n$ matrix. The following statements are logically equivalent (the truth of any one implies the truth of every other).

a. For each $\vec{b}$ in $\mathbb{R}^m$, the matrix equation $A\vec{x} = \vec{b}$ is consistent.
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. The matrix A has a pivot position in every row.

The proof of this theorem is in the Appendix of this lecture note. You are invited to study it on your own to increase your comprehension of the relations among the concepts in the four statements, and a review of methods of proof (direct and contrapositive).

Computation of $A\vec{x}$. We defined the product of an $m \times n$ matrix A and a vector $\vec{x}$ in $\mathbb{R}^n$ by the linear combination of the columns of A with weights being the entries of $\vec{x}$.

A careful look at this product reveals a row-column pattern for the product.

Example. For

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 2 & 6 & -1 \\ 1 & 4 & -5 \end{bmatrix} \text{ and } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

we have

$$A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 = \begin{bmatrix} 3x_1 - x_2 + 2x_3 \\ 2x_1 + 6x_2 - x_3 \\ x_1 + 4x_2 - 5x_3 \end{bmatrix}.$$

The first entry of $A\vec{x}$ is the first row of A "times" the column vector $\vec{x}$, i.e., the sum of products of the corresponding entries of the first row of A with the entries of the column $\vec{x}$.

The same pattern appears in the second and third entries of $A\vec{x}$.

Example Is there a 3×3 matrix I such that $I\vec{x} = \vec{x}$?

Yes, the **identity matrix**

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

satisfies $I\vec{x} = \vec{x}$ for all $\vec{x}$ in $\mathbb{R}^n$.

Properties of the Matrix-Vector Product $A\vec{x}$. The matrix-vector product enjoys several familiar-looking algebraic properties.

Theorem 5. If A is an $m \times n$ matrix, $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar (in $\mathbb{R}$), then

a. $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$, and
b. $A(c\vec{u}) = c(A\vec{u})$.

Proof. (a) Write $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$ and

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Then by the linear combination definition of the matrix-vector product and the algebraic properties of vectors, we have

$$\begin{aligned} A(\vec{u}+\vec{v}) &= \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{bmatrix} \\ &= (u_1+v_1)\vec{a}_1 + (u_2+v_2)\vec{a}_2 + \cdots + (u_n+v_n)\vec{a}_n \\ &= u_1\vec{a}_1 + u_2\vec{a}_2 + \cdots + u_n\vec{a}_n + v_1\vec{a}_1 + v_2\vec{a}_2 + \cdots + v_n\vec{a}_n \\ &= A\vec{u} + A\vec{v}. \end{aligned}$$

(b) Here, again using the linear combination definition of the matrix-vector product, and algebraic properties of vectors, we have

$$\begin{aligned} A(c\vec{u}) &= \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix} \\ &= (cu_1)\vec{a}_1 + (cu_2)\vec{a}_2 + \cdots + (cu_n)\vec{a}_n \\ &= c\big[u_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + u_n\vec{a}_n\big] \\ &= c(A\vec{u}). \end{aligned}$$

This completes the proof. $\square$

Appendix. Here we give a proof of Theorem 4.

The logical equivalence of parts (a), (b), and (c) follow from the linear combination definition of matrix-vector product $A\vec{x}$ and what it means for a set of vectors to span $\mathbb{R}^m$.

It remains to show that (a) and (d) are logically equivalent.

Let $\vec{b}$ be a vector in $\mathbb{R}^m$.

If U is an echelon form for A, then row reduction of the augmented matrix $\begin{bmatrix} A & | & \vec{b} \end{bmatrix}$ leads to the row equivalent augmented matrix $\begin{bmatrix} U & | & \vec{d} \end{bmatrix}$ for some $\vec{d}$ in $\mathbb{R}^m$.

Suppose (d) is true, that is, that every row of A has a pivot position in it.

Now ask yourself if there can be a pivot position in the augmented part of the augmented matrix $\begin{bmatrix} U & | & \vec{d} \end{bmatrix}$.

The answer is no, because the left most nonzero entry in each nonzero row of $\begin{bmatrix} U & | & \vec{d} \end{bmatrix}$ is the pivot position, and this occurs in the U part of the augmented matrix.

So $A\vec{x} = \vec{b}$ has a solution by the Existence Theorem, and (a) is true, that is we have proved that the truth of (d) implies the truth of (a).

To achieve the converse, i.e., the truth of (a) implies the truth of (d), we use the contrapositive, i.e., we show that the falseness of (d) implies the falseness of (a).

Suppose (d) is false, that is, not every row of A has a pivot position in it.

Then not every row of U has a pivot position in it, and so the last row of U must be all zeros.

Let $\vec{d}$ be a vector with a nonzero entry in its last entry, so that the augmented column of $\begin{bmatrix} U & | & \vec{d} \end{bmatrix}$ is a pivot column.

Thus $U\vec{x} = \vec{d}$ is inconsistent by the Existence Theorem.

Reversing the row operations that reduced $\begin{bmatrix} A & | & \vec{b} \end{bmatrix}$ to $\begin{bmatrix} U & | & \vec{d} \end{bmatrix}$ results in a $\vec{b}$ for which $A\vec{x} = \vec{b}$ is also inconsistent.

Thus $A\vec{x} = \vec{b}$ is not consistent for every $\vec{b}$ in $\mathbb{R}^m$.

We have proved that the falseness of (d) leads to the falseness of (a), and this completes the proof. $\square$