1 Section 1.4. Matrix Equations

1.1 Products of matrices and vectors

Definition 1 *Let* $A = [\vec{a}_1, \vec{a}_2, ..., \vec{a}_n]$ *be a* $m \times n$ *matrix whose ith column is* $\vec{a}_i \in R^m$, $i = 1, 2, ..., n$, *and* $\vec{x}$ *a (unknown) vector in* R^n :

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}$$

. *Notice that the number of columns in* $A =$ *the number of rows in* $\vec{x}$.

The product $A\vec{x}$ *is defined as the linear combination of* $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$ *with the weight* $x_1, x_2, ..., x_n$, *i.e.,*

$$A\vec{x} = [\vec{a}_1, \vec{a}_2, ..., \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i \vec{a}_i = x_1\vec{a}_1 + x_2\vec{a}_2 + ... + x_n\vec{a}_n.$$

If A *is* $1 \times n$ *matrix, i.e.,* $A = [a_1, a_2, ..., a_n]$, *then*

$$A\vec{x} = [a_1 \ a_2 \ ... \ a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1a_1 + x_2a_2 + ... + x_na_n$$

$$= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \textit{(dot product)}$$

Example 2 *(a)*

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1\cdot 4 + 2\cdot 3 + (-1)\cdot 7 \\ 2\cdot 4 + 5\cdot 3 + 3\cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 44 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 2 & -2 & 3 \\ 8 & 0 & 1 \\ -5 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 1\begin{bmatrix} 2 \\ 8 \\ -5 \\ 0 \end{bmatrix} + 3\begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 10 \\ 1 \\ 7 \end{bmatrix}$$

Note that for product (b), one can also redo it using "dot product" as

$$\begin{bmatrix} 2 & -2 & 3 \\ 8 & 0 & 1 \\ -5 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 8 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} -5 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 8 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} -5 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 2 \\ 10 \\ 1 \\ 7 \end{bmatrix}$$

The matrix-vector product has the following **PROPERTIES**:

$$A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}, \quad A(\lambda\vec{x}) = \lambda A\vec{x}$$
$$(A+B)\vec{x} = A\vec{x} + B\vec{x}, \ (\lambda A)\vec{x} = \lambda(A\vec{x})$$

These relations can be summarized so-called linearity principle:

$$(aA+bB)(\lambda\vec{x}+\delta\vec{y}) = a\lambda A\vec{x} + a\delta A\vec{y} + b\lambda B\vec{x} + b\delta B\vec{y}.$$

1.2 Matrix Equations

The system of linear equation, or its equivalent vector equation

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

can now be rewritten as **Matrix equation:**

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

Theorem 3 *Let $A = [\vec{a}_1, \vec{a}_2, ..., \vec{a}_n]$ be a $m \times n$ matrix whose ith column is $\vec{a}_i \in R^m$, $i = 1, 2, ..., n$, and $\vec{x}$ an unknown vector in R^n. The following statements are equivalent:*

1. *For each vector $\vec{b} \in R^m$, the vector equation $A\vec{x} = \vec{b}$ has a solution.*
2. *Each vector $\vec{b} \in R^m$ is a linear combination of $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$, the columns of A.*
3. *Column vectors of A span R^m.*
4. *The matrix A has a pivot position in each row.*

Proof. (1) $\Longrightarrow$ (2) : (1) implies that for each $\vec{b}$, there exists a solution $x_i = \lambda_i$, $i = 1, ..., n$ such that

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + ... + \lambda_n \vec{a}_n = \vec{b}.$$

Therefore, each $\vec{b}$ is a linear combination of $\vec{a}_i$.

(2) $\Longrightarrow$ (3) : Since each vector $\vec{b}$ in R^m is a linear combination, by definition, each $\vec{b}$ in R^m is in the subset spanned by $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$, i.e., $R^m \subset \text{Span}\{\vec{a}_1, \vec{a}_2, ..., \vec{a}_n\} \subset R^m$.

(3) $\Longrightarrow$ (4) : Suppose (3) is true but (4) is NOT true. then the reduced Echelon form of A has at least one zero row, i.e., A is row equivalent to E by a series of elementary row reductions:

$$A \to E = \begin{bmatrix} * & * & ... & * \\ ... & ... & ... & ... \\ * & * & ... & * \\ 0 & 0 & ... & 0 \end{bmatrix}$$

We now consider the augmented matrix

$$\left[E \vdots \vec{e}_m \right] = \begin{bmatrix} * & * & ... & * & 0 \\ ... & ... & ... & ... & ... \\ * & * & ... & * & 0 \\ 0 & 0 & ... & 0 & 1 \end{bmatrix}, \tag{1}$$

and reverse completely all the row reductions we did from A to E. This leads to

$$\left[E \vdots \vec{e}_m \right] \to \left[A \vdots \vec{b} \right],$$

where $\vec{b}$ is a vector in R^m. Since the system corresponding to (1) is inconsistent, $A\vec{x} = \vec{b}$ for this particular vector $\vec{b}$ is inconsistent. In other words, $\vec{b}$ cannot be possibly a linear combination of $\vec{a}_i$. This contradicts to (3).

For instance,

$$A = \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix} \xrightarrow{R_1 + R_2 \to \tilde{R}_2} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = E$$

Then (with $m = 2$)

$$\left[E \vdots \vec{e}_m \right] = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\tilde{R}_2 - R_1 \to R_2} \begin{bmatrix} 1 & 2 & 0 \\ -1 & -2 & 1 \end{bmatrix}.$$

Thus

$$A\vec{x} = \vec{b}, \quad \vec{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ is inconsistent.}$$

(4) $\Longrightarrow$ (1) : Suppose for simplicity, A has the reduced Echelon form

$$A \to E = \begin{bmatrix} 1 & * & ... & * & ... & * \\ 0 & 1 & ... & * & ... & ... \\ ... & ... & ... & ... & ... & * \\ 0 & 0 & ... & 1 & ... & * \end{bmatrix}.$$

Then, for each $\vec{b}$, $\left[A \vdots \vec{b}\right]$ must have the reduced Echelon form

$$\left[A \vdots \vec{b}\right] \rightarrow \left[\begin{array}{cccccc:c} 1 & * & ... & * & ... & & * \\ 0 & 1 & ... & * & ... & & * \\ ... & ... & ... & ... & ... & & * \\ 0 & 0 & ... & 1 & ... & & * \end{array}\right].$$

The augmented matrix on the right is consistent. Therefore, $A\vec{x} = \vec{b}$ is consistent. ∎

Example 4 *Let*

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}.$$

1. *Determine if $A\vec{x} = \vec{b}$ has a solution for ALL $\vec{b} \in R^3$.*
2. *Describe $Span\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$, where $A = [\vec{a}_1, \vec{a}_2, \vec{a}_3]$.*

Solution. (1) By row operation, we have

$$\begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix} \xrightarrow[R_3 + 3R_1 \rightarrow R_3]{R_2 + 4R_1 \rightarrow R_2} \begin{bmatrix} 1 & 3 & 4 \\ 0 & 14 & 10 \\ 0 & 7 & 5 \end{bmatrix} \xrightarrow{R_3 - R_2/2 \rightarrow R3} \begin{bmatrix} 1 & 3 & 4 \\ 0 & 14 & 10 \\ 0 & 0 & 0 \end{bmatrix}.$$

By (4) in Theorem, $A\vec{x} = \vec{b}$ doesn't has a solution for all $\vec{b}$.

Solution. (2) We need to find for what $\vec{b}$,

$$\vec{b} \in Span\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}?$$

In other words, we need to describe all $\vec{b}$ such that $A\vec{x} = \vec{b}$ has a solution. In other words, we need to find conditions on $\vec{b}$ under which

$$A\vec{x} = \vec{b} \text{ is consistent.}$$

To this end, we perform exactly the same row operations as above for the augmented matrix $\left[A \vdots \vec{b}\right]$:

$$\left[A \vdots \vec{b}\right] = \begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \xrightarrow[R_3+3R_1 \to R_3]{R_2+4R_1 \to R_2} \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2+4b_1 \\ 0 & 7 & 5 & b_3+3b_1 \end{bmatrix}$$
$$\xrightarrow{R_3 - R_2/2 \to R3} \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2+4b_1 \\ 0 & 0 & 0 & b_3+3b_1-b_2/2-2b_1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & 4b_1+b_2 \\ 0 & 0 & 0 & b_1-b_2/2+b_3 \end{bmatrix}.$$

We know that $A\vec{x} = \vec{b}$ has a solution iff the last column is not pivot, i.e.,

$$b_1 - b_2/2 + b_3 = 0, \text{ or } b_3 = -b_1 + b_2/2.$$

We conclude that $\vec{b} \in$Span$\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ iff $b_3 = -b_1 + b_2/2$, and

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ -b_1+b_2/2 \end{bmatrix} = \begin{bmatrix} b_1+0b_2 \\ 0b_1+b_2 \\ -b_1+b_2/2 \end{bmatrix}$$
$$= \begin{bmatrix} b_1 \\ 0b_1 \\ -b_1 \end{bmatrix} + \begin{bmatrix} 0b_2 \\ b_2 \\ b_2/2 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 1/2 \end{bmatrix}$$

for any b_1 and b_2. Later, we call it "parametric representation". Therefore,

$$Span\{\vec{a}_1, \vec{a}_2, \vec{a}_3\} = Span \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1/2 \end{bmatrix} \right\}.$$

Homework 1.4:

#5,7,9,11,13,17,21,23