The Matrix Equation $A\mathbf{x} = \mathbf{b}$

A linear combination of vectors $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ can be expressed as the product of a matrix A and vector $\mathbf{x}$ such that $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Definition (The Product $A\mathbf{x}$).

If A is a $m \times n$ matrix with n columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and the vector $\mathbf{x} \in \mathbb{R}^n$, then $A\mathbf{x}$ is the linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ with weights $x_1, \ldots, x_n$. That is, the product of A and $\mathbf{x}$ is the linear combination of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of A using the corresponding entries $x_1, \ldots, x_n$ in $\mathbf{x}$ as weights.

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$$

Note that, as consequence of the preceding definition, the product of a matrix A and vector $\mathbf{x}$, $A\mathbf{x}$, is defined if and *only* if the number of columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in A equals the number of entries $x_1, \ldots, x_n$ in $\mathbf{x}$.

Example. For $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^m$, express the linear combination $3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$ as an $m \times n$ matrix A multiplied by a vector.

Let the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^m$ be the n columns of the $m \times n$ matrix A and let the weights $(3, -5, 7)$ be the entries (x_1, x_2, x_3) of a vector $\mathbf{x}$. Then $A\mathbf{x}$, the product of a matrix A multiplied by a vector $\mathbf{x}$, can be expressed by the following.

$$A\mathbf{x} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = 3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$$

■

Observe that, since a linear system can be expressed by a vector equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ involving a linear combination of vectors, and since a linear combination of vectors $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ can be expressed as a matrix A multiplied by a vector $\mathbf{x}$, a linear system can also be expressed by a **matrix equation** of the general form $A\mathbf{x} = \mathbf{b}$, where A is the matrix of coefficients, $[A \quad \mathbf{b}]$ is the augmented matrix, and $\mathbf{x}$ is the solution vector.

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{1}$$

Theorem (Solution Set of the Matrix Equation $A\mathbf{x} = \mathbf{b}$).

If A is an $m \times n$ matrix with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ and the vector $\mathbf{b} \in \mathbb{R}^m$, then the matrix equation $A\mathbf{x} = \mathbf{b}$ has the *same* solution set as the vector equation $x_1\mathbf{a}_1 + \cdots x_n\mathbf{a}_n = \mathbf{b}$ which, in turn, has the *same* solution set as the linear system whose augmented matrix is $\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}$.

Existence of Solutions

Recall that a vector $\mathbf{b} \in \operatorname{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ if and only if $\mathbf{b}$ is a linear combination of $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$, that is, if there exists a solution $(x_1, \dots, x_n)$ to the vector equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$. Therefore, since $A\mathbf{x}$ is defined as the linear combination of the columns $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ in a matrix A using the corresponding entries $x_1, \dots, x_n$ of a vector $\mathbf{x} \in \mathbb{R}^n$ as weights, the following is fact obtained.

Fact (Existence of a Solution $\mathbf{x}$ to the Matrix Equation $A\mathbf{x} = \mathbf{b}$).

The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ in A.

Example. Consider the matrix A and vector $\mathbf{b}$ defined below. Is the matrix equation $A\mathbf{x} = \mathbf{b}$ consistent for all possible b_1, b_2, b_3?

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

The augmented matrix of the matrix equation $A\mathbf{x} = \mathbf{b}$, for the given matrix A and vector $\mathbf{b}$, can be row reduced as follows.

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

Observe that the third entry in column four $b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) = b_1 - \frac{1}{2}b_2 + b_3$ and, therefore, because some choices of $\mathbf{b} = (b_1, b_2, b_3)$ can make $b_1 - \frac{1}{2}b_2 + b_3 \neq 0$, the matrix equation $A\mathbf{x} = \mathbf{b}$ is *not* consistent for all possible $\mathbf{b} = (b_1, b_2, b_3)$. ■

Note. In the preceding example, the matrix equation $A\mathbf{x} = \mathbf{b}$ fails to be consistent for all $\mathbf{b}$ because the echelon form of A has a row of zeros. Observe that if A were to have a pivot in each row, the calculations of the augmented column $\mathbf{b}$ would not matter because, in such a case, an echelon form of the augmented matrix $[A \quad \mathbf{b}]$ could *not* have a row of the form $\begin{bmatrix} 0 & \cdots & 0 & b \end{bmatrix}$.

In general, a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \in \mathbb{R}^m$ spans $\mathbb{R}^m$ if every vector $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \in \mathbb{R}^m$, that is, if $\text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \mathbb{R}^m$. Therefore, if the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of an $m \times n$ matrix A span $\mathbb{R}^m$, then every vector $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in A.

Theorem (Properties of the Coefficient Matrix A).

Let A be a $m \times n$ matrix. Then for a particular A, the following statements are *logically equivalent*, that is, they are either *all* true or *all* false.

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

b. Each $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in A.

c. The columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in A span $\mathbb{R}^m$.

d. A has a pivot position in *every* row.

Properties of the Matrix-Vector Product $A\mathbf{x}$

Consider the matrix $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, and scalar c. Observe that if the matrix-vector product $A(\mathbf{u} + \mathbf{v})$ is computed as a linear combination of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in A using the corresponding entries $u_1 + v_1, \ldots, u_n + v_n$ in $\mathbf{u} + \mathbf{v}$ as weights, then $A(\mathbf{u} + \mathbf{v})$ can be expressed as follows.

$$\begin{aligned} A(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} &= (u_1 + v_1)\,\mathbf{a}_1 + \cdots + (u_n + v_n)\,\mathbf{a}_n \\ &= (u_1\mathbf{a}_1 + \cdots + u_n\mathbf{a}_n) + (v_1\mathbf{a}_1 + \cdots + v_n\mathbf{a}_n) = A\mathbf{u} + A\mathbf{v} \end{aligned}$$

Also observe that if the matrix-vector product $A(c\mathbf{u})$ is computed as a linear combination of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in A using the corresponding entries $cu_1, \ldots, cu_n$ in $c\mathbf{u}$ as weights, then $A(c\mathbf{u})$ can be expressed as follows.

$$\begin{aligned} A(c\mathbf{u}) = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} cu_1 \\ \vdots \\ cu_n \end{bmatrix} &= (cu_1)\,\mathbf{a}_1 + \cdots + (cu_n)\,\mathbf{a}_n \\ c\,(u_1\mathbf{a}_1) + \cdots + c\,(u_n\mathbf{a}_n) &= c\,(u_1\mathbf{a}_1 + \cdots + u_n\mathbf{a}_n) = c(A\mathbf{u}) \end{aligned}$$

Theorem (Properties of the Matrix-Vector Product $A\mathbf{x}$).

If A is an $m \times n$ matrix, c is a scalar, and the vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, then:

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

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