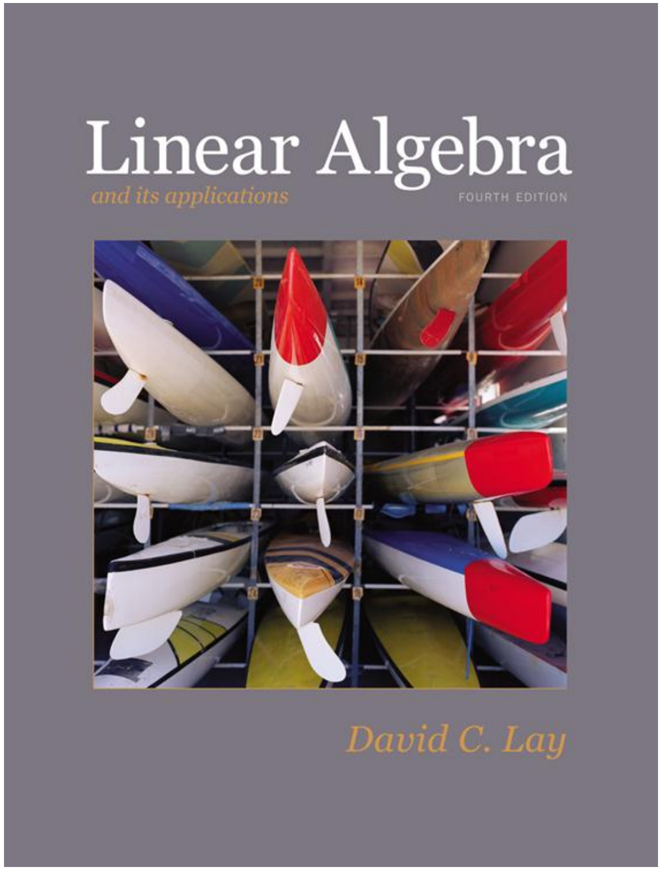
1 Linear Equations in Linear Algebra

1.4

THE MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$



MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- **Definition:** If A is an $m \times n$ matrix, with columns $\mathbf{a}_1$, …, $\mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of** A **and x**, denoted by $A\mathbf{x}$, **is the linear combination of the columns of** A **using the corresponding entries in x as weights**; that is,

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

- $A\mathbf{x}$ is defined only if the number of columns of A equals the number of entries in $\mathbf{x}$.

MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- **Example 1:** For $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ in $\mathbf{R}^m$, write the linear combination $3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$ as a matrix times a vector.
- **Solution:** Place $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ into the columns of a matrix A and place the weights $3, -5$, and 7 into a vector $\mathbf{x}$.
- That is,

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

MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- Now, write the system of linear equations as a vector equation involving a linear combination of vectors.
- For example, the following system

$$x_1 + 2x_2 - x_3 = 4 \qquad \text{----(1)}$$
$$-5x_2 + 3x_3 = 1$$

is equivalent to

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \qquad \text{----(2)}$$

MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- As in the given example (1), the linear combination on the left side is a matrix times a vector, so that (2) becomes

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \qquad \text{----(3)}$$

- Equation (3) has the form $A\mathbf{x} = \mathbf{b}$. Such an equation is called a **matrix equation**.

MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- **Theorem 3:** If A is an $m \times n$ matrix, with columns a1, …, an, and if b is in $\mathbf{R}^m$, then the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n a_n = \mathbf{b},$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

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

EXISTENCE OF SOLUTIONS

- The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if **b** is a linear combination of the columns of A.
- **Theorem 4:** Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.
  a. For each **b** in $\mathbf{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
  b. Each **b** in $\mathbf{R}^m$ is a linear combination of the columns of A.
  c. The columns of A span $\mathbf{R}^m$.
  d. A has a pivot position in every row.

PROOF OF THEOREM 4

- Statements (a), (b), and (c) are logically equivalent.
- So, it suffices to show (for an arbitrary matrix A) that (a) and (d) are either both true or false.
- Let U be an echelon form of A.
- Given **b** in $\mathbf{R}^m$, we can row reduce the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ to an augmented matrix $\begin{bmatrix} U & \mathbf{d} \end{bmatrix}$ for some **d** in $\mathbf{R}^m$:

$$\begin{bmatrix} A & \mathbf{b} \end{bmatrix} \sim \ldots \sim \begin{bmatrix} U & \mathbf{d} \end{bmatrix}$$

- If statement (d) is true, then each row of U contains a pivot position, and there can be no pivot in the augmented column.

PROOF OF THEOREM 4

- So $A\mathbf{x} = \mathbf{b}$ has a solution for any **b**, and (a) is true.
- If (d) is false, then the last row of U is all zeros.
- Let **d** be any vector with a 1 in its last entry.
- Then $\begin{bmatrix} U & \mathbf{d} \end{bmatrix}$ represents an *inconsistent* system.
- Since row operations are reversible, $\begin{bmatrix} U & \mathbf{d} \end{bmatrix}$ can be transformed into the form $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$.
- The new system $A\mathbf{x} = \mathbf{b}$ is also inconsistent, and (a) is false.

COMPUTATION OF *A*x

- **Example 2:** Compute $A\mathbf{x}$, where $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.
- **Solution:** From the definition,

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

COMPUTATION OF $A\mathbf{x}$

$$= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix} \text{---}(1)$$

$$= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}.$$

- The first entry in the product $A\mathbf{x}$ is a sum of products (*a dot product*), using the first row of A and the entries in $\mathbf{x}$.

COMPUTATION OF *A*x

- That is, $\begin{bmatrix} 2 & 3 & 4 \\ & & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ \\ \end{bmatrix}.$

- Similarly, the second entry in *A***x** can be calculated by multiplying the entries in the second row of *A* by the corresponding entries in **x** and then summing the resulting products.

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

ROW-VECTOR RULE FOR COMPUTING $A\mathbf{x}$

- Likewise, the third entry in $A\mathbf{x}$ can be calculated from the third row of A and the entries in $\mathbf{x}$.
- If the product $A\mathbf{x}$ is defined, then the ith entry in $A\mathbf{x}$ is the sum of the products of corresponding entries from row i of A and from the vertex $\mathbf{x}$.
- The matrix with 1s on the diagonal and 0s elsewhere is called an **identity matrix** and is denoted by I.
- For example, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is an identity matrix.

PROPERTIES OF THE MATRIX-VECTOR PRODUCT $A\mathbf{x}$

- **Theorem 5:** If A is an $m \times n$ matrix, u and v are vectors in $\mathbf{R}^n$, and c is a scalar, then
  a. $A(\mathrm{u}+\mathrm{v}) = A\mathrm{u} + A\mathrm{v}$;
  b. $A(c\mathrm{u}) = c(A\mathrm{u})$.
- **Proof:** For simplicity, take $n = 3$, $A = \begin{bmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \mathrm{a}_3 \end{bmatrix}$, and **u**, **v** in $\mathbf{R}^3$.
- For $i = 1, 2, 3$, let u_i and v_i be the ith entries in **u** and **v**, respectively.

PROPERTIES OF THE MATRIX-VECTOR PRODUCT *A*x

- To prove statement (a), compute $A(\mathbf{u}+\mathbf{v})$ as a linear combination of the columns of A using the entries in $\mathbf{u}+\mathbf{v}$ as weights.

$$A(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix}$$

Entries in $\mathbf{u}+\mathbf{v}$

$$= (u_1+v_1)\mathbf{a}_1 + (u_2+v_2)\mathbf{a}_2 + (u_3+v_3)\mathbf{a}_3$$

Columns of A

$$= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3)$$

$$= A\mathbf{u} + A\mathbf{v}$$

PROPERTIES OF THE MATRIX-VECTOR PRODUCT $A\mathbf{x}$

- To prove statement (b), compute $A(c\mathbf{u})$ as a linear combination of the columns of A using the entries in $c\mathbf{u}$ as weights.

$$A(c\mathbf{u}) = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} = (cu_1)\mathbf{a}_1 + (cu_2)\mathbf{a}_2 + (cu_3)\mathbf{a}_3$$

$$= c(u_1\mathbf{a}_1) + c(u_2\mathbf{a}_2) + c(u_3\mathbf{a}_3)$$

$$= c(u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3)$$

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