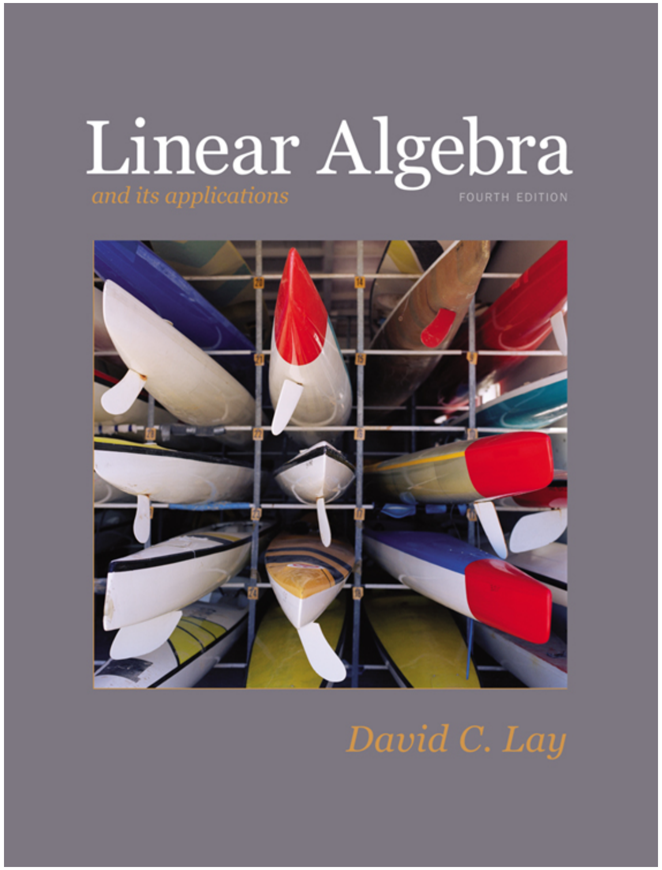
1 Linear Equations in Linear Algebra

1.7

LINEAR INDEPENDENCE



PEARSON

LINEAR INDEPENDENCE

- **Definition:** An indexed set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\mathrm{v}_1 + x_2\mathrm{v}_2 + \ldots + x_p\mathrm{v}_p = 0$$

has only the trivial solution. The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + \ldots + c_p\mathrm{v}_p = 0 \quad \text{-----}(1)$$

LINEAR INDEPENDENCE

- Equation (1) is called a **linear dependence relation** among $\mathbf{v}_1, \ldots, \mathbf{v}_p$ when the weights are not all zero.

- An indexed set is linearly dependent if and only if it is not linearly independent.

- **Example 1:** Let $\mathrm{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathrm{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, and $\mathrm{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

a. Determine if the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.
b. If possible, find a linear dependence relation among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

- **Solution:** We must determine if there is a nontrivial solution of the following equation.

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

LINEAR INDEPENDENCE

- Row operations on the associated augmented matrix show that

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- x_1 and x_2 are basic variables, and x_3 is free.
- Each nonzero value of x_3 determines a nontrivial solution of (1).
- Hence, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent.

LINEAR INDEPENDENCE

b. To find a linear dependence relation among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, row reduce the augmented matrix and write the new system:

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

- Thus, $x_1 = 2x_3$, $x_2 = -x_3$, and x_3 is free.
- Choose any nonzero value for x_3—say, $x_3 = 5$.
- Then $x_1 = 10$ and $x_2 = -5$.

LINEAR INDEPENDENCE

- Substitute these values into equation (1) and obtain the equation below.

$$10v_1 - 5v_2 + 5v_3 = 0$$

- This is one (out of infinitely many) possible linear dependence relations among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

LINEAR INDEPENDENCE OF MATRIX COLUMNS

- Suppose that we begin with a matrix $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ instead of a set of vectors.

- The matrix equation $A\mathbf{x} = \mathbf{0}$ can be written as

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = \mathbf{0}.$$

- *Each linear dependence relation among the columns of A corresponds to a nontrivial solution of* $A\mathbf{x} = \mathbf{0}$.

- Thus, the columns of matrix A are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

SETS OF ONE OR TWO VECTORS

- A set containing only one vector – say, **v** – is linearly independent if and only if **v** is not the zero vector.
- This is because the vector equation $x_1\mathbf{v} = \mathbf{0}$ has only the trivial solution when $\mathbf{v} \neq \mathbf{0}$.
- The zero vector is linearly dependent because $x_1\mathbf{0} = \mathbf{0}$ has many nontrivial solutions.

SETS OF ONE OR TWO VECTORS

- A set of two vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent if at least one of the vectors is a multiple of the other.

- The set is linearly independent if and only if neither of the vectors is a multiple of the other.

SETS OF TWO OR MORE VECTORS

- **Theorem 7:** Characterization of Linearly Dependent Sets
- An indexed set $S = \{v_1, \ldots, v_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others.
- In fact, if S is linearly dependent and $v_1 \neq 0$, then some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors, $\mathbf{v}_1, \ldots, v_{j-1}$.

SETS OF TWO OR MORE VECTORS

- **Proof:** If some $\mathbf{v}_j$ in S equals a linear combination of the other vectors, then $\mathbf{v}_j$ can be subtracted from both sides of the equation, producing a linear dependence relation with a nonzero weight (-1) on $\mathbf{v}_j$.
- [For instance, if $\mathrm{v}_1 = c_2\mathrm{v}_2 + c_3\mathrm{v}_3$, then $0 = (-1)\mathrm{v}_1 + c_2\mathrm{v}_2 + c_3\mathrm{v}_3 + 0\mathrm{v}_4 + \ldots + 0\mathrm{v}_p$.]
- Thus S is linearly dependent.
- Conversely, suppose S is linearly dependent.
- If $\mathbf{v}_1$ is zero, then it is a (trivial) linear combination of the other vectors in S.

SETS OF TWO OR MORE VECTORS

- Otherwise, $\mathbf{v}_1 \neq \mathbf{0}$, and there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p = \mathbf{0}.$$

- Let j be the largest subscript for which $c_j \neq 0$.

- If $j = 1$, then $c_1\mathbf{v}_1 = \mathbf{0}$, which is impossible because $\mathbf{v}_1 \neq \mathbf{0}$.

SETS OF TWO OR MORE VECTORS

- So $j > 1$, and

$$c_1\mathbf{v}_1 + \ldots + c_j\mathbf{v}_j + 0\mathbf{v}_j + 0\mathbf{v}_{j+1} + \ldots + 0\mathbf{v}_p = 0$$

$$c_j\mathbf{v}_j = -c_1\mathbf{v}_1 - \ldots - c_{j-1}\mathbf{v}_{j-1}$$

$$\mathbf{v}_j = \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \ldots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1}.$$

SETS OF TWO OR MORE VECTORS

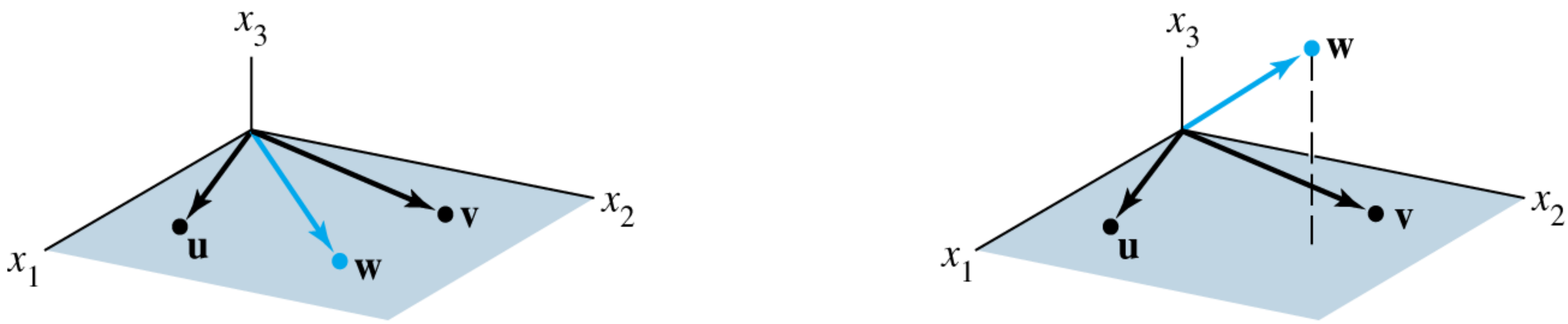
- Theorem 7 does *not* say that *every* vector in a linearly dependent set is a linear combination of the preceding vectors.
- A vector in a linearly dependent set may fail to be a linear combination of the other vectors.
- **Example 2:** Let $\mathrm{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $\mathrm{v} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describe the set spanned by **u** and **v**, and explain why a vector **w** is in Span $\{\mathbf{u}, \mathbf{v}\}$ if and only if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

SETS OF TWO OR MORE VECTORS

- **Solution:** The vectors **u** and **v** are linearly independent because neither vector is a multiple of the other, and so they span a plane in $\mathbb{R}^3$.
- Span {**u**, **v**} is the x_1x_2-plane (with $x_3 = 0$).
- If **w** is a linear combination of **u** and **v**, then {**u**, **v**, **w**} is linearly dependent, by Theorem 7.
- Conversely, suppose that {**u**, **v**, **w**} is linearly dependent.
- By theorem 7, some vector in {**u**, **v**, **w**} is a linear combination of the preceding vectors (since $\mathbf{u} \neq \mathbf{0}$).
- That vector must be **w**, since **v** is not a multiple of **u**.

SETS OF TWO OR MORE VECTORS

- So **w** is in Span {**u**, **v**}. See the figures given below.



Linearly dependent, **w** in Span{**u**, **v**}

Linearly independent, **w** *not* in Span{**u**, **v**}

- Example 2 generalizes to any set {**u**, **v**, **w**} in $\mathbb{R}^3$ with **u** and **v** linearly independent.
- The set {**u**, **v**, **w**} will be linearly dependent if and only if **w** is in the plane spanned by **u** and **v**.

SETS OF TWO OR MORE VECTORS

- **Theorem 8:** If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.
- **Proof:** Let $A = \begin{bmatrix} \mathrm{v}_1 & \cdots & \mathrm{v}_p \end{bmatrix}$.
- Then A is $n \times p$, and the equation $A\mathrm{x} = 0$ corresponds to a system of n equations in p unknowns.
- If $p > n$, there are more variables than equations, so there must be a free variable.

SETS OF TWO OR MORE VECTORS

- Hence $A\mathrm{x} = 0$ has a nontrivial solution, and the columns of A are linearly dependent.
- See the figure below for a matrix version of this theorem.

$$\begin{array}{c} \\ n \end{array}\overset{\displaystyle p}{\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}}$$

If $p > n$, the columns are linearly dependent.

- Theorem 8 says nothing about the case in which the number of vectors in the set does *not* exceed the number of entries in each vector.

SETS OF TWO OR MORE VECTORS

- **Theorem 9:** If a set $S = \{\mathrm{v}_1, \ldots, \mathrm{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

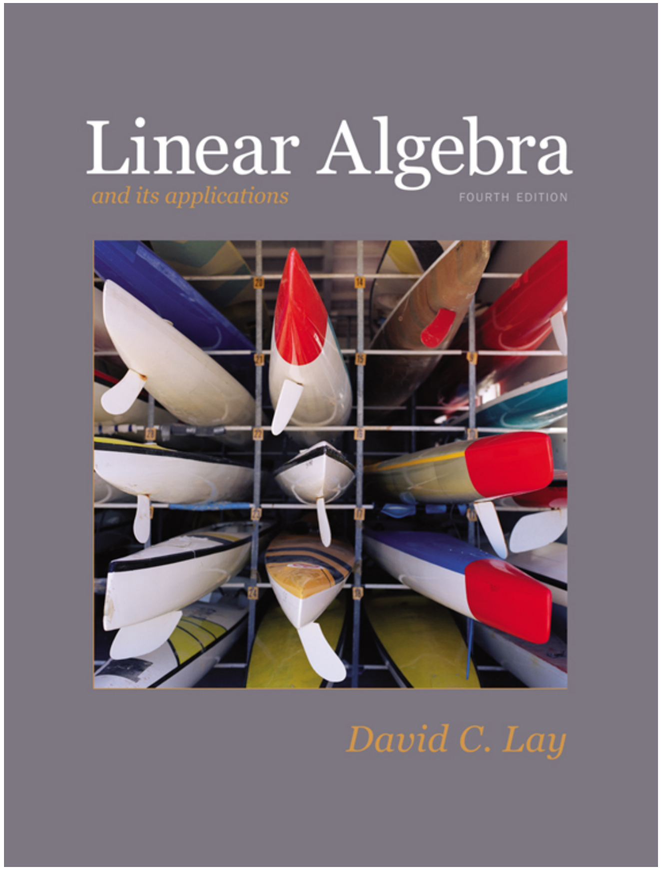
- **Proof:** By renumbering the vectors, we may suppose $\mathrm{v}_1 = 0$.

- Then the equation $1\mathrm{v}_1 + 0\mathrm{v}_2 + \ldots + 0\mathrm{v}_p = 0$ shows that S in linearly dependent.

2 Matrix Algebra

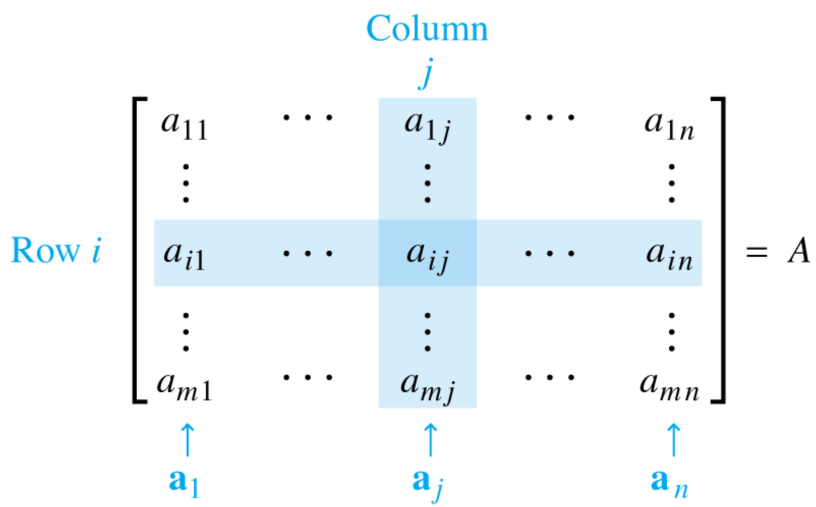
2.1

MATRIX OPERATIONS



MATRIX OPERATIONS

- If A is an $m \times n$ matrix—that is, a matrix with m rows and n columns—then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j)-entry of A. See the figure below.
- Each column of A is a list of m real numbers, which identifies a vector in $\mathbb{R}^m$.



Matrix notation.

MATRIX OPERATIONS

- The columns are denoted by $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and the matrix A is written as

$$A = \begin{bmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \cdots & \mathrm{a}_n \end{bmatrix}.$$

- The number a_{ij} is the ith entry (from the top) of the jth column vector $\mathbf{a}_j$.
- The **diagonal entries** in an $m \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ are a_{11}, a_{22}, a_{33}, …, and they form the **main diagonal** of A.
- A **diagonal matrix** is a sequence $n \times m$ matrix whose nondiagonal entries are zero.
- An example is the $n \times n$ identity matrix, I_n.

SUMS AND SCALAR MULTIPLES

- An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0.

- The two matrices are **equal** if they have the same size (*i.e.*, the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal.

- If A and B are $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns in A and B.

SUMS AND SCALAR MULTIPLES

- Since vector addition of the columns is done entrywise, each entry in $A+B$ is the sum of the corresponding entries in A and B.

- The sum $A+B$ is defined only when A and B are the same size.

- **Example 1:** Let $A=\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, B=\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix},$

  and $C=\begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$. Find $A+B$ and $A+C$.

SUMS AND SCALAR MULTIPLES

- **Solution:** $A+B=\begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$ but $A+C$ is not defined because A and C have different sizes.

- If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose columns are r times the corresponding columns in A.

- **Theorem 1:** Let A, B, and C be matrices of the same size, and let r and s be scalars.
  a. $A+B=B+A$

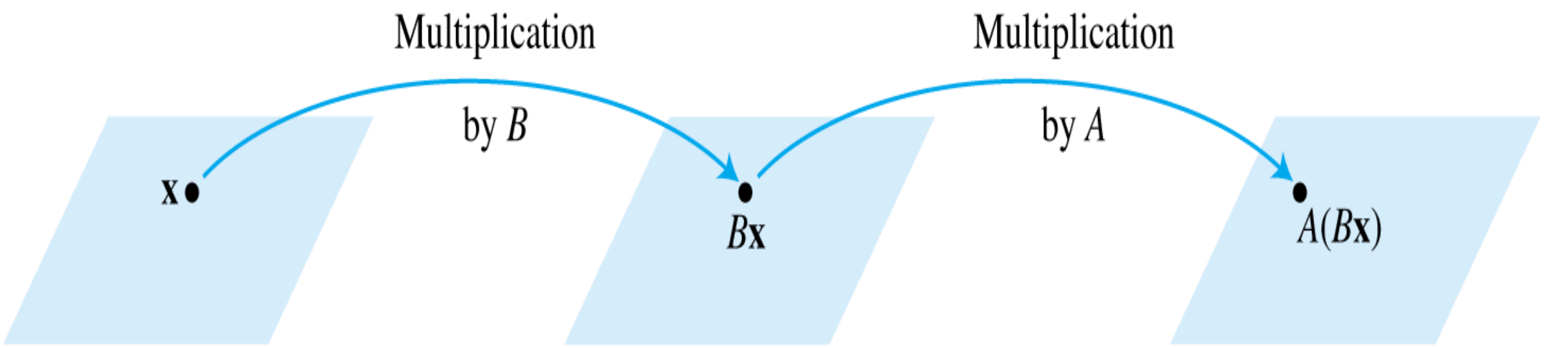
SUMS AND SCALAR MULTIPLES

b. $(A+B)+C=A+(B+C)$
c. $A+0=A$
d. $r(A+B)=rA+rB$
e. $(r+s)A=rA+sA$
f. $r(sA)=(rs)A$

- Each quantity in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal.

MATRIX MULTIPLICATION

- When a matrix B multiplies a vector $\mathbf{x}$, it transforms $\mathbf{x}$ into the vector $B\mathbf{x}$.
- If this vector is then multiplied in turn by a matrix A, the resulting vector is $A(B\mathbf{x})$. See the Fig. below.

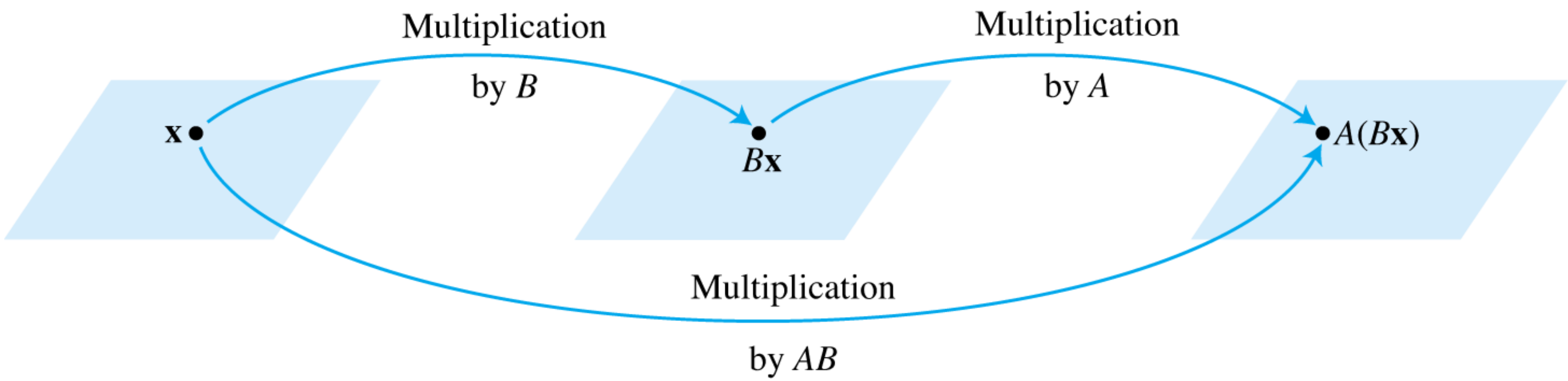


Multiplication by B and then A.

- Thus $A(B\mathbf{x})$ is produced from x by a composition of mappings—the linear transformations.

MATRIX MULTIPLICATION

- Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by AB, so that $A(B\mathrm{x})=(\mathrm{AB})\mathrm{x}$. See the figure below.



Multiplication by AB.

- If A is $m \times n$, B is $n \times p$, and $\mathbf{x}$ is in $\mathbb{R}^p$, denote the columns of B by $\mathbf{b}_1, \ldots, \mathbf{b}p$ and the entries in $\mathbf{x}$ by $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

MATRIX MULTIPLICATION

- Then

$$B\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_p\mathbf{b}_p$$

- By the linearity of multiplication by A,

$$A(B\mathbf{x}) = A(x_1\mathbf{b}_1) + \ldots + A(x_p\mathbf{b}_p)$$

$$= x_1A\mathbf{b}_1 + \ldots + x_pA\mathbf{b}_p$$

- The vector $A(B\mathbf{x})$ is a linear combination of the vectors $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$, using the entries in $\mathbf{x}$ as weights.
- In matrix notation, this linear combination is written as

$$A(B\mathbf{x}) = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}\mathbf{x}$$

MATRIX MULTIPLICATION

- Thus multiplication by $\begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$ transforms $\mathbf{x}$ into $A(B\mathbf{x})$.

- **Definition:** If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$.

- That is,

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

- *Multiplication of matrices corresponds to composition of linear transformations.*

MATRIX MULTIPLICATION

- **Example 2:** Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 9 \\ 1 & -2 & 3 \end{bmatrix}$.

- **Solution:** Write $B = \begin{bmatrix} \mathrm{b}_1 & \mathrm{b}_2 & \mathrm{b}_3 \end{bmatrix}$, and compute:

MATRIX MULTIPLICATION

$$Ab_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad Ab_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad Ab_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

- Then

$$AB = A\begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$$Ab_1 \quad Ab_2 \quad Ab_3$$

MATRIX MULTIPLICATION

- Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

- Row—column rule for computing AB
- If a product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B.
- If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj}.$$

PROPERTIES OF MATRIX MULTIPLICATION

- **Theorem 2:** Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.
  a. $A(BC) = (AB)C$ (associative law of multiplication)
  b. $A(B + C) = AB + AC$ (left distributive law)
  c. $(B + C)A = BA + CA$ (right distributive law)
  d. $r(AB) = (rA)B = A(rB)$ for any scalar r
  e. $I_m A = A = AI_n$ (identity for matrix multiplication)

PROPERTIES OF MATRIX MULTIPLICATION

- **Proof:** Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known that the composition of functions is associative.
- Let $C = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_p \end{bmatrix}$
- By the definition of matrix multiplication,

$$BC = \begin{bmatrix} B\mathbf{c}_1 & \cdots & B\mathbf{c}_p \end{bmatrix}$$

$$A(BC) = \begin{bmatrix} A(B\mathbf{c}_1) & \cdots & A(B\mathbf{c}_p) \end{bmatrix}$$

PROPERTIES OF MATRIX MULTIPLICATION

- The definition of AB makes $A(B\mathbf{x}) = (AB)\mathbf{x}$ for all $\mathbf{x}$, so

$$A(BC) = \begin{bmatrix} (AB)\mathbf{c}_1 & \cdots & (AB)\mathbf{c}_p \end{bmatrix} = (AB)C$$

- The left-to-right order in products is critical because AB and BA are usually not the same.
- Because the columns of AB are linear combinations of the columns of A, whereas the columns of BA are constructed from the columns of B.
- The position of the factors in the product AB is emphasized by saying that A is *right-multiplied* by B or that B is *left-multiplied* by A.

PROPERTIES OF MATRIX MULTIPLICATION

- If $AB = BA$, we say that A and B **commute** with one another.

- **Warnings:**
  1. In general, $AB \neq BA$.
  2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$.
  3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$.

POWERS OF A MATRIX

- If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k copies of A:

$$A^k = \underbrace{A \cdots A}_{k}$$

- If A is nonzero and if $\mathbf{x}$ is in $\mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times.
- If $k = 0$, then $A^0\mathbf{x}$ should be $\mathbf{x}$ itself.
- Thus A^0 is interpreted as the identity matrix.

THE TRANSPOSE OF A MATRIX

- Given an $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

Theorem 3: Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$

b. $(A+B)^T = A^T + B^T$

c. For any scalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$

THE TRANSPOSE OF A MATRIX

- The transpose of a product of matrices equals the product of their transposes in the *reverse* order.