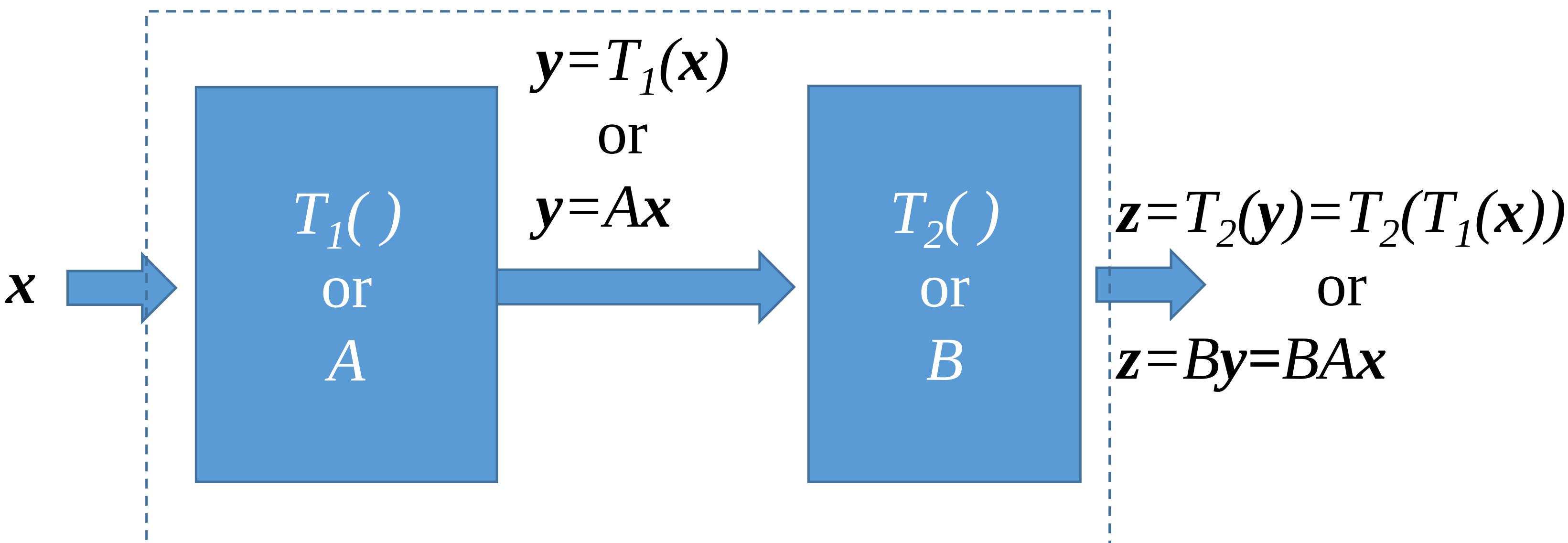
Chapter 2: Matrices and Linear Transformations



- A garden supply store sells three mixtures of grass seed

$$\begin{array}{cccc} & \text{deluxe} & \text{standard} & \text{economy} & \\ B = & \left[\begin{array}{c} .80 \\ .20 \end{array}\right. & \begin{array}{c} .60 \\ .40 \end{array} & \left.\begin{array}{c} .40 \\ .60 \end{array}\right] & \begin{array}{c} \text{bluegrass} \\ \text{rye} \end{array} \end{array}$$

- $v = \begin{bmatrix} 60 \\ 50 \\ 30 \end{bmatrix}$, the number of pounds of each mixture in stock.
- The germination rates of bluegrass and rye seeds under both wet and dry conditions:

$$\begin{array}{cccc} & \text{bluegrass} & \text{rye} & \\ A = & \left[\begin{array}{c} .80 \\ .60 \end{array}\right. & \left.\begin{array}{c} .70 \\ .40 \end{array}\right] & \begin{array}{c} \text{wet} \\ \text{dry} \end{array} \end{array}$$

  - ✓ The amounts of bluegrass and rye seed are respectively

$$B\boldsymbol{v} = \begin{bmatrix} .80 & .60 & .40 \\ .20 & .40 & .60 \end{bmatrix} \begin{bmatrix} 60 \\ 50 \\ 30 \end{bmatrix} = \begin{bmatrix} 90 \\ 50 \end{bmatrix}.$$

  - ✓ The amounts of seed expected to germinate:

$$A(B\boldsymbol{v}) = \begin{bmatrix} .80 & .70 \\ .60 & .40 \end{bmatrix} \begin{bmatrix} 90 \\ 50 \end{bmatrix} = \begin{bmatrix} 107 \\ 74 \end{bmatrix} \begin{array}{c} \text{wet} \\ \text{dry} \end{array}$$

2.1 Matrix Multiplication

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix.

- For any $p \times 1$ vector $\boldsymbol{v}$, the product $B\boldsymbol{v}$ is an $n \times 1$ vector, and hence $A(B\boldsymbol{v})$ is an $m \times 1$ vector.
- Is there an $m \times p$ matrix C such that $A(B\boldsymbol{v}) = C\boldsymbol{v}$ for every $p \times 1$ vector $\boldsymbol{v}$?

$A(B\boldsymbol{v}) = A\left(v_1\boldsymbol{b}_1 + v_2\boldsymbol{b}_2 + \cdots + v_p\boldsymbol{b}_p\right)$ by the definition of matrix-vector product

$= A(v_1\boldsymbol{b}_1) + A(v_2\boldsymbol{b}_2) + \cdots + A(v_p\boldsymbol{b}_p)$ by Theorem 1.3 (a)

$= v_1A\boldsymbol{b}_1 + v_2A\boldsymbol{b}_2 + \cdots + v_pA\boldsymbol{b}_p$ by Theorem 1.3 (b)

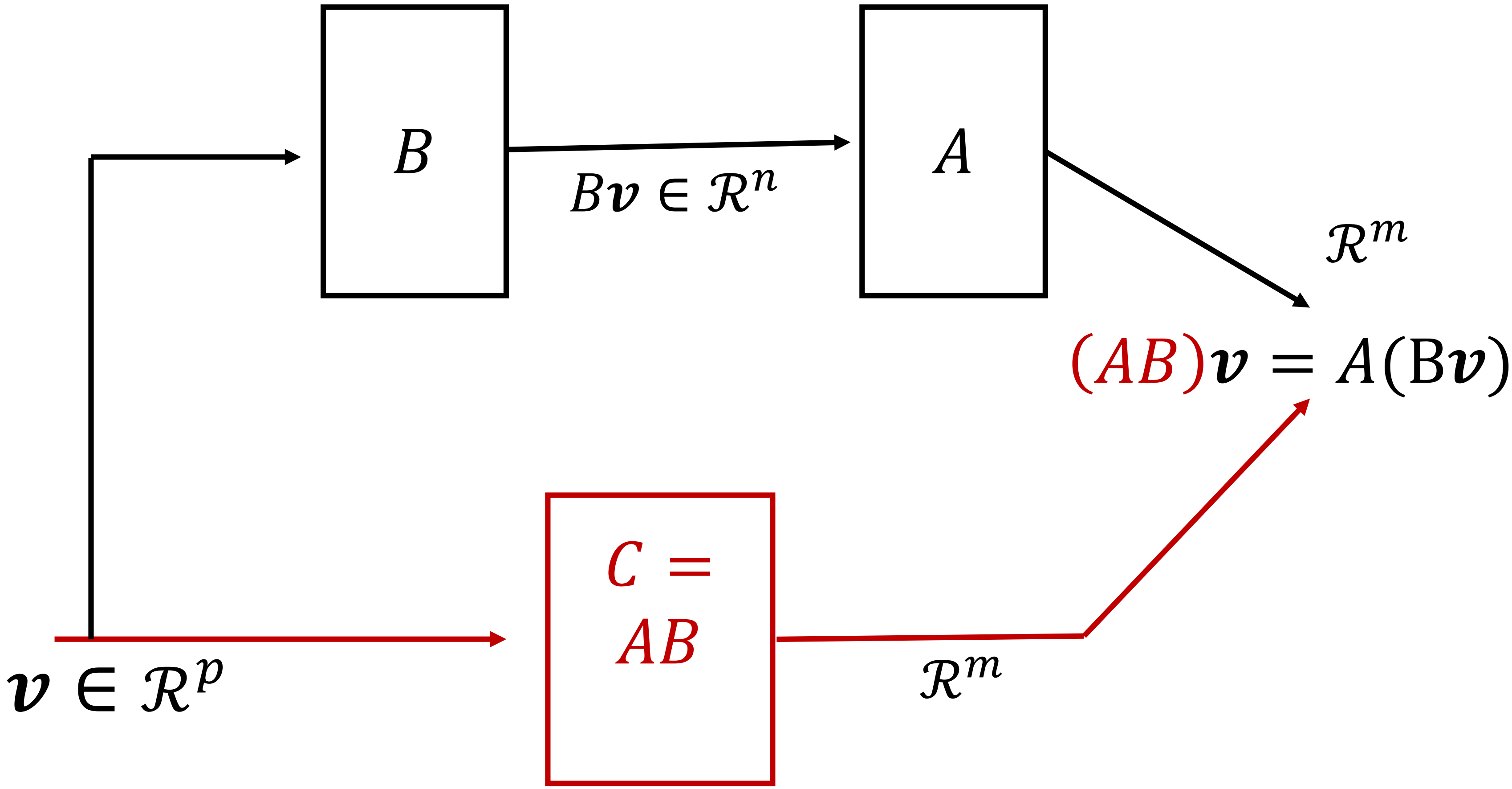
$= [A\boldsymbol{b}_1 \quad A\boldsymbol{b}_2 \quad \cdots \quad A\boldsymbol{b}_p]\boldsymbol{v}$ by the definition of matrix-vector product

Let C be the $m \times p$ matrix $\begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_p \end{bmatrix}$; that is $\boldsymbol{c}_j = A\boldsymbol{b}_j$. Then, $A(B\boldsymbol{v})=C\boldsymbol{v}$ for all $\boldsymbol{v}$ in R^p.

Furthermore, C is the only matrix with this property by Theorem 1.3(e).

Definition Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Define the (**matrix**)**product** AB to be the $m \times p$ matrix whose j-th column is $A\boldsymbol{b}_j$. That is,

$$C=[A\boldsymbol{b}_1 \;\; A\boldsymbol{b}_2 \; \cdots \; A\boldsymbol{b}_p].$$



The associate law of matrix multiplication

$$AB = A[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_p] = [A\boldsymbol{b}_1 \quad A\boldsymbol{b}_2 \quad \cdots \quad A\boldsymbol{b}_p]$$

Associate law for the product of two matrices and a vector:

For any $m\times n$ matrix A, any $n\times p$ matrix B, and any $p\times 1$ vector $\mathbf{v}$,

$$(AB)\mathbf{v}=A(B\mathbf{v}).$$

$$
\begin{array}{cc} & n \\ m & \begin{bmatrix} (m\times n) \end{bmatrix} \end{array}
\begin{array}{cc} & p \\ n & \begin{bmatrix} (n\times p) \end{bmatrix} \end{array}
=
\begin{array}{cc} & p \\ m & \begin{bmatrix} (m\times p) \end{bmatrix} \end{array}
$$

Matrix multiplication is not commutative

For arbitrary matrices A and B, AB need not equal BA.

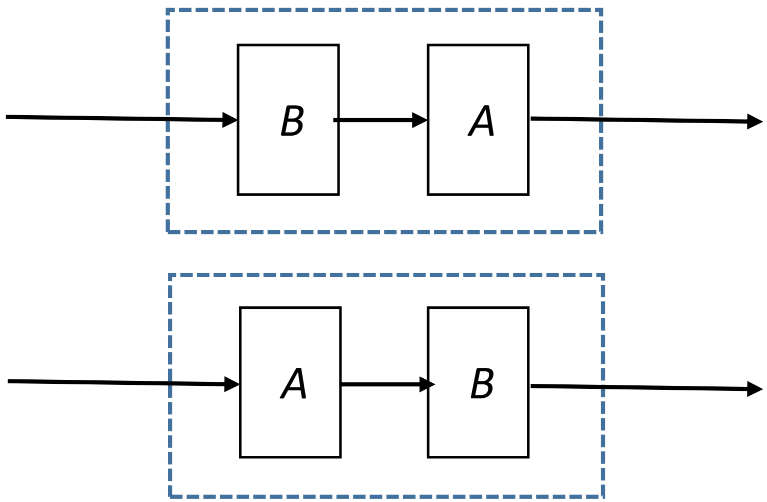
Example:

Let $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Then

- $(AB)\boldsymbol{e}_1 = A(B\boldsymbol{e}_1) = A\boldsymbol{e}_2 = \boldsymbol{e}_2$.
- $(BA)\boldsymbol{e}_1 = B(A\boldsymbol{e}_1) = B\boldsymbol{e}_3 = \boldsymbol{e}_3$.

Hence $AB \neq BA$.



A *row-column* rule is one way to find only some specific entries of the product AB.

- The (i,j)-entry of AB is the i-th component of its j-th column, $A\boldsymbol{b}_j$.
- This entry equals $\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$, which is the product of row i of A and column j of B.

$$
\begin{array}{cc}
 & \text{Column } j \text{ of } B \\
\text{Row } i \text{ of } A &
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}
\begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2p} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{bmatrix}
\end{array}
$$

Row-Column Rule for the (i,j)-Entry of a Matrix Product

To compute the (i,j)-entry of AB, locate the i-th row of A and the j-th column of B as in the preceding diagram. Moving across the i-th row of A and down the j-th column of B, multiply each entry of the row by the corresponding entry of the column. Then sum these products to obtain the (i,j)-entry of AB. In symbol, the (i,j)-entry of AB is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}. \qquad \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)$$

Theorem 2.1 Let A and B be $k \times m$ matrices, C be an $m \times n$ matrix, and P and Q be $n \times p$ matrices. Then the following statements are true:

a) $s(AC)=(sA)C=A(sC)$ for any scalar s.
b) $A(CP)=(AC)P$. (Associate law of matrix multiplication)
c) $(A+B)C=AC+BC$. (Right distributive law)
d) $C(P+Q)=CP+CQ$. (Left distributive law)
e) $I_k A=A=AI_m$.
f) The product of any matrix and a zero matrix is a zero matrix.
g) $(AC)^T=C^T A^T$.

- Theorem 2.1 (b) allows us to omit parentheses when writing products of matrices. $\rightarrow ACP$
- If A is an $n \times n$ matrix, we use the exponential notation A^k to denote the product of A with itself k times. $\rightarrow \overbrace{AA\cdots A}^{k} = A^k$
- By convention, $A^1=A$ and $A^0=I_n$. $\rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Proof of Theorem 2.1

a) $s(AC) = s[A\boldsymbol{c_1} \quad A\boldsymbol{c_2} \quad \cdots \quad A\boldsymbol{c_n}]$ by the definition of matrix product
$= [s(A\boldsymbol{c_1}) \quad s(A\boldsymbol{c_2}) \quad \cdots \quad s(A\boldsymbol{c_n})] \leftarrow$ scalar multiple of a matrix
$= [(sA)\boldsymbol{c_1} \quad (sA)\boldsymbol{c_2} \quad \cdots \quad (sA)\boldsymbol{c_n}]$ or $[A(s\boldsymbol{c_1}) \quad A(s\boldsymbol{c_2}) \quad \cdots \quad A(s\boldsymbol{c_n})] \leftarrow$ by theorem 1.3(b)
$= (sA)C$ or $A(sC) \leftarrow$ by the definition of matrix product.

b) Both $A(CP)$ and $(AC)P$ are $k \times p$ matrices. $\leftarrow$ the same matrix size
Let $\boldsymbol{u_j} = C\boldsymbol{p_j}, j = 1,2,\cdots,p$, the j-th column of CP. Then,
$A\boldsymbol{u_j} = A\left(C\boldsymbol{p}_j\right) = (AC)\boldsymbol{p}_j \ \forall\, j \in \{1,2,\cdots,p\} \leftarrow$ by the boxed result on slide 6.
Therefore $A(CP) = (AC)P$. *(Usually written as ACP).*

c) Both $(A+B)C$ and $AC + BC$ are $k \times n$ matrices.
$(A+B)C = [(A+B)\boldsymbol{c_1} \quad (A+B)\boldsymbol{c_2} \quad \cdots \quad (A+B)\boldsymbol{c_n}]$ by the definition of matrix product ..... (1)
Similarly,
$AC = [A\boldsymbol{c_1} \quad A\boldsymbol{c_2} \quad \cdots \quad A\boldsymbol{c_n}]$ and $BC = [B\boldsymbol{c_1} \quad B\boldsymbol{c_2} \quad \cdots \quad B\boldsymbol{c_n}]$. Then
$AC + BC = [(A+B)\boldsymbol{c_1} \quad (A+B)\boldsymbol{c_2} \quad \cdots \quad (A+B)\boldsymbol{c_n}]$ by Theorem 1.3(c) and matrix sum.
Thus, $(A+B)C = AC + BC$.

If A and B are matrices with the same number of rows, then the matrix $[A\ B]$, whose columns are the columns of A followed by the columns of B, is called an **augmented matrix**.

- For example, if $A = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 3 & 1 \end{bmatrix}$, then $[A\ \ B] = \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 3 & 1 \end{bmatrix}$.
- If P is an $m \times n$ matrix and A and B are matrices with n rows, then

$$P[A\ \ B] = [PA\ \ PB].$$

Special Matrices

- The (i,j)-entry of a matrix A is called a **diagonal entry** if $i=j$.
- The diagonal entries form the **diagonal** of A.
- A *square* matrix A is called a **diagonal matrix** if all its nondiagonal entries are all zeros.
  - ✓ Identity matrices and square zero matrices are diagonal matrices.
- If A and B are $n \times n$ diagonal matrices, then AB is also an $n \times n$ diagonal matrix. Moreover, the diagonal entries of AB are the products of the corresponding diagonal entries of A and B.
- A square matrix A is called **symmetric** if $A^T=A$.
  - ✓ Any diagonal matrix is equal to its transpose, and is hence symmetric.

2.2* Applications of Matrix Multiplication (Self-study)

Examples of diagonal matrices and a symmetric matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

- $AB = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$ and $BA = ?$

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

- $A^T = A$

2.3 Invertibility and Elementary Matrices

For any real number $a \neq 0$, there is a unique real number b, called the *multiplicative inverse* of a, such that $ab=ba=1$.

In the context of matrices:

- The identity matrix I_n is a *multiplicative identity*;
- For what matrices A does there exist a matrix B such that $AB=BA=I_n$.

  (The above is possible only if both A and B are $n \times n$ matrices)

Examples:

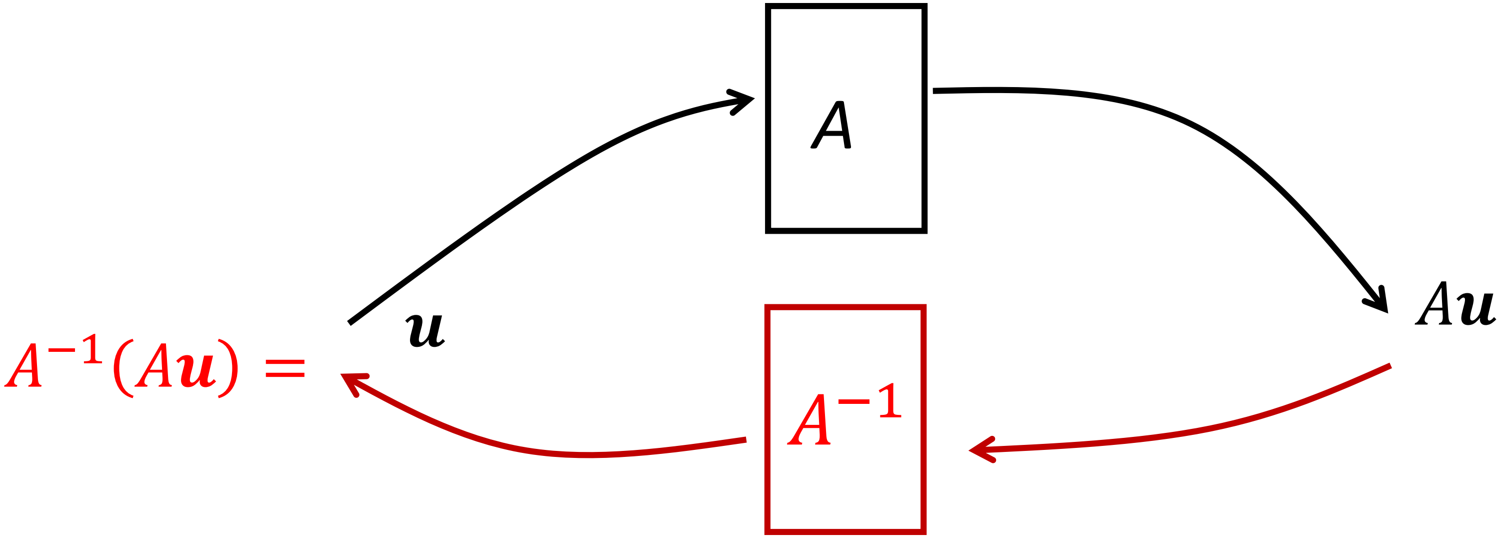
- Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$. Then $AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$.
- Let $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ and any $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $AB \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Definition An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called an **inverse** of A.

- If A is invertible, then its inverse is unique; the unique inverse of A is denoted by A^{-1}.
- The $n \times n$ zero matrix O has no inverse because $OB = O \neq I_n$ for any $n \times n$ matrix B.

If A is an invertible $n \times n$ matrix, then for every $\boldsymbol{b}$ in R^n, $A\boldsymbol{x} = \boldsymbol{b}$ has the unique solution $A^{-1}\boldsymbol{b}$.

$$A^{-1}(A\boldsymbol{x}) = A^{-1}\boldsymbol{b} \rightarrow (A^{-1}A)\boldsymbol{x} = A^{-1}\boldsymbol{b} \rightarrow I_n\, \boldsymbol{x} = A^{-1}\boldsymbol{b} \rightarrow \boldsymbol{x} = A^{-1}\boldsymbol{b}$$



The rotation matrix $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$:

- For $\theta = 0^\circ$, $A_\theta = I_2$
- $A_\alpha A_{-\alpha} = A_{\alpha+(-\alpha)} = A_{-\alpha}A_\alpha = A_{0^\circ} = I_2$.

Theorem 2.2

Let A and B be $n \times n$ matrices.

a) If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

b) If A and B are invertible, then AB is invertible and $(AB)^{-1} = B^{-1} A^{-1}$.

c) If A is invertible, then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.

Let $A_1, A_2, \cdots, A_k$ be $n \times n$ invertible matrices. Then the product $A_1 A_2 \cdots A_k$ is invertible, and
$(A_1 A_2 \cdots A_k)^{-1} = (A_k)^{-1} (A_{k-1})^{-1} \cdots (A_1)^{-1}$. (by Theorem 2.2 (b))

Proof of Theorem 2.2

(a) Suppose that A is invertible, with inverse A^{-1}.

Then $AA^{-1} = A^{-1}A = I_n$.

Hence, the inverse of A^{-1}, denoted by $(A^{-1})^{-1}$, is A. ← by the definition of matrix inverse.

(b) Suppose that A and B are invertible.

Then $AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$.

Similarly,$(B^{-1}A^{-1})AB = \ldots = I_n$.

Hence $(AB)^{-1} = B^{-1}A^{-1}$ ← by the definition of matrix inverse.

(c) Suppose that A is invertible.

Then $A^{-1}A = AA^{-1} = I_n$.

Applying Theorem 2.1 (g), we have

$A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$. Similarly

$(A^{-1})^TA^T = (AA^{-1})^T = I_n^T = I_n$.

Thus, $(A^T)^{-1} = (A^{-1})^T$, by the definition of matrix inverse.

Elementary Matrices

Every elementary row operation can be performed by matrix multiplication.

For example, let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$; Then

- Multiply row 2 of A by the scalar k through the matrix product $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$.
- Interchange rows 1 and 2 of A through the matrix product $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$.
- Add k times row 1 of A to row 2 through the matrix product $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ka + c & kb + d \end{bmatrix}$

An $n \times n$ matrix E is called an **elementary matrix** if it can be obtained from I_n by a single elementary row operation.

For example, the following elementary matrix E and the associated elementary row operation on A:

- $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{2\boldsymbol{r}_1 + \boldsymbol{r}_3 \to \boldsymbol{r}_3} E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$
- $A \xrightarrow{2\boldsymbol{r}_1 + \boldsymbol{r}_3 \to \boldsymbol{r}_3} E\mathrm{A}$

Let A be an $m \times n$ matrix, and let E be an $m \times m$ elementary matrix resulting from an elementary row operation on I_m. Then the product EA can be obtained from A by the identical elementary row operation on A.

The concept of a *reverse* operation leads to a way of obtaining the inverse of an elementary matrix.

- Consider the preceding matrix E and let $F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ be the elementary matrix obtained from I_3.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{-2\boldsymbol{r}_1+\boldsymbol{r}_3\rightarrow\boldsymbol{r}_3} F$$

We see that $FE=EF=I_3$. $\rightarrow$ Hence E is invertible and $E^{-1}=F$.

Every elementary matrix is invertible. Furthermore, the inverse of an elementary matrix is also an elementary matrix.

If A is an $m \times n$ matrix with reduced row echelon R, there exist elementary matrices $E_1, E_2, \cdots, E_k$ such that $E_k E_{k-1} \cdots E_1 A = R$.

Let $P = E_k E_{k-1} \cdots E_1$. Then

- P is a product of elementary matrices. So, P is invertible, with $P^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$.
- $PA = R$.

$$\Downarrow$$

Theorem 2.3

Let A be an $m \times n$ matrix with reduced row echelon form R. Then there exists an invertible $m \times m$ matrix P such that $PA = R$.

Corollary

The matrix equation $A\boldsymbol{x}=\boldsymbol{b}$ has the same solutions as $R\boldsymbol{x}=\boldsymbol{c}$ where $[R\ \boldsymbol{c}]$ is the reduced row echelon form of the augmented matrix $[A\ \boldsymbol{b}]$.

Proof:

$\exists$ an invertible matrix *P* such that $P[A \quad \boldsymbol{b}] = [PA \quad P\boldsymbol{b}] = [R \quad \boldsymbol{c}]$

$\leftarrow$by Theorem 2.3.

Therefore, $PA = R$ and $P\boldsymbol{b} = \boldsymbol{c}$.

Since *P* is invertible , $A = P^{-1}R$ and $\boldsymbol{b} = P^{-1}\boldsymbol{c}$.

$\Rightarrow$

Suppose that ***v*** is a solution of $A\boldsymbol{x} = \boldsymbol{b}$. …. $\rightarrow$ $R\boldsymbol{v} = \boldsymbol{c}$.

$\Leftarrow$

Suppose that ***v*** is a solution of $R\boldsymbol{x} = \boldsymbol{c}$. …. $\rightarrow$ $A\boldsymbol{v} = \boldsymbol{b}$.

The Column Correspondence Property

Let $S=\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_6\}$, where $\boldsymbol{a}_i \in R^4$ is the i-th column of A in the following.

- Since S, with $|S|=6$, contains more than n vectors, $n=4$, the set S is linearly dependent.
  - ✓ At least one of the vectors in S is a linear combination of the others.
  - ✓ $A\boldsymbol{x}=\boldsymbol{0}$ has nonzero solutions.
- Which vector in S is a linear combination of other vectors?

$$A = [\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \cdots \ \ \boldsymbol{a}_6] = \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ -1 & -2 & 1 & 2 & 3 & 6 \\ 2 & 4 & -3 & 2 & 0 & 3 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix}$$

$$R = [\boldsymbol{r}_1 \ \boldsymbol{r}_2 \ \cdots \ \boldsymbol{r}_6] = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- R is the reduced row echelon form of A. Then $A\boldsymbol{x}=\boldsymbol{0}$ and $R\boldsymbol{x}=\boldsymbol{0}$ have the same solutions.
- Let $S' = \{\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_6\}$. We see such a correspondence:

$$\boldsymbol{r}_2 = 2\boldsymbol{r}_1 \quad \leftrightarrow \quad \boldsymbol{a}_2 = 2\boldsymbol{a}_1$$

$$\boldsymbol{r}_5 = -\boldsymbol{r}_1 + \boldsymbol{r}_4 \quad \leftrightarrow \quad \boldsymbol{a}_5 = -\boldsymbol{a}_1 + \boldsymbol{a}_4$$

Column Correspondence Property

Let A be a matrix and R its reduced row echelon form. If column j of R is a linear combination of other columns of R, then the column j of A is a linear combination of the corresponding columns of A using the same coefficients and vice versa.

Proof

$\Rightarrow$

$\exists$ An invertible matrix P such that $PA = R$ (by Theorem 2.3)

Hence $P\boldsymbol{a}_i = \boldsymbol{r}_i \ \forall\, i$.

Suppose that column j of A is a linear combination of other columns of A.

That is,

$\boldsymbol{a}_j = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k$ for some scalars $c_1, c_2, \cdots, c_k$.

Therefore

$$
\begin{aligned}
\boldsymbol{r}_j = P\boldsymbol{a}_j &= P(c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k) \\
&= c_1P\boldsymbol{a}_1 + c_2P\boldsymbol{a}_2 + \cdots + c_kP\boldsymbol{a}_k \\
&= c_1\boldsymbol{r}_1 + c_2\boldsymbol{r}_2 + \cdots + c_k\boldsymbol{r}_k
\end{aligned}
$$

$\Leftarrow$ (Apply $\boldsymbol{a}_j = P^{-1}\boldsymbol{r}_j$)

Properties of a Matrix in Reduced Row Echelon form:

Let R be an $m \times n$ matrix in reduced row echelon form:

a) A column of R is a pivot column iff it is nonzero and not a linear combination of the preceding columns of R.

b) The j-th pivot column of R is $\boldsymbol{e}_j$.
   - ✓ The pivot columns of R are linearly independent.

c) Suppose that $\boldsymbol{r}_j$ is not a pivot column of R, and there are k pivot columns of R preceding it. Then $\boldsymbol{r}_j$ is a linear combination of the k preceding pivot columns.
   - ✓ The coefficients of the linear combination are the first k entries of $\boldsymbol{r}_j$.
   - ✓ The other entries of $\boldsymbol{r}_j$ are zeros.

Theorem 2.4

The following statements are true for any matrix A:

a) The pivot columns of A are linearly independent.

b) Each nonpivot column of A is a linear combination of the previous pivot columns of A, where the coefficients of the linear combination are the entries of the corresponding column of the reduced row echelon form of A.

Example 4:

The reduced row echelon form of A= $\begin{bmatrix} 1 & ? & 2 & ? \\ 2 & ? & 2 & ? \\ 1 & ? & 3 & ? \end{bmatrix}$ is $R = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Then, by the column correspondence property

- $\boldsymbol{a}_2 = 2\boldsymbol{a}_1 = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$ since $\boldsymbol{r}_2 = 2\boldsymbol{r}_1$, and
- $\boldsymbol{a}_4 = -\boldsymbol{a}_{\mathbf{1}} + \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ since $\boldsymbol{r}_4 = -\boldsymbol{r}_{\mathbf{1}} + \boldsymbol{r}_3$

2.4 The Inverse of A Matrix

Theorem 2.5

Let A be an $n \times n$ matrix. Then A is invertible iff the reduced row echelon form of A is I_n.

Proof of Theorem 2.5

$\Rightarrow$

Suppose that A is invertible.

- Then $A\boldsymbol{x}=\boldsymbol{0}$ has the unique solution $\boldsymbol{x} = A^{-1}\mathbf{0} = \mathbf{0}$ (on slide 16)
  - ✓ $\mathrm{Rank}\, A = n$. (by Theorem 1.8)
  - ✓ The reduced row echelon form of A is I_n. (on slide 37 of chapter 1)

$\Leftarrow$

Suppose that the reduced row echelon form of A is I_n.

- $\exists$ an invertible $n \times n$ matrix P such that $PA = I_n$ (by Theorem 2.3).
  - ✓ $A = (P^{-`1}P)A = P^{-1}(PA) = P^{-1}I_n = P^{-1}$.
  - ✓ P^{-1} is invertible (by Theorem 2.2), and therefore A is invertible.

Let R be the reduced row echelon form of $n \times n$ matrix A by means of elementary row operations, that is $PA = R$, where $P = E_k E_{k-1} \cdots E_1$. By applying the same elementary row operations to the $n \times 2n$ augmented matrix $[A \;\; I_n]$, we obtain an $n \times 2n$ augmented matrix $[R \;\; B]$ for some $n \times n$ matrix B.

Hence, $\exists$ an invertible matrix P such that

$$[R \quad B] = P[A \;\; I_n] = [PA \;\; PI_n] = [PA \quad P\,]$$

Algorithm for Matrix Inversion

Let A be an $n \times n$ matrix. Use elementary row operations to transform $[A\ I_n]$ into the form $[R\ B]$, where R is a matrix in reduced row echelon form. Then either

a) $R=I_n$, in which case A is invertible and $B=A^{-1}$; or

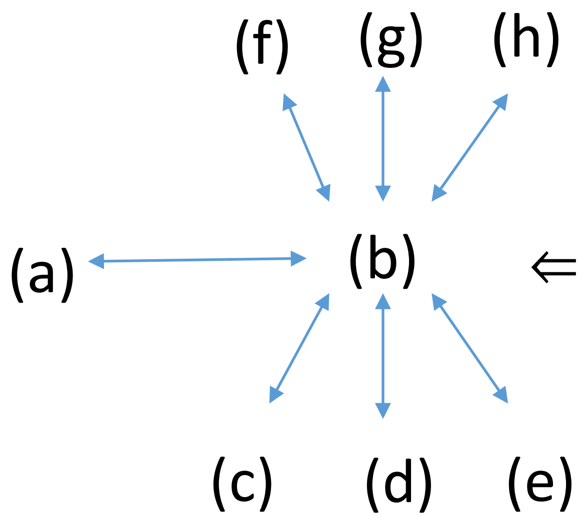
b) $R \neq I_n$, in which case A is not invertible.

Theorem 2.6 (**Invertible Matrix Theorem**)

Let A be an $n \times n$ matrix. The following statements are equivalent:

a) A is invertible.
b) The reduced row echelon form of A is I_n.
c) The rank of A equals n.
d) The span of the columns of A is R^n.
e) The equation $A\mathbf{x}=\boldsymbol{b}$ is consistent for every $\boldsymbol{b}$ in R^n.
f) The nullity of A equals zero.
g) The columns of A are linearly independent.
h) The only solution of $A\mathbf{x}=\boldsymbol{0}$ is $\boldsymbol{0}$.
i) There exists an $n \times n$ matrix B such that $BA=I_n$.
j) There exists an $n \times n$ matrix C such that $AC=I_n$.
k) A is a product of elementary matrices.

Proof:



$\Leftarrow$ By Theorem 1.8

$\Leftarrow$ By Theorem 2.5

$\Leftarrow$ By Theorem 1.6

$(a) \overset{?}{\leftrightarrow} (k)$ proved by $(a) \rightarrow (k)$ using Theorems 2.3 and 2.5 and then $(k) \rightarrow (a)$

$(a) \overset{?}{\leftrightarrow} (i)$ proved by $(a) \rightarrow (i) \rightarrow (h) \rightarrow (a)$

$(a) \overset{?}{\leftrightarrow} (j)$ proved by $(a) \rightarrow (j) \rightarrow (e) \rightarrow (a)$

Algorithm for Computing $A^{-1}B$

Let A be an invertible $n \times n$ matrix and B be an $n \times p$ matrix. Suppose that the $n \times (n+p)$ matrix $[A \;\; B]$ is transformed by means of elementary row operations into the matrix $[I_n \;\; C]$ in reduced row echelon form. Then $C = A^{-1}B$.

$\exists$ an invertible matrix P such that

$$[I_n \;\; C] = P[A \;\; B] = [PA \;\; PB]$$

- $PA = I_n$
- $C = PB = A^{-1}B$

2.5 Partitioned matrices and block multiplication

By drawing horizontal and vertical lines within a matrix, the matrix is divided into an array of submatrices called **blocks**.

- The resulting array is called a **partition** of the matrix.
- The process of forming these blocks is called **partitioning**.

$$A = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 6 & 8 & 5 & 0 \\ -7 & 9 & 0 & 5 \end{array}\right] = \begin{bmatrix} I_2 & O \\ B & 5I_2 \end{bmatrix} \quad \text{where } I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 6 & 8 \\ -7 & 9 \end{bmatrix}$$

Then $A^2 = \begin{bmatrix} I_2 & O \\ B & 5I_2 \end{bmatrix}\begin{bmatrix} I_2 & O \\ B & 5I_2 \end{bmatrix} = \begin{bmatrix} I_2 & O \\ 6B & 5^2 I_2 \end{bmatrix}$ and A^3=?

Block Multiplication

Suppose two matrices A and B are partitioned into blocks so that the number of blocks in each row of A is the same as the number of blocks in each column of B. Then the matrices can be multiplied according to the usual rules for matrix multiplication, treating the blocks as if they were scalars, provided that the individual products are defined.

Two additional methods for computing a matrix product AB

Assume that A is an $m \times n$ matrix with rows $\boldsymbol{a}_1', \boldsymbol{a}_2', \cdots, \boldsymbol{a}_m'$ and B is an $n \times p$ matrix with rows $\boldsymbol{b}_1', \boldsymbol{b}_2', \cdots, \boldsymbol{b}_n'$.

1. **By rows** The i-th row of AB is obtained by multiplying the i-th row of A by B; that is , $\boldsymbol{a}_i' B$.
2. **By outer products** The matrix AB is sum of matrix products of each column of A with the corresponding row of B; that is,

$$AB = \sum_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{b}_i'$$

Suppose that the product AB is defined.

- By columns:

$$AB = A[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_p] = [A\boldsymbol{b}_1 \quad A\boldsymbol{b}_2 \quad \cdots \quad A\boldsymbol{b}_p]$$

- By rows: $AB = \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_m' \end{bmatrix} B = \begin{bmatrix} \boldsymbol{a}_1' B \\ \boldsymbol{a}_2' B \\ \vdots \\ \boldsymbol{a}_m' B \end{bmatrix}$

- By outer products

$$AB = [\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots \quad \boldsymbol{a}_n] \begin{bmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \\ \vdots \\ \boldsymbol{b}_n' \end{bmatrix} = \boldsymbol{a}_1 \boldsymbol{b}_1' + \boldsymbol{a}_2 \boldsymbol{b}_2' + \cdots + \boldsymbol{a}_n \boldsymbol{b}_n'$$

Let $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$. Their outer product is $\boldsymbol{v}\boldsymbol{w}^{\boldsymbol{T}} = \begin{bmatrix} v_1 \boldsymbol{w}^T \\ v_2 \boldsymbol{w}^T \\ \vdots \\ v_m \boldsymbol{w}^T \end{bmatrix}$.

✓ Rank ($\boldsymbol{v}\boldsymbol{w}^{\boldsymbol{T}}$), or each rank ($\boldsymbol{a}_i \boldsymbol{b}_j'$) , ≤1.

2.6 The LU decomposition of a matrix

- **Upper triangular matrix**,
  The entries below and to the left of the *diagonal* entries are zeros.
- **Lower triangular matrix**,
  The entries above and to the right of the *diagonal* entries are zeros.
- **Unit lower triangular matrix**,
  A lower triangular matrix whose *diagonal* entries are all ones.

$$\begin{bmatrix} 2 & 0 & 1 & -1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 3 & 4 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}$$

Suppose that an $m\times n$ matrix A can be transformed into a matrix U in row echelon form without the use of row interchange. Then,

$$U = E_k E_{k-1} \cdots E_1 A,$$

where $E_1, \cdots, E_{k-1}, E_k$ are the elementary matrices corresponding to the elementary row operations that transform A into U.

Thus,

$$A = LU,$$

where $$L = (E_k E_{k-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

- U is an $m\times n$ matrix, an upper triangular matrix.
- L is an $m\times m$ matrix, a unit lower triangular matrix.

$$c\boldsymbol{r}_i + \boldsymbol{r}_j \rightarrow \boldsymbol{r}_j \ :$$

$$E = \begin{array}{r} \\ \\ row\ i \rightarrow \\ \\ row\ j \rightarrow \\ \\ \end{array} \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 1 & & & \\ & & \vdots & \ddots & & \\ & & c & & \ddots & 0 \\ 0 & 0 & & & 0 & 1 \end{bmatrix}$$

$$-c\boldsymbol{r}_i + \boldsymbol{r}_j \rightarrow \boldsymbol{r}_j$$

$$E^{-1} = \begin{array}{r} \\ \\ row\ i \rightarrow \\ \\ row\ j \rightarrow \\ \\ \end{array} \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 1 & & & \\ & & \vdots & \ddots & & \\ & & -c & & \ddots & 0 \\ 0 & 0 & & & 0 & 1 \end{bmatrix}$$

Definition

For any matrix A, a factorization $A=LU$, where L is a unit lower triangular matrix and U is an upper triangular matrix, is called an ***LU* decomposition** of A.

- If a matrix has an LU decomposition and is also invertible, then the LU decomposition is unique (proof ?).

The *LU* Decomposition of an $m \times n$ matrix A

(a) Use step 1, 3, and 4 of Gaussian elimination (in Sec. 1.4) to transform A into a matrix U in row echelon form by means of elementary row operations. If this is impossible, then A has no LU decomposition.

(b) While performing (a), create an $m \times m$ matrix L as follows

(i) Each diagonal entry of L is 1.

(ii) If some elementary row operation in (a) adds c times row j of a matrix to row i, then $l_{ij} = -c$; otherwise $l_{ij} = 0$.

Using an LU decomposition to solve a system of linear equations of a matrix

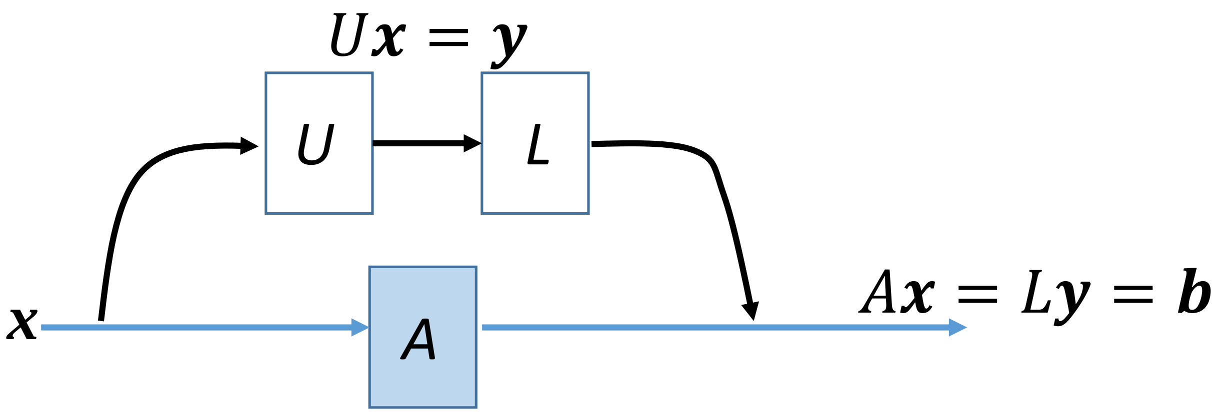
Given a system of linear equations, $A\boldsymbol{x} = \boldsymbol{b}$, where A has an LU decomposition $A=LU$. Then,

$$A\boldsymbol{x} = LU\boldsymbol{x} = L(U\boldsymbol{x}) = \boldsymbol{b}.$$

We can set

$$U\boldsymbol{x} = \boldsymbol{y} \text{ and } L\boldsymbol{y} = \boldsymbol{b}$$

and solve $\boldsymbol{y}$ first and then $\boldsymbol{x}$.

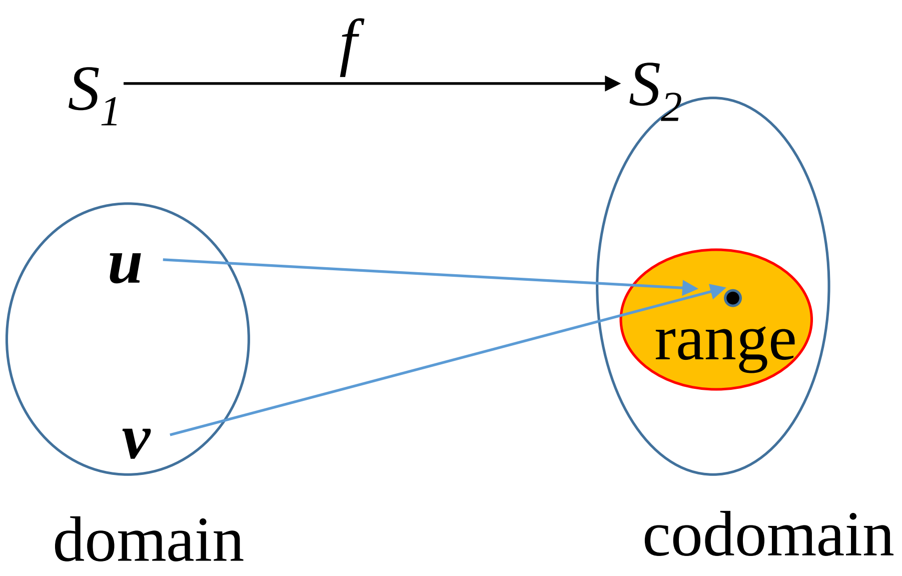


What if a matrix has no LU decomposition?

(Self-study)

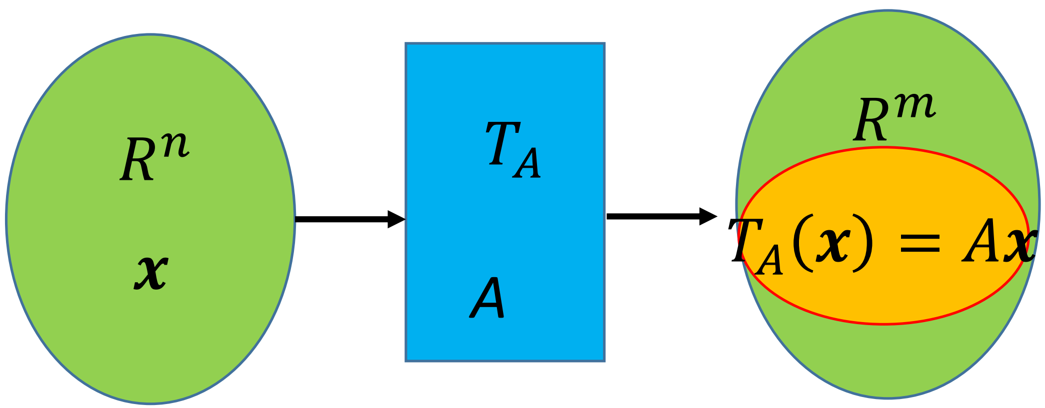
2.7 Linear Transformations and Matrices

Definition Let S_1 and S_2 be subsets of R^n and R^m, respectively. A **function** f from S_1 to S_2, written $f : S_1 \to S_2$, is a rule that assigns to each vector $\boldsymbol{v}$ in S_1 a unique vector $f(\boldsymbol{v})$ in S_2. The vector $f(\boldsymbol{v})$ is called the **image** of $\boldsymbol{v}$ (under f). The set S_1 is called the **domain** of a function f, and the set S_2 is called the **codomain** of f. The **range** of f is defined to be the set of images $f(\boldsymbol{v})$ for all $\boldsymbol{v}$ in S_1.



Definition Let A be an $m \times n$ matrix. The function $T_A : R^n \to R^m$ defined by $T_A(\boldsymbol{x})=A\boldsymbol{x}$ for all $\boldsymbol{x}$ in R^n is called the **matrix transformation induced** by A.

Let $T_A : R^3 \to R^3$ be defined by $T_A\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$. Thus, $T_A(\boldsymbol{u})$ is the *orthogonal projection* of $\boldsymbol{u}$ on the *xy*-plane.



Theorem 2.7.

For any $m \times n$ matrix A and any vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in R^n, the following statements are true:

a) $T_A(\boldsymbol{u}+\boldsymbol{v})=T_A(\boldsymbol{u})+T_A(\boldsymbol{v})$.

b) $T_A(c\boldsymbol{u})=cT_A(\boldsymbol{u})$ for every scalar c.

???: (Example 1) f: $\mathcal{R}^3 \to \mathcal{R}^2$, defined by the rule $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1^2 \end{bmatrix}$, does not satisfy either of the above properties.

Definition A function $T : R^n \to R^m$ is called a **linear transformation** (or simply **linear**) if, for all vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in R^n and all scalars c, both of the following conditions hold:

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$. (In this case, T **preserves vector addition**.)

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$. (In this case, T **preserves scalar multiplication**.)

Every matrix transformation is linear (by Theorem 2.7).

Two special L.T. (Linear transformations):

- The **identity transformation** $I : R^n \to R^n$, defined by $I(\mathbf{x})=\mathbf{x}\ \ \forall \mathbf{x} \in R^n$.
- The **zero transformation** $T_0 : R^n \to R^m$, defined $T_0(\mathbf{x})=\mathbf{0}\ \ \forall \mathbf{x} \in R^n$.

Theorem 2.8

For any linear transformation $T : R^n \to R^m$, the following statements are true:

a) $T(\mathbf{0})=\mathbf{0}$.

b) $T(-\mathbf{u}) = -T(\mathbf{u}) \ \forall \mathbf{u} \in R^n$.

c) $T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) \ \forall \mathbf{u}, \mathbf{v} \in R^n$

d) $T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v}) \ \forall \mathbf{u}, \mathbf{v} \in R^n$ and all scalars a and b.

Let $T : R^n \to R^m$ be a L.T. If $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ are vectors in R^n and $a_1, a_2, \cdots, a_k$ are scalars, then

$T(a_1\mathbf{u}_1+a_2\mathbf{u}_2+\cdots+a_k\mathbf{u}_k) = a_1T(\mathbf{u}_1) + a_2T(\mathbf{u}_2) + \cdots + a_kT(\mathbf{u}_k)$.

Every L.T. with domain R^n and codomain R^m , denoted by T, is a matrix transformation, denoted by T_A.

Theorem 2.9

Let $T : R^n \to R^m$ be linear. Then there is a unique $m \times n$ matrix

$$A = [T(\boldsymbol{e}_1)T(\boldsymbol{e}_2)\cdots T(\boldsymbol{e}_n)],$$

whose columns are the images under T of the standard vectors for R^n, such that $T(v)=Av$ $\forall \boldsymbol{v} \in R^n$.

Proof ?

- Let $T : R^n \to R^m$ be an L.T.. The following $m \times n$ matrix
$$A = [T(\boldsymbol{e}_1)T(\boldsymbol{e}_2)\cdots T(\boldsymbol{e}_n)]$$
is called the **standard matrix** of T.
- By Theorem 2.9, the standard matrix A of T has the property that $T(\boldsymbol{v})=A\boldsymbol{v}$ for every $\boldsymbol{v} \in R^n$

2.8 Composition and Invertibility of Linear Transformations

Use standard matrix to study some basic properties of an L.T.

- Onto and one-to-one
- The existence and uniqueness of solutions of systems of linear equations

Example:

Suppose that $T : R^3 \to R^2$, a L.T, is defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 - 4x_2 \\ 2x_1 + x_3 \end{bmatrix}.$$

The standard matrix of T is (by Theorem 2.9)

$$A = [T(\boldsymbol{e_1})\ T(\boldsymbol{e_2})\ T(\boldsymbol{e_3})] = \begin{bmatrix} 3 & -4 & 0 \\ 2 & 0 & 1 \end{bmatrix} .$$

Now $\boldsymbol{w}$ is in the range of T iff $\boldsymbol{w}=T(\boldsymbol{v})$ for some $\boldsymbol{v}$ in R^3.

- $\Leftarrow$ Write $\boldsymbol{v}=v_1\boldsymbol{e}_1+v_2\boldsymbol{e}_2+v_3\boldsymbol{e}_3$. Then

$$\boldsymbol{w} = T(\boldsymbol{v}) = T(v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2+v_3\boldsymbol{e}_3) = v_1T(\boldsymbol{e}_1) + v_2T(\boldsymbol{e}_2)+v_3T(\boldsymbol{e}_3),$$

which is a linear combination of the columns of A is in the range of T.

- $\Rightarrow$ Likewise, every linear combination of the columns A is in the range of T.

Conclusion: The range of T equals the span of $\left\{\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$. $\Downarrow$

The range of an L.T. equals the span of the columns of its standard matrix.

Definition A function $f : R^n \to R^m$ is said to be **onto** if its range is all of R^m; that is, if every vector in R^m is an image.

- An L.T. is onto (surjective) iff the columns of its standard matrix form a generating set for its codomain.
- If standard matrix A is an $m \times n$ matrix, the columns of A form a generating set for R^m iff rank $A=m$ (by Theorem 1.6).

Theorem 2.10

Let $T : R^n \to R^m$ be an L.T. with standard matrix A.
The following conditions are equivalent:

a) T is *onto*; that is, the range of T is R^m.
b) The columns of A form a generating set for R^m.
c) Rank $A = m$.

Definition A function $f : R^n \to R^m$ is said to be **one-to-one** if every pair of distinct vectors in R^n has distinct images. That is, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are distinct vectors in R^n, then $f(\boldsymbol{u})$ and $f(\boldsymbol{v})$ are distinct vectors in R^m.

- Suppose that $T : R^n \to R^m$ is a one-to-one L.T. If $\boldsymbol{w}$ is a nonzero vector in R^n, then $T(\boldsymbol{w}) \neq T(\boldsymbol{0}) = \boldsymbol{0}$. Hence $\boldsymbol{0}$ is the only vector in R^n whose image under T is the zero vector of R^m.
- Conversely, if $\boldsymbol{0}$ is the only vector whose image under T is the zero vector, then T must be one-to-one.

$\Leftarrow$ Suppose that there are vectors $\boldsymbol{u} \neq \boldsymbol{v}$ in R^n such that $T(\boldsymbol{u}) = T(\boldsymbol{v})$. Then $T(\boldsymbol{u}-\boldsymbol{v}) = T(\boldsymbol{u}) - T(\boldsymbol{v}) = \boldsymbol{0}$, and hence there exists a nonzero vector $(\boldsymbol{u}-\boldsymbol{v}) \in R^{\text{n}}$ such that $T((\boldsymbol{u}-\boldsymbol{v})) = \boldsymbol{0}$! (A contradiction)

Definition Let $T: R^n \to R^m$ be linear. The **null space** of T is the set of all $\boldsymbol{v}$ in R^n such that $T(\boldsymbol{v})=\boldsymbol{0}$.

An L.T. is one-to-one iff its null space contains only $\boldsymbol{0}$.

Theorem 2.11

Let $T: R^n \to R^m$ be an L.T. with standard matrix A. Then the following statements are equivalent:

a) T is one-to-one.
b) The null space of T consists only of the zero vector.
c) The columns of A are linearly independent.
d) rank $A=n$.

(Refer to Theorem 1.8 for the above.)

Three related topics:

- Systems of linear equations, $A\boldsymbol{x}=\boldsymbol{b}$, where A is an $m\times n$ matrix, $\boldsymbol{x}$ a vector in R^n, and $\boldsymbol{b}$ a vector in R^m. It can written in the form $T_A(\boldsymbol{x})=\boldsymbol{b}$.
- Standard matrix A.
- Linear transformation, T_A
  a) $A\boldsymbol{x}=\boldsymbol{b}$ has a solution iff $\boldsymbol{b}$ is in the range of T_A.
  b) $A\boldsymbol{x}=\boldsymbol{b}$ has a solution for every $\boldsymbol{b}$ iff T_A is onto (surjective).
  c) $A\boldsymbol{x}=\boldsymbol{b}$ has at most one solution for every $\boldsymbol{b}$ iff T_A is one-to-one (injective).

Composition of Linear Transformations

If $f : S_1 \to S_2$ and $g : S_2 \to S_3$, then the *composition* $g \circ f : S_1 \to S_3$ is defined by $(g \circ f)(\mathbf{u}) = g(f(\mathbf{u}))\ \forall\, \mathbf{u} \in S_1$.
In linear algebra we drop the "circle" notation.

Suppose that A is an $m \times n$ matrix and B a $p \times m$ matrix. Then BA is a $p \times n$ matrix.

- The corresponding matrix transformations are $T_A : R^n \to R^m$, $T_B : R^m \to R^p$, and $T_{BA} : R^n \to R^p$.
- For every $\mathbf{v}$ in R^n,

$$\mathrm{T}_{\mathrm{BA}}(\mathbf{v}) = BA(\mathbf{v}) = B(A(\mathbf{v})) = B(T_A(\mathbf{v})) = T_B(T_A(\mathbf{v})) = T_B T_A(\mathbf{v}).$$

Thus, $T_B T_A = T_{BA}$.

Theorem 2.12

If $T: R^n \to R^m$ and $U: R^m \to R^p$ are linear transformations with standard matrices A and B, respectively, the composition $UT: R^n \to R^p$ is also linear, and its standard matrix is BA.

A function f: $\mathcal{S}_1 \to \mathcal{S}_2$ is *invertible* if there exists a function g:$\mathcal{S}_2 \to \mathcal{S}_1$ such that $g(f(\boldsymbol{v}))=\boldsymbol{v}$ $\forall \boldsymbol{v} \in \mathcal{S}_1$ and $f(g(v))=\boldsymbol{v}$ $\forall \boldsymbol{v} \in \mathcal{S}_2$.

- If f is invertible, then the function g is unique and is called the *inverse* of f, denoted as f^{-1}.

- A function is invertible iff it is one-to-one (injective) and onto (surjective).
- Suppose that A is an $n \times n$ invertible matrix.
  - ✓ Then, $\forall \boldsymbol{v} \in \mathcal{R}^n$, we have $T_A T_{A^{-1}}(\boldsymbol{v}) = T_A(A^{-1}\boldsymbol{v}) = AA^{-1}\boldsymbol{v} = I_n\boldsymbol{v} = \boldsymbol{v}$.
  - ✓ Likewise, $T_{A^{-1}} T_A(\boldsymbol{v}) = \boldsymbol{v}$.

  Thus T_A is invertible and $T_A^{-1} = T_{A^{-1}}$.
- If $T : R^n \to R^n$ is linear and invertible, then it is also one-to-one.
  - ✓ Its standard matrix A has rank n (by Theorem 2.11)and is thus invertible.
  - ✓ $T^{-1} = T_{A^{-1}}$. T^{-1} is a matrix transformation and thus linear.

Theorem 2.13

Let $T : R^n \to R^n$ be an L.T. with standard matrices A. Then T is invertible iff A is invertible, in which case $T^{-1} = T_{A^{-1}}$. Thus, T^{-1} is linear, and its standard matrix is A^{-1}.

(Injection + Surjection $\Leftrightarrow$ Bijection)

Let $T: R^n \to R^m$ be an L.T. with standard matrix A, which has size $m \times n$. Properties listed in the same row of the following table are *equivalent*.

Property of T	**The number of solutions of $Ax=b$**	**Property of the columns of A**	**Property of the rank of A**
T is onto.	$A\mathbf{x}=\mathbf{b}$ has at least one solution for every $\mathbf{b}$ in R^m.	The columns of A are a generating set of R^m.	rank $A = m$
T is one-to-one.	$A\mathbf{x}=\mathbf{b}$ has at most one solution for every $\mathbf{b}$ in R^m.	The columns of A are linearly independent.	rank $A = n$
T is invertible.	$A\mathbf{x}=\mathbf{b}$ has a unique solution for every $\mathbf{b}$ in R^m.	The columns of A are a linearly independent generating set for R^m.	rank $A = m = n$