

Chapter 2: Matrices and Determinants

March 16, 2009

1 Linear transformations

Lecture 7

Definition 1.1. Let X and Y be nonempty sets. A **function** from X to Y is a rule $f : X \rightarrow Y$ such that each element x in X is assigned a unique element y in Y , written as

$$y = f(x);$$

the element y is called the **image** of x under f , and the element x is called the **preimage** of $f(x)$. Functions are also called **maps**, or **mappings**, or **transformations**.

Definition 1.2. A function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called a **linear transformation** if for any vectors \mathbf{u}, \mathbf{v} in \mathbb{R}^n and scalar c ,

(a) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$,

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

Example 1.1. (a) The function $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, defined by $T(x_1, x_2) = (x_1 + 2x_2, x_2)$, is a linear transformation, see Figure 1

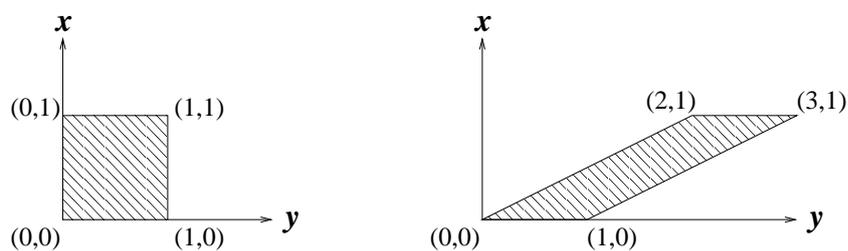


Figure 1: The geometric shape under a linear transformation.

(b) The function $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, defined by $T(x_1, x_2) = (x_1 + 2x_2, 3x_1 + 4x_2, 2x_1 + 3x_2)$, is a linear transformation.

(c) The function $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, defined by $T(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 3x_1 + 2x_2 + x_3)$, is a linear transformation.

Example 1.2. The transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by $T(\mathbf{x}) = A\mathbf{x}$, where A is an $m \times n$ matrix, is a linear transformation.

Example 1.3. The map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, defined by $T(\mathbf{x}) = \lambda\mathbf{x}$, where λ is a constant, is a linear transformation, and is called the **dilation** by λ .

Example 1.4. The reflection $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ about a straight line through the origin is a linear transformation.

Example 1.5. The rotation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ about an angle θ is a linear transformation, see Figure 3.

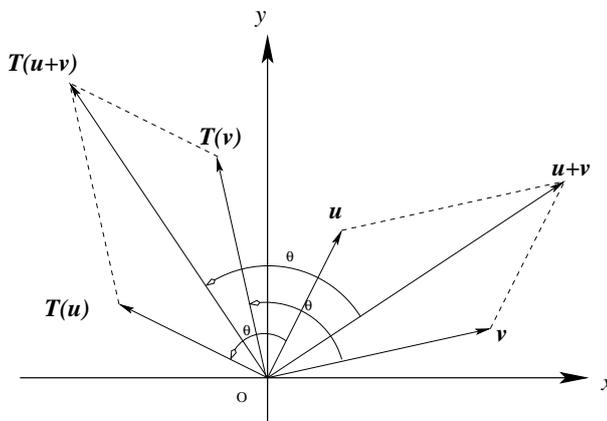


Figure 2: Rotation about an angle θ .

Proposition 1.3. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Then for any vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in \mathbb{R}^n and scalars c_1, c_2, \dots, c_k , we have

$$T(c_1\mathbf{v}_1 + \mathbf{v}_2 + \dots + c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \dots + c_kT(\mathbf{v}_k).$$

In particular,

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

Theorem 1.4. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in \mathbb{R}^n .

- (a) If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly dependent, then $T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_k)$ are linearly dependent.
- (b) If $T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_k)$ are linearly independent, then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent.

2 Standard matrix of linear transformation

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Consider the following vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

in the coordinate axis of \mathbb{R}^n . It is clear that for every vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \dots, x_n]^T$ is a

linear combination of $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, i.e.,

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n.$$

Then it follows from linearity that

$$T(\mathbf{x}) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n).$$

This means that $T(\mathbf{x})$ is completely determined by the images $T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)$. The ordered set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is called the **standard basis** of \mathbb{R}^n .

Definition 2.1. The **standard matrix** of a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the $m \times n$ matrix

$$A = [T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)] = [T\mathbf{e}_1, T\mathbf{e}_2, \dots, T\mathbf{e}_n].$$

Proposition 2.2. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation whose standard matrix is A . Then

$$T(\mathbf{x}) = A\mathbf{x}.$$

Proof.

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= [T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)]\mathbf{x} = A\mathbf{x}. \end{aligned}$$

□

Example 2.1. (a) The linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T(x_1, x_2) = (x_1 + 2x_2, 3x_1 + 4x_2)$, can be written as the matrix form

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(b) Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $T(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 3x_1 + 2x_2 + 1x_3)$. Then T is a linear transformation and its standard matrix is given by

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) &= \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 3x_1 + 2x_2 + x_3 \end{bmatrix} \\ &= x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned}$$

Example 2.2. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a rotation about an angle θ counterclockwise. Then

$$T(\mathbf{e}_1) = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}.$$

Thus

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Lecture 8

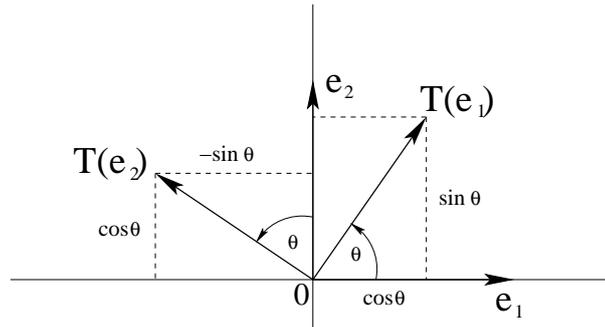


Figure 3: Rotation about an angle θ .

Proposition 2.3. Let $F, G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be linear transformations with the standard matrices A and B , respectively, that is,

$$F(\mathbf{x}) = A\mathbf{x}, \quad G(\mathbf{x}) = B\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n.$$

Then $F \pm G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are linear transformation with standard matrix $A \pm B$, that is,

$$(F \pm G)(\mathbf{x}) = F(\mathbf{x}) \pm G(\mathbf{x}) = (A \pm B)\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n;$$

and for any scalar c , $cF : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation with the standard matrix cA , that is,

$$(cF)(\mathbf{x}) = cF(\mathbf{x}) = (cA)\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n.$$

Example 2.3. Let $F, G : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be linear transformations defined by (writing in coordinates)

$$F(x_1, x_2, x_3) = (x_1 - x_2, x_2 - x_3),$$

$$G(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3).$$

Writing in vectors,

$$\begin{aligned} F \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) &= \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} G \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) &= \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned}$$

We then have

$$\begin{aligned}
(F + G) \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) &= \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \end{bmatrix} + \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \\
(F - G) \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) &= \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \end{bmatrix} - \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix} \\
&= \begin{bmatrix} -2x_2 \\ -2x_3 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.
\end{aligned}$$

Hence

$$\begin{aligned}
\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}; \\
\begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.
\end{aligned}$$

Definition 2.4. The **composition** of a function $f : X \rightarrow Y$ and a function $g : Y \rightarrow Z$ is a function

$$g \circ f : X \rightarrow Z, \quad (g \circ f)(x) = g(f(x)), \quad x \in X.$$

Theorem 2.5. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation with standard matrix A . Let $G : \mathbb{R}^p \rightarrow \mathbb{R}^n$ be a linear transformation with standard matrix B . Then the composition $F \circ G : \mathbb{R}^p \rightarrow \mathbb{R}^m$ is a linear transformation with the standard matrix AB . Symbolically, we have

$$\mathbb{R}^p \xrightarrow{B} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m \implies \mathbb{R}^p \xrightarrow{AB} \mathbb{R}^m.$$

Proof. For vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^p$ and scalars $a, b \in \mathbb{R}$, we have

$$\begin{aligned}
(F \circ G)(a\mathbf{u} + b\mathbf{v}) &= F(G(a\mathbf{u} + b\mathbf{v})) \\
&= F(aG(\mathbf{u}) + bG(\mathbf{v})) \\
&= aF(G\mathbf{u}) + bF(G\mathbf{v}) \\
&= a(F \circ G)(\mathbf{u}) + b(F \circ G)(\mathbf{v}).
\end{aligned}$$

Then the composition map $F \circ G$ is a linear transformation.

Write $A = [a_{ij}]_{m \times n}$, $B = [b_{jk}]_{n \times p}$. Let $C = [c_{ik}]_{m \times p}$ be the standard matrix of $F \circ G$. Let us write

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix}.$$

Then the linear transformations F , G , and $F \circ G$ can be written as

$$\mathbf{y} = G(\mathbf{x}) = B\mathbf{x}, \quad \mathbf{z} = F(\mathbf{y}) = A\mathbf{y}, \quad \mathbf{z} = (F \circ G)(\mathbf{x}) = C\mathbf{x}.$$

Writing in coordinates, we have

$$\begin{aligned} y_j &= \sum_{k=1}^p b_{jk}x_k, \quad 1 \leq j \leq n, \\ z_i &= \sum_{j=1}^n a_{ij}y_j = \sum_{k=1}^p c_{ik}x_k, \quad 1 \leq i \leq m. \end{aligned}$$

Substitute $y_j = \sum_{k=1}^p b_{jk}x_k$ into $z_i = \sum_{j=1}^n a_{ij}y_j$. We obtain

$$z_i = \sum_{j=1}^n a_{ij} \sum_{k=1}^p b_{jk}x_k = \sum_{k=1}^p \left(\sum_{j=1}^n a_{ij}b_{jk} \right) x_k.$$

Comparing the coefficients of the variables x_k , we conclude that

$$c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}, \quad 1 \leq i \leq m, \quad 1 \leq k \leq p.$$

For instance, let $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and $G : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be linear transformations given by

$$F(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad G(\mathbf{y}) = B\mathbf{y} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Then

$$G(\mathbf{y}) = \begin{bmatrix} b_{11}y_1 + b_{12}y_2 \\ b_{21}y_1 + b_{22}y_2 \end{bmatrix}.$$

Thus

$$\begin{aligned} (F \circ G)(\mathbf{y}) &= F(G(\mathbf{y})) = F \left(\begin{bmatrix} b_{11}y_1 + b_{12}y_2 \\ b_{21}y_1 + b_{22}y_2 \end{bmatrix} \right) \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11}y_1 + b_{12}y_2 \\ b_{21}y_1 + b_{22}y_2 \end{bmatrix} \\ &= \begin{bmatrix} a_{11}(b_{11}y_1 + b_{12}y_2) + a_{12}(b_{21}y_1 + b_{22}y_2) \\ a_{21}(b_{11}y_1 + b_{12}y_2) + a_{22}(b_{21}y_1 + b_{22}y_2) \end{bmatrix} \\ &= \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21})y_1 + (a_{11}b_{12} + a_{12}b_{22})y_2 \\ (a_{21}b_{11} + a_{22}b_{21})y_1 + (a_{21}b_{12} + a_{22}b_{22})y_2 \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right) \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \end{aligned}$$

□

Example 2.4. Consider the linear transformations

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad F(\mathbf{x}) = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ -x_1 + x_2 \\ x_1 + x_2 \end{bmatrix},$$

$$G: \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad G(\mathbf{x}) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}.$$

Then $F \circ G: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and $G \circ F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ are both linear transformations, and

$$\begin{aligned} (F \circ G)(\mathbf{x}) &= F(G(\mathbf{x})) = F\left(\begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}\right) \\ &= \begin{bmatrix} x_1 - x_3 \\ -x_1 + x_3 \\ x_1 + 2x_2 + x_3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} (G \circ F)(\mathbf{x}) &= G(F(\mathbf{x})) = G\left(\begin{bmatrix} x_1 - x_2 \\ -x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} 0 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

Hence

$$\begin{aligned} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \\ \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

Proposition 2.6. Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Write $B = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p]$. Then

$$AB = A[\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p] = [A\mathbf{b}_1, A\mathbf{b}_2, \dots, A\mathbf{b}_p].$$

Example 2.5. Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$. Write $B = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$. Then

$$A\mathbf{b}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix},$$

$$A\mathbf{b}_2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix},$$

$$A\mathbf{b}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

and

$$AB = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 4 \\ 5 & 8 & 5 \end{bmatrix}.$$

Example 2.6. Let $D_r : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a dilation by a real number r . The standard matrix of D_r is the diagonal matrix

$$\text{Diag}[r, \dots, r] = \begin{bmatrix} r & 0 & \cdots & 0 \\ 0 & r & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r \end{bmatrix}.$$

Example 2.7. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation with standard matrix A . Then

$$T \circ D_r = D_r \circ T.$$

The standard matrices of $T \circ D_r$ and $D_r \circ T$ are equal to rA .

Theorem 2.7. (a) If A is an $m \times n$ matrix, B and C are $n \times p$ matrices, then

$$A(B + C) = AB + AC.$$

(b) If A and B are $m \times n$ matrices, C is an $n \times p$ matrix, then

$$(A + B)C = AC + BC.$$

(c) If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then for any scalar a ,

$$a(AB) = (aA)B = A(aB).$$

(d) If A is an $m \times n$ matrix, then

$$I_m A = A = A I_n.$$

Definition 2.8. The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose (i, j) -entry is the (j, i) -entry of A , that is,

$$A^T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nm} \end{bmatrix}.$$

Example 2.8.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 6 \\ 5 & 6 & 7 & 8 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 5 & 7 \\ 4 & 6 & 8 \end{bmatrix}.$$

Proposition 2.9. Let A be an $m \times n$ matrix and B an $n \times p$ matrix. Then

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

Proof. Let c_{ij} be the (i, j) -entry of AB . Let b'_{ik} be the (i, k) -entry of B^T and a'_{kj} the (k, j) -entry of A^T . Then $b'_{ik} = b_{ki}$ and $a'_{kj} = a_{jk}$. Thus the (i, j) -entry of $(AB)^T$ is

$$c_{ji} = \sum_{k=1}^n a_{jk} b_{ki} = \sum_{k=1}^n b'_{ik} a'_{kj},$$

which is the (i, j) -entry of $B^T A^T$ by the matrix multiplication. \square

Theorem 2.10. Let $h : X \rightarrow Y$, $g : Y \rightarrow Z$, $f : Z \rightarrow W$ be functions. Then

$$(f \circ g) \circ h = f \circ (g \circ h).$$

Proof. For any $x \in X$,

$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x))) = f((g \circ h)(x)) = (f \circ (g \circ h))(x).$$

□

Corollary 2.11. Let A be an $m \times n$ matrix, B an $n \times p$ matrix, and C an $p \times q$ matrix. Then $(AB)C$ and $A(BC)$ are $m \times q$ matrices, and

$$(AB)C = A(BC).$$

Proof. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $F(\mathbf{x}) = A\mathbf{x}$, $G : \mathbb{R}^p \rightarrow \mathbb{R}^n$ with $G(\mathbf{x}) = B\mathbf{x}$, and $H : \mathbb{R}^q \rightarrow \mathbb{R}^p$ with $H(\mathbf{x}) = C\mathbf{x}$. Then

$$\mathbb{R}^q \xrightarrow{H} \mathbb{R}^p \xrightarrow{G} \mathbb{R}^n \xrightarrow{F} \mathbb{R}^m \implies \mathbb{R}^q \xrightarrow{F \circ (G \circ H)} \mathbb{R}^m \quad \text{and} \quad \mathbb{R}^q \xrightarrow{(F \circ G) \circ H} \mathbb{R}^m.$$

Since $F \circ (G \circ H) = (F \circ G) \circ H$, we thus have $A(BC) = (AB)C$. □

Lecture 9

Definition 2.12. Let $f : X \rightarrow Y$ be a function.

- (a) The function f is called **one-to-one** if distinct elements in X are mapped to distinct elements in Y .
- (b) The function f is called **onto** if every element y in Y is the image of some element x in X under f .

Example 2.9. (a) The transformation $F : \mathbb{R} \rightarrow \mathbb{R}^2$ by $F(x) = \begin{bmatrix} 2x \\ x \end{bmatrix}$ is 1-1, but not onto.

(b) The transformation $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ given by

$$F\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is 1-1, but not onto. The image set

$$F(\mathbb{R}^2) = \left\{ s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\}$$

is the xz -plane of \mathbb{R}^3 through the origin, not the whole space \mathbb{R}^3 .

Example 2.10. The transformation $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ given by

$$F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is not 1-1, but onto. The image set

$$\begin{aligned} F(\mathbb{R}^3) &= \left\{ x_1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\} \\ &= \left\{ s \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \end{bmatrix} : s, t \in \mathbb{R} \right\} = \mathbb{R}^2. \end{aligned}$$

Theorem 2.13 (Characterization of One-to-One and Onto). *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation with standard matrix A . Then*

- (a) T is one-to-one
 - $\iff T(\mathbf{x}) = \mathbf{0}$ has the only trivial solution.
 - \iff The column vectors of A are linearly independent.
 - \iff Every column of A has a pivot position.
- (b) T is onto
 - \iff The column vectors of A span \mathbb{R}^m .
 - \iff Every row of A has a pivot position.

Example 2.11. (a) The linear transformation $T_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, defined by

$$T_1 \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

is one-to-one but not onto. The column vectors are linearly independent, but can not span \mathbb{R}^3 .

(b) The linear transformation $T_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, defined by

$$T_2 \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

is onto but not one-to-one. The column vectors span \mathbb{R}^2 and but are linearly dependent.

(c) The linear transformation $T_3 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, defined by

$$T_3 \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

is neither one-to-one nor onto. The column vectors are linearly dependent and can not span \mathbb{R}^3 .

(e) The linear transformation $T_5 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, defined by

$$T_4 \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

is both one-to-one and onto. The column vectors are linearly independent and span \mathbb{R}^3 .

Corollary 2.14. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be defined by $T(\mathbf{x}) = A\mathbf{x}$.*

- (a) *If T is one-to-one, then $n \leq m$.*
- (b) *If T is onto, then $n \geq m$.*
- (c) *If T is one-to-one and onto, then $m = n$.*

Proof. (a) Since T is one-to-one, the system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Then the reduced row echelon form of A must be of the form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

This shows that every column of A has a pivot position. Thus $n \leq m$.

(b) Since T is onto, the reduced row echelon form B for A can not have zero rows. This means that every row of B has a pivot position. Note that every pivot positions must be in different columns. Hence $n \geq m$. \square

Corollary 2.15. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. The following statements are equivalent.*

- (a) T is one-to-one.
- (b) T is onto.
- (c) T is one-to-one and onto.

3 Invertible matrices

Lecture 10

Definition 3.1. A function $f : X \rightarrow Y$ is said to be **invertible** if there is a function $g : Y \rightarrow X$ such that

$$\begin{aligned} g \circ f(x) &= x \quad \text{for all } x \in X, \\ f \circ g(y) &= y \quad \text{for all } y \in Y. \end{aligned}$$

If so, the function g is called the **inverse** of f , and is denoted by $g = f^{-1}$. The function $\text{Id} : X \rightarrow X$, defined by $\text{Id}(x) = x$, is called the **identity function**.

Theorem 3.2. *A function $f : X \rightarrow Y$ is invertible if and only if f is one-to-one and onto.*

Proof. Assume that f is invertible. For two distinct elements u and v of X , if $f(u)$ and $f(v)$ are not distinct, i.e., $f(u) = f(v)$, then $u = g(f(u)) = g(f(v)) = v$, a contradiction. This means that f is 1-1. For any element w of Y , consider the element $u = g(w)$ of X . We have $f(u) = f(g(w)) = w$. This means that f is onto.

Conversely, assume that f is 1-1 and onto. For any element y of Y , there exists a unique element x in X such that $f(x) = y$. Define $g : Y \rightarrow X$ by $g(y) = x$, where $f(x) = y$. Then g is the inverse of f . \square

Note. If a linear transformation $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is invertible, then F must be one-to-one and onto. Hence $m = n$.

Definition 3.3. A linear transformation $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is said to be **invertible** if there is a linear transformation $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that

$$F \circ G = \text{Id}_n \quad \text{and} \quad G \circ F = \text{Id}_n,$$

where $\text{Id}_n : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the **identity transformation** defined by $\text{Id}_n(\mathbf{x}) = \mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$. The standard matrix of Id_n is the **identity matrix**

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Definition 3.4. A square $n \times n$ matrix A is said to be **invertible** if there is a square $n \times n$ matrix B such that

$$AB = I_n = BA.$$

If A is invertible, the matrix B is called the **inverse** of A , and is denoted by $B = A^{-1}$.

Theorem 3.5. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation.

- (a) If T is 1-1, then T is invertible.
- (b) If T is onto, then T is invertible.

In order to find out whether an $n \times n$ matrix A is invertible or not, it is to decide whether the matrix equation

$$AX = I_n$$

has a solution, and if it has a solution, the solution matrix is the inverse of A . Write $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$. Then $AX = I_n$ is equivalent to solving the following linear systems

$$A\mathbf{x}_1 = \mathbf{e}_1, \quad A\mathbf{x}_2 = \mathbf{e}_2, \quad \dots, \quad A\mathbf{x}_n = \mathbf{e}_n.$$

Perform row operation to the corresponding augmented matrices; we have

$$[A | \mathbf{e}_1] \xrightarrow{\sim} [I_n | \mathbf{b}_1], \quad [A | \mathbf{e}_2] \xrightarrow{\sim} [I_n | \mathbf{b}_2], \quad \dots, \quad [A | \mathbf{e}_n] \xrightarrow{\sim} [I_n | \mathbf{b}_n].$$

The corresponding solutions $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ can be obtained simultaneously by applying the same row operations to $[A | I]$, i.e.,

$$[A | I] = [A | \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n] \xrightarrow{\sim} [A | \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n] = [I | B].$$

Then B is the inverse of A .

Example 3.1. The matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$ is invertible.

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 4 & 0 & 0 & 1 \end{array} \right] \begin{array}{l} R_2 - 2R_1 \\ \rightsquigarrow \\ R_3 - R_1 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -3 & -2 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array} \right] \begin{array}{l} R_3 + R_2 \\ \rightsquigarrow \\ (-1)R_2 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 2 & -1 & 0 \\ 0 & 0 & -1 & -3 & 1 & 1 \end{array} \right] \begin{array}{l} R_2 + 3R_3 \\ \rightsquigarrow \\ R_1 + 2R_3 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & -5 & 2 & 2 \\ 0 & 1 & 0 & -7 & 2 & 3 \\ 0 & 0 & -1 & -3 & 1 & 1 \end{array} \right] \begin{array}{l} R_1 - R_2 \\ \rightsquigarrow \\ (-1)R_3 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 0 & -1 \\ 0 & 1 & 0 & -7 & 2 & 3 \\ 0 & 0 & 1 & 3 & -1 & -1 \end{array} \right],$$

$$\left[\begin{array}{ccc} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 4 \end{array} \right]^{-1} = \left[\begin{array}{ccc} 2 & 0 & -1 \\ -7 & 2 & 3 \\ 3 & -1 & -1 \end{array} \right].$$

Example 3.2. The matrix $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 4 & 1 \end{bmatrix}$ is not invertible.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 4 & 1 & 0 & 0 & 1 \end{array} \right] \begin{array}{l} R_2 - 2R_1 \\ \rightsquigarrow \\ R_3 - 3R_1 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -4 & -2 & 1 & 0 \\ 0 & -2 & -8 & -3 & 0 & 1 \end{array} \right] \begin{array}{l} (-1)R_2 \\ \rightsquigarrow \\ R_3 - 2R_2 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 4 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array} \right].$$

Lecture 11

Proposition 3.6. If A and B are $n \times n$ invertible matrices, then

$$(A^T)^{-1} = (A^{-1})^T,$$

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof. Since taking transposition reverses the order of multiplication, we have

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I,$$

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

By definition of inverse matrix, we conclude $(A^T)^{-1} = (A^{-1})^T$.

Since

$$B^{-1}A^{-1}AB = B^{-1}B = I = AA^{-1} = ABB^{-1}A^{-1}.$$

By definition of invertible matrix, we conclude $(AB)^{-1} = B^{-1}A^{-1}$. □

Proposition 3.7. For any 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, if $\det A := ad - bc \neq 0$, then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Proof. Assume $ad - bc = 0$, that is, $ad = bc$. If $ad = 0$, then $a = 0$ or $d = 0$; $b = 0$ or $c = 0$. The matrix contains either a zero row or zero column; so it is not invertible. If $ad \neq 0$, then a, b, c , and d are all nonzero. Thus $a/b = c/d$. This means that the two columns are linearly dependent. So the matrix is not invertible.

Assume $ad - bc \neq 0$; we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

□

Theorem 3.8 (Characterization of Invertible Matrices). Let A be an $n \times n$ square matrix. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $T(\mathbf{x}) = A\mathbf{x}$. Then the following statements are equivalent.

- (a) A is invertible.
- (b) The reduced row echelon form of A is the identity matrix I_n .
- (c) A has n pivot positions.
- (d) $A\mathbf{x} = \mathbf{0}$ has the only trivial solution.
- (e) The column vectors of A are linearly independent.
- (f) The linear transformation T is one-to-one.
- (g) The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for any $\mathbf{b} \in \mathbb{R}^n$.
- (h) The column vectors of A span \mathbb{R}^n .
- (i) The linear transformation T to \mathbb{R}^n is onto.
- (j) A^T is invertible.
- (k) There is an $n \times n$ matrix C such that $CA = I_n$.
- (l) There is an $n \times n$ matrix D such that $AD = I_n$.

Proof. It is straightforward that

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (f) \Rightarrow (g) \Rightarrow (h) \Rightarrow (i) \Rightarrow (T \text{ is invertible}) \Rightarrow (a).$$

Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be define by $F(\mathbf{x}) = C\mathbf{x}$. Then $CA = I_n \Leftrightarrow F \circ T = \text{Id}_n \Leftrightarrow T$ is 1-1.

Let $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be define by $G(\mathbf{x}) = D\mathbf{x}$. Then $AD = I_n \Leftrightarrow T \circ G = \text{Id}_n \Leftrightarrow T$ is onto. □

4 Determinants

Lecture 12

Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation with the standard matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Recall that the determinant of a 2×2 matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is defined by

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

It is easy to see that for vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^2$, we have

$$\det[a\mathbf{u} + b\mathbf{v}, \mathbf{w}] = a \det[\mathbf{u}, \mathbf{w}] + b \det[\mathbf{v}, \mathbf{w}],$$

$$\det[\mathbf{u}, a\mathbf{v} + b\mathbf{w}] = a \det[\mathbf{u}, \mathbf{v}] + b \det[\mathbf{u}, \mathbf{w}].$$

This means that \det is linear in each column vector variables. Moreover, switching two columns (rows) changes the sign, i.e.,

$$\det[\mathbf{u}, \mathbf{v}] = -\det[\mathbf{v}, \mathbf{u}].$$

Let $A = [a_{ij}]$ be an $n \times n$ square matrix. For a fixed (i, j) , where $1 \leq i \leq m$ and $1 \leq j \leq n$, let

$$A_{ij} = (n-1) \times (n-1) \text{ submatrix of } A \text{ obtained by deleting the } i\text{th row and } j\text{th column of } A.$$

Definition 4.1. Let $A = [a_{ij}]$ be an $n \times n$ square matrix. The **determinant** of A is a number $\det A$, inductively defined by

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{n+1} a_{1n} \det A_{1n} \\ &= a_{11} \det A_{11} - a_{21} \det A_{21} + \cdots + (-1)^{n+1} a_{n1} \det A_{n1}. \end{aligned}$$

The motivation of this definition is as follows: Write $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$, Note that $\mathbf{a}_1 = \mathbf{a}_{11} + \mathbf{a}_{21} + \cdots + \mathbf{a}_{n1}$, where

$$\mathbf{a}_{11} = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{a}_{21} = \begin{bmatrix} 0 \\ a_{21} \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{a}_{n1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ a_{n1} \end{bmatrix}.$$

Since $\det A$ is linear in each column vectors of the matrix A , i.e., $\det[a\mathbf{u} + b\mathbf{v}, \mathbf{a}_2, \dots, \mathbf{a}_n] = a \det[\mathbf{u}, \mathbf{a}_2, \dots, \mathbf{a}_n] + b \det[\mathbf{v}, \mathbf{a}_2, \dots, \mathbf{a}_n]$, we have

$$\begin{aligned} \det A &= \det[\mathbf{a}_{11} + \mathbf{a}_{21} + \cdots + \mathbf{a}_{n1}, \mathbf{a}_2, \dots, \mathbf{a}_n] \\ &= \det[\mathbf{a}_{11}, \mathbf{a}_2, \dots, \mathbf{a}_n] + \det[\mathbf{a}_{21}, \mathbf{a}_2, \dots, \mathbf{a}_n] + \cdots + \det[\mathbf{a}_{n1}, \mathbf{a}_2, \dots, \mathbf{a}_n] \\ &= \det \begin{bmatrix} a_{11} & * \\ 0 & A_{11} \end{bmatrix} - \det \begin{bmatrix} a_{21} & * \\ 0 & A_{21} \end{bmatrix} + \cdots + (-1)^{n-1} \det \begin{bmatrix} a_{n1} & * \\ 0 & A_{n1} \end{bmatrix} \\ &= a_{11} \det A_{11} - a_{21} \det A_{21} + \cdots + (-1)^{n+1} a_{n1} \det A_{n1}. \end{aligned}$$

Theorem 4.2. Let $A = [a_{ij}]$ be an $n \times n$ matrix. The (i, j) -**cofactor** of A ($1 \leq i, j \leq n$) is the number

$$C_{ij} = (-1)^{i+j} \det A_{ij}.$$

Then

$$\begin{aligned} \det A &= a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} \\ &= a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}. \end{aligned}$$

Example 4.1. For any 3×3 matrix,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Theorem 4.3 (Cofactor Expansion Formula). For any $n \times n$ square matrix $A = [a_{ij}]$,

$$\begin{aligned} \det A &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \\ &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}. \end{aligned}$$

In other words,

$$\det A^T = \det A.$$

Proposition 4.4.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22} \cdots a_{nn}.$$

Theorem 4.5. Determinant satisfies the following properties.

- Adding a multiple of one row (column) to another row (column) does not change the determinant.
- Interchanging two rows (columns) changes the sign of the determinant.
- If two rows (columns) are the same, then the determinant is zero.
- If one row (column) of A is multiplied by a scalar γ to produce a matrix B , i.e., $A \xrightarrow{\gamma R_k} B$, then

$$\det B = \gamma \det A.$$

Proof. For 2×2 case, we have

(a)

$$\begin{aligned} \begin{vmatrix} a & b \\ c + \gamma a & d + \gamma b \end{vmatrix} &= a(d + \gamma b) - b(c + \gamma a) \\ &= ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}; \\ \begin{vmatrix} a + \gamma c & b + \gamma d \\ c & d \end{vmatrix} &= (a + \gamma c)d - (b + \gamma d)c \\ &= ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}. \end{aligned}$$

(b)

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - da = -(ad - bc) = - \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Part (c) follows from Part (a). As for Part (d), we have

$$\begin{vmatrix} a & b \\ \gamma c & \gamma d \end{vmatrix} = a\gamma c - b\gamma d = \gamma(ad - bc) = \gamma \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

For $n \geq 2$, the properties can be verified by Mathematical Induction on n . □

Theorem 4.6 (The Algorithm for Determinant). *Let $A = [a_{ij}]$ be an $n \times n$ matrix. If*

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{array}{l} \text{interchanging rows} \\ \rightsquigarrow \\ \text{adding a multiple of} \\ \text{one row to another} \end{array} \begin{bmatrix} a_1 & * & \cdots & * \\ 0 & a_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}.$$

Then

$$\det A = (-1)^k a_1 a_2 \cdots a_n,$$

where k is the number times of interchanging two rows.

Example 4.2. The determinant

$$\begin{vmatrix} 1 & 2 & 4 \\ 2 & 3 & 5 \\ 3 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & -3 \\ 0 & -2 & -4 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & -2 & -4 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{vmatrix} = -2.$$

Theorem 4.7. *Let A be an $n \times n$ matrix, and let C_{ij} be the (i, j) -cofactor of A , i.e.,*

$$C_{ij} = (-1)^{i+j} \det A_{ij}.$$

Then we have

$$\sum_{k=1}^n a_{ik} C_{jk} = a_{i1} C_{j1} + a_{i2} C_{j2} + \cdots + a_{in} C_{jn} = \begin{cases} \det A & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Proof. For $i = j$, it follows from the cofactor expansion formula. For $i \neq j$, let us verify for 3×3 matrices. For instance, for $i = 1$ and $j = 3$, we have

$$\begin{aligned} a_{11}C_{31} + a_{12}C_{32} + a_{13}C_{33} &= a_{11} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{12} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{13} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = 0. \end{aligned}$$

□

Definition 4.8. For an $n \times n$ matrix $A = [a_{ij}]$, the **classical adjoint matrix** of A is the $n \times n$ matrix $\text{adj } A$ whose (i, j) -entry is the (j, i) -cofactor C_{ji} , that is,

$$\text{adj } A := \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}^T.$$

Theorem 4.9. For any $n \times n$ matrix A ,

$$A(\text{adj } A) = (\det A) I_n.$$

Proof. We verify for the case of 3×3 matrices.

$$\begin{aligned} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^3 a_{1k} C_{1k} & \sum_{k=1}^3 a_{1k} C_{2k} & \sum_{k=1}^3 a_{1k} C_{3k} \\ \sum_{k=1}^3 a_{2k} C_{1k} & \sum_{k=1}^3 a_{2k} C_{2k} & \sum_{k=1}^3 a_{2k} C_{3k} \\ \sum_{k=1}^3 a_{3k} C_{1k} & \sum_{k=1}^3 a_{3k} C_{2k} & \sum_{k=1}^3 a_{3k} C_{3k} \end{bmatrix} \\ &= \begin{bmatrix} \det A & 0 & 0 \\ 0 & \det A & 0 \\ 0 & 0 & \det A \end{bmatrix} \end{aligned}$$

□

Theorem 4.10. An $n \times n$ matrix A is invertible if and only if $\det A \neq 0$.

5 Elementary matrices

Definition 5.1. The following $n \times n$ matrices are called **elementary matrices**.

$$\begin{aligned} E(c(i)) &:= \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & c & \\ & & & & 1 \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \text{ith, } c \neq 0; \\ E(i, j) &:= E(j, i) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & \cdots & 1 \\ & & \vdots & \ddots & \vdots \\ & & & 1 & \cdots & 0 \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \text{ith} \\ \\ \\ \text{jth} \end{matrix}; \\ E(j + c(i)) &:= \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \cdots & 0 \\ & & \vdots & \ddots & \vdots \\ & & c & \cdots & 1 \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \text{ith} \\ \\ \\ \text{jth} \end{matrix}. \end{aligned}$$

Let I_n be the $n \times n$ identity matrix. We have

$$\begin{aligned} I_n &\xrightarrow{cR_i} E(c(i)), \\ I_n &\xrightarrow{R_i \leftrightarrow R_j} E(i, j) = E(j, i), \\ I_n &\xrightarrow{R_j + cR_i} E(j + c(i)). \end{aligned}$$

Proposition 5.2. *Let A be an $m \times n$ matrix. Then*

- (a) $E(c(i))A$ is the matrix obtained from A by multiplying c to the i th row.
- (b) $E(i, j)A$ is the matrix obtained from A by interchanging the i th row and the j th row.
- (c) $E(j + c(i))A$ is the matrix obtained from A by adding the c multiple of the i th row to the j th row.

$$\begin{aligned} S(j + c(i))A &= \begin{matrix} \textit{ith} \\ \textit{jth} \end{matrix} \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ a_{j1} & \cdots & a_{ji} & \cdots & a_{jj} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ a_{j1} + ca_{i1} & \cdots & a_{ji} + ca_{ii} & \cdots & a_{jj} + ca_{ij} & \cdots & a_{jn} + ca_{in} \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} \end{aligned}$$

Proposition 5.3. *Elementary matrices are invertible. Moreover,*

$$\begin{aligned} E(c(i))^{-1} &= E\left(\frac{1}{c}(i)\right), \\ E(i, j)^{-1} &= E(j, i) = E(i, j), \\ E(j + c(i))^{-1} &= E(j - c(i)). \end{aligned}$$

Theorem 5.4. *A square matrix A is invertible if and only if it can be written as product*

$$A = E_1 E_2 \cdots E_p,$$

of some elementary matrices E_1, E_2, \dots, E_p .

Theorem 5.5. *A square matrix A is invertible if and only if $\det A \neq 0$.*

Proposition 5.6. *Let A and B $n \times n$ matrices. Then*

$$\det AB = (\det A) \cdot (\det B).$$

Proof. Case 1: $\det A = 0$. Then A is not invertible by Theorem 4.10. We claim that AB is not invertible. Otherwise, if AB is invertible, then there is a matrix C such that $(AB)C = I$; thus $A(BC) = I$; this means that A is invertible, contradictory to that $\det A = 0$. Again by Theorem 4.10, $\det(AB) = 0$. Therefore $\det AB = \det A \det B$.

Case 2: $\det A \neq 0$. It is easy to check directly that

$$\det EB = \det E \det B$$

for any elementary matrix E . Let A be written as $A = E_1 E_2 \cdots E_k$ for some elementary matrices E_1, E_2, \dots, E_k . Then

$$\begin{aligned} \det AB &= \det(E_1 E_2 \cdots E_k B) = (\det E_1) \det(E_2 \cdots E_k B) \\ &= (\det E_1) (\det E_2) \det(E_3 \cdots E_k B) = \cdots \\ &= \det E_1 \det E_2 \cdots \det E_k \det B = \det A \det B. \end{aligned}$$

□

6 Interpretation of determinant

Lecture 13

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ be a 2×2 matrix. Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the standard basis of \mathbb{R}^3 . Set

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ 0 \end{bmatrix}.$$

Then the area of the parallelepiped spanned by the two vectors \mathbf{a}_1 and \mathbf{a}_2 is the length of the cross product $\mathbf{a}_1 \times \mathbf{a}_2$, that is,

$$\begin{aligned} |\mathbf{a}_1 \times \mathbf{a}_2| &= |(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2) \times (a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2)| \\ &= |a_{11}a_{22}\mathbf{e}_3 - a_{21}a_{12}\mathbf{e}_3| \\ &= |a_{11}a_{22} - a_{21}a_{12}|. \end{aligned}$$

Let $A = [a_{ij}] = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$ be a 3×3 matrix. The volume of the parallelotope spanned by the three vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ is the length of the triple $(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3$, that is,

$$\begin{aligned} |(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3| &= |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)| \\ &= |a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31}|. \end{aligned}$$

Example 6.1. The Vandermonde determinant is

$$\begin{aligned} \begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix} &= \begin{vmatrix} 1 & 1 & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 \\ 0 & a_2^2 - a_2 a_1 & a_3^2 - a_3 a_1 \end{vmatrix} \\ &= \begin{vmatrix} a_2 - a_1 & a_3 - a_1 \\ a_2(a_2 - a_1) & a_3(a_3 - a_1) \end{vmatrix} \\ &= (a_3 - a_2)(a_3 - a_1)(a_2 - a_1). \end{aligned}$$

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{j>i} (a_j - a_i).$$

Example 6.2.

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 7 \end{vmatrix} \begin{array}{l} R_2 - 2R_1 \\ = \\ R_4 + 3R_1 \\ \\ R_3 + 4R_2 \\ = \\ R_4 - R_2 \end{array} \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 1 \\ \\ 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & -7 & 9 \\ 0 & 0 & 0 & -1 \end{vmatrix} = 7.$$

Theorem 6.1 (Cramer's Rule). *Let $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ be an $n \times n$ invertible matrix with the column vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Then for any $\mathbf{b} \in \mathbb{R}^n$, the unique solution of the system*

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

is given by

$$x_i = \frac{\det A_i}{\det A}, \quad i = 1, 2, \dots, n$$

where

$$A_i = [\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n], \quad 1 \leq i \leq n.$$

Proof. If a vector \mathbf{x} is the solution of the system, then

$$\begin{aligned} A[\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{x}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n] &= [A\mathbf{e}_1, \dots, A\mathbf{e}_{i-1}, A\mathbf{x}, A\mathbf{e}_{i+1}, \dots, A\mathbf{e}_n] \\ &= [\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n]. \end{aligned}$$

Taking determinant of both sides, we have

$$(\det A) x_i = \det A_i, \quad 1 \leq i \leq n.$$

□

Example 6.3. For the case $n = 3$,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

let the second column of A be replaced by \mathbf{b} . Then

$$\begin{aligned} A[\mathbf{e}_1, \mathbf{x}, \mathbf{e}_3] &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & x_1 & 0 \\ 0 & x_2 & 0 \\ 0 & x_3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{11}x_1 + a_{12}x_2 + a_{13}x_3 & a_{13} \\ a_{21} & a_{21}x_1 + a_{22}x_2 + a_{23}x_3 & a_{23} \\ a_{31} & a_{31}x_1 + a_{32}x_2 + a_{33}x_3 & a_{33} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{bmatrix} = [\mathbf{a}_1, \mathbf{b}, \mathbf{a}_3]. \end{aligned}$$

Note that

$$\begin{vmatrix} 1 & x_1 & 0 \\ 0 & x_2 & 0 \\ 0 & x_3 & 1 \end{vmatrix} = \begin{vmatrix} x_2 & 0 \\ x_3 & 1 \end{vmatrix} = x_2.$$

On the one hand,

$$\det(A[\mathbf{a}_1, \mathbf{x}, \mathbf{a}_3]) = (\det A)(\det[\mathbf{a}_1, \mathbf{x}, \mathbf{a}_3]) = (\det A) x_2.$$

On the other hand, $\det(A[\mathbf{a}_1, \mathbf{x}, \mathbf{a}_3]) = \det[\mathbf{a}_1, \mathbf{b}, \mathbf{a}_3]$. It follows that

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

7 Ratio of areas by transformation

Let P be a parallelogram spanned by vectors \mathbf{u}, \mathbf{v} . Set $A = [\mathbf{u}, \mathbf{v}]$. Then

$$\text{Area of } P = |\det A|.$$

Proof.

$$\begin{aligned} \text{Area} &= (a+c)(b+d) - 2bc - ab - cd \\ &= ad - bc = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = |\det A|. \end{aligned}$$

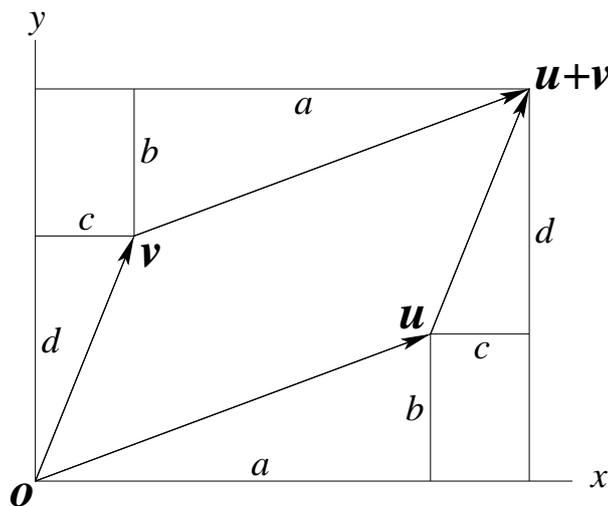


Figure 4: Area of a parallelogram.

□

Theorem 7.1. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation with standard matrix A , i.e., $T(\mathbf{x}) = A\mathbf{x}$. Let P be a parallelogram in \mathbb{R}^2 . Then

$$\{\text{area of } T(P)\} = |\det A| \cdot \{\text{area of } P\}$$

Proof. Let the parallelogram P be spanned by two vectors \mathbf{b}_1 and \mathbf{b}_2 ; i.e., $P = \{s\mathbf{b}_1 + t\mathbf{b}_2 : s, t \in [0, 1]\}$. Set $B = [\mathbf{b}_1, \mathbf{b}_2]$. Then

$$T(P) = \{T(s\mathbf{b}_1 + t\mathbf{b}_2) : s, t \in [0, 1]\} = \{sT(\mathbf{b}_1) + tT(\mathbf{b}_2) : s, t \in [0, 1]\}.$$

Thus

$$\begin{aligned} \{\text{area of } T(P)\} &= |\det[T(\mathbf{b}_1), T(\mathbf{b}_2)]| = |\det[A\mathbf{b}_1, A\mathbf{b}_2]| \\ &= |\det(A[\mathbf{b}_1, \mathbf{b}_2])| = |(\det A) \cdot (\det B)| \\ &= |\det A| \cdot \{\text{area of } P\}. \end{aligned}$$

□

Theorem 7.2. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear transformation with standard matrix A , i.e., $T(\mathbf{x}) = A\mathbf{x}$. Let P be a parallelepiped in \mathbb{R}^3 . Then

$$\{\text{volume of } T(P)\} = |\det A| \cdot \{\text{volume of } P\}$$

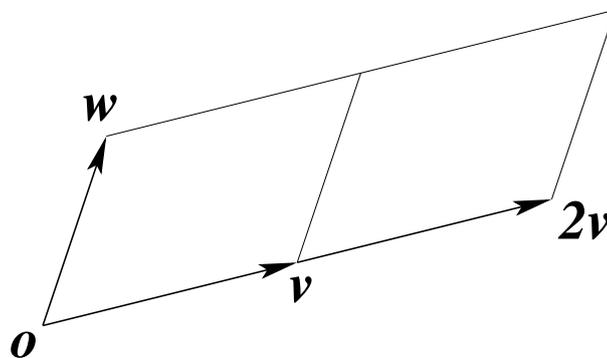


Figure 5: Area of a parallelogram.

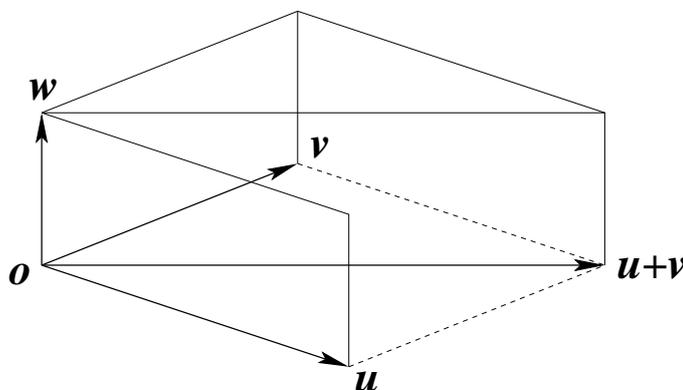


Figure 6: Area of a parallelogram.