Chapter 24: Linear Transformations

Geometric interpretation of "multiplication by A" = Transformation of $\mathbb{R}^n$.

A: $n \times n$ matrix

⊛ The key properties:

(1) $A\vec{0} = \vec{0}$,

(2) $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$, for any $\vec{u}, \vec{v} \in \mathbb{R}^n$,

(3) $A(r\vec{u}) = r(A\vec{u})$, for any $\vec{u} \in \mathbb{R}^n$, $r \in \mathbb{R}$.

These properties imply that the four points of a parallelogram with vertices $\vec{0}, \vec{u}, \vec{v}, \vec{u}+\vec{v}$ are sent to the four points $\vec{0}, A\vec{u}, A\vec{v}, A\vec{u}+A\vec{v}$, which again define the corners of a parallelogram.

Definition: Let U and V be vector spaces. A "linear transformation" T is a map from U to V satisfying

(1) For all $\vec{u}, \vec{v} \in U$, $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}) \in V$

(2) For all $\vec{u} \in U$ and $r \in \mathbb{R}$, $T(r\vec{u}) = rT(\vec{u}) \in V$

⊛ T is a black box (or formula or rule) which accepts a vector of U as input and produces a uniquely determined vector in V as output. It also takes sums to sums and scalar mutiples to scalar multiples

Examples :

(1) multiplication by a square matrix A is a linear transformation.

(2) Let A be an $m \times n$ matrix. Define the map $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by

$$T_A(\vec{u}) = A\vec{u},$$

then T_A is also a linear transformation.

proof:

1- $\vec{u}, \vec{v} \in \mathbb{R}^n \Rightarrow T_A(\vec{u}+\vec{v}) = A(\vec{u}+\vec{v})$

$$= A\vec{u} + A\vec{v} = T_A(\vec{u}) + T_A(\vec{v}),$$

as required.

2- $\vec{u} \in \mathbb{R}^n, r \in \mathbb{R}, T_A(r\vec{u}) = A(r\vec{u}) = r(A\vec{u})$

$$= rT_A(\vec{u}),$$

as required.

Therefore, T_A is a linear transformation for any m and n.

(3) Show that the map $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by $T(x,y) = (x+1, xy)$ is not a linear transformation:

1- $T(\vec{u}+\vec{v}) = T(u_1+v_1, u_2+v_2) = (u_1+v_1+1, (u_1+v_1)(u_2+v_2))$

$T(\vec{u}) + T(\vec{v}) = (u_1+1, u_1u_2) + (v_1+1, v_1v_2) = (u_1+u_2+2, u_1u_2+v_1v_2)$

$\Rightarrow T(\vec{u}+\vec{v}) \neq T(\vec{u}) + T(\vec{v})$

2- $T(r\vec{u}) = T(ru_1, u_2) = (ru_1+1, (ru_1)(ru_2))$
$$= (ru_1+1, r^2 u_1 u_2)$$

$rT(\vec{u}) = r(u_1+1, u_1u_2) = (ru_1+1, ru_1u_2)$

$\Rightarrow T(r\vec{u}) \neq rT(\vec{u})$ (in general.)

So, T is not a linear transformation.

Theorem (24.1): Determination of Linear Transformations on a Basis.

(1) Suppose $T: U \to V$ is a linear transformation and $\{\vec{u}_1, \dots, \vec{u}_n\}$ is a basis for U. Then, T is completely determined by the vectors $T(\vec{u}_1), \dots, T(\vec{u}_n)$.

(2) Suppose $\{\vec{u}_1, \dots, \vec{u}_n\}$ is a basis for U and $\{\vec{v}_1, \dots, \vec{v}_n\}$ are any n vectors in V (even possibly dependant or $\vec{0}$). Then, there is a unique linear transformation T, which satisfies $T(\vec{u}_i) = \vec{v}_i$, for all i.

Proof: please refer to the textbook.

⊛ This theorem implies that every linear transformation is just matrix multiplication in disguise!

Theorem (24.2): The standard matrix of a linear transformation

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then, there is an $m \times n$ matrix A such that

$$T(\vec{x}) = A\vec{x},$$

for all $\vec{x} \in \mathbb{R}^n$. More precisely, if $\{\vec{e}_1, \dots, \vec{e}_n\}$ is the standard basis for $\mathbb{R}^n$, then the matrix A is given by

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

The matrix A is called the standard matrix of T.

proof: Let $\vec{u} = (u_1, u_2, \dots, u_n) = u_1\vec{e}_1 + u_2\vec{e}_2 + \dots + u_n\vec{e}_n$.

By theorem (24.1), then

$$T(\vec{u}) = u_1 T(\vec{e}_1) + u_2 T(\vec{e}_2) + \dots + u_n T(\vec{e}_n)$$

$$= [T(\vec{e}_1), T(\vec{e}_2), \dots, T(\vec{e}_n)] \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = A\vec{u}.$$

⊛ Kernels and Images:

Definition: Let $T: U \to V$ be a linear transformation. Then,

(1) The Kernel of T, denoted by $\ker(T)$, is the set of all elements of U, which are sent to $\vec{0}$ by T, that is

$$\boxed{\ker(T) = \{\vec{u} \in U \mid T(\vec{u}) = \vec{0}\}.}$$

(2) The image of T, denoted by $\operatorname{im}(T)$, is the set of all elements of V which are equal to $T(\vec{u})$, for some $\vec{u} \in U$. that is

$$\boxed{\operatorname{im}(T) = \{\vec{v} \in V \mid \vec{v} = T(\vec{u}), \text{ for some } \vec{u} \in U\}}$$

Theorem (24.3): (Kernels and Images of the standard matrix).

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

$$\boxed{\ker(T) = \operatorname{Null}(A)} \quad \text{and} \quad \boxed{\operatorname{im}(T) = \operatorname{Col}(A)}$$

⊛ Rank-nullity theorem for a linear transformation $T: U \to V$:

$$\boxed{\dim(\ker(T)) + \dim(\operatorname{im}(T)) = n,}$$

where $n = \dim U$. If T sends U onto a subspace of V of dimension equal to U, then the kernel must be zero. If the image is smaller than U, then those missing dimensions had to go somewhere; in fact, they went to zero.

A remark about the projection matrix:

Recall that one method to find the projection onto a subspace is to:

(1) Create a matrix B, such that $\mathrm{Col}(B) = W$,

(2) Solve $(B^T B)\vec{x} = B^T \vec{b}$,

(3) $\mathrm{proj}_W(\vec{b}) = B\vec{x}$.

Suppose B has linearly independent columns so that $B^T B$ is invertible then,

$$\mathrm{proj}_W(\vec{b}) = B(B^T B)^{-1} B^T \vec{b},$$

So, the projection is given by multiplication by the matrix

$$B(B^T B)^{-1} B^T,$$

which is the standard matrix of T.