線型写像と行列の積

$\mathbb{K}^m$ を m 次元列ベクトル全体の集合とします。

$\vec{a}_1, \cdots, \vec{a}_n \in \mathbb{K}^m$ に対して $\vec{a}_1, \cdots, \vec{a}_n$ を束ねた行列 A を（m 行 n 列の）

$$A = (\vec{a}_1 \cdots \vec{a}_j \cdots \vec{a}_n) = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

と定めます。$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ を右から A にかけることができます。

$$A\vec{x} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \begin{pmatrix} \vdots \\ x_1 a_{i1} + \cdots + x_j a_{ij} + \cdots + x_n a_{in} \\ \vdots \end{pmatrix}$$

となります。（$A\vec{x} \in \mathbb{K}^m$ はすぐに分かるでしょう）上の A は行ベクトル $\mathbf{a}_i = (a_{i1} \cdots a_{ij} \cdots a_{in})$ を用いて

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_m \end{pmatrix}, \quad A\vec{x} = \begin{pmatrix} \mathbf{a}_1\vec{x} \\ \vdots \\ \mathbf{a}_i\vec{x} \\ \vdots \\ \mathbf{a}_m\vec{x} \end{pmatrix} \in \mathbb{K}^m$$

と表現できます。これを用いると A は写像

$$f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$$
$$\vec{x} \longmapsto A\vec{x}$$

を定めます。これを A が定める線型写像と呼びます。何故線型と呼ぶかというと $\vec{x}, \vec{y} \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$ に対して

$$\begin{cases} f_A(\vec{x} + \vec{y}) = f_A(\vec{x}) + f_A(\vec{y}) \\ f_A(\lambda\vec{x}) = \lambda f_A(\vec{x}) \end{cases}$$

が成立するからです。

実際、

$$f_A(\vec{x}+\vec{y}) = \underline{A(\vec{x}+\vec{y})} = (\vec{a}_1 \cdots \vec{a}_n)\begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix}$$

$$= (x_1+y_1)\vec{a}_1 + \cdots + (x_n+y_n)\vec{a}_n$$

$$= x_1\vec{a}_1 + y_1\vec{a}_1 + \cdots + x_n\vec{a}_n + y_n\vec{a}_n$$

$$= x_1\vec{a}_1 + \cdots + x_n\vec{a}_n + y_1\vec{a}_1 + \cdots + y_n\vec{a}_n$$

$$= \underline{A\vec{x} + A\vec{y}} = f_A(\vec{x}) + f_A(\vec{y})$$

$$f_A(\lambda\vec{x}) = (\vec{a}_1 \cdots \vec{a}_n)\begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} = (\lambda x_1)\vec{a}_1 + \cdots + (\lambda x_n)\vec{a}_n$$

$$= \lambda x_1\vec{a}_1 + \cdots + \lambda x_n\vec{a}_n$$

$$= \lambda(x_1\vec{a}_1 + \cdots + x_n\vec{a}_n) = \lambda(A\vec{x}) = \lambda f_A(\vec{x})$$

と示せます. 帰納的に $\vec{x}_1, \dots, \vec{x}_l \in K^n$ に対し

$$f_A(\vec{x}_1 + \cdots + \vec{x}_l) = f_A(\vec{x}_1) + \cdots + f_A(\vec{x}_l)$$

$$A(\vec{x}_1 + \cdots + \vec{x}_l) = A\vec{x}_1 + \cdots + A\vec{x}_l$$

が示せ

$$f_A(\lambda_1 x_1 + \cdots + \lambda_l \vec{x}_l)$$

$$= f_A(\lambda_1\vec{x}_1) + \cdots + f_A(\lambda_l\vec{x}_l) = \lambda_1 f_A(\vec{x}_1) + \cdots + \lambda_l f_A(\vec{x}_l)$$

すなわち

$$A(\lambda_1\vec{x}_1 + \cdots + \lambda_l\vec{x}_l) = \lambda_1(A\vec{x}_1) + \cdots + \lambda_l(A\vec{x}_l)$$

が成立します. λ_j は 変数っぽくないので c_j とすると

⊛ $$A(z_1\vec{x}_1+\cdots+z_\ell\vec{x}_\ell) = z_1(A\vec{x}_1)+\cdots+z_\ell(A\vec{x}_\ell)$$

とします。$\vec{x}_j=\vec{e}_j\in\mathbb{K}^n$ とします。$B=(\vec{e}_1\cdots\vec{e}_\ell)$ を n 行 ℓ 列の行列と定めると $\vec{z}=\begin{pmatrix}z_1\\ \vdots\\ z_\ell\end{pmatrix}\in\mathbb{K}^\ell$ に対して

$$B\vec{z} = z_1\vec{e}_1+\cdots+z_\ell\vec{e}_\ell\in\mathbb{K}^n$$

となります。⊛ から

$$A(z_1\vec{e}_1+\cdots+z_\ell\vec{e}_\ell)=z_1(A\vec{e}_1)+\cdots+z_\ell(A\vec{e}_\ell)$$

従って

$$A(B\vec{z}) = (A\vec{e}_1\cdots A\vec{e}_\ell)\vec{z}.$$

が成立します。ここで m 行 n 列の A と n 行 ℓ 列の B の積を

$$AB=(A\vec{e}_1\cdots A\vec{e}_\ell)$$

と（$A\vec{e}_j\in\mathbb{K}^m$ なので）m 行 ℓ 列の行列として定義します。すると

$$\boxed{f_A(f_B(\vec{z})) = f_{AB}(\vec{z})}$$

が成立することが分かります

$$\begin{tikzcd}
\mathbb{K}^\ell \arrow[r, "f_B"] \arrow[rr, bend right] & \mathbb{K}^n \arrow[r, "f_A"] & \mathbb{K}^m
\end{tikzcd}$$

$$f_A\circ f_B=f_{AB}.$$

注 $f_1:X\to Y,\ f_2:X\to Y$ に対して

$$f_1=f_2 \iff \forall x\in X\ \ f_1(x)=f_2(x)$$

次に進む準備をします.

$$\vec{e}_j \in \mathbb{K}^n \text{ を } \quad \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots$$

と定義します. $A = (\vec{a}_1 \cdots \vec{a}_n)$ を m 行 n 列とすると

$$A\vec{e}_j = \vec{a}_j \quad (第 j 列)$$

であることが分かります. F, G を m 行 n 列とするとき

$$F\vec{x} = G\vec{x} \quad (\forall \vec{x} \in \mathbb{K}^n) \Longleftrightarrow F = G$$

すなわち $\vec{f}_j = \vec{g}_j$

が分かります. ($\Rightarrow$) $\vec{x} = \vec{e}_j$ とすると $\vec{f}_j = \vec{g}_j$ が $(j = 1, \cdots, n)$

従います. これはすぐに応用します.

A, B は上の通り m 行 n 列の行列, B が n 行 l 列の行列とします.

さらに C を l 行 k 列の行列とします. このとき

$$\mathbb{K}^k \xrightarrow{f_C} \mathbb{K}^l \xrightarrow{f_B} \mathbb{K}^n \xrightarrow{f_A} \mathbb{K}^m$$

に対して一般論から (写像の結合則)

$$f_A \circ (f_B \circ f_C) = (f_A \circ f_B) \circ f_C$$

$$f_A \circ f_{BC} = f_{AB} \circ f_C$$

さらに

$$f_{A(BC)} = f_{(AB)C}$$

$\forall \vec{w} \in \mathbb{K}^k$ に対して $\quad (A(BC))\vec{w} = ((AB)C)\vec{w}$

から $\boxed{A(BC) = (AB)C}$ が従います. 結合則.

m行n列の行列全体の集合を $M_{m,n}(\mathbb{K})$ にします。

また n行n列の行列（n次正方行列）全体の集合を $M_n(\mathbb{K})$ とします。

$$\boxed{\begin{array}{l}\underline{\text{定理}}\quad A \in M_{m,n}(\mathbb{K}),\ B \in M_{n,l}(\mathbb{K}),\ C \in M_{l,k}(\mathbb{K}) \\ \text{に対して} \\ \qquad (AB)C = A(BC)\end{array}}$$

これを直接示しましょう

$$\begin{aligned}(AB)C &= ((AB)\vec{c}_1 \cdots (AB)\vec{c}_k) \\ &= (A(B\vec{c}_1) \cdots A(B\vec{c}_k)) \\ &= A(B\vec{c}_1 \cdots B\vec{c}_k) = A(BC)\end{aligned}$$

ここで

$$f_A \circ f_B = f_{AB} \text{ から } \forall \vec{z} \in \mathbb{K}^l \quad A(B\vec{z}) = (AB)\vec{z}$$

が成立することを用いました。

$\mathrm{Im}(f_A) = ?$

$A = (\vec{a}_1 \cdots \vec{a}_n) \in M_{m,n}(\mathbb{K})$ とします。このとき

$$f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$$

となります。$\vec{a} \in \mathbb{K}^m$ に対して

$$\vec{a} \in \mathrm{Im}(f_A) \iff \exists \vec{x} \in \mathbb{K}^n \quad f_A(\vec{x}) = \vec{a}$$

ここで $f_A(\vec{x}) = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n$ から

$$\vec{a} \in \mathrm{Im}(f_A) \Longleftrightarrow \vec{a} \in L(\vec{a}_1, \cdots, \vec{a}_n)$$

$$:= \{x_1\vec{a}_1 + \cdots + x_n\vec{a}_n \in \mathbb{K}^m ; x_j \in \mathbb{K}\}$$

がわかります.