行列の積に関してよくある勘違い

・A の列の数と B の行の数が合っていないのに積 AB を考えようとする.

とくに, $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ のような書き方をする人がよくいるのですが (下記も参照), これは行列の積のつもりで誤って書いているのか, あるいは内積など別の意味でこのように書いているのか, よく分からないことがあります.

・行列 AB のサイズが合わない.

A が $m \times n$ 行列, B が $n \times l$ 行列ならば AB は $m \times l$ 行列にならなければおかしいのですが, 何かを間違えて $m \times l$ 以外のサイズになってしまう人がいます. 計算して $m \times l$ 行列にならなかったとすれば, それはどこかが間違っていることを意味するので, 計算を見直しましょう.

・成分ごとの掛け算にしてしまう.

特に, 正方行列のベキ積 $A^n = \underbrace{A \cdots A}_{n}$ をこのように計算してしまう人が多いような気がします:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 \overset{?}{=} \begin{pmatrix} a^2 & b^2 \\ c^2 & d^2 \end{pmatrix} \qquad (\leftarrow \text{間違い})$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & db+d^2 \end{pmatrix} \qquad (\leftarrow \text{正しくはこう})$$

他に多いのが, 2 つの n 項ベクトルの「積」を

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 \\ \vdots \\ b_nb_n \end{pmatrix} \qquad (\leftarrow\ !?\ )$$

のように計算する人です. これは行列の積としては意味を持ちません[1]. ただ, これは内積のつもりでこのように書いてしまうことの方が多いかもしれません. ここで,「内積としてはこれで合っているのでは？」と思った人は内積の定義を復習し直してください. **内積はスカラーです**:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\cdot\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + \cdots + a_nb_n.$$

[1] もう少し進んだ代数学ではこのような積が意味をもつこともあります→環の直積 (直和)

二つのベクトル $\boldsymbol{a}, \boldsymbol{b}$ の内積を表したい場合は $\boldsymbol{a} \cdot \boldsymbol{b}$ のように中点を使うか, カッコをつけて $(\boldsymbol{a}, \boldsymbol{b})$ と書くのが誤解を避ける道です (私はどちらかというと中点の方が好みです). 単に $\boldsymbol{ab}$ と書いたのでは, どの積を考えているのか分からず[2], 読み手に無用の混乱を与えてしまいます.

数学 (に限らずどの分野でもそうですが) の文章を書くときには, 読み手に誤解を与えないように, ということに結構気を使う必要があります. 内積を表すのに $(\boldsymbol{a}, \boldsymbol{b})$ というちゃんとした記号を使っていてさえ, 記号だけみると例えば直積集合 $K^n \times K^n$ の元と区別がつきません. 読み手が誤解する恐れがあると思ったら注意書きをつけるとか, 文脈でそれと分かるように書かないといけなかったりする場合もあるわけです.

・普通の数の計算と同様に次が成立すると思いこんでしまう.

(1) $AB = BA$
(2) $AB = O \Rightarrow A = O$ または $B = O$

どちらも一般には正しくありません. 反例を作ってみれば分かるので, これは演習問題にしましょう.

補足: 行列の積をあのように定義する理由. 一言でいえば, あの定義は 1 次変換の合成からきています.

例えば $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ を表現行列とする 2 つの 1 次変換

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}$$
$$g : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_{11}x + b_{12}y \\ b_{21}x + b_{22}y \end{pmatrix}$$

を考えてみます. ベクトル $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ に g を施した後で f を施すと,

$$\boldsymbol{x} \overset{g}{\mapsto} B\boldsymbol{x} \overset{f}{\mapsto} A(B\boldsymbol{x})$$

となりますが, これをまとめてひとつの 1 次変換とみたものを f と g の合成といいます[3]. 行先を実際に計算してみましょう:

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &\overset{g}{\mapsto} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_{11}x + b_{12}y \\ b_{21}x + b_{22}y \end{pmatrix} \\ &\overset{f}{\mapsto} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11}x + b_{12}y \\ b_{21}x + b_{22}y \end{pmatrix} = \begin{pmatrix} a_{11}(b_{11}x + b_{12}y) + a_{12}(b_{21}x + b_{22}y) \\ a_{21}(b_{11}x + b_{12}y) + a_{22}(b_{21}x + b_{22}y) \end{pmatrix} \end{aligned}$$

[2]ベクトルの積には他に外積もありますし

[3]記号 $f \circ g$ で表します. 教科書の 3.6 節 (第 3 章付録) を参照.

$$
\begin{aligned}
&= \begin{pmatrix} (a_{11}b_{11} + a_{12}b_{21})x + (a_{11}b_{12} + a_{12}b_{22})y \\ (a_{21}b_{11} + a_{22}b_{21})x + (a_{21}b_{12} + a_{22}b_{22})y \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.
\end{aligned}
$$

これの表現行列が積 AB になるように, 行列の積を定義したかったわけです:

$$
AB = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.
$$

端的にいえば, $A(B\boldsymbol{x}) = (AB)\boldsymbol{x}$ が成り立つように AB を定義したかった, と.

ふつう線形代数を勉強するときは, 行列の積を先に定義して, そして積の結合法則を証明して— 1 次変換を勉強するときには $A(B\boldsymbol{x}) = (AB)\boldsymbol{x}$ は当たり前になっている—そういう順番で教科書を書いた方がすっきりするのでそうなっているわけですが, やはり 1 次変換を念頭におかないと真の理解を得るのは難しいのかも知れません. 天下り的な定義を見たときに戸惑いや拒絶感を感じるのは正常なことですが, それが数学を勉強する上での躓きになることも多いものです. 行列の積を成分ごとの積で計算してしまう人などは「なんでこんな複雑な定義をするんだ！」という憤りのようなものを感じているのかもしれません. そういう人にこそ, ここに書いたような事情を分かってほしくて, まあ, そのあたりを補完するためにこういう文章を書いているわけです.

さて, 正方行列でない行列の積についても同様です. 例えば, 2×3 行列 $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$ と 3×2 行列 $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$ を表現行列とする 2 つの変換

$$
\begin{aligned}
f : \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
g : \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}
$$

を考えて, 「g を施した後 f を施す」という合成を考えてみてください. 実際に計算してみれば, ちゃんと表現行列が積 AB の定義と一致するはずです

なお, これは「なぜ A の列の数と B の行の数が同じでなければならないか」ということを理解するためのきっかけにもなるかもしれません. 上の例の場合, f は 3 項ベクトルに適用する変換なので, g の行先が 3 項ベクトルでなければなりません. (f と g がそういう関係でなければ合成など考えることはできませんよね？)