MA 110:
Lecture 08

Saurav Bhaumik,
Department of Mathematics
IIT Bombay

Spring 2025

Corollary

Let $\mathbf{A}$ be a square matrix.

(i) If two rows of $\mathbf{A}$ are interchanged, then $\det \mathbf{A}$ gets multiplied by -1.

(ii) Addition of a multiple of a row to another row of $\mathbf{A}$ does not alter $\det \mathbf{A}$.

(iii) Multiplication of a row of $\mathbf{A}$ by a scalar α results in the multiplication of $\det \mathbf{A}$ by α.

Proof. Since the columns of $\mathbf{A}^{\mathsf{T}}$ are the rows of $\mathbf{A}$, and since $\det \mathbf{A} = \det \mathbf{A}^{\mathsf{T}}$, these results follow from the previous proposition. □

The above corollary can be used to find $\det \mathbf{A}$ as follows. Transform $\mathbf{A}$ to $\mathbf{A}'$ by EROs of type I and type II, where $\mathbf{A}'$ is in REF, keeping track of the number of row interchanges.

Now $\mathbf{A}'$ is an upper triangular matrix. Let p be the number of row interchanges, and let $a'_{11}, \ldots, a'_{nn}$ be the diagonal entries of $\mathbf{A}'$. Then $\det \mathbf{A} = (-1)^p \det \mathbf{A}' = (-1)^p a'_{11} \cdots a'_{nn}$.

Corollary

A square matrix $\mathbf{A}$ is invertible if and only if $\det \mathbf{A} \neq 0$.

Proof. Note that $\mathbf{A}$ is invertible if and only if $\mathbf{A}'$ is. Now as $\mathbf{A}'$ is a square matrix in REF, it is invertible if and only if there are no zero rows. If there are no zero rows then each column is pivotal with pivots given by precisely the diagonal entries. Thus the formula above shows that $\det \mathbf{A} = \det \mathbf{A}' \neq 0$. Conversely, if the k-th row is zero in $\mathbf{A}'$, then the diagonal term $a_{kk} = 0$. Then the same formula above shows that $\det \mathbf{A} = \det \mathbf{A}' = 0$. $\square$

Example

Let $\mathbf{A} := \begin{bmatrix} 0 & 2 & 0 & -1 \\ 1 & 2 & 1 & -1 \\ 0 & 0 & 3 & 2 \\ 1 & -2 & 1 & -2 \end{bmatrix} \xrightarrow{R_1 \longleftrightarrow R_2} \begin{bmatrix} 1 & 2 & 1 & -1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 3 & 2 \\ 1 & -2 & 1 & -2 \end{bmatrix}$

$$\xrightarrow{R_4 - R_1} \begin{bmatrix} 1 & 2 & 1 & -1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 3 & 2 \\ 0 & -4 & 0 & -1 \end{bmatrix} \xrightarrow{R_4 + 2R_2} \begin{bmatrix} 1 & 2 & 1 & -1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -3 \end{bmatrix} = \mathbf{A}'.$$

Since there is only one row interchange while transforming $\mathbf{A}$ to $\mathbf{A}'$, since $\mathbf{A}'$ is in REF, and since $\det \mathbf{A}' = 1 \cdot 2 \cdot 3 \cdot (-3) = -18$, we see that $\det \mathbf{A} = (-1)(-18) = 18$.

Remark We have given several criteria for the invertibility of an $n \times n$ matrix $\mathbf{A}$. We list them in the order we proved them.

(i) $\text{nullity}(\mathbf{A}) = 0$.
(ii) There is a matrix $\mathbf{B}$ such that either $\mathbf{BA} = \mathbf{I}$ or $\mathbf{AB} = \mathbf{I}$.
(iii) The RCF of $\mathbf{A}$ is $\mathbf{I}$.
(iv) $\operatorname{rank} \mathbf{A} = n$.
(v) $\det \mathbf{A} \neq 0$.

Determinant and Rank

We now relate the rank of a matrix with determinants of its submatrices.

Lemma

Let $\mathbf{A}$ *be an* $m \times n$ *matrix, and* $r \in \mathbb{N}$*. Then*

$$\operatorname{rank} \mathbf{A} \geq r \iff \exists \text{ an } r \times r \text{ submatrix } \mathbf{B} \text{ of } \mathbf{A} \text{ with } \det \mathbf{B} \neq 0$$

Proof. Suppose $\operatorname{rank} \mathbf{A} \geq r$. Since $\operatorname{rank} \mathbf{A}$ equals the column rank of $\mathbf{A}$, there are r linearly independent columns of $\mathbf{A}$. Let $\mathbf{C}$ denote the $m \times r$ submatrix of $\mathbf{A}$ consisting of these r columns. Then the column rank of $\mathbf{C}$ is r, and so the row rank of $\mathbf{C}$ is also r. Hence there are r linearly independent rows of $\mathbf{C}$. Let $\mathbf{B}$ denote the $r \times r$ submatrix of $\mathbf{C}$ consisting of these r rows. These r rows of $\mathbf{B}$ are linearly independent, and so $\operatorname{rank} \mathbf{B} = r$. Hence $\mathbf{B}$ is invertible, and so $\det \mathbf{B} \neq 0$.

Conversely, suppose $\mathbf{B}$ is an $r \times r$ submatrix of $\mathbf{A}$ such that $\det \mathbf{B} \neq 0$. Then $\mathbf{B}$ is invertible, and so $\operatorname{rank} \mathbf{B} = r$. Hence the r rows of $\mathbf{B}$, and consequently, the corresponding r rows of $\mathbf{A}$ are linearly independent. Hence $\operatorname{rank} \mathbf{A} \geq r$. $\square$

Corollary (Determinantal Characterization of Rank)

Let $\mathbf{A}$ be an $m \times n$ matrix, and $r \in \mathbb{N}$. Then $r = \operatorname{rank} \mathbf{A}$ if and only if the following two conditions are satisfied.

(i) there is an $r \times r$ submatrix $\mathbf{B}$ of $\mathbf{A}$ such that $\det \mathbf{B} \neq 0$,

(ii) $\det \mathbf{C} = 0$ for every $(r+1) \times (r+1)$ submatrix $\mathbf{C}$ of $\mathbf{A}$.

Proof. This is an immediate consequence of the above lemma.

We remark that although the above result is of theoretical interest, it does not give a practically useful method for finding the rank of a matrix $\mathbf{A}$. On the other hand, transformation of $\mathbf{A}$ to a Row Echelon Form quickly reveals its rank.

Example

Let $\mathbf{A} := \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix}$. Since $\det \begin{bmatrix} 3 & 0 \\ -6 & 42 \end{bmatrix} \neq 0$ and since the determinants of all 3×3 submatrices of $\mathbf{A}$ are equal to 0, we see that $\operatorname{rank} \mathbf{A} = 2$.

This also follows by noting that $\mathbf{A}$ can be transformed by EROs to

$$\mathbf{A}' = \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which is in REF, and by noting that $\operatorname{rank} \mathbf{A}$ is equal to the row rank of $\mathbf{A}'$, which is 2.

We now consider another classical method of finding solutions of $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is invertible.

Proposition (Cramer's Rule)

Let $\mathbf{A} := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_k & \cdots & \mathbf{c}_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ be invertible, and let $\mathbf{b} \in \mathbb{R}^{n \times 1}$. For $k = 1, \ldots, n$, let $\mathbf{B}_k := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{b} & \cdots & \mathbf{c}_n \end{bmatrix}$ be the matrix obtained by replacing the kth column $\mathbf{c}_k$ of $\mathbf{A}$ by the right side $\mathbf{b}$ of $\mathbf{Ax} = \mathbf{b}$. Then the unique solution $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}}$ of the linear system $\mathbf{Ax} = \mathbf{b}$ is given by

$$x_k := \frac{\det \mathbf{B}_k}{\det \mathbf{A}} \quad \text{for } k = 1, \ldots, n.$$

Proof. If $\mathbf{Ax} = \mathbf{b}$, then $\mathbf{b} = x_1\mathbf{c}_1 + \cdots + x_k\mathbf{c}_k + \cdots + x_n\mathbf{c}_n$. By the first two crucial properties of the determinant function, $\det \mathbf{B}_k = \det \begin{bmatrix} \mathbf{c}_1 & \cdots & x_1\mathbf{c}_1 + \cdots + x_k\mathbf{c}_k + \cdots + x_n\mathbf{c}_n & \cdots & \mathbf{c}_n \end{bmatrix}$ $= x_k \det \mathbf{A}$ for $k = 1, \ldots, n$. Since $\mathbf{A}$ is invertible, $\det \mathbf{A} \neq 0$. Hence the result. □

Let $\mathbf{A} := \begin{bmatrix} 3 & -2 & 1 \\ -2 & 1 & 4 \\ 1 & 4 & -5 \end{bmatrix}$, $\mathbf{b} := \begin{bmatrix} 13 \\ 11 \\ -31 \end{bmatrix}$. Then $\det \mathbf{A} = -60$.

Also,

$$\det \mathbf{B}_1 = \det \begin{bmatrix} 13 & -2 & 1 \\ 11 & 1 & 4 \\ -31 & 4 & -5 \end{bmatrix} = -60,$$

$$\det \mathbf{B}_2 = \det \begin{bmatrix} 3 & 13 & 1 \\ -2 & 11 & 4 \\ 1 & -31 & -5 \end{bmatrix} = 180,$$

$$\det \mathbf{B}_3 = \det \begin{bmatrix} 3 & -2 & 13 \\ -2 & 1 & 11 \\ 1 & 4 & -31 \end{bmatrix} = -240.$$

Hence the unique solution of the linear system $\mathbf{Ax} = \mathbf{b}$ is given by $x_1 := 1,\ x_2 = -3,\ x_3 = 4$, that is, $\mathbf{x} = \begin{bmatrix} 1 & -3 & 4 \end{bmatrix}^{\mathsf{T}}$.

Note: Cramer's Rule is rarely used for solving linear systems; the preferred method is the GEM. But Cramer's Rule is of theoretical interest, especially in solutions of differential eqns.

Formula for the Inverse of a Matrix

Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{n\times n}$ with $n \geq 2$. Recall that $\mathbf{A}_{jk}$ denotes the submatrix of $\mathbf{A}$ obtained by deleting the jth row and the kth column of $\mathbf{A}$, and $M_{jk} := \det \mathbf{A}_{jk}$, the (j,k)th minor of $\mathbf{A}$, for $j,k = 1,\ldots,n$. We define $C_{jk} := (-1)^{j+k} M_{jk}$, $j,k = 1,\ldots,n$. It is called the **cofactor** of the entry a_{jk}. Then the expansion of $\det \mathbf{A}$ in terms of the kth column is given by

$$\det \mathbf{A} = \sum_{\ell=1}^{n} a_{\ell k} C_{\ell k}, \quad \text{where } k \in \{1,\ldots,n\}.$$

Define $\mathbf{C} := [C_{jk}] \in \mathbb{R}^{n\times n}$. It is called the **cofactor matrix** of $\mathbf{A}$.

Theorem

Let $\mathbf{A}$ be a square matrix. Then $\mathbf{C}^{\mathsf{T}}\mathbf{A} = (\det \mathbf{A})\mathbf{I} = \mathbf{A}\mathbf{C}^{\mathsf{T}}$. In particular, if $\det \mathbf{A} \neq 0$, then $\mathbf{A}$ is invertible and

$$\mathbf{A}^{-1} = \mathbf{C}^{\mathsf{T}} / \det \mathbf{A}.$$

Proof. Let $\mathbf{D} := \mathbf{C}^{\mathsf{T}}\mathbf{A} = [d_{jk}]$ say. By the definition of matrix multiplication, the (j,k)th entry of $\mathbf{D}$ is $d_{jk} = \sum_{\ell=1}^{n} C_{\ell j} a_{\ell k}$.

If $j = k$, then $d_{kk} = \sum_{\ell=1}^{n} C_{\ell k} a_{\ell k} = \det \mathbf{A}$, being the expansion in terms of its kth column of $\mathbf{A}$.

Let now $j \neq k$. Write $\mathbf{A} := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_k & \cdots \ \mathbf{c}_j & \cdots \mathbf{c}_n \end{bmatrix}$ in terms of its columns, and let $\mathbf{B}$ denote the matrix obtained by replacing the jth column $\mathbf{c}_j$ by the kth column $\mathbf{c}_k$ of $\mathbf{A}$, that is, $\mathbf{B} := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_k & \cdots \ \mathbf{c}_k & \cdots \mathbf{c}_n \end{bmatrix} = [b_{jk}]$, say. Then $\det \mathbf{B} = 0$ since two columns are identical. Expanding $\det \mathbf{B}$ in terms of its jth column, $\det \mathbf{B} = \sum_{\ell=1}^{n} b_{\ell j} C_{\ell j} = \sum_{\ell=1}^{n} a_{\ell k} C_{\ell j}$. Thus $d_{jk} = \sum_{\ell=1}^{n} C_{\ell j} a_{\ell k} = \det \mathbf{B} = 0$ if $j \neq k$.

This shows that $\mathbf{C}^{\mathsf{T}}\mathbf{A} = (\det \mathbf{A})\mathbf{I}$. Similarly, we can prove $\mathbf{A}\mathbf{C}^{\mathsf{T}} = (\det \mathbf{A})\mathbf{I}$, and so $\mathbf{C}^{\mathsf{T}}\mathbf{A} = (\det \mathbf{A})\mathbf{I} = \mathbf{A}\mathbf{C}^{\mathsf{T}}$.

In case $\det \mathbf{A} \neq 0$, we see that $\mathbf{A}$ is invertible, and

$$\mathbf{A}^{-1} = \mathbf{C}^{\mathsf{T}} / \det \mathbf{A}.$$

$\square$

Multiplicativity of Determinant Function

Proposition

Let $\mathbf{A}$, $\mathbf{B}$ be $n \times n$ matrices. Then $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$.

Proof. Suppose first $\mathbf{A}$ is not invertible. Then $(\det \mathbf{A}) = 0$. Also, $\mathbf{AB}$ is not invertible; otherwise there would be $\mathbf{C}$ such that $(\mathbf{AB})\mathbf{C} = \mathbf{I}$, that is, $\mathbf{A}(\mathbf{BC}) = \mathbf{I}$, which is impossible since $\mathbf{A}$ is not invertible. Hence $\det(\mathbf{AB}) = 0 = (\det \mathbf{A})(\det \mathbf{B})$.

Next, suppose $\mathbf{A}$ is invertible. Then we can transform $\mathbf{A}$ to a diagonal matrix $\mathbf{A}'$ (having nonzero diagonal elements) by elementary row transformations of type I and type II. Then $\det \mathbf{A}' = \det \mathbf{A}$ if an even number of row interchanges are involved, and $\det \mathbf{A}' = -\det \mathbf{A}$ otherwise.

We observe that the same elementary row operations transform $\mathbf{AB}$ to $\mathbf{A}'\mathbf{B}$.

To see this, we can use Q. 2.3 in Tut Sheet 2: Making an elementary row operation is equivalent to multiplying on the left by the corresponding elementary matrix! And of course $\mathbf{E}(AB) = (EA)B$ where E is any elementary matrix.

Hence $\det \mathbf{A}'\mathbf{B} = \det \mathbf{AB}$ if an even number of row interchanges are involved, and $\det \mathbf{A}'\mathbf{B} = -\det \mathbf{AB}$ otherwise.

Thus it is enough to show that

$$\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B}) \quad \text{when } \mathbf{A} \text{ is a diagonal matrix.}$$

But this is easily seen because

$$\mathbf{A} := \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix} \text{ and } \mathbf{B} := \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{bmatrix} \implies \mathbf{AB} := \begin{bmatrix} \alpha_1\mathbf{b}_1 \\ \alpha_2\mathbf{b}_2 \\ \vdots \\ \alpha_n\mathbf{b}_n \end{bmatrix},$$

where $\mathbf{b}_1, \ldots, \mathbf{b}_n$ denote the rows of $\mathbf{B}$. Hence

$$\det(\mathbf{AB}) = \alpha_1\alpha_2\cdots\alpha_n \det \mathbf{B} = (\det \mathbf{A})(\det \mathbf{B}). \qquad \square$$

Corollary

If $\mathbf{A}$ is invertible, then $\det \mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}}$.

Proof. $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \implies (\det \mathbf{A})(\det \mathbf{A}^{-1}) = \det \mathbf{I} = 1.$ $\square$

Example

Let $\mathbf{A} := \begin{bmatrix} 13 & 0 & 0 \\ 11 & 1 & 0 \\ -31 & 22 & -5 \end{bmatrix}$. Then $\det \mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}} = -\dfrac{1}{65}$.

Linear Transformations

Just as we can define a continuous function from a subset of $\mathbb{R}^n$ to $\mathbb{R}^m$, we now define a 'linear' function from a subspace of $\mathbb{R}^{n\times 1}$ to $\mathbb{R}^{m\times 1}$.

Let V be a subspace of $\mathbb{R}^{n\times 1}$, and let W be a subspace of $\mathbb{R}^{m\times 1}$. A **linear transformation** or a **linear map** from V to W is a function $T : V \to W$ which 'preserves' the operations of addition and scalar multiplication, that is, for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$,

$$T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \quad \text{and} \quad T(\alpha\,\mathbf{x}) = \alpha\, T(\mathbf{x}).$$

It follows that if $T : V \to W$ is linear, then $T(\mathbf{0}) = \mathbf{0}$, and T 'preserves' linear combinations of vectors in V, that is,

$$T(\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k) = \alpha_1 T(\mathbf{x}_1) + \cdots + \alpha_k T(\mathbf{x}_k)$$

for all $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$.

Model Example

Let $V := \mathbb{R}^{n\times 1}$, $W := \mathbb{R}^{m\times 1}$ and $\mathbf{A}$ be an $m\times n$ matrix, that is, $\mathbf{A} \in \mathbb{R}^{m\times n}$. Define $T_{\mathbf{A}} : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ by

$$T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\,\mathbf{x} \quad \text{for } \mathbf{x} \in V.$$

The properties of matrix multiplication show that $T_{\mathbf{A}}$ is linear.

Conversely, suppose $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ is linear. We show that $T = T_{\mathbf{A}}$ for some $\mathbf{A} \in \mathbb{R}^{m\times n}$. Let $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n\times 1}$. Then $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the basic column vectors in $\mathbb{R}^{n\times 1}$. Since T is linear, we obtain

$$T(\mathbf{x}) = x_1\,T(\mathbf{e}_1) + \cdots + x_n\,T(\mathbf{e}_n).$$

Define $\mathbf{c}_k := T(\mathbf{e}_k)$ for $k = 1, \ldots, n$, and $\mathbf{A} := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix}$. Then $T(\mathbf{x}) = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \mathbf{A}\,\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^{n\times 1}$. Thus $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $T = T_{\mathbf{A}}$. (Note: kth column of $\mathbf{A}$ is $T(\mathbf{e}_k)$.)

Thus every linear transformation $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ is given by

$$T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) := \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} \quad \text{for} \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1},$$

where $a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn} \in \mathbb{R}$.

Similarly, one can define a linear map $T : \mathbb{R}^{1\times n} \to \mathbb{R}^{1\times m}$, and find that for $\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \in \mathbb{R}^{1\times n}$,

$$T\left(\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}\right) := \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n & \cdots & a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Remark: Let D be an open subset of $\mathbb{R}^{1\times 2}$, $[x_0, y_0] \in D$, and let a function $f : D \to \mathbb{R}$ be differentiable at $[x_0, y_0]$. Then the total derivative of f at $[x_0, y_0]$ is a linear map (which depends on f) given by $T([x,\ y]) = \alpha x + \beta y$ for $[x,\ y] \in \mathbb{R}^{1\times 2}$, where $\alpha := f_x(x_0, y_0)$ and $\beta := f_y(x_0, y_0)$.

Let $\mathbf{A}$, $\mathbf{B} \in \mathbb{R}^{m\times n}$ and $\alpha, \beta \in \mathbb{R}$. Then $\alpha\mathbf{A} + \beta\mathbf{B} \in \mathbb{R}^{m\times n}$ and

$$T_{\alpha\mathbf{A}+\beta\mathbf{B}}(\mathbf{x}) = (\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha T_{\mathbf{A}}(\mathbf{x}) + \beta T_{\mathbf{B}}(\mathbf{x})$$

for $\mathbf{x} \in \mathbb{R}^{n\times 1}$. We write this as follows:

$$T_{\alpha\mathbf{A}+\beta\mathbf{B}} = \alpha T_{\mathbf{A}} + \beta T_{\mathbf{B}}.$$

Next, let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Then $\mathbf{AB} \in \mathbb{R}^{m\times p}$, and

$$T_{\mathbf{AB}}(\mathbf{x}) = (\mathbf{AB})\mathbf{x} = \mathbf{A}(\mathbf{Bx}) = T_{\mathbf{A}}(\mathbf{Bx}) = T_{\mathbf{A}}(T_{\mathbf{B}}(\mathbf{x})) = T_{\mathbf{A}} \circ T_{\mathbf{B}}(\mathbf{x})$$

for $\mathbf{x} \in \mathbb{R}^{p\times 1}$ by the associativity of matrix multiplication. Thus

$$T_{\mathbf{AB}} = T_{\mathbf{A}} \circ T_{\mathbf{B}}.$$

This says that the linear map associated with the product $\mathbf{AB}$ of matrices $\mathbf{A}$ and $\mathbf{B}$ is the composition of the linear maps associated with $\mathbf{A}$ and associated with $\mathbf{B}$ in the same order. This partially justifies the definition of matrix multiplication.

Examples

Let $\mathbf{A} \in \mathbb{R}^{2\times 2}$. Then $T_{\mathbf{A}} : \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$.

(i) Let $\mathbf{A} := \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Then $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ stretches each vector by a factor of 2.

(ii) Let $\mathbf{A} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} x_2 & x_1 \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ is the reflection in the line $x_1 = x_2$.

(iii) Let $\mathbf{A} := \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Then $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \mapsto \begin{bmatrix} -x_1 & -x_2 \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ is the reflection in the origin.

(iv) Let $\mathbf{A} := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, where $\theta \in (-\pi, \pi]$. Then

$T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} x_1\cos\theta - x_2\sin\theta & x_1\sin\theta + x_2\cos\theta \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ is the rotation through an angle θ.

These are geometric interpretations of matrices.