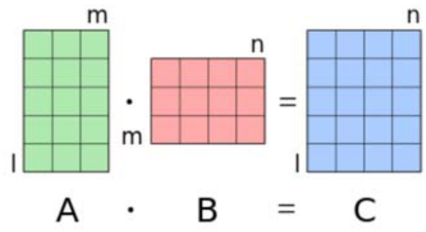
WIKIPEDIA

Matrix multiplication

In mathematics, particularly in linear algebra, **matrix multiplication** is a binary operation that produces a matrix from two matrices. For matrix multiplication, the number of columns in the first matrix must be equal to the number of rows in the second matrix. The resulting matrix, known as the **matrix product**, has the number of rows of the first and the number of columns of the second matrix. The product of matrices A and B is then denoted simply as AB.[1][2]



For matrix multiplication, the number of columns in the first matrix must be equal to the number of rows in the second matrix. The result matrix has the number of rows of the first and the number of columns of the second matrix.

Matrix multiplication was first described by the French mathematician Jacques Philippe Marie Binet in 1812,[3] to represent the composition of linear maps that are represented by matrices. Matrix multiplication is thus a basic tool of linear algebra, and as such has numerous applications in many areas of mathematics, as well as in applied mathematics, statistics, physics, economics, and engineering.[4][5] Computing matrix products is a central operation in all computational applications of linear algebra.

Contents

Notation

This article will use the following notational conventions: matrices are represented by capital letters in bold, e.g. $\mathbf{A}$; vectors in lowercase bold, e.g. $\mathbf{a}$; and entries of vectors and matrices are italic (since they are numbers from a field), e.g. A and a. Index notation is often the clearest way to express definitions, and is used as standard in the literature. The i, j entry of matrix $\mathbf{A}$ is indicated by $(\mathbf{A})_{ij}$, A_{ij} or a_{ij}, whereas a numerical label (not matrix entries) on a collection of matrices is subscripted only, e.g. $\mathbf{A}_1$, $\mathbf{A}_2$, etc.

Definition

If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

the *matrix product* $\mathbf{C} = \mathbf{AB}$ (denoted without multiplication signs or dots) is defined to be the $m \times p$ matrix[6][7][8][9]

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix}$$

such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj},$$

for $i = 1, ..., m$ and $j = 1, ..., p$.

That is, the entry c_{ij} of the product is obtained by multiplying term-by-term the entries of the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$, and summing these n products. In other words, c_{ij} is the dot product of the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$.[1]

Therefore, $\mathbf{AB}$ can also be written as

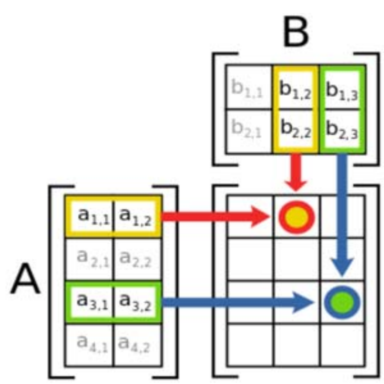
$$\mathbf{C} = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1p} + \cdots + a_{mn}b_{np} \end{pmatrix}$$

Thus the product $\mathbf{AB}$ is defined if and only if the number of columns in $\mathbf{A}$ equals the number of rows in $\mathbf{B}$,[2] in this case n.

In most scenarios, the entries are numbers, but they may be any kind of mathematical objects for which an addition and a multiplication are defined, that are associative, and such that the addition is commutative, and the multiplication is distributive with respect to the addition. In particular, the entries may be matrices themselves (see block matrix).

Illustration

The figure to the right illustrates diagrammatically the product of two matrices $\mathbf{A}$ and $\mathbf{B}$, showing how each intersection in the product matrix corresponds to a row of $\mathbf{A}$ and a column of $\mathbf{B}$.



$$\overset{4\times 2 \text{ matrix}}{\begin{bmatrix} a_{11} & a_{12} \\ \cdot & \cdot \\ a_{31} & a_{32} \\ \cdot & \cdot \end{bmatrix}} \overset{2\times 3 \text{ matrix}}{\begin{bmatrix} \cdot & b_{12} & b_{13} \\ \cdot & b_{22} & b_{23} \end{bmatrix}} = \overset{4\times 3 \text{ matrix}}{\begin{bmatrix} \cdot & c_{12} & c_{13} \\ \cdot & \cdot & \cdot \\ \cdot & c_{32} & c_{33} \\ \cdot & \cdot & \cdot \end{bmatrix}}$$

The values at the intersections marked with circles are:

$$\begin{aligned} c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{33} &= a_{31}b_{13} + a_{32}b_{23} \end{aligned}$$

Fundamental applications

Historically, matrix multiplication has been introduced for facilitating and clarifying computations in linear algebra. This strong relationship between matrix multiplication and linear algebra remains fundamental in all mathematics, as well as in physics, engineering and computer science.

Linear maps

If a vector space has a finite basis, its vectors are each uniquely represented by a finite sequence of scalars, called a coordinate vector, whose elements are the coordinates of the vector on the basis. These coordinate vectors form another vector space, which is isomorphic to the original vector space. A coordinate vector is commonly organized as a column matrix (also called *column vector*), which is a matrix with only one column. So, a column vector represents both a coordinate vector, and a vector of the original vector space.

A linear map A from a vector space of dimension n into a vector space of dimension m maps a column vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

onto the column vector

$$\mathbf{y} = A(\mathbf{x}) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}.$$

The linear map A is thus defined by the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

and maps the column vector $\mathbf{x}$ to the matrix product

$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

If B is another linear map from the preceding vector space of dimension m, into a vector space of dimension p, it is represented by a $p \times m$ matrix $\mathbf{B}.$ A straightforward computation shows that the matrix of the composite map $B \circ A$ is the matrix product $\mathbf{BA}.$ The general formula $(B \circ A)(\mathbf{x}) = B(A(\mathbf{x}))$) that defines the function composition is instanced here as a specific case of associativity of matrix product (see § Associativity below):

$$(\mathbf{BA})\mathbf{x} = \mathbf{B}(\mathbf{Ax}) = \mathbf{BAx}.$$

System of linear equations

The general form of a system of linear equations is

$$\begin{matrix} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{matrix}.$$

Using same notation as above, such a system is equivalent with the single matrix equation

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

Dot product, bilinear form and inner product

The dot product of two column vectors is the matrix product

$$\mathbf{x}^\mathsf{T}\mathbf{y},$$

where $\mathbf{x}^\mathsf{T}$ is the row vector obtained by transposing $\mathbf{x}$ and the resulting 1×1 matrix is identified with its unique entry.

More generally, any bilinear form over a vector space of finite dimension may be expressed as a matrix product

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{y},$$

and any inner product may be expressed as

$$\mathbf{x}^\dagger\mathbf{A}\mathbf{y},$$

where $\mathbf{x}^\dagger$ denotes the conjugate transpose of $\mathbf{x}$ (conjugate of the transpose, or equivalently transpose of the conjugate).

General properties

Matrix multiplication shares some properties with usual multiplication. However, matrix multiplication is not defined if the number of columns of the first factor differs from the number of rows of the second factor, and it is non-commutative,[10] even when the product remains definite after changing the order of the factors.[11][12]

Non-commutativity

An operation is commutative if, given two elements $\mathbf{A}$ and $\mathbf{B}$ such that the product $\mathbf{A}\mathbf{B}$ is defined, then $\mathbf{B}\mathbf{A}$ is also defined, and $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$.

If $\mathbf{A}$ and $\mathbf{B}$ are matrices of respective sizes $m \times n$ and $p \times q$, then $\mathbf{A}\mathbf{B}$ is defined if $n = p$, and $\mathbf{B}\mathbf{A}$ is defined if $m = q$. Therefore, if one of the products is defined, the other is not defined in general. If $m = q \neq n = p$, the two products are defined, but have different sizes; thus they cannot be equal. Only if $m = q = n = p$, that is, if $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same size, are both products defined and of the same size. Even in this case, one has in general

$$\mathbf{AB} \neq \mathbf{BA}.$$

For example

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

but

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

This example may be expanded for showing that, if $\mathbf{A}$ is a $n \times n$ matrix with entries in a field F, then $\mathbf{AB} = \mathbf{BA}$ for every $n \times n$ matrix $\mathbf{B}$ with entries in F, if and only if $\mathbf{A} = c\,\mathbf{I}$ where $c \in F$, and $\mathbf{I}$ is the $n \times n$ identity matrix. If, instead of a field, the entries are supposed to belong to a ring, then one must add the condition that c belongs to the center of the ring.

One special case where commutativity does occur is when $\mathbf{D}$ and $\mathbf{E}$ are two (square) diagonal matrices (of the same size); then $\mathbf{DE} = \mathbf{ED}$.[10] Again, if the matrices are over a general ring rather than a field, the corresponding entries in each must also commute with each other for this to hold.

Distributivity

The matrix product is distributive with respect to matrix addition. That is, if $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ are matrices of respective sizes $m \times n$, $n \times p$, $n \times p$, and $p \times q$, one has (left distributivity)

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC},$$

and (right distributivity)

$$(\mathbf{B} + \mathbf{C})\mathbf{D} = \mathbf{BD} + \mathbf{CD}.$$ [10]

This results from the distributivity for coefficients by

$$\sum_k a_{ik}(b_{kj} + c_{kj}) = \sum_k a_{ik}b_{kj} + \sum_k a_{ik}c_{kj}$$
$$\sum_k (b_{ik} + c_{ik})d_{kj} = \sum_k b_{ik}d_{kj} + \sum_k c_{ik}d_{kj}.$$

Product with a scalar

If $\mathbf{A}$ is a matrix and c a scalar, then the matrices $c\mathbf{A}$ and $\mathbf{A}c$ are obtained by left or right multiplying all entries of $\mathbf{A}$ by c. If the scalars have the commutative property, then $c\mathbf{A} = \mathbf{A}c$.

If the product $\mathbf{AB}$ is defined (that is, the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$), then

$$c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B} \text{ and } (\mathbf{AB})c = \mathbf{A}(\mathbf{B}c).$$

If the scalars have the commutative property, then all four matrices are equal. More generally, all four are equal if c belongs to the center of a ring containing the entries of the matrices, because in this case, $c\mathbf{X} = \mathbf{X}c$ for all matrices $\mathbf{X}$.

These properties result from the bilinearity of the product of scalars:

$$c\left(\sum_k a_{ik}b_{kj}\right) = \sum_k (ca_{ik})b_{kj}$$
$$\left(\sum_k a_{ik}b_{kj}\right)c = \sum_k a_{ik}(b_{kj}c).$$

Transpose

If the scalars have the commutative property, the transpose of a product of matrices is the product, in the reverse order, of the transposes of the factors. That is

$$(\mathbf{AB})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$$

where $^\mathrm{T}$ denotes the transpose, that is the interchange of rows and columns.

This identity does not hold for noncommutative entries, since the order between the entries of $\mathbf{A}$ and $\mathbf{B}$ is reversed, when one expands the definition of the matrix product.

Complex conjugate

If $\mathbf{A}$ and $\mathbf{B}$ have complex entries, then

$$(\mathbf{AB})^* = \mathbf{A}^*\mathbf{B}^*$$

where * denotes the entry-wise complex conjugate of a matrix.

This results from applying to the definition of matrix product the fact that the conjugate of a sum is the sum of the conjugates of the summands and the conjugate of a product is the product of the conjugates of the factors.

Transposition acts on the indices of the entries, while conjugation acts independently on the entries themselves. It results that, if $\mathbf{A}$ and $\mathbf{B}$ have complex entries, one has

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger,$$

where † denotes the conjugate transpose (conjugate of the transpose, or equivalently transpose of the conjugate).

Associativity

Given three matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, the products $(\mathbf{AB})\mathbf{C}$ and $\mathbf{A}(\mathbf{BC})$ are defined if and only if the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$, and the number of columns of $\mathbf{B}$ equals the number of rows of $\mathbf{C}$ (in particular, if one of the products is defined, then the other is also defined). In this case, one has the associative property

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

As for any associative operation, this allows omitting parentheses, and writing the above products as $\mathbf{ABC}$.

This extends naturally to the product of any number of matrices provided that the dimensions match. That is, if $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_n$ are matrices such that the number of columns of $\mathbf{A}_i$ equals the number of rows of $\mathbf{A}_{i+1}$ for $i = 1, ..., n - 1$, then the product

$$\prod_{i=1}^{n} \mathbf{A}_i = \mathbf{A}_1 \mathbf{A}_2 \cdots \mathbf{A}_n$$

is defined and does not depend on the order of the multiplications, if the order of the matrices is kept fixed.

These properties may be proved by straightforward but complicated summation manipulations. This result also follows from the fact that matrices represent linear maps. Therefore, the associative property of matrices is simply a specific case of the associative property of function composition.

Complexity is not associative

Although the result of a sequence of matrix products does not depend on the order of operation (provided that the order of the matrices is not changed), the computational complexity may depend dramatically on this order.

For example, if $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices of respective sizes 10×30, 30×5, 5×60, computing $(\mathbf{AB})\mathbf{C}$ needs $10 \times 30 \times 5 + 10 \times 5 \times 60 = 4{,}500$ multiplications, while computing $\mathbf{A}(\mathbf{BC})$ needs $30 \times 5 \times 60 + 10 \times 30 \times 60 = 27{,}000$ multiplications.

Algorithms have been designed for choosing the best order of products, see Matrix chain multiplication. When the number n of matrices increases, it has been shown that the choice of the best order has a complexity of $O(n \log n)$.

Application to similarity

Any invertible matrix $\mathbf{P}$ defines a similarity transformation (on square matrices of the same size as $\mathbf{P}$)

$$S_{\mathbf{P}}(\mathbf{A}) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

Similarity transformations map product to products, that is

$$S_{\mathbf{P}}(\mathbf{AB}) = S_{\mathbf{P}}(\mathbf{A}) S_{\mathbf{P}}(\mathbf{B}).$$

In fact, one has

$$\mathbf{P}^{-1}(\mathbf{AB})\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}(\mathbf{PP}^{-1})\mathbf{BP} = (\mathbf{P}^{-1}\mathbf{AP})(\mathbf{P}^{-1}\mathbf{BP}).$$

Square matrices

Let us denote $\mathcal{M}_n(R)$ the set of $n\times n$ square matrices with entries in a ring R, which, in practice, is often a field.

In $\mathcal{M}_n(R)$, the product is defined for every pair of matrices. This makes $\mathcal{M}_n(R)$ a ring, which has the identity matrix $\mathbf{I}$ as identity element (the matrix whose diagonal entries are equal to 1 and all other entries are 0). This ring is also an associative R-algebra.

If $n > 1$, many matrices do not have a multiplicative inverse. For example, a matrix such that all entries of a row (or a column) are 0 does not have an inverse. If it exists, the inverse of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^{-1}$, and, thus verifies

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

A matrix that has an inverse is an invertible matrix. Otherwise, it is a singular matrix.

A product of matrices is invertible if and only if each factor is invertible. In this case, one has

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

When R is commutative, and, in particular, when it is a field, the determinant of a product is the product of the determinants. As determinants are scalars, and scalars commute, one has thus

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det(\mathbf{A})\det(\mathbf{B}).$$

The other matrix invariants do not behave as well with products. Nevertheless, if R is commutative, $\mathbf{AB}$ and $\mathbf{BA}$ have the same trace, the same characteristic polynomial, and the same eigenvalues with the same multiplicities. However, the eigenvectors are generally different if $\mathbf{AB} \neq \mathbf{BA}$.

Powers of a matrix

One may raise a square matrix to any nonnegative integer power multiplying it by itself repeatedly in the same way as for ordinary numbers. That is,

$$\begin{aligned}\mathbf{A}^0 &= \mathbf{I},\\ \mathbf{A}^1 &= \mathbf{A},\\ \mathbf{A}^k &= \underbrace{\mathbf{AA}\cdots\mathbf{A}}_{k\text{ times}}.\end{aligned}$$

Computing the kth power of a matrix needs $k-1$ times the time of a single matrix multiplication, if it is done with the trivial algorithm (repeated multiplication). As this may be very time consuming, one generally prefers using exponentiation by squaring, which requires less than $2\log_2 k$ matrix multiplications, and is therefore much more efficient.

An easy case for exponentiation is that of a diagonal matrix. Since the product of diagonal matrices amounts to simply multiplying corresponding diagonal elements together, the kth power of a diagonal matrix is obtained by raising the entries to the power k:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}^k = \begin{pmatrix} a_{11}^k & 0 & \cdots & 0 \\ 0 & a_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^k \end{pmatrix}.$$