BSc Mathematics for Computer Scientists 2: III. Matrices and Matrix arithmetic

Péter Hajnal

Bolyai Institute, TTIK, SZTE, Szeged, Hungary

2026 Spring semester

Tables of Numbers

• In a table there are positions. The positions are structured: they are arranged into rows and columns. Let R be the set of rows, C the set of columns. $R \times C$ is the set of positions.

• The shape/type of a table is $k \times \ell$, where $k = |R|, \ell = |C|$.

Definition

The set of $k \times \ell$ tables containing real numbers is $\mathbb{R}^{k \times \ell}$. Its elements are called $k \times \ell$ real matrices.

Clarification

Two tables are equal if and only if, first, their shapes coincide. That is, they have the same number of rows (with rows correspondingly matched) and the same number of columns (with columns correspondingly matched). Thus, the positions in the two matrices correspond. Second, the number in every corresponding position are the same in the two tables/matrices.

• In summary: Specifying a matrix means giving its size and the element in each of its positions.

Special Matrices

Definition: Zero Matrix

The zero matrix $0 = 0_{k,\ell} \in \mathbb{R}^{k\times\ell}$ is the matrix whose every position contains $0 \in \mathbb{R}$.

Definition: Square Matrices

A matrix $M \in \mathbb{R}^{k\times\ell}$ is called square if $k = \ell$, that is, the number of its rows equals the number of its columns, so its shape resembles a square.

Definition: Identity Matrix

The identity matrix $1 = I = 1_{\ell,\ell} = I_{\ell,\ell} \in \mathbb{R}^{\ell\times\ell}$ is the matrix whose (i, j) position contains 0 if $i \neq j$ and 1 if $i = j$.

Special Matrices (continued)

For a square matrix $M \in \mathbb{R}^{\ell \times \ell}$, the elements $M_{i,i}$ are called diagonal elements. The position of $M_{i,j}$ is said to be below the main diagonal if $i > j$. The position of $M_{i,j}$ is said to be above the main diagonal if $i < j$.

Definition: Diagonal Matrices

A matrix $M \in \mathbb{R}^{\ell \times \ell}$ is diagonal if all elements outside the main diagonal are 0.

Definition: Lower Triangular Matrices

A matrix $M \in \mathbb{R}^{\ell \times \ell}$ is lower triangular if all elements above the main diagonal are 0.

Definition: Upper Triangular Matrices

A matrix $M \in \mathbb{R}^{\ell \times \ell}$ is upper triangular if all elements below the main diagonal are 0.

Matrices: Transposition

Definition

The transpose of a matrix $M \in \mathbb{R}^{k\times\ell}$ is the matrix of shape $\ell \times k$ whose j-th element of the i-th row is $M_{j,i}$.
The transpose of M is denoted by M^{T}.

Definition

M is symmetric if $M^{\mathsf{T}} = M$. That is, M is square (why?) and $M_{i,j} = M_{j,i}$.

Matrices: Vectors

Definition

If we list numbers in an ordered sequence, the resulting mathematical object is called a vector. The elements of the sequence are called components/coordinates. If the length of the sequence v is n (the numbers listed by v are $v_1, v_2, \ldots, v_n$, where $v_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$), then we say the vector is n-dimensional. In the case of real numbers, the set of n-dimensional vectors is denoted by $\mathbb{R}^n$.

- Naturally, everything we know about matrices can also be said about vectors. In particular, we can speak about the zero vector.

Example: 4-Dimensional Zero Vectors

$$\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Matrices: Vectors (continued)

Listing numbers can be regarded as forming a matrix. In fact, we have two possibilities.

Definition

If we write the components in a row, then the result can be regarded as a $1 \times n$ matrix. In this case we speak of a row vector: $\mathbb{R}^n \simeq \mathbb{R}^{1\times n}$.

Definition

If we write the components in a column, then the result can be regarded as an $n \times 1$ matrix. In this case we speak of a column vector: $\mathbb{R}^n \simeq \mathbb{R}^{n\times 1}$.

Observation

If v is a row vector, then v^{T} is a column vector. If u is a column vector, then u^{T} is a row vector.

- A matrix $M \in \mathbb{R}^{n\times m}$ can be viewed as a collection of m column vectors:

$$M = \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \dots & c_m \\ | & | & & | \end{pmatrix},$$

where $c_i \in \mathbb{R}^n$.

- A matrix $M \in \mathbb{R}^{n\times m}$ can also be viewed as a collection of n row vectors:

$$M = \begin{pmatrix} \text{---} & r_1^{\mathsf{T}} & \text{---} \\ \text{---} & r_2^{\mathsf{T}} & \text{---} \\ & \vdots & \\ \text{---} & r_n^{\mathsf{T}} & \text{---} \end{pmatrix}$$

where $r_i \in \mathbb{R}^m$.

Break

Matrices: Addition

Definition

Let A and B be two matrices (that is, for suitable k, ℓ, m, n we have $A \in \mathbb{R}^{k\times\ell}$, $B \in \mathbb{R}^{m\times n}$). A and B can be added if and only if their shapes coincide ($k = m$ and $\ell = n$). In this case, the matrix $A + B$ has the same shape, and for every (i, j) position ($1 \le i \le k$, $1 \le j \le \ell$)

$$(A + B)_{i,j} = A_{i,j} + B_{i,j}.$$

• Notice that in $A + B$ the plus sign connects two matrices. This is a new kind of "arithmetic," which we define. In $A_{i,j} + B_{i,j}$, the plus sign connects two real numbers. This addition was already learned in high school.

Example

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} + \begin{pmatrix} -1 & -4 & -7 & -11 \\ -2 & -5 & -8 & -12 \\ -3 & -6 & -9 & -13 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -4 & -7 \\ 3 & 1 & -1 & -4 \\ 6 & 4 & 2 & -1 \end{pmatrix}$$

Matrices: Scalar Multiplication

Definition

Let A be a matrix (that is, for suitable k, ℓ we have $A \in \mathbb{R}^{k\times\ell}$. Then the matrix A can be multiplied by a scalar $\lambda \in \mathbb{R}$. The matrix λA has the same shape, and for every (i, j) position $(1 \le i \le k,\ 1 \le j \le \ell)$

$$(\lambda \cdot A)_{i,j} = \lambda \cdot A_{i,j}.$$

- Notice that in $\lambda \cdot A$ the multiplication sign connects a number and a matrix. This is a new kind of "arithmetic," which we define. In $\lambda \cdot A_{i,j}$ the multiplication connects two real numbers. This multiplication was already learned in high school.

Example

$$2 \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 & 8 \\ 10 & 12 & 14 & 16 \\ 18 & 20 & 22 & 24 \end{pmatrix}$$

Systems of Vectors: Linear Combinations

Let $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ be d-dimensional real vectors. We may regard them as matrices (by default, as column matrices). Thus we can add them and multiply them by scalars.

Definition

Given $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$, take $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$. The expressions of the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

are called linear combinations of the vectors $v_1, v_2, \ldots, v_n$. The numbers $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ are called the coefficients of the linear combination.

Systems of Vectors: Linear Combinations (continued)

- Choosing a linear combination of the vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ means choosing the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$. The choice $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ is always possible:

$$0 \cdot v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n.$$

Definition

The representation

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \cdot v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n$$

is called the trivial representation of the zero vector (as a linear combination).

Systems of Vectors: Generated Subspace

Definition

The subspace generated by the vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ is

$$[v_1, v_2, \ldots, v_n]_{\text{lin}} = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n : \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}\} \subset \mathbb{R}^d.$$

Definition

A subset of $\mathbb{R}^d$ is called a subspace if the sum of any two of its elements and any scalar multiple of any element of it belongs to it as well. I.e. if any linear combination of its elements is also in the subset.

Observation

A linear combination of linear combinations is again a linear combination.

Systems of Vectors: Linear Independence, Spanning, Basis in $\mathbb{R}^d$

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ span $\mathbb{R}^d$ if

$$[v_1, v_2, \ldots, v_n]_{\text{lin}} = \mathbb{R}^d.$$

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ are linearly independent if $0 \in \mathbb{R}^d$ can be written only in the trivial way as their linear combination.

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ form a basis of $\mathbb{R}^d$ if they span the space and are linearly independent.

Systems of Vectors: Linear Independence, Spanning, Basis (again)

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ span $\mathbb{R}^d$ if every vector in $\mathbb{R}^d$ can be written (in at least one way) as a linear combination of $v_1, v_2, \ldots, v_n$.

Proposition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ are linearly independent if and only if every vector in $\mathbb{R}^d$ that can be written as a linear combination of $v_1, v_2, \ldots, v_n$ has a unique representation.

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ form a basis of $\mathbb{R}^d$ if every vector in $\mathbb{R}^d$ can be written in exactly one way as a linear combination of $v_1, v_2, \ldots, v_n$.

Systems of Vectors: Spanning and Basis in a Subspace

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ span the subspace A if

$$[v_1, v_2, \ldots, v_n]_{\text{lin}} = A.$$

Definition

The vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$ form a basis of the subspace A if they span the subspace and are linearly independent.

Theorem

Let A be a subspace and let $\mathcal{B}_1, \mathcal{B}_2$ be two bases of A. Then the two systems have the same cardinality, a natural number in $[0, d]$.

Definition

The common cardinality in the above theorem is called the dimension of the subspace A and is denoted by $\dim A$.

Subspaces: Examples

- Two subspaces of $\mathbb{R}^d$ are $\mathbb{R}^d$ itself and $\{0\}$.

Example

dim $\mathbb{R}^d = d$ and dim $\{0\} = 0$.

- The standard basis of $\mathbb{R}^d$ is $e_1, e_2, \ldots, e_d$, where

$$e_i = (\overset{1}{\overbrace{0}}, \ldots, \overset{i-1}{\overbrace{0}}, \overset{i}{\overbrace{1}}, \overset{i+1}{\overbrace{0}}, \ldots, \overset{d}{\overbrace{0}}).$$

Proposition

$H = \{x \in \mathbb{R}^d :$ the sum of its coordinates is $0\}$ is a subspace and dim $H = d - 1$.

Definition

A set V is called a (real) vector space if addition and (real) scalar multiplication are defined on it and satisfy the following rules:

(1) For all $u, v, w \in V$: $(u + v) + w = u + (v + w)$ and $u + v = v + u$,

(2) There exists $0_V \in V$ such that for all $v \in V$, $v + 0_V = 0_V + v = v$, and for all $v \in V$ there exists v^- with $v + v^- = v^- + v = 0_V$.

(3) For all $\alpha, \beta \in \mathbb{R}$ and $v \in V$: $(\alpha + \beta)v = \alpha v + \beta v$, $(\alpha\beta)v = \alpha(\beta v)$, $1v = v$.

(4) For all $\alpha \in \mathbb{R}$ and $u, v \in V$: $\alpha(u + v) = \alpha u + \alpha v$.

Proposition

For all $v \in V$ and $s \in \mathbb{R}$ we have $0 \cdot v = 0_V$, $s \cdot 0_V = 0_V$, $(-1) \cdot v = v^-$ and $s \cdot v = 0_V$ implies that $s = 0$ or $v = 0_V$.

Abstract Vector Spaces: Examples

Main example

$V = \mathbb{R}^d$ with the usual addition and scalar multiplication.

Further examples

- Polynomials of Degree at Most Two:
$$\{ax^2 + bx + c : a, b, c \in \mathbb{R}\}.$$
- Matrices:
$$\mathbb{R}^{k\times\ell}.$$
- Functions:
$$\{f : \mathbb{R} \to \mathbb{R}\}.$$
- Polynomials of Degree at Most Two with a Given Root:
$$\{p = ax^2 + bx + c : a, b, c \in \mathbb{R}, p(-2) = 0\}.$$
- Periodic Functions with Constant Period:
$$\{f : \mathbb{R} \to \mathbb{R}, \forall x \in \mathbb{R}\ f(x) = f(x + 2\pi)\}.$$

Abstract Vector Spaces: $\mathbb{R}^d$ as a Universal Example

Reminder

$\mathbb{R}^d$ is a d-dimensional real vector space with standard basis $e_1, e_2, \ldots, e_d$.

Let V be a d-dimensional vector space. Then we can choose a basis $\mathcal{B} = \{b_1, b_2, \ldots, b_d\}$.

Theorem

For every vector $v \in V$ there exists exactly one coordinate vector $v_\mathcal{B} = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{R}^d$ such that $v = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_d b_d$.

This correspondence can be reversed: we can establish a bijection between coordinate vectors and vectors. The pairing is compatible with the operations ($\rightarrow$ isomorphism). The two structures are identical up to naming.

Vectors: Scalar Product

Reminder

Historically: $d = 2, 3$ geometric structures: plane and space. The standard basis (e_1, e_2, e_3) is a natural choice of pairwise orthogonal unit vectors.

Definition: Scalar/Inner Product in $\mathbb{R}^d$

For $u = (u_1, u_2, \ldots, u_d)$ and $v = (v_1, v_2, \ldots, v_d)$

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \ldots + u_d v_d.$$

Definition: Norm/Length in $\mathbb{R}^d$

For $v = (v_1, v_2, \ldots, v_d)$

$$|v| = \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2 + \ldots + v_d^2}.$$

Theorem: Cauchy–Schwarz Inequality

For $u = (u_1, u_2, \ldots, u_d)$ and $v = (v_1, v_2, \ldots, v_d)$

$$\langle u, v \rangle \leq |u| \cdot |v|, \quad \text{that is} \quad \frac{\langle u, v \rangle}{|u| \cdot |v|} \in [-1, 1].$$

Vectors: Angles

Definition

The angle between vectors $u, v \in \mathbb{R}^d$ is

$$(u,v)\angle = \arccos \frac{\langle u, v\rangle}{|u| \cdot |v|}.$$

Definition

Two vectors in $\mathbb{R}^d$ are orthogonal if

$$(u,v)\angle = \frac{\pi}{2}, \quad \text{that is} \quad \langle u, v\rangle = 0.$$

• In the plane and in space these notions coincide with the usual geometric concepts. For $d = 4, 5, 6, \ldots$ these notions make it possible to introduce geometry and become familiar with it.

Hyperplanes

Definition: Hyperplane in $\mathbb{R}^d$

Let $\nu \in \mathbb{R}^d - \{0\}$ be a nonzero vector and $v_0 \in \mathbb{R}^d$ a point. Then

$$\mathcal{H}_{\nu,v_0} = \{x \in \mathbb{R}^d : \langle \nu, x - v_0 \rangle = 0\}$$

is a hyperplane in $\mathbb{R}^d$.

- For the hyperplane $\mathcal{H}_{\nu,v_0}$ defined above, the vectors lying in it and the point v_0 are orthogonal to the vector ν. $\mathcal{H}_{\nu,v_0}$ is the hyperplane through v_0 perpendicular to ν. ν is a normal vector of the hyperplane.

- Let $\nu = (\nu_1, \nu_2, \ldots, \nu_d) \in \mathbb{R}^d - \{0\}$ be a nonzero sequence of coefficients. The solution set of

$$\nu_1 x_1 + \nu_2 x_2 + \ldots + \nu_d x_d = b$$

is geometrically a hyperplane in $\mathbb{R}^d$.

Break

Definition

Let $A \in \mathbb{R}^{n \times k}$ and $B \in \mathbb{R}^{\ell \times m}$ be two matrices.

(o) These two matrices can be multiplied if and only if $k = \ell$.

(i) In this case the product matrix has size $n \times m$, that is, $AB \in \mathbb{R}^{n \times m}$.

(ii) Moreover, the (i, j)-entry is

$$(AB)_{i,j} = \sum_{s=1}^{\ell} A_{i,s} B_{s,j}.$$

Example

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1\cdot 5 + 2\cdot 7 & 1\cdot 6 + 2\cdot 8 \\ 3\cdot 5 + 4\cdot 7 & 3\cdot 6 + 4\cdot 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}.$$

Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 3 + 3\cdot 5 & 1\cdot 2 + 2\cdot 4 + 3\cdot 6 \\ 4\cdot 1 + 5\cdot 3 + 6\cdot 5 & 4\cdot 2 + 5\cdot 4 + 6\cdot 6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix}.$$

Matrices: Multiplication: Scalar Product of Vectors

Notice that the scalar product of vectors can also be interpreted as matrix multiplication.
If $u, v \in \mathbb{R}^d$ are defined as column vectors, then

$$\langle u, v \rangle = u^{\mathsf{T}} v.$$

In algebraic arguments involving matrices we prefer this notation of the scalar product. The notation $\langle u, v \rangle$ emphasizes the geometric viewpoint.

Matrix Multiplication: Viewpoint I

• Let $A \in \mathbb{R}^{n\times\ell}$ and $B \in \mathbb{R}^{\ell\times m}$ be two matrices that can be multiplied. Decompose the first factor into row vectors: $r_i \in \mathbb{R}^\ell$, and the second factor into column vectors: $c_j \in \mathbb{R}^\ell$.

$$A = \begin{pmatrix} \text{---} & r_1^\mathsf{T} & \text{---} \\ \text{---} & r_2^\mathsf{T} & \text{---} \\ & \vdots & \\ \text{---} & r_n^\mathsf{T} & \text{---} \end{pmatrix} \qquad B = \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \dots & c_m \\ | & | & & | \end{pmatrix}$$

• Multiplicability $\equiv$ the row vectors of the first factor and the column vectors of the second factor belong to spaces of the same dimension.

• Then $(AB)_{i,j} = r_i^\mathsf{T} c_j$, that is, the (i,j)-entry of the product matrix is obtained by taking the scalar product of the i-th row of A and the j-th column of B.

• The product matrix is a kind of "multiplication table" for the two vector systems obtained from the factors with respect to the scalar product.

Matrix Multiplication: Viewpoint II

• Let $A \in \mathbb{R}^{n\times\ell}$ and $B \in \mathbb{R}^{\ell\times m}$ and decompose both into row vectors. For A let these be $\alpha^\mathsf{T}, \beta^\mathsf{T}, \ldots$, and for B let them be $s_i \in \mathbb{R}^m$:

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_\ell \\ \beta_1 & \beta_2 & \ldots & \beta_\ell \\ \vdots & \vdots & \ldots & \vdots \\ \omega_1 & \omega_2 & \ldots & \omega_\ell \end{pmatrix} \qquad B = \begin{pmatrix} — & s_1^T & — \\ — & s_2^T & — \\ & \vdots & \\ — & s_\ell^T & — \end{pmatrix}$$

• Multiplicability $\equiv$ each row of the first factor contains as many entries as the second factor has rows.

• Start computing according to the definition:

$$(AB)_{1,1} = \alpha_1(s_1^\mathsf{T})_1 + \alpha_2(s_2^\mathsf{T})_1 + \ldots + \alpha_\ell(s_\ell^\mathsf{T})_1$$
$$(AB)_{1,2} = \alpha_1(s_1^\mathsf{T})_2 + \alpha_2(s_2^\mathsf{T})_2 + \ldots + \alpha_\ell(s_\ell^\mathsf{T})_2$$
$$\vdots$$
$$(AB)_{1,m} = \alpha_1(s_1^\mathsf{T})_m + \alpha_2(s_2^\mathsf{T})_m + \ldots + \alpha_\ell(s_\ell^\mathsf{T})_m$$

Matrix Multiplication: Viewpoint II (continued)

- Thus the first row of the product matrix is

$$\alpha_1 s_1^{\mathsf{T}} + \ldots + \alpha_\ell s_\ell^{\mathsf{T}}.$$

- In summary

$$AB = \begin{pmatrix} — & \alpha_1 s_1^{\mathsf{T}} + \ldots + \alpha_\ell s_\ell^{\mathsf{T}} & — \\ — & \beta_1 s_1^{\mathsf{T}} + \ldots + \beta_\ell s_\ell^{\mathsf{T}} & — \\ & \vdots & \\ — & \omega_1 s_1^{\mathsf{T}} + \ldots + \omega_\ell s_\ell^{\mathsf{T}} & — \end{pmatrix}$$

- In particular, each row of the product matrix is a linear combination of the vectors $s_1^{\mathsf{T}}, \ldots, s_\ell^{\mathsf{T}}$ (the rows of the second factor).

Matrix Multiplication: Viewpoint III

• Let $A \in \mathbb{R}^{n\times\ell}$ and $B \in \mathbb{R}^{\ell\times m}$ and decompose both into column vectors. For A let these be $o_i \in \mathbb{R}^n$, and for B let the columns be given by coefficients $\alpha, \beta, \ldots$:

$$A = \begin{pmatrix} | & | & & | \\ o_1 & o_2 & \ldots & o_\ell \\ | & | & & | \end{pmatrix} \qquad B = \begin{pmatrix} \alpha_1 & \beta_1 & \ldots & \omega_1 \\ \alpha_2 & \beta_2 & \ldots & \omega_2 \\ \vdots & \vdots & \ldots & \vdots \\ \alpha_\ell & \beta_\ell & \ldots & \omega_\ell \end{pmatrix}$$

• Multiplicability $\equiv$ each column of the second factor contains as many entries as the first factor has columns.

Matrix Multiplication: Viewpoint III (continued)

Similarly, the product matrix is

$$AB = \begin{pmatrix} \Big| & \Big| & & \Big| \\ \alpha_1 o_1 + \alpha_2 o_2 + \ldots + \alpha_\ell o_\ell & \beta_1 o_1 + \beta_2 o_2 + \ldots + \beta_\ell o_\ell & \ldots & \omega_1 o_1 + \omega_2 o_2 + \ldots + \omega_\ell o_\ell \\ \Big| & \Big| & & \Big| \end{pmatrix}$$

• In particular, each column of the product matrix is a linear combination of the column vectors $o_1, \ldots, o_\ell$ (the columns of the first factor).

Theorem

Assume the following matrix products are defined. Then

(a)

$$A(BC) = (AB)C,$$

(b)

$$A(B + C) = AB + AC, \qquad (A + B)C = AC + BC,$$

(c)

$$(\lambda A)B = \lambda(AB) = A(\lambda B),$$

(d) $A, B \in \mathbb{R}^{n \times n}$

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

(d)* $A, B \in \mathbb{R}^{k \times n}$

$$\left(AB^{\mathsf{T}}\right)^{T} = BA^{\mathsf{T}}.$$

Matrices: Algebraic Rules: Non-Commutativity

Observation

Let $A, B \in \mathbb{R}^{2\times 3}$. Then neither AB nor BA is defined.

Observation

If $A \in \mathbb{R}^{2\times 2}$ and $B \in \mathbb{R}^{2\times 3}$, then AB is defined, but BA is not.

Observation

If $A \in \mathbb{R}^{2\times 3}$ and $B \in \mathbb{R}^{3\times 2}$, then AB is defined and is a 2×2 matrix. Also BA is defined and is a 3×3 matrix. Since AB and BA have different sizes, $AB \neq BA$.

Example

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ 24 & 10 \end{pmatrix} \neq \begin{pmatrix} 10 & 16 \\ 6 & 10 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Matrices: Algebraic Rules: Square Matrices

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a square real matrix. Let $0, I \in \mathbb{R}^{n\times n}$. Then

$$AI = IA = A, \qquad A0 = 0A = 0.$$

Example

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

Example

$$\begin{pmatrix}1 & 2\\ 3 & 7\end{pmatrix}\begin{pmatrix}7 & -2\\ -3 & 1\end{pmatrix} = \begin{pmatrix}7 & -2\\ -3 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 7\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Matrices: Algebraic Rules: Inverse

Definition

Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. If there exists a matrix $A^{-1} \in \mathbb{R}^{n\times n}$ such that

$$A^{-1}A = AA^{-1} = I,$$

then we say that A is invertible. For such matrices we denote the inverse by A^{-1}.

Theorem

If $A, B \in \mathbb{R}^{n\times n}$ are invertible matrices, then AB, A^{T} and A^{-1} are also invertible, moreover

$$(AB)^{-1} = B^{-1}A^{-1}, \quad (A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}, \quad (A^{-1})^{-1} = A.$$

Matrices: Algebraic Rules: Inverse: Questions

Problem

Does every matrix have an inverse? Does every non-0 matrix have an inverse?

• The answer: No. Examples of non-invertible matrices:

$$\begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Problem

Is the inverse of an invertible matrix unique?

• The answer: Yes. If A is invertible ($\rightarrow \ A^{-1}$) and B is also an inverse, then

$$AB = I \quad \Rightarrow \quad A^{-1}AB = A^{-1}I \quad \Rightarrow \quad B = A^{-1}.$$

Problem

Given a matrix, how can we decide whether it is invertible? If it is invertible, how can we compute its inverse?

Matrices: Algebraic Rules: Inverse: $\mathbb{GL}(n, \mathbb{R})$

Definition: $\mathbb{GL}(n, \mathbb{R})$

$$\mathbb{GL}(n, \mathbb{R}) = \{A \in \mathbb{R}^{n\times n} : \ A \text{ is invertible}\}$$

Theorem

Multiplication and inverse are defined in $\mathbb{GL}(n, \mathbb{R})$. Moreover, $\mathbb{GL}(n, \mathbb{R})$ is closed under transposition.

Thank you for your attention!