european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
INVESTMENTS IN EDUCATION DEVELOPMENT

Linear algebra – vectors, matrices, determinants

Mathematics – FRDIS

MENDELU

Podpořeno projektem Průřezová inovace studijních programů Lesnické a dřevařské fakulty MENDELU v Brně (LDF) s ohledem na disciplíny společného základu http://akademie.ldf.mendelu.cz/cz (reg. č. CZ.1.07/2.2.00/28.0021) za přispění finančních prostředků EU a státního rozpočtu České republiky.

Simona Fišnarová Brno 2014

Vectors in $\mathbb{R}^n$

Definition (Vectors in $\mathbb{R}^n$)

By $\mathbb{R}^n$ we denote the set of all ordered n-tuples of real numbers, i.e.,

$$\mathbb{R}^n = \{(a_1, a_2, \ldots, a_n); a_1 \in \mathbb{R}, a_2 \in \mathbb{R}, \ldots, a_n \in \mathbb{R}\}.$$

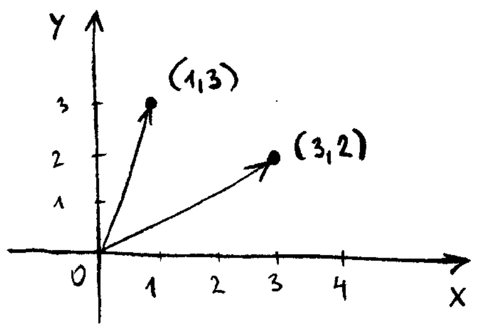
Elements of the set $\mathbb{R}^n$, the ordered n-tuples, are called **(algebraic) vectors**.
Vectors are denoted by an arrow symbol over the variable: $\vec{a} = (a_1, a_2, \ldots, a_n)$.
The numbers $a_1, \ldots, a_n$ are called **components** of the vector $\vec{a} = (a_1, a_2, \ldots, a_n)$ and the number n is called a **dimension** of the vector $\vec{a}$.
Vector $\vec{a}$ can be also written as the so-called **column vector** in the form $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$.

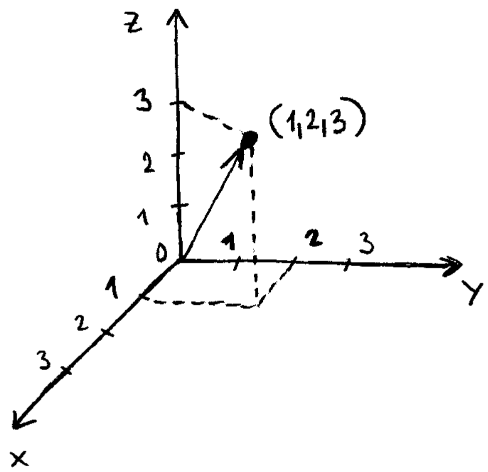
Example

$$\vec{a} = (-1, 6) \in \mathbb{R}^2, \quad \vec{b} = (2, 9, -1) \in \mathbb{R}^3, \quad \vec{c} = (2, 5, 0, -8, 9) \in \mathbb{R}^5$$

Remark (Geometric description of $\mathbb{R}^2$ and $\mathbb{R}^3$)

- $\mathbb{R}^2$ can be regarded as the set of all points in the plane, since each point in the plane is determined by an ordered pair of numbers. A vector (a_1, a_2) is represented geometrically by the point (a_1, a_2) (sometimes with arrow from the origin $(0, 0)$ included for visual clarity).
- Similarly, vectors in $\mathbb{R}^3$ can be regarded as points in a 3-dimensional coordinate space.





Definition (Vector addition and multiplication by a real number)

For $k \in \mathbb{R}$ and the vectors $\vec{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$, $\vec{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$ we define the operations **vector addition** and **multiplication by a real number**:

$$\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$
$$k\vec{a} = (ka_1, ka_2, \ldots, ka_n).$$

Example

Let $\vec{a} = (2, -1, 3),\ \vec{b} = (1, 0, 6),\ \vec{c} = (5, 6, -2, 1)$.

$$\begin{aligned} \vec{a} + \vec{b} &= (2, -1, 3) + (1, 0, 6) = (3, -1, 9) \\ -5\vec{c} &= -5 \cdot (5, 6, -2, 1) = (-25, -30, 10, -5) \end{aligned}$$

The sum $\vec{a} + \vec{c}$ is not defined, since the vectors $\vec{a}$, $\vec{c}$ are not of the same dimension ($\vec{a} \in \mathbb{R}^3$, $\vec{c} \in \mathbb{R}^4$).

Definition

- Two vectors $(a_1, a_2, \ldots, a_n),\ (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$ are called **equal** if $a_1 = b_1,\ a_2 = b_2, \ldots, a_n = b_n$.
- The vector $\vec{o} := (0, 0, \ldots, 0)$ is called the **zero vector**.
- The **difference** $\vec{a} - \vec{b}$ of two vectors $\vec{a} \in \mathbb{R}^n$ and $\vec{b} \in \mathbb{R}^n$ is defined by $\vec{a} - \vec{b} = \vec{a} + (-\vec{b})$, where the vector $-\vec{b} = (-1)\vec{b}$ is called the **negative** of $\vec{b}$.

Example

Let $\vec{a} = (2, -1, 3),\ \vec{b} = (1, 0, 6),\ \vec{c} = (5, 6, -2, 1)$.

$$\begin{aligned} -\vec{a} &= (-2, 1, -3) \\ -\vec{b} &= (-1, 0, -6) \\ \vec{a} - \vec{b} &= (2, -1, 3) - (1, 0, 6) = (1, -1, -3) \end{aligned}$$

The difference $\vec{a} - \vec{c}$ is not defined, since the vectors $\vec{a}$, $\vec{c}$ are not of the same dimension.

Definition (Scalar product)

The **scalar product** of the vectors $\vec{a} = (a_1, a_2, \ldots, a_n),\ \vec{b} = (b_1, b_2, \ldots, b_n)$ is the real number

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Example

Let $\vec{a} = (2, -1, 3, 2),\ \vec{b} = (1, 0, 6, -3)$.

$$\begin{aligned} \vec{a} \cdot \vec{b} &= 2 \cdot 1 + (-1) \cdot 0 + 3 \cdot 6 + 2 \cdot (-3) \\ &= 2 + 0 + 18 - 6 = 14. \end{aligned}$$

Definition (Linear combination of vectors)

Let $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ $(k \in \mathbb{N})$ be vectors in $\mathbb{R}^n$, $t_1, t_2, \ldots, t_k$ be real numbers. The vector

$$\vec{b} = t_1 \vec{a}_1 + t_2 \vec{a}_2 + \cdots + t_k \vec{a}_k,$$

is called a **linear combination** of the vectors $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$. The numbers $t_1, t_2, \ldots, t_k$ are called **coefficients of the linear combination**.

Example

Let $\vec{a} = (3, 2, -1),\ \vec{b} = (1, 0, -3),\ \vec{c} = (2, -1, -1)$.
Some examples of linear combinations of vectors $\vec{a},\ \vec{b},\ \vec{c}$ are

$$\begin{aligned} \vec{d} &= 3\vec{a} - 2\vec{b} + \vec{c} = (9, 5, 2) \\ \vec{o} &= 0\vec{a} + 0\vec{b} + 0\vec{c} = (0, 0, 0). \end{aligned}$$

The zero vector can be always written as a linear combination of given vectors, since it can be written as the so-called **trivial linear combination** when all coefficients of the linear combination equal zero.

Definition (Linear (in)dependence of vectors)

We say that vectors $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ are **linearly dependent** if there exist real numbers $t_1, t_2, \ldots, t_k$, **not all zero**, such that

$$t_1\vec{a}_1 + t_2\vec{a}_2 + \cdots + t_k\vec{a}_k = \vec{o}.$$

In the opposite case we say that these vectors are **linearly independent**.

Remark

It follows from the definition:

- Given vectors are linearly dependent if and only if at least one of these vectors can be written as a linear combination of the others.
- Given vectors are linearly independent if the trivial linear combination is the only possibility how to write the zero vector as a linear combination of these vectors.

Remark (Special cases of linearly (in)dependent vectors)

- Two vectors are linearly dependent if and only if one of these vectors is a constant multiple of the other.

A set of vectors (of the same dimension) is linearly dependent whenever at least one of the following statements holds:

- The set contains a zero vector.
- At least one of the vectors in the set is a constant multiple of another one.
- The number of vectors in the set is greater than a dimension of each vector.

Generally we will be able to decide about linear (in)dependence after introducing the concept of the rank of a matrix.

Example

- Vectors $(1,2,0,-3),\ (-2,-4,0,6)$ are linearly dependent, since
$$(-2,-4,0,6) = -2 \cdot (1,2,0,-3).$$
- Vectors $(1,5,0,-2),\ (5,6,-1,-1)$ are linearly independent.
- Vectors $(1,3,8),\ (0,0,0),\ (-1,0,3)$ are linearly dependent, since there is a zero vector.
- Vectors $(1,2,3),\ (3,7,1),\ (2,4,6)$ are linearly dependent since
$$(2,4,6) = 2 \cdot (1,2,3).$$
- Vectors $(1,3),\ (2,1),\ (-3,2)$ are linearly dependent since there are three vectors in the set and the dimension of the vectors is only two.
- We are not able to decide (at first sight) about linear dependence/independence of the vectors $(1,3,8),\ (1,0,-1),\ (9,3,-4)$.

Matrix-basic concepts

Definition (Matrix)

A rectangular array with m rows and n columns

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

where $a_{ij} \in \mathbb{R}$ for $i = 1,\dots,m,\ j = 1,\dots,n$, is called an $m \times n$ **matrix**. Shortly we write $A = (a_{ij})$. The set of all $m \times n$ matrices is denoted by $\mathbb{R}^{m\times n}$. The entries a_{ii} form the **main diagonal** of A. If $m = n$, then a matrix is called a **square matrix**.

Remark

The rows and the columns of a matrix can be viewed as vectors. Hence we speak about addition, multiplication by a real number, linear combination, linear (in)dependence, etc. of the rows (columns).

Zero matrix

A matrix whose entries are all zero is called a **zero matrix** and is denoted by 0. The size of a zero matrix is usually clear from the context.

Identity matrix

A square $n \times n$ matrix with the numbers 1 in the main diagonal and the numbers 0 outside this diagonal is called an $n \times n$ **identity matrix** and is denoted by I_n or shortly by I.

Example

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The transpose of a matrix

Let $A = (a_{ij})$ be an $m \times n$ matrix. The **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A, i.e., $A^T = (a_{ji})$.

Example

Let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 6 \\ 3 & -2 & 8 \\ 5 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & -2 \\ 0 & -2 & 7 \end{pmatrix}.$$

Then

$$A^T = \begin{pmatrix} 1 & -1 & 3 & 5 \\ 2 & 3 & -2 & 1 \\ 0 & 6 & 8 & 1 \end{pmatrix}, \quad B^T = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & -2 \\ 0 & -2 & 7 \end{pmatrix}.$$

We have $B = B^T$. A matrix with this property is called a **symmetric matrix**. Another example of a symmetric matrix is an identity matrix.

Matrix operations

Definition (Matrix addition and multiplication by a real number)

- Let $A = (a_{ij}),\ B = (b_{ij})$ be $m \times n$ matrices. The **sum of the matrices** A and B is the $m \times n$ matrix $C = (c_{ij})$, where
$$c_{ij} = a_{ij} + b_{ij}$$
for all i, j. We write $C = A + B$.
- Let $A = (a_{ij})$ be an $m \times n$ matrix, $t \in \mathbb{R}$. The **product of the number** t **and the matrix** A is the $m \times n$ matrix $D = (d_{ij})$, where
$$d_{ij} = t \cdot a_{ij}$$
for all i, j. We write $D = tA$.

Remark

As with vectors, we define $-A$ to mean $(-1)A$, and we write $A - B$ instead of $A + (-1)B$.

Example

Let

$$A = \begin{pmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & -3 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned} A + B &= \begin{pmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{pmatrix}, \\ 3C &= \begin{pmatrix} 6 & -9 \\ 0 & 3 \end{pmatrix}, \\ A - 2B &= \begin{pmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{pmatrix}. \end{aligned}$$

The sum $A + C$ is not defined since A and C have different sizes.

Definition (Matrix multiplication)

Let $A = (a_{ij})$ be an $m \times n$ matrix, $B = (b_{ij})$ be an $n \times p$ matrix. The **product of the matrices** A and B (in this order) is the $m \times p$ matrix $C = (c_{ij})$, where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

for $i = 1, \ldots, m,\ j = 1, \ldots, p$. We write $C = AB$.

Remark (Explanation of the previous definition)

- The number of the columns of A must be equal to the number of the rows of B. Otherwise, AB is not defined. AB has the same number of rows as A and the same number of columns as B.
- The entry c_{ij} in C is the scalar product of the i-th row of A and the j-th column of B.
- As we will see from examples, **matrix multiplication is not a commutative operation**, i.e., in general, $AB \neq BA$. The position of the factors in the product AB is emphasized by saying that A is **right-multiplied** by B or that B is **left-multiplied** by A.

Example

Let

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 3 & 0 \\ 1 & 3 \end{pmatrix}.$$

Then

$$\begin{aligned} AB &= \begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 0 \\ 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 3 \cdot 2 + 2 \cdot 3 - 1 \cdot 1 & 3 \cdot 1 + 2 \cdot 0 - 1 \cdot 3 \\ 0 \cdot 2 + 1 \cdot 3 + 2 \cdot 1 & 0 \cdot 1 + 1 \cdot 0 + 2 \cdot 3 \\ -2 \cdot 2 + 0 \cdot 3 + 1 \cdot 1 & -2 \cdot 1 + 0 \cdot 0 + 1 \cdot 3 \end{pmatrix} = \begin{pmatrix} 11 & 0 \\ 5 & 6 \\ -3 & 1 \end{pmatrix} \\ BA &= \begin{pmatrix} 2 & 1 \\ 3 & 0 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{pmatrix} \quad \text{is not defined.} \\ A^2 &= \begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 11 & 8 & 0 \\ -4 & 1 & 4 \\ -8 & -4 & 3 \end{pmatrix} \end{aligned}$$

Theorem (Properties of matrix multiplication)

Let A, B, C, I *have sizes for which the indicated sums and products are defined,* $r \in \mathbb{R}$.

1. $A(BC) = (AB)C$ *(associative law)*
2. $A(B + C) = AB + AC$ *(left distributive law)*
3. $(B + C)A = BA + CA$ *(right distributive law)*
4. $r(AB) = (rA)B = A(rB)$
5. $IA = AI = A$ *(identity for matrix multiplication)*

Warnings

1. In general, $AB \neq BA$.
2. The cancellation laws do not hold for matrix multiplication, i.e., $AB = AC \not\Rightarrow B = C$ in general.
3. In general, $AB = 0 \not\Rightarrow A = 0$ or $B = 0$.

Row echelon form

Definition (Row echelon matrix)

We say that a matrix is in **row echelon form** if it has the following properties:

- All zero rows (if exist any) are at the bottom of the matrix.
- If two successive rows are nonzero, then the second row starts with more zeros than the first one.

Example

The following matrices are in row echelon form:

$$\begin{pmatrix} 1 & 0 & 7 & 3 \\ 0 & 2 & 5 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 5 & 7 & 3 & 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 4 & 5 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{pmatrix}.$$

The leading (leftmost nonzero) entries form an echelon pattern that moves down and to the right through the matrix.

Rank of a matrix

Definition (Rank of a matrix)

Let A be a matrix. The **rank of the matrix** A, denoted by $\operatorname{rank} A$, is the maximal number of linearly independent rows of A.

Theorem

The rank of a matrix in row echelon form equals to the number of the nonzero rows of this matrix.

Example

Let

$$A = \begin{pmatrix} 0 & 5 & 7 & 3 & 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 4 & 5 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 7 & 3 \\ 0 & 1 & 5 & 2 \\ 0 & 3 & 1 & 8 \\ 7 & 5 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A is in row echelon form and $\operatorname{rank} A = 4$.
B is not in row echelon form, and hence we are not able to find $\operatorname{rank} B$ at first sight.

Definition (Equivalent row operations)

The following operations

1. multiplying a row by a nonzero constant,
2. interchanging rows in arbitrary order,
3. adding a multiple of one row to a nonzero multiple of another row,
4. omitting a zero row or a row which is a constant multiple of another row,

are called **equivalent row operations**. The fact that a matrix A is transformed into a matrix B applying a sequence of these operations is denoted by $A \sim B$ and these matrices are said to be **equivalent**.

Theorem

(i) *Any nonzero matrix can be transformed by equivalent row operations into its row echelon form.*

(ii) *The equivalent row operations preserve the rank of matrices, i.e., equivalent matrices have the same rank.*

Remark

- To find the rank of a matrix A, we apply equivalent row operations onto A in order to convert A into its row echelon form. The rank of A is then equal to the number of the nonzero rows of this echelon form.
- Any nonzero matrix may be transformed by elementary row operations into more than one matrix in row echelon form, using different sequences of row operations.
- In the following, by the **pivot** we mean a nonzero number that is used to create zeros via row operations. The row and the column containing the pivot are called **pivot row** and **pivot column**, respectively.

Example

Find the rank of the matrix

$$A = \begin{pmatrix} 3 & 6 & -1 & -2 & 4 \\ \boxed{1} & 3 & 2 & -1 & 2 \\ 0 & 2 & 1 & 0 & -1 \\ -1 & 1 & 0 & 1 & -4 \end{pmatrix} \begin{matrix} \leftarrow \\ \leftarrow \\ \\ \\ \end{matrix}$$

$$\sim \begin{pmatrix} \boxed{1} & 3 & 2 & -1 & 2 \\ 3 & 6 & -1 & -2 & 4 \\ 0 & 2 & 1 & 0 & -1 \\ -1 & 1 & 0 & 1 & -4 \end{pmatrix} \begin{matrix} -3 \\ \leftarrow + \\ \\ \leftarrow + \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 3 & 2 & -1 & 2 \\ 0 & \boxed{-3} & -7 & 1 & -2 \\ 0 & 2 & 1 & 0 & -1 \\ \cancel{0} & \cancel{4} & \cancel{2} & \cancel{0} & \cancel{-}2 \end{pmatrix} \begin{matrix} \\ 2 \\ |3 \leftarrow + \\ \\ \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 3 & 2 & -1 & 2 \\ 0 & -3 & -7 & 1 & -2 \\ 0 & 0 & -11 & 2 & -7 \end{pmatrix}$$

- The matrix in row echelon form has three nonzero rows, hence

$$\text{rank}(A) = 3.$$

The process of conversion of a matrix into its row echelon form described in the previous example can be summarized as follows:

Conversion a matrix into its row echelon form

1. Begin with the leftmost nonzero column – the so-called pivot column. Select a nonzero entry in the pivot column as a pivot (the best choice: 1 or -1).
2. Move the pivot row to the top.
3. Use row operations to create zeros in all positions below the pivot (i.e., add an appropriate constant multiple of the pivot row to an appropriate constant multiple of each row below). Optionally, use additional row operations that may simplify the matrix.
4. Cover (or ignore) the pivot row and all rows, if any, above it. Apply steps 1–3 to the submatrix that remains. Repeat the process until the matrix is in a row echelon form.

Theorem

The rank of a matrix and the rank of its transpose are the same, i.e., $\mathrm{rank}A = \mathrm{rank}A^T$.

Remark

It follows from the last theorem that all facts concerning the rank of a matrix which are valid for rows can be reformulated for columns. Hence the rank of a matrix can be also regarded as the maximal number of linearly independent columns of the matrix.

Remark (Linear (in)dependence of vectors)

We can conclude from the definition of the rank of a matrix (and from the last theorem) that m given vectors are

- linearly dependent whenever $\mathrm{rank}A < m$,
- linearly independent whenever $\mathrm{rank}A = m$,

where A is a matrix whose rows (or columns) are formed from the given vectors.

Example

The vectors

$$(3, 6, -1, -2, 4),\ (1, 3, 2, -1, 2),\ (0, 2, 1, 0, -1),\ (-1, 1, 0, 1, -4)$$

are linearly dependent since the rank of the matrix

$$\begin{pmatrix} 3 & 6 & -1 & -2 & 4 \\ 1 & 3 & 2 & -1 & 2 \\ 0 & 2 & 1 & 0 & -1 \\ -1 & 1 & 0 & 1 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 2 & -1 & 2 \\ 0 & -3 & -7 & 1 & -2 \\ 0 & 0 & -11 & 2 & -7 \end{pmatrix}$$

equals 3, see the example above.

Example

The vectors

$$(0,2,0,3,0),\ (1,0,0,0,0),\ (0,0,0,5,-1),\ (0,0,1,0,-4)$$

are linearly independent since the rank of the matrix

$$\begin{pmatrix} 0 & 2 & 0 & 3 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & -1 \\ 0 & 0 & 1 & 0 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 5 & -1 \end{pmatrix}$$

equals 4.

Inverse matrix

Definition (Inverse matrix)

Let A be an $n \times n$ square matrix. We say that A is **invertible** if there exists an $n \times n$ matrix A^{-1} such that

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

The matrix A^{-1} is called the **inverse** of A.

If A is invertible, than the inverse A^{-1} is determined uniquely by A. But not every matrix is invertible.

Theorem

- *A matrix A is invertible if and only if A can be transformed into I by equivalent row operations.*
- *Any sequence of equivalent row operations that transforms A into I also transforms I into A^{-1}.*

The last theorem leads to a method for finding the inverse of a matrix.

An algorithm for finding A^{-1}

Let A be a square matrix.

1. We form the "partitioned" matrix $(A|I)$.
2. We apply to $(A|I)$ the equivalent row operations that converts A into I. (We create zeros in all positions below and above the pivots.)
3. If A is converted into I (the left-hand side of the resulting matrix is I), then A^{-1} appears on the right-hand side of the resulting matrix, i.e.,

$$(A|I) \sim \cdots \sim (I|A^{-1}).$$

4. If A cannot be converted to I (if there appears a zero row on the left-hand side), then A does not have an inverse.

! We cannot use any column operations.

Example

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 0 & 2 & -2 \end{pmatrix} \qquad A^{-1} = ?$$

$$\left(\begin{array}{ccc|ccc} \boxed{1} & -1 & 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 0 & 0 & 1 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & \boxed{2} & -3 & -2 & 1 & 0 \\ 0 & 2 & -2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} | \cdot 2 \\ \\ \\ \end{array}$$

$$\sim \left(\begin{array}{ccc|ccc} 2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 2 & -3 & -2 & 1 & 0 \\ 0 & 0 & \boxed{1} & 2 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 2 & 0 & 1 \\ 0 & 2 & 0 & 4 & -2 & 3 \\ 0 & 0 & 1 & 2 & -1 & 1 \end{array}\right) \begin{array}{l} | : 2 \\ | : 2 \\ \\ \end{array}$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 2 & -1 & \frac{3}{2} \\ 0 & 0 & 1 & 2 & -1 & 1 \end{array}\right) \implies A^{-1} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 1 \\ 4 & -2 & 3 \\ 4 & -2 & 2 \end{pmatrix}.$$

Example

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 4 & 3 & 7 \end{pmatrix} \qquad B^{-1} = ?$$

$$\left(\begin{array}{ccc|ccc} \boxed{1} & 1 & 2 & 1 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 & 0 \\ 4 & 3 & 7 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \cdot(-2)\ \ \cdot(-4) \\ \leftarrow + \\ \leftarrow\ \ \ \ \ \ \ \ + \end{array} \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & -1 & -1 & -4 & 0 & 1 \end{array}\right).$$

It follows from the last matrix that B^{-1} does not exist.

Determinants

Determinant of a matrix

Let $A = (a_{ij})$ be an $n \times n$ matrix. The **determinant** of A is a real number $\det A$ (also denoted by $|A|$), i.e.,

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

which is assigned to A “in a certain way”.

Determinants of small matrices

- If $n = 1$, i.e., $A = a_{11}$, then $\det A = a_{11}$.
- If $n = 2$:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12},$$

- If $n = 3$:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{12}a_{23}a_{31} \\ -a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{21}a_{12}a_{33}. \end{matrix}$$

Sarus' rule for determinant of 3×3 matrix

The formula for calculation determinants of 3×3 matrices can be remembered as follows:

- We write once more the first and the second row below the determinant.
- Then we multiply each three elements in the main diagonal and below (these products are with the sign $+$).
- Next we multiply the elements in the diagonal $a_{31} - a_{22} - a_{13}$ and below (these products are with the sign $-$),
- and we sum up all these products.

$$\begin{array}{c} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ \hline \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix} \end{array} = \begin{matrix} a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ -a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} \end{matrix}$$

- **This trick does not generalize in any way to 4×4 or larger matrices!**

When computing determinants of 4×4 or larger matrices we can use the following Laplace expansion.

Theorem (The Laplace expansion)

Let $n \geq 2$ and assume that the determinant of $(n-1) \times (n-1)$ matrix has been defined. Denote by M_{ij} the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i-th row and the j-the column. Then

$$\det A = (-1)^{i+1} a_{i1} M_{i1} + (-1)^{i+2} a_{i2} M_{i2} + \cdots + (-1)^{i+n} a_{in} M_{in}$$
$$\text{for} \quad i = 1, \ldots, n,$$

*(**the so-called Laplace expansion along the i-th row**)*

or in an alternative way:

$$\det A = (-1)^{1+j} a_{1j} M_{1j} + (-1)^{2+j} a_{2j} M_{2j} + \cdots + (-1)^{n+j} a_{nj} M_{nj}$$
$$\text{for} \quad j = 1, \ldots, n.$$

*(**the so-called Laplace expansion along the j-th column**)*

Example (Laplace expansion)

When using Laplace expansion for computing the determinant of a matrix, we choose the row or column that contains many zeros.

- The best choice in the following determinant is the third row.
- The obtained two determinants of 3×3 matrices can be computed using the Sarus rule.

$$\begin{vmatrix} -1 & 2 & 5 & -1 \\ -1 & -2 & -4 & 2 \\ 2 & 0 & 1 & 0 \\ -3 & 3 & -7 & -1 \end{vmatrix}$$

$$= (-1)^{3+1} 2 \begin{vmatrix} 2 & 5 & -1 \\ -2 & -4 & 2 \\ 3 & -7 & -1 \end{vmatrix} + (-1)^{3+3} 1 \begin{vmatrix} -1 & 2 & -1 \\ -1 & -2 & 2 \\ -3 & 3 & -1 \end{vmatrix}$$

$$= 2[8 - 14 + 30 - (12 - 28 + 10)] + [-2 + 3 - 12 - (-6 - 6 + 2)] = 59$$

Note that to calculate the determinant of $n \times n$ matrix we need to calculate n determinants of $(n-1) \times (n-1)$ matrices. To calculate each of these n determinants we need to calculate $n-1$ determinants of $(n-2) \times (n-2)$ matrices, and so on.
This means that the Laplace expansion is not suitable for computing determinants of large matrices with not many zeros.
The following statement says how to simplify computations by creating more zeros in the matrix.

Theorem (Properties of determinants)

Let A be an $n \times n$ square matrix.

1. *The determinant of a matrix and the determinant of its transpose are the same, i.e.,* $\det A = \det A^T$.
2. *If a multiple of a row is added to another row, the value of the determinant is unchanged.*
3. *If any two rows (columns) of A are interchanged, the determinant of the new matrix equals $-\det A$.*
4. *If a row (column) of A is multiplied by a number k, the determinant of the new matrix is $k \det A$.*
5. *If a row (column) of A is divided by a number $k \neq 0$, the determinant of the new matrix is $\frac{1}{k} \det A$.*
6. *The determinant of a matrix with zeros below the main diagonal is equal to the product of the entries on the main diagonal, i.e.,* $\det A = a_{11}a_{22} \cdots a_{nn}$.

The last theorem leads to other methods for computing the determinant of a matrix using the row or column operations:

Method 1

We convert the matrix into the matrix which has zeros below the main diagonal. Then we find the product of the entries of the main diagonal.

Method 2

We convert the matrix into the matrix which has at most one nonzero entry in some row (or column). Then we use the Laplace expansion along this row (or column).

We have to be careful when using the row (or column) operations, since some of these operations change the value of the determinant, namely:

- changing rows (columns),
- multiplying (dividing) the non-pivot rows (columns).

Theorem

Let A be an $n \times n$ square matrix. Then $\det A = 0$ whenever at least one of the following conditions is satisfied:

1. *A contains a zero row (or column).*
2. *Two rows (columns) of A are identical.*
3. *One row (column) is a constant multiple of another row (column).*

Regular and singular matrix

Theorem

Let A be an $n \times n$ square matrix. Then the following statements are equivalent:

1. *A is invertible, i.e., A^{-1} exists.*
2. $\det A \neq 0$.
3. $\mathrm{rank} A = n$.
4. *The rows (columns) of A are linearly independent.*

Definition (Regular and singular matrix)

We say that a square matrix is **regular (or non-singular)**, if it has the properties stated in the last theorem. In the opposite case, the matrix is said to be **singular**.

Using the computer algebra systems

Wolfram Alpha:

`http://www.wolframalpha.com/`

Example

Find the product of the matrices using the Wolfram Alpha:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 4 \end{pmatrix}.$$

Solution:

{{1,2,2},{2,1,3}}*{{1,2},{3,1},{1,4}}

Example

Using the Wolfram Alpha fing the rank, determinant and the inverse matrix of the matrix

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

Solution:

1. rank:

   rank{{1,2,3},{2,0,1},{3,2,1}}

2. determinant:

   det{{1,2,3},{2,0,1},{3,2,1}}

3. inverse matrix:

   inv{{1,2,3},{2,0,1},{3,2,1}}