Vector spaces

Linear algebra

Lectures 6-7

Gábor V. Nagy

Bolyai Intitute

Szeged, 2020.

Definition. $\mathbb{R}^n$ denotes the set of $n \times 1$ (real) column vectors:

$$\mathbb{R}^n = \left\{ [a_1, a_2, \ldots, a_n]^T : a_1, a_2, \ldots, a_n \in \mathbb{R} \right\}.$$

For the sake of readability, we write the column vector $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ as $[a_1, a_2, \ldots, a_n]^T$ in the lecture slides.

Definition. $\mathbb{R}^n$ denotes the set of $n \times 1$ (real) column vectors:

$$\mathbb{R}^n = \left\{ [a_1, a_2, \ldots, a_n]^T : a_1, a_2, \ldots, a_n \in \mathbb{R} \right\}.$$

Definition. We define 'addition' and 'multiplication by scalar λ' operations for vectors in $\mathbb{R}^n$ in the same way as we did for $n \times 1$ matrices:

$$[a_1, \ldots, a_n]^T + [b_1, \ldots, b_n]^T := [a_1 + b_1, \ldots, a_n + b_n]^T$$
$$\lambda \cdot [a_1, \ldots, a_n]^T := [\lambda a_1, \ldots, \lambda a_n]^T.$$

For example, in $\mathbb{R}^3$,

$$[3, -1, 4]^T + [1, 2, 3]^T = [4, 1, 7]^T$$
$$3 \cdot [2, 5, -6]^T = [6, 15, -18]^T.$$

Definition. $\mathbb{R}^n$ denotes the set of $n \times 1$ (real) column vectors:

$$\mathbb{R}^n = \left\{ [a_1, a_2, \ldots, a_n]^T : a_1, a_2, \ldots, a_n \in \mathbb{R} \right\}.$$

Definition. We define 'addition' and 'multiplication by scalar λ' operations for vectors in $\mathbb{R}^n$ in the same way as we did for $n \times 1$ matrices:

$$[a_1, \ldots, a_n]^T + [b_1, \ldots, b_n]^T := [a_1 + b_1, \ldots, a_n + b_n]^T$$
$$\lambda \cdot [a_1, \ldots, a_n]^T := [\lambda a_1, \ldots, \lambda a_n]^T.$$

For example, in $\mathbb{R}^3$,

$$[3, -1, 4]^T + [1, 2, 3]^T = [4, 1, 7]^T$$
$$3 \cdot [2, 5, -6]^T = [6, 15, -18]^T.$$

Notations. $\mathbf{0}$ denotes the zero vector $[0, 0, \ldots, 0]^T$ in $\mathbb{R}^n$.
$-\mathbf{v}$ denotes the vector $(-1) \cdot \mathbf{v}$ for any vector $\mathbf{v} \in \mathbb{R}^n$.

Theorem. For all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and for all scalars $\lambda, \mu \in \mathbb{R}$,

(1) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$,
(2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$,
(3) $\mathbf{v} + \mathbf{0} = \mathbf{v}$,
(4) $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$,
(5) $\lambda \cdot (\mathbf{u} + \mathbf{v}) = \lambda \cdot \mathbf{u} + \lambda \cdot \mathbf{v}$,
(6) $(\lambda + \mu) \cdot \mathbf{v} = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v}$,
(7) $(\lambda\mu) \cdot \mathbf{v} = \lambda \cdot (\mu \cdot \mathbf{v})$,
(8) $1 \cdot \mathbf{v} = \mathbf{v}$.

Theorem. For all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and for all scalars $\lambda, \mu \in \mathbb{R}$,

(1) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$,

(2) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$,

(3) $\mathbf{v}+\mathbf{0} = \mathbf{v}$,

(4) $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$,

(5) $\lambda \cdot (\mathbf{u}+\mathbf{v}) = \lambda \cdot \mathbf{u} + \lambda \cdot \mathbf{v}$,

(6) $(\lambda+\mu) \cdot \mathbf{v} = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v}$,

(7) $(\lambda\mu) \cdot \mathbf{v} = \lambda \cdot (\mu \cdot \mathbf{v})$,

(8) $1 \cdot \mathbf{v} = \mathbf{v}$.

Mathematical structures (equipped with 'addition' and 'scalar multiplication' operations) that satisfy properties (1)-(8) are called vector spaces. (We make this more precise on the next side.)

$\mathbb{R}^n$ is the most important example of vector spaces.

Definition. A (real) vector space is a nonempty set V (whose elements will be called vectors), equipped with a binary operation $\oplus$ (addition of vectors) and for each real number λ, a unary operation $\lambda\odot$ (multiplication by scalar λ) that satisfy the following properties for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and for all scalars $\lambda, \mu \in \mathbb{R}$:

(1) $(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = \mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w})$,

(2) $\mathbf{u} \oplus \mathbf{v} = \mathbf{v} \oplus \mathbf{u}$,

(3) There exists a vector $\mathbf{0} \in V$ such that $\mathbf{v} \oplus \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$,

(4) For any vector $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ for which $\mathbf{v} \oplus (-\mathbf{v}) = \mathbf{0}$,

(5) $\lambda \odot (\mathbf{u} \oplus \mathbf{v}) = \lambda \odot \mathbf{u} \oplus \lambda \odot \mathbf{v}$,

(6) $(\lambda + \mu) \odot \mathbf{v} = \lambda \odot \mathbf{v} \oplus \mu \odot \mathbf{v}$,

(7) $(\lambda\mu) \odot \mathbf{v} = \lambda \odot (\mu \odot \mathbf{v})$,

(8) $1 \odot \mathbf{v} = \mathbf{v}$.

Example. $\mathbb{R}^n$ is a vector space (when $\oplus$ is the addition of vectors in $\mathbb{R}^n$, and $\lambda\odot$ is the multiplication of vectors in $\mathbb{R}^n$ by scalar λ). However, there exist other vector spaces, too.

Example. $\mathbb{R}^n$ is a vector space (when $\oplus$ is the addition of vectors in $\mathbb{R}^n$, and $\lambda\odot$ is the multiplication of vectors in $\mathbb{R}^n$ by scalar λ). However, there exist other vector spaces, too.

Convention. Instead of $\oplus$ and $\lambda\odot$, the addition and scalar multiplication operations will be denoted by $+$ and $\lambda\cdot$ in all vector spaces (like in $\mathbb{R}^n$).

Example. $\mathbb{R}^n$ is a vector space (when $\oplus$ is the addition of vectors in $\mathbb{R}^n$, and $\lambda\odot$ is the multiplication of vectors in $\mathbb{R}^n$ by scalar λ). However, there exist other vector spaces, too.

Convention. Instead of $\oplus$ and $\lambda\odot$, the addition and scalar multiplication operations will be denoted by $+$ and $\lambda\cdot$ in all vector spaces (like in $\mathbb{R}^n$).

Remark. In the context of a vector space, scalar just means 'real number'. It is used when it is important to emphasize "not a vector".

Example. $\mathbb{R}^n$ is a vector space (when $\oplus$ is the addition of vectors in $\mathbb{R}^n$, and $\lambda\odot$ is the multiplication of vectors in $\mathbb{R}^n$ by scalar λ). However, there exist other vector spaces, too.

Convention. Instead of $\oplus$ and $\lambda\odot$, the addition and scalar multiplication operations will be denoted by $+$ and $\lambda\cdot$ in all vector spaces (like in $\mathbb{R}^n$).

Remark. In the context of a vector space, scalar just means 'real number'. It is used when it is important to emphasize "not a vector".

Claim. The followings are true in any vector space V:

(i) $0 \cdot \mathbf{v} = \mathbf{0}$, for any vector $\mathbf{v} \in V$.

(ii) $\lambda \cdot \mathbf{0} = \mathbf{0}$, for any scalar $\lambda \in \mathbb{R}$.

(iii) $(-1)\mathbf{v} = -\mathbf{v}$ for any vector $\mathbf{v} \in V$.

Definition. Let V be a vector space. The subset W of V is a subspace of V, if

- $\mathbf{0} \in W$ (i.e. the zero vector belongs to W),
- whenever $\mathbf{u}, \mathbf{v} \in W$, it follows that $\mathbf{u} + \mathbf{v} \in W$, and
- whenever $\mathbf{u} \in W$, it follows that $\lambda \cdot \mathbf{u} \in W$ for all scalars $\lambda \in \mathbb{R}$.

Remark. A subspace W is also a vector space, contained within the (larger) vector space V.

Definition. Let V be a vector space. The subset W of V is a subspace of V, if

- $\mathbf{0} \in W$ (i.e. the zero vector belongs to W),
- whenever $\mathbf{u}, \mathbf{v} \in W$, it follows that $\mathbf{u} + \mathbf{v} \in W$, and
- whenever $\mathbf{u} \in W$, it follows that $\lambda \cdot \mathbf{u} \in W$ for all scalars $\lambda \in \mathbb{R}$.

Remark. A subspace W is also a vector space, contained within the (larger) vector space V.

Example 1. If W_1 denotes the set of all vectors in $\mathbb{R}^3$ whose first coordinate is 1, then W_1 is not a subspace of $\mathbb{R}_3$. This is because, for example, the vectors $[1, 2, 3]^T$ and $[1, 5, 6]^T$ are in W_1, but their sum, $[2, 7, 9]^T$ is not in W_1, violating the second condition. (Our we could argue that the zero vector $[0, 0, 0]^T$ does not belong to W_1.)

Definition. Let V be a vector space. The subset W of V is a subspace of V, if

- $\mathbf{0} \in W$ (i.e. the zero vector belongs to W),
- whenever $\mathbf{u}, \mathbf{v} \in W$, it follows that $\mathbf{u} + \mathbf{v} \in W$, and
- whenever $\mathbf{u} \in W$, it follows that $\lambda \cdot \mathbf{u} \in W$ for all scalars $\lambda \in \mathbb{R}$.

Example 2. If W_0 denotes the set of all vectors in $\mathbb{R}^3$ whose first coordinate is 0, then W_0 is a subspace of $\mathbb{R}_3$:

- The zero vector $[0, 0, 0]^T$ belongs to W_0, because its first coordinate is 0.
- If we add any two vectors with first coordinate 0, then the we obtain a vector with first coordinate 0.
- Every scalar multiple of a vector with first coordinate 0 has also first coordinate 0.

Definition. In a vector space V, the linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ with (scalar) coefficients $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ is the vector

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k.$$

Definition. In a vector space V, the linear combination of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ with (scalar) coefficients $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ is the vector

$$\lambda_1 \cdot \mathbf{v}_1 + \dots + \lambda_k \cdot \mathbf{v}_k.$$

Example. In $\mathbb{R}^2$, the linear combination of vectors $[1,5]^T$, $[4,-2]^T$, $[0,3]^T$ with coefficients $2, 3, -1/2$ is the vector

$$2 \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} + 3 \cdot \begin{bmatrix} 4 \\ -2 \end{bmatrix} + (-1/2) \cdot \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \\ 5/2 \end{bmatrix}.$$

Definition. In a vector space V, the linear combination of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ with (scalar) coefficients $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ is the vector

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k.$$

Example. In $\mathbb{R}^2$, the linear combination of vectors $[1,5]^T$, $[4,-2]^T$, $[0,3]^T$ with coefficients $2, 3, -1/2$ is the vector

$$2 \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} + 3 \cdot \begin{bmatrix} 4 \\ -2 \end{bmatrix} + (-1/2) \cdot \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \\ 5/2 \end{bmatrix}.$$

Def + Claim. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be vectors in a vector space V. Span$(\mathbf{v}_1, \dots, \mathbf{v}_k)$ denotes the set of all linear combinations of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, that is,

$$\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \{\lambda_1 \cdot \mathbf{v}_1 + \dots \lambda_k \cdot \mathbf{v}_k : \lambda_1, \dots, \lambda_k \in \mathbb{R}\}.$$

Then Span$(\mathbf{v}_1, \dots, \mathbf{v}_k)$ is a subspace of V, and it is called the subspace generated by the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$.

In the vector space $\mathbb{R}^n$, for any given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathbf{u}$, we can find (the coefficients of) all linear combinations of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ that yields $\mathbf{u}$. This can be done by solving a linear system.

In the vector space $\mathbb{R}^n$, for any given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathbf{u}$, we can find (the coefficients of) all linear combinations of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ that yields $\mathbf{u}$. This can be done by solving a linear system.

Let's demonstrate this with an example in $\mathbb{R}^3$:

Exercise. Find all triples of coefficients $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, for which

$$\lambda_1 \cdot \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} 3 \\ -5 \\ 6 \end{bmatrix} + \lambda_3 \cdot \begin{bmatrix} 7 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ 2 \end{bmatrix}$$

holds.

Exercise. Find all triples of coefficients $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, for which

$$\lambda_1 \cdot \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} 3 \\ -5 \\ 6 \end{bmatrix} + \lambda_3 \cdot \begin{bmatrix} 7 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ 2 \end{bmatrix}$$

holds.

Solution. By inspecting the first, second and third coordinates, the condition can be rewritten as

$$\begin{cases} 5\lambda_1 + 3\lambda_2 + 7\lambda_3 = 8 \\ \lambda_1 - 5\lambda_2 + 4\lambda_3 = -1 \\ 3\lambda_1 + 6\lambda_2 + 7\lambda_3 = 2. \end{cases}$$

This means that we need to find the solutions of the linear system

$$\begin{cases} 5x_1 + 3x_2 + 7x_3 = 8 \\ x_1 - 5x_2 + 4x_3 = -1 \\ 3x_1 + 6x_2 + 7x_3 = 2 \end{cases} \longleftrightarrow \left[\begin{array}{ccc|c} 5 & 3 & 7 & 8 \\ 1 & -5 & 4 & -1 \\ 3 & 6 & 7 & 2 \end{array}\right].$$

The solution to this system is $(3, 0, -1)$, so the coefficients are $\lambda_1 = 3$, $\lambda_2 = 0$, $\lambda_3 = -1$. $\square$

Exercise. Find all triples of coefficients $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, for which

$$\lambda_1 \cdot \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} 3 \\ -5 \\ 6 \end{bmatrix} + \lambda_3 \cdot \begin{bmatrix} 7 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ 2 \end{bmatrix}$$

holds.

Solution. By inspecting the first, second and third coordinates, the condition can be rewritten as

$$\begin{cases} 5\lambda_1 + 3\lambda_2 + 7\lambda_3 = 8 \\ \lambda_1 - 5\lambda_2 + 4\lambda_3 = -1 \\ 3\lambda_1 + 6\lambda_2 + 7\lambda_3 = 2. \end{cases}$$

This means that we need to find the solutions of the linear system

$$\begin{cases} 5x_1 + 3x_2 + 7x_3 = 8 \\ x_1 - 5x_2 + 4x_3 = -1 \\ 3x_1 + 6x_2 + 7x_3 = 2 \end{cases} \longleftrightarrow \left[\begin{array}{ccc|c} \boxed{\begin{matrix} 5 \\ 1 \\ 3 \end{matrix}} & \boxed{\begin{matrix} 3 \\ -5 \\ 6 \end{matrix}} & \boxed{\begin{matrix} 7 \\ 4 \\ 7 \end{matrix}} & \boxed{\begin{matrix} 8 \\ -1 \\ 2 \end{matrix}} \end{array}\right].$$

The COLUMN vectors of the augmented matrix are exactly the vectors in the exercise, keeping their order! . . . $\square$

In general, we have the following:

Theorem. For any given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathbf{u}$ in $\mathbb{R}^n$, the coefficients $(\lambda_1, \ldots, \lambda_k)$ for which

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k = \mathbf{u}$$

holds (where $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$), are precisely the solutions of the linear system with augmented matrix

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{v}_1} & \boxed{\mathbf{v}_2} & \boxed{\mathbf{v}_3} & \ldots & \boxed{\mathbf{v}_k} & \boxed{\mathbf{u}} \end{array}\right].$$

This theorem says that finding the coefficients of those linear combinations of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ which yield the vector $\mathbf{u}$ is equivalent to solving the above linear system. (Which has 0, 1 or infinitely many solutions, see the last slide of Lecture 5.)

In view of the previous theorem, in the vector space $\mathbb{R}^n$, we can easily decide whether a vector $\mathbf{u}$ is contained in the spanned subspace Span$(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, i.e. whether $\mathbf{u}$ can be obtained as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$:

We only have to decide whether the linear system

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{v}_1} & \boxed{\mathbf{v}_2} & \boxed{\mathbf{v}_3} & \ldots & \boxed{\mathbf{v}_k} & \boxed{\mathbf{u}} \end{array}\right].$$

has a solution or not. And a linear system has solution if and only if its row-echelon form has no contradicting rows (see the last slide of Lecture 5).

Example. See the solution of Exercise 4.5.9.

Definition. The linear combination

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k$$

is trivial if all coefficients are zero:

$$\lambda_1 = \cdots = \lambda_k = 0.$$

Definition. The linear combination

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k$$

is trivial if all coefficients are zero:

$$\lambda_1 = \cdots = \lambda_k = 0.$$

Definition. In a vector space V, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ are said to be linearly dependent if there is a nontrivial linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ that yields the zero vector, i.e. if there exists coefficients $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ such that not all of them are zero and

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k = \mathbf{0}.$$

Definition. The linear combination

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k$$

is trivial if all coefficients are zero:

$$\lambda_1 = \cdots = \lambda_k = 0.$$

Definition. In a vector space V, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ are said to be linearly dependent if there is a nontrivial linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ that yields the zero vector, i.e. if there exists coefficients $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ such that not all of them are zero and

$$\lambda_1 \cdot \mathbf{v}_1 + \cdots + \lambda_k \cdot \mathbf{v}_k = \mathbf{0}.$$

Example. The vectors $\mathbf{v}_1 = [1, 2, 1]^T$, $\mathbf{v}_2 = [-1, -1, 2]^T$ and $v_3 = [5, 7, -4]^T$ are linearly dependent in $\mathbb{R}^3$, because

$$2 \cdot \mathbf{v}_1 + (-3) \cdot \mathbf{v}_2 + (-1) \cdot \mathbf{v}_3 = \mathbf{0}.$$

Definition. If the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ are not linearly dependent, then we say that they are linearly independent.

In other words, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ are linearly independent, iff the only linear combination of these vectors that gives the zero vector is the trivial linear combination.

Definition. If the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ are not linearly dependent, then we say that they are linearly independent.

Example. The vectors $\mathbf{v}_1 = [1, 2, 2]^T$, $\mathbf{v}_2 = [-1, -1, 2]^T$ and $v_3 = [5, 7, -4]^T$ are linearly independent in $\mathbb{R}^3$. To see this, find all triples $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that

$$\lambda_1 \cdot \mathbf{v}_1 + \lambda_2 \cdot \mathbf{v}_2 + \lambda_3 \cdot \mathbf{v}_3 = \mathbf{0},$$

by solving the linear system

$$\left[\begin{array}{ccc|c} 1 & -1 & 5 & 0 \\ 2 & -1 & 7 & 0 \\ 2 & 2 & 4 & 0 \end{array}\right].$$

This linear system has the only solution $(0, 0, 0)$, which means that the trivial linear combination

$$0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + 0 \cdot \mathbf{v}_3$$

is the only linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ that yields $\mathbf{0}$.

In general, we have the following:

Theorem. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ are linearly independent if and only if the linear system with augmented matrix

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{v}_1} & \boxed{\mathbf{v}_2} & \boxed{\mathbf{v}_3} & \ldots & \boxed{\mathbf{v}_k} & \boxed{\mathbf{0}} \end{array}\right].$$

has no other solution than $(0, 0, 0, 0, \ldots, 0)$.

Definition. The vector system $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V is said to be a generator system of V if every vector in V can be obtained as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

In other words, $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ is a generator system of V if

$$\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = V.$$

Definition. The vector system $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V is said to be a generator system of V if every vector in V can be obtained as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Definition. A linearly independent generator system of V is called a basis of V.

Definition. The vector system $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V is said to be a generator system of V if every vector in V can be obtained as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Definition. A linearly independent generator system of V is called a basis of V.

Example. Let $V = \mathbb{R}^n$. The vector system

$$\begin{aligned} \mathbf{e}_1 &= [1, 0, 0, 0, \ldots, 0]^T \\ \mathbf{e}_2 &= [0, 1, 0, 0, \ldots, 0]^T \\ \mathbf{e}_3 &= [0, 0, 1, 0, \ldots, 0]^T \\ &\vdots \\ \mathbf{e}_n &= [0, 0, 0, 0, \ldots, 1]^T \end{aligned}$$

is a basis of the vector space $\mathbb{R}^n$. It is called the standard basis of $\mathbb{R}^n$.

Example. Let $V = \mathbb{R}^n$. The vector system

$$\begin{aligned} \mathbf{e}_1 &= [1, 0, 0, 0, \ldots, 0]^T \\ \mathbf{e}_2 &= [0, 1, 0, 0, \ldots, 0]^T \\ \mathbf{e}_3 &= [0, 0, 1, 0, \ldots, 0]^T \\ &\vdots \\ \mathbf{e}_n &= [0, 0, 0, 0, \ldots, 1]^T \end{aligned}$$

is a basis of the vector space $\mathbb{R}^n$. It is called the standard basis of $\mathbb{R}^n$.

Observe that in the vector e_i the i'th element is 1, and all other elements are 0.

Example. Let $V = \mathbb{R}^n$. The vector system

$$\begin{aligned}
\mathbf{e}_1 &= [1, 0, 0, 0, \ldots, 0]^T \\
\mathbf{e}_2 &= [0, 1, 0, 0, \ldots, 0]^T \\
\mathbf{e}_3 &= [0, 0, 1, 0, \ldots, 0]^T \\
&\vdots \\
\mathbf{e}_n &= [0, 0, 0, 0, \ldots, 1]^T
\end{aligned}$$

is a basis of the vector space $\mathbb{R}^n$. It is called the standard basis of $\mathbb{R}^n$.

Observe that in the vector $\mathbf{e}_i$ the i'th element is 1, and all other elements are 0.

Now we will prove why the vector system $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$. For concreteness, we will discuss the case $n = 3$, but the same argument works for any number n.

Theorem. $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$.

Theorem. $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$.

Proof (for $n = 3$). The point is that the linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ with coefficients $\alpha_1, \alpha_2, \alpha_3$ is the vector $[\alpha_1, \alpha_2, \alpha_3]^T$, as

$$\alpha_1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_3 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \qquad (*)$$

Theorem. $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$.

Proof (for $n = 3$). The point is that the linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ with coefficients $\alpha_1, \alpha_2, \alpha_3$ is the vector $[\alpha_1, \alpha_2, \alpha_3]^T$, as

$$\alpha_1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_3 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \qquad (*)$$

1. The vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are linearly independent, because the linear combination in $(*)$ is equal to the $[0, 0, 0]^T$ zero vector if and only if the coefficients $\alpha_1, \alpha_2, \alpha_3$ are all 0, because

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \alpha_1 = 0,\ \alpha_2 = 0,\ \alpha_3 = 0.$$

Theorem. $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$.

Proof (for $n = 3$). The point is that the linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ with coefficients $\alpha_1, \alpha_2, \alpha_3$ is the vector $[\alpha_1, \alpha_2, \alpha_3]^T$, as

$$\alpha_1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_3 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \qquad (*)$$

1. The vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are linearly independent, because the linear combination in $(*)$ is equal to the $[0, 0, 0]^T$ zero vector if and only if the coefficients $\alpha_1, \alpha_2, \alpha_3$ are all 0, because

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \alpha_1 = 0,\ \alpha_2 = 0,\ \alpha_3 = 0.$$

2. The vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ form a generator system, since by $(*)$, ANY vector $[a_1, a_2, a_3]^T$ in $\mathbb{R}^3$ is a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

$$[a_1, a_2, a_3]^T = a_1 \cdot \mathbf{e}_1 + a_2 \cdot \mathbf{e}_2 + a_3 \cdot \mathbf{e}_3. \qquad \square$$

$\vdots$

2. The vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ form a generator system, since by $(*)$, ANY vector $[a_1, a_2, a_3]^T$ in $\mathbb{R}^3$ is a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

$$[a_1, a_2, a_3]^T = a_1 \cdot \mathbf{e}_1 + a_2 \cdot \mathbf{e}_2 + a_3 \cdot \mathbf{e}_3. \qquad \square$$

An illustration of **2.**:

$$\begin{bmatrix} 3 \\ 8 \\ -2 \end{bmatrix} = 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 8 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + (-2) \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

and every vector in $\mathbb{R}^3$ can be obtained analogously as a linear combination of the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

Definition. A vector space is finitely generated if it has a finite set of generators.

Definition. A vector space is finitely generated if it has a finite set of generators.

Theorem + Definition. In a finitely generated vector space V any two bases have the same number of elements. This common number of elements is called the dimension of V, and it is denoted by $\dim(V)$.

Definition. A vector space is finitely generated if it has a finite set of generators.

Theorem + Definition. In a finitely generated vector space V any two bases have the same number of elements. This common number of elements is called the dimension of V, and it is denoted by $\dim(V)$.

Example. The dimension of the vector space $\mathbb{R}^n$ is n, because the standard basis has n elements.

Definition. A vector space is finitely generated if it has a finite set of generators.

Theorem + Definition. In a finitely generated vector space V any two bases have the same number of elements. This common number of elements is called the dimension of V, and it is denoted by $\dim(V)$.

Example. The dimension of the vector space $\mathbb{R}^n$ is n, because the standard basis has n elements.

Theorem + Definition. If $\mathbf{f}_1, \ldots, \mathbf{f}_n$ is a basis in V, then for each vector $\mathbf{v} \in V$ there are uniquely determined coefficients $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that $\mathbf{v} = \lambda_1 \cdot \mathbf{f}_1 + \cdots + \lambda_n \cdot \mathbf{f}_n$.
The vector $[\lambda_1, \ldots, \lambda_n]^T \in \mathbb{R}^n$ is called the coordinate vector of $\mathbf{v}$ with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$.

In the vector space $\mathbb{R}^n$, the coordinate vector of $\mathbf{v}$ with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ can be determined by solving the linear system with augmented matrix

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{f}_1} & \boxed{\mathbf{f}_2} & \boxed{\mathbf{f}_3} & \ldots & \boxed{\mathbf{f}_k} & \boxed{\mathbf{v}} \end{array}\right],$$

using the Theoreom on slide 6/14 .

In the vector space $\mathbb{R}^n$, the coordinate vector of $\mathbf{v}$ with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ can be determined by solving the linear system with augmented matrix

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{f}_1} & \boxed{\mathbf{f}_2} & \boxed{\mathbf{f}_3} & \ldots & \boxed{\mathbf{f}_k} & \boxed{\mathbf{v}} \end{array}\right],$$

using the Theoreom on slide 6/14 .

Example. The coordinate vector of $[1,2,1]^T$ with respect to the basis $[-1,2,1]^T$, $[1,2,3]^T$, $[-1,1,1]^T$ is $[3/2,1/2,-2]^T$, because the solution of the linear system

$$\left[\begin{array}{ccc|c} -1 & 1 & -1 & 1 \\ 2 & 2 & 1 & 2 \\ 1 & 3 & 1 & 1 \end{array}\right].$$

is $(3/2, 1/2, -2)$. $\square$

In the vector space $\mathbb{R}^n$, the coordinate vector of $\mathbf{v}$ with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ can be determined by solving the linear system with augmented matrix

$$\left[\;\boxed{\mathbf{f}_1}\;\boxed{\mathbf{f}_2}\;\boxed{\mathbf{f}_3}\;\ldots\;\boxed{\mathbf{f}_k}\;\middle|\;\boxed{\mathbf{v}}\;\right],$$

using the Theoreom on slide 6/14 .

Example. The coordinate vector of $[1, 2, 1]^T$ with respect to the basis $[-1, 2, 1]^T$, $[1, 2, 3]^T$, $[-1, 1, 1]^T$ is $[3/2, 1/2, -2]^T$, because ...

Check of solution.

$$(3/2) \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} + (1/2) \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + (-2) \cdot \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

In the vector space $\mathbb{R}^n$, the coordinate vector of $\mathbf{v}$ with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ can be determined by solving the linear system with augmented matrix

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{f}_1} & \boxed{\mathbf{f}_2} & \boxed{\mathbf{f}_3} & \ldots & \boxed{\mathbf{f}_k} & \boxed{\mathbf{v}} \end{array}\right],$$

using the Theoreom on slide 6/14 .

Example 2. The coordinate vector of a vector $\mathbf{v} \in \mathbb{R}^n$ with respect to the standard basis is the vector $\mathbf{v}$ itself. We have already seen this, when we proved that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is indeed a basis in $\mathbb{R}^n$ (check step **2** in that proof again):

$$\begin{bmatrix} 3 \\ 8 \\ -2 \end{bmatrix} = 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 8 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + (-2) \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

In the vector space $\mathbb{R}^n$, the coordinate vector of $\mathbf{v}$ with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ can be determined by solving the linear system with augmented matrix

$$\left[\begin{array}{ccccc|c} \boxed{\mathbf{f}_1} & \boxed{\mathbf{f}_2} & \boxed{\mathbf{f}_3} & \ldots & \boxed{\mathbf{f}_k} & \boxed{\mathbf{v}} \end{array}\right],$$

using the Theoreom on slide 6/14 .

Worked-out example 1. CLICK HERE

Worked-out example 2. CLICK HERE

Theorem. Let V be a finitely generated vector space. Then

(a) Every linearly independent vector system in V has at most $\dim(V)$ elements.

(b) Every generator system of V has at least $\dim(V)$ elements.

(c) If a linearly independent vector system in V has $\dim(V)$ elements, then it is a basis.

(d) If a generator system of V has $\dim(V)$ elements, then it is a basis.

Definition. The rank of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V is the dimension of the spanned subspace $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. It is denoted by $\mathrm{rank}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

Definition. The rank of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V is the dimension of the spanned subspace $\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. It is denoted by $\text{rank}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

Proposition. If $\text{rank}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = r$, then there exists r vectors among $\mathbf{v}_1, \ldots, \mathbf{v}_k$ that are linearly independent, but any $r + 1$ of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly dependent.

Definition. The rank of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in a vector space V is the dimension of the spanned subspace $\mathrm{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. It is denoted by $\mathrm{rank}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proposition. If $\mathrm{rank}(\mathbf{v}_1, \dots, \mathbf{v}_k) = r$, then there exists r vectors among $\mathbf{v}_1, \dots, \mathbf{v}_k$ that are linearly independent, but any $r + 1$ of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly dependent.

Theorem. In $\mathbb{R}^n$, the rank of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ can be calculated as follows:

1. Create the $k \times n$ matrix whose rows are the given vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$.
2. Transform this matrix into row-echelon form by Gaussian elimination.
3. Then the number of non-zero rows in the row-echelon form gives $\mathrm{rank}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Exercise. Determine the rank of the following vectors in $\mathbb{R}^4$:

$$[1, 3, -1, 2]^T, \quad [2, 8, 1, 3]^T, \quad [1, 1, -4, 7]^T, \quad [-3, -5, 9, 2]^T.$$

Exercise. Determine the rank of the following vectors in $\mathbb{R}^4$:

$$[1,3,-1,2]^T, \quad [2,8,1,3]^T, \quad [1,1,-4,7]^T, \quad [-3,-5,9,2]^T.$$

Solution. Write the given vectors into the rows of a matrix, and transform the matrix into row-echelon form.

$$\begin{bmatrix} 1^* & 3 & -1 & 2 \\ 2 & 8 & 1 & 3 \\ 1 & 1 & -4 & 7 \\ -3 & -5 & 9 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 2^* & 3 & -1 \\ 0 & -2 & -3 & 5 \\ 0 & 4 & 6 & 8 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 3 & -1 \\ 0 & 0 & 0 & 4^* \\ 0 & 0 & 0 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 3 & -1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the row-echelon form has 3 non-zero rows, the rank of the given vectors is 3. $\square$

Note. The previous theorem has an interesting consequence: All row-echelon forms of a matrix have the same number of non-zero rows.

Note. The previous theorem has an interesting consequence: All row-echelon forms of a matrix have the same number of non-zero rows.

We saw that the rank of vectors in $\mathbb{R}^n$ can be easily determined by Gaussian elimination. The good news is that the rank can be used to test whether the vectors are linearly independent, form a generator system, or form a basis. This is discussed in the next theorem . . .

Theorem. Let $\mathbf{v}_1, \dots \mathbf{v}_k$ be vectors in $\mathbb{R}^n$. (So k is the number of vectors; and n is the dimension of $\mathbb{R}^n$, that is, the number of components of a given vector.) Let r denote the rank of vectors $\mathbf{v}_1, \dots \mathbf{v}_k$. Then the followings hold.

(a) The vectors $\mathbf{v}_1, \dots \mathbf{v}_k$ are linearly independent if and only if $r = k$.

(b) The vectors $\mathbf{v}_1, \dots \mathbf{v}_k$ form a generator system of $\mathbb{R}^n$ if and only if $r = n$.

(c) The vectors $\mathbf{v}_1, \dots \mathbf{v}_k$ form a basis of $\mathbb{R}^n$ if and only if $r = k = n$.

Worked-out example 1. CLICK HERE

Worked-out example 2. CLICK HERE

Definition. Let A be a matrix. A submatrix of A is a matrix formed by selecting some rows and some columns of A, and forming a new matrix by using those entries, in the same relative positions, that appear in both the rows and columns of those selected.

Example.

$$\begin{bmatrix} 2 & 3 & -1 & 6 & 1 & 8 \\ 4 & -4 & 2 & 6 & 3 & 0 \\ 0 & 7 & 5 & 2 & -1 & 6 \\ 6 & 2 & 1 & 0 & 7 & 3 \\ 9 & -5 & 0 & 8 & -8 & 5 \\ 4 & 2 & -7 & 3 & 0 & 1 \end{bmatrix} \supseteq \begin{bmatrix} -4 & 2 & 6 & 0 \\ 2 & 1 & 0 & 3 \\ -5 & 0 & 8 & 5 \end{bmatrix}.$$

Definition + Theorem. Let A be an $m \times n$ matrix.

- The row rank $\text{rank}_r(A)$ of A is the rank of the m row vectors of A (as vectors in $\mathbb{R}^n$).
- The column rank $\text{rank}_c(A)$ of A is the rank of the n column vectors of A (as vectors in $\mathbb{R}^m$).
- The determinant rank $\text{rank}_d(A)$ of A is the largest number r for which A contains a submatrix of size $r \times r$ with non-zero determinant.

Then $\boxed{\text{rank}_r(A) = \text{rank}_c(A) = \text{rank}_d(A)}$. This common number is called the rank of A, and it is denoted by $\text{rank}(A)$.

Definition + Theorem. Let A be an $m \times n$ matrix.

- The row rank $\text{rank}_r(A)$ of A is the rank of the m row vectors of A (as vectors in $\mathbb{R}^n$).
- The column rank $\text{rank}_c(A)$ of A is the rank of the n column vectors of A (as vectors in $\mathbb{R}^m$).
- The determinant rank $\text{rank}_d(A)$ of A is the largest number r for which A contains a submatrix of size $r \times r$ with non-zero determinant.

Then $\boxed{\text{rank}_r(A) = \text{rank}_c(A) = \text{rank}_d(A)}$. This common number is called the rank of A, and it is denoted by $\text{rank}(A)$.

The theorem on determining rank of vectors implies the following:

Theorem. The rank of a matrix is equal to the number of non-zero rows in its row-echelon form.

Theorem. The rank of a matrix is equal to the number of non-zero rows in its row-echelon form.

Example. Find the rank of $A = \begin{bmatrix} 1 & 2 & -4 & 1 \\ 2 & 3 & 1 & 5 \\ 0 & -1 & 9 & 3 \end{bmatrix}$.

Theorem. The rank of a matrix is equal to the number of non-zero rows in its row-echelon form.

Example. Find the rank of $A = \begin{bmatrix} 1 & 2 & -4 & 1 \\ 2 & 3 & 1 & 5 \\ 0 & -1 & 9 & 3 \end{bmatrix}$.

Solution. The Gaussian elimination gives

$$\begin{bmatrix} 1^* & 2 & -4 & 1 \\ 2 & 3 & 1 & 5 \\ 0 & -1 & 9 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -4 & 1 \\ 0 & -1^* & 9 & 3 \\ 0 & -1 & 9 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -4 & 1 \\ 0 & -1 & 9 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The number of non-zero rows in the row-echelon form is 2, hence the rank of A is 2. $\square$