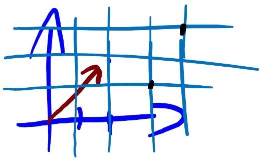
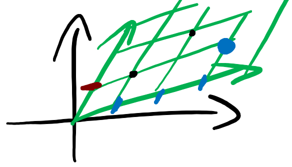




$\rightsquigarrow$ 3 "λ_1" and $1 = \lambda_2$

Proposition & Definition 3.23. Coefficients with respect to a basis

*Let $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ be a basis of a subspace $V \subset \mathbb{R}^n$. Each $\mathbf{x} \in V$ can be written as a linear combination $\lambda_1\mathbf{v}_1 + \dots + \lambda_k\mathbf{v}_k$ where the coefficients $\lambda_1, \dots, \lambda_k$ are **unique**. They are called the coordinates of $\mathbf{x}$ with respect to $\mathcal{B}$.*

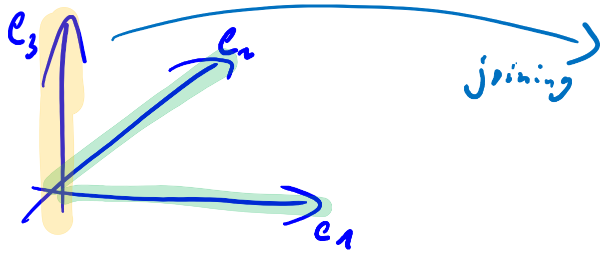
Example: $V = \mathbb{R}^n$, $\mathcal{B} = (e_1, \dots, e_n)$, $x \in V$ "=" $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

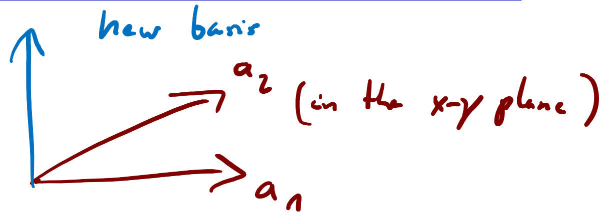
$x = x_1 e_1 + \dots + x_n e_n$

Coordinates of x w.r.t. the standard basis of $\mathbb{R}^n$.

Theorem 3.24. Steinitz's theorem → building new bases

Consider a basis $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ of a subspace $V \subset \mathbb{R}^n$ and a linearly independent set of vectors $\mathcal{A} = (\mathbf{a}_1, \dots, \mathbf{a}_\ell) \subset V$. Then we can extend $\mathcal{A}$ to a basis of V by adding $k - \ell$ elements of $\mathcal{B}$.





Sketch of the proof. Pack $\mathcal{B}$ and $\mathcal{A}$ together to a linearly dependent set, and remove vectors (starting with elements of $\mathcal{B}$) until it is linearly independent. One has to show now, that the resulting set has again k elements, and that $\mathcal{A}$ remains untouched. □

$\rightsquigarrow$ mathematical induction

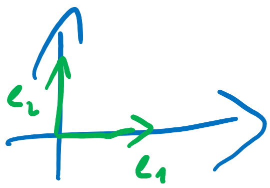
Now, we can record that all bases of V have the same number of elements.

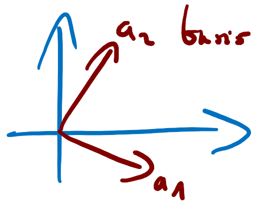
Corollary 3.25.

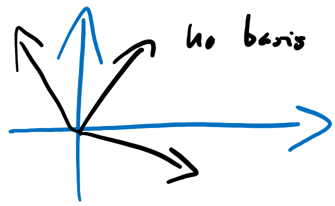
Let V be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ be a basis of V. Then:

(a) Each family $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ consisting of vectors from V where $m > k$ is linearly dependent.

(b) Each basis of V has exactly k elements.







So we can define:

$\mathrm{span}(\emptyset) = \{o\}$ zero elements $\qquad \emptyset$ is a basis of $\{o\}$

Definition 3.26. Dimension of a linear subspace

Let V be a subspace of $\mathbb{R}^n$, and let $\mathcal{B}$ be a chosen basis of V. The number of elements in $\mathcal{B}$ is well-defined and called the dimension *of V, written as $\dim(V)$. As a special case, we set $\dim(\{\mathbf{o}\}) = 0$.*

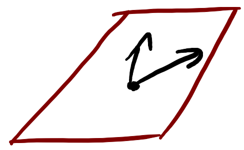
The unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in $\mathbb{R}^n$ form a basis. The linear independency can be seen by:

$$x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = \mathbf{o} \quad \Rightarrow \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{o} \quad \Rightarrow \quad x_1 = \ldots = x_n = 0.$$

We obtain, as expected:

$$\dim(\mathbb{R}^n) = n.$$

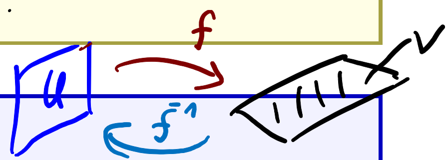
n-dim.



Rule of thumb:

The dimension of a vector space V says how many independent degrees of freedom are needed to build linear combinations of all vectors in V.

Theorem 3.27. Important

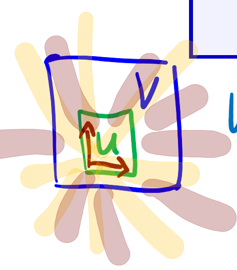


Let U and V be two linear subspaces of $\mathbb{R}^n$.

(i) *One has $\dim(U) = \dim(V)$ if and only if there exists a linear bijective map between U and V.*

$\Leftrightarrow$ [dim $\leftrightsquigarrow$ bij. linear map]

(ii) *If $U \subset V$ and $\dim(U) = \dim(V)$, then $U = V$.*



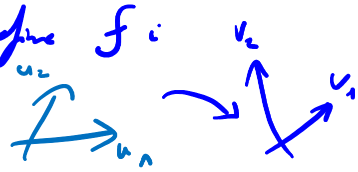
$U = V$

Proof: (i) ($\Rightarrow$) Assume $\dim(U) = \dim(V)$. This means that

$\mathcal{B} = (u_1, \ldots, u_k)$ basis of U

$\mathcal{C} = (v_1, \ldots, v_k)$ basis of V.

Now define f:



$f : U \to V, \quad f(u_i) := v_i$ for all $i = 1, \ldots, k$.

$x \in U: \quad f(x) = f(\underbrace{\lambda_1 u_1 + \cdots + \lambda_k u_k}_{\text{unique linear combination}})$

$= f(\lambda_1 u_1) + \cdots + f(\lambda_k u_k) = \underbrace{\lambda_1 f(u_1) + \cdots + \lambda_k f(u_k)}_{=: f(x)}$ with $f(u_1) = v_1$, $f(u_k) = v_k$

f is linear ✓, $\quad f^{-1} : V \to U, \quad f^{-1}(v_i) = u_i$ (same)

Now we get $(f^{-1} \circ f)(x) = x, \quad (f \circ f^{-1})(y) = y$.

$\rightsquigarrow f$ is bijective ✓

($\Leftarrow$) Assume there is an $f : U \to V$ bij. + lin., (inj. and sur.)

$\mathcal{B} = (u_1, \ldots, u_k)$ basis of U $\Rightarrow$ $(f(u_1), \ldots, f(u_k))$ basis in V ??

Exercise $\qquad \rightsquigarrow$ Exercise

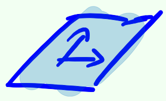
lin. ind. (since injective) $\qquad \mathrm{span}(f(u_1), \ldots, f(u_k)) = V$ (since surjective)

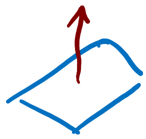
$\Rightarrow (f(u_1), \dots, f(u_k))$ is a basis of V $\Rightarrow \underbrace{\dim(U)}_{k} = \dim(V) = k$

Example 3.28.

The following subspaces of $\mathbb{R}^n$ are very important:

- *The trivial subspace* $\{0\}$ *with* $\dim(\{0\}) = 0$
- *Lines L (through the origin):* $\dim(L) = 1$
- *Planes P (through the origin):* $\dim(P) = 2$
- *Hyperplanes H (through the origin):* $\dim(H) = n - 1$ $\to$ having inner product $\to$ normal vector





The dimension of an affine subspace $W = \mathbf{u}_0 + U$ (where U is a linear subspace) is usually set to the dimension of U.

Corollary 3.29.

A family consisting of more than n vectors in $\mathbb{R}^n$ is always linearly dependent.

Proof. Use Corollary 3.25. $\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}2\\4\\5\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}$ basis of $\mathbb{R}^3$? $\square$

3.8 Identity and inverses

For each $n \in \mathbb{N}$, we define the *identity matrix* $\mathbb{1}_n$ by

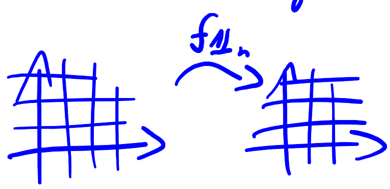
I, E, Id, id

$$\mathbb{1}_n := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & \vdots \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

neutral element w.r.t. matrix multipl.

$\mathbb{1}_n \cdot B = B$

$f_{\mathbb{1}_n}$



$f_{\mathbb{1}_n} : \mathbb{R}^n \to \mathbb{R}^n$, $(x \mapsto \mathbb{1}_n \cdot x = x)$, $id = f_{\mathbb{1}_n}$

$x \mapsto x$

Inverses w.r.t. the matrix multiplication: $A \cdot \tilde{A} = \mathbb{1}_n$, $\tilde{A} \cdot A = \mathbb{1}_n$

$\to \tilde{A} =: A^{-1}$ inverse of A.

Definition 3.30. Invertible Matrix, A^{-1}

We call a square matrix $A \in \mathbb{R}^{n\times n}$ invertible or nonsingular if the corresponding linear map $f_A : \mathbb{R}^n \to \mathbb{R}^n$ is bijective. Otherwise, we call A singular. A matrix $\tilde{A}$ with $f_{\tilde{A}} = (f_A)^{-1}$ is called the inverse of A and is usually denoted by A^{-1}. We have

$$f_{A^{-1}} \circ f_A = id \quad \text{and} \quad f_A \circ f_{A^{-1}} = id\,,$$

which means $f_{A^{-1}} = (f_A)^{-1}$.
For the matrices, this means:

$$A^{-1}(A\mathbf{x}) = \mathbf{x} \quad \text{and} \quad A(A^{-1}\mathbf{x}) = \mathbf{x} \quad \text{for all} \quad \mathbf{x} \in \mathbb{R}^n.$$

In short: $A^{-1}A = \mathbb{1}$ and $AA^{-1} = \mathbb{1}$.

$\tilde{A}^{-1}$ inverse of A.

solve $\quad \overset{=\mathbb{1}}{(A^{-1}(A)}x) = A^{-1}b$

If A is invertible, the linear system $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Theorem 3.31.

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

$$f_A \text{ injective} \Leftrightarrow f_A \text{ surjective}$$

Hence, if one of these cases holds, then f_A is already bijective, i.e., invertible.

Proof. This is a classical *dimension argument*. $\to (e_1, \dots, e_n) \Rightarrow (f(e_1), \dots, f(e_n))$ $\to$ basis of $\mathbb{R}^n$

f_A

($\Rightarrow$): If f_A is injective, then n linearly independent vectors form a basis for $\mathbb{R}^n$. This means that f_A is surjective.

($\Leftarrow$): If f_A is surjective, then each n vectors that span the $\mathbb{R}^n$ form a basis for $\mathbb{R}^n$, so f_A is injective. ☐

For each $y \in \mathbb{R}^n$, you find an $x \in \mathbb{R}^n$ with $f(x) = y$.

$y = f(x) = x_1 f(e_1) + \dots + x_n f(e_n)$

basis of $\mathbb{R}^n$ $\to$ lin. independent $\to$ f_A inj.

For two invertible matrices A and B we have the formula:

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

f_B $\quad f_A$

$f_A \circ f_B = f_{AB}$

f $\quad f^{-1}$

Remark:

If $f : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map that is bijective, then $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is also a linear map.

$f^{-1}(\lambda y) = f^{-1}(\lambda f(x)) \overset{f \text{ lin.}}{=} f^{-1}(f(\lambda x)) = \lambda x = \lambda f^{-1}(y)$ ✓

[There is exactly one x with $f(x) = y$]

3.9 Transposition

→ changing the roles of columns and rows

We already know transposition of column vectors:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}^T = (a_1 \ \dots \ a_n)$$

and similarly, we can define:

$$(a_1 \ \dots \ a_n)^T = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

Then we have the simple formula $(\mathbf{a}^T)^T = \mathbf{a}$.

For a matrix, we can do the same:

Definition 3.32. Transpose

For a matrix $A \in \mathbb{R}^{m \times n}$, we define a matrix $A^T \in \mathbb{R}^{n \times m}$ and call it the transpose of A. The i^{th} column of A becomes the i^{th} row of A^T and the j^{th} row of A becomes the j^{th} column of A^T:

$$\text{For} \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \text{we define} \quad A^T := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

Example 3.33. (a)

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & 3 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 4} \qquad \Rightarrow \qquad A^T = \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 3 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 2}.$$

(b)

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2} \qquad \Rightarrow \qquad A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

(c)

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2} \qquad \Rightarrow \qquad A^T = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

(d)

$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^{3\times 1} \qquad \Rightarrow \qquad A^T = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \in \mathbb{R}^{1\times 3}.$$

(e)

$$A = \begin{pmatrix} 4 & 5 & 6 & 7 \end{pmatrix} \in \mathbb{R}^{1\times 4} \qquad \Rightarrow \qquad A^T = \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} \in \mathbb{R}^{4\times 1}.$$

Since we have exchanged the roles of rows and columns, the order of multiplication changes, too:

$$(A\mathbf{x})^T = \mathbf{x}^T A^T \qquad \mathbf{x}^T A = (A^T\mathbf{x})^T.$$

Just as with matrix-vector multiplication, transposition reverses the order of matrix-matrix multiplication:

$$(AB)^T = B^T A^T.$$

In particular, if A is invertible, then

$$\underline{\mathbb{1}} = \mathbb{1}^T = (A^{-1}A)^T = \underline{A^T(A^{-1})^T} \quad \Rightarrow \quad A^T \text{ is invertible and } (A^T)^{-1} = (A^{-1})^T.$$

Example 3.34. We find

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 0 & -2 & 2 \\ 4 & -1 & 2 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 11 & -4 & 9 \\ 26 & -13 & 24 \end{pmatrix}$$

and

$$\begin{pmatrix} 0 & 4 & 1 \\ -2 & -1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 11 & 26 \\ -4 & -13 \\ 9 & 24 \end{pmatrix}.$$

Proposition 3.35. Some rules for transposition

(a) For all $A, B \in \mathbb{R}^{m\times n}$ we have: $(A+B)^T = A^T + B^T$.

(b) For all $A \in \mathbb{R}^{m\times n}$ and for all $\lambda \in \mathbb{R}$, we have: $(\lambda \cdot A)^T = \lambda \cdot A^T$.

"linearity"

(c) For all $A \in \mathbb{R}^{m\times n}$ we have: $(A^T)^T = A$.

(d) For all $A \in \mathbb{R}^{m\times n}$ and for all $B \in \mathbb{R}^{n\times r}$ we have: $(A \cdot B)^T = B^T \cdot A^T$.

(e) If $A \in \mathbb{R}^{n\times n}$ is invertible, then A^T is also invertible and we get $(A^T)^{-1} = (A^{-1})^T$.

(f) For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ we have: $\mathbf{u}^T \cdot \mathbf{v} = \mathbf{v}^T \cdot \mathbf{u} = \langle \mathbf{u}, \mathbf{v} \rangle$.

Proof. If we denote the entries of a matrix A by A_{ij}. Then we have

$$(A^T)_{ij} = A_{ji} \quad \text{for all } i, j$$

and from this one can prove all properties. For example for showing (d), we see

$$(B^T \cdot A^T)_{ij} = \sum_k (B^T)_{ik}(A^T)_{kj} = \sum_k A_{jk}B_{ki} = (A \cdot B)_{ji} = ((A \cdot B)^T)_{ij}$$

for all i, j. □

Proposition 3.36. What has A^T to do with the inner product?

For $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$, we have for the standard inner product:

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^T\mathbf{y} \rangle.$$

Proof. We already know that for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T\mathbf{v}$. Hence, we conclude that for $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$, the following holds $\langle A\mathbf{x}, \mathbf{y} \rangle = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \langle \mathbf{x}, A^T\mathbf{y} \rangle$. □

Moreover, A^T is the only matrix in $B \in \mathbb{R}^{n\times m}$ that satisfies the equation $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, B\mathbf{y} \rangle$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. Therefore, some people use this as the definition for A^T.

Definition 3.37. Symmetric and skew-symmetric matrices

One typical notation for quadratic matrices: $\in \mathbb{R}^{n\times n}$

- *If $A^T = A$, then A is called* symmetric.
- *If $A^T = -A$, then A is called* skew-symmetric.

Example 3.38. (a)

$$A = \begin{pmatrix} 1 & 3 & -4 \\ 3 & 0 & 5 \\ -4 & 5 & 3 \end{pmatrix} \quad \text{is symmetric since} \quad A^T = \begin{pmatrix} 1 & 3 & -4 \\ 3 & 0 & 5 \\ -4 & 5 & 3 \end{pmatrix} = A.$$

(b)

$$A = \begin{pmatrix} 0 & 3 & 4 \\ -3 & 0 & -5 \\ -4 & 5 & 0 \end{pmatrix} \quad \text{is skew-symmetric since} \quad A^T = \begin{pmatrix} 0 & -3 & -4 \\ 3 & 0 & 5 \\ 4 & -5 & 0 \end{pmatrix} = -A.$$

By definition, all skew-symmetric matrices have only zeros on the diagonal.

3.10 The kernel, range and rank of a matrix

$f_A: \mathbb{R}^n \to \mathbb{R}^m$

Definition 3.39. Range and kernel of matrices

Let $A \in \mathbb{R}^{m \times n}$. The set

$$\mathrm{Ran}(A) := \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m$$

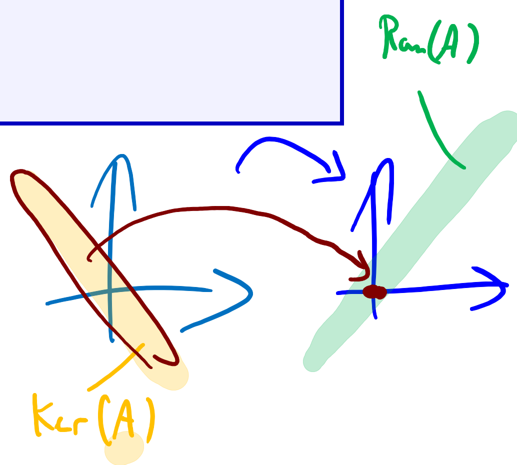
is called the range or image of the matrix A.
The set

$$\mathrm{Ker}(A) := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{o}\} \subset \mathbb{R}^n$$

is called the kernel or nullspace of the matrix A.

$f_A: \mathbb{R}^n \to \mathbb{R}^m$, $\mathrm{Ran}(f_A) =: \mathrm{Ran}(A)$

$x \mapsto Ax$, $f_A^{-1}(\{0\}) =: \mathrm{Ker}(A)$



Both are linear subspaces!

Since our ultimate goal is to understand linear systems of the formal

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

we would like to know more about $\mathrm{Ran}(A)$ (because it tell us, for which $\mathbf{b}$ our system has a solution) and $\mathrm{Ker}(A)$ (because it tells us about the uniqueness of solutions).

Definition 3.40. Rank of a matrix

Let $A \in \mathbb{R}^{m \times n}$. The number

$$\mathrm{rank}(A) := \dim(\mathrm{Ran}(A)) = \dim(\mathrm{Span}\,(\mathbf{a}_1, \ldots, \mathbf{a}_n)).$$

is called the rank of the matrix A.

We obviously have:

$$\mathrm{rank}(A) \leq \min\{m, n\}$$

A is said to have full rank, if $\mathrm{rank}(A) = \min\{m, n\}$.