Univ. of Washington

Lecture : Linear Dependence, Span, Range, and Nullspace

Winter 2021

Lecturer: Dan Calderone

Linear combinations and linear dependence

For the following define

$$x \in \mathbb{R}^n, \qquad y \in \mathbb{R}^m, \qquad A = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix}$$

- A vector y is **linear dependent** on the columns of a matrix $A \in \mathbb{R}^{m\times n}$

$$\text{if } \exists\, x \in \mathbb{R}^n \quad \text{s. t.} \quad y = Ax \qquad \text{or (equivalently) } y = \sum_i A_i x_i$$

- A set of vectors (the columns of A) is **linearly dependent** if at least one vector is dependent on the others.

$$A_i = \sum_{j\neq i} A_j x'_j$$

for some i and $\{x_j\}_{j\neq i}$. A useful characterization is the columns of A are **linearly dependent**

$$\text{if } \exists\, x \in \mathbb{R}^n, x \neq 0 \quad \text{s. t.} \quad Ax = 0$$

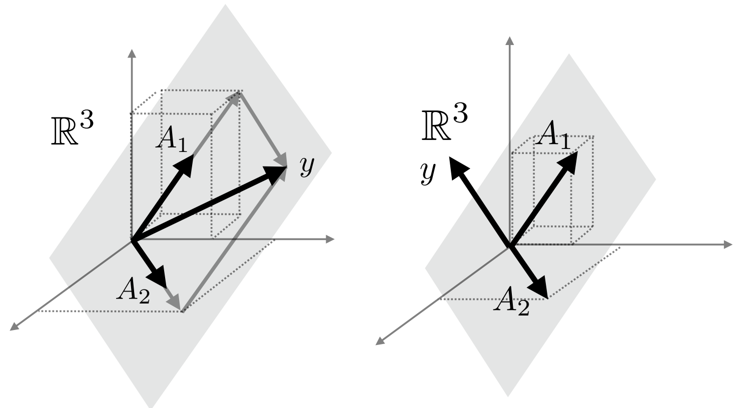
Indeed for some i

$$A_i x_i = -\sum_{j\neq i} A_j x_j, \qquad \Rightarrow \qquad A_i = \sum_{j\neq i} A_j \tfrac{-x_j}{x_i} = \sum_{j\neq i} A_j x'_j$$

- A vector y is **linearly independent** on the columns of A, if it is not linearly dependent on them, i.e. there does not exist $x \in \mathbb{R}^n$ such that $y = Ax$.
- A set of vectors (cols of A) is **linear independent** if none of the columns are linearly dependent on the others. A useful characterization is

$$Ax = 0 \qquad \Rightarrow \qquad .x = 0$$

- All linear combinations of a set of vectors is the **span** of those vectors.



y, **linearly dependent** on A_1, A_2.

y, **linearly independent** from A_1, A_2.

Rank

- **Column rank**: # number of linearly independent columns.
- **Row rank**: # number of linearly independent columns.
- **Rank**: **Column rank=Row rank=Rank**

The column-rank of a matrix is the dimension of the column space of A, $\mathcal{R}(A)$. The row-rank of a matrix is the dimension of the row space, $\mathcal{R}(A^T)$. The column rank and row rank are always equal and are simply called the rank of A, denoted $\text{rk}(A)$.

Column rank = row rank

Proof: Let the column rank be denoted by k and the row rank be denoted by r.

If $A \in \mathbb{R}^{m \times n}$ has column rank k then there exists $C \in \mathbb{R}^{m \times k}$ with linearly independent columns such that

$$A = CV$$

where the columns of $V \in \mathbb{R}^{k \times n}$ are the coordinates of the columns of A with respect to the (basis) columns of C. Thinking of the rows of C as coefficients of linear combinations of the rows of V and realizing that V has k rows, we have that the dimension of $\mathcal{R}(A^T)$ is at most k. Thus we have that $r \leq k$.

If $A \in \mathbb{R}^{m \times n}$ has row rank r then there exists $R \in \mathbb{R}^{r \times n}$ with linearly independent rows such that

$$A = WR$$

where the rows of $W \in \mathbb{R}^{m \times r}$ are the coordinates of the rows of A with respect to the (basis) rows of R. Thinking of the columns or R as coefficients of linear combinations of the columns of W and realizing that W has r columns, we have that the dimension of $\mathcal{R}(A)$ is at most r. Thus we have that $k \leq r$.

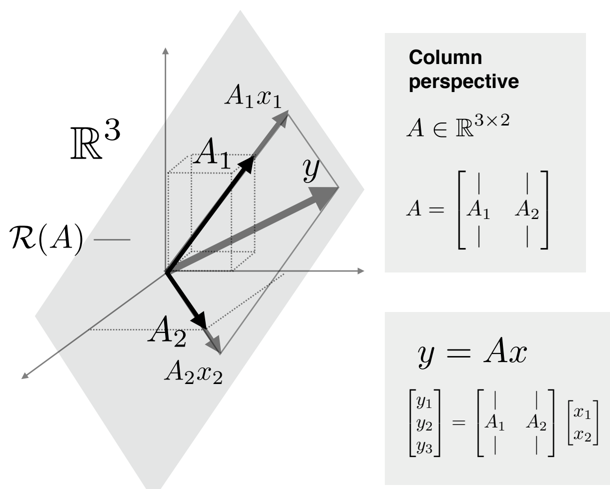
Combining the two inequalities, $r \leq k$ and $k \leq r$ gives that $k = r$, ie. the column and row ranks are equal.

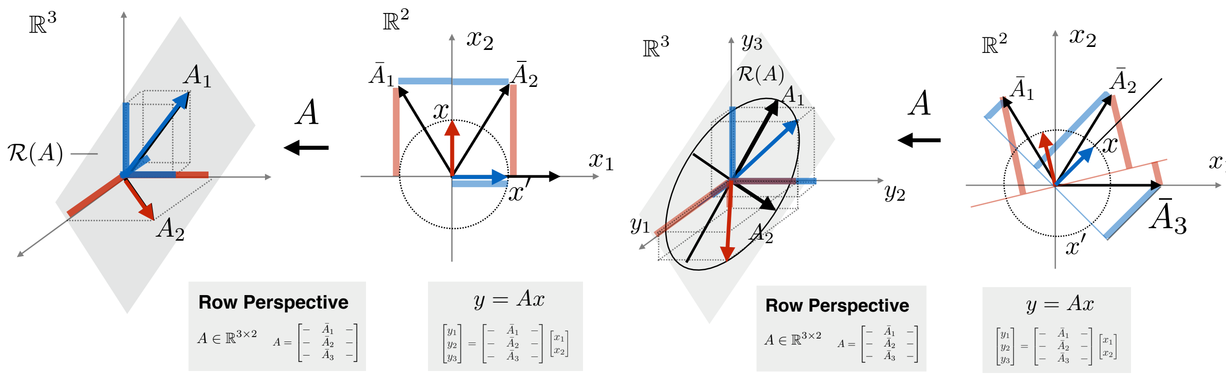
Range and Nullspace

Range Space

A matrix $A \in \mathbb{R}^{m \times n}$ represents a linear map from $\mathbb{R}^n$ which is called the *domain* to $\mathbb{R}^m$ which is called the *co-domain*. The span of the columns of A is a subspace of the co-domain called the *range of* A sometimes denoted $\mathcal{R}(A)$. Note this is equivalent to the definition.

$$\mathcal{R}(A) = \{y \in \mathbb{R}^m \mid y = Ax, \text{ for some } x \in \mathbb{R}^n\} \tag{1}$$





Null Space

The *nullspace of* A, sometimes denoted $\mathcal{N}(A)$, is the subspace of the domain such that

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\} \tag{2}$$

- **Orthogonal to the rows of** A

  $\mathcal{N}(A) \perp \mathcal{R}(A^T)$

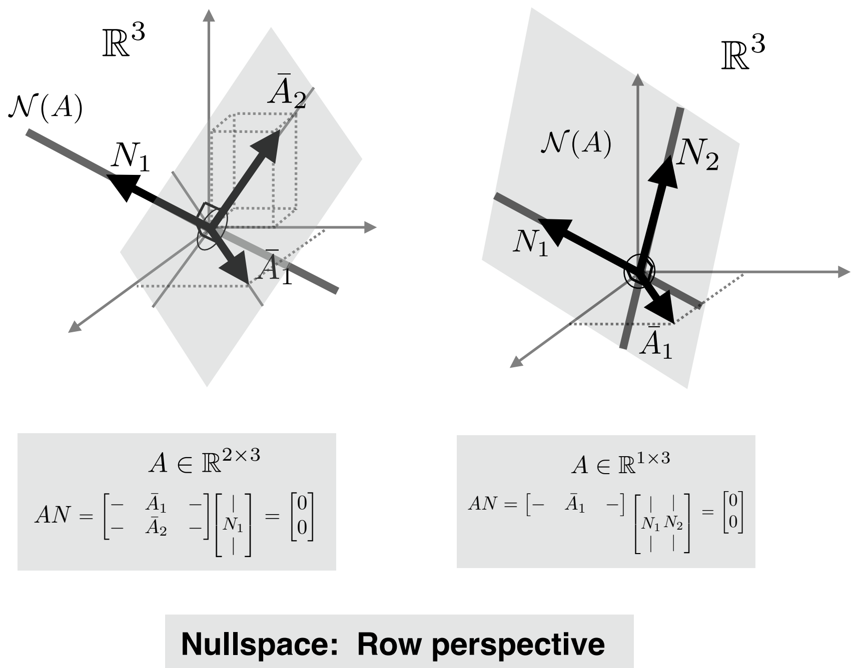
  $\mathcal{R}(A^T)$ is the span of the rows of A. Decompose $A \in \mathbb{R}^{m\times n}$ into rows as

$$A = \begin{bmatrix} - & \bar{a}_1^T & - \\ & \vdots & \\ - & \bar{a}_1^T & - \end{bmatrix}$$

  If $Ax = 0$, then

$$Ax = \begin{bmatrix} - & \bar{a}_1^T & - \\ & \vdots & \\ - & \bar{a}_m^T & - \end{bmatrix} x = \begin{bmatrix} \bar{a}_1^T x \\ \vdots \\ \bar{a}_m^T x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

  ie. if $x \in \mathcal{N}(A)$ then x is orthogonal to each row of A. $\mathcal{N}(A) \perp \mathcal{R}(A^T)$ and similarly $\mathcal{N}(A^T) \perp \mathcal{R}(A)$.



- **Basis Construction:**

  Suppose A has column rank k, ie. $\mathcal{R}(A)$ has dimension k. Assume (without loss of generality) that the first k columns of A are linearly independent (and thus span the range). (A similar construction can be done with any k linearly independent columns of A.) Let

$$A = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_k \\ | & & | \end{bmatrix}$$

  where $B \in \mathbb{R}^{n\times k}$ is (the first) k linearly independent columns of A. A can then be written as

$$A = \begin{bmatrix} B & BD \end{bmatrix}$$

  where the columns of $D \in \mathbb{R}^{n\times(n-k)}$ are the coordinates of the remaining $n-k$ columns of A with respect to the columns of B.

$$\begin{bmatrix} | & & | \\ A_{k+1} & \cdots & A_n \\ | & & | \end{bmatrix} = BD = \underbrace{\begin{bmatrix} | & & | \\ A_1 & \cdots & A_k \\ | & & | \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} | & & | \\ D_{k+1} & \cdots & D_n \\ | & & | \end{bmatrix}}_{D}$$

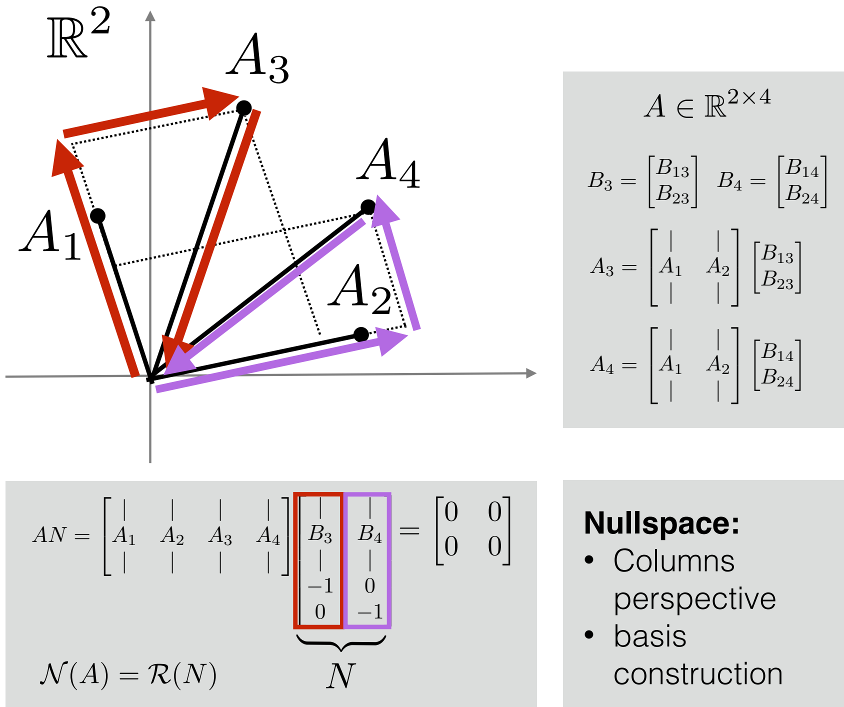
  Let $N \in \mathbb{R}^{n\times(n-k)}$ be given by

$$N = \begin{bmatrix} -D \\ I \end{bmatrix}$$

Note that

$$AN = \begin{bmatrix} B & BD \end{bmatrix} \begin{bmatrix} -D \\ I \end{bmatrix} = 0$$

We have also that the columns of N form a basis for the nullspace of A



Proof:

– **Span $\mathcal{N}(A)$:**
Suppose $Ax = 0$

$$Ax = \begin{bmatrix} B & BD \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{aligned} Bx_1 &= -BDx_2 \\ \big((B^TB)^{-1}B^T\big)Bx_1 &= -\big((B^TB)^{-1}B^T\big)BDx_2 \\ x_1 &= -Dx_2 \end{aligned}$$

Note that the linear independence of the columns of B guarantees that B^TB is invertible. Plugging in then gives

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -Dx_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} -D \\ I \end{bmatrix} x_2 = Nx_2$$

showing that x is a linear combination of the columns of N.

– **Linear independence:**
Suppose $Nx_2 = 0$

$$Nx_2 = \begin{bmatrix} -D \\ I \end{bmatrix} x_2 = \begin{bmatrix} -Dx_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

It follows that $x_2 = 0$.

Rank-Nullity Theorem

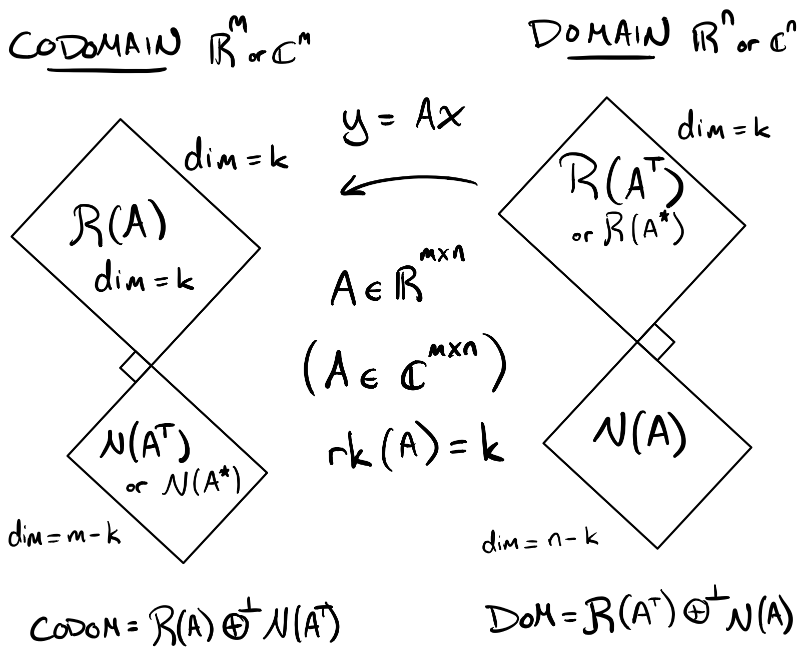
The explicit construction of a basis for the nullspace given above shows that if a matrix has (column) rank k then the nullspace has dimension $n - k$. The dimension of the nullspace is known as the nullity and we have the rank-nullity theorem

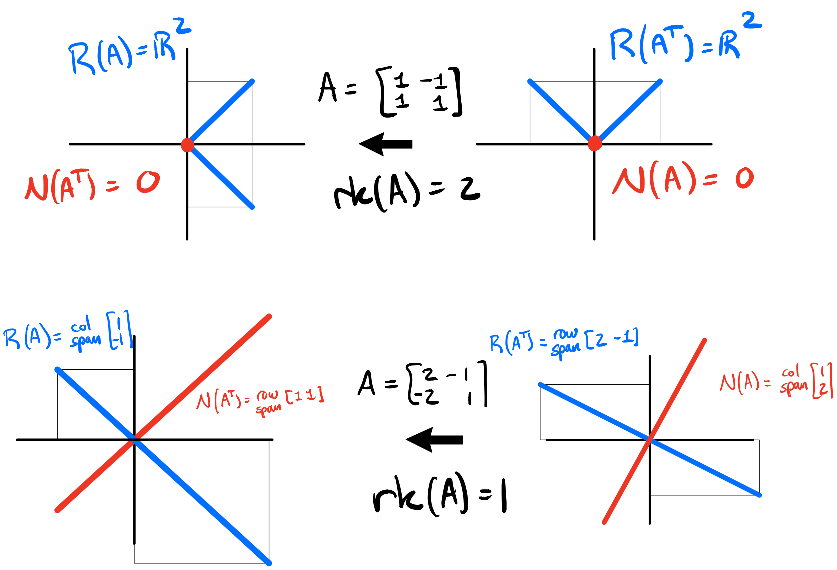
$$\dim(\mathcal{R}(A)) + \dim(\mathcal{N}(A)) = n$$
$$\text{rk}(A) + \dim(\mathcal{N}(A)) = n$$

1 Fundamental Theorem of Linear Algebra

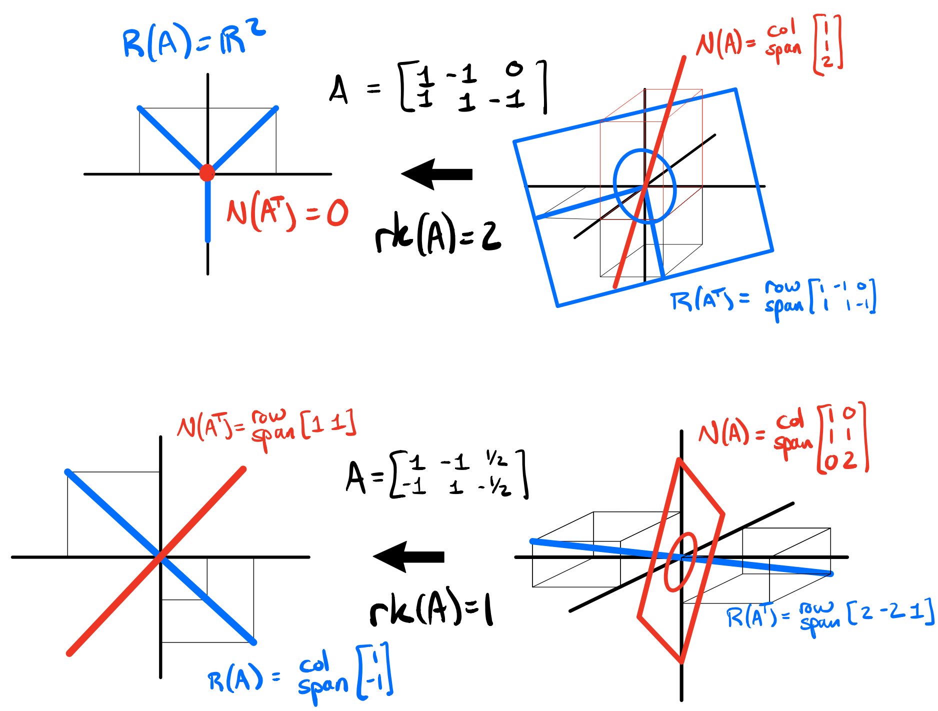
$\mathcal{R}(A^T)$ and $\mathcal{N}(A)$ are *orthogonal subspaces* of the domain, meaning that any vector in one is orthogonal to any vector in the other. In addition, together $\mathcal{N}(A)$ and $\mathcal{R}(A^T)$ span all of the domain $\mathbb{R}^n$. Similarly, $\mathcal{R}(A)$ and $\mathcal{N}(A^T)$ are orthogonal subspaces of the co-domain and together they span the co-domain.

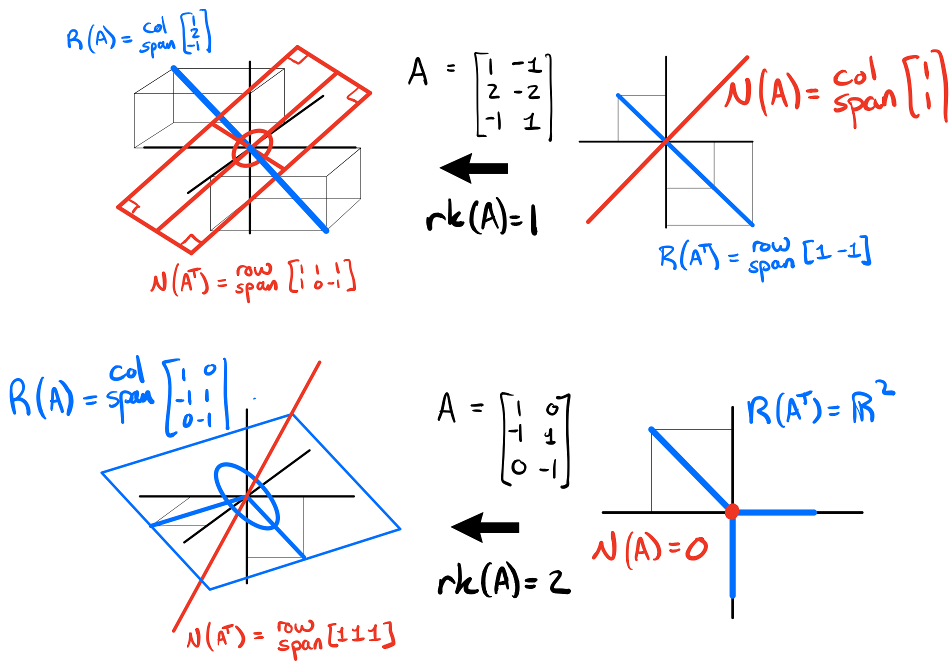
Fundamental Theorem of Linear Algebra Diagram



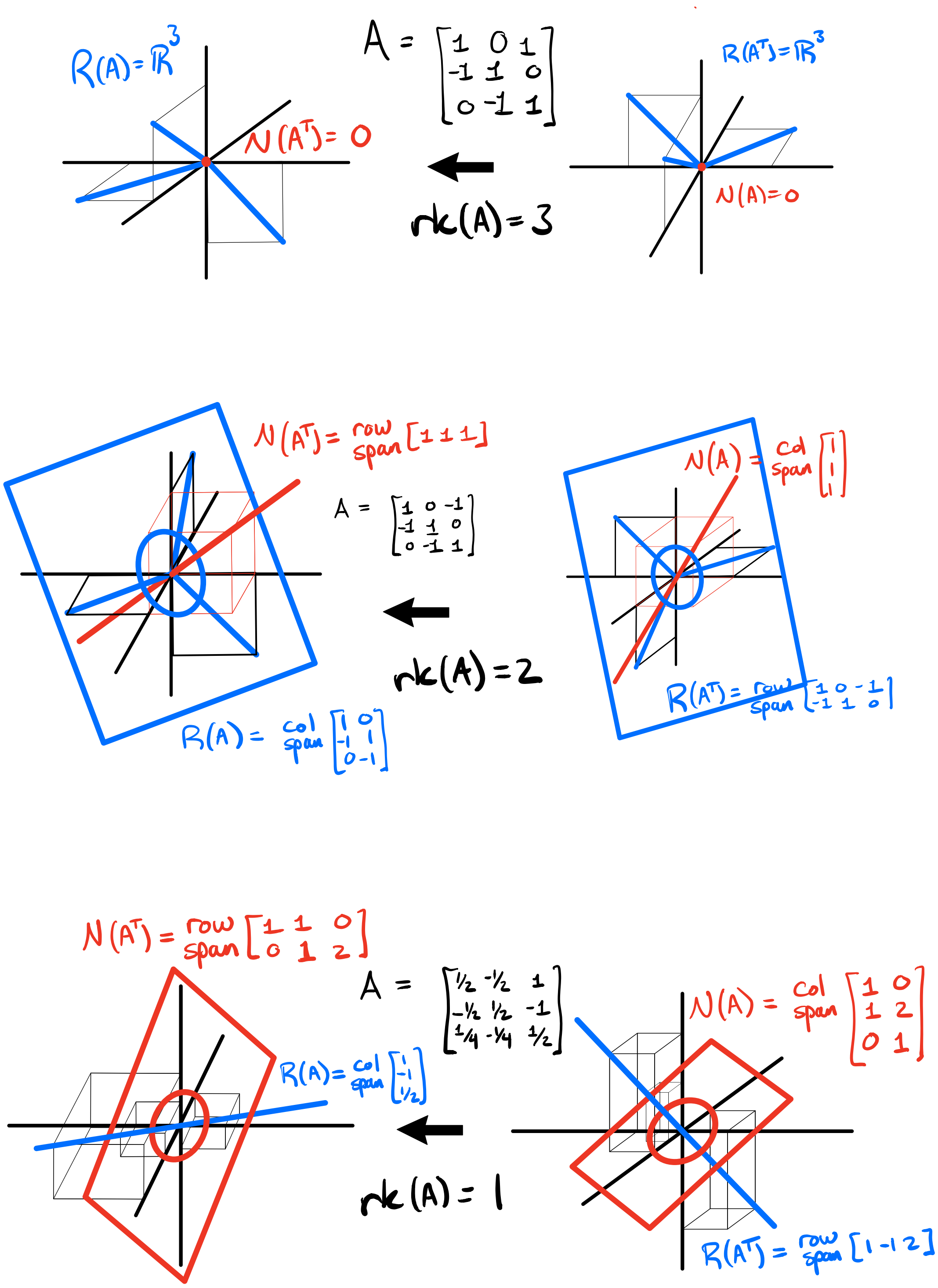


2×2 **examples of rank 2 and 1.**





2×3 **examples of rank 2 and 1.** 3×2 **examples of rank 2 and 1.**



3×3 **examples of rank 3,2, and 1.**