Column Space and Row Space of A

Nullspaces of A and A^{T}

Row Rank = Column Rank

Gilbert Strang, MIT

$A\boldsymbol{x} = \boldsymbol{b}$ in engineering and physics

$A =$ square $\boldsymbol{n}$ **by** $\boldsymbol{n}$ matrix—usually invertible

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

$A\widehat{\boldsymbol{x}} \approx \boldsymbol{b}$ in statistics and data science

$A =$ rectangular $\boldsymbol{m}$ **by** $\boldsymbol{n}$ matrix—no 2-sided inverse

The **pseudoinverse** $\boldsymbol{A}^{+}$ is the best we can do : $\widehat{\boldsymbol{x}} = \boldsymbol{A}^{+}\boldsymbol{b}$

Four Fundamental Subspaces $\mathsf{C}(A), \mathsf{C}(A^{\mathrm{T}}), \mathsf{N}(A), \mathsf{N}(A^{\mathrm{T}})$

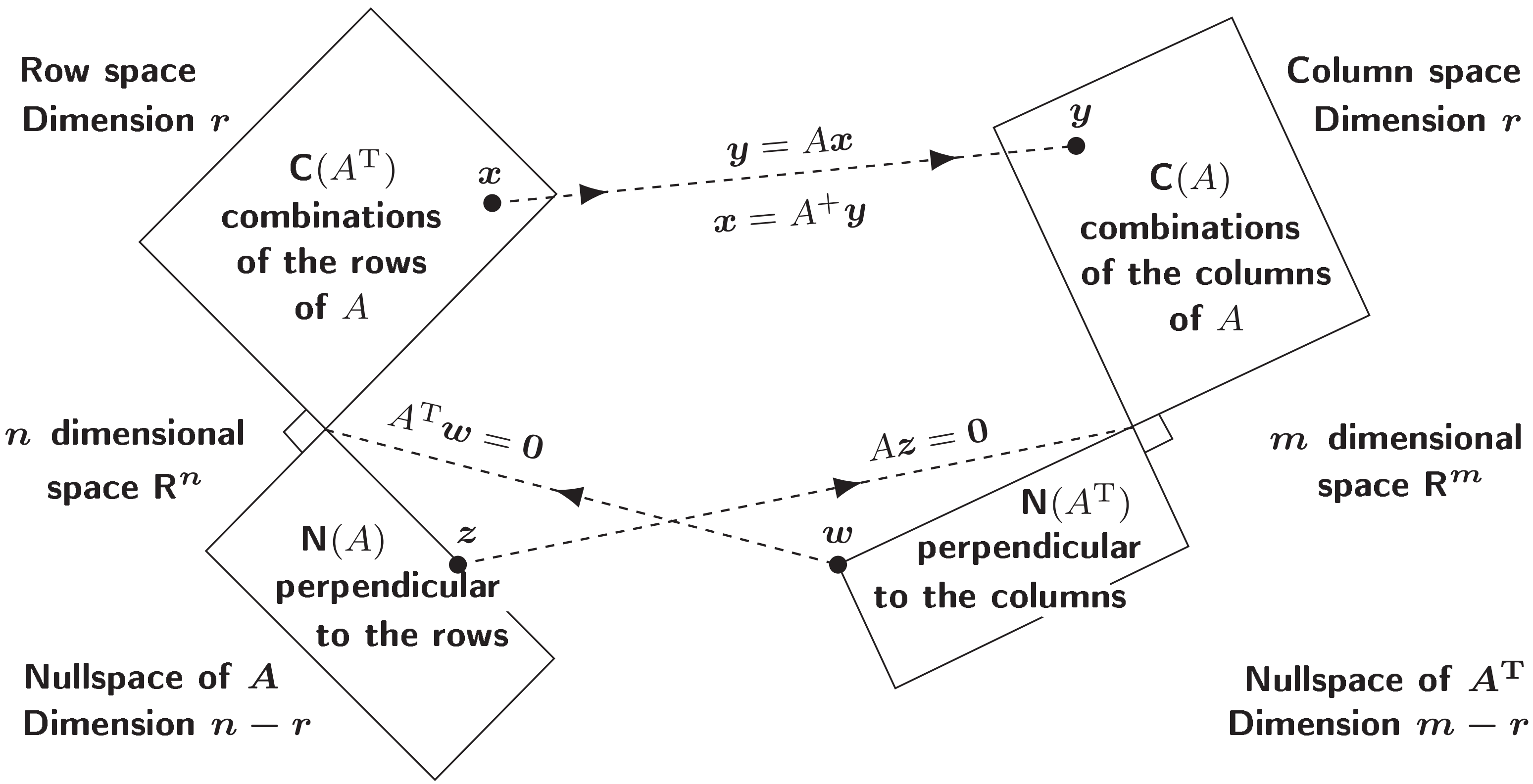


Figure: **The two perpendicular subspaces in n dimensions and m dimensions**

The Column Space $\mathbf{C}(\boldsymbol{A})$ of a Matrix $\boldsymbol{A}$

$\mathbf{C}(\boldsymbol{A})$ contains all combinations $\boldsymbol{Av}$
of the columns of A (all v_1, v_2, v_3).

$$\boldsymbol{Av} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + v_2 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + v_3 \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

For this matrix A, the column space $\mathbf{C}(\boldsymbol{A})$ is the
whole $\boldsymbol{xy}$ plane

We only need to use the first 2 columns (set $v_3 = 0$)

Solve $\begin{array}{l} 1v_1 + 2v_2 = x \\ 4v_1 + 5v_2 = y \end{array}$ to find v_1 and v_2 for any point (x, y)

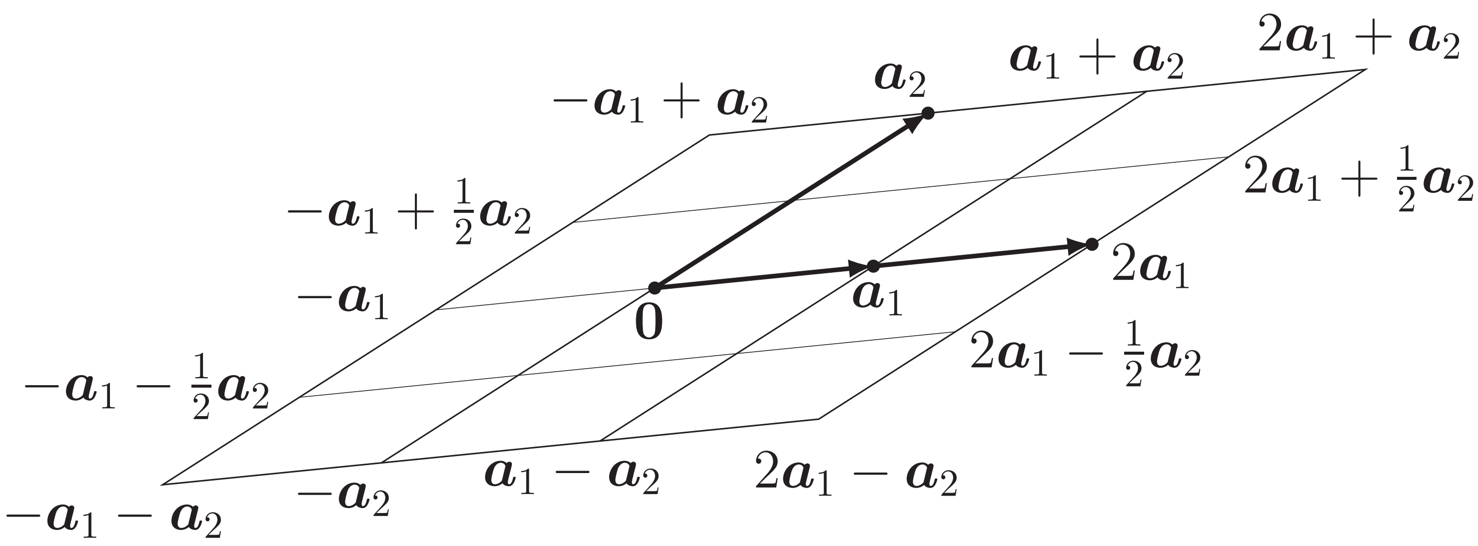
The Row Space $\mathrm{C}(A^{\mathrm{T}})$ of a Matrix A

Columns of A^{T} = transpose of A are the **rows of A.**
The center point is the zero vector :

$$A^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \quad \boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \boldsymbol{a}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad \textbf{zero vector} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Linear combination $= A^{\mathrm{T}}v = v_1 a_1 + v_2 a_2$
for any numbers v_1 and v_2

When v_1 and v_2 are negative, $A^{\mathrm{T}}v$ will reverse direction : right to left. Also very important, v_1 and v_2 can involve fractions. Here is a picture with 20 combinations.



The combinations $ca_1 + da_2$ fill a whole plane—the column space of A^{T}. It is a 2-dimensional infinite plane inside 3-dimensional space. By using more and more fractions and decimals v_1 and v_2, we fill in the plane!

The Nullspace of A

The nullspace $\mathbf{N}(\boldsymbol{A})$ contains all solution vectors $\boldsymbol{x}$ to the m equations $\boldsymbol{Ax} = \mathbf{0}$:

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Those equations say: $\boldsymbol{x}$ is perpendicular to each row of A

The nullspace N(A) is perpendicular to the row space (plane)

$$A\boldsymbol{x} = \mathbf{0} \text{ for } \boldsymbol{x} = \boldsymbol{c} \begin{bmatrix} \mathbf{1} \\ \mathbf{-2} \\ \mathbf{1} \end{bmatrix} = \textbf{line of vectors in N}(A)$$

The Nullspace of A^{T}

$\mathbf{N}(\boldsymbol{A}^{\mathrm{T}})$ contains all solution vectors $\boldsymbol{w}$ to $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w} = \mathbf{0}$:

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1w_1 + 4w_2 \\ 2w_1 + 5w_2 \\ 3w_1 + 6w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

In this case the only solution has $\boldsymbol{w_1} = \mathbf{0}$ and $\boldsymbol{w_2} = \mathbf{0}$

The nullspace of this example A^{T} is the **zero vector**

Column space $\mathbf{C}(\boldsymbol{A}) =$ whole x-y plane

Nullspace $\mathbf{N}(\boldsymbol{A}^{\mathrm{T}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Four Fundamental Subspaces $\mathbf{C}(A), \mathbf{C}(A^{\mathrm{T}}), \mathbf{N}(A), \mathbf{N}(A^{\mathrm{T}})$

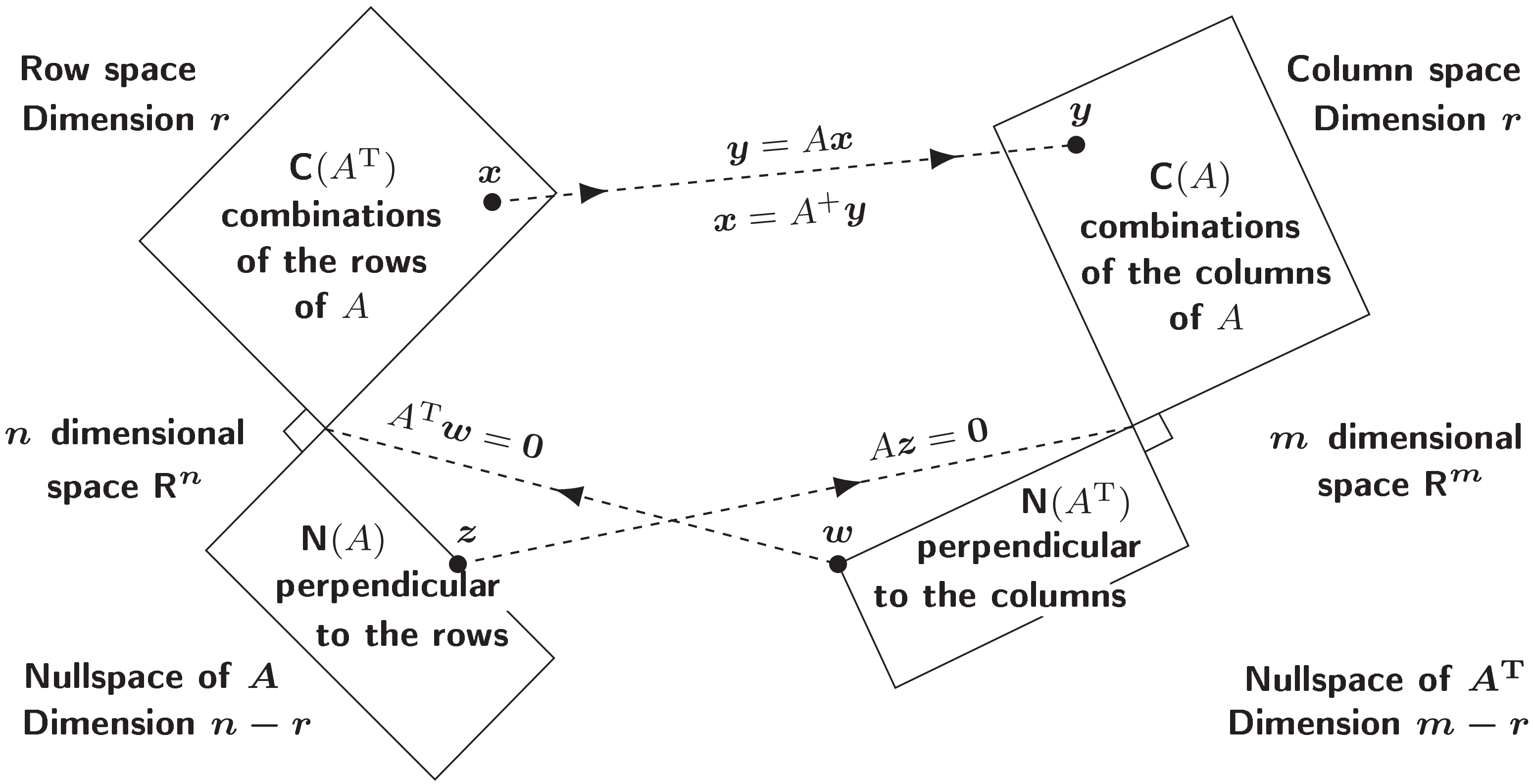


Figure: **The two perpendicular subspaces in n dimensions and m dimensions**

Column Rank = Row Rank = "Dimension" of C($\boldsymbol{A}$) and C($\boldsymbol{A}^{\mathrm{T}}$)

Number of independent columns $= \boldsymbol{r} =$ Number of independent rows $=$ Rank of A

One proof: Every $\boldsymbol{A} = \boldsymbol{CR} =$ (r independent columns of A in C) $\times$ (coefficients in R to produce all columns of A)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} = \boldsymbol{CR}$$

2 rows of R are independent. From $A = \boldsymbol{CR}$, the rows of R span the row space of A. Dimension of the column space of $A = 2 =$ Dimension of row space of A

From its row space to its column space, every $\boldsymbol{A}$ is invertible!

Reason: For v and w in the row space, suppose $A\boldsymbol{v} = A\boldsymbol{w}$ in the column space.

Then $A(\boldsymbol{v} - \boldsymbol{w}) = \mathbf{0}$. So $\boldsymbol{v} - \boldsymbol{w}$ is in both the **row space** and **nullspace** of A.

Those 2 spaces are perpendicular so $\boldsymbol{v} - \boldsymbol{w} =$ **zero vector**.

If rows $\boldsymbol{v}_1$ to $\boldsymbol{v}_r =$ basis for $\mathbf{C}(\boldsymbol{A}^{\mathrm{T}})$**, then columns** $\boldsymbol{A}\boldsymbol{v}_1$ **to** $\boldsymbol{A}\boldsymbol{v}_r =$ **basis for** $\mathbf{C}(\boldsymbol{A})$.

The **pseudoinverse** $\boldsymbol{A}^{+}$ brings each $A\boldsymbol{v}$ in $\mathbf{C}(A)$ back to $\boldsymbol{v}$ in the row space.

“Regression” in statistics = “Least squares” in linear algebra

Minimizing $||\boldsymbol{Ax} - \boldsymbol{b}||^2$ leads to $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$

If A has rank m (independent rows), then $\widehat{\boldsymbol{x}}$ = best $\boldsymbol{x}$

If $\boldsymbol{v}$ is in the nullspace of A, then also
$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(\widehat{\boldsymbol{x}} + \boldsymbol{v}) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$

$\widehat{\boldsymbol{x}}$ is the **minimum norm least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$**

$\boldsymbol{Ax} = \boldsymbol{b}$ leads to $\widehat{\boldsymbol{x}} = \boldsymbol{A}^{+}\boldsymbol{b}$
$\boldsymbol{A}^{+}$ is the “pseudoinverse” of $\boldsymbol{A}$