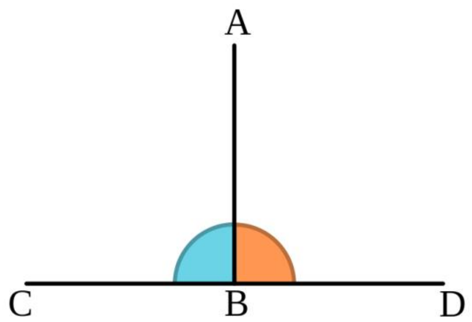
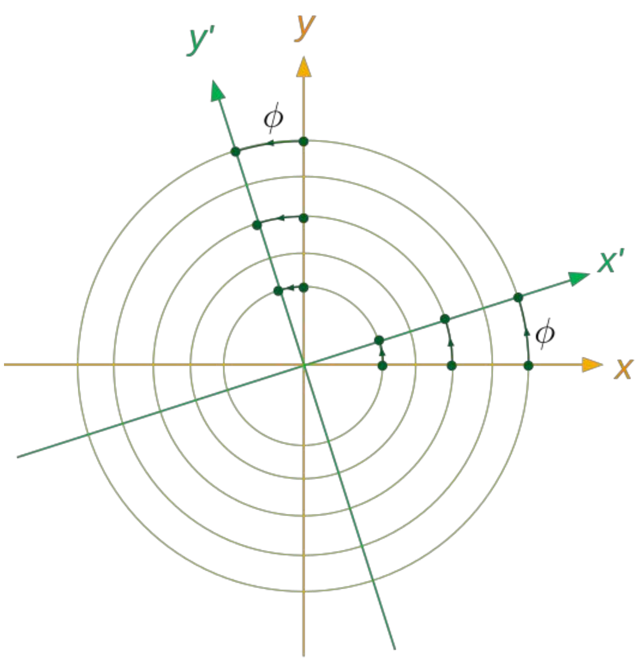
3 Orthogonality （正交性）



3.1 ORTHOGONAL VECTORS AND SUBSPACES（正交向量与子空间）

Orthogonal vectors

Orthogonal subspaces



I. Orthogonal vectors (正交向量)

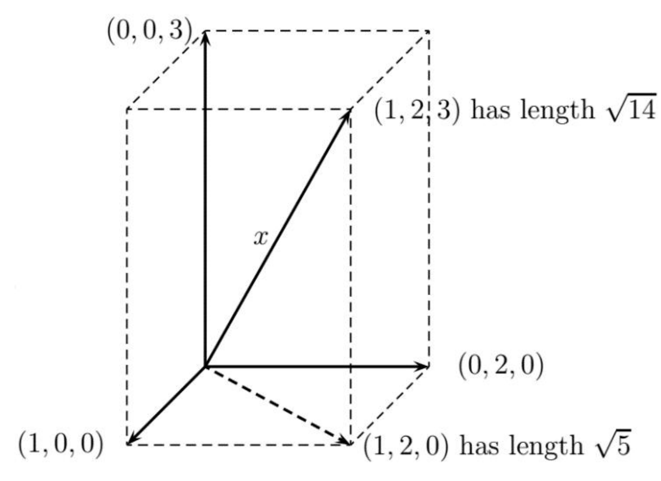
Definition 1 Let $\boldsymbol{u} = (a_1, a_2, \ldots, a_n)^{\mathrm{T}}$ and $\boldsymbol{v} = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$ be vectors of a vector space $\mathbf{R}^n$. Then the **inner product (内积)** of $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined as

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = \underset{1\times n}{\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}} \underset{n\times 1}{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}} = a_1b_1 + a_2b_2 + \cdots a_nb_n .$$

also called the **dot product (点积)** of $\boldsymbol{u}$ and $\boldsymbol{v}$.

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} = a_1^{\ 2} + a_2^{\ 2} + \cdots + a_n^{\ 2}$$

is the length squared of $\boldsymbol{u}$.



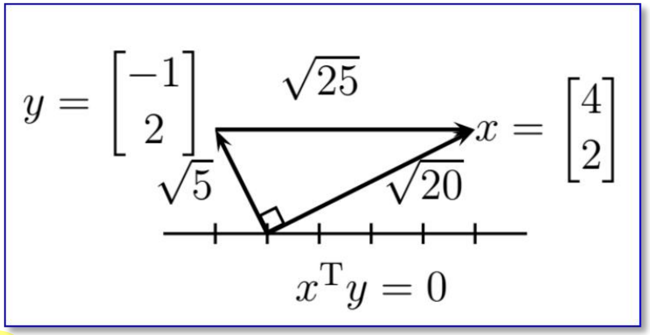
The length of $\boldsymbol{u}$ (向量$\boldsymbol{u}$的长度) is denoted by $||\boldsymbol{u}||$.

Notes:

- If $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0$, then $\boldsymbol{u}$, $\boldsymbol{v}$ are **orthogonal (正交)**, also called **perpendicular (垂直)**.

How can we decide whether vectors $\mathbf{x}$ *and* $\mathbf{y}$ *are perpendicular?*

$$||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2 = ||\boldsymbol{x}-\boldsymbol{y}||^2.$$

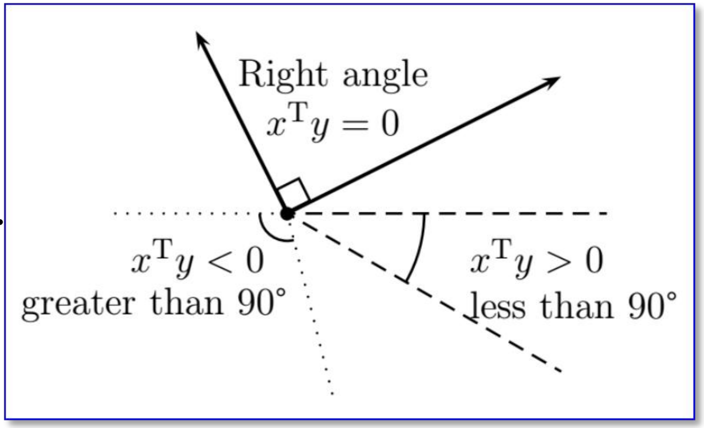


$$(x_1^2 + x_2^2 + \cdots + x_n^2) + (y_1^2 + y_2^2 + \cdots + y_n^2)$$
$$= (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2$$
$$= (x_1^2 + x_2^2 + \cdots + x_n^2) - 2\underline{(x_1y_1 + x_2y_2 + \ldots + x_ny_n)} + (y_1^2 + y_2^2 + \cdots + y_n^2)$$

$$\boxed{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = 0}$$

Notes:

- The only vector with length zero—the only vector orthogonal to itself—is the zero vector.
- The zero vector is orthogonal to every vector in $\mathbf{R}^n$.
- If $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} > 0$, their angle is less than 90^{o};
- if $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} < 0$, their angle is greater than 90^{o}.



Lemma *If nonzero vectors* $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ *are mutually orthogonal* (i.e., *every vector is perpendicular to every other*), *then they are linearly independent.* ~~(不含零向量的正交向量组一定线性无关)~~

不含零向量

Proof Suppose that $\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_k\boldsymbol{v}_k = \boldsymbol{0}$ for some scalars λ_i. Then

$$
\begin{aligned}
0 &= \boldsymbol{v}_i^{\mathrm{T}}\,\boldsymbol{0} \\
&= \boldsymbol{v}_i^{\mathrm{T}}(\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_k\boldsymbol{v}_k) \\
&= \lambda_1\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_2 + \cdots + \lambda_k\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_k \\
&= \lambda_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_i \\
&= \lambda_i||\boldsymbol{v}_i||^2.
\end{aligned}
$$

Since $\boldsymbol{v}_i \neq \boldsymbol{0}$, it follows that $\lambda_i = 0$.

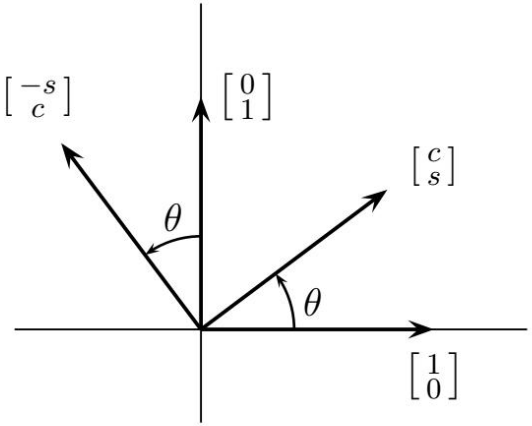
Therefore, the k vectors are linearly independent.

For $\mathbf{R}^n$, the basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ satisfy

$$\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_j = 1 \text{ or } 0 \text{ depending on } i = j \text{ or } i \neq j, \text{ respectively.}$$

- They are the most important orthogonal vectors
- They are the columns of the identity matrix
- They form the simplest basis for $\mathbf{R}^n$
- They are **unit vectors** (单位向量, 即长度为1的向量), $\|\boldsymbol{e}_i\|=1$
- They point along the coordinate axes; every $\boldsymbol{e}_i$ is perpendicular to every other $\boldsymbol{e}_j$

In general, a basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is called **orthonormal** (标准正交, 单位正交) if $\|\boldsymbol{v}_i\| = 1$, and $\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_j = 0$ for $i \neq j$.



Rotation through θ

In $\mathbf{R}^2$, if these axes are rotated, the result is a new **orthonormal basis**: a new system of *mutually orthogonal unit vectors*:

$$\boldsymbol{v}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}.$$

II. Orthogonal subspaces (正交子空间)

Definition 2 Let V, W be two subspaces of a vector space $\mathbf{R}^n$.
If *every* vector of $\boldsymbol{v}$ in V is orthogonal to *every* vector $\boldsymbol{w}$ in W (i.e., $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = 0$ for all $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$), then V is said to be **orthogonal** (正交) to W, denoted by

$$V \perp W.$$

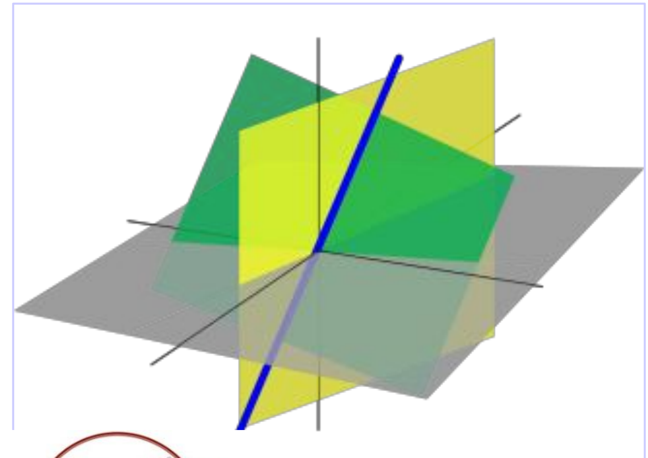
Example 1. Let $V = \{(x_1, x_2, 0, 0)^{\mathrm{T}} \mid x_1, x_2 \in \mathbf{R}\}$,

and $W = \{(0, 0, x_3, x_4)^{\mathrm{T}} \mid x_3, x_4 \in \mathbf{R}\}$.

Then $V \perp W$.

The orthogonality of two subspaces:

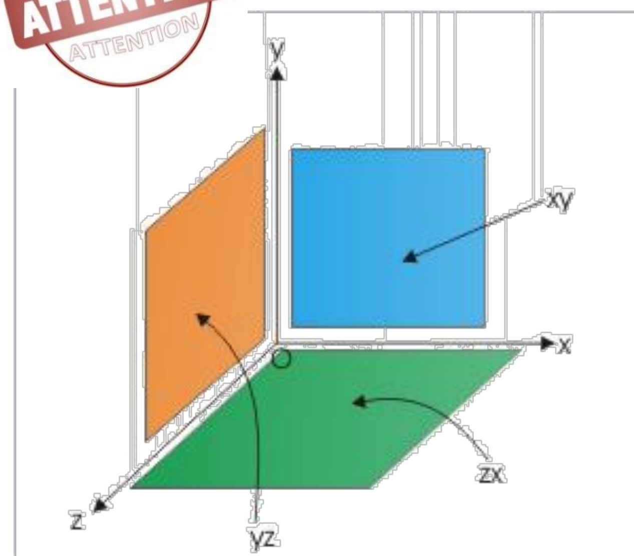
- **Every vector** *in one subspace must be orthogonal to* **every vector** *in the other subspace*.
- Subspaces of $\mathbf{R}^3$ can have dimension 0, 1, 2, or 3. The subspaces are represented by lines or planes through the origin—and in the extreme cases, by the origin alone or the whole space.
- The subspace $\{\boldsymbol{0}\}$ is orthogonal to all subspaces.



- A line can be orthogonal to another line, or it can be orthogonal to a plane, but –

Two planes in $\mathbf{R}^3$ *cannot be orthogonal.*

By our definition, the line along the corner is in *both* walls, and it is certainly not orthogonal to itself.



判断 basis 是否正交

Example 2. Suppose V is the plane spanned by $\boldsymbol{v}_1 = (1,0,0,0)^{\mathrm{T}}$ and $\boldsymbol{v}_2 = (1,1,0,0)^{\mathrm{T}}$.

If W is the line spanned by $\boldsymbol{w} = (0,0,4,5)^{\mathrm{T}}$,

then $\boldsymbol{w}$ is orthogonal to both $\boldsymbol{v}$'s.

The line W will be orthogonal to the whole plane V.

In this case, with subspaces of dimension 2 and 1 in $\mathbf{R}^4$, there is room for a third subspace.

The line L through $\boldsymbol{z} = (0,0,5,-4)^{\mathrm{T}}$ is perpendicular to V and W.

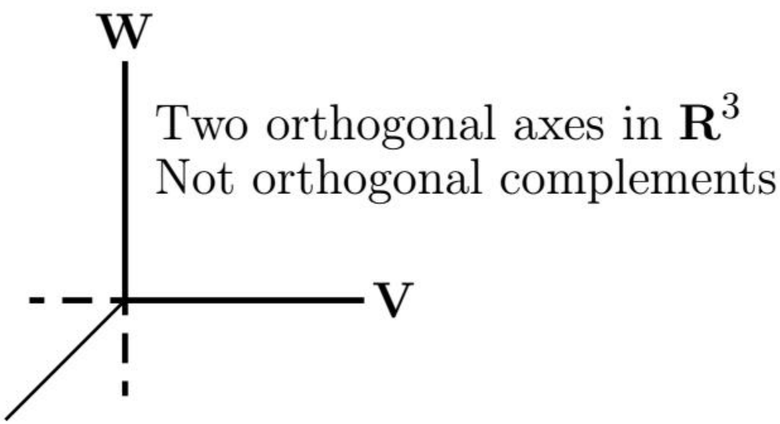
Then the dimensions add to $2 + 1 + 1 = 4$.

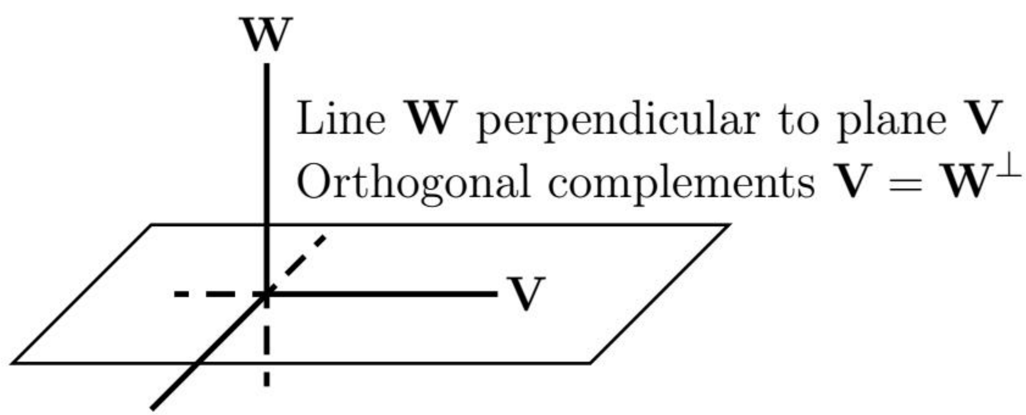
Orthogonal complement

Given a subspace V of $\mathbf{R}^n$, the space of ***all*** vectors orthogonal to V is called the **orthogonal complement** (正交补) of V. It is denoted by $W = V^{\perp} =$ "V perp."

Note:

Two subspaces can be orthogonal without being orthogonal complements.





Splitting $\mathbf{R}^n$ into orthogonal parts will split every vector into

$$\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{w}.$$

The vector $\boldsymbol{v}$ is the *projection* onto the subspace V.

The orthogonal component $\boldsymbol{w}$ is the *projection* of $\boldsymbol{x}$ onto W.

The four subspaces related to the matrix $A \in \mathbf{R}^{m\times n}$

$\boldsymbol{C}(A)$ = column space of A; dimension r.	$\subseteq \mathbf{R}^m$
$\boldsymbol{N}(A)$ = nullspace of A; dimension $n-r$.	$\subseteq \mathbf{R}^n$
$\boldsymbol{C}(A^{\mathrm{T}})$ = row space of A; dimension r.	$\subseteq \mathbf{R}^n$
$\boldsymbol{N}(A^{\mathrm{T}})$ = left nullspace of A; dimension $m-r$.	$\subseteq \mathbf{R}^m$

Fundamental Theorem of Linear Algebra, Part I

Theorem *Let* $\boldsymbol{A}$ *be a matrix. Then the row space* $\mathrm{C}(\boldsymbol{A}^{\mathrm{T}})$ *is orthogonal to the nullspace* $\mathrm{N}(\boldsymbol{A})$*, and the column space* $\mathrm{C}(\boldsymbol{A})$ *is orthogonal to the left nullspace* $\mathrm{N}(\boldsymbol{A}^{\mathrm{T}})$*. That is to say,*

$$\mathrm{C}(A^{\mathrm{T}}) \perp \mathrm{N}(A), \quad \textit{and} \quad \mathrm{C}(A) \perp \mathrm{N}(A^{\mathrm{T}}).$$

Those four subspaces are perpendicular in pairs, two in $\mathbf{R}^n$ and two in $\mathbf{R}^m$.

$\mathrm{Nul}(A) = (\mathrm{Row}(A))^{\perp} = (\mathrm{Col}(A^T))^{\perp}$

$\mathrm{Nul}(A^T) = (\underline{\mathrm{Col}(A)})^{\perp}$

First proof Suppose $\boldsymbol{x} \in \mathrm{N}(A)$. Then $A\boldsymbol{x} = \boldsymbol{0}$, and this system of m equations can be written out as rows of A multiplying $\boldsymbol{x}$:

$\boldsymbol{x}$ is orthogonal to every row

$$A\boldsymbol{x} = \begin{bmatrix} \cdots & \text{row } 1 & \cdots \\ \cdots & \text{row } 2 & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \text{row } m & \cdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Therefore $\boldsymbol{x}$ is orthogonal to every *combination* of the rows.
Each $\boldsymbol{x}$ in the nullspace is orthogonal to each vector in the row space, so $\mathrm{N}(A) \perp \mathrm{C}(A^{\mathrm{T}})$.

The other pair of orthogonal subspaces comes from $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$, or $\boldsymbol{y}^{\mathrm{T}}A = \boldsymbol{0}$:

$$\boldsymbol{y}^{\mathrm{T}}A = \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} \begin{bmatrix} \text{column } 1 & \cdots & \text{column } n \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}.$$

$\boldsymbol{y}$ is orthogonal to every column (and every *combination* of the columns)

So $\mathrm{N}(A^{\mathrm{T}}) \perp \mathrm{C}(A)$.

Second proof (It shows a more "abstract" method of reasoning.)

If $\boldsymbol{x}$ is in the nullspace then $\boldsymbol{Ax} = \boldsymbol{0}$.

(向量$\boldsymbol{x}$在零空间中, 故 $\boldsymbol{Ax} = \boldsymbol{0}$)

If $\boldsymbol{v}$ is in the row space, it is a combination of the rows: $\boldsymbol{v} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}$ for some vector $\boldsymbol{z}$.

(向量$\boldsymbol{v}$在行空间中, 故它可以写成行向量的线性组合, 即: 存在某个向量$\boldsymbol{z}$, 使得 $\boldsymbol{v} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}$)

Now, in one line (一言以蔽之):

$$\underline{\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{z}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{z}^{\mathrm{T}}\boldsymbol{0} = 0.}$$

That is to say,

$$\textbf{Nullspace} \perp \textbf{Row space}.$$

Similarly,

$$\textbf{Left nullspace} \perp \textbf{Column space}.$$

Example 3 Suppose $\boldsymbol{A}$ has rank 1:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{bmatrix},$$

so its row space and column space are lines.

The rows are multiples of $(1,3)^{\mathrm{T}}$.

The nullspace contains $\boldsymbol{x} = (-3,1)^{\mathrm{T}}$, which is orthogonal to all the rows.

The nullspace and row space are perpendicular lines in $\mathbf{R}^2$.

In contrast, the other two subspaces are in $\mathbf{R}^3$.

The column space is the line through $(1,2,3)^{\mathrm{T}}$.

The left nullspace must be the *perpendicular plane* $y_1 + 2y_2 + 3y_3 = 0$.

That equation is exactly the content of $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A} = \mathbf{0}$.

$\boldsymbol{C}(A) =$ column space of A; dimension $\underline{r}$. $\subseteq \mathbf{R}^m$

$\boldsymbol{N}(A) =$ nullspace of A; dimension $\underline{n-r}$. $\subseteq \mathbf{R}^n$

$\boldsymbol{C}(A^{\mathrm{T}}) =$ row space of A; dimension $\underline{r}$. $\subseteq \mathbf{R}^n$

$\boldsymbol{N}(A^{\mathrm{T}}) =$ left nullspace of A; dimension $\underline{m-r}$. $\subseteq \mathbf{R}^m$

$$r + (n - r) = n$$

$$r + (m - r) = m$$

Fundamental Theorem of Linear Algebra, Part II

Theorem *Let **A** be a matrix. Then the row space* $\mathrm{C}(\boldsymbol{A}^{\mathrm{T}})$ *is orthogonal to the nullspace* $\mathrm{N}(\boldsymbol{A})$, *and the column space* $\mathrm{C}(\boldsymbol{A})$ *is orthogonal to the left nullspace* $\mathrm{N}(\boldsymbol{A}^{\mathrm{T}})$. *Moreover, if **A** has size* $m \times n$, *then*

$$\mathrm{N}(\boldsymbol{A}) = (\mathrm{C}(\boldsymbol{A}^{\mathrm{T}}))^{\perp}, \quad \text{and} \quad \mathrm{N}(\boldsymbol{A}^{\mathrm{T}}) = (\mathrm{C}(\boldsymbol{A}))^{\perp}.$$

In other words: The nullspace is the *orthogonal complement* of the row space in $\mathbf{R}^n$. The left nullspace is the *orthogonal complement* of the column space in $\mathbf{R}^m$.

The row space contains *everything* orthogonal to the nullspace.

The column space contains *everything* orthogonal to the left nullspace.

We notice the following equivalence:

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$$

$\Leftrightarrow$ $\boldsymbol{b} \in \mathrm{C}(\boldsymbol{A})$ ($\boldsymbol{b}$ is in the column space of $\boldsymbol{A}$)

$\Leftrightarrow$ $\boxed{\boldsymbol{b} \perp \mathrm{N}(\boldsymbol{A}^{\mathrm{T}})}$

$\Leftrightarrow$ $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0$ for all $\boldsymbol{y} \in \mathrm{N}(\boldsymbol{A}^{\mathrm{T}})$

(i.e., $\boldsymbol{b}$ is perpendicular to the left nullspace)

$\Leftrightarrow$ $\boxed{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0 \text{ for all } \boldsymbol{y}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{0}.}$

It implies that following conclusion.

Lemma. *A system* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *has solutions* (*or called 'solvable'*) *if and only if* $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0$ *whenever* $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{0}$.

Theorem *As a linear transformation, each matrix **A** transforms its row space* $C(A^T)$ *to its column space* $C(A)$.

Proof Let A be an $(m\times n)$-matrix.

Then A maps vectors of $\mathbf{R}^n$ to vectors of $\mathbf{R}^m$ by $A : \boldsymbol{x} \mapsto A\boldsymbol{x}$, since for each $\boldsymbol{x} \in \mathbf{R}^n$, the image $A\boldsymbol{x} \in \mathbf{R}^m$.

Each vector $\boldsymbol{x} \in \mathbf{R}^n$ can be written as

$$\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n,$$

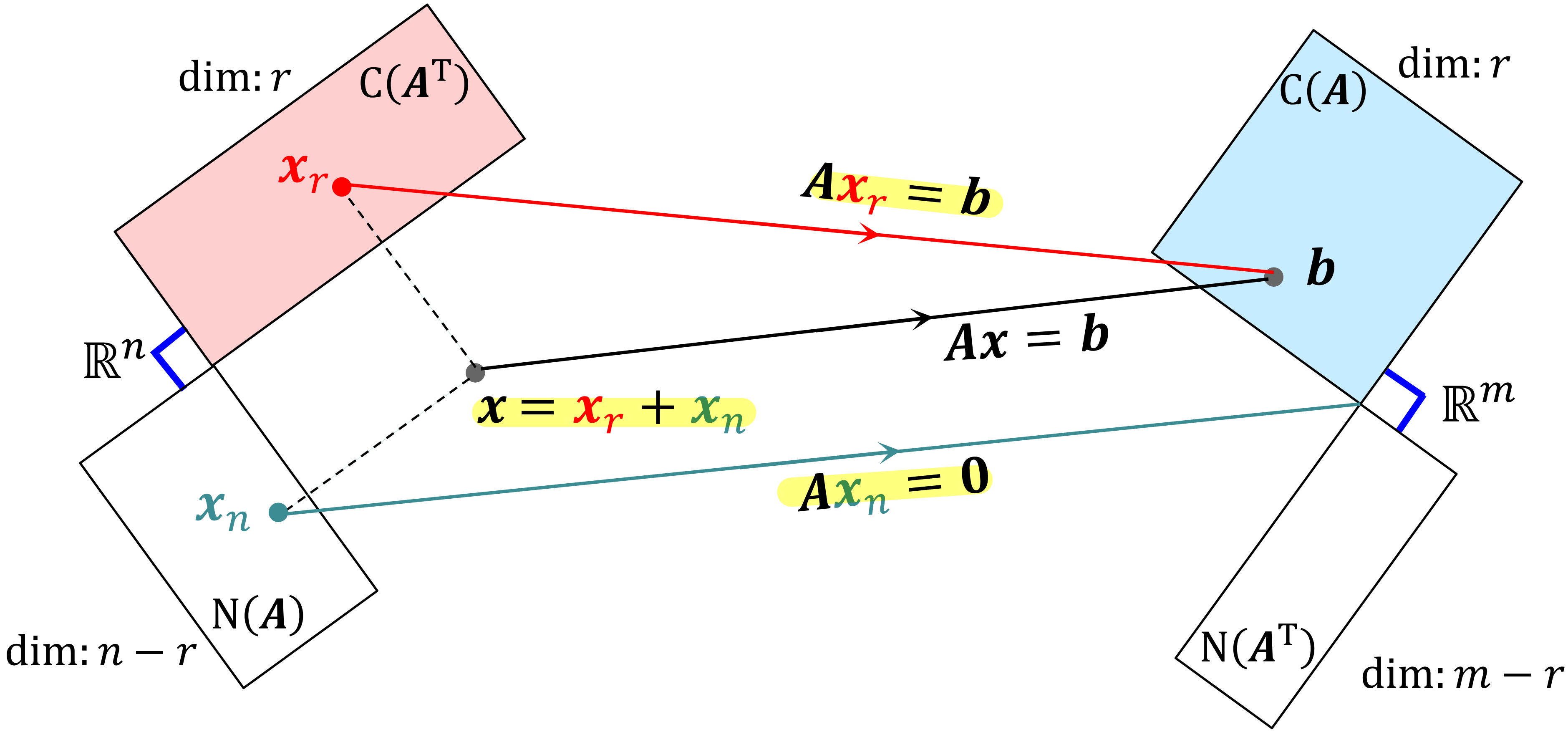
where $\boldsymbol{x}_r \in C(A^T)$ and $\boldsymbol{x}_n \in N(A)$. Since $A\boldsymbol{x}_n = \boldsymbol{0}$, we have

$$A\boldsymbol{x} = A(\boldsymbol{x}_r + \boldsymbol{x}_n) = A\boldsymbol{x}_r + A\boldsymbol{x}_n = A\boldsymbol{x}_r.$$

On the other hand, let $\boldsymbol{x}_r = (k_1, k_2, \ldots, k_n)^T$, and let $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n$ be the columns of A, i.e., $A = [\boldsymbol{c}_1\ \ \boldsymbol{c}_2\ \ \ldots\ \ \boldsymbol{c}_n]$. Then

$$A\boldsymbol{x}_r = k_1\boldsymbol{c}_1 + k_2\boldsymbol{c}_2 + \cdots + k_n\boldsymbol{c}_n \in C(A).$$

Therefore, as a linear transformation, A transforms the row space $C(A^T)$ to the column space $C(A)$.



The fundamental subspaces meet at right angles.

Every matrix *A* transforms its row space onto its column space.

Actually, A transforms vectors of $\mathbf{R}^n$ to the column space $C(A)$.
Since A transforms vectors of the nullspace $N(A)$ to $\boldsymbol{0}$,
A actually transforms $C(A^T)$ onto $C(A)$.

Note: Every vector $\boldsymbol{b} \in \mathrm{C}(\boldsymbol{A})$ comes from exactly one vector $\boldsymbol{x}_r \in \mathrm{C}(\boldsymbol{A}^{\mathrm{T}})$.

Proof If $\boldsymbol{x}_r, \boldsymbol{x}'_r \in \mathrm{C}(\boldsymbol{A}^{\mathrm{T}})$ and $\boldsymbol{A}\boldsymbol{x}_r = \boldsymbol{A}\boldsymbol{x}'_r = \boldsymbol{b}$,

then

$$\boldsymbol{A}(\boldsymbol{x}_r - \boldsymbol{x}'_r) = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{0},$$

$$(\boldsymbol{x}_r - \boldsymbol{x}'_r) \in \mathrm{N}(\boldsymbol{A}) \cap \mathrm{C}(\boldsymbol{A}^{\mathrm{T}}),$$

$\boldsymbol{x}_r - \boldsymbol{x}'_r$ is orthogonal to itself since $\mathrm{N}(\boldsymbol{A}) \perp \mathrm{C}(\boldsymbol{A}^{T})$. Therefore

$$\boldsymbol{x}_r = \boldsymbol{x}'_r,$$

i.e., every vector $\boldsymbol{b} \in \mathrm{C}(\boldsymbol{A})$ comes from *exactly* one vector $\boldsymbol{x}_r \in \mathrm{C}(\boldsymbol{A}^{\mathrm{T}})$.

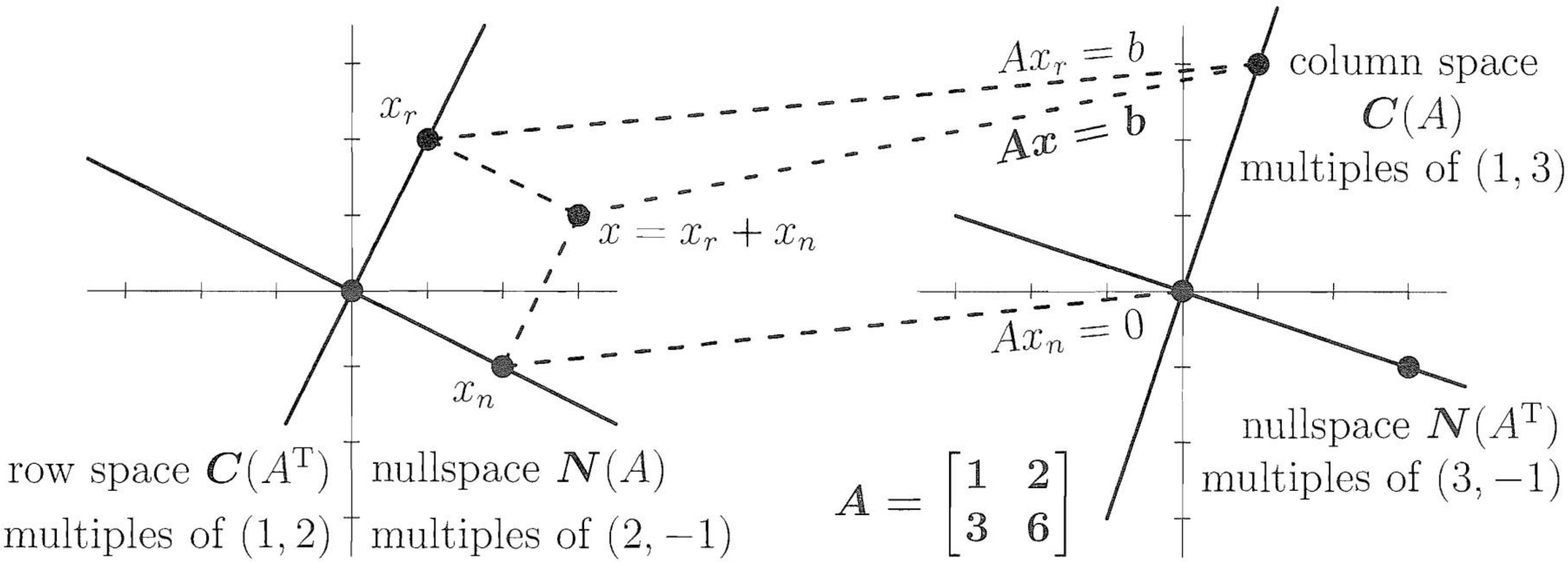
P149#12(hard cover): Let $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix}$.

Split the vector $\boldsymbol{x}$ =(3,3,3) into a row space component $\boldsymbol{x}_r$ and a nullspace component $\boldsymbol{x}_n$.

Example 4 $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, $m = n = 2$, $r = 1$. **Singular matrix**

Column space: $Span\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\}$ Nullspace: $Span\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}\}$

Row space: $Span\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\}$ Left nullspace: $Span\{\begin{bmatrix} 3 \\ -1 \end{bmatrix}\}$



Key words:

Orthogonal vectors

Orthogonal subspaces

Fundamental Theorem of Linear Algebra

Homework

See Blackboard