# Definitions

① Vector space $\mathcal{V}$ or $\mathbb{R}^n$:

collection of all vectors with n components

Example: $\mathbb{R}^3$

② Subspace $\mathcal{S}$ of $\mathcal{V}$:

collection of all vectors satisfying the following:

• for any vectors $\vec{x}$ & $\vec{y}$ inside $\mathcal{S}$, all of their linear combinations are also inside $\mathcal{S}$ (which also implies that $\mathcal{S}$ includes the zero vector)

Examples: plane, line or zero vector in $\mathbb{R}^3$ (through the origin)

③ Complementary subspaces $\mathcal{S}1$ & $\mathcal{S}2$ of $\mathcal{V}$ satisfy following conditions:

• intersect only at 0

• every vector $\vec{x}$ in $\mathcal{V}$ can be written uniquely as $\vec{x} = \vec{x}1 + \vec{x}2$ with $\vec{x}1$ in $\mathcal{S}1$ and $\vec{x}2$ in $\mathcal{S}2$

or, equivalently, $\mathcal{S}1$ & $\mathcal{S}2$ span $\mathcal{V}$

④ Orthogonal subspaces $\mathcal{S}1$ & $\mathcal{S}2$ satisfy following condition:

every vector in $\mathcal{S}1$ is orthogonal to every vector in $\mathcal{S}2$

Note:

• Complementary does not imply orthogonal

• Orthogonal does not imply complementary ($\mathcal{S}1$ & $\mathcal{S}2$ may not span $\mathcal{V}$)

• $\mathcal{S}2$ as the set of ALL vectors orthogonal to all vectors in $\mathcal{S}1$ does imply complementary ($\mathcal{S}1$ & $\mathcal{S}2$ span $\mathcal{V}$)

• If $\mathcal{S}1$ & $\mathcal{S}2$ are complementary, most vectors do not lie entirely in $\mathcal{S}1$ or $\mathcal{S}2$: they have components in both $\mathcal{S}1$ & $\mathcal{S}2$

• Complementary is a dimension condition:

$\dim(\mathcal{S}1) + \dim(\mathcal{S}2) = \dim(\mathcal{V})$ and $\mathcal{S}1 \cap \mathcal{S}2 = \{0\}$

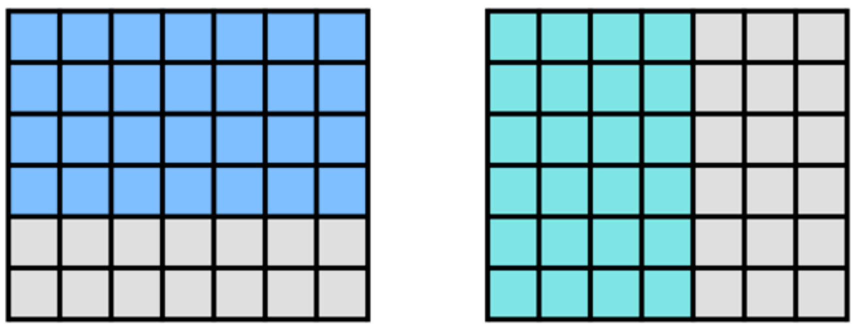
The four subspaces of A: introduction

Suppose A is an $(m=6)\times(n=7)$ matrix with rank $r=4$

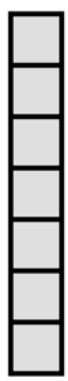
‣ r independent rows highlighted (left)

‣ r independent columns highlighted (right):

independent rows & columns may appear in any order



Matrix-vector product $A\times\vec{x}$ is defined for all $n\times 1$ vectors $\vec{x}$



⦿ Domain of A is the set of all $\vec{x}$ where $A\times\vec{x}$ is defined

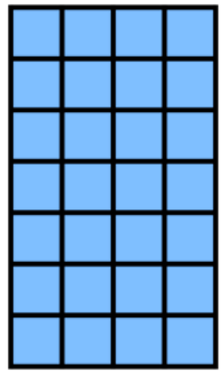
$\Leftrightarrow$

$\mathbb{R}^n$

Domain of A or $\mathbb{R}^n$ can be decomposed into two complementary subspaces:

① All linear combinations of independent rows of A,

shown as an r-dimensional subspace of $\mathbb{R}^n$



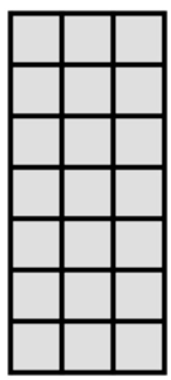
called row space of A

⇔

independent rows of A provide a basis for this subspace

② Subset of all $\vec{x}$ where $A \times \vec{x} = 0$, called null space of A

(here each vector represents a direction: the whole line $\{ c \times \vec{x} : c \in \mathbb{R} \}$)



- null space ⊕ row space = $\mathbb{R}^n$ or the two subspaces are complements of each other

This statement means the following is true:

‣ their dimensions add up to n

‣ they do not overlap or, equivalently, the only intersection is zero vector

‣ every vector $\vec{x}$ in $\mathbb{R}^n$ has a unique representation as ($\vec{x}_{||} \in$ row space) + ($\vec{x}_{\perp} \in$ null space)

- Note that for every $\vec{x}$ in the null space, $A \times \vec{x} = 0$

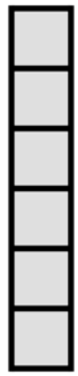
⇓

$a^i \cdot \vec{x} = 0$ ($a^i$ are rows of A)

⇩

every vector in null space is orthogonal to every vector in row space

⦿ Codomain of A is the set of all vectors $\vec{b}$ of size m



⇩

$\mathbb{R}^m$

Codomain of A or $\mathbb{R}^m$ also can be decomposed into two complementary subspaces:

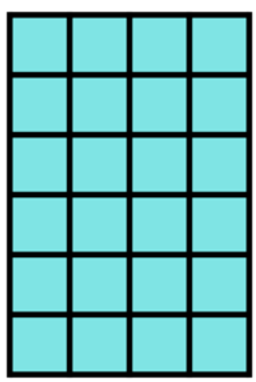
③ Range (column space) of A: subset of $\vec{b}$ that can be written as $\vec{b} = A \times \vec{x}$
where any $\vec{b}$ is a linear combination of columns of A with coefficients from $\vec{x}$

⇩

Range of A is the set of all linear combinations of columns of A

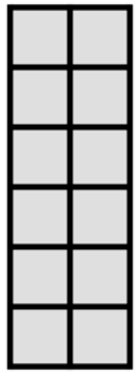
⇔

columns of A provide a basis for this subspace



④ Left null space of A: subset of $\vec{b}$ where $\vec{b}^T \times A = 0$

shown as (m-r)-dimensional subspace of $\mathbb{R}^m$



- column space ⊕ left null space = $\mathbb{R}^m$ or the two subspaces are complements of each other

This statement means the following is true:

‣ their dimensions add up to m $\left(r + (m-r) = m\right)$

‣ they do not overlap or, equivalently, the only intersection is zero vector

‣ every vector $\vec{b}$ in $\mathbb{R}^m$ has a unique representation as ($\vec{b}_{\parallel}$ ∈ column space) + ($\vec{b}_{\perp}$ ∈ left null space)

- Note that for every $\vec{b}$ in the left null space, $\vec{b}^T \times A = 0$

⇩

$$\vec{b}^T \cdot a^i = 0 \quad \left(a^i \text{ are columns of A}\right)$$

⇩

every vector in left null space is orthogonal to every vector in column space

---

Geometric interpretation of the four subspaces of A

Suppose we have a 3×2 matrix A = $[\, \vec{a}_1 \mid \vec{a}_2 \,]$
where $\vec{a}_1$ & $\vec{a}_2$ are linearly independent vectors in $\mathbb{R}^3$,
so rank(A) = 2

① Domain(A) = $\mathbb{R}^2$, an abstract 2D coefficient space, or an abstract plane

② null space(A) or all directions in the domain ($\mathbb{R}^2$) that produce $\vec{0}$ output:
For the given A, only the $\vec{0}$ vector:
• Dimension: n-r=0 in $\mathbb{R}^2$
• Every vector from the domain produces a unique output
⇔
Transformation is one-to-one

③ row space(A)
While rows of A cannot be directly visualized in the same image as columns,
row space is the set of all vectors in the domain that are orthogonal to null space,
or, equivalently,
set of all vectors that have 0 component in the null space
For the given A, all vectors in $\mathbb{R}^2$ satisfy the condition
• Dimension: r=2 in $\mathbb{R}^2$

④ Codomain(A) is the set of all vectors in $\mathbb{R}^3$

⑤ column space(A) or all linear combinations of $\vec{a}_1$ & $\vec{a}_2$
or the plane spanned by the two vectors
in $\mathbb{R}^3$ and expressed in 3D coordinates
• Dimension: r=2 in $\mathbb{R}^3$
• Output does not fill the codomain
⇔
Transformation is not onto

⑥ ℓ-null space(A) or all directions in the codomain ($\mathbb{R}^3$) orthogonal to column space:
all vectors along the orthogonal line, shown in purple
• Dimension: m-r=1 in $\mathbb{R}^3$

⑦ Matrix A

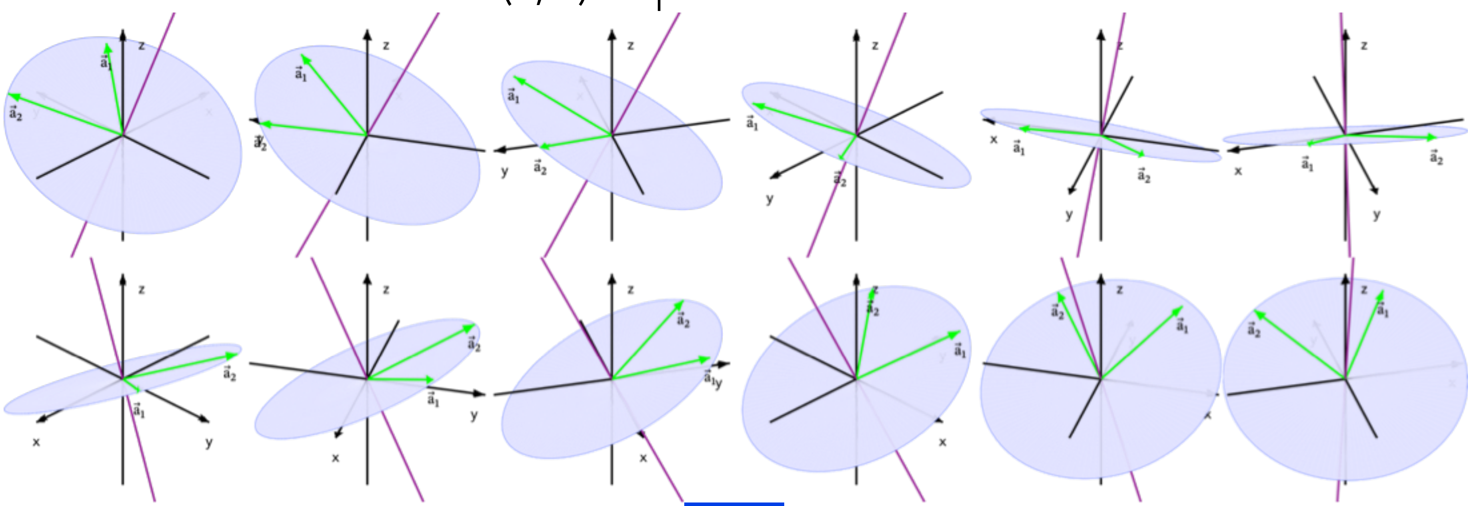
- takes a set of all vectors in $\mathbb{R}^2$ and transforms them in a way that
  - (1, 0) maps to $\vec{a}_1$
  - (0, 1) maps to $\vec{a}_2$
- all other vectors $\vec{x}$ follow as $x_1\vec{a}_1 + x_2\vec{a}_2$, which fills the blue plane in $\mathbb{R}^3$

⑧ Matrix $A^T$

- takes vectors in $\mathbb{R}^3$ and transforms them in a way that
  - $\vec{a}_1$ maps to $A^TA$ column 1
  - $\vec{a}_2$ maps to $A^TA$ column 2
- all other 3D vectors $\vec{y}$ 'follow suit' and are now expressed in 2D coordinates

⑨ Matrix $A^TA$ acts entirely inside $\mathbb{R}^2$ in a way that

- (1, 0) maps to $A^TA$ column 1
- (0, 1) maps to $A^TA$ column 2



---

The four subspaces of A: summary table

- Domain ($\mathbb{R}^n$) = row space ⊕ null space

- Codomain ($\mathbb{R}^m$) = column space ⊕ ℓ-null space

| Space | Lives in | Dim | Defined by | ⊥ Complement | How it arises | Interpretation |
|---|---|---|---|---|---|---|
| null space(A) | $\mathbb{R}^n$ | n−r | all $\vec{x}$ satisfying $A\,\vec{x} = \vec{0}$ | row space(A) | Computed from property: $A\vec{x} = \vec{0}$ | Input directions that produce $\vec{0}$ output |
| row space(A) | $\mathbb{R}^n$ | r | span{rows of A} | null space(A) | Inferred as orthogonal to null(A) | Input directions that have no component in null(A) |
| col space(A) | $\mathbb{R}^m$ | r | span{cols of A} | ℓ-null space(A) | Geometrically constructed from columns | Output directions that can be produced by A |
| ℓ-null space(A) | $\mathbb{R}^m$ | m−r | all $\vec{y}$ satisfying $A^T\vec{y} = \vec{0}$ or all $\vec{y}$ satisfying $\vec{y}^T A = \vec{0}^T$ | col space(A) | Inferred as orthogonal to col(A) | Output directions that have no component in column space |

Note how the 'deduced' space orthogonal to null(A) becomes row space: suppose

- $B = A^T$ and $B^T = A$
- row space(A) = col space (B) & col space (B) = row space(A)
- vector $\vec{y}$ satisfying $\vec{y}^T B = \vec{0}^T$ is in the ℓ-null space (B)
- same vector $\vec{y}$ also satisfies $B^T\,\vec{y} = \vec{0}$, same as $A\,\vec{y} = \vec{0}$
- same vector $\vec{y}$ belongs to null space (A)
- orthogonal complement of ℓ-null space (B) is column space (B), same as row space (A)

⇓

The set of all vectors orthogonal to null space(A) is precisely the row space(A)

We propose starting with the two primary spaces:

1. column space(A) in $\mathbb{R}^m$ (what A can produce, can be visualized directly)

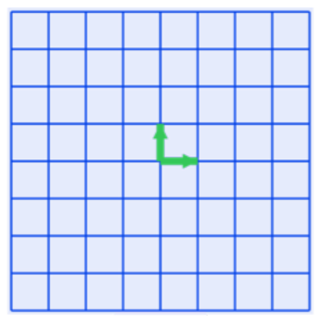
2. null space(A) in $\mathbb{R}^n$ (solutions to $A\vec{x} = 0$, can be computed directly)

The remaining two subspaces are inferred as orthogonal complements to the primary ones

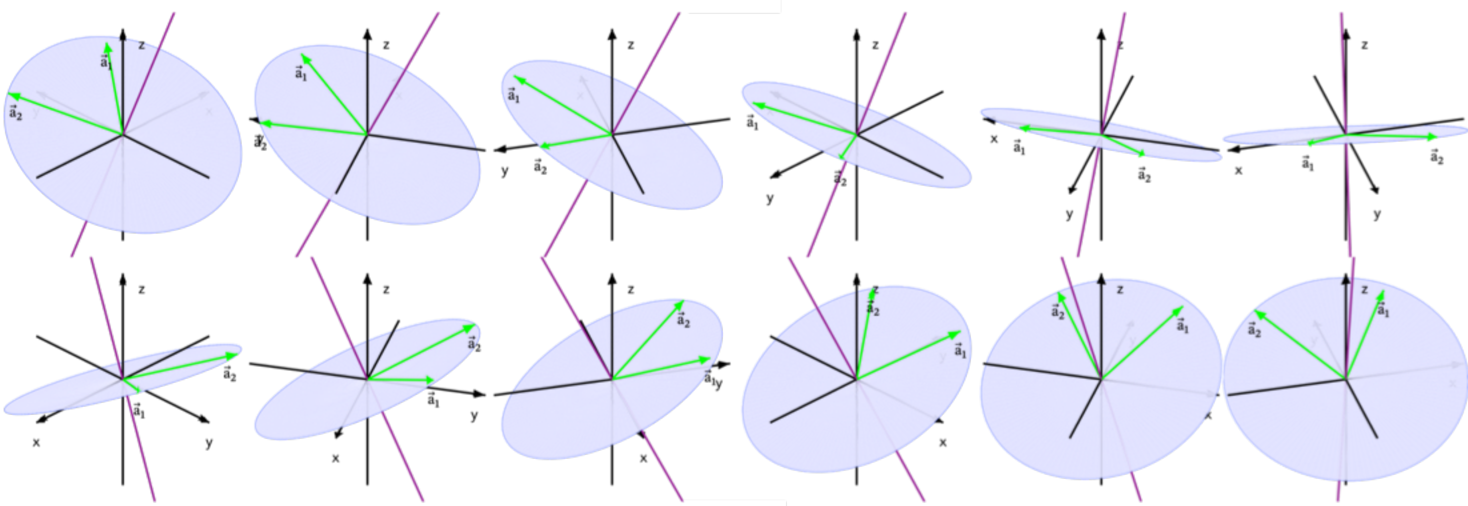
Geometric view of transformations by A, $A^T$ & $A^TA$

We are looking at the same 3×2 matrix $A = [\ \vec{a}_1 \mid \vec{a}_2\ ]$
with linearly independent $\vec{a}_1$ and $\vec{a}_2$ in $\mathbb{R}^3$ and rank(A) = 2

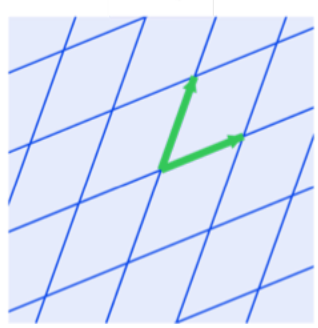
Transformations by A, followed by $A^T$



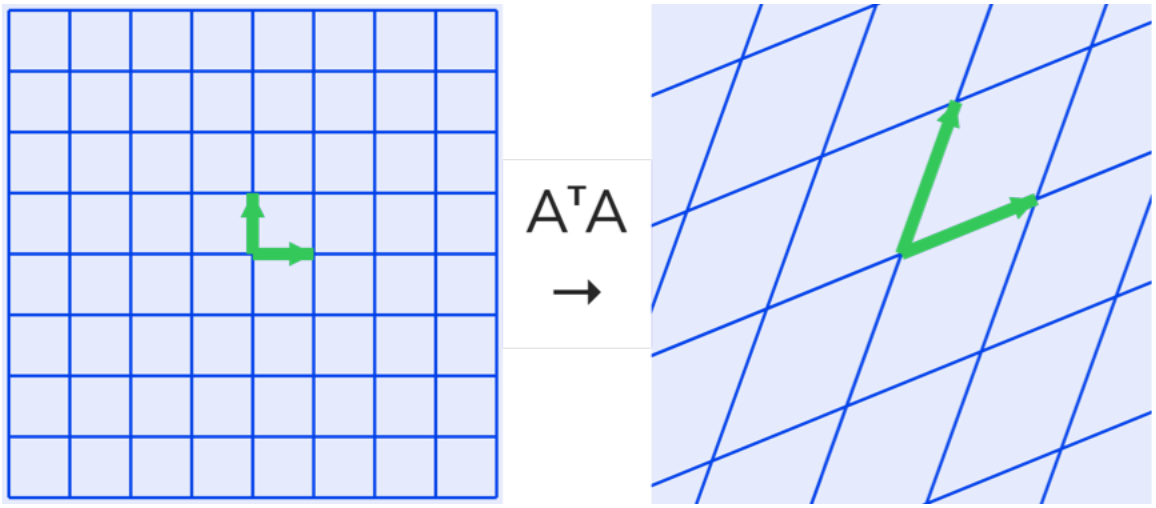
A ⇓



$A^T$ ⇓



Direct transformation by $A^TA$




Note changing relationship between $\vec{a}_1$ & $\vec{a}_2$:

- They start as orthogonal
- After transformation by A, the grid loses squareness
- $A^TA$ brings this non-square grid back to $\mathbb{R}^2$ and

the final $A^TA$ grid looks more distorted than the grid seen in the embedded plane in $\mathbb{R}^3$

This is because $A^TA$ applies the same distortion twice:
once to embed the grid via A and once pull that geometry back via $A^T$

Suppose $\vec{a}_1$ & $\vec{a}_2$ are orthogonal with length = 1:
In this case $A^TA$ is identity and the grid returns to original state

This is how to measure 'grid distortion' by A:
Suppose we have vectors $\vec{x}$ & $\vec{y}$ in the original plane, so their dot product = $\vec{x}^T \cdot \vec{y}$
After transformation by A, dot product becomes
$(A\,\vec{x})^T \cdot (A\,\vec{y}) = \vec{x}^T A^TA \cdot \vec{y} = \vec{x}^T \cdot (A^TA\,\vec{y})$

The four subspaces of $A^T$

We now apply the same construction to the transpose matrix $A^T$:
rows of $A^T$ are columns of A, and columns of $A^T$ are rows of A

| Space | Lives in | Dim | Spanned by | ⊥ Complement |
|---|---|---|---|---|
| row space($A^T$) | $\mathbb{R}^m$ | r | Rows of $A^T$ (Cols of A) | null space($A^T$) |
| null space($A^T$) | $\mathbb{R}^m$ | m−r | all $\vec{x}$ satisfying $A^T \vec{x} = 0$ | row space($A^T$) |
| col space($A^T$) | $\mathbb{R}^n$ | r | Cols of $A^T$ (Rows of A) | ℓ-null space($A^T$) |
| ℓ-null space($A^T$) | $\mathbb{R}^n$ | n−r | all $\vec{b}$ satisfying $A \vec{b} = 0$ | col space($A^T$) |

Connecting the null spaces of A and $A^T$

① null(A) = ℓ-null($A^T$)

Proof:
Let $\vec{x}$ be a vector in $\mathbb{R}^n$

Recall the transpose identity for matrix-vector multiplication:
$(A \vec{x})^T = \vec{x}^T A^T$

Now prove set inclusion in both directions

- If $\vec{x} \in$ null(A), then $A \vec{x} = 0$
Taking transpose gives $(A \vec{x})^T = 0^T = 0$
Using the identity above, $\vec{x}^T A^T = 0$
So $\vec{x} \in$ ℓ-null($A^T$)

- If $\vec{x} \in$ ℓ-null($A^T$), then $\vec{x}^T A^T = 0$

Taking transpose gives $(\vec{x}^T A^T)^T = 0$
But $(\vec{x}^T A^T)^T = A\,\vec{x}$
So $A\,\vec{x} = 0$, hence $\vec{x} \in$ null(A)

Since each set is contained in the other, the two sets are equal

② $\ell$-null(A) = null($A^T$)

Proof:
Let $\vec{y}$ be a vector in $\mathbb{R}^m$

Recall the transpose identity:
$(\vec{y}^T A)^T = A^T \vec{y}$

‣ If $\vec{y} \in \ell$-null(A), then $\vec{y}^T A = 0$
Taking transpose gives $A^T \vec{y} = 0$
So $\vec{y} \in$ null($A^T$)

‣ If $\vec{y} \in$ null($A^T$), then $A^T \vec{y} = 0$
Taking transpose gives $\vec{y}^T A = 0$
So $\vec{y} \in \ell$-null(A)

Thus the two spaces are also equal

---

The four subspaces of $A^TA$

Consider any m×n matrix A

- $A^TA$ is n×n matrix, which maps $\mathbb{R}^n$ to $\mathbb{R}^n$

$\Leftrightarrow$

Domain of $A^T A$ is identical to its codomain

- $A^T A$ is symmetric: $(A^T A)^T = A^T (A^T)^T = A^T A$

① $\text{null}(A^T A) = \text{null}(A)$

Suppose vector $\vec{x} \in$ null space(A)

or

$A\,\vec{x} = \vec{0}$

Then

$(A^T A)\,\vec{x} = A^T(A\,\vec{x}) = \vec{0}$

$\Downarrow$

null space(A) $\subseteq$ null space($A^T A$) (is a subset of)

To prove the reverse inclusion,

suppose vector $\vec{y} \in$ null space($A^T A$)

or

$(A^T A)\,\vec{y} = \vec{0}$

Multiply on the left by $\vec{y}^T$:

$\vec{y}^T (A^T A) \vec{y} = 0$

$\vec{y}^T (A^T A) \vec{y} = (A\,\vec{y})^T (A\,\vec{y}) = \|A\,\vec{y}\|^2$

So $\|A\,\vec{y}\|^2 = 0 \Rightarrow A\,\vec{y} = \vec{0}$

$\Downarrow$

null space($A^T A$) $\subseteq$ null(A)

Together:

null space($A^T A$) = null space(A)

② $A^T A$ is symmetric

$\Downarrow$

- row space($A^T A$) = col space($A^T A$)
- $\ell$-null space($A^T A$) = null space($A^T A$)

$\Downarrow$

four subspaces of A reduce to two

| Space | Lives in | Dim | Spanned by | ⊥ Complement |
| --- | --- | --- | --- | --- |
| row space($A^TA$) = col space($A^TA$) | $\mathbb{R}^n$ | r | rows (same as cols) of $A^TA$ | null space($A^TA$) = ℓ-null space($A^TA$) |
| null space($A^TA$) = ℓ-null space($A^TA$) = null space(A) | $\mathbb{R}^n$ | n−r | all $\vec{x}$ satisfying $(A^TA)\,\vec{x} = \vec{0}$ or $\vec{x}^T\,(A^TA) = \vec{0}^T$ | row space($A^TA$) = col space($A^TA$) |

Suppose A is a 'tall' matrix with m>n and full column rank

⇓

$A^TA$ has full rank = n

⇓

$(A^TA)\,\vec{x} = A^T\vec{b}$ has a single solution

Importance of this concept will be illustrated on the next page

---

Why do we need this abstract concept?

Suppose we have an inconsistent system of equations $A\,\vec{x} = \vec{b}$ where

- $\vec{b}$ cannot be presented as a linear combination of columns of A

  ⇔

  there is no $\vec{x}$ that satisfies the equation
- A has full column rank ⇔ $(A^TA)^{-1}$ is defined

So we set the goal of finding the closest $\vec{x}$

Recall from the page on the four subspaces that every $\vec{b}$ in $\mathbb{R}^m$ (or co-domain of A) can be written as sum of $(\vec{b}_{\parallel} \in \text{col}(A)) + (\vec{b}_{\perp} \in \ell\text{-null}(A))$

or

$$\vec{b} = \vec{b}_{\parallel} + \vec{b}_{\perp}$$

Also recall from previous page that $\ell$-null(A) = null($A^T$)

$$\Downarrow$$

$$\vec{b}_{\perp} \in \text{null}(A^T)$$

or, by definition,

$$A^T \vec{b}_{\perp} = \vec{0}$$

To define the goal of finding the closest solution,
we want to solve for such vector that

- belongs to col(A)
- is closest to $\vec{b}$

The vector we are looking for happens to be $\vec{b}_{\parallel}$,
which takes us to solving

$$A\,\vec{x} = \vec{b}_{\parallel}$$

or, equivalently,

$$A\,\vec{x} = \vec{b} - \vec{b}_{\perp}$$

We can multiply both sides by $A^T$:

$$A^TA\,\vec{x} = A^T(\vec{b} - \vec{b}_{\perp}) = A^T\vec{b} - A^T\,\vec{b}_{\perp} = A^T\vec{b} - \vec{0}$$

A has full column rank, therefore $(A^TA)^{-1}$ is defined, and
we can multiply both sides by $(A^TA)^{-1}$:

$$(A^TA)^{-1}\,A^TA\,\vec{x} = (A^TA)^{-1}\,A^T\,\vec{b}$$

or

$$\vec{x} = (A^TA)^{-1}\,A^T\,\vec{b}$$

Image for (m=3)×(n=2) matrix A with linearly independent columns: r=2
The blue plane is col(A) = span($\vec{a}_1$, $\vec{a}_2$)
The 'unreachable' red vector $\vec{b}$ is outside span of the columns

We decompose $\vec{b}$ as $\vec{b} = \vec{b}_{\|}$(teal) + $\vec{b}_{\perp}$(magenta)
• $\vec{b}_{\|} \in$ col(A)
• $\vec{b}_{\perp} \perp$ col(A)

So $\vec{b}_{\|}$ is the closest vector to $\vec{b}$ that lies in col(A)
and the least squares condition is $A\,\vec{x} = \vec{b}_{\|}$

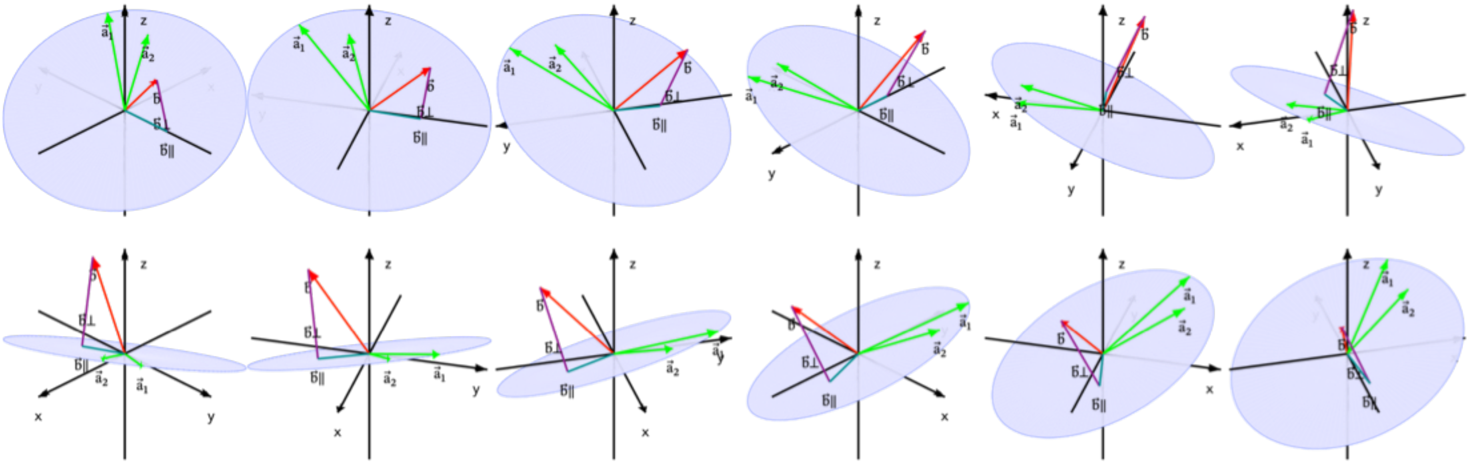
Note that the image shows the following structures:
1. Entire $\mathbb{R}^3$
2. Plane as subspace of $\mathbb{R}^3$, or $\mathbb{R}^2$ space in its own right with its own coordinate system
• Matrix A inputs vectors represented in coordinates of $\mathbb{R}^2$
and represents them in coordinates of $\mathbb{R}^3$
⇔

maps $\mathbb{R}^2 \rightarrow \mathbb{R}^3$

• Matrix $A^T$ inputs vectors represented in coordinates of $\mathbb{R}^3$
and represents them in coordinates of $\mathbb{R}^2$
⇔

maps $\mathbb{R}^3 \rightarrow \mathbb{R}^2$

⇓
left-multiplying by $A^T$ takes $\vec{b}$ and outputs a 2D vector
whose entries are dot products with directions $\vec{a}_1$ and $\vec{a}_2$
These numbers describe how $\vec{b}_{\|}$ aligns with the spanning directions of the plane
They are not yet the expansion coefficients:
The expansion coefficients $\vec{x}$ are obtained by solving $A^TA\,\vec{x} = A^T\vec{b}$ as demonstrated above




And we leave the n-dimensional geometric picture to the reader's imagination... 🙂

---

## Tying it all together

| Condition | Equivalent to |
|---|---|
| null space = $\{\vec{0}\}$ | • $A\vec{x} = \vec{0}$ has only trivial solution<br>• no free variables<br>• if $A\vec{x}1 = A\vec{x}2$ then $\vec{x}1 = \vec{x}2$<br>• transformation is one-to-one |
| null space ≠ $\{\vec{0}\}$ | • $A\vec{x} = \vec{0}$ has nontrivial solutions<br>• free variables exist<br>• if $A\vec{x}1 = A\vec{x}2$ then $\vec{x}1$ and $\vec{x}2$ may be different<br>• transformation is not one-to-one |
| ℓ-null space = $\{\vec{0}\}$ | • column space = codomain<br>• every $\vec{b}$ in codomain has a solution to $A\vec{x} = \vec{b}$<br>• no impossible outputs<br>• transformation is onto |
| ℓ-null space ≠ $\{\vec{0}\}$ | • column space ≠ codomain<br>• some $\vec{b}$ in codomain have no solution to $A\vec{x} = \vec{b}$<br>• outputs with nonzero ℓ-null component cannot be produced<br>• transformation is not onto |