① rank(AB) ≤ rank(B)

② rank(AB) = rank(B) − dim(col(B) ∩ null(A))

① Will use a simple example where B = [$\vec{b}_1$ | $\vec{b}_2$ | $s\vec{b}_1 + t\vec{b}_2$]

↓

AB = [$A\vec{b}_1$ | $A\vec{b}_2$ | $sA\vec{b}_1 + tA\vec{b}_2$]

From here, we see that column 3 of AB is a linear combination of columns 1 & 2

And in general,

any linearly dependent columns of B

produce linearly dependent columns of AB

↓

Rank(AB) ≤ Rank(B)

② The converse:

'any linearly independent $\vec{b}$ will produce a linearly independent column of AB'

is NOT true as shown in the numerical example below where

• B has rank=2 with linearly independent $\vec{b}_1$ & $\vec{b}_2$ & $\vec{b}_3 = \vec{b}_1 + \vec{b}_2$

• C has rank=1

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Notice that $A\,\vec{b}_2 = \vec{0}$

or equivalently

$\vec{b}_2$ belongs to the null space of A, or the combined 'blind spot' of rows of A

↓

$(AB)_2 = \vec{0}$

↓

rank (AB) drops by 1

In general:

the rank loss equals the dimension of the intersection

col(B) ∩ null(A)

↓

$$\text{rank}(AB) \leq \text{rank}(B) - \dim\big(\text{col}(B) \cap \text{null}(A)\big)$$

③ It remains to be proven that if

- B has full column rank r
- col(B) ∩ null(A) = $\{\vec{0}\}$

↓

rank(AB) = rank(B)

Proof by contradiction is used in this case

Let $\vec{b}_1, \ldots, \vec{b}_r$ be r linearly independent columns of B

Assume rank(AB) < rank(B) & B has full column rank

this implies

$A\vec{b}_1 \ldots A\vec{b}_r$ are linearly dependent

Apply the definition of linear dependence:

$$s_1 A\vec{b}_1 + \ldots + s_r A\vec{b}_r = \vec{0}$$

Then

$$A(s_1\vec{b}_1 + \ldots + s_r\vec{b}_r) = \vec{0}$$

So the vector

$$\vec{v} = s_1\vec{b}_1 + \ldots + s_r\vec{b}_r$$

belongs to col(B) and also to null(A)

This contradicts our asumption of col(B) ∩ null(A) = $\{\vec{0}\}$

and proves that, in this particular case,

rank(AB) = rank(B)

Conclusion

From ② we have

$$\text{rank}(AB) \leq \text{rank}(B) - \dim\big(\text{col}(B) \cap \text{null}(A)\big)$$

Proof ③ shows that
whenever col(B) ∩ null(A) = $\{\vec{0}\}$,
no additional rank loss occurs
↓
inequality ② becomes equality

$$\text{rank}(AB) = \text{rank}(B) - \dim\left(\text{col}(B) \cap \text{null}(A)\right)$$

Alternative view of the same concept of
rank loss = overlap between col(B) & null(A):

- B is a set of vectors $\vec{b}$ written as columns
- A is a set of instructions for vectors of B
- A rotates and deforms column space(B)
- Those vectors $\vec{b}$ that lie in the null space(A) map to $\vec{0}$

Visual illustration of

$$\text{rank}(AB) = \text{rank}(B) - \dim\left(\text{col}(B) \cap \text{null}(A)\right)$$

This images below show decomposition of
domain(A) = codomain(B) = $\mathbb{R}^n$
①
- row space(A) in blue
- null space(A) in grey,

as well as
②

- col space(B) in green
- ℓ-null space(B) in gray

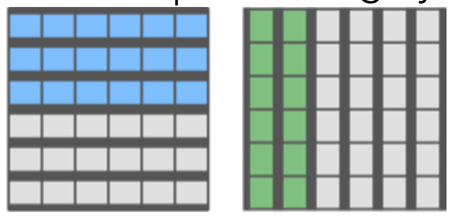


Image 1: col space(B) is contained within row space(A)

↓

AB has 2 linearly independent columns

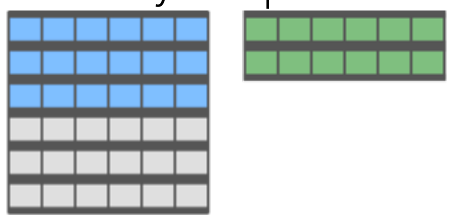


Image 2: col space(B) straddles row space(A) & null space(A)

↓

AB has 1 linearly independent column

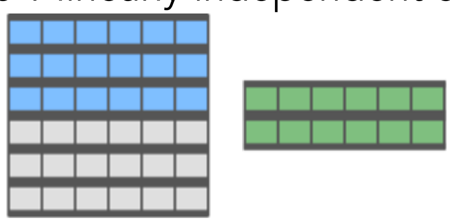
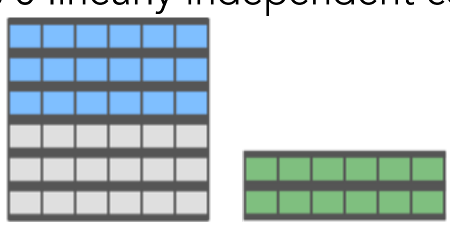


Image 3: col space(B) is contained within null space(A)

↓

AB has 0 linearly independent columns



Extension of

$$\operatorname{rank}(AB) = \operatorname{rank}(B) - \dim\big(\operatorname{col}(B) \cap \operatorname{null}(A)\big)$$

For three matrices A, B, C

• C has columns that must be seen by B

$$\downarrow$$

$$\operatorname{rank}(BC) = \operatorname{rank}(C) - \dim\big(\operatorname{col}(C) \cap \operatorname{null}(B)\big)$$

• BC has columns that must be seen by A

$$\downarrow$$

$$\operatorname{rank}(ABC) = \operatorname{rank}(BC) - \dim\big(\operatorname{col}(BC) \cap \operatorname{null}(A)\big) =$$

$$\operatorname{rank}(C) - \dim\big(\operatorname{col}(C) \cap \operatorname{null}(B)\big) - \dim\big(\operatorname{col}(BC) \cap \operatorname{null}(A)\big)$$

column space and null space propagate from right to left in the following fashion:

• column space is 'filtered': $\operatorname{col}(AB) \subseteq \operatorname{col}(A)$

• null space 'grows' from B: $\operatorname{null}(B) \subseteq \operatorname{null}(AB)$

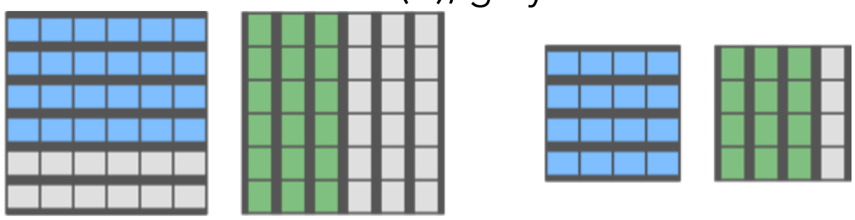
Image below shows serial multiplication

$A(m\times 6) \times B(6\times 4) \times C(4\times n)$

with spaces in reading order (left → right), while multiplication propagates right → left

• square 4: $\mathbb{R}^4$ (where B meets C) partitioned into
  - col(C), green
  - ℓ-null(C), gray
• square 3: $\mathbb{R}^4$ (where B meets C) partitioned into

    - row(B), blue
    - null(B), gray
- square 2: $\mathbb{R}^6$ (where A meets BC) partitioned into
    - col(BC), green
    - ℓ-null(BC), gray
- square 1: $\mathbb{R}^6$ (where A meets BC) partitioned into
    - row(A), blue
    - null(A), gray



In the example above,

- dim(null(B) = 0)

$$\downarrow$$

rank(BC) = rank(C) = 3

- rank(ABC) is determined by the overlap between col(BC) and null(A)

$$\text{rank}(AB) \leq \text{rank}(A) \ \&$$

$$\text{rank}(AB) = \text{rank}(A) - \dim\big(\,\text{row}(A) \cap \ell\text{-null}(B)\,\big)$$

We already showed:

$$\text{rank}(AB) = \text{rank}(B) - \dim\big(\,\text{col}(B) \cap \text{null}(A)\,\big)$$

Now the derive the parallel formula that starts with rank(A)

Let $C = AB$

Define:

- $M = A^T$
- $N = B^T$
- $O = C^T$

Transpose reverses multiplication order:

$O = C^T = (AB)^T = B^T A^T = NM$

Apply the proved equality to NM:

$$\text{rank}(NM) = \text{rank}(M) - \dim\big(\text{col}(M) \cap \text{null}(N)\big)$$

Convert each term back:

- $\text{rank}(NM) = \text{rank}(O) = \text{rank}(C) = \text{rank}(AB)$
- $\text{rank}(M) = \text{rank}(A^T) = \text{rank}(A)$
- $\text{col}(M) = \text{col}(A^T) = \text{row}(A)$
- $\text{null}(N) = \text{null}(B^T) = \ell\text{-null}(B)$

Therefore:

$$\text{rank}(AB) = \text{rank}(A) - \dim\big(\text{row}(A) \cap \ell\text{-null}(B)\big)$$

This demonstrates a general proof technique:
if we proved a property for the right factor in AB,
we can use transpose multiplication rule
to find equivalent property for the left factor

Multiplication by invertible matrix

① Suppose A is an $n \times n$ invertible matrix with rank $r=n$

Will use the previously derived identity

$$\text{rank}(AB) = \text{rank}(B) - \dim\left(\text{col}(B) \cap \text{null}(A)\right)$$

$$\dim\left(\text{null}(A) = 0\right)$$

$$\downarrow$$

$$\text{rank}(AB) = \text{rank}(B)$$

Now suppose B is invertible

Will use the transpose proof method used on the previous page:

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

Since B^T is invertible,

$$\text{rank}(B^T A^T) = \text{rank}(A^T)$$

Transpose preserves rank:

$$\text{rank}\left((AB)^T\right) = \text{rank}(AB)$$

Therefore:

$$\text{rank}(AB) = \text{rank}(A)$$

Geometric view of rank (AB)

Here, we will provide a simple illustration for

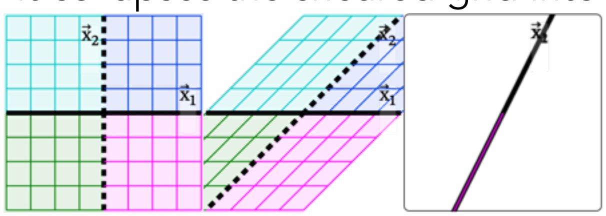
$$\text{rank}(AB) \le \min\big(\text{rank}(A), \text{rank}(B)\big)$$

For this illustration, will use the following matrices:

$$\text{rank-1 } A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{rank-2 } B \text{ (shear)} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

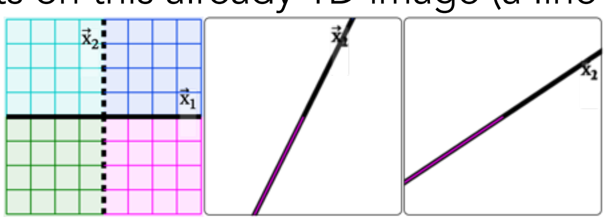
Transformation AB:

- B shears the grid
- A has rank 1, so it collapses the sheared grid into a single direction



Transformation BA:

- A has rank 1, so it collapses the grid into a single direction
- B then acts on this already 1D image (a line stays a line)



Notice:

Two of the intermediate images (the collapsed line after A) are identical in AB and BA:

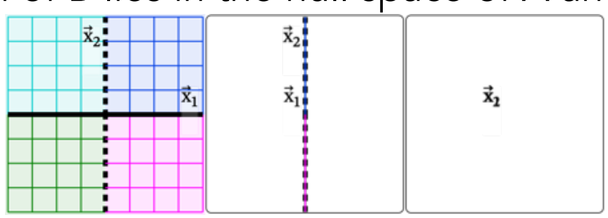
this is the direction of column space of A

Geometric view of rank (AB) $<$ min $\big(\text{rank}(A), \text{rank}(B)\big)$

Will use 2×2 example

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

- Transformation by B projects everything on the only column of B: $(0, 1)^T$
  - The only column of B lies in the null space of A and is projected to $\vec{0}$



rank($A^T A$) = rank(A)

• start with the identity we showed before:

$$\text{rank}(MN) = \text{rank}(N) - \dim\big(\text{col}(N) \cap \text{null}(M)\big)$$

• substitute A^T for M

• substitute A for N

• obtain: $\text{rank}(A^T A) = \text{rank}(A) - \dim\big(\text{col}(A) \cap \text{null}(A^T)\big)$ (eq 1)

• recall $\text{null}(B) = \ell\text{-null}(B^T)$

↓

• $\text{null}(A^T) = \ell\text{-null}(A)$

• equation 1 becomes:

$$\text{rank}(A^T A) = \text{rank}(A) - \dim\big(\text{col}(A) \cap \ell\text{-null}(A)\big)$$

• recall definition of column space and ℓ-null space:

col(A) is orthogonal to ℓ-null(A)

↓

$\text{col}(A) \cap \ell\text{-null}(A) = \vec{0}$

↓

$\text{rank}(A^T A) = \text{rank}(A)$

Another way to think about this concept:

• A produces vectors in the col(A) and orthogonal to ℓ-null(A)

• A^T can only remove directions in ℓ-null(A)

↓

no directions are removed

This property allows $(A^TA)^{-1}$ to be used for Least Squares solutions

This proof is essentially a restatement of the classical argument based on the identity $\vec{x}^TA^TA\vec{x} = \|A\vec{x}\|^2$

Both rely on the same geometric fact that

$\text{col}(A) \cap \ell\text{-null}(A) = \{\vec{0}\}$

rank(A + B) ≤
rank(A) + rank(B) − dim(col(A) ∩ col(B)

For any m×n A & B,
$(A + B)\vec{x} = A\vec{x} + B\vec{x}$
or equivalently
Matrix addition represents a linear combination of transformations

Will use a simple numerical example of rank(A + B) = rank(A) + rank (B):

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

for comparison, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Suppose
- A has r(A) linearly independent columns
- B has r(B) linearly independent columns

Let U = col(A), V = col(B)
Since U and V share q independent directions,
- dim(U ∩ V) = q
- col(A + B) ⊆ U + V

Therefore
- rank(A + B) ≤ dim(U + V)

- rank(A + B) ≤ r(A) + r(B) − q

If U = col(A) and V = col(B) share q independent directions,
then q directions of V already lie inside U
Adding B does not introduce those q directions again

Below is an example of rank(C + D) < rank (C) + rank(D)
due to shared dimension q = 1

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{rank=1} \qquad D = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{rank=1}$$

$$C + D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{rank=0}$$

Comparison between AB & A + B

Operation	AB	A + B
Mental Model	Filter	Combine
Propagation of directions	Directions from B are filtered by A	Directions from A and B accumulate unless overlapping
Rank Behavior	$\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$	$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$
Rank loss mechanism	Non-trivial intersection $\text{col}(B) \cap \text{null}(A)$	• Overlap of column spaces $\text{col}(A) \cap \text{col}(B)$ • Cancellation of columns