Vv417 Lecture 14

Jing Liu

UM-SJTU Joint Institute

October 24, 2019

Theorem

For any matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$,

$$\operatorname{rank}(\mathbf{A})$$

is equal to the maximum number of linearly independent columns of $\mathbf{A}$.

Proof

- $\operatorname{rank}(\mathbf{A})$ is defined as the number of pivots in $\operatorname{rref}(\mathbf{A})$. So let $\mathbf{E} \in \mathbb{R}^{m\times m}$

$$\mathbf{EA} = \mathbf{R}, \qquad \text{where} \qquad \mathbf{R} = \operatorname{rref}(\mathbf{A})$$

- Thus we only need to show the number of pivot columns in $\mathbf{R}$ is equal to the maximum number of linearly independent columns of $\mathbf{A}$, that is, to show

$$\operatorname{rank}(\mathbf{A}) = \dim\big(\operatorname{col}(\mathbf{A})\big)$$

- We apply column operations, which will not alter the column space, so that a non-pivot column $\mathbf{r}_k$ is to the right of all j pivot columns

$$\mathbf{R} = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_j & \cdots & \mathbf{r}_k & \cdots\end{bmatrix}$$

Proof

- Note pivot columns can only be the standard vectors,

$$\mathbf{e}_j$$

and $\mathbf{r}_k$ can only be a linear combination of those pivot columns, otherwise it would be a pivot column, so

$$\text{rank}(\mathbf{A}) = \dim\big(\,\text{col}\,(\mathbf{R})\,\big)$$

- Consider any selection of columns of $\mathbf{A}$,

$$\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots \alpha_n\mathbf{a}_n = \mathbf{0}$$

$$\alpha_1\mathbf{E}\mathbf{a}_1 + \alpha_2\mathbf{E}\mathbf{a}_2 + \cdots \alpha_n\mathbf{E}\mathbf{a}_n = \mathbf{E}\mathbf{0}$$

$$\alpha_1\mathbf{r}_1 + \alpha_2\mathbf{r}_2 + \cdots \alpha_n\mathbf{r}_n = \mathbf{0}, \qquad \text{since } \mathbf{EA} = \mathbf{R},$$

where $\mathbf{r}_i$ are the corresponding columns vectors in $\mathbf{R}$ of $\mathbf{a}_i$ in $\mathbf{R}$.

Proof

- Now consider the linear independence of any columns $\mathbf{r}_i$ of $\mathbf{R}$,

$$\beta_1 \mathbf{r}_1 + \beta_2 \mathbf{r}_2 + \cdots \beta_n \mathbf{r}_n = \mathbf{0}$$

$$\beta_1 \mathbf{E}^{-1}\mathbf{a}_1 + \beta_2 \mathbf{E}^{-1}\mathbf{a}_2 + \cdots \beta_n \mathbf{E}^{-1}\mathbf{a}_n = \mathbf{E}^{-1}\mathbf{0}$$

$$\beta_1 \mathbf{a}_1 + \beta_2 \mathbf{a}_2 + \cdots \beta_n \mathbf{a}_n = \mathbf{0}, \qquad \text{since } \mathbf{A} = \mathbf{E}^{-1}\mathbf{R},$$

  where $\mathbf{a}_i$ are the corresponding columns vectors in $\mathbf{A}$ of $\mathbf{r}_i$ in $\mathbf{R}$.

- This shows a set of columns of $\mathbf{A}$ is linearly independent if and only if the corresponding set of columns in $\mathbf{R}$ is linearly independent, hence

$$\dim\big(\operatorname{col}(\mathbf{A})\big) = \dim\big(\operatorname{col}(\mathbf{R})\big)$$

- Therefore,

$$\operatorname{rank}(\mathbf{A}) = \dim\big(\operatorname{col}(\mathbf{A})\big) \quad \square$$

Theorem

The row space and the column space of a matrix $\mathbf{A}$ have the same dimension.

Proof

- We have effectively shown

$$\operatorname{rank}(\mathbf{A}) = \dim(\operatorname{col}(\mathbf{R})) = \dim(\operatorname{col}(\mathbf{A})), \qquad \text{where } \mathbf{R} = \operatorname{rref}(\mathbf{A})$$

- It is clear that the column and row space must share the same dimension

$$\dim(\operatorname{col}(\mathbf{R})) = \dim(\operatorname{row}(\mathbf{R}))$$

- Since we have shown row operations do not alter the row space

$$\begin{aligned} \operatorname{row}(\mathbf{A}) = \operatorname{row}(\mathbf{R}) \implies \dim(\operatorname{row}(\mathbf{A})) &= \dim(\operatorname{row}(\mathbf{R})) \\ &= \operatorname{rank}(\mathbf{A}) \\ &= \dim(\operatorname{col}(\mathbf{A})) \quad \square \end{aligned}$$

- In the light of the last two theorems, the rank of a matrix $\mathbf{A}$ is essentially the common dimension of the row space and column space of $\mathbf{A}$.

Definition

If $\mathcal{X}$ and $\mathcal{Y}$ are subspaces of a vector space $\mathcal{V}$, then the sum of $\mathcal{X}$ and $\mathcal{Y}$ is defined to be the set of all possible sums of vectors from $\mathcal{X}$ with vectors from $\mathcal{Y}$. That is,

$$\mathcal{X} + \mathcal{Y} = \left\{ \mathbf{x} + \mathbf{y} \mid \mathbf{x} \in \mathcal{X} \text{ and } \mathbf{y} \in \mathcal{Y} \right\}$$

Theorem

If $\mathcal{X}$ and $\mathcal{Y}$ are subspaces of a vector space $\mathcal{V}$, then

$$\dim(\mathcal{X} + \mathcal{Y}) = \dim \mathcal{X} + \dim \mathcal{Y} - \dim (\mathcal{X} \cap \mathcal{Y})$$

Exercise

Use the above theorem to show

$$\operatorname{rank}(\mathbf{A} + \mathbf{B}) \leq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B}), \qquad \text{for any matrices } \mathbf{A}, \mathbf{B} \in \mathcal{M}_{m \times n}.$$

Solution

- If $\mathbf{b} \in \mathrm{col}(\mathbf{A}+\mathbf{B})$, then there exists a vector $\mathbf{x}$ such that

$$(\mathbf{A}+\mathbf{B})\mathbf{x} = \mathbf{b}$$

$$\underbrace{\mathbf{A}\mathbf{x}}_{\mathrm{col}(\mathbf{A})} + \underbrace{\mathbf{B}\mathbf{x}}_{\mathrm{col}(\mathbf{B})} = \mathbf{b} \in \mathrm{col}(\mathbf{A}) + \mathrm{col}(\mathbf{B})$$

$$\Longrightarrow \mathrm{col}(\mathbf{A}+\mathbf{B}) \subset \mathrm{col}(\mathbf{A}) + \mathrm{col}(\mathbf{B})$$

- Invoke the theorem on $\boxed{\textbf{??}}$ and $\boxed{6}$, then the dimensions must satisfy

$$\begin{aligned}
\dim\Big(\mathrm{col}(\mathbf{A}+\mathbf{B})\Big) &\le \dim\Big(\mathrm{col}(\mathbf{A}) + \mathrm{col}(\mathbf{B})\Big) \\
&= \dim\Big(\mathrm{col}(\mathbf{A})\Big) + \dim\Big(\mathrm{col}(\mathbf{B})\Big) - \dim\Big(\mathrm{col}(\mathbf{A}) \cap \mathrm{col}(\mathbf{B})\Big) \\
&\le \dim\Big(\mathrm{col}(\mathbf{A})\Big) + \dim\Big(\mathrm{col}(\mathbf{B})\Big)
\end{aligned}$$

$$\Longrightarrow \mathrm{rank}(\mathbf{A}+\mathbf{B}) \le \mathrm{rank}(\mathbf{A}) + \mathrm{rank}(\mathbf{B})$$

Theorem

If $\mathbf{A}$ is a matrix of $m \times n$ and $\mathbf{B}$ is a matrix of $n \times r$, then

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

Proof

- The main idea is to find dimensions of each vectors spaces involved

$$\text{col}(\mathbf{AB}) \qquad \text{col}(\mathbf{B}) \qquad \text{null}(\mathbf{A}) \cap \text{col}(\mathbf{B})$$

  and confirm the the equality is true.
- We do that by consider the basis for each vector space involved.
- Let $\mathcal{S} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_s\}$ be a basis for

$$\text{null}(\mathbf{A}) \cap \text{col}(\mathbf{B})$$

Proof

- Assume $\mathrm{rank}(\mathbf{B}) = \dim(\mathrm{col}(\mathbf{B})) = s + t$, then there exist an extension set,

$$\mathcal{S}_{\text{ext}} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_t\}$$

  such that

$$\mathcal{B} = \{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{z}_1, \ldots, \mathbf{z}_t\}$$

  is a basis for $\mathrm{col}(\mathbf{B})$.

- We need to show

$$\mathrm{rank}(\mathbf{AB}) = \dim(\mathrm{col}(\mathbf{AB})) = t$$

- This is done by showing

$$\mathcal{W} = \{\mathbf{Az}_1, \mathbf{Az}_2, \ldots, \mathbf{Az}_t\}$$

  is a basis for $\mathrm{col}(\mathbf{AB})$.

Proof

- $\mathcal{W}$ spans $\text{col}(\mathbf{AB})$ if $\mathbf{b}$ can be represented as a linear combination of $\mathbf{Az}_j$'s,

  where $\quad \mathbf{b} = (\mathbf{AB})\mathbf{y} \quad$ for any vector $\mathbf{y}$ that makes the product defined.

- Since $\mathbf{By} \in \text{col}(\mathbf{B})$, there must exist α_i's and β_j's such that

$$\mathbf{By} = \sum_{i=1}^{s} \alpha_i \mathbf{x}_i + \sum_{j=1}^{t} \beta_j \mathbf{z}_j$$

- So

$$\mathbf{b} = (\mathbf{AB})\mathbf{y} = \sum_{i=1}^{s} \alpha_i \mathbf{A}\mathbf{x}_i + \sum_{j=1}^{t} \beta_j \mathbf{A}\mathbf{z}_j$$

- Since $\mathbf{x}_i$'s form a basis for $\text{null}(\mathbf{A}) \cap \text{col}(\mathbf{B})$.

Proof

- Since $\mathbf{x}_i$'s form a basis for $\mathrm{null}(\mathbf{A}) \cap \mathrm{col}(\mathbf{B})$, thus

$$\sum_{i=1}^{s} \alpha_i \mathbf{A}\mathbf{x}_i = \mathbf{0} \implies \mathbf{b} = \sum_{j=1}^{t} \beta_j \mathbf{A}\mathbf{z}_j$$

$$\implies \mathcal{W} \text{ spans } \mathrm{col}(\mathbf{AB}).$$

- $\mathcal{W}$ is also linearly independent for if

$$\mathbf{0} = \sum_{j=1}^{t} \gamma_j \mathbf{A}\mathbf{z}_j = \mathbf{A} \sum_{j=1}^{t} \gamma_j \mathbf{z}_j$$

- It is clear that $\sum_{j=1}^{t} \gamma_j \mathbf{z}_j \in \mathrm{null}(\mathbf{A})$, but $\sum_{j=1}^{t} \gamma_j \mathbf{z}_j \in \mathrm{col}(\mathbf{B})$ for $\mathbf{z}_j \in \mathcal{B}$,

Proof

- Therefore $$\sum_{j=1}^{t} \gamma_j \mathbf{z}_j \in \mathrm{null}(\mathbf{A}) \cap \mathrm{col}(\mathbf{B})$$
- So there must be λ_i such that

$$\sum_{j=1}^{t} \gamma_j \mathbf{z}_j = \sum_{i=1}^{s} \lambda_i \mathbf{x}_i \implies \sum_{j=1}^{t} \gamma_j \mathbf{z}_j - \sum_{i=1}^{s} \lambda_i \mathbf{x}_i = \mathbf{0}$$

- γ_j's and λ_i's must all be zero because

$$\mathcal{B} = \{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{z}_1, \ldots, \mathbf{z}_t\}$$

  is a basis for $\mathrm{col}(\mathbf{B})$ and thus linearly independent.
- So $\mathcal{W}$ is a basis for $\mathrm{col}(\mathbf{AB})$, and we have all the bases and their dimensions.

$$\mathrm{rank}(\mathbf{B}) = s + t = \dim\Big(\mathrm{null}(\mathbf{A}) \cap \mathrm{col}(\mathbf{B})\Big) + \mathrm{rank}(\mathbf{AB}) \quad \square$$

Definition

The dimension of the null space of $\mathbf{A}$ is called the nullity of $\mathbf{A}$ and is denoted by

$$\mathrm{nullity}(\mathbf{A})$$

- The fundamental relationship between the rank and the nullity is given by

Dimension theorem for matrices

If $\mathbf{A}$ is a matrix with n columns, then

$$\mathrm{rank}(\mathbf{A}) + \mathrm{nullity}(\mathbf{A}) = n$$

- Loosely speaking, it is kind of conservation law, it states as the amount of content in $\mathrm{col}(\mathbf{A})$ increases, the amount in $\mathrm{null}(\mathbf{A})$ drops, and vice versa.
- Any matrix $\mathbf{A}_{m\times n}$ seems to define some division of the vector space $\mathbb{R}^n$.

Proof

- Since $\mathbf{A}$ has n columns, so the corresponding linear system

$$\mathbf{Ax} = \mathbf{0} \quad \text{has } n \text{ unknowns.}$$

- These unknowns fall into two distinct categories:
  1. Leading variables, let's assume we have p leading variables.
  2. Free variables, then there must be $n - p$ free variables.
- p is the same as the number of pivots in $\mathrm{ref}(\mathbf{A})$

$$p = \mathrm{rank}(\mathbf{A})$$

- $n - p$ is the number of parameters in the general solution of $\mathbf{Ax} = \mathbf{0}$,

$$n - p = \dim(\mathrm{null}(\mathbf{A})) = \mathrm{nullity}(\mathbf{A})$$

- Thus

$$\mathrm{rank}(\mathbf{A}) + \mathrm{nullity}(\mathbf{A}) = n \quad \square$$

Q: Are there any relationships among the solutions of a linear system

$$\mathbf{Ax} = \mathbf{b}$$

and the row space, column space, and null space of the coefficient matrix $\mathbf{A}$?

- Using matrix multiplication by columns $\mathbf{c}_i$ of $\mathbf{A}$, the product $\mathbf{Ax}$,

$$\mathbf{Ax} = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n$$

- A linear system, $\mathbf{Ax} = \mathbf{b}$, of m equations in n unknowns can be written as

$$x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \mathbf{b}$$

- From which we conclude $\mathbf{Ax} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is a linear combination of the column vectors of $\mathbf{A}$.

Theorem

A system of linear equations

$$\mathbf{Ax} = \mathbf{b}$$

is consistent if and only if $\mathbf{b}$ is in the column space of $\mathbf{A}$.

Theorem

The general solution of a consistent linear system $\mathbf{Ax} = \mathbf{b}$ can be obtained by adding any particular solution of $\mathbf{Ax} = \mathbf{b}$ to the general solution of $\mathbf{Ax} = \mathbf{0}$.

Proof

- Let $\mathbf{x}_\mathsf{p}$ be any particular solution of

$$\mathbf{Ax} = \mathbf{b},$$

- Let $\mathbf{x}_\mathsf{c} \in \operatorname{null}(\mathbf{A})$ denote a solution to the corresponding equation

$$\mathbf{Ax} = \mathbf{0},$$

- Let the following denote

$$\mathbf{x}_\mathsf{p} + \operatorname{null}(\mathbf{A})$$

the set of all vectors that resulted by adding $\mathbf{x}_\mathsf{p}$ to each vector in $\operatorname{null}(\mathbf{A})$

Proof

- We must show if
$$\mathbf{x} \in \mathbf{x}_\mathsf{p} + \mathrm{null}(\mathbf{A}),$$
then $\mathbf{x}$ is a solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, and conversely that every solution of
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$
is in $\mathbf{x}_\mathsf{p} + \mathrm{null}(\mathbf{A})$.
- Assume $\mathbf{x}$ is in $\mathbf{x}_\mathsf{p} + \mathrm{null}(\mathbf{A})$, then
$$\begin{aligned}\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{x}_\mathsf{p} + \mathbf{x}_\mathsf{c}) = \mathbf{A}\mathbf{x}_\mathsf{p} + \mathbf{A}\mathbf{x}_\mathsf{c} = \mathbf{b} + \mathbf{0} = \mathbf{b}& \\ \implies \mathbf{x} \text{ is a solution of } \mathbf{A}\mathbf{x} = \mathbf{b}.&\end{aligned}$$
- Assume $\mathbf{x}$ is the general solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, and $\mathbf{x}_\mathsf{p}$ is any particular $\mathbf{x}$
$$\begin{aligned}\mathbf{A}(\mathbf{x} - \mathbf{x}_\mathsf{p}) = \mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_\mathsf{p} = \mathbf{b} + \mathbf{b} = \mathbf{0} &\implies (\mathbf{x} - \mathbf{x}_\mathsf{p}) \in \mathrm{null}(\mathbf{A}) \\ &\implies \mathbf{x} \in \mathbf{x}_\mathsf{p} + \mathrm{null}(\mathbf{A}) \quad \square\end{aligned}$$

Theorem

If $\mathbf{x}_\mathsf{p}$ is any solution of a consistent linear system $\mathbf{Ax} = \mathbf{b}$, and if

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$$

is a basis for the null space of $\mathbf{A}$, then

every solution of $\mathbf{Ax} = \mathbf{b}$ can be expressed in the form

$$\mathbf{x} = \underbrace{\mathbf{x}_\mathsf{p}}_{\text{particular solution}} + \underbrace{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \cdots + \alpha_k \mathbf{v}_k}_{\text{general solution}}$$

Conversely, for all choices of scalars

$$\alpha_1, \alpha_2, \ldots, \alpha_k$$

the vector $\mathbf{x}$ in this formula is a solution of $\mathbf{Ax} = \mathbf{b}$.

Q: What does this theorem mean geometrically?

Equivalence Theorem

If $\mathbf{A}$ is an $n \times n$ matrix, then the following statements are equivalent,

1. $\mathbf{A}$ is invertible.
2. $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced echelon form of $\mathbf{A}$ is $\mathbf{I}_n$.
4. $\mathbf{A}$ is expressible as a product of elementary matrices.
5. $\mathbf{Ax} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
6. $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
7. $\det(\mathbf{A}) \neq 0$.
8. The column vectors of $\mathbf{A}$ are linearly independent.
9. The row vectors of $\mathbf{A}$ are linearly independent.
10. $\mathbf{A}$ has rank n
11. $\mathbf{A}$ has nullity 0.