3. Matrix Rank

Tuesday, April 11, 2023 4:41 PM

Contents:

1) The rank
   - defn
   - How to compute?
2) Rank-Nullity Theorem
   - Theorem
   - Inequalities
3) Rank & Invertible matrices
4) Transpose
   - Defn & properties
   - Symmetric matrices
5) Why do we care?
   - Low rank matrices.

Reference: Strang 2 & 3

Rank of a family of vectors:

We define the rank of a family $x_1, \dots, x_k$ of vectors of $\mathbb{R}^n$ as the dimension of it's span:

$$\text{rank}(x_1, \dots, x_k) \overset{\text{def}}{=} \dim(\text{Span}(x_1, \dots, x_k))$$

Rank of a matrix: Definition:

Let $M \in \mathbb{R}^{n \times m}$. Let $c_1, \dots, c_m \in \mathbb{R}^n$ be its colums.

We define:

$$\text{rank}(M) = \text{rank}(\underbrace{c_1, \dots, c_m}_{\text{span of image space = span of columns of matrix.}}) = \dim(\text{Im}(M))$$

Example:

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \end{bmatrix} \quad 2\times 3$$

$c_1 \quad c_2 \quad c_3$

We are going from: $\mathbb{R}^3 \longrightarrow \mathbb{R}^2$

c_1, c_2, c_3 span the vector space of $\text{Image}(M) \longrightarrow$ so

Why? because

$\text{Span}((c_1, c_2, c_3)) = ?$

Rk: $\text{Span}((c_1, c_2, c_3))$ is a subspace of $\mathbb{R}^2$

- Are they Linearly dependent or Independent? $\boxed{c_3 = \frac{1}{2} c_1 + \frac{5}{4} c_2}$ Thus c_1, c_2, c_3 are NOT lin. Ind.
- $\underbrace{c_1 \,\&\, c_2 \text{ are LI}}$: vectors of $\mathbb{R}^2$ $\longrightarrow$ Thus basis
- (c_1, c_2) is a basis of $\mathbb{R}^2$.
- $\text{Span}((c_1, c_2)) = \mathbb{R}^2$
- $\text{Span}((c_1, c_2, c_3)) = \mathbb{R}^2$

$$\text{rank } M = \dim(\text{Span}((c_1, c_2, c_3))) = 2$$

rank ... (...

Proposition :

Rank of columns = Rank of rows.

Let $M \in \mathbb{R}^{n\times m}$. Let $r_1, \dots, r_n \in \mathbb{R}^m$ be rows of M
& $c_1, \dots, c_m \in \mathbb{R}^n$ be its columns.

Then we have :

$$\operatorname{rank}(r_1, \dots, r_n) = \operatorname{rank}(c_1, \dots, c_m) = \operatorname{rank}(M).$$

$$M = \begin{pmatrix} - r_1 - \\ - r_2 - \\ \vdots \\ - r_n \end{pmatrix} = \begin{pmatrix} | & & | \\ c_1 & \cdots & c_m \\ | & & | \end{pmatrix}$$

Implies

$$\boxed{\dim(\operatorname{Span}(r_1 \dots r_n)) = \dim(\operatorname{Span}(c_1, \dots, c_m))}$$

a.k.a row-column equivalence property or the rank theorem

Proof?

Rank Intuition from Data Science

Consider a Matrix M of size 1000×500 :

$$M = \begin{pmatrix} - r_1 - \\ \vdots \\ - r_{1000} - \end{pmatrix}$$

What does it mean to say that « $\operatorname{rank}(M) = 5$ » ?

- $\dim(\operatorname{Span}(r_1, \dots, r_{1000})) = 5$

  means to express all those 1000 rows we actually only need 5 of them.

- there exist 5 rows, which can generate (through linear combinations) all the 1000 rows

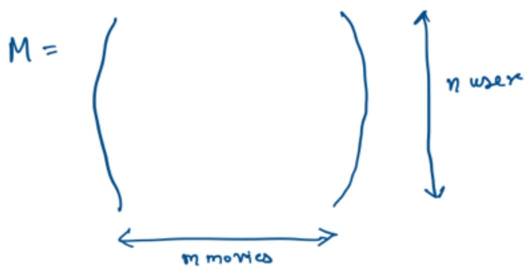
  In the sense that there is some dependence between rows
  & we only need 5 (unique) rows.

Example:

Imagine now that

- The rows of M correspond to Netflix users
- The column of M correspond to Netflix's movies
- The entry M_{ij} is rating of the movie j by the user i,

assuming that all the users have rated the movies.



We claim that this matrix has <u>low rank</u>.

What does that mean?

→ it has a lot of similar movies or users.

- The ratings of a user can be obtained as a <u>linear combination</u> of a small number of « profiles »

- In practise, we don't have access to the full matrix, so we can use this assumption to predict the missing entries.

How do we compute the rank?

For $v_1, \dots, v_k \in \mathbb{R}^n$, and $\alpha \in \mathbb{R} \setminus \{0\}$ ($\neq 0$), $\beta \in \mathbb{R}$ we have

$$\operatorname{rank}(v_1, \dots, v_k) = \begin{cases} \operatorname{rank}(v_1, \dots, v_{i-1}, \alpha v_i, v_{i+1} \dots, v_k) & \text{multiply } v_i \text{ by } \alpha \\ \operatorname{rank}(v_1, \dots, v_{i-1}, v_i + \beta v_j, v_{i+1}, \dots, v_k) & \text{replace } v_i \text{ by } v_i + \beta v_j \end{cases}$$

As a consequence, the Gaussian Elimination method keeps the rank of matrix unchanged.

Rank Nullity Theorem

Theorem:

Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then:

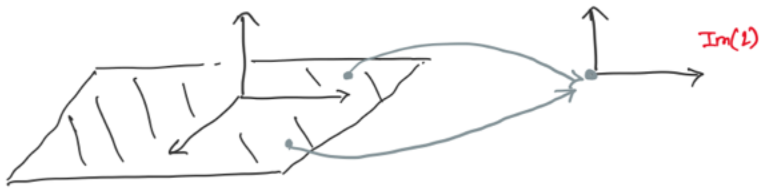
$$\boxed{\text{rank}(L) + \dim(\text{Ker}(L)) = m}$$

matrix of $n \times m$

dimension of input space

Intuition : "Conservation of dimension"

Suppose $\mathbb{R}^3 \xrightarrow{L} \mathbb{R}^2$



But note that

$\dim(\text{Im}(L)) = \text{rank}(L)$

" 2 out of 3 dimensions of $\mathbb{R}^3$ that are mapped to 0."

Let us solve the Linear system $Ax = 0$ characterizing $x \in \text{Ker}(A)$

$$\underbrace{\begin{bmatrix} 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ -1 & 5 & 2 & 0 \end{bmatrix}}_{3\times 4} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\mathbb{R}^4 \to \mathbb{R}^3$

"Augmented Matrix"

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 0 \\ 2 & 0 & 1 & -1 & 0 \\ -1 & 5 & 2 & 2 & 0 \end{array}\right) \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix}$$

$$\Rightarrow \left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 0 \\ 0 & 2 & 1 & -3 & 0 \\ 0 & 0 & 0 & 7 & 0 \end{array}\right) \Rightarrow \begin{cases} x_1 - x_2 + 0x_3 + x_4 = 0 \\ 0x_1 + 2x_2 + x_3 - 3x_4 = 0 \\ 0 + 0 + 0 + 7x_4 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x_1 = -1/2\, x_3 \\ x_2 = -1/2\, x_3 \\ x_4 = 0 \end{cases}$$

only 1 free variable

"set of solution"

$S = \text{Ker } A = \{ (-t/2, -t/2, t, 0) \mid t \in \mathbb{R} \}$

$= \text{span}\big(\underbrace{(-1/2, -1/2, 1, 0)}\big)$

will serve as basis of Kernel

Since only 1 vector —

$\dim(\text{Ker } A) = 1$

Dimension of Kernel ?

$\hookrightarrow$ 1 (Span of single vector)

$\dim(\text{Ker } A) = 1$

$\dim(\text{Ker } A) = 1$ $\qquad\qquad\qquad\qquad \dim(\text{Ker } A) = \ldots$

By Rank Nullity Theorem:

$$\dim(\text{Im}(A)) = \text{Input space dimension} - \dim(\text{Ker } A)$$

$$= 4 - 1$$

$$\boxed{\dim(\text{Im}(A)) = 3}$$

Important Inequalities:

Proposition:

Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times k}$. Then the following holds:

1.) $\boxed{\underbrace{\text{Rank}(A)}_{\dim(\text{Im}(A))} \leq \min(m, n)}$

2) $\boxed{\text{Rank}(A,B) \leq \min(\text{rank}(A), \text{rank}(B))}$

Proof:

1) $\text{Rank}(A) \leq \min(m,n)$

$$A = \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} — r_1 — \\ \vdots \\ — r_m — \end{pmatrix}$$

Recall that rank of A = $\underbrace{\underbrace{\text{rank of rows}}_{\text{can be at max } m} = \underbrace{\text{rank of columns}}_{\text{can be at max } n}}_{\text{thus} \leq \min(m,n)}$.

$$\text{rank } A = \dim(\text{Span}(c_1, \ldots, c_n)) \leq n$$

$$= \dim(\text{Span}(r_1, \ldots, r_m)) \leq m$$

$\therefore$ conclusion: $\text{rank } A \leq \min(m,n)$

2) $\boxed{\text{Rank}(AB) \leq \min(\text{rank } A, \text{rank } B)}$

Show that $\begin{cases} \text{Rank}(AB) \leq \text{rank}(A) \\ \text{Rank}(AB) \leq \text{rank}(B) \end{cases}$

$A \in R^{n \times m}$ $\quad B \in R^{m \times k}$

$AB \in R^{n \times k}$ $\quad \rightarrow \text{rank} = n$

$\therefore \text{rank}(AB) \leq \min(n, k)$

But

$\text{rank}(A) \leq \min(\overset{3}{n}, \overset{2}{m})$

& $\quad \text{rank}(B) \leq \min(\overset{2}{m}, \overset{1}{k})$

Permutations:

if n smallest.

if $n > m$
$\text{rank}(A) = m$

if $m < k$
$\text{rank}(B) \leq m$.

if $m > k$
$\text{rank}(B) \leq k$.

Rank & Invertible Matrices

Theorem:

Let $M \in R^{n \times n}$. Then the following are equivalent:

1. M is invertible
2. $\text{rank}(M) = n$
3. $\text{Ker}(M) = \{0\}$.
4. For all $y \in R^n$, there exists a unique $x \in R^n$ such that $Mx = y$.

Recall:

M is invertible if there exists M^{-1} such that

$MM^{-1} = M^{-1}M = Id_n$

Note:

Point 2 and 3 are equivalent

Because: Rank Nullity Theorem:
$\text{Rank}(M) + \dim(\text{Ker}(M)) = n$

$\therefore$ if $\text{Rank}(M) = n$,

$n + \dim(\text{Ker}(M)) = n$

We get,

$\boxed{\dim(\mathrm{Ker}(M)) = 0}$

Proof :

1) M is invertible $\implies$ (3) $\dim(\mathrm{Ker}(M)) = 0$

if $x \in \mathrm{Ker}(M)$, $\quad Mx = 0$

using the assumption that M is invertible,

$$M^{-1}Mx = 0$$
$$Ix = 0$$
$$x = 0$$

$\mathrm{Ker}\, M \subset \{0\}$

And $\{0\} \subset \mathrm{Ker}\, M$

So $\mathrm{Ker}\, M = \{0\} \iff \dim \mathrm{Ker}\, M = 0$

2) $\mathrm{rank}\, M = n$

3) $\dim(\mathrm{Ker}\, M) = 0$

4) exists unique x solution $Mx = y$ for any y in $\mathbb{R}^n$.

we have proved (4) given (1) & (2)

- $\mathrm{rank}\, M = n$ what can we say about $\mathrm{Im}(M)$?
  - $\mathrm{Im}(M)$ is a subspace of $\mathbb{R}^n$
  - $\mathrm{rank}(M) = \dim \mathrm{Span}(\ldots) = \dim \mathrm{Im}(M) = n$
  - $\mathrm{Im}(M) = \mathbb{R}^n$.
- Any y in $\mathbb{R}^n$ belong to $\mathrm{Im}(M)$, so there exists a solution of x.
- Solution is unique because $\mathrm{Ker}(M) = \{0\}$.

Let's also prove the other way round to establish equivalence.

For every $y \in \mathbb{R}^n$, unique x $\quad Mx = y$

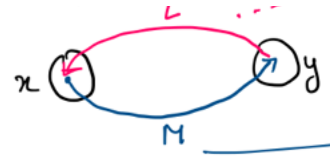
Let's define :

$$L : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ y & \longrightarrow & x \end{array}$$

with $\tilde{\mathcal{I}}$ canonical

... can be written

L is a Linear Transform that converts y to x



→ x can be as $L(y) = x$

→ we know $Mx = y$.

Claim: L is a linear transformation

$$\mathbb{R}^{n \times n} \quad M \overbrace{L(y)}^{\text{linear transformation}} = Mx = y$$

in terms of composition of linear maps:

$$\underbrace{M \circ L}_{\text{composition of linear maps.}} (y) = y$$

Thus,

$$\boxed{M \circ L = Id_n}$$

$$M\tilde{L} = Id_n$$

$\Rightarrow$ M is invertible and $\hat{L} = M^{-1}$.

Transpose of a Matrix

Definition:

Let $M \in \mathbb{R}^{n \times m}$. We define its transpose $M^T \in \mathbb{R}^{m \times n}$ by

$$(M^T)_{i,j} = M_{j,i}$$

for all $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$

Remarks:

- We have $(M^T)^T = M$
- The mapping $M \longrightarrow M^T$ is linear. → $\left\{ \begin{array}{l} (A+B)^T = A^T + B^T \\ (\alpha A)^T = \alpha A^T \end{array} \right.$

Properties of the Transpose:

Proposition:

For all $A \in R^{n \times m}$, $\boxed{\operatorname{rank}(A) = \operatorname{rank}(A^T)}$ } → rank of family of rows = rank of family of columns.

Proposition:

Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times k}$. Then:

$$\boxed{(AB)^T = B^T A^T}$$

Symmetric Matrices

Definition: A square matrix $A \in R^{n \times n}$ is said to be symmetric if:

$$\forall\, i,j \in \{1, \dots, n\},\ A_{i,j} = A_{j,i}$$

or, equivalently if $A = A^T$.

Remark: For all $M \in R^{n \times m}$ the matrix MM^T is symmetric